# An empirical and theoretical study of stereoscopic illusory contours and surfaces

Thesis submitted by
Samuel Huf BPsych(Hons)
in June 2002

for the degree of Doctor of Philosophy
in the School of Psychology
James Cook University

## STATEMENT OF ACCESS

I, the undersigned, the author of this thesis, understand that James Cook University of North Queensland will make it available for use within the University Library and, by microfilm or other photographic means, allow access to users in other approved libraries. All users consulting this thesis will have to sign the following statement:

"In consulting this thesis I agree not to copy or closely paraphrase it in whole or in part without the written consent of the author; and to make proper written acknowledgement for any assistance which I have obtained from it."

Beyond this, I do not wish to place any restriction on access to this thesis.

24/06/03

(signature) (date)

## Declaration

I declare that this thesis is a record of original work and that it contains no material which has been accepted for the award of any other degree or diploma in any University.

To the best of my Knowledge and belief, this thesis contains no material previously published or written by any other person, except where due reference is given in the text of the thesis.

I consent to this thesis being made available for photocopying or loan.

Sam Huf
14 /05/03

## Acknowledgements

I wish to thank my academic supervisors Prof. Colin Ryan (JCU) and Prof. Barbara Gillam (UNSW) for their efforts. Special thanks to Dr. Chris Woodruff (DSTO) and Assoc. Prof. Rick van der Zwan (JCU) for helpful comments. I am also very grateful to Colin Cressy (JCU) for his assistance with early programming (and as the teacher who inspired my first interest in Psychology).

**But most thanks are to my daughter Isabel for waiting!**

## Contents

# List of Figures

# Part 1

## Background to stereoscopic illusory contours and surfaces

*Summary: The first part of the thesis outlines major concepts important for later analyses.*

*Chapter 1 introduces concepts of stereoscopic vision. Topics include retinal coordinate geometry, retinal disparity, its neural substrates, disparity gradients, the correspondence problem and a contemporary information-processing model of binocular vision.*

*Chapter 2 presents some of the diverse literature that has addressed illusory contours in 2-D stimuli. The chapter defines illusory contours, discusses the notions of modal and amodal completion and the experiments that have revealed the psychophysical character of illusory contours. Suggested neural substrates and Grossberg's explanation of 2-D illusory contours are also presented.*

*In Chapter 3, recent research in stereoscopic illusory contours and surfaces is examined. Stimuli including the stereoscopic Kanizsa square, Gulick and Lawson's sparse texture matrices, stereoscopic contrast spreading, and stereo capture effects are discussed. The chapter develops theoretical distinction between a Surface Heuristic view and a Form Computation view of the mechanisms that underpin stereoscopic illusory contours and surfaces.*

# 1. Stereoscopic vision

*Summary: This chapter introduces key concepts of binocular vision used in subsequent analysis and discussion of stereoscopic illusory surfaces. Retinal coordinate geometry, the topographic organisation of the visual system, relative horizontal and vertical disparities, orientation disparity, the correspondence problem, and Panum's limiting case are briefly addressed.*

## 1.1 Theoretical ambiguity of retinal coordinate measures

### 1.1.1. Monocular geometry and the optic array

The retinal images are geometric products of rectilinear projection of light from surfaces and substances inhabiting the visual field. Euclid, around 300 B.C. may have first identified perspective geometry of visual space.

At least since Euclid, it has been recognised that the visual system extracts measures of the retinal images to help work out the spatial arrangement of the visual field (Pizzlo, Rosenfeld and Wiess, 1997; Howard and Rogers, 1995; Gulick and Lawson, 1976). To do this the system must accurately access 2-D proportions of the images. The retina provides such a device. It is structured in terms of spatial coordinates. This structure is the basic sensory tool for depth perception. A monocular coordinate matrix is described in Fig 1.1.

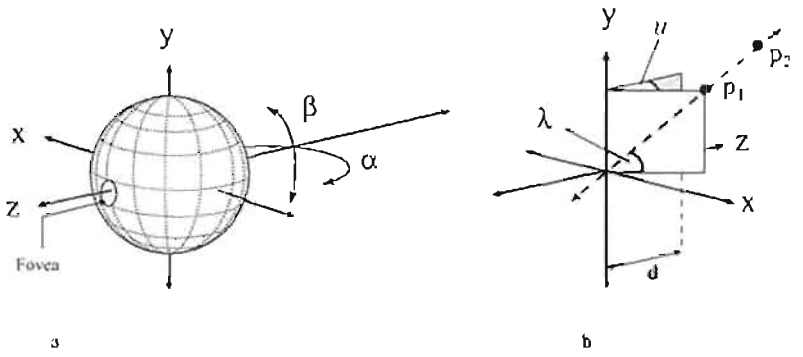



Fig 1.1. A monocular coordinate model

In Fig 1.1a, a sphere represents the eye. The retina lines the nearer hemisphere. All optic projections pass through the centre of the sphere that acts as the origin of a coordinate system based on the arrangement outlined in 1.1b. The coordinate origin is centred around a visual line passing though the centre of the fovea and the system origin, x and y = 0. Viewing distance is some measure along the axis z. In this way, a retinotopic map of the optic array is captured. A narrow central visual core of the array is focussed on the fovea, the site of most detailed information extraction. Points can be described by $p_1$ azimuth ( $u$ ) and elevation ( $\lambda$ ) at a certain distance from the system origin (adapted from Howard and Rogers, 1995).

The human visual system is topographically organised. This means that the coordinate structure of the retina, the spatial arrangement of sensory elements, is maintained with great precision at the visual cortex. Indeed precise retinal topography is reproduced in many different regions of the striate and extra-striate cortices (Hubel, 1980; Kandel, Jessell and Schwartz, 1991).

Topographic and parallel organisation allows the visual system to exploit the geometry of an array of projections. However, coordinate geometry alone is usually thought to explain little of our actual experience of visual space because all visual angles are theoretically ambiguous. For example, in Fig 1.1b, point $P_2$ can occupy the same coordinate position on the retina as point $P_1$. Moreover, there are an infinite number of points at different distances that could fall along the same projection line.

Accommodation of the crystalline lens is one purely proprioceptive source of information that contributes to interpretation of retinal coordinate measures. Proprioceptive feedback from accommodation, however, is not sufficient to explain the perception of space. A strong *sense of space* is very easy to invoke using simple 2-D pictures, that is, with accommodation, visual angles and perspective projection essentially fixed (Rogers, 1995a).

### 1.1.2 Pictorial cues and monocular perspective

Renaissance painters were among the first to pictorially represent a near-natural sense of space and depth. They achieved this by the systematic study and exploitation of visual perspective. Leonardo da Vinci (1452-1519), for example, simulated linear, aerial, texture perspective, size and height in field, perceptual grouping, interposition, shadowing, lightness, and so on, by developing pictorial cues which evoked the experience of a natural setting.

All the monocular coordinate system can do in spatial terms is to define the direct spatial position of points of light on a retinal coordinate map. As the eyes, the head, the body and the objects in the environment move, the monocular coordinate system must be referenced against environmental invariants such as horizon, relative motion, head and eye position (Howard and Rogers, 1995).

Superficially, at least, the same monocular coordinate data term can literally represent many possible alternatives, that is, visual angles are ambiguous. The ambiguity of retinal coordinate measures is reduced by stereoscopic vision.

## 1.2 Stereoscopic Vision

### 1.2.1 *Stereopsis*

*Basic sensory topography*

Stereoscopic vision requires two forward-looking eyes situated close to each other at the front of the head, so that their visual fields substantially overlap. The arrangement is sketched in Fig 1.2. In humans, the monocular coordinate geometry of each eye contributes to an overall reconstruction of retinal images at the visual cortex. One half of the field is represented at each hemisphere. This is achieved as described in Fig 1.2. Passage of neural pathways through the optic chiasm enables the system to represent the spatial layout of the whole visual field at the striate cortex (Ogle, 1950). This involves combining the monocular images while maintaining the layout of the entire visual field.

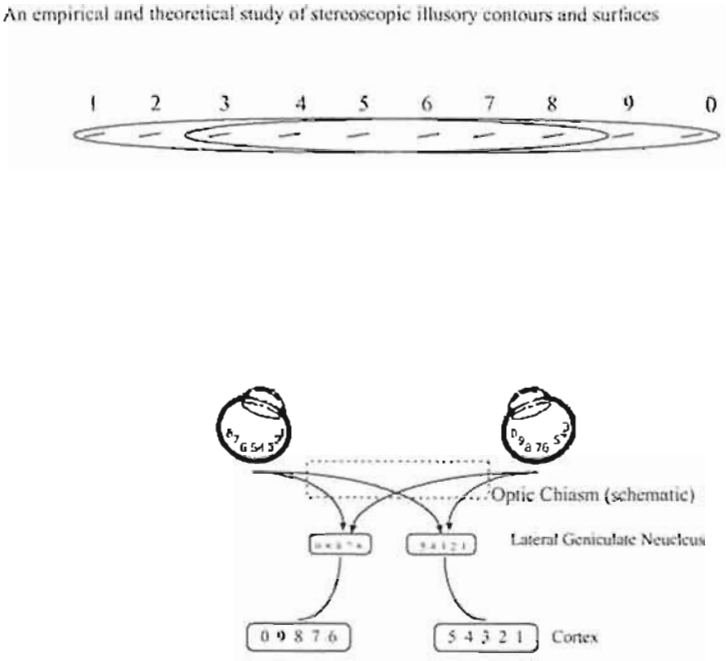



**Fig 1.2. Schematic organisation of the binocular vision system**
Adapted from Netrauli and Haskell (1988).

Image differences are available to the system because the optical arrangement of the eyes creates binocular parallax. This means that the system captures two slightly different views of nearby objects. Stereopsis is the recovery of visual depth from binocular parallax. This involves fusion of two perspective views to yield one singular view.

*Identical retinal points and visual direction*

Stereopsis requires that the organisation of the array of retinal sensors must be precisely coordinated. Hering first described this coordinated structure in terms of visual direction. Hering's *Retinal Point Mapping* and the *Law of Identical Visual Directions*, proposed in the 19th century, suggested that the system analysed depth by generating three "space feelings" - height, breadth and depth. Hering claimed that the perception of an object in depth carried the average of the three space feelings (Gulick and Lawson, 1976).

This was perhaps the first insight into what has sometimes been termed the cyclopean view or *cyclopean retina* (see for example Julesz, 1971). Hering argued that seen depth was actually the average direction of an object seen by two eyes. Thus, the union or fusion of sensations from the two retinae defined the visual direction of a point, not its position in either retina *per se*. Figure 1.3 demonstrates the notion of the cyclopean view.

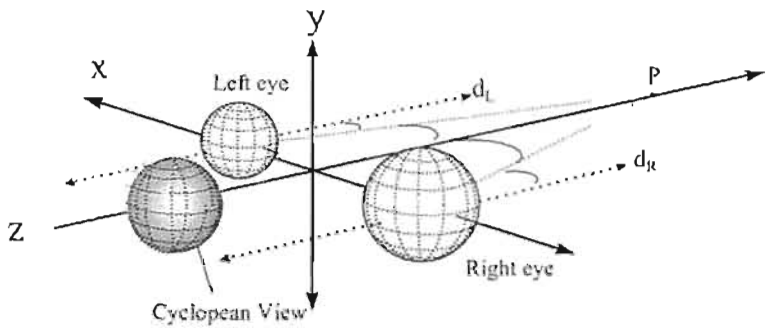



Fig 1.3. The cyclopean view
Hering's notion of visual direction placed the ego-centre of the stereoscopic depth percept between the two eyes. This arrangement has subsequently been described as a *cyclopean view* (Julesz, 1986; Garding, Porrill, Mayhew and Frisby, 1994; Gregory, 1998).

*Retinal disparity*

In the 1830s, Charles Wheatstone demonstrated that 'visual disparity' resulting from the projection to disparate rather than identical retinal points was a key factor in deriving depth from the fusion of the two retinal images (Gulick and Lawson, 1976; Bruce, Green and Georgeson, 1996; Gregory, 1998). Wheatstone was able to demonstrate that:

> ... when the half views of a solid object are exact replicas of the monocular views of that object *[in the binocular viewing situation]*, then binocular combination of the half-views yields a percept of the solid object *(cited in Gulick and Lawson, 1976, p. 20)*.

Wheatstone presented subjects with sketches such as those in Fig 1.4. Evidently perspective projection captured from two slightly different directions provided by interocular separation provides substantial difference information from which a 3-D cyclopean percept can be generated.

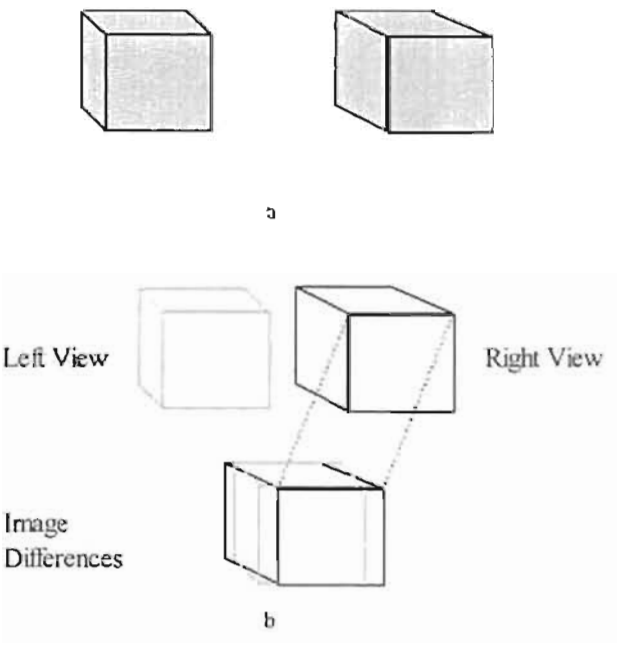



**Fig 1.4. Disparate perspective views**
Wheatstone demonstrated that the two monocular views of an object could be fused to generate a three dimensional view of that object. He also showed that reversing the two images resulted in an opposite signed depth percept. Crossed fusion of the images in (a) yields the percept of a three dimensional rectangular prism. The approximate perspective differences are shown in (b). Note that orientation differences as well as point position differences are available (adapted from Gregory, 1998).

*Object position, convergence and disparity*

The simplest explanation of disparity created by binocular parallax describes the relative position of two points in space. Fig 1.5 demonstrates that a relative distance separation (d) between points $P_1$ and $P_2$, can be derived as a function of the interocular distance (i) and viewing distance (v):

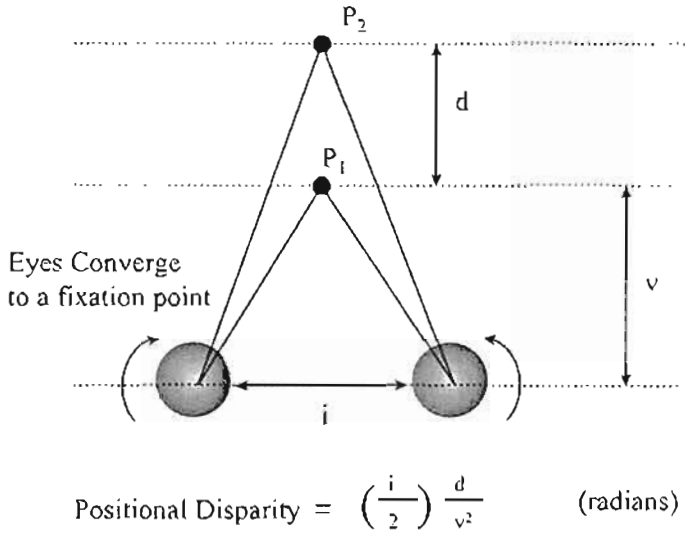



**Fig 1.5. Positional disparity and convergence**
Adapted from Ritter (1979).

There is evidence that convergence, the coordinated inward rotation of the eyes to fixate upon an object, is monitored to aid in scaling the absolute magnitude of retinal disparity (Riter, 1977, 1979; Wallach and Zuckerman, 1963).

*A binocular coordinate system*

This section describes a convenient binocular coordinate system used to explain subsequent issues in stereopsis. Assuming that a monocular coordinate system describes detection of a monocular array of projections, it is possible to combine two such systems to derive precise geometric analysis of object position. The critical assumption behind this notion is that fixating on a point in the field with two eyes aligns monocular coordinate systems, as suggested by Hering in 19$^{th}$ century (Gulick and Lawson, 1976).

Theoretically, differences in the two views can be prescribed as differences in the angles of binocular subtense in relation to that fixation point. Angular disparity arises from differences in azimuth ($\alpha$), which will be termed horizontal disparity ($\Delta\alpha$), and differences in elevation ($\beta$), to be termed vertical disparity ($\Delta\beta$).

The binocular coordinate system described in Fig 1.6, can theoretically derive the position of a point P or $P_1$ in space according to their relative positions in each retina, given that the convergence of the eyes is monitored proprioceptively.

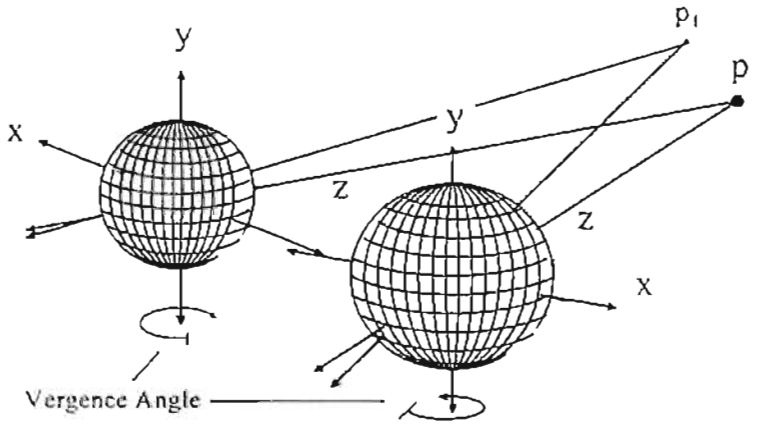



**Fig 1.6. A binocular coordinate system**
Appropriate eye movements align the two coordinate systems in fixating on P. The position of $P_1$ can therefore be described according to its elevation (Y-axis and X-axis, hence its distance Z can be derived) (adapted from Howard and Rogers, 1995).

### 1.2.2 The nature of retinal disparities

This section is concerned with a geometric definition of retinal disparities of various types, their proposed neural substrates and their role in depth perception.

*Horizontal retinal disparity and the longitudinal horopter*

Fig 1.7 illustrates the optical geometry underpinning horizontal retinal disparity. The coordinate origins are represented as head-centric directions. When the eyes converge to fix on a single point, $P_1$, its visual projections pass through the optic centres of both eyes. A theoretical arc prescribes the distance of other points for which binocular subtense will be equal. This arc is termed the longitudinal horopter or the Vieth-Muller Circle. Points $P_1$ and $P_2$, in 1.7a, are therefore subtended by the same visual angles with reference to each optic axis. Further, in each eye, visual projections from $P_2$ will intersect the same relative horizontal coordinate positions in the left and right eyes.

Fig 1.7b then, demonstrates the idealised theoretical derivation of retinal disparity. Points $P_1$ and $P_3$ will subtend different binocular angles. The magnitude of the difference is relative to the distance of $P_3$ from the Vieth-Muller circle. Point $P_3$ lies beyond the horopter, so all points with the same binocular subtense will fall upon arcs describing the same disparity, hence these arcs are termed isodisparity circles.

The horopter represents a 'zero-disparity', isodisparity circle. Points positioned along different isodisparity circles, such as $P_2$ and $P_3$, define a disparity relative to $P_1$ of opposite sign.

Points falling on nearer isodisparity circles (eg. $P_2$ in Fig 1.7b) prescribe *convergent* or *crossed* disparity. Similarly, points lying on isodisparity circles beyond the fixation point, (eg. $P_3$ in Fig 1.7b) represent *divergent* or *uncrossed* disparity.

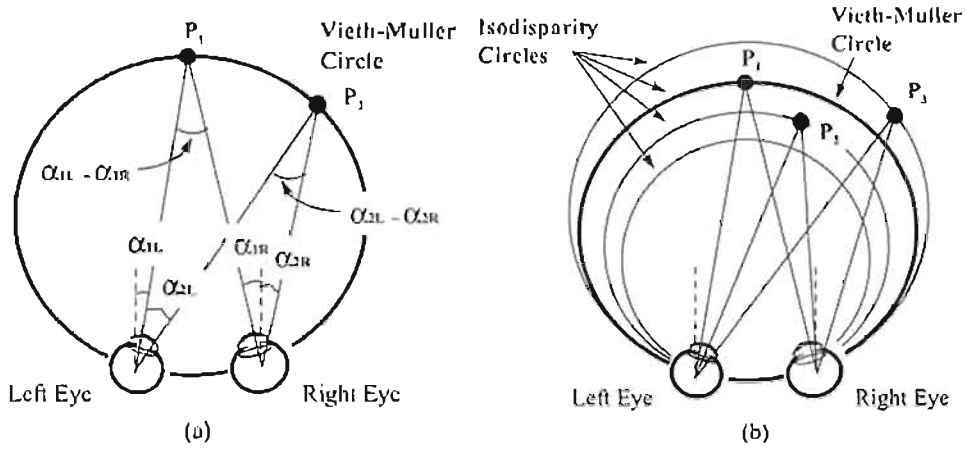



**Fig 1.7. Vieth-Muller and isodisparity circles**
The absolute azimuth of two points P1 and P2 are specified with respect to aligned axes in the two eyes. The binocular subtense of each point corresponds to the difference in their absolute azimuth (Adapted from Howard and Rogers, 1995, pp. 248-250).

Horizontal disparity between a pair of points is the difference between the absolute angles of binocular subtense. Howard and Rogers express the relationship as follows:

$$\text{Horizontal Disparity, } (\Delta\alpha) = (\alpha_{1L} - \alpha_{1R}) - (\alpha_{2L} - \alpha_{2R})$$

*The relative nature of horizontal disparity*

The utility of horizontal disparity as a depth cue lies in the optical characteristics of binocular parallax. Disparity magnitude will be a geometric function of the inter-ocular distance, the distance of points from the eyes, and the spatial separation of those points. It is evident in simple textbook demonstrations of disparity, such as those in Fig 1.8a, that, if the line pairs A and B are cross fused, the thick line will stand behind the thin line in the cyclopean view[1]. Fusing pairs B and C reverses the effect.

It is interesting to note that experienced fusers can shift fixation between the two fused pairs as well. This is interesting, because it seems to require the system to *double fuse* the

1 - For experienced free-fusers, it is possible to shift fixation from one line to the other, track the eyes along the fixated line, and move one's head slightly, yet have the two lines remain in a similar depth relationship despite changes in the absolute disparity measures.

image pair B. A simple demonstration can verify this. Try pointing a figure at one of the pairs while it is binocularly fused. Your finger will appear double while the cyclopean view is stable. Fig 1.8b schematically illustrates the disparate views relative to the projection plane which, in this case, is the surface of the page.

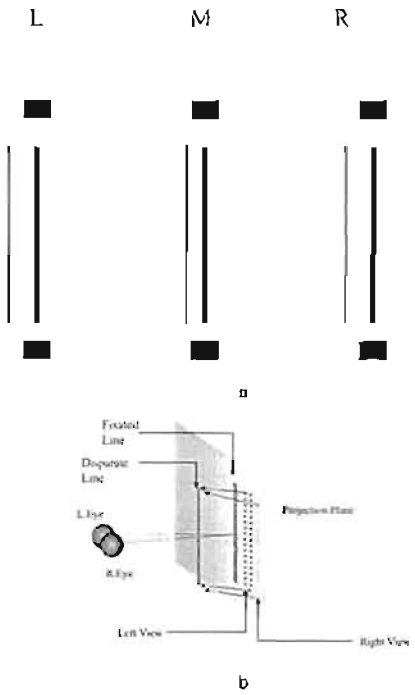



**Fig 1.8. Relative horizontal disparity and the perception of depth**
The depth percept derived from retinal disparity. In (a), cross fusing L and M yields the sense that the thin line is nearer. Cross fusing M and R gives the impression that the thin line is more distant (the effect will be reversed for divergent fusers). This effect is demonstrated pictorially in (b).

*Binocular fusion and the limits of disparity processing*

Two points of similar quality, but representing different retinal coordinate positions, can be fused together to give the experience of a single point lying at some position in depth. This combination of disparate retinal images is termed binocular fusion. A threshold region of fusion was first established by Panum (1858, Howard and Rogers, 1995; Gulick and Lawson, 1976; Gillam, Blackburn, and Cook, 1995).

Panum's fusion zones define the limits of *single vision*. Points outside those zones cannot be fused and will appear double—a phenomenon known as *diplopia*. If the eyes fixate a given point, the region of space around that point within which fusion (single vision) is possible is about 0.1° visual angle. This means that at 57 mm fusional space is +/- 9mm (Bruce, Green

and Georgeson, 1996). The scale of these limits appears to be increased for larger objects and increases to a certain extent in the far peripheries of the visual field (Ogle, 1950). There is evidence that these limitations are dynamic and, according to Julesz (1986), fusion limits can be *stretched* after fusion to many times greater than traditionally accepted limits.

A maximum disparity gradient threshold of single vision also exists (Burt and Julesz, 1980). *Disparity gradient* is the difference in coordinate position between points as a ratio of their horizontal separation. It describes change in horizontal disparity across an inclined surface. The gradient limit of fusion is 1, that is, when the disparity evident between two points equals their horizontal separation (Howard and Rogers, 1995).

*Physiological substrates of disparity detection*

The concept behind retinal disparity as a useful information source is that experience or judgment of depth must be derived from some neural response by the system that assigns a distance in depth to positional differences between corresponding points in each eye's coordinate system. But disparity does not give depth on its own.

It has been shown that many neural cells in the lower visual cortex are binocular, that is, their receptive fields monitor receptor output in both eyes. There is substantial evidence that these cells somehow compare the position of corresponding points in each monocular part of the receptive field. Some simple cells of area V1 and V2 of the striate cortex appear to be directly sensitive to signals applied to horizontally disparate regions of their binocular visual fields (Poggio and Poggio, 1984); to vertically disparate regions (Tyler, 1975); and to differences in orientation (Gillam and Rogers, 1989; Cagenello and Rogers, 1989; Ninio, 1985).

Evidence also suggests that three channels of disparity processing may be in operation. These include *near* (crossed disparity sensitive), *far* (uncrossed disparity sensitive) and a so-called *tuned* (network modulated sensitivity) disparity detectors. The activity of these channels may be mutually inhibitory, that is, they may operate in an opponent mechanism (Cormack, Stevenson and Schor, 1993).

*Vertical disparity*

Differences in the angle of elevation subtended by each eye theoretically prescribe vertical disparities (see Fig 1.9).

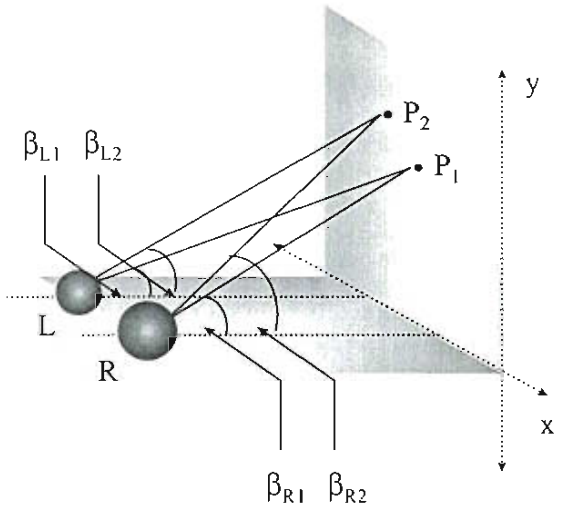



**Fig 1.9. Theoretical vertical disparity**
Due to the horizontal separation of the eyes, points at different eccentricities will yield different binocular angles of subtense.

From Fig 1.9, theoretical vertical disparity can be expressed as:

$$\text{Vertical Disparity } (\Delta\beta) = (\beta_{1L} - \beta_{1R}) - (\beta_{2L} - \beta_{2R})$$

The role of vertical disparities in the perception of depth is not presently well understood. Mayhew and Longuet-Higgins (1982) demonstrated theoretically that depth might be recovered using relative horizontal and vertical disparities across any three points in space, that is in the absence of proprioceptive cues.

In psychophysical terms, vertical disparity is a subtle source of distance information. It theoretically yields valuable information about gaze angle, and may guide vergence control, that is, vergence may be set by minimising vertical disparities (Poggio and Pogio, 1994). The role of vertical disparity may be as part of the overall pattern of disparities across surface features, that is, as vertical size ratios (VSRs) rather than as a discrete derivative of positional difference between corresponding retinal points (Rogers and Bradshaw, 1993, 1995).

### 1.2.3 The correspondence problem and surface texture

A key assumption of contemporary theories of stereopsis, is that the system identifies matched pairs of features in the surface textures captured at each image so as to or arrive at coordinate position differences between each retina.

*Disparity alone as a cue to distance*

Kenneth Ogle saw stereopsis as primarily a physiological determination of the geometric relations between recognisably corresponding contours in the field. He stated:

> We must stress the importance of contours , those lines of demarcation between the "figure" and the "background". In every case stereoscopic depth depends on the disparity between the images of identifiable contours (Ogle, 1959, p. 380).

According to Gulick and Lawson, Wheatstone had paid little attention to the correspondence problem. He did note, however, that a matching process identified *similar* features in the case of ambiguity. Figure 1.10 demonstrates his similarity principle. In this figure, there are two possible matches for the single line. A line of similar thickness is chosen as the *correct match*. A *similarity constraint* such as this underpins virtually all contemporary theories of binocular matching.

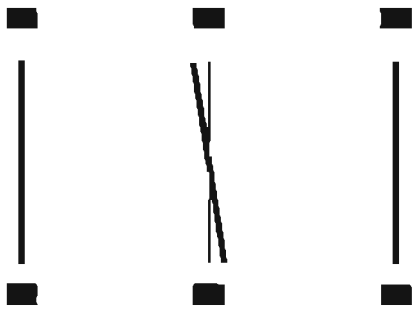


**Fig 1.10. Wheastone's Similarity Principal**
Crossed fusion of the left pair of images yields a percept of the thick line rotated obliquely. With the thin line attributed to the depth of the fixation point determined by fusion of the upper and lower black rectangles (Adapted from Gulick and Lawson, 1976).

In summary, early notions of correspondence were that features readily recognisable in each eye's view were aligned or fused in generating a 3-D view. This idea proved to be naïve to the physics of retinal images and to the remarkable capacity of the binocular system to synthesise correspondence patterns from complex visual stimuli. As technology with which to explore binocular vision improved, in the later half of this century, the ability of the system to extract disparity information from complex surface textures was recognised.

*Julesz and the RDS*

Bela Julesz first introduced the concept that point disparities within a matrix of apparently random points induced a sense of depth in the absence of any contours recognisable to each single eye. Julesz saw this as evidence of a pure Cyclopean perception. A simple example is given in Fig 1.11. Note that the random texture matrix is almost identical in each eye's view, that is, each dot in one eye's view of the matrix has a matched pair in the other. The difference is that a central square figure has been shifted horizontally in one of the eye's views relative to the other, generating a uniform pattern of disparity amongst the dot pairs. Resulting vacant regions of the matrix are filled in by unpaired random-dots.

Julesz's (1964) suggestion was that, in the complete absence of identifiable monocular contours (luminance differences available in a single retinal image), the system was capable of interpreting depth from an array which gives rise to contours only available once point matches were decomposed. Julesz's work, has had enormous impact. He produced a series of demonstrations of cyclopean perception, evidence that binocular depth could be purely disparity driven and could constrain illusions such as the Muller-Lyer and Poggendorf line length effects.

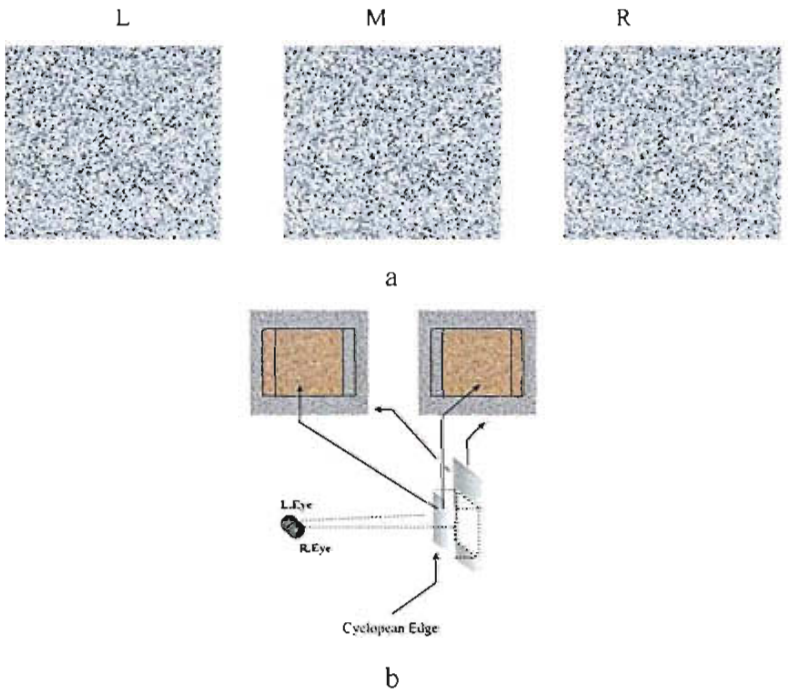



**Fig 1.11. Random dot stereograms**
In the absence of any monocularly identifiable contour, Julesz demonstrated that disparity within a region of other wise random points could generate a sense of depth. In 1.15a, crossed fusion of L and M yields the percept of a central square standing forward of a textured background. Fusion of M and R sees the square standing behind a larger textured ground. This is demonstrated pictorially in 1.15b.

*Cyclopean Perception*

Julesz argued that disparity extraction must proceed at some cortical locale beyond the Lateral Geniculate Nucleus (LGN). His general processing scheme is presented in Fig 1.12.

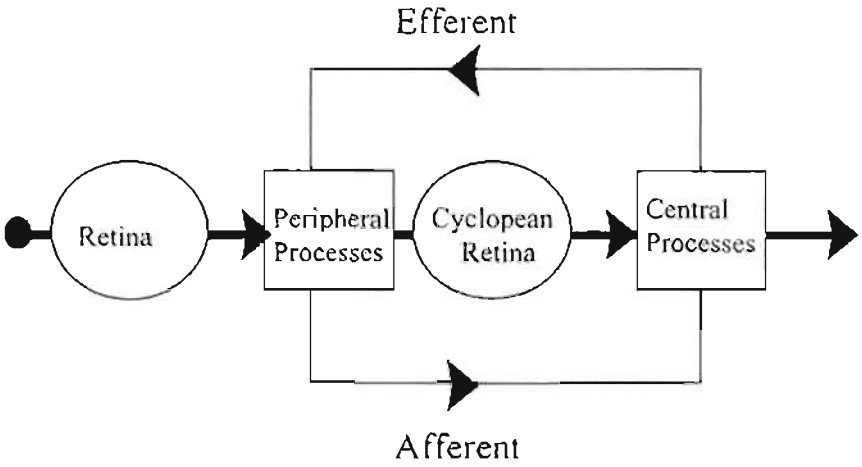



**Fig 1.12. Cyclopean processing of RDS figures**
Julesz cyclopean scheme was a classic *Black Box* account of information processing. Retinal stimulation was influenced by peripheral factors such as feedback from a central processor. Image difference patterns constituted a cyclopean retina. (Julesz ,1971, p27).

The actual matching primitives in Julesz's original cyclopean scheme were texture elements themselves (dots), defined by luminance contours. The propensity of the system to achieve a 3-D experience from the RDS introduced another layer of uncertainty to the problem of stereoscopic vision.

In an RDS matrix, an infinite number of possible matches could theoretically be achieved as the system attempted to identify correspondence amongst image pairs. Such images take some considerable time to *fuse*, which would seem to undermine their utility in describing natural vision. Saye and Frisby (1975) demonstrated that insertion of monocularly conspicuous contours provide vergence control which greatly speeds the time taken to fuse RDS figures.

*Computation of disparities and constraint of point to point matching*

Julesz (1986) argued that stereopsis could be broken into local and global degrees of scale. Local correspondence recovery was the lowest scale process at which fine-grained textures

were analysed and correspondence determined. Local correspondence decomposition was guided and revised through cooperative linkages with a global stereoscopic scheme, which addressed issues of correspondence between large features such as luminance boundaries. The cyclopean shape derived from a matching solution verified correct matches.

Many other attempts to constrain matching ambiguity have followed. It is not useful to review all of these in the present context since in virtually all images generating illusory contours correspondence is not a difficult issue. Note that Howard and Rogers outline some fourteen of the commonly imposed matching constraints.

Marr and Poggio (1979) argued that the major issue for stereoscopic vision was stage-wise elimination of *false- targets*. They suggested that both retinal images were subjected to four scales of spatial frequency filtering, and up to twelve orientation sensitive filters. These were seen as independent processing channels, which produced a memory buffer termed the 2½D sketch. This was essentially a map of intensity values across the two retina. Patterns of retinal correspondence were then decomposed by cooperation between filtering devices and orientation sensitive units. Low-frequency channels controlled vergence shifts (equivalent to Julesz's global stereopsis) which would bring finer grained texture near to correspondence (Equates to local level processing). The matching primitives in Marr and Poggio's theory were points of contrast inflection between intensity peaks, termed zero-crossings. Their primary assumptions were:

> [1]... a given point on a physical surface has a unique position in space at any one time.
>
> [2]...matter is cohesive, it is separated into objects, and the surfaces of objects are generally smooth compared with their distance from the viewer
>
> [3] ... when correspondence is achieved , it is held and written down in the 2½D sketch.
>
> [4] ... there is a backwards relation between the memory and ... the control of eye movements (Marr and Poggio, 1979, p. 302).

Mayhew and Frisby (1981) altered this scheme using intensity peaks as well as zero-crossings—and included a continuity constraint which checked candidate feature matches against *figural* possibilities presented by images in the other. This essentially represented an

ecological validity check of the components involved. Figure 1.13, below, illustrates the concepts of zero-crossings and intensity peaks in a square wave patterned stereogram.

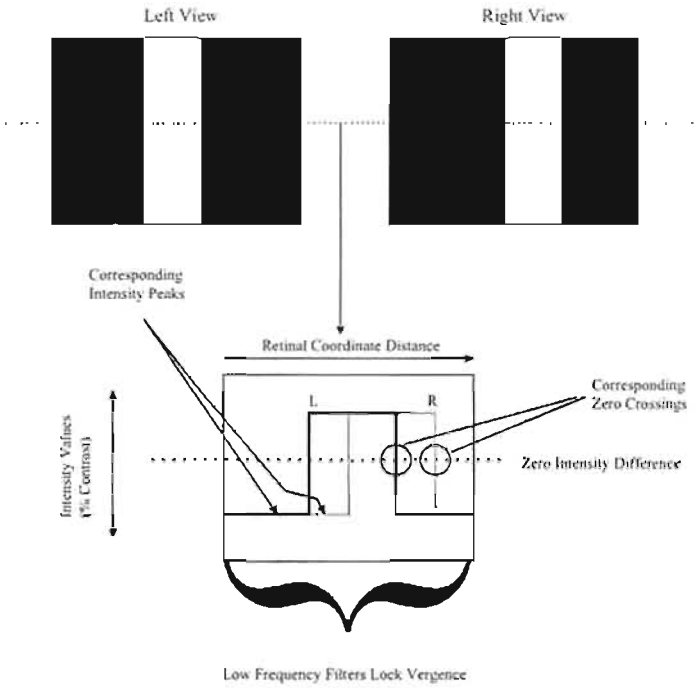



**Fig 1.13. Intensity values as matching primitives**
Zero-crossings and intensity peaks have commonly been utilised as matching primitives in stereopsis.

*Problems for point-matching*

The concept of stereopsis as a point matching or contour matching process has dominated the field since. This section reviews three observations that pose problems for this view. The first is stereopsis from shading. Secondly, Panum's Limiting Case shows that it is possible for more than one feature to correspond with a feature in the other eye. Third, occlusion yields regions of the field that are non-correspondent and therefore theoretically unmatchable. Each is dealt with in turn.

**Stereo and Shading**

Bulthoff and Mallot (1988) have demonstrated that stereoscopic depth can be induced in the complete absence of zero-crossings, or of steep intensity peaks, using stereograms with disparate slopes of smooth intensity change (shading). Stereopsis in such stereograms is predicted by the disparate rate of intensity change. Figure 1.14, below, demonstrates this concept. There are no corresponding features in these images, except for the outer perimeters

of the stereograms, which will facilitate vergence fixation, or what has been termed fusion locking (by Gulick and Lawson, 1976; Anderson , 1997).

Mallott, Arndt and Bulthoff (1995) argue that minimisation of mean squared intensity differences can account quantitatively for depth perceptions evoked by these smoothly shaded images. The amount of depth derived from surface regions with no explicit contour differentiation is said to be an average of the true disparities.

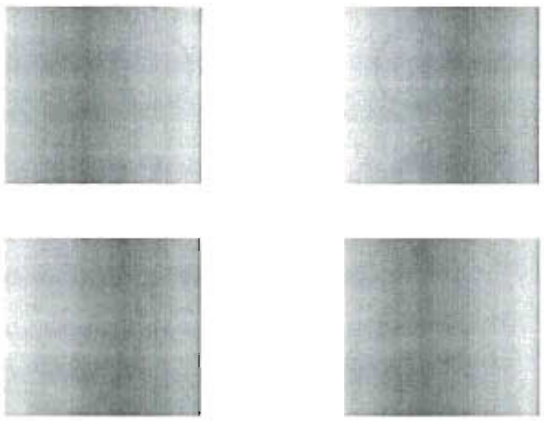


**Fig 1.14. Stereopsis from Shading**
Crossed fusion of the top pair sees a portion of the surface standing forward of the projection. The effect is reversed in the lower pair (Note that these stereograms are only approximations of stereo-intensity effect for demonstration).

*Panum's Limiting Case*

Another difficult issue for computational theorists is a classic effect, once again, from the 19th century termed Panum's limiting case (Panum, 1858, cited in Gillam, Blackburn and Cook, 1995). Panum's observation was that, when an unpaired line presented in one eye was combined with two in the other, fusion could be achieved, and a percept invoked, that positioned the single line forward or behind the paired line. Fig 1.15 demonstrates Panum's Limiting Case.

There are diverse explanations of this effect. Several authors have explained it as a vergence error (e.g., Howard and Ohmi, 1993), or an occlusion configuration dealt with by non-stereoscopic processes (Anderson and Nakayama, 1994; Ono, Shimojo and Shibuta, 1992). Gillam et al established that the effect results from the same processes which access retinal disparity, generates very similar robust distance information, and therefore must involve double

fusion of the half images. This view directly contradicts the uniqueness constraint (Mckee, Bravo and Smallmen, 1995).

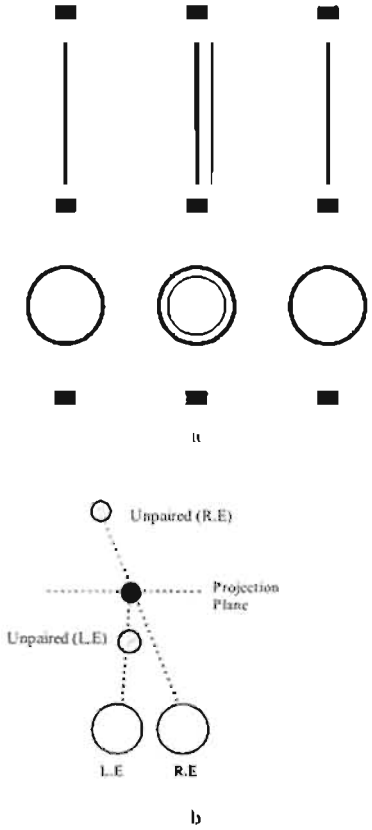



Fig 1.15. Panum's limiting case
Crossed fusion of the left top pair (a) will reveal the thin line to be nearer with depth sign reversed when fusing the right pair. Crossed fusion of the bottom left pair will yield the percept of a slanted inner circle with left edge forward in the left pair and right edge forward in the left pair. This will be reversed for un-crossed fusion. The geometric arrangement of Panum's limiting case is demonstrated schematically in (b) These figures were adapted from Gillam, Blackburn and Cook (1995) and Howard and Rogers (1995).

### Unmatchable regions of the visual field

Another problem for the achievement of retinal correspondence is the case of unpaired image features at occlusion where a near surface overlaps a distant surface in the visual field. Binocular parallax causes small regions of the distant surface to be obscured from the view of one eye. Fig 1.16 outlines the visual projection geometry underpinning this arrangement. The regions obscured from one eye are termed monocular zones.

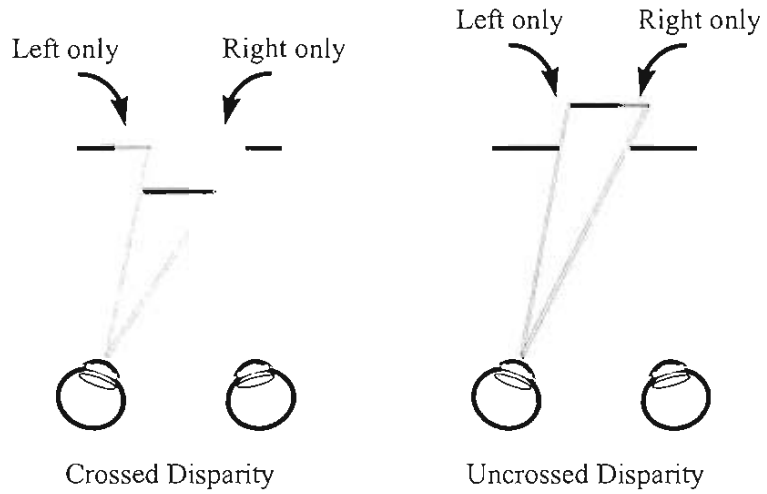



**Fig 1.16. Monocular zones in stereoscopic vision**
Binocular perspective projection means that occlusion generates unpaired regions of the distant surface where one surface overlaps another in the visual field.

Julesz' cyclopean perception dealt with these unpaired zones by invoking an *a priori* constraint termed *the most distant surface rule* (Julesz, 1971). Subsequently, most computational theorists have treated unpaired zones as residual components of the point matching process.

Nakayama and his colleagues have recently revisited this issue. Nakayama and Shimojo (1990), for example, have demonstrated what they termed *da Vinci stereopsis* a scheme in which monocular zones contribute substantively to surface perception.

In summary, this section has reviewed the capacity of binocular vision to utilise retinal disparities available due to binocular parallax—the disparate position of the vantage points from which eye accesses the optic array.

## 1.3 Patterns of retinal disparity across surfaces

This section deals with the implications of binocular parallax for the perception of textured surfaces. Once again, the optical arrangement of binocular parallax captures arrays of projection from a particular surface from slightly different directions. Because of perspective projection, precise patterns of disparity are available to stereopsis. In fact, as outlined earlier, Burt and Julesz established that the system somehow responds directly to gradients of disparity.

### 1.3.1 Stereoscopic slant

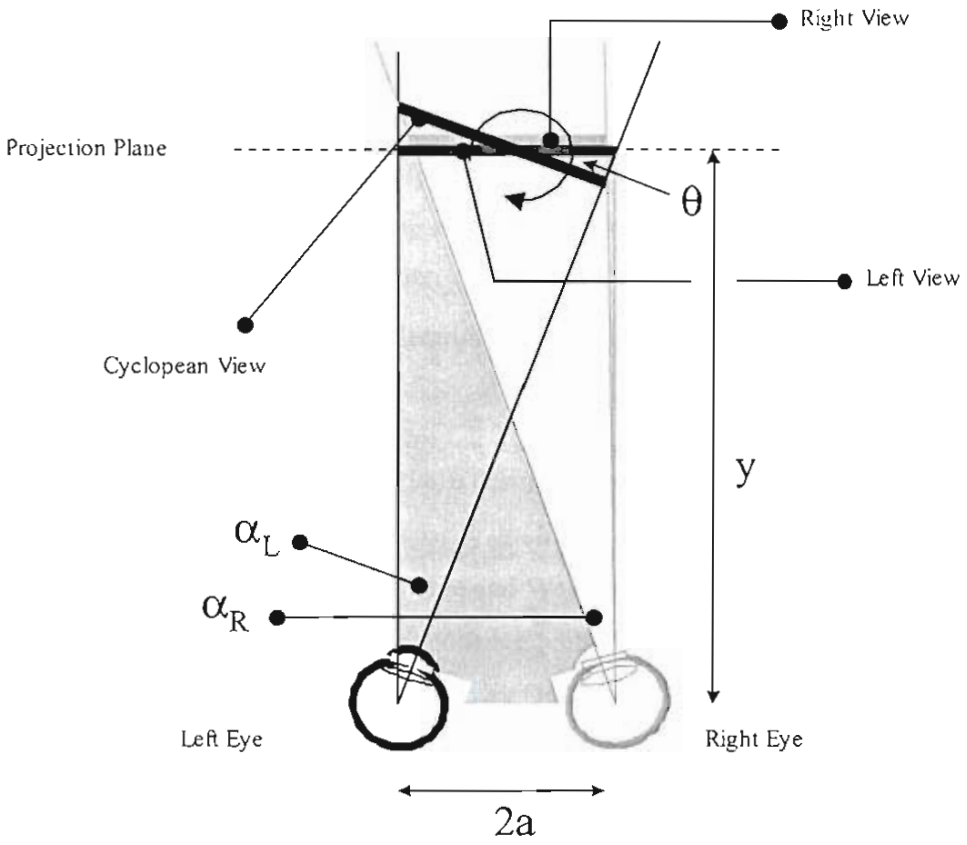
Ogle (1950) showed that magnification of one eye's view in relation to the other generated predictable stereoscopic slant: the *Geometric Effect*. The Geometric effect is a derivative of horizontal magnification of one eye's view, while the induced effect stems from a relative vertical magnification. The significance of these manipulations was that relative magnifications generated retinal disparities by systematic distortion of binocular perspective so that predictable patterns of point disparity are created. The perceptual result of relative magnification is a rotation of the cyclopean view: stereoscopic slant.

*Geometric effect*

Ogle's geometric assessment of slant using a lens over one eye led him to the following equation.

$$\tan \theta = (M-1)/2M \cdot y/a \qquad \ldots\ldots\ldots (1)$$

Where tan θ is the angle of slant in degrees, M is a magnification factor defining the differential in $\alpha_R:\alpha_L$, $y$ is the observation distance and $a$ is half the interocular distance. Fig 1.17 demonstrates the application of Ogle's formula to a textured fronto-parallel rectangle whose image is magnified in the left eye.



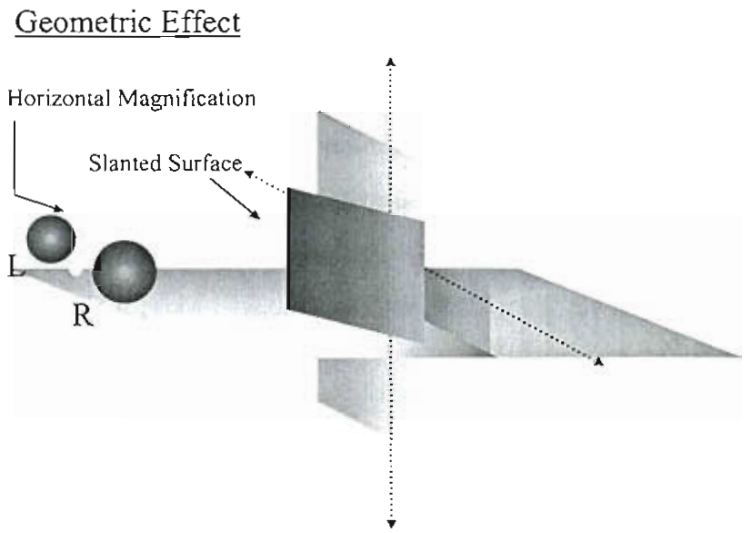




**Fig 1.17. Ogle's geometric effect**
Ogle's formula defines the relative difference in binocular angles of subtense between the left and right view required to stereoscopically rotate a surface plane about the vertical axis.

Ogle's geometric effect demonstrates that point disparities alone do not predict seen slant. The orientation of surface configuration patterns, shape, and perspective/disparity conflicts, impact upon the precision with which stereoscopic slant is resolved (Cagenello and Rogers, 1993; Rogers and Graham, 1983; Stevens and Brooks, 1988; Gillam, 1968; Gillam, Chambers and Roso, 1988; Gillam, Flagg and Finlay, 1984; Gillam and Ryan, 1992; Ryan and Gillam, 1994; Mitcheson and Mckee, 1990; Ninio, 1985; Mitcheson and Westheimer, 1984). These issues will be reviewed in some detail in Chapter 3.

*Disparity gradients*

As mentioned earlier, Burt and Julesz (1980) described evidence that the system was responsive to a gradient of horizontal disparity. Disparity gradient is a ratio of horizontal disparity to some unit of retinal distance (Bulthoff, Fahle and Wegmann, 1991). A disparity gradient essentially represents a pattern of disparity change predicted by a slanted or curved surface. In effect the whole field magnification described by Ogle must generate a disparity gradient across surface texture.

According to Howard and Rogers (1995), patterns of disparity can theoretically describe a map of image differences in terms of three orders of spatial derivative:

i) Absolute disparity, determined by the difference in angles of subtense ($\Delta\alpha$ or $\Delta\beta$) of two points is termed zero order disparity.

ii) First Order disparity, refers to a disparity gradient. This is a ratio of differences in binocular subtense between two points ($\Delta\alpha$) relative to their average separation or what is termed their *cyclopean* separation ($\Delta\phi$). So they can defined as follows.

Horizontal disparity gradient

$$\frac{\Delta\alpha}{\Delta\phi} = \frac{(\alpha_{1L} - \alpha_{1R}) - (\alpha_{2L} - \alpha_{2R})}{\{(\alpha_{1L} - \alpha_{1R}) + (\alpha_{2L} - \alpha_{2R})\}/2}$$

Vertical disparity gradient

iii) Second order disparity, also called disparity curvature, compares the disparity gradient of surface segments divided by their cyclopean separation. There are four components:

a) $\partial^2\alpha/\partial\phi_h^2$ (horizontal change of horizontal gradient in horizontal direction)

b) $\partial^2\alpha/\partial\phi_h\partial\phi_v$ (vertical change of horizontal gradient in horizontal direction)

c) $\partial^2\alpha/\partial\phi_v^2$ (vertical change of horizontal gradient in vertical direction)

d) $\partial^2\alpha/\partial\phi_v\partial\phi_h$ (horizontal change of horizontal gradient in vertical direction)

Vertical disparities are also presumed to be important in scaling horizontal gradients (Gillam and Lawergren, 1983; Rogers and Bradshaw, 1993; Rogers and Bradshaw, 1995). Further, Rogers and Cagenello (1989), argued that disparity curvatures or 2nd order disparities could be utilised in binocular vision by comparing the disparity gradient across two patches of a surface. Figure 1.18 below schematically demonstrates some examples of 1st and 2nd order disparity. 1st order gradient describes a planar surface, while 2nd order gradient describes change in 1st order gradients across a surface.

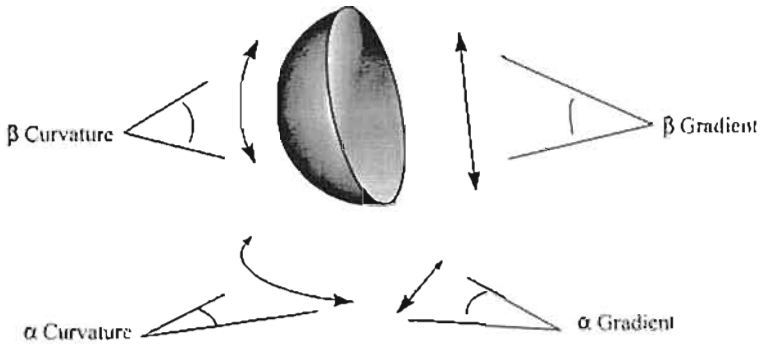



**Fig 1.18. Disparity curvature**
Three orders of disparity describe difference information available across a surface. Disparity gradients provide a measure of disparity change across a surface texture in a given vertical or horizontal dimension.

### 1.3.2 A Contemporary Model of Disparity Processing

This chapter has presented a very brief summary of relevant issues in stereoscopic vision, focusing on the information contained within Wheatstone's perspective drawings. In natural scenes stereoscopic vision evidently has access to an array of possible difference measures. An attempt to model the full utility of pictorial depth cues, proprioception, vertical and horizontal disparities in stereopsis has been recently attempted by Garding, Porrill, Mayhew and Frisby (1995- see Fig 1.19).

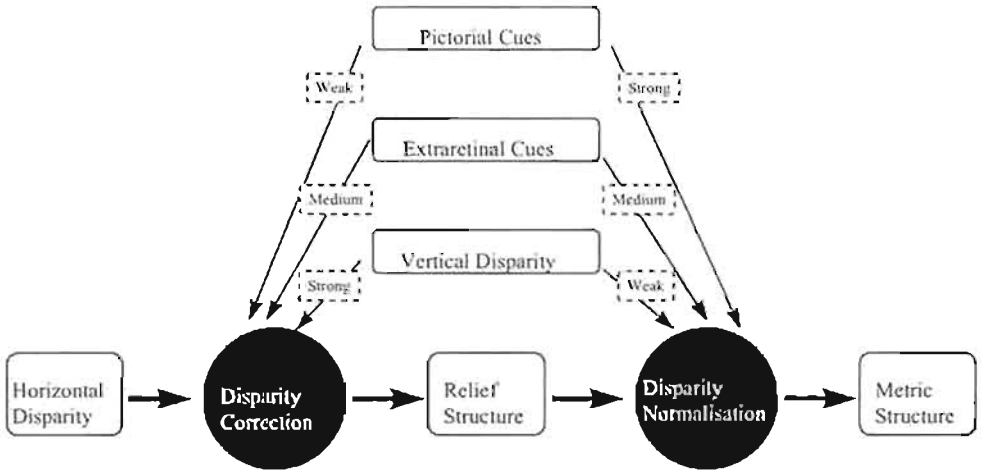



Fig 1.19. Regional disparity correction model

Garding, Porrill, Mayhew and Frisby propose a sequential decomposition of disparity information in to two separate processing schemes: *disparity correction* (computation of shape up to a basic relief estimation) and *disparity normalisation* (resolution of the ambiguity of this estimation).

The model of Garding et al integrates many advances in the understanding of stereoscopic vision over the last century. The validity of these kinds of models will be subject to extensive testing and revision of course. It is now clear that stereoscopic vision responds to a complex pattern of 2-D retinal disparities. The depth percept is not computed from point-disparities alone. My point is that recent theoretical approaches to stereoscopic illusory contours and surfaces may have taken an overly simplistic view of the nature of retinal disparity processing.

# 2. 2-D Illusory Contours

*Summary: This chapter introduces concepts surrounding the perception of illusory contours studied in 2-D. An enormous amount of research has examined these phenomena. A review of the phenomenological, psychophysical and physiological debates most relevant to later discussion is presented.*

## 2.1 Perceptual organisation in the induction of illusory contours

Kanizsa (1955, 1976, 1979) has suggested that illusory contours are accompanied firstly by regional change in lightness or appearance (Erscheinungweise); and secondly by displacement in depth generated by the appearance of figure-ground interpolation or occlusion. He argued that, when the stimulus configuration was 'optimal', image characteristics acquired modal character, that is, the boundaries of an occluding figure became visible (in the way that the rectangles in Fig 2.1 are visible).

Kanizsa presented two crucial integrative processes: *modal* and *amodal completion*. These were drawn from traditional Gestalt theory. Modal completion involved *production* of visible contours in the absence of localised retinal stimulation. Kanizsa described amodal completion as *the type of perceptual existence that is not verified by any sensory modality.* Figure 2.1 demonstrates modal and amodal completion in the generation of illusory contours. In Fig 2.1a, completion *per se* is not a difficult issue. The system can detect retinal patterns that segregate four circular regions of opposite contrast polarity to the ground luminance.

When a grey rectangle is drawn so that it's corners overlap those disks, as in Fig 2.1b, the rectangle clearly appears to lie in front of those disks. Note that the circular shape of the disk is now a matter of interpretation — a *complete* circle is not physically evident. The pattern of stimulation that these contours induce on the retinae is not strictly circular, though most observers will describe them as such. This is the notion of 'amodal completion'. Likewise, in Fig 2.1c, when disks are drawn overlaying the rectangle, it is amodally completed. Koffka termed these processes perceptual judgments— immediate recognition of an organisation or structure in a simple stimulus. Completion is, then, a rapid perceptual inference.

In Fig 2.1d, a white rectangle has been drawn to overlay the disks. In terms of retinal stimulation, there is no physical difference between the luminance of the rectangle and its background. The disk cut-outs somehow signal, or are interpreted, in such a way that they

generate visible connections across physically homogeneous space, resulting in perception of a rectangular figure defined by contours of enhanced lightness. These mark the boundaries of a rectangle. The figure exemplifies modal completion. If one fixates for sometime on the illusory boundary of the rectangle, it quickly fades.

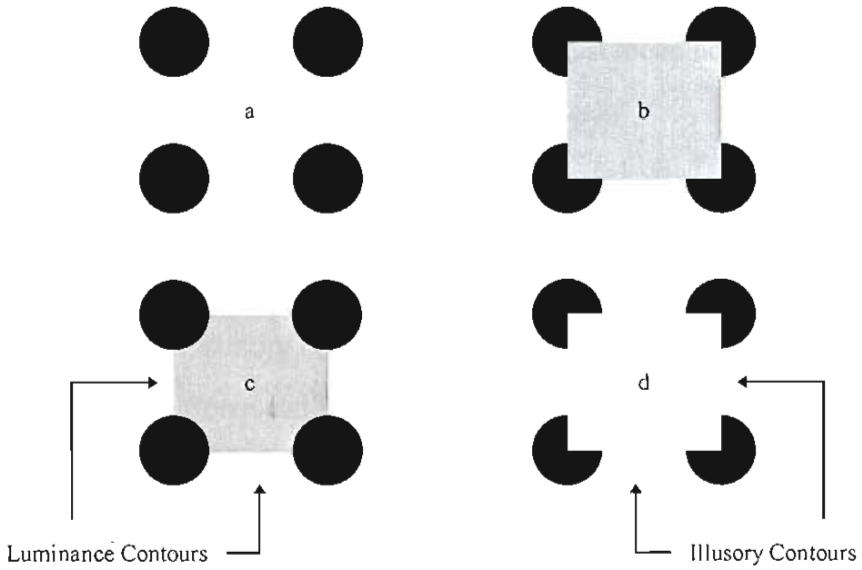



**Fig 2.1. Interposition amodal and modal completion**
In (a), the disks are complete while, in (b), they tend to be amodally completed. Here the square is most often described as occluding the grey disks. The order of interposition is phenomenally reversed in (c) while ,in (d), illusory contours modally complete an interposed rectangle.

In summary, Kanizsa claimed that illusory contours arose as the disks in Fig 2.1d were completed behind interposed cut-out sectors. In the same *inferential* and organisational process, the cut-outs were modally completed to yield visible borders bounding a figure lighter than its surrounds. Kanizsa did not identify specific stages or levels at which this organisation might be achieved. Inference was not stage or process specific, and surface interpolation invoked illusory contours conditional upon stimulus conditions that supported the amodal completion of occluded features. Interposition, here, refers to the relative stratification of objects or forms in depth:

> In my view, the phenomena of amodal completion are of special interest because they constitute an area particularly adapted for studying how the optical system goes beyond the information given. If we consider amodal completion as the result of a process of inference ... the analysis of the ways in which this is realised may allow us to discover the "logic" that these inferential processes follow. Or, as I would prefer to say, the logic of the phenomenal construction of reality. Such logic is perhaps not the same as the logic that the mind employs in making true inference (Kanizsa, 1979, p.7).

Subsequent research strongly suggests that completion and interpolation are not necessary to the generation of illusory contours.

### 2.1.1 On the independence of illusory contours and interposition

Coren and his colleagues (1972; Coren and Porac, 1983) used the Kanizsa square to show that illusory contours are reliably judged to bound regions of "near" figure, as opposed to the "far" ground of the inducing disks. Coren suggested that the cut-out sectors in Kanizsa's figures represented 'cues-to-depth', that is, implicit cues to interposition. He described such a cue as:

> ... some aspect of a configuration which can be defined as consistent with a given spatial arrangement of objects at different relative distances (Coren and Porac, 1983, p. 365).

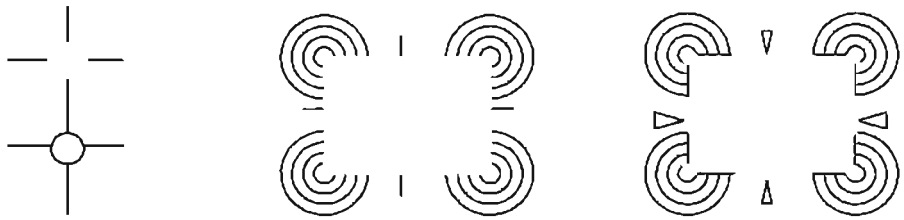
In this view, illusory contours arise from perceptual stratification in depth between figure and ground. Coren (1972) placed a small disk on the illusory surface between cut-outs. To his subjects, the disk appeared slightly smaller than one placed on the adjacent ground. Further, the difference in perceived size was greater for illusory versus luminance contours. Size-distance scaling was apparently triggered by the depth differential between figure and ground. In other words, the disk appeared nearer, its retinal subtense of the disk was interpreted as relatively smaller.

Kennedy (1975, 1976, 1981), Jory and Day (1979), and Day and Kasperczyk (1983) demonstrated that illusory contours can be seen when the inducing elements and bounded figures are phenomenally *coplanar*, that is, where there is no detectable depth step. There is good evidence, then, that illusory contours and concomitant illusory brightness differentials are

not contingent upon interposition (Ramachandran, Ruskin, Cobb, and Rogers-Ramachandran, 1994).

### 2.1.2 On the independence of illusory contours and amodal completion

Ehrenstein (1941, cited in Spillman and Dresp, 1995) and Kanizsa (1979), demonstrated that lightness enhancement is negated in lined stimuli by removing the possibility of amodal completion at the inducing edges of the configuration. In Fig 2.2a, enclosing the central region precludes lightness enhancement. The same is demonstrated in Figs 2.2b and 2.2c. These figures also demonstrate that the illusory effects are not contingent upon large regions of contrast differentiation.



a b c

**Fig 2.2. Lightness differentiation is obviated by prohibiting amodal completion**

(a) The Ehrenstein cross yields lightness enhancement that is obviated by luminance boundaries.

(b) Concentric line inducers generate clear lightness enhancement.

(c) Lightness enhancement is once again obviated by luminance boundaries.

However, many examples of illusory contours and lightness enhancement effects have shown that amodal completion may not be the necessary condition for illusory contour generation. In Fig 2.3, three examples of line arrangement effects are given.

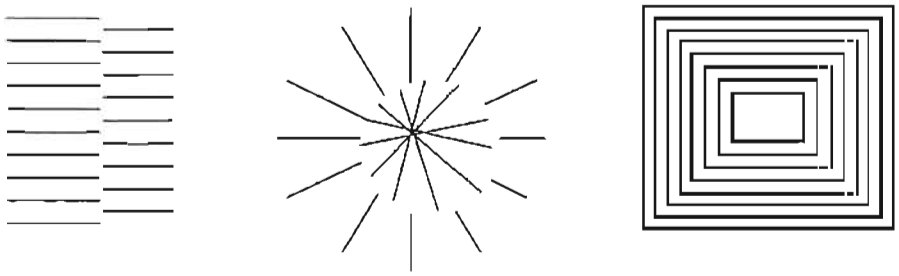



Fig 2.3. Illusory contours in line stimuli
(a) Similarly misaligned abutting line-ends yield strong illusory contours
(b) Misalignment of a central section of line ends
(c) Contours prescribed by a change in orientation

A series of influential studies by Day and Jory putatively demonstrated that the shape and intensity of illusory lightness need not involve interposition or completion. *Local* phenomenal interactions with dots placed within a Kofka cross clearly affect the dispersion or *spread* of illusory contours and the perceived lightness in these figures (see Fig 2.4).

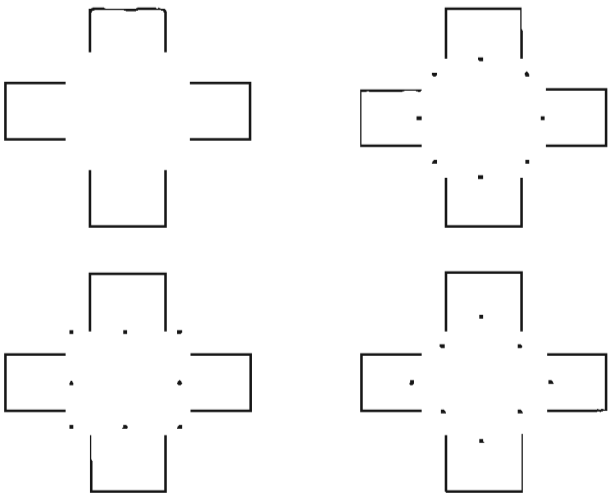


Fig 2.4. Lightness effects derived from the positioning of dots
The position of luminance dots within the Koffka cross manipulates the shape of the contour suggesting that local interactions are in operation
(adapted from Day, 1987).

In a similar vein, Kennedy, (1975, 1976, 1981, 1987, 1988, 1991) extensively explored the effects of shape and arrangement of line-ends on the perceived shape of illusory contours

beyond those of higher order completion. Selected examples of Kennedy's demonstrations are shown in Fig 2.5.

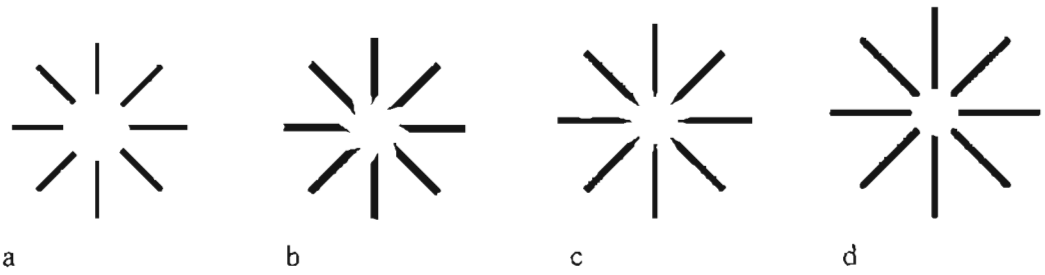



Fig 2.5. Kennedy's line-end effects
(a) Curvilinear, (b) angled (c) symmetrically pointed and (d) rounded line-ends generate markedly different effects. (Adapted from Kennedy 1987, 1991).

Thus, contrary to Kanizsa's original explanation, lightness enhancement in a physically homogeneous region may be independent of, but complementary to, the mechanisms of completion (Purghe and Coren, 1992; Purge, 1991; Albert, 1993; Davi, Pinna and Sambin, 1992; Sambin, 1987; Minguzzini, 1987). There are clearly local aspects of the configuration, for example, the shape of inducing boundaries and local non-connected luminance boundaries, which modulate the appearance of illusory contours.

### 2.1.3 Higher-order postulates

Gregory has consistently maintained that Kanizsa's figures induce *fictional* contours (1972, 1973, 1974, 1980, 1986, 1987, 1992, 1998). He argues they are the result of low-level cognition. According to this view, a local absence of physical stimulation (ie a gap) is interpreted as an object or edge partially overlaying the inducing elements— a cognitive solution to fragmentary stimuli (see also, Rock and Anson, 1979; Rock, 1987; Bradely, 1987). Bonaiuto, Giannini and Bonaiuto (1991), for example, used sketches of active human forms to generate illusory contours. They argued that only cognitive interpretation of the forms and their actions could have resulted in those contours. These configurations do not rule out the possibility that continuation of luminance boundaries might be responsible.

It has been argued, to the contrary, that illusory contours are generated before any such inferences could possibly be achieved (Ullman, 1976; Marr 1982; Grossberg 1994). Still, higher order influences have been shown to impact upon the detection of illusory contours *in certain images*. Both attention and perceptual set have been clearly implicated in moderating

detection thresholds of the contours in Kanizsa's configurations (Halpern 1985; Coren, 1987; Coren, Porac, Theodor, 1986; Tsal, Meiran and Lavie, 1994).

## 2.2 Psychophysical Parameters of Illusory Contours

It is important to note that that illusory contours and surface segmentation effects are not all-or-none phenomena. Rather they appear to systematically differ in 'strength' (degree of lightness differentiation) and 'persistence' (longevity of lightness differentiation), depending upon quantifiable physical parameters of the stimulus and *time*. These issues are addressed below.

### 2.2.1 Strength of Illusory Contours

The apparent strength of illusory contours depend on the degree of contrast variation between inducing elements and the homogenous surrounds; their number; their width; and the magnitude of the 'gap' separating them (Lesher and Mingolla, 1993; Petry et al 1983; Purghe and Katsaras 1991; Siegal and Petry 1991; Shipley, 1988). Shipley and Kellman (1992) have argued that the clarity, or distinctiveness, of illusory contours depends on the *support ratio* inherent in a stimulus configuration. The support ratio is the ratio of edge length to the total distance between inducing elements.

### 2.2.2 Onset and Persistence of Illusory Contours

Contour onset is very rapid. They are detectable at somewhere between 30 and 100ms, depending upon the spatial configuration (Reynolds, 1981). Reynolds suggests that illusory contours develop over 100ms or less; are stable for 200ms; and then fade. Meyer and Ming (1988) have demonstrated that contour persistence is directly proportional to presentation time up to 200ms. However, subtle relationships between inducing features are apparent: Kojo, Liinasuo and Rovamo (1993) have shown that increases in the proportional size of inducers increases the persistence of contours significantly. Siegal and Petry (1991) found that contour strength and persistence is maximal at a flicker rate of 5-7 Hz and at about 3degrees of peripheral eccentricity.

In summary, psychophysical findings suggest that inducing area algebraically contribute to the strength and persistence. So if the area of an interposed illusory figure is maintained and the size of the inducers is reduced the relative intensity of the contours will be reduced. An extreme example of this is demonstrated in Fig 2.6.

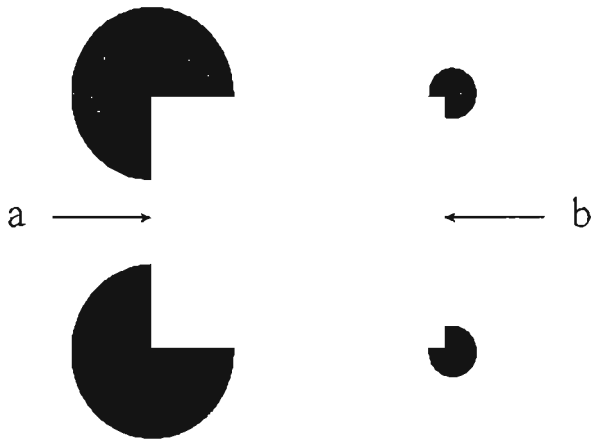



Fig 2.6. Inducing area and the strength of illusory contours
The intensity of the illusory contour at (a) is greater than that at (b). This has been explained in terms of the ratio of gap size to the inducing area width.

### 2.2.3 A caveat for the induction area hypothesis

While there is evidence that the ratio of gap size to the width of the inducing region is important, the issue is substantially more complex. Gillam (1987) has shown that perceptual segregation of surface planes is stronger when abutting grids are comprised of irregular—rather than regular—lines (see Fig 2.7).

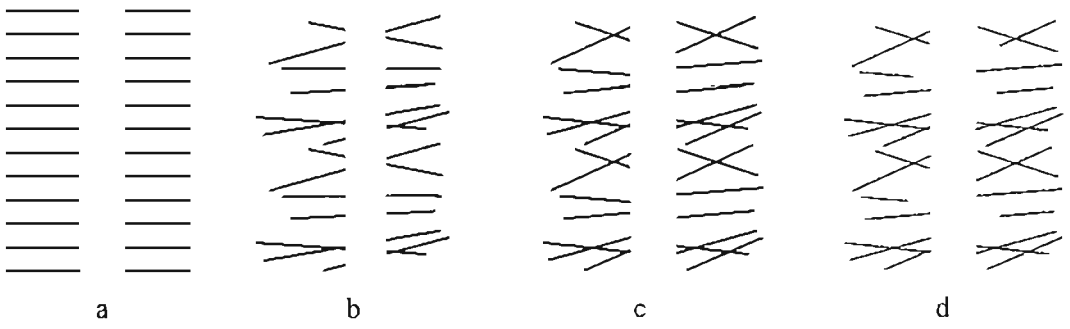



Fig 2.7. Irregular lines enhance the sense of contour
Gillam showed that regions differentiated by random-lines generated strong illusory contours. Compared to evenly arranged lines separated across a gap, as in (a), edges of random line surfaces appear more distinct (see 12.8b). In (c), co linearity across the gap is no longer maintained. In (d), despite the misalignment of irregular lines along edges, the bounding contour completes as a straight edge.

## 2.3 Illusory Contours and contrast interactions

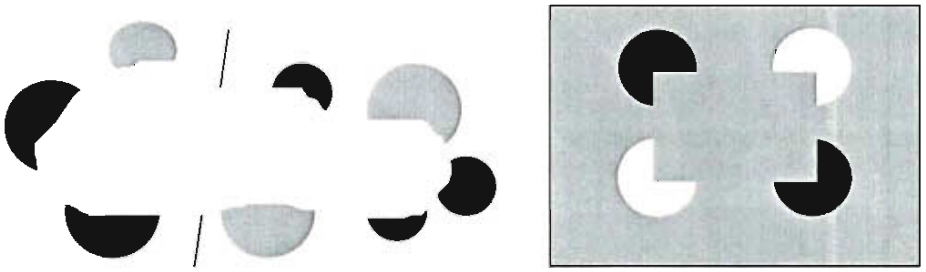
Both regional lightness enhancement and illusory contours disappear when equiluminant colour contrasts are used. When the Kanizsa triangle is isoluminant red and green (no mean luminance difference) no lightness enhancement results (Brigner and Gallagher 1974; Frisby and Clatworthy 1975; Spillman and Dresp 1995; Gregory 1977; Livingston and Hubel, 1987). Chao-Yi and Kun (1995) claim that a minimum luminance contrast required to induce visible contour was 2.3 % for the Kanizsa figure and 5.3% for abutting gratings.

This suggests that illusory brightness may be induced, or at least initiated, very early in the visual system by a luminance-contrast mechanism. This may happen in a similar way to simultaneous brightness-contrast—that is, by lateral inhibition. However, the Kanizsa and Ehrenstein figures can be perceived with dichoptic presentation of their component parts (Gregory 1972; Spillman et al 1976). This suggests that illusory lightness is, or can be, induced beyond the level of monocular visual processing, that is, in the early cortical regions. Hence, area V2 of the striate cortex—the earliest cortical region in which monocular retinal output is combined—is often nominated as the site of illusory contour generation. This possibility is explored in a later section, which presents findings on the proposed physiological substrates of illusory contour generation.

Gegenfurtner, Brown and Rieger (1996) have recently shown that form processing may not be interrupted by the lack of lightness differentiation in isoluminant figure-ground stimuli. They argue that a coarse-to-fine boundary segmentation process may be in operation. This is in line with computational theorists such as Marr (1982). Using a visual search paradigm, these authors found that without the presence of distinct illusory contours in Kanizsa figures (at isoluminance), the configurations were still rapidly detected as segmented figure and ground. This finding suggests that mechanisms segregating figure and ground are responsive to the structural arrangement of image features and not to luminance contrast alone.

### 2.3.2 The problem of contrast polarity

Further evidence of closure or completion mechanisms in some way independent of contrast detection mechanisms is provided by Prazdny (1983). Prazdny demonstrated clear contours bridging gaps between elements of opposite contrast polarity. This observation is not congruent with the traditional understanding of simultaneous contrast. Two examples are shown in Fig 2.8.



a b

**Fig 2.8. Illusory contours and contrast polarity**
Both Kanizsa (a) and Prazdny (b) have shown that disks of opposite contrast polarity yield clear illusory connections.

Dresp, Salvano-Pardieu and Bonnet, (1996) have also recently demonstrated that illusory boundaries are evident in abutting grids with inducers of opposite contrast polarity using phase shifted line gratings. Dresp has also claimed that the figure-ground context does not play an important role in the generation of lightness enhancement (Dresp, 1993). Dresp and Bonnet (1991) and Dresp, Lorenceau and Bonnet (1990) found a degree of lightness enhancement at the centre of a Kanizsa square in the complete absence of the cut-out disks (as in Fig 2.1a).

In summary, as Shapley and Gordon (1987) have suggested, form (completion) and lightness differential must be processed separately—or at least are able to be generated independently. In other words, modal completion is arguably the result of a confluence of integrative and local level interactions which can be blind to contrast polarities as such.

### 2.3.3 Spatio-temporal effects

Bradely and Lee (1982), and Kellman and Cohen, (1984) have demonstrated that illusory figures can be detected when inducing features are displayed in rapid succession at differing orientations. This work suggested that a spatio-temporal correlation was possible between temporally disparate edge boundaries. Prazdny (1986), using flowing random dot fields instead of contrasting inducers, generated clearly identifiable illusory shapes. Fig 2.9 demonstrates temporally disparate stimuli.

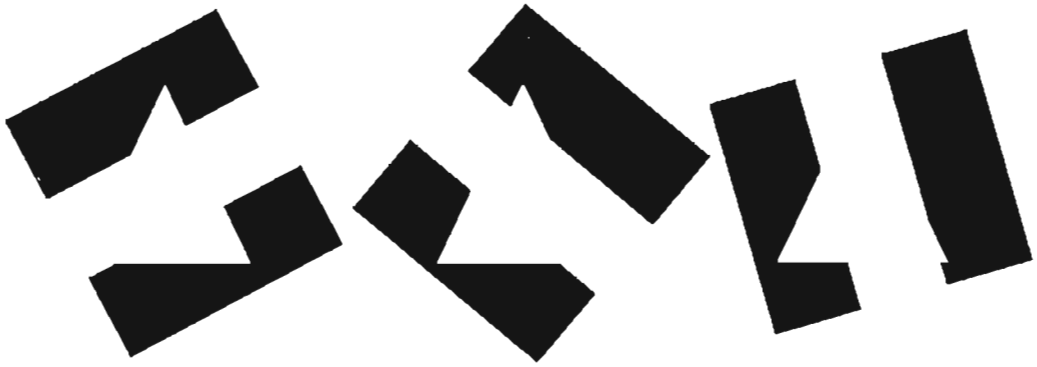


**Fig 2.9. Edge correlation in time**
Presentation of these three stimuli in rapid succession yields the percept of a strong illusory triangle.

Further, Prazdny demonstrated that, in a random dot field, motion of either the inducers or the triangle would result in the appearance of a strongly segregated object. Clearly, illusory contours can be generated by luminance contrast or motion.

Ramachandran (1986), Ramachandran and Cavanagh (1985) have shown that illusory figures can capture certain bounded elements in depth and/or in motion. Further, capture breaks down in isoluminant stimuli. Some examples used to show capture effects are shown in Fig 2.10.

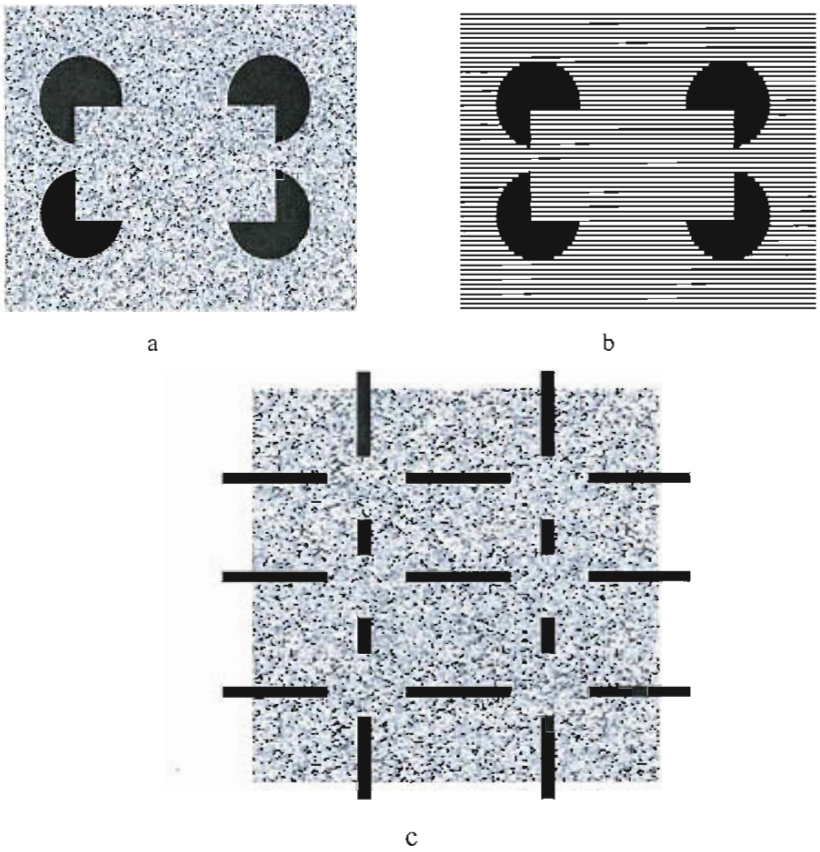



**Fig 2.10. Surface capture effects**

**(a)** Prazdny (1986), using inducers of opposite contrast sign, demonstrated that apparent motion of a surface can be seen when illusory contours are minimal

**(b)** Ramachandran (1986) demonstrated that apparent motion or stereopsis 'captured' the background patterns to the depth of the cut-out sectors.

**(c)** Spillman 1977, demonstrated that dragging an Ehrestein Grid across a random surface led to the sense that the dots moved with the cut-out portions of the grid and, viewed statically, were perceived as qualitatively different from their surrounds.

Ramachandran has interpreted these phenomena as evidence that the brain takes a series of short-cuts when interpreting 'captured' dots, lines or other surface features in motion, and in depth. Figure 2.11 illustrates a stimulus used by Ramachandran in his tests of stroboscopic movement of the illusory square. Observers report a form moving between these two background configurations.

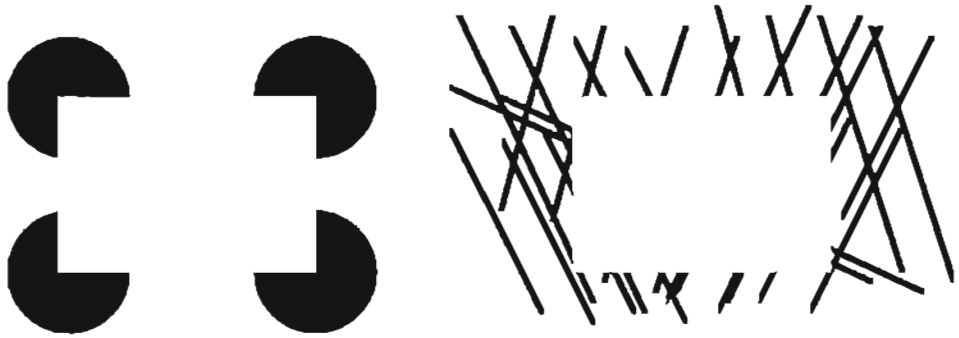
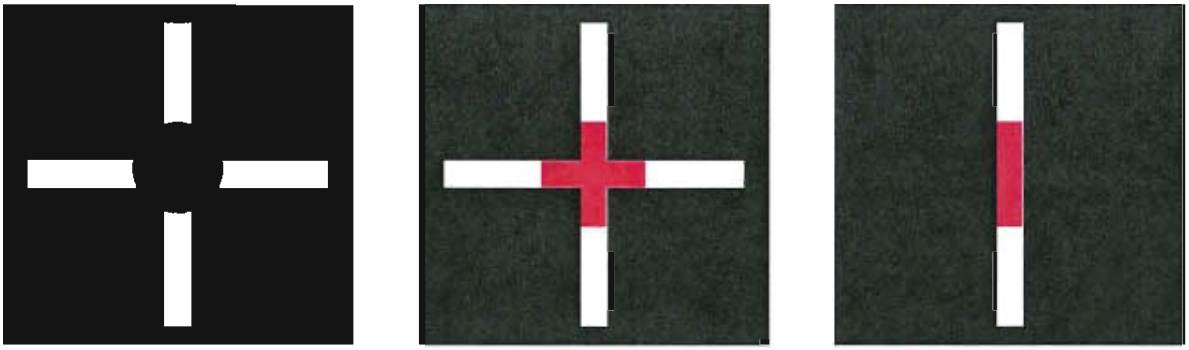


Fig 2.11. Illusory contours and apparent motion
Ramachandran demonstrated that stroboscopic motion was perceived between different types of inducing elements.

There are clearly various domains of surface segmentation which can contribute to the extraction of form from the stimuli that Kanizsa presented: in combination or in isolation. Some regional differentiation in contrast, motion, distance or complexity is clearly required to generate illusory surfaces, although none of these can be considered to be necessary. An information *primitive* has evaded specification and generality. This strongly suggests that a constructive visual boundary mechanism is in operation which utilises physical, spatial and temporal attributes to evoke Gestalt-like organisations. The mechanism can utilise relative motion or colour, rather than steep luminance gradients as a basis for figure-ground segregation.

Even figure-ground segregation though is insufficient to explain the various local contrast interactions (Shapely and Gordon, 1987). It seems that illusory contours represent a class of perceptual phenomena which may be evoked at many possible levels and domains of visual analysis (Day, 1987). Consider *neon colour spreading*, first demonstrated by van Tuijl (1975). Redies and Spillman (1981) showed that, when a red cross was inserted in the centre of an Ehrenstein figure which yields illusory contour, neon colour spreading or neon flank results (after Takeichi, Shimojo and Watanabe, 1992). Fig 2.12 demonstrates these effects (though *neon flank* is a very subtle effect, not easily reproduced).



a b c

**Fig 2.12. Neon colour spreading and neon flank**
(a) An Ehrenstein cross yields illusory boundaries

(b) Neon colour spreading arises when the centre of the crossed is filled with colour or luminance contrast. The colour appears to spread into the region previously bounded by the illusory contour. This is suggestive of completion

(c) Neon flank occurs where only one arm of the Ehrenstein cross is presented. Colour is still perceived to spread into its immediate surrounds. This is suggestive of a discrete local interaction

In summary, mechanisms which generate illusory contours apparently are blind to the polarity of contrast at their boundaries. It appears that a more abstract notion of regional differentiation may be necessary to found a comprehensive theory of illusory contours. The next section outlines a recent attempt to broach this issue.

## 2.4 Figure-ground splitting: perceptual unit formation

Kellman and Shipley (1991), recently presented an extensive theory of *Visual Interpolation* which attempts to integrate some of the diverse findings about illusory figures. They tried account for Gestalt principles of grouping and the closure. Kellman and Shipley argue that specifiable geometric principles guide a set of abstract mechanistic process in what they term *unit formation*.

Kellman and Shipley believe that continuous geometric projection across the occluded space, between spatially *relatable* contours, is the key to understanding how this unique identity is constructed or recognised. Such visual events are, in a sense, reminiscent of the classic reversible figure-ground stimuli of Rubin. In reversibly interpolated figures, however, a switch is possible between modal and amodal completion (as in Fig 2.13).


**Fig 2.13. Two-dimensional unit formation**
Kellman and Shipley presented figure (a), a 'spontaneously splitting figure. In (b), Rubin's classic reversible figure demonstrates similar segmentation processes. Staring at each figure for some time will result in a reversal of figure-ground, but the contours defining boundaries remain invariant.

Unit formation can result in visible contours when spatial discontinuity conditions are optimal. Kellman and Shipley suggest that modal and amodal completion are in effect two aspects of a single completion or 'unit' formation process. Where these processes are interchangeable, the figures will reverse, where they are not, stable illusory contours may form. The difference is that these percepts are in a sense *driven* by the implications of different relative depth relationships (see Fig 2.14)

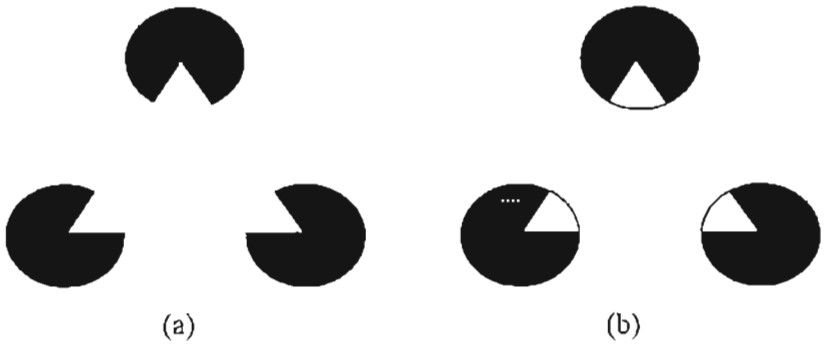



**Fig 2.14. Modal and amodal completion: Unit formation**
Kellman and Shipley argue that modal completion (a) and amodal completion (b) are alternative products of continuation of luminance contours between discontinuities rather than integration and synthesis of form.

Kellman and Shipley's theory is briefly summarised below. Their argument hinges upon three concepts: spatial discontinuity, spatial relatability and the monotonicity constraint.

*Spatial Discontinuity*

Spatial continuity refers to curved uniform edges or straight lines such as the boundaries of a circle, an oblong or some similar figure. An example of discontinuity might be the corners of a square, or the boundaries of the cut-out disks in Figure 2.14a above. Discontinuity is considered to be a necessary but insufficient feature to generate unit formation / segregation. A further spatial condition is required.

*Spatial Relatability*

*Spatial relatability* refers to the reciprocal concept of *spatial discontinuity*. This arises where one edge is 'relatable' to another by a smooth linear extension: across a 'spatial continuity'. The modal (Figure 2.14a) or amodal (Figure 2.14b) connections between the adjacent inner borders of the cut-out sectors in the Kanizsa triangle are examples of relatable contours. Edges are relatable only if their extensions intersect at 90 degrees or less. However, this condition, together with the notion of spatial discontinuity, are again insufficient to generate the perception of interpolation.

*Monotonicity Constraint*

The final condition required to yield an occlusion is what is termed the *monotonicity constraint*. It amounts to an assumption about what contours will do when they cross above or behind an occluded or occluding surface. The assumption is that they will adhere to 'spatial continuity', that is, they will not loop back on themselves (as is entirely possible but unlikely in natural scenes). Monotonicity is therefore an assumption that relatability is verifiable and does not support the kind of occlusion seen in Fig 2.15.

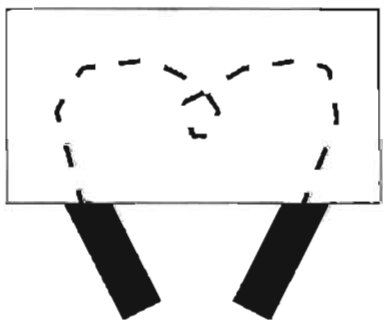


**Fig 2.15. The Monotonicity Constraint**
In Kellman and Shipley's theory, continuation is constrained by the monotonicity constraint which precludes continuation beyond $90^0$.

## 2.4.2 Advantages of interpolation theory

Kellman and Shipley's is a remarkably simple explanation of illusory form, and interpolation processes in general. It is a very powerful description of the characteristics of the Kanizsa figures, and other occlusion effects which appear to represent *emergent forms*. For example, in Fig 2.16, spatial continuity and discontinuity can explain how it is that (a) a figure can appear partly modally and partly amodally completed; (b) can be generated in the absence of a closed figure; in (c) the process of relatability can resolve indeterminate texture features; and in (d) can resolve occlusion where amodal completion of the 'background' disks is indeterminate as well

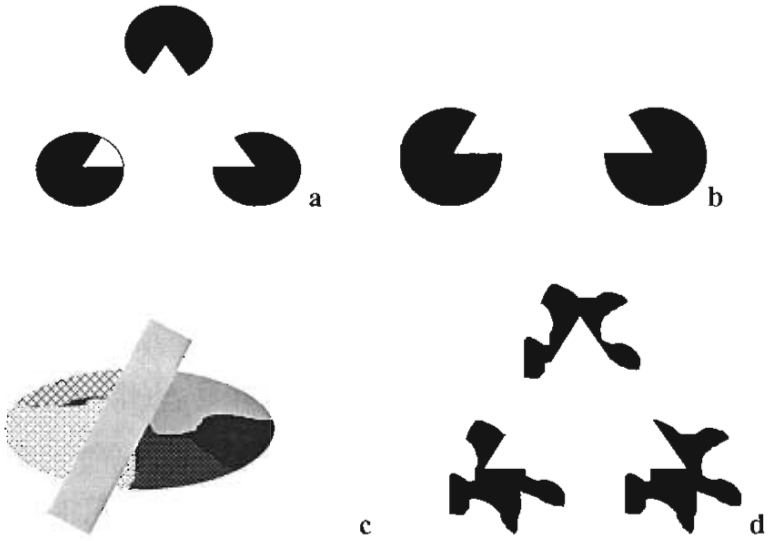



**Fig 2.16. Spatial relations and continuation**
In (a), a complete figure can be interpreted from a combination of modal and amodal contours while, in (b), continuity between inducers generates contours in the absence of completion. In (c), figure-ground splitting arises regardless of local contrast data. Figure (d) demonstrates that irregular shapes generate illusory contours in figures where continuity arises despite abstract irregular inducing features.

Shipley and Kellman (1990, 1992) have argued that unit formation and discontinuities contribute to every instance of illusory contour formation. A single boundary interpolation process unifies partly occluded objects and illusory figures: discontinuities in the first derivative of projected edges are initiating conditions for unit formation. By this relatability approach, physical edges can be connected to non-physical. Some criticisms of this view from Sekulear, Palmer and Flynn (1994) suggest that the observations of unit formation cannot be explained without global or configurational interactions. Perhaps most intriguing is the work of Anderson (1994), Anderson and Julesz (1995) and Anderson (1997) where, in the domain of stereopsis, it is possible to derive illusory contours where no relatable contours are present. Recall, also,

Gillam's demonstration that certain texture patterns can intensify the strength of illusory contours and that continuity is not necessary to invoke the sense of illusory boundaries. Once again, it appears that illusory contours defy general explanation.

## 2.5 Physiological approaches to illusory contours

### 2.5.1 Spatial frequency analysis and figural completion

The assumption that illusory contours arose in the absence of stimulation was hotly disputed by Ginsburg (1975, 1987). Ginsburg claimed that spatial frequency filtering of visual images prescribed by Kanizsa's figures creates a physical intensity pattern that stimulates the system's low spatial frequency detection mechanisms directly.

Spatial filtering is thought to be a function of the receptive field structure at the retina and higher level orientation-sensitive cells of the visual system (Goldstein, 1996). In computational simulations of early vision, Ginsburg (1975) was able to demonstrate in that a Fourier transform of low spatial frequency luminance intensities would parse an equivalent shape to that prescribed by the Kanizsa triangle. Ginsburg concluded that low-spatial-frequency attenuation in the visual system might aid in the induction of illusory objects by a mechanistic information-reduction scheme. These arguments have been highly influential in a number object recognition schemes based on a computational resolution of intensity values across the visual field (*cf.* Marr, 1982; Marr and Poggio, 1979).

The spatial frequency approach has been criticised, however, because many figures have been produced which can not be resolved by Fourier analysis (Tyler 1977, Becker and Knopp 1978; Parks and Pendergrass 1982, Parks, 1983). An example is shown in Fig 2.17. Parks and Pendergrass demonstrated that low spatial frequency filtering can not resolve the corner arrangements of the simple figure below.

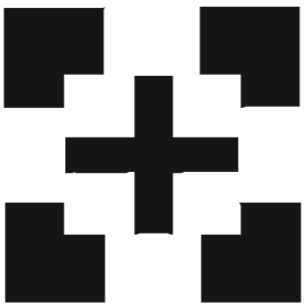


Fig 2.17. An image which defies spatial frequency analysis.
(Adapted from Parks and Pendergrass, 1982).

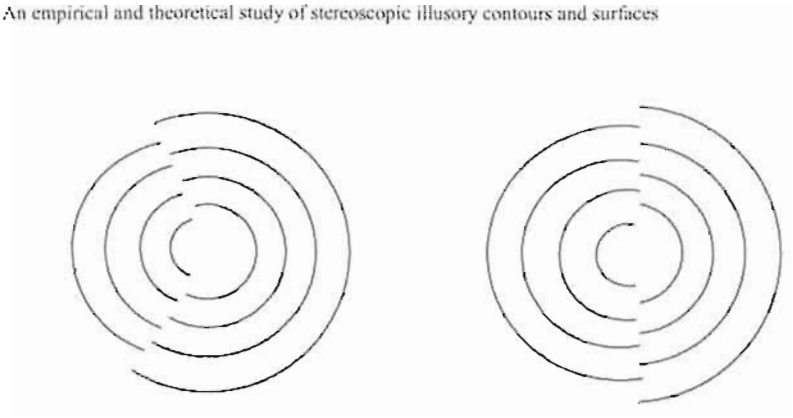
### 2.5.2 Neural corollaries of illusory contour induction

Specific physiological correlates of illusory contour perception have been described by Peterhans and von der Heydt (eg. 1991). Studying responses of area V1 and V2 in alert monkeys, they found that no V1 cells of the striate cortex responded to stimuli that might generate illusory contours, yet 40% of V2 cells do. These were classified as *contour neurones*. Such neurones, respond to edges prescribed by collinear line ends in offset abutting gratings (see Figure 2.19a). The activity of these cells in response to 'gaps' is very similar to their response to oriented contours defined by steep luminance modulation (Peterhans and von der heydt 1991, von der Heydt and Peterhans 1989). Similarly, their activity is inhibited by inserting small closing lines, or when only one part of the stimulus is presented to the receptive field. These findings suggest that aggregation of neural mechanisms sensitive to orientation may constitute the basis of a specific category of illusory-contour percepts invoked by abutting line-gratings (Grosof, Shapley, Hawken, 1993).

Peterhans and von der Heydt suggest that at one, or both, ends of the receptive field of their *contour neurones* there is an inhibitory zone that suppresses the cell's response when a stimulus line or edge is longer then the length of the main receptive field. A reliable means of detecting edges (and illusory contours) could result from linking together line-end-sensitive mechanisms.

### 2.5.3 Functional equivalence of luminance contours and illusory contours

Most contemporary explanations of illusory contours, based on known physiology of the visual system, maintain that illusory contours are functionally equivalent to luminance defined edges. However, studies of tilt after effects (TAEs) by Smith and Over (1975, 1976), suggest nontrivial differences in the perception of illusory contours and luminance borders. Fig 2.18 exemplifies stimuli used to assess TAEs.



a b

Fig 2.18. Tilt After effects

After fixating upon the *adapting* stimulus in (a), for several seconds, then shifting gaze to the test stimulus (b) subjects typically judge the illusory contours in (b) to be oriented non-vertically.

Adaptation to a tilted stimulus typically will cause a proportional error in the estimate of a subsequently viewed non-tilting line. Tilt after effects are often cited as discriminating the orientation sensitive processes at area V2. Smith and Over have demonstrated that illusory contours yield a reduced, but reliable, TAE. However, whereas luminance contour TAEs are colour selective - illusory contours are not. Illusory contour TAEs are also interrupted by binocular rivalry, while luminance contours are not. In contrast, subthreshold summation effects (displaying a line of the same contrast polarity below detection threshold) lower the threshold for the detection of the target line. Thus, illusory and real lines are likely to be processed by the same mechanism (Dresp and Bonnet, 1995; Mcourt and Paulsen, 1994; van der Zwaan and Wenderoth, 1994, 1995).

While it seems there may a general 'contour invariance' across illusory and other contours (Laurie, Wann, Dember and Frank, 1994; Berkley, Debruyn and Orban, 1994) subtle differences in the activation of TAEs by illusory contours suggest some caution is required. Perhaps luminance and illusory contours are *organised* by the same visual processes rather than equivalent detection devices. For example luminance contours *can* represent boundaries while it seems likely that illusory contours must represent.

Illusory contours and figures appear to result, then, from the operation of standard edge-detectors in the visual system. Inferential processes may still have an important part to play in the perception. As Pradiso, Shimojo and Nakayama (1989, p.1212) note:

> While we feel that our data combined with the physiological results ... clearly demonstrate an early neural process contributing to the perception of subjective contours this conclusion does not nullify the possibility that something "inferential" is involved in their perception. Knowledge about the visual world may, through learning, be incorporated into the organisation of the visual system ... In this sense the neuronal response to a subjective contour can be thought of as an "inference" even though it is made at a very early stage of visual processing as such, the neural interconnections in [the] prestriate cortex may embody an early component of what is commonly considered cognition .

This contemporary notion of neural operation shows that the physiological correlates of illusory contour induction which might result may be a product of feedback into the lowest-cortical regions. This is a major problem for contemporary notions of neural interconnectivity. Further, it is not clear how specific edge detection devices can explain lightness spreading or the area proportionate ratio account of the *strength* of illusory contours. One obvious way in which illusory contours are functionally non-equivalent to luminance contours: they tend to fade when fixated.

Nakayama and Shimojo (1992) suggest that neural inference-like processes in very early visual processing—a mechanistic analogue of Hemholtz's notion of unconscious inference—may respond to broad patterns of stimulation. This was assumed to be based on image sampling during locomotion and previous associative learning, where a population of neurons in the cortex will respond to part induction of a previously experienced pattern of activity at the receptors. Experiential shaping of interneuron connections are the critical issue here. Previous experience means that, when faced with ambiguity as in most inducing configurations of illusory contours, the system responds to a 'generic' pattern of stimulation, rather than that which might be induced by an 'accidental' view of a particular object. Nakayama's ideas, particularly those pertaining to stereoscopic vision, will be addressed in detail in Chapter 3.

## 2.6 Computational models of illusory contour induction

Based primarily upon contemporary knowledge of the human and other primate visual cortical architecture, several computational models have made a useful contribution to our understanding of illusory-form perception In essence, illusory figures are derived from the same processes used to segment luminance defined forms from the visual scene (e.g. Marr

1982; Cohen and Grossberg 1984; Grossberg and Mingolla 1985, 1987, 1993; Grossberg and Todorovic 1988; and Grossberg 1994; Francis and Grossberg (1996). Illusory percepts are regarded as being critically important issues in understanding integrative visual information processing. As Grossberg and Mingolla (1987, p171) have explained:

> [these] ... paradoxical percepts are expressions of adaptive brain designs aimed at achieving informative visual representations of the external world. Paradoxical percepts may therefore be used as probes and tests of the mechanisms that are hypothesised to instantiate these adaptive brain designs. Illusory contour percepts in particular, provide numerous clues and constraints for a theory of boundary formation and textural segmentation, because they involve subtle interactions of form and colour processing ... our theory makes precise the sense in which perception of illusory contours - or contour percepts that do not correspond to one dimensional luminance differences in a scenic image - and percepts of 'real contours' are both synthesised by the same mechanisms.

The computational approach therefore assumes that illusory figures arise from the same mechanisms which account for form in natural visual scenes. In Grossberg's terms, form and contour are processed by separate neural network sub-systems. Interaction of these systems generates a global three dimensional representation and visual arises from combinations of feature extraction routines—that is, second order contrast interactions derived from orientation selective neural networks.

### 2.6.1 Grossberg's Computational / Physiological Approach

Grossberg's (1987) model of cortical dynamics has had a seminal influence. It is very briefly summarised here, and his contribution to the perception of three-dimensional forms will be addressed in detail in Chapter 3. His work provides an extensive treatment of the physiological processes that may underpin traditional Gestalt organisational rules and illusory percepts. Most importantly, he suggests that form, colour, and brightness characteristics of visual images are 'emergent' from the interactions of several discrete—though interacting— processing systems operating in parallel.

The pivotal assumption of Grossberg's work is that three particular systems separately deal with visual organisation and feature emergence in perception. These are a Boundary Contour System (BCS) a Feature Contour System (FCS) and an Object Recognition

System(ORS). The BCS synthesises visual form from the combinations of oriented 'feature elements' extracted very early in visual processing, yielding three dimensional closure, or segmentation of portions of a scene.

*Segmentation*, here, refers to the perceptual phenomena of uniqueness, that is, phenomenal separateness allocated to a region of the field. This is an indirect reference to the Gestalt idea of "closure" or wholeness which is evident in many figure-ground stimuli. The outcome of BCS processing is perceptually invisible. Visible percepts are a function of the FCS. The FCS extracts colour and brightness information in parallel, but separately from the BCS. So these FCS signals interact with the BCS system to control filling-in, of an illusory occluding surface for example. Subsequently these filling in processes lead to visible percept of lightness, colour and form in depth. At the final stage of functioning of the FCS the ORS, a higher order object recognition mechanism specifies the salience and meaning of the representations previously developed. As such, the FCS and BCS networks are said to be pre-attentive while the ORS system is considered a directed, post-attentive feedback system. Figure 2.19 below illustrates Grossberg's model.

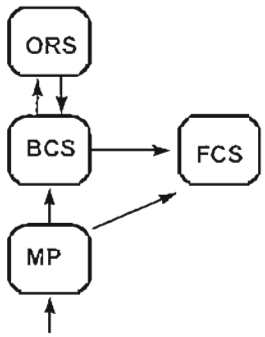



**Fig 2.19. Grossberg's processing stages**
Monocular pre-processed signals (MP) are sent independently to both the Boundary and Feature Contour Systems. The Boundary System also exchanges template like information from an Object Recognition System.

In terms of this model, Kanizsa's square is a three-dimensional representation derived from the construction of a discrete visual entity synthesised in BCS processing. Visible phenomenal characteristics of the percept are accomplished by the FCS—for example contrast interactions, colour and depth effects outlined earlier. Subsequent filling in of the contrasting disks, the various lightness effects identified earlier, opacity, or transparency of the overlayed

figure are a function of interaction of the feedback between these two systems and the ORS device.

Grossberg (1987), in effect, describes his integrative processing model as adaptive resolution of uncertainty, where *uncertainty* refers to the interpretation of retinal and pre-attentive measurements made by banks of detection devices. He endeavours to model the way an ill-posed, or ambiguous, visual world is deciphered to derive form, exploring mechanisms that can resolve multiple sources of information and uncertainty in that information pre-attentively.

The BCS activity represents a completion mechanism similar to continuation, as described by Kellman and Shipley (1991). It portrays modal and amodal completion as two possible outcomes of the same process, each dependent upon FCS filling-in.

In subsequent analyses, Gove, Francis, Grossberg, and Mingolla (1994) have extended the model to the cortical dynamics of form and motion integration, and Grossberg, and Mingolla, (1995), have attempted to map the possible connections between binocular vision, illusory contours, and brightness: a neural model used to demonstrate proposed LGN feedback integration of these elements.

## 2.7 Concluding Remarks on Perception of Illusory Contours

This chapter presented a review of research surrounding illusory contours. The phenomena are a product of processes, which differentiate regions of the visual field from one another. Regional differentiation appears to be maximal when perceptual closure is invoked in concert with steep contrast discontinuities (such as in the typical black and white Kanizsa Square). However it is also clear that closure of a perceived form, interpolation and other higher-order factors are not necessary for illusory contours. As yet, no satisfactory explanation of illusory contours has emerged from the literature. The next chapter is concerned with stereoscopic illusory contours.

# 3. 3-D illusory contours and surfaces

*Summary: This chapter surveys the diverse and somewhat discordant literature dealing with the perception of 3-D illusory contours and surfaces. The percepts generated by SKS will be addressed. Other notable stereograms and theoretical developments in this field are reviewed. In so doing a convenient partition between the various explanations of these percepts will be developed. In a very broad sense, stereoscopic illusory contours have been explained in two ways. The two approaches will be termed the Surface Heuristic and Form Computation.*

## 3.1 A stereoscopic Kanizsa square

A SKS highlights important issues to be addressed in the chapter. Figure 3.1 demonstrates some intriguing 3-D phenomenology. When faced with apparently sparse point-disparity information the visual system seems to construct an illusory 3-D percept.

Some characteristics of the 3-D Kanizsa percept are well established. First, note that stereoscopic illusory contours between the pacmen are stable and distinct. In fact they are more stable and distinct than those generated in the absence of disparity (Gregory, 1972; Bloomfield, 1973; Lawson and Gulick, 1967; Simmonds, 1975). Second, when the sign of disparity is changed, the sites at which illusory contours occur also changes. With uncrossed disparity at the pacmen, the disks take on the appearance of portholes through which an *amodal* square is perceived (Gregory and Harris, 1974; Fujita, 1993). The term amodal is used as a short hand description. In addition to these aspects, the emergence of illusory contours appears to be related to the assignment of homogeneous regions of the stereogram to different depth planes.

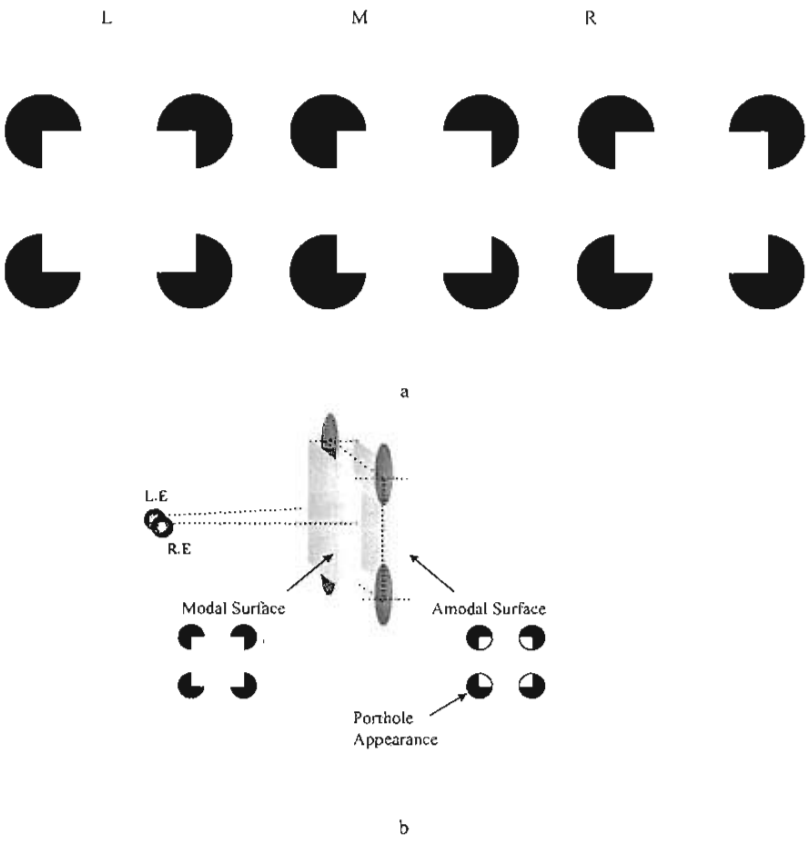



**Fig 3.1. The stereoscopic Kanizsa square**
In 3.1a crossed fusion of the left (L) and middle (M) images yields the percept of a distinctive illusory square standing forward of the disks. The white region surrounded by the illusory boundaries appears to be *captured* to the depth of those boundaries. Crossed fusion of (M) and (R) demonstrates the reverse disparity sign. An amodal square results. Disks appear like 'portholes' through which the square is partly seen. These effects are pictorially represented in (b). The sequence will be reversed for divergent fusers.

Clearly, stereopsis has dramatic implications for the perception of a 3-D Kanizsa square. Given that local disparities within the SKS image pairs are very sparse in comparison to most natural textured surfaces, the part played by stereopsis is of theoretical interest. Stereoscopic illusory contours perhaps offer a window onto the interpretative mechanisms of binocular vision ie. 3-D perceptual organisation (Nakayama, 1996).

## 3.2 Stereoscopic contours and surface textures

### 3.2.1 Stereoscopic contours in sparse texture matrices

Gulick and Lawson (1976) gave an early account of stereoscopic illusory contours. However, they did not call them illusory, but *stereoscopic*. Their argument was that stereoscopic contours

result when the sharp surface boundaries apparent in RDSs are translated onto sparse textures such as those in Fig 3.2. Horizontal interocular differences in the position of large gaps in those matrices provided what they termed *form disparity*. They defined form disparity as:

> ...the disparate views of a form which differentially obstructed elements of the ground matrix (1976, p110).

Gulick and Lawson showed that sparse matrices (see Fig 3.2a) generated sharp visible boundaries between matrix features in a similar fashion to the RDS. Perceptual organisation left specific features unmatched. These non-corresponding elements of such a matrix were important in differentiating an occluding form from its background. Unpaired features were interpreted as occluded from the view of one eye. This is shown pictorially in Fig 3.2b.

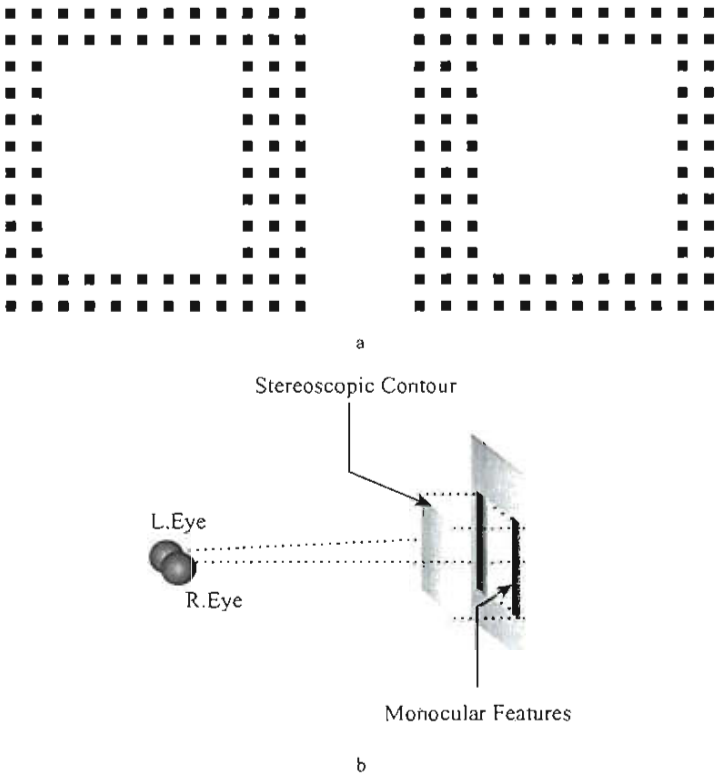



**Fig 3.2. Form disparity**
Gulick and Lawson interpreted the scission of surface planes as evidence of 'form disparity'. Fig (a) yields the appearance that a central opaque region is extracted as a bounded form. The pictorial depiction (b) shows that monocular features must be grouped with the background surface to achieve the effect.

Gulick and Lawson found that increasing the *global density* of the matrix qualitatively increased the "sense" that the homogeneous space, the large gap in the matrix, was separated

from the plane of the surrounding matrix. This, they explained, was due to reduction of ambiguity in the non-corresponding portions of the matrix. It is interesting to note that Kellman and Shipley (1992a) would attribute this effect to the size of the inducing area (see Chapter 2). Note also that Julesz invoked a *most distant surface rule* to describe how monocular features in RDS were left on the occluded plane.

Another achievement of Gulick and Lawson was their demonstration of a *spreading* stereoscopic form where edges were not explicitly demarcated by unpaired features. The effect is shown in Fig 3.3. Most observers report a misshapen circle. The point being that the illusory sector does not just join luminance dots but actually passes before them.

This demonstration shows that neither the closure of fragmentary figures (after Gregory, 1998) nor propagation of relatable contours across luminance discontinuities (Kellman and Shipley, 1991; 1992a) are sufficient to explain stereoscopic illusory contours. Some kind of *surface spreading* appears to be involved. In Fig 3.3 the white form *spreads* across the near depth plane rather than aligning with contrast boundaries. Gulick and Lawson attributed this to perceptual aggregation of background texture.

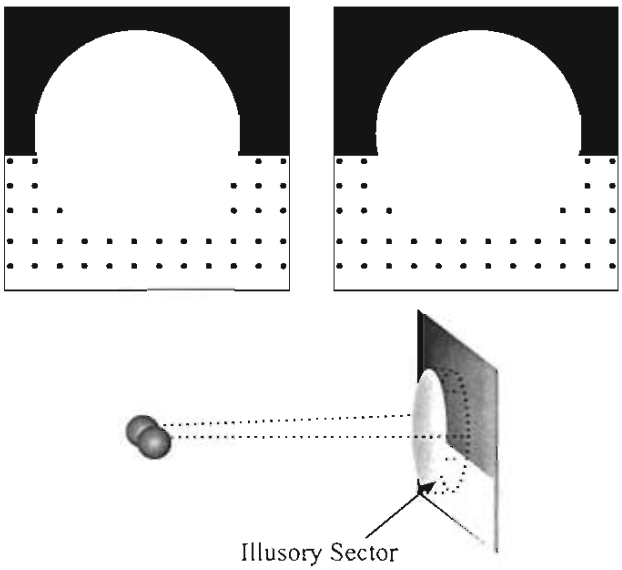



**Fig 3.3. Stereoscopic contours across homogeneous regions**
When disparity is introduced to a semi-circular central figure, its stereoscopic boundaries become strongly enhanced. Note that the contour of the lower portion of the 'circle' is misshapen and its edge is not distinct. Gulick and Lawson claimed that this was because of a lack of non-corresponding features and hence the denigration of *form disparity* (Adapted from Gulick and Lawson, 1976).

In summary, Gulick and Lawson identified many of the issues that are of interest to this thesis: the role of non-corresponding features, the extraction of 3-D form, and the significance

of both sharply defined and more indeterminate stereoscopic boundaries - stereoscopic illusory contours. Their conclusion was that the contours were related to Julesz cyclopean contours by the relative density and thus the degree of perceptual aggregation of texture features.

### 3.2.2. Ambiguous textures: The wallpaper effect, ground capture and its failure

The so-called *wallpaper effect*, observed by Brewster, 1844 (cited in Tyler, 1991) arises when a repetitive textured transparent film is overlayed across a monocular Ehrenstein grid or vice. The textured elements of the film tend to be captured by the plane of the grid (Spillman and Dresp, 1995).

Prazdny (1985), Ramachandran (1986, 1987, 1988), and Ramachandran and Cavanagh (1985) have demonstrated similar effects with the Kanizsa square in the stereoscopic domain as well as in motion. Regularly dispersed dots, horizontal and vertical lines can be *captured* forward by regions bounded by disparate illusory contours.

Ramachandran explains these effects as ecologically useful 'tricks' or 'shortcuts' by which the brain makes assumptions about segmentation of a sparsely textured scene. But capture is not an all-or-none effect. The percepts can be surprising. For example, there is a subtle anisotropy in the distinctiveness of capture between horizontal and vertical lines. Further, with random-line texture patterns, 'capture' disappears entirely (Howard and Rogers, 1995). The percept is a peculiar transparent glass-like plane. The same figure against a random-dot texture produces no capture, and in fact stereopsis causes the illusory contours to vanish. Figure 3.4 illustrates these effects.

Vallortigara and Bressan (1994) argue that the key to stereo capture is ambiguous occlusion geometry. They suggest that a conflict exists between the depth plane of the elements indicating occlusion (the inducing disks), and those with zero disparity (the periodic patterns susceptible to the wallpaper effect). When the texture of a surface, or a group of elements provides only ambiguous stereoscopic information, the texture or elements are drawn to the near surface.

Wantanabe and Cavanagh (1992, 1993) have also found that the classic Kanizsa square captured texture more strongly than isoluminant figures and those defined by line ends. Evidently then there are relative degrees or strengths of closure that contribute toward illusory forms in the stereoscopic domain. It also appears that texture patterns vary in the degree to which they represent a perceptually robust surface. These in issues will be addressed in detail in later chapters.

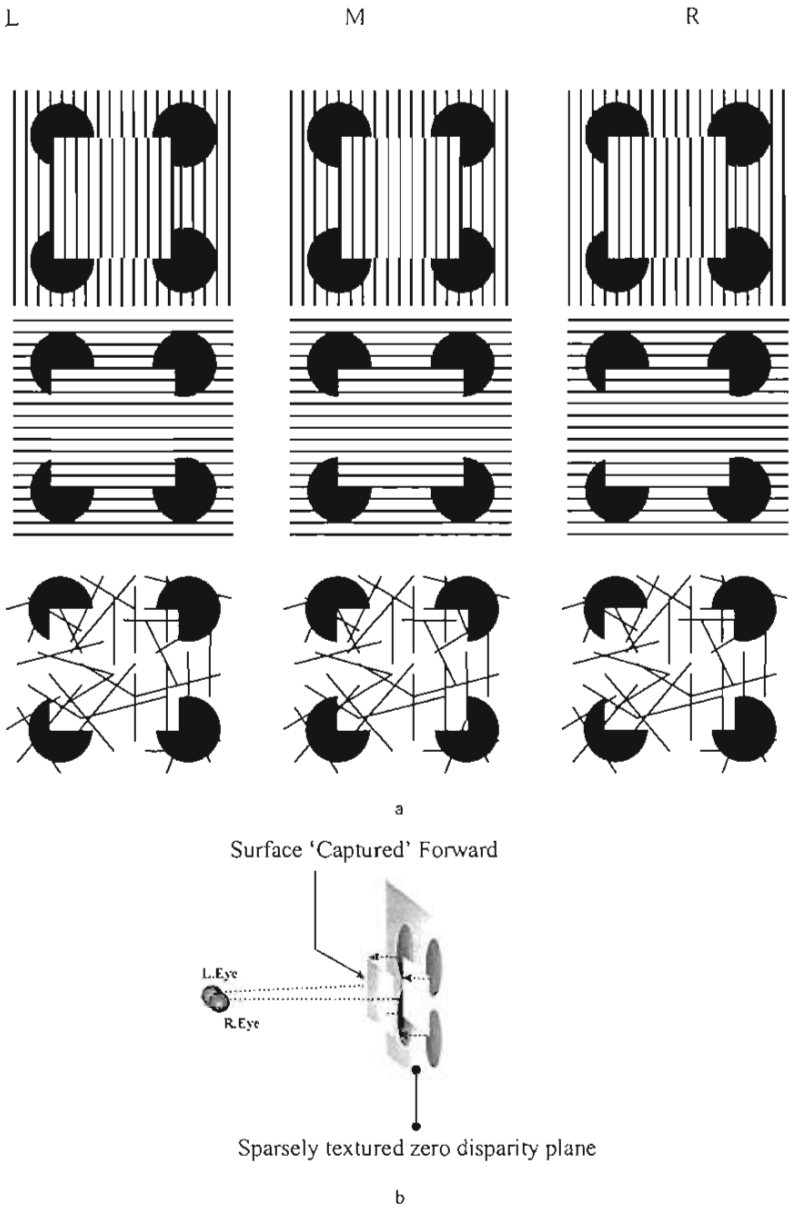



**Fig 3.4. Capture and the failure of capture in illusory figures**
The stimuli in (a) have precisely the same dimensions at both signs of disparity against three different backgrounds. Crossed fusion of the (L) and (M) pairs gives the percept that horizontal or vertical lines are captured to the plane of the illusory surface. However, the sparsely textured random line background yields a glass-like transparency. With uncrossed fusion, (M) and (R) pairs, all backgrounds show that features bounded by the disks are captured to the distant plane except in the case of random lines. Capture is shown pictorially in (b).

Gillam (1995) has also shown that random-line stimuli generate very distinctive stereoscopic illusory contours at their boundaries. As mentioned, such random-line stimuli resist capture. They apparently represent robust surfaces. In the same way that densely textured RDS surfaces are robust, the organisation of random lines must be highly

unambiguous to stereopsis. Both contour orientation and irregular shaped spaces between the contours generate local disparities (that is, zero disparity) conflicting against capture on to the illusory plane.

### 3.2.3 Illusory contours in densely textured stimuli

This section examines illusory contours in dense random-dot textures. Prazdny (1985) has demonstrated that a cyclopean Kanizsa square defined by motion, flicker (rapid contrast polarity alteration), or stereopsis alone, did not produce illusory contours (see Fig 3.5a). Some of his more interesting static effects are shown in Fig 3.5.

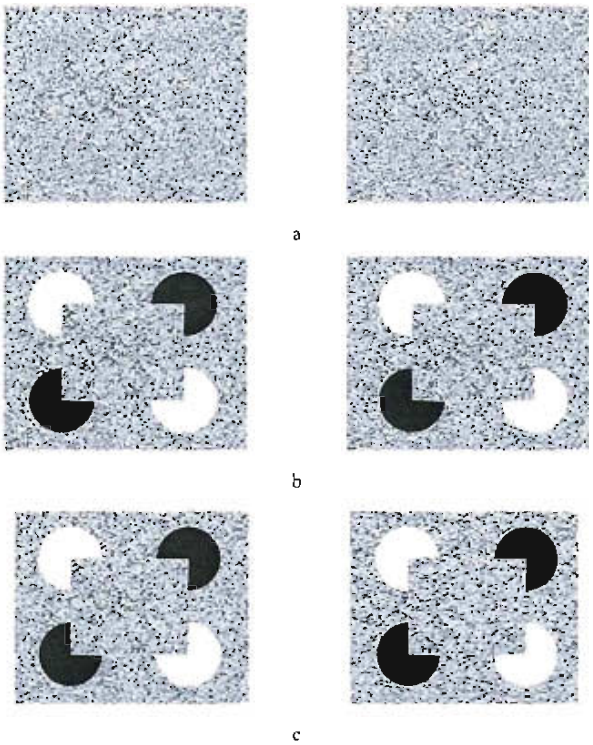



**Fig 3.5. Illusory contours against a random dot field**
In the absence of regional contrast, cyclopean inducers generate no contour effects in flicker, motion, or stereo studies (3.5a). In 3.5b contrasting disks against a random-dot background yield illusory contours only when they are stereoscopically set behind the surface (but they are monocularly visible). Inducers in 3.5c are set against an uncorrelated texture. Stereoscopic illusory figures can be seen (adapted from Prazdny, 1985).

Clearly, Kanizsa squares set against a densely textured background generate illusory contours in the monocular domain but not in the stereoscopic domain. Pradzny's findings signify a relationship, yet to be defined, between the relative ambiguity of local level and larger scale features in the perception of stereoscopic illusory contours. Densely textured surfaces

have what might be termed a robust surface. Each local texture feature is highly *unambiguous*. By unambiguous is meant that stereopsis provides robust information about the position of local features relative to each other (see also Howard and Rogers, 1995).

## 3.3 Surface spreading effects

This section presents a series of intriguing percepts that will collectively be termed *spreading effects*. These include *disparity propagation* and *contrast spreading*, or *neon colour spreading*.

### 3.3.1 Spreading of disparity signals within illusory boundaries

Consider Fig 3.6a. Takeichi, Watanabe and Shimojo (1992) argued that uncrossed disparity signals generated by the position of the small dots in each image propagated outward within an illusory boundary to yield the percept of a bounded surface behind the Kanizsa triangle (and the Ehrenstein cross in 3.6b). When the central dots fall behind the depth of the disks, they seem to capture the region bounded by illusory contours to their depth. This is not an all or none effect, the spreading can be ambiguous.

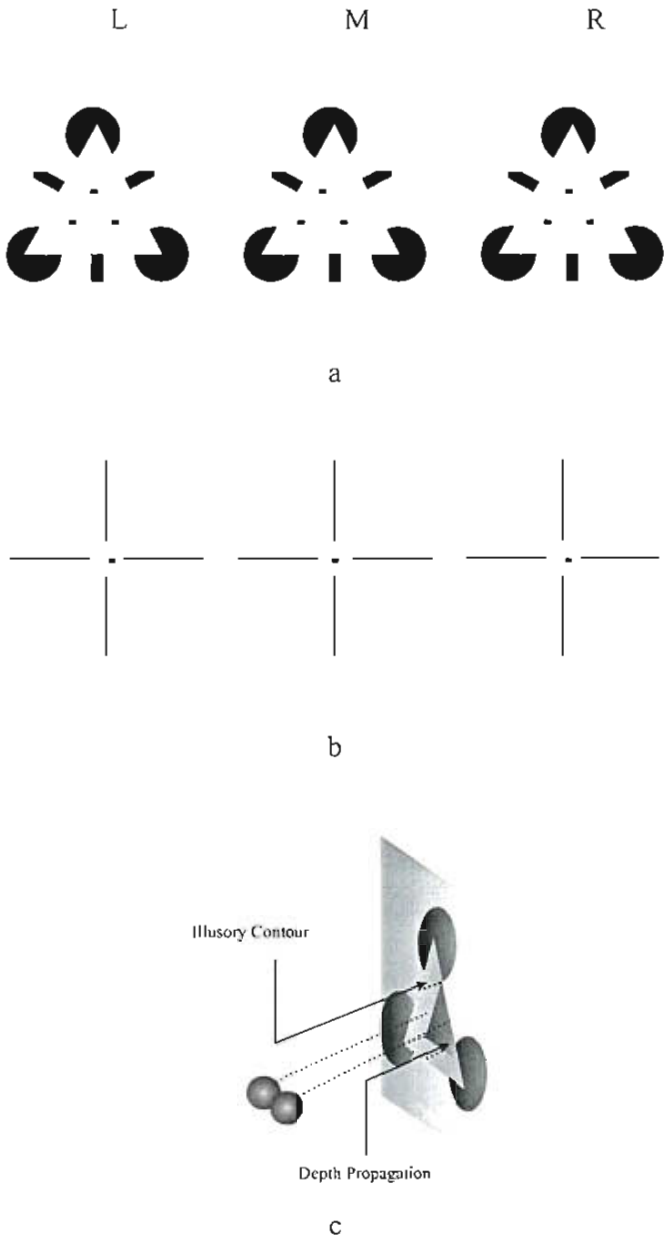



**Fig 3.6. Depth propagation**
A so-called *depth propagation* effect is seen when the middle and left pairs in both sets of half images (a) and (b) are crossed fused. In the right and middle images, no such effect is evident when the small set of features stands forward of the disks. As the pictorial depiction is meant to show, the propagation arises behind the picture plane, that is, behind the illusory boundary, not forward of it.

Taikeichi et al argued that the dot disparity signals spread outwards within the boundaries of the illusory figure. But why doesn't the same disparity spreading occur when the dots stand forward of the inducing discs or the cross? Secondly, why don't the inducing disks in (a), or the arms of the cross in (b), modally complete as they do in the SKS with uncrossed

disparity? These questions appear to suggest that local issues involving segmentation of surface layers have an asymmetric character between the two signs of disparity.

### 3.3.2 Surface spreading: contrast, colour and transparency effects

A phenomenon that is intimately related to the role of stereopsis in segmenting apart surface planes is *neon colour spreading* (van Tuij, 1975). Neon spreading occurs when a coloured cross is set into the central sector of an Ehrenstein figure. If disparity is applied to the position of the coloured sectors in each eye's view, a transparent coloured disk[2] floats forward of the cross as shown in Fig 3.7. When the sign of disparity is reversed, the disk stands behind the black regions. It now appears opaque; no colour spreading results (Nakayama and Shimojo, 1990).

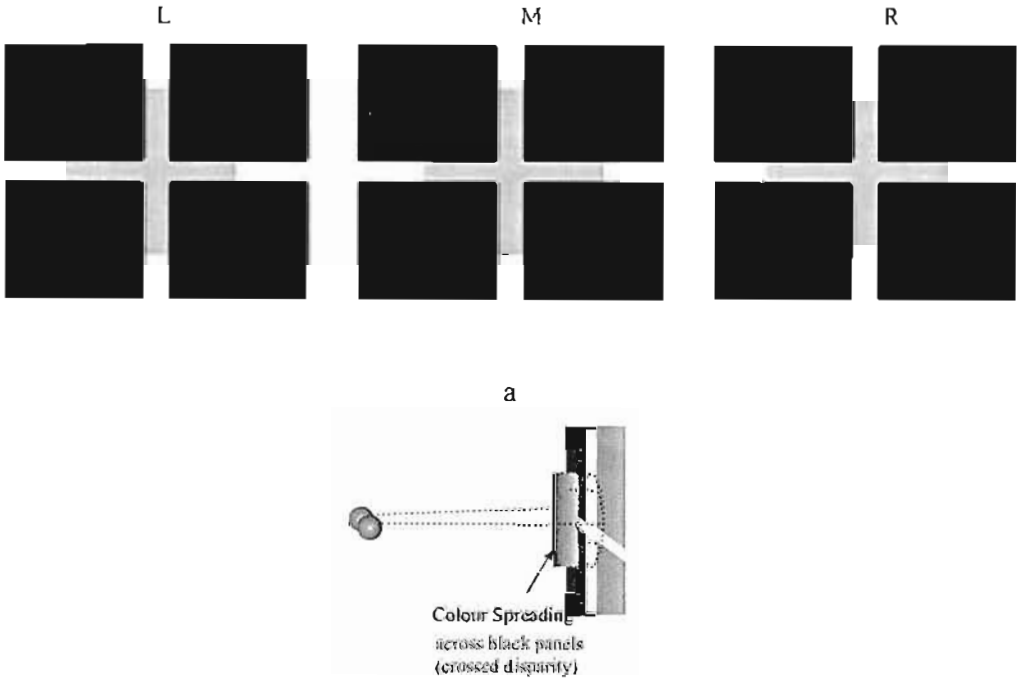



**Fig 3.7. Neon colour (contrast) spreading in stereopsis.**
van Tuij's neon colour spreading effect is vastly enhanced in the stereoscopic domain where a red disk is inserted into the central region of an Ehrenstein figure. The sign of disparity has critical implications for the phenomenal quality of the figure. Crossed fusion of the left and middle pair in (a) sees the central figure standing forward of the Ehrenstein cross - this is pictorially presented in (b). Fusion of the right and middle pair in (a) yields no spreading.

2 - Note that the figures presented in this thesis use non-chromatic contrasts; the effects are very similar (Anderson, 1997).

The transparent nature of the figure, as it stands forward of the plane of the cross, means that the cross boundaries are still visible through the spreading disk. In the opposite disparity case, reversal of the figure reveals no such transparency. The black regions are now the interposed surface. Watanabe and Sato (1989); Watanabe, Takeichi and Shimojo (1990); Watanabe and Shimojo (1990); Watanabe, Nanez and Moreno (1995) have argued that these remarkable effects are due to the parallel operation of Grossberg's (1987) Boundary and Feature Contour mechanisms in the stereoscopic domain. The colour spreading percepts are another example of perceptual asymmetry between the signs of disparity.

There has been a lot of recent interest in these and the contrast spreading figures (Anderson and Julesz, 1995; Anderson, 1997; Nakayama, Shimojo and Ramachandran, 1989). Only a very brief summary of related figures is presented. Nakayama (1996) has argued that transparency is due to contrast ordering. The disk, being of intermediate contrast between the white spaces and the black panels completes as a by spreading across the black panels. Mertelli (1974) first introduced the relationship between luminance order and transparency.

An alternative view of Anderson (1997) is that the luminance order constrains the *scission* or separation into two *causal layers*. When two contours are aligned, a change in contrast that does not disrupt contrast polarity can result in the lower contrast region being *decomposed* into two separate layers. Transparency results when two surface layers can represent a change in contrast along aligned contours. A shift in contrast polarity, in this context, means a shift from high contrast to low contrast, or vice verse.

Demonstrations of Anderson's concept of scission are presented below in Fig 3.8. Though effects are not entirely new (see Nakayama, Shimojo and Ramachandran, 1989), Anderson's work has clearly identified that in order to explain illusory contours in stereoscopic vision any theory will need also to be applicable to the spreading of contrast layers in untextured stereograms generally. Anderson's theoretical approach will be addressed in some detail in a later section.

In summary, this section has flagged that a general relationship exists between stereoscopic contrast spreading and illusory contours. There are evidently localised issues of contrast polarity which constrain the *qualitative* dimensions of contrast spreading (Anderson, 1997). But in Fig 3.8d it is demonstrated that features which are unambiguously resolved at the background plane will generate transparency as well (the oblique line can be any detectible luminance).

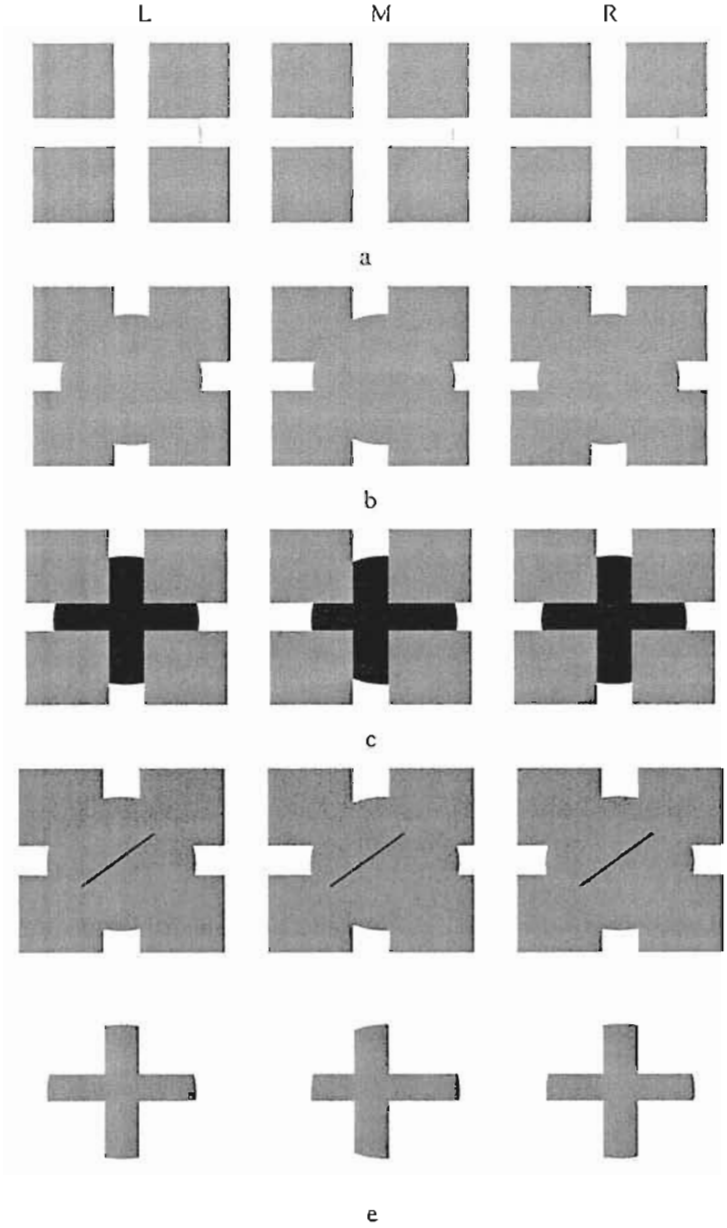



**Fig 3.8 Surface scission, spreading and luminance order**
Anderson (1997, and see also Anderson and Julesz, 1995) demonstrated the importance of contrast interactions in the perception of transparency and its relationship to illusory contours. In 3.10a a transparent form is seen: spreading of the intermediate-contrast sector boundaries occurs. Anderson describes this as contrast scission. In 3.10b with no contrast difference between the central cross and the panels, no scission arises. The system segments out an opaque, modally complete, disk. 3.10c, with the luminance of the disk darker than the side panels, scission does not arise. The disk still completes. But there is no spreading of contrast. The result is an illusory boundary crossing the panels. The disk takes the appearance of glass. In (d) transparency can be achieved by setting a single unambiguous oblique line to the background plane. Stereoscopic ambiguity may also be important where no intermediate form or luminance term is available (as in e).

The summary of the perceptual phenomena to this point suggests that perception of 3-D illusory contours is related to the manner in which the system partitions the depth percept into

surface layers. Accordingly the discussion now turns to address the role of stereopsis in resolving occlusion, that is, where one surface layer partly occludes another.

## 3.4 Stereopsis and occlusion

Binocular parallax means that when a near surface partly occludes a distant surface, more of the distant surface will be visible to one eye than the other. There is mounting evidence that stereopsis plays a critical role in resolution of occlusion, that is, where a near surface partly obscures another. Unpaired features will occur on the distant surface that will be left over from the matching process. The perceived depth of such regions has been treated by a simple computational rules such as the *most distant surface* rule (Julesz, 1971). However, the unpaired features may turn out to be useful depth cues not just elements left over after point-matching.

For example, Gillam, Findlay and Flagg (1984) have found that disparity discontinuities, such as those that arise at object boundaries provide important information on depth relationships. Slant latencies (time to binocularly fuse and recover stereoscopic slant) in grid matrix patterns were much reduced when disparity discontinuity was available. Fig 3.9 demonstrates what is meant by disparity discontinuity.

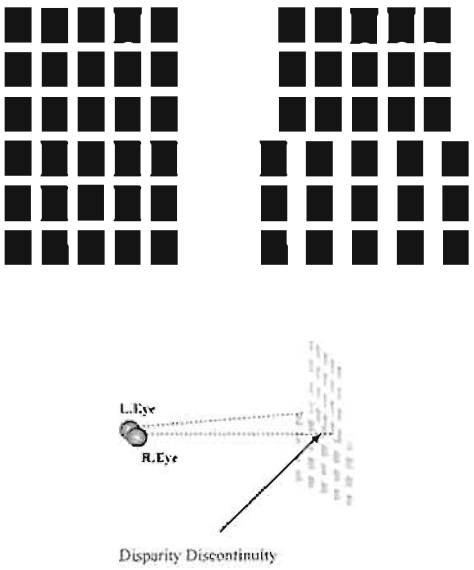



**Fig 3.9. Disparity discontinuity as a primary cue for stereopsis**
Gillam demonstrated that disparity discontinuities reduce the time taken to fuse image pairs dramatically.

Further, Gillam and Borsting (1988) have shown that clearly demarcated monocular occlusion zones reduce slant latencies in RDS figures as well. It has become clear that surface discontinuities and even near reference frames can powerfully disambiguate the geometry of the disparity field.

### 3.4.1 The Geometry of Monocular Occlusion Zones

Figure 3.10, presents a pattern of linear projections that would create monocular zones in stereopsis. Howard and Rogers (1995) presented four basic rules that describe these regions:

1. The monocular regions in each eye are on the temporal side of the occluding object S1.
2. A monocular zone due to occlusion is more distant than the binocular object that creates it (surface S1, above).
3. Eye movements affect the physical size of the monocular zone only slightly.
4. The angle of subtense of a monocular zone is inversely proportional to the distance of the occluded object. For a binocular object at a given distance, the angle f above increases with the distance between surfaces S1 and S2 increases.

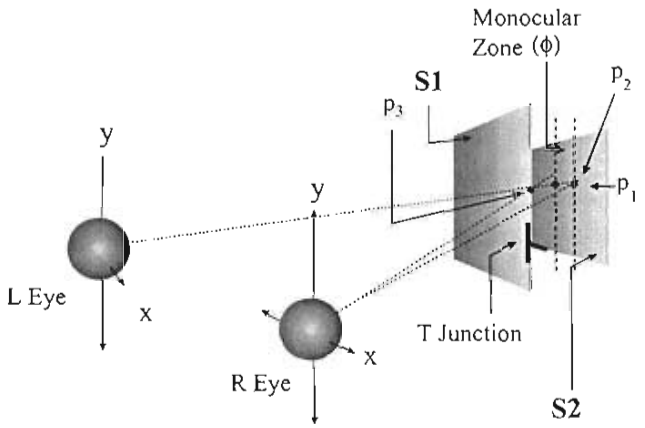



**Fig 3.10. Interposition and monocular occlusion**
Surface S1 here occludes S2 resulting in a monocular zone on S2. The point $P_1$ is visible to both eyes while $P_2$ is visible only to the right eye. The result is a monocular zone of magnitude $\phi$. If the surfaces are opaque, stereoscopic T-junctions form at the point of interposition. The Top of the T is the edge of surface s1 and the stem: s2 separated by the distance between $P_3$ and $P_1$.

The geometry of monocular zones in natural vision is quite straight-forward. Their utility as a possible depth cue is also clear. However the manner in which the system accesses monocular zones is not well established.

### 3.4.2 Ecological Validity of Monocular Occlusion Zones

According to the geometric rules above, when monocular zones are presented in the 'wrong eye' binocular rivalry may result. Nakayama and Shimojo (1990) illustrated these rules in action. In Fig 3.11a, the eyes' views of the black disk occlude two crescent shapes meant to represent monocular zones. When the images are cross-fused, the nasal side crescent is clearly defined in depth. It stands behind the disk as part of the textured background plane. The temporal side crescent rivals, and may be suppressed. In (b), the crescents are in precisely the same position, but the textured plane is shifted forward of the black disks. These arrangements will generate rivalry because they are inappropriately positioned, at least according to Nakayama and Shimojo (1990).

Nakyama and Shimojo (1990) claimed that such effects signify the ecological utility of monocular occlusion as a low-level information primitive. Where monocular occlusion breaks the rules of nature, learned through experience, the resulting percept is not permissible and, therefore, rivals. The eye of origin and the position in the retina where these features arise, are deemed to be critical information. Nakayama, (1996) claims that this means the processing for such features must be occur in area V1 of the striate cortex, just prior to the earliest binocular processing. As a result, the treatment of such zones was thought to be non-stereoscopic. Nakayama and his colleagues have extended their understanding of monocular occlusion zones to explain the generation of occluding edges in untextured stereograms.

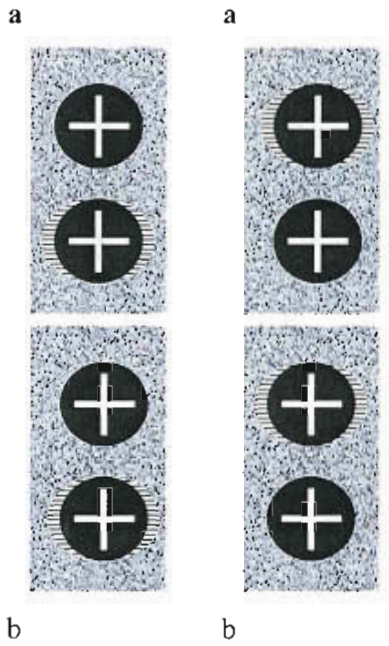



**Fig 3.11. Ecological appropriateness of monocular zones**
This figure (from Nakayama and Shimojo, 1990) shows the importance of the organisation of occlusion configurations for the binocular resolution of monocular zones (see text for detail).

### 3.4.3 Da Vinci Stereopsis: processing specificity of monocular image regions

Nakayama and Shimojo's theory of da Vinci Stereopsis (1990) suggested that unmatchable image features, could provide a source of unambiguous ordinal information specific to their eye of origin which accounted for the generation of sharp occluding edges. Da Vinci Stereopsis suggests that unmatchable regions are processed separately from corresponding features. This is demonstrated in Fig 3.12. In an almost identical image, where no monocular features exist, there is no illusory contour induced.

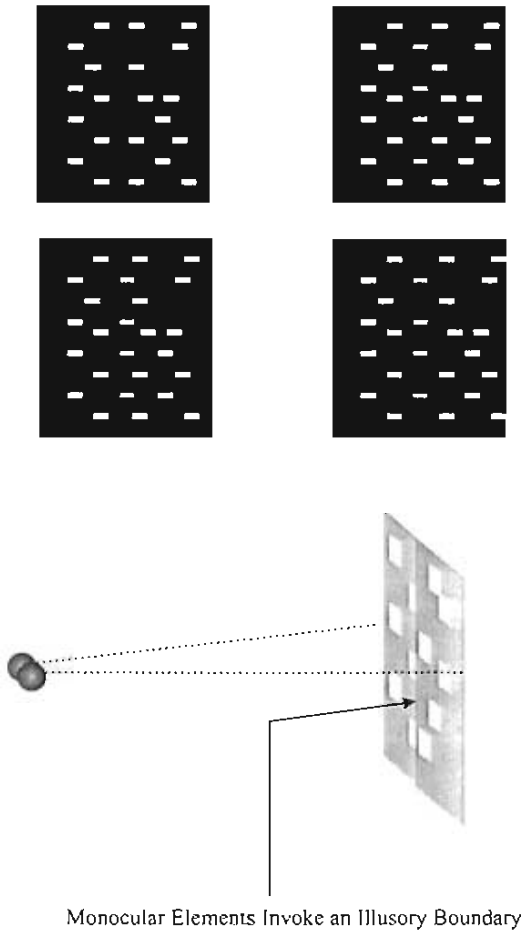



**Fig 3.12. Da Vinci stereopsis**
Nakyama and Shimjo (1990) demonstrated that surface segmentation and the stratification of surface features was intimately related to monocular features (grey dots). In the top pair, crossed fusion yields an illusory contour occluding the monocular zones from one eye. In the lower pair, the grey features do have corresponding features. No occluding edge is detected.

### 3.4.4 Quantitative stereoscopic depth from monocular occlusion

The work of Nakayama and his colleagues has led to the understanding that unpaired regions of stereoscopic half-images can invoke a 3-D percept in the absence of retinal disparity *per se* (Anderson and Julesz, 1995; Nakayama, 1996). Similarly, Liu, Stevensen and Schor (1995) presented a series of effects, which they claimed generated stereoscopic depth in the complete absence of disparity information. In their demonstrations (see Fig 3.13) a white rectangle intrudes into a black sector. When fused, a rivalrous but clear segregation of surface planes is perceived. The white rectangle can be seen floating above (or behind) a black rectangle, dependent upon the sign of image differences (crossed or uncrossed).

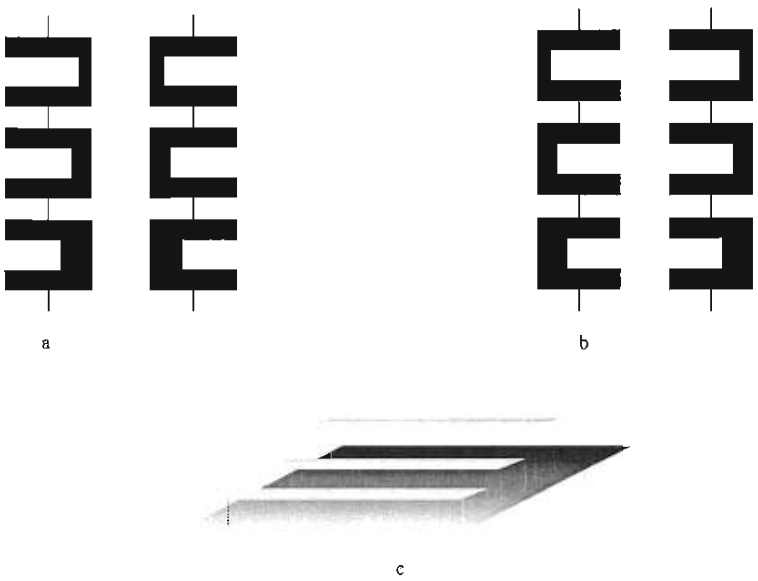



**Fig 3.13. Stereopsis from monocular occlusion**
Crossed fusion of the pairs at (a) yield the percept that a central white rectangle stands behind the black. The distance of separation increases downwards in this set of images. The set of pairs (b) reverses the direction of distance perception as the white rectangle stands forward of the black. (adapted form Lui et al. 1994). The crossed fused case is schematically presented in (c).

Gillam (1995) argued that these images yield horizontal disparity and so could not be considered to represent stereopsis from monocular zones alone. Gillam's explanation is illustrated in Fig 3.14. Horizontal disparity is present in the magnitude of the horizontal junctions between the black and white regions.

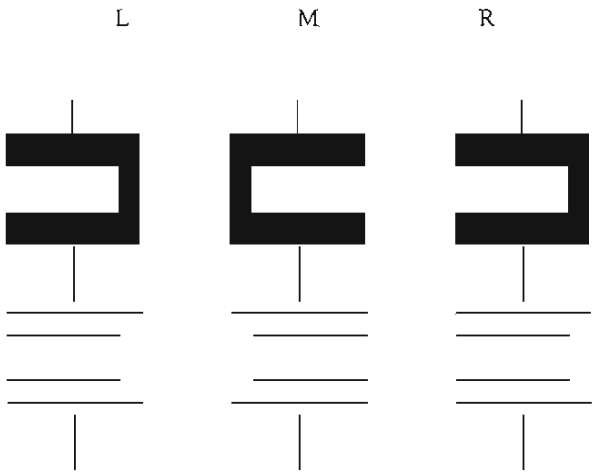



**Fig 3.14. Disparities available surrounding occlusion patterns**
By replacing Lui's stimuli with lines of the same horizontal extent, Gillam showed that a standard horizontal disparity could explain the effects.

In summary, the geometry of occlusion is, rightly, beginning to be recognised as an integral part of binocular perspective. Occlusion represents a problem for the traditional understanding of stereopsis because, by definition, unpaired features arise where an occluding surface obscures a region of the background from one eye's view.

Da Vinci stereopsis suggests that monocular zones are the crucial factor in disambiguating surface interposition. Lui et al. have attempted to show that quantitative distance information can be recovered from the magnitude of monocular regions in untextured stimuli. However, Gillam has demonstrated that disparities in the dimensions of larger scale features of the stereoscopic half-images must be considered when interpreting the role of putatively monocular elements.

## 3.5 Theoretical explanations of illusory percepts in stereopsis.

The next section focuses on the most recent theoretical positions on illusory stereoscopic phenomena. These include a Form Computation View and a Surface Heuristic View.

### 3.5.1 FACADE Theory: Form Computation

Grossberg's (1994) FACADE (Form-And-Colour-And-Depth) theory argues that stereoscopic illusory contours and the spreading effects result from figure-ground segmentation by two magnocellular streams, the Boundary Contour System (BCS) and Feature Contour System (FCS). A BCS generates *emergent segmentation* reconstituted from edge, texture, shading, and stereo data. The FCS fills-in phenomenological properties such as brightness, colour and depth in interaction with the BCS.

*Emergent Segmentation*

By emergent segmentation, Grossberg means (1994, p.68) 'partitioning of an image into boundary structures'. These boundaries need not correlate directly with the physical character of the image data, but can *emerge* indirectly through the response of various stages of BCS processing. In that sense, "all line ends are illusory" (Grossberg, 1994, p.68). Completion of fragmentary shape is performed by what is termed a cooperative-competitive feedback loop (CC-Loop). This feature is designed to find continuity within noisy image features, that is, it searches for coherence of contrast orientation. Very early image data consists of the orientation selectivity of cortical simple cells. These orientation responses are later subjected to *end-cut*

refinements. These define the critical end points of lines, groups of lines or shading features. Illusory contours are generated by these end-cut mechanisms.

*Filling-In*

In the stereoscopic domain, these processes lead to a *multiplexed representation* in the final stages of the FCS through a binocular *filling-in-domain* (FIDO). This notion stems from Grossberg's assumption that the visual system is able accommodate distortions or gaps in texture and structure by filling-in segmented regions at later stages. In the case of binocular images, boundaries are created by pools of cells sensitive to, for example, a certain range of disparity signals. The activity of these cells, via the CC-Loop, can yield emergent boundaries. Surface feature data, derived from Near-zero disparity information, pools then fills-in the surface percept. Regions not bounded by the BCS system at the filling in stage, can suffer feature flow, as the filling-in process spills out of the semi-contained region.

*FACADE theory and da Vinci stereopsis*

Grossberg uses the term *Da Vinci stereopsis* to encapsulate the processes by which monocularly visible features are integrated with binocular features of the image data. This is a slightly different view to that of Nakayama and Shimojo (1990). He couches his explanation in terms of allelotropia—the distortion required to be applied to each retinal image to achieve binocular fusion–and, hence, the 3-D percept. He argues that, in order to achieve allelotropia, there is cooperation and filling-in of gaps in the binocular fusion process achieved by the Boundary Contour System. In summary, disparity sensitive cells are able to call upon a pool of Zero-Disparity activity which enables the BCS to, in a sense, fill-in the gaps that might exist in one eye's view of a scene. Figure 3.15 illustrates some of the key issues with which Grossberg deals.

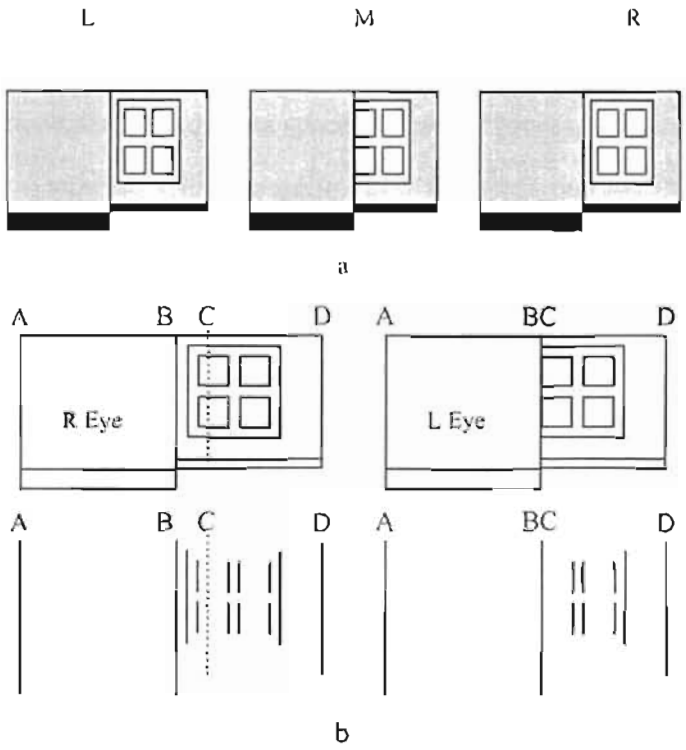



**Fig 3.15. Da Vinci stereopsis according to Grossberg**
Crossed fusion of the top left and middle pair in (a) yields the percept of a wall at left standing forward of and occluding windows (divergent fusers use the right and middle). This is despite the fact that only parts of some windows are visible to the left eye. Diagrams in (b) define corresponding regions in each eye. The region BC in the right eye's image is monocular (Adapted from Grossberg, 1994).

In the figure above, disparity signals can be divided into several scales of magnitude. A large scale disparity signal can resolve the broad patterns ABD. It is the smaller scale region at the window frames which are problematic for allelotropea. The small scale signals between BC cannot be directly fused. These features are assumed to be relegated to the distant surface: part of a Near Zero Disparity Pool. Allelotropea can be achieved for BC from the activity of cells in this near-zero disparity pool and so are filled in by activity termed a binocular syncytium, which utilises boundaries from either eye to formulate a discrete object.

*Neon colour and contrast spreading*

Neon colour spreading is an example of the continuation and segmentation processes in Grossberg's theory. The BCS utilises the CC-Loop to integrate vertical and horizontal end-cuts at the vertical and horizontal extents of the central cross. Integration takes the form of a smoothly rounded appearance that unites disparity signals indicating a nearer surface. An FCS

binocular FIDO then fills this bounded region from the zero-disparity pool and so the colour or contrast spreads out within the confounds of the circular boundaries. Figure 3.16 illustrates this process. Note that the black panels and white spaces below are processed in a similarly discrete fashion.

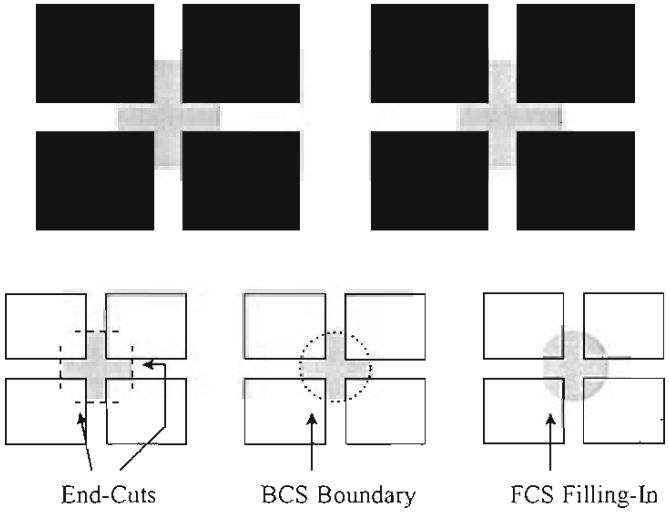



**Fig 3.16. Colour spreading according to Grossberg**
Crossed fusion of the stereo pair reveals a grey disk, standing forward of the black panels. This percept is the result of a sequence of end-cut, Boundary completion (by the CC-Loop) and finally, filling in by the FCS's binoeular FIDO.

In the case of the reversed disparity sign in Fig 3.16, the black panel boundaries now inhibit the filling-in process and so no spreading arises.

FACADE theory is arguably the pinnacle of physiological pre-attentive processing models. It seeks to explain visual phenomenology in terms of hypothetical processes of computation derived from current physiological models. As such, it is a valuable yardstick against which to test psychological approaches to the issue of perception. The theory is continually evolving to account for recent research findings.

One criticism from Anderson and Julesz has been that it under emphasises the distance information that is available at monocular zones. Also, in terms of Panum's limiting case (Gillam, Blackburn and Cook, 1995), multiple fusions can mean that the same contour can yield different disparity values– and, accordingly, different disparity pools. Further, it is possible to generate contours and interpolation effects that are not constrained by bounded forms, as such.

The view that a monocular region represents a gap in the cyclopean percept would seem to be problematic in that light. Again, this is because of the implicit assumption that resolution of correspondence is the pivotal task of stereoscopic vision. One advantage of Grossberg's

theory, is that it deals with the inter-retinal magnitude of broad regions of the stimulus. My project utilises these broad dimensions of untextured stereograms to describe how 3-D percepts can be interpreted.

## 3.6 Surface Heuristic approaches

Nakayama and Shimojo (1992) and Anderson and Julesz (1995) have developed somewhat different theoretical approaches to these illusory stereoscopic phenomena. Their ideas will be loosely categorised as a Surface Heuristic approach. The term heuristic indicates that their understanding is generally that the system uses certain features of stereoscopic images as a rule of thumb to infer 3-D the layout of perceived surfaces.

### 3.6.1 Nakayama's Principle of Generic Sampling

Nakayama and Shimojo (1992) have presented a Bayesian inferential framework, borrowed from computational approaches to perception, that relates directly to the issues addressed here. To understand Nakayama's approach, it is helpful to briefly summarise this influential understanding of visual processes.

*Bayes' Theorem as a framework for understanding visual perception*

Bulthoff (eg. 1993) has developed an influential framework for theoretical issues in perception based on Bayesian Decision Theory. Bayes' theorem is a general statistical in which event outcomes are explained in terms of the likelihood of contributing factors giving rise to them. In vision, the factors contributing to the final percept are said to be the outputs of various detection mechanisms. In sum, the system utilises a set of competing prior assumptions about the visual scene to combine visual cues to construct a particular percept.

The underlaying assumption of this approach is that the visual scene is 'ill-posed' at the retina. The retinal image is an 'arbitrarily complicated function' of field of view, providing insufficient information on which to found a representation. Bayes' formula is used to demonstrate how the system might constrain the possible views inferred from a particular image. Bayes' theorem, in this application, states that:

$$P(S|I) = \frac{P(I|S)P(S)}{P(I)}$$

Where S signifies the scene; I, the image; and so P(I|S) is a *likelihood function* that describes the statistical probability of deriving an image I from a scene S. P(S) therefore represents the prior probability of a set or sequence of scenes occurring in the natural world. So P(S/I) is a description of the probability of a scene being S when a certain image I is detected.

*Integration of Visual Cues*

In Yuille and Bulthoff's view, the percept represents *fusion* of information derived from the output of modules detecting discrete visual cues (eg. disparity or shading). Unlike the computational approaches of Marr (1982) and Marr and Poggio (1982), their argument is that *strong coupling* exists between the activity of these modules. Prior assumptions indeed operate at the level of individual cues, and each may conflict with, or be made redundant by, other visual cues. For example, in Julesz's RDS figures, disparity information will override luminance evidence that no depth differential exists. However, each module does not necessarily bear a static or deterministic likelihood function. Higher level interpretative inference can have a top-down effect on the eventual decision: the percept.

*POGS and stereoscopic illusory phenomena*

Nakayama and Shimojo argued that binocular vision is not a hard-wired process. Rather, binocular vision is inferential in nature and acts as a surface-representation mechanism. Inference, in this sense, is not "cognitive" or "problem-solving" in nature, but results from low-level neural processing networks whose activities are shaped by perceptual experience.

Their suggestion is that, in locomoting around the environment, the visual system develops an inherently inferential neural architecture that is biased toward certain interpretations of a scene. The system learns that ambiguous image patterns can be created by objects when they are viewed from a particular vantage point. This is termed *image sampling*. A vantage point from which extensive unambiguous three-dimensional information can be recovered is termed a *Generic View*. The static image that might be captured by the retina at any particular instance of such a view is termed an *accidental view*. The Principle of Generic Sampling (POGS) is, therefore, the notion that:

> When faced with more than one surface interpretation of an image, the visual system assumes it is viewing the scene from a generic, not an accidental, vantage point (Nakayama and Shimojo, 1992).

This framework is complicated by the fact that accidental views potentially can be associated with several different generic views. As Fig 3.17 demonstrates, the image of a square ($I_3$) can theoretically represent the generic view of a square ($S_2$) or an accidental view of a cube ($S_3$). According to the POGS the system will assume the image detected was a generic view.

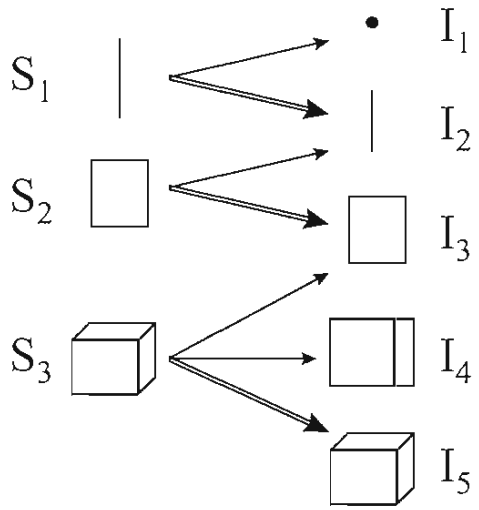



**Fig 3.17. Inverse ecological optics, generic and accidental views**
Dependent upon the position from which the objects $S_1$, $S_2$, and $S_3$ are viewed, they can produce *Accidental* (thin arrows) or *Generic* views (thick arrows).

Consider a view of the cube ($S_3$). In moving around the environment, the likelihood of achieving $I_5$ (three sided image) is greater than that for the two sided corner $I_4$, and still greater than the square accidental image $I_4$. With increases in viewing distance $P(I_5)$ approaches 1 and both $P(I_4)$ and $P(I_3)$ approach 0. It is logical, therefore, that when the system is presented with $I_2$ it will perceive $S_1$. Similarly, in detecting image $I_3$ , it will be interpreted as $S_2$. It is only when $I_5$ is detected that a cube will be interpreted.

*Implications of POGS for surface perception*

Nakayama has demonstrated the implications of POGS for illusory boundaries and transparency. In Fig 3.18a, a paradoxical asymmetry is perceived, given the reversal of disparity between two simple cruciform stimuli. Crossed fusion of the left and middle cruciform sees the

horizontal arms of the cross folded forward of the vertical arm. Fusion of the middle and right images yields the percept of a horizontal arm occluded by the vertical arm. These percepts are presented schematically in Fig 3.18c. The former percept, folded horizontal arms, is that predicted by simple horizontal disparity, while the later percept is paradoxical given a literal interpretation of disparity. The configurations predicted by disparity alone are those illustrated in Fig 3.18b, where vertical lines have been used to indicate the important corresponding elements of the images. But why does the horizontal arm not appear folded in the second case? Note that both a folded, and an occluding, configuration were thought to represent two possible resolutions of binocular fusion. These are therefore ambiguous stimuli.

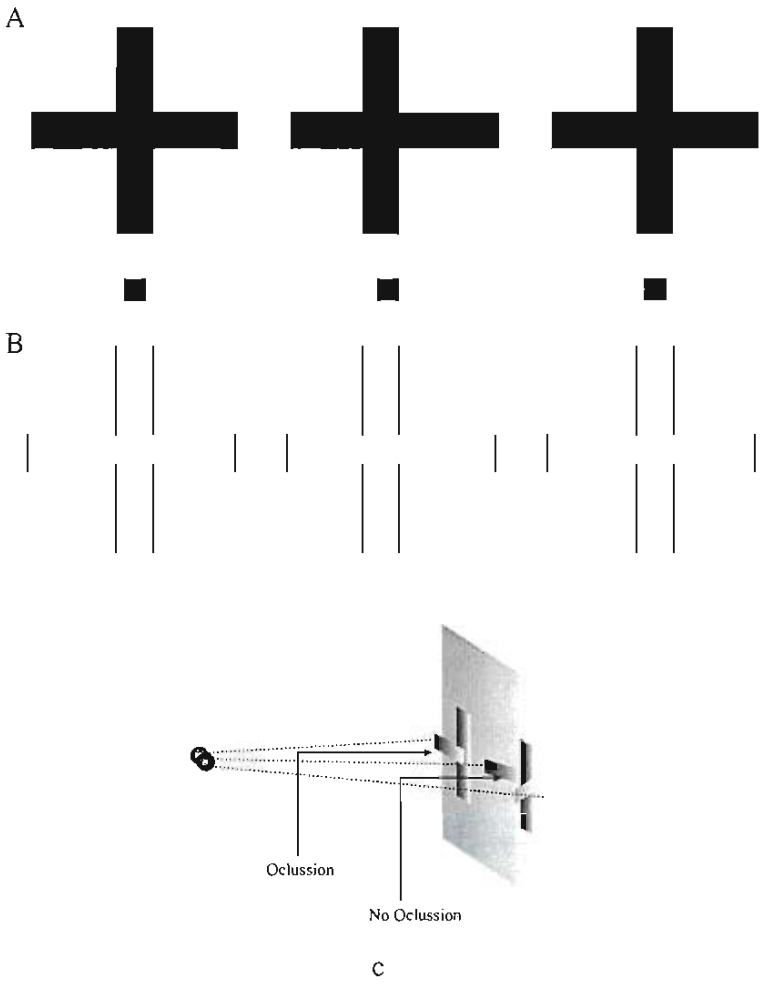



**Fig 3.18. Image Ambiguity in Untextured Stereograms**
In (A) crossed fusion of left and middle images compared to fusion of the right and middle produces a paradoxical asymmetry in the percepts invoked. (B) presents the simple disparity information available. The figural asymmetry is presented pictorially in (C).

Nakayma explained that the key to the asymmetry above is a difference in image sampling possibilities. An observer would need to be in a particular position relative to the stimulus to detect the folded form. Any other position would yield a different organisation, either an occluding configuration, or a change in the slant of the arms. Therefore, this folded figure is considered to be an accidental view. However, in the occluded case, any position taken by an observer will not reveal a qualitatively different percept. Therefore, the occluded figure is considered generic.

Fig 3.19 presents two other examples from Nakayama and Shimojo, demonstrating the POGS processing scheme.

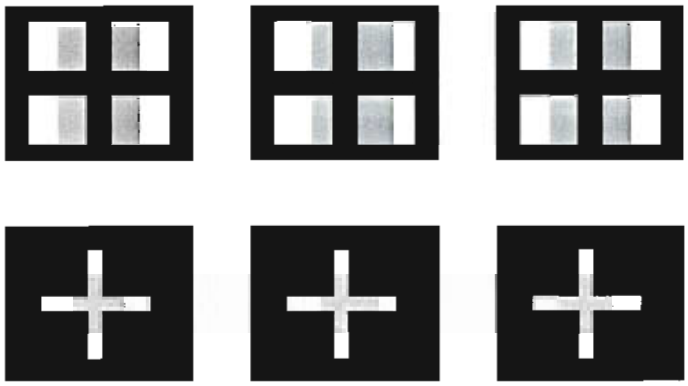


**Fig 3.19. Occlusion configurations from generic sampling**
Crossed fusion of these stereograms yields symmetric transparent occluding or occlusion formations despite the alternative view of folded central panels.

*Local Primitives: Stereoscopic T-Junctions*

Nakayama and Shimojo considered the involvement of local features in this scheme. They argued that the junctions along which contrast boundaries were aligned were distinctive patterns invoking occlusion, that is a T-junction. In monocular views, it is not feasible that an object will yield an occluding surface at the top of the T. Their suggestion was that the stereoscopic interpretation of these junctions contributed to the process of image sampling. Accordingly, such junctions were interpreted as being generic to the occlusion configuration generated. This kind of paradox, they claimed, was evidence of what they termed inverse ecological optics. A stereoscopic T-junction is presented in Fig 3.21 below.

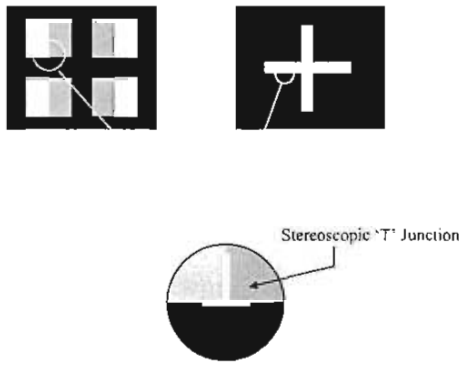



**Fig 3.21. Stereoscopic 'T'-junctions**

A T-junction represents a strong monocular cue to occlusion. In the case of transparencies observed by Nakayama and Shimojo, the stereoscopic version is an important factor in the resolution of the occlusion relationship as well. This is despite the fact that its manifestation in the stereoscopic case flies in the face of previous monocular theory.

Nakayama and Shimojo's Bayesian interpretation of this image sampling notion is as follows. The eventual percept is designated $P(I_m|S_n)$, which is the probability of a given image $I_m$ given a real world layout $S_n$. This conditional probability is 'represented' in the architecture of visual processes as 'an associative strength'. This is expressed in a likelihood function derived from perceptual learning. Nakayama came up with a modification of Bayes' theorem where $P(S_n \mid I_m) = P(S_n|I_m) / [P(I_m|S_1) + P(I_m|S_2) + \ldots P(I_m|S_n)]$; which essentially does away with the prior probability estimations.

The key message from Nakayama and Shimojo (1992) is that the percepts from in stereoscopic untextured stereograms are those for which alternative possibilities are founded in perceptual learning. One advantage of POGS, is that it removes the resolution of the percept from linear notions of disparity processing as used in many computational models. As such, POGS represents a major advance in the understanding of binocular vision. Binocular cues relating to the interpolation of surface features appear to be used in conjunction with other information about a scene such as luminance orders.

The major criticism of this approach comes from Anderson and Julesz (1995), who argue that it is not possible to partial out Bayesian *priors* in this way. This argument will be developed in the next section, which presents a brief review of the Anderson and Julesz paper.

### 3.6.2 Junction Signatures and Surface Interpolation

An alternative approach, also couched within a general Bayesian framework of cue integration, has been provided by Anderson (1994) and Anderson and Julesz (1995). For them, the principal benefit conferred by the Bayesian model is that it can be used to explain the recovery of 3-D surfaces from a range of image data largely independent of conventional retinal disparity. All visual cues are computed in parallel. Surface properties are resolved in the context of overall data, not a dominant stereoscopic data term *per se*. The observer's experience of the world is therefore treated as a 3-D construct that gives the best fit with image data. Their's is an ecological view: the resolution of untextured images reflects normal binocular image processing in an ambiguous context. Their argument is that instantaneous information regarding accretion and deletion is available from the domain of binocular parallax depicted at junction patterns in the context of prior learning of junction structures.

The ambiguities present in untextured stereograms are due to sparse disparity information and indeterminate monocular zones. Both of these factors arise from the paucity of texture elements. Rather than removing object-specific knowledge from the Bayesian scheme (as Nakayama and Shimojo have done in POGS), Anderson and Julesz argue that local image junction *priors* play a critical role. Two parallel processes are emphasised: constraint of retinal correspondence; and recovery of occlusion geometry from non-corresponding image features. Correspondence between matchable features reveals the relative depth of these features, while unmatchable features reveal the structure of occlusion. They argue that illusory contours are invoked by end-cut mechanisms responding to the breakdown in correspondence between the two images, and so:

> ...the decomposition of untextured stereograms into matchable and unmatchable features is responsible for generating the appearance of illusory surfaces*(Anderson and Julesz, 1995, p4)*.

*The Epi-polar Constraint*

The epi-polar constraint limits matching possibilities to points falling on horizontal slices through the picture plane; only horizontal coordinate differences can be used to reveal distance measures. Anderson (1994) argued that the system does not respond to vertical disparities in

this context. Vertical disparities are either not used, or provide only very weak binocular information. The epi-polar constraint is described in Fig 3.22.

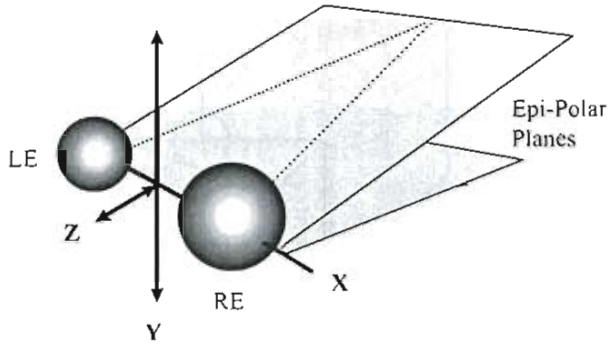



**Fig 3.22. The Epi-Polar Constraint**
Epi-polar planes are those passing through the lateral poles of the eyes, through the optic centres as shown. The Epi-polar constraint limits point matching in the Anderson and Julesz (1995) argument.

Given that the epi-polar constraint limits inter-retinal correlation, unmatchable regions are interpreted by the system as being occluded. There is, therefore, a key role in the Anderson and Julesz scheme for vertical image differences in the disambiguation of occlusion geometry. All vertical image differences are excluded from possible matches by the epi-polar constraint. Therefore vertical differences between the retinae can represent unambiguously occluded features.

*Resolution of Unmatchable Features*

Anderson and Julesz argue that the particular patterns of matchable and non-matchable features at contrast junctions contain rich information about occlusion. These patterns contain specific occlusion neighbourhood structures which betray the arrangement of surfaces. Inspiration for this approach comes from Guzman (1968), who argued that, in 2-D line drawings, it is possible for particular shapes of object junctions to be used as computational primitives (See Fig 3.23) to determine object shapes. Anderson and Julesz argue that a similar scheme applies in the case of stereoscopic junctions. However, in the stereoscopic domain, the signature junctions can define a 'neighbourhood' of surface relationships.

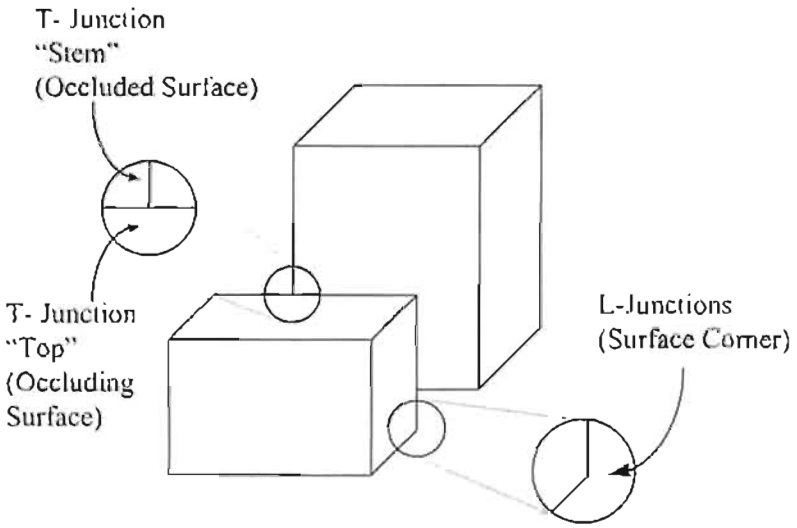



**Fig 3.23. Contrast junctions in 2-D images**
Guzman (1968) identified the importance of image junction shapes. T-junctions identify occlusion with the top being an occluding contour and stem being occluded. L-junctions represent the corners or cusp of a surface.

(adapted from Anderson and Julesz, 1995).

*A key role for vertical image differences*

Anderson and Julesz explored vertical image differences as key unambiguous instances of what they termed 'partial occlusion'. In Fig 2.24a, the occluding arms of an oblique cross are completed in a symmetric fashion. This percept is at odds with the horizontally - vertically aligned arms of the cruciform presented by Nakayama and Shimojo (1992), in which an asymmetry resulted in folded arms (see Fig 3.22). Anderson and Julesz claim that there is no disparity at these junctions. They are unambiguous L-junctions and as such are perceptually stable. So, binocular disparities are apparently insufficient to explain the induction of an occluding illusory contour by modal completion. The system parses two edges at different strata from the scene.

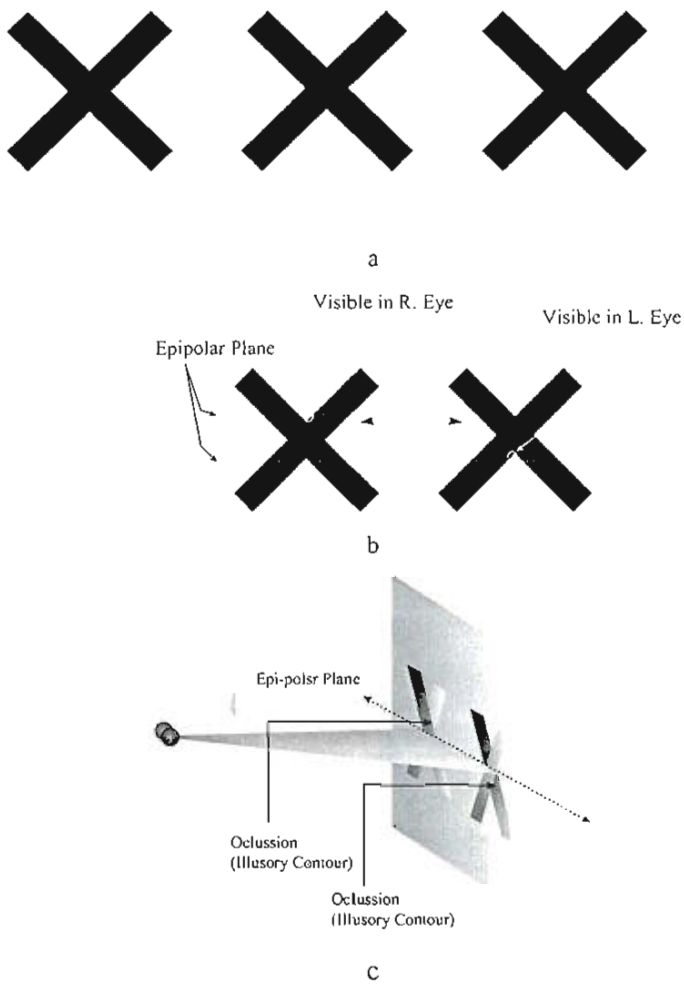



**Fig 3.24. Vertical image differences disambiguate occlusion**
In (a) crossed fusion of the left pair of images yields a percept of the black bar at $-45^0$ occluding the other. Fusion of the right pair reverses the occlusion. In both cases, where the two bars intersect, a stable illusory contour defines the boundary of the occluding bar. (B) demonstrates the position of so-called unpaired regions. The vertical image differences which yield non-matchable surfaces. (c) demonstrates the percept pictorially.

After Guzman (1968), Anderson and Julesz identify three specific stereoscopic junctions: L, l and T-junctions. These data suggest, that demonstrate that junction signatures, within the constraint outlined above, are critical for the interpretation of untextured stereograms.

*Stereoscopic L-junctions*

The L-junction defines an intersection between two surfaces which have precisely the same luminance. At such an intersection, there is no contrast difference which might constrain binocular matching. It is therefore critical that the system recognise the pattern of matching geometry. Figure 3.25 describes the eight L-junctions which must be resolved in the oblique cross. At these intersections, the matchable region of an oblique arm is relegated to the near

surface. In the figure below, this is the arm of the cross forming the + $45^0$ diagonal. Anderson and Julesz's description of neighbourhood structure is presented to the right of each image. Surfaces are termed F (far) and N (near). Note that the far segment in all of these descriptions is left unpaired. The eye of origin is designated L (left) and R (right). The dotted line signifies modal continuation of the near surface.

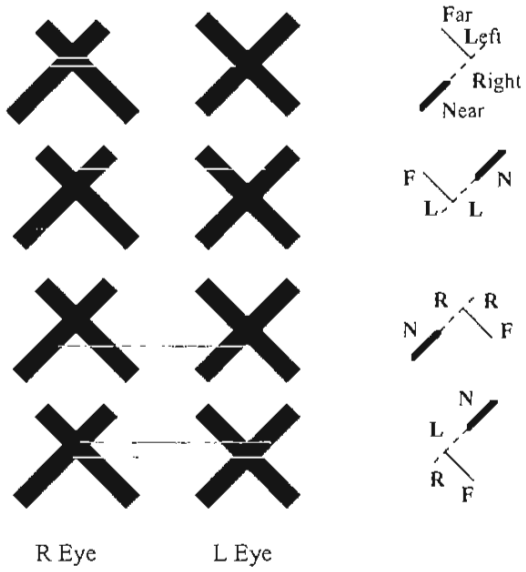



**Fig 3.25. Neighbourhood structure at L-Junctions**

Anderson and Julesz presented a series of invariant relationships arising at the intersection of surfaces that invoked illusory contours.

*Stereoscopic I-junctions*

Consider the stereoscopic I-junction, first demonstrated by Anderson (1994). In this case, the vertical displacement of a line-end in one eye with reference to the other reveals a bounding occluding illusory contour. Given the epi-polar constraint on point matching, the vertical differences in line lengths were left unmatched, and so were treated as being occluded. This scheme generated figure-ground segmentation which hides the unmatchable segment from the view of one eye. The unmatchable region is interpreted as a monocular zone. A selection of these effects is presented in Fig 3.26. The shape, or orientation, of the occluding illusory contour is a function of the eye of origin of the unmatchable region.

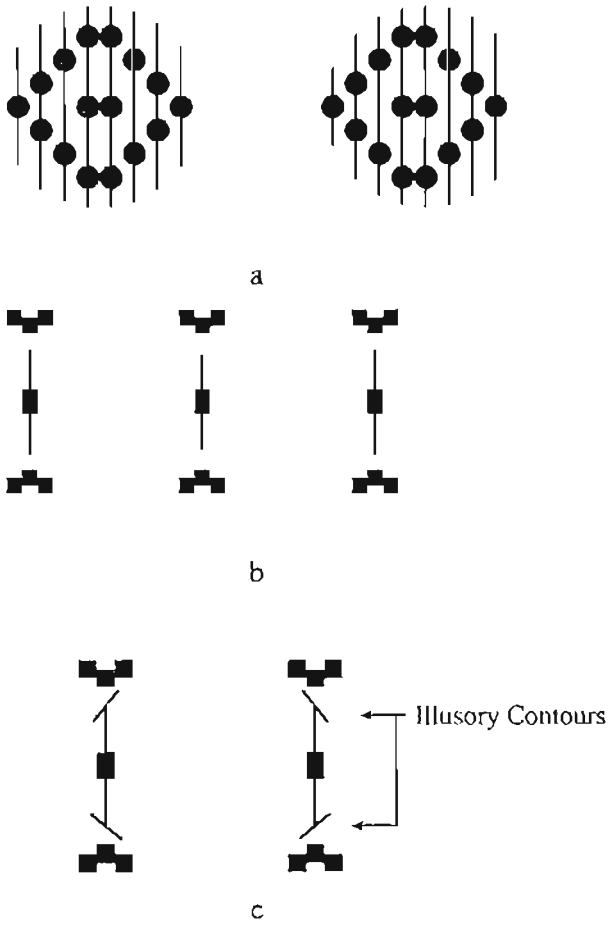



**Fig 3.26 Vertical partial occlusions**
Crossed fusion of these instances of partial occlusion, matched under the Epi-polar constraint, were claimed to yield illusory contours in isolation from any relatable contour.

Induction of illusory contours, in this way, is said to involve end-cut mechanisms initiated by disambiguation of the occlusion formation. Contours have both stereoscopic depth and orientation and people are highly sensitive to them. The disambiguation of matching geometry means that the contours invoked are highly stable.

*Stereoscopic T-junctions and Transparency*

The final junction structure identified by Anderson and Julesz relates to the notion of transparency as previously demonstrated in neon colour and contrast spreading. Once again, the neighbourhood structure of the T-junction bears rich information regarding surface interpolation. We already have demonstrated the various transparency effects identified by Anderson and Julesz. They demonstrate that junction structures not only reveal the binocular depth stratification process, but they also dictate border ownership of the surrounding contrast

regions. Arguably, the mechanisms invoking illusory contours in transparent percepts displays are *epiphenomena*. They come from mechanisms organised to decompose matchable and unmatchable features in high contrast figures. Figure 3.27 demonstrates the generality of these effects.

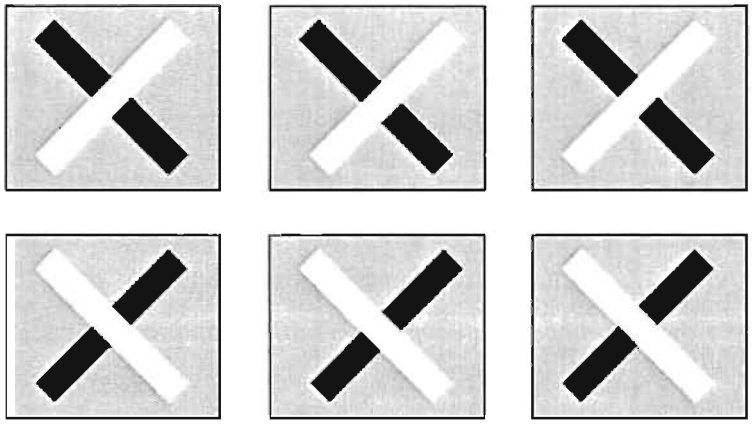


**Fig 3.27. Transparency and T-junctions**
Crossed fusion of these images demonstrates the phenomenological consequences of luminance order.

Anderson and Julesz (1995) have couched illusory contours, and the recovery of surface interpolation, in the terms of the correspondence problem. Their explanation of illusory contours relies upon *prior* junction structures and *end-cut* mechanisms. For example, they have demonstrated that the l-junction induces stereoscopic illusory contours in the absence of Kellman and Shipley's spatial relatability. This demonstration is said to be the pure case of illusory contour induction generated by partial occlusion. However, they do not attempt to reconcile their theory with evidence that *continuation* of contours is fundamental. Gillam's demonstrations, using random line stimuli, have shown that local end cut mechanisms are themselves not sufficient to explain illusory contour induction. Gillam's stimuli and Gulick and Lawson's sparse texture matrices, suggest that the issue of local junctions needs to be extended to segmentation at a broader scale.

In summary, it is interesting to note that both POGS and *partial occlusion* are terms borrowed from the artificial vision/computational vision literature. These approaches emphasise the ill-posed nature of the visual scene. This leaves such theories in danger of neglecting higher order information available in the image itself or in the spatio-temporal

patterns contained within perspective array. Bayesian approaches arguably have neglected some of the spatial information available in untextured images, such as orientation disparities, larger scale angular image difference and the possibility of double, and even multiple, fusion schemes. It will be argued that these theories of stereopsis have tended to underestimate the system's access to binocular perspective information.

## 3.7 Isomorphism in Illusory Contours

Carmen and Welsh (1993) have recently taken an alternative approach to illusory figures in stereopsis. Carmen and Welsh emphasise neural 3-D object representation. This they termed *morphism*. They claim:

> ... perception of both real and illusory forms results from a visual process which constructs a representation of three-dimensional geometry from various spatial cues (1993, p586).

This idea revisits functional equivalence of *real* and *illusory* contours. They suggest that some separate, as yet undisclosed, neural substrate underpins the detection and reconstruction of the three dimensional geometry of such figures. One function of this neural activity is the construction of a model object representation. Carmen and Welsh (1993) have identified what they term *View Stability* and *Morphic Generality* in these illusory forms that is shared by *real* objects. These terms describe the findings that observers can reliably identify the same illusory object across manipulations of view point, and can recognise or classify generic shapes or morphologies across subtle manipulations. An adaptation of the Carmen and Welsh 3-D illusory surfaces is provided in Fig 3.28. In many ways their argument is a re-statement of Kanizsa's Gestalt notion of *completion*, but in three dimensions and with a neural basis.

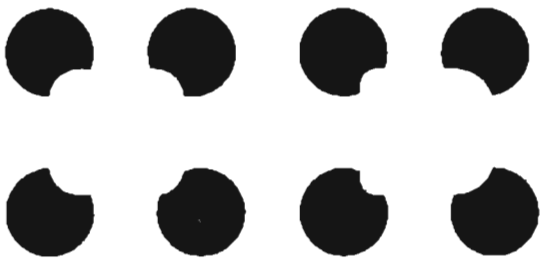


**Fig 3.28. Three-dimensional illusory surface**
Carmen and Welsh explained that the figure above revealed the operations of a neural substrate for 3-D representation. Crossed fusion yields the percept of an illusory surface curved in three dimensions.

The work of Carmen and Welsh brings this review of arguments surrounding illusory contours full circle. In demonstrations of 2-D illusory figures; in the stereoscopic spreading phenomena; in the perceptual asymmetries described throughout this chapter; and in Anderson and Julesz's work, there is much evidence to contradict their representational approach. For instance, many instances of apparently 'formless' contours and ambiguous surface effects have been demonstrated. It is logical that when such boundaries are perceptually stable they will be perceived in the same manner as any unambiguous luminance *edge*. On the other hand when boundaries are unstable, or in some way ambiguous, incomplete, or in conflict, then Morphism may be degraded.

## 3.8 A research problem

This chapter has presented a review of research in the general area of 3-D illusory percepts in stereopsis. Stereoscopic illusory contours and related effects appear to be evidence that binocular vision plays a key role in perceptual organisation and image segmentation beyond point-disparity computation.

Two main accounts, the Form Computation and Surface Heuristic approaches have been identified. These approaches take somewhat different views on the theoretical meaning of stereoscopic illusory percepts. There appears to be a gap in these literatures regarding the relationship between the geometric structure of half-images and the percepts reported by observers. The remainder of this thesis presents an examination of that relationship.

# Part 2 Exploring binocular vision processes underlying perception of stereoscopic illusory contours and surfaces

***Summary:*** *The second part of the thesis presents studies designed to explore the binocular vision processes that extract information from the 2-D layout of half-images in a way that yields 3-D illusory contours and surfaces. The aim of the series of experiments was to reveal binocular image processing mechanisms able to recover stereoscopic depth in the absence of texture.*

***Chapter 4*** *describes an initial experiment designed to establish the relationship between the phenomenal characteristics of a 3-D illusory figure and stereopsis. Experiment 1 employed three metrics: seen depth, seen slant and lightness judgement. These three metrics suggest that the appearance of an illusory 3-D percept might be a product of binocular image processing mechanisms.* ***Chapter 5*** *proposes mechanisms that could account for the stereoscopic Kanizsa percepts. These are tested in* ***Chapter 6*** *and* ***Chapter 7****. Finally, the generality of the mechanisms is tested in several experiments concerning stereoscopic Ehrenstein figures in* ***Chapter 8****.*

# 4. Processing binocular geometry and the perception of a stereoscopic Kanizsa square

***Summary:*** *This chapter describes the established phenomenological character of SKS percepts. It then examines the stereoscopic response to untextured stereograms in some detail and proposes that the SKS percepts may be a product of some basic response to disparate subtense of large-scale image regions. An exploratory experiment investigated the manner in which an illusory 3-D percept is related to the 2-D layout of Kanizsa square half-images. To test the relationship, the experiment employed three metrics: seen slant, depth and lightness. Collectively, these metrics characterise the important properties of the percept ie. separation of surface layers, and perceived depth.*

## 4.1 Perception of a 3-D Kanizsa square

When observers fuse disparate Kanizsa square half-images they report stereoscopic depth perception. This is despite retinal point-disparity being confined to just a few luminance contours. Experiment 1 sought to clarify the relationship between disparity in the SKS half-images and the 3-D percepts achieved when SKS half-images are fused.

### 4.1.1 Phenomenology associated with perception of a SKS

Since Julesz invented the RDS, the 3-D experience provided by stereopsis has mostly been thought of as detection and computation of depth from retinal point-disparity. For an illusory percept, the concept of computational stereopsis is less clear. The percepts appear to be underdetermined by physical visual information, and in particular, retinal disparity.

Figure 4.1 revisits the phenomenology of a SKS to clarify aspects of the 3-D percept established previously (see for example Bloomfield, 1973; Gregory, 1972; Lawson and Gulick, 1967; Simmonds, 1975, Grossberg, 1994; Anderson and Julesz, 1995; Gregory and Harris, 1974; Fujita, 1993). Five phenomenal characteristics of the SKS percepts can be summarised as follows:

i. **Illusory contours:** When crossed disparity is applied to the Kanizsa square[1] (cross fuse L-M) binocular fusion yields *illusory* contours that cross physically homogenous regions between the mouths of the pacmen.

---

[1] The half-images in Fig 4.1 were constructed by drawing a square, the same luminance as the page, to partly obscure four black circles. The configuration was repeated to create three Kanizsa square half-images whose pacmen were identical. Disparity was then introduced by shifting the position of the overlayed square in half-image (M) by about 2mm, in the horizontal, relative to (L) and (R) (see Fig. 4.1b). This changes the shape of the mouths of the pacmen in (M). A conventional retinal disparity is present at the vertical contours bounding the mouths of the pacmen.

ii. **Illusory form:** Collectively, these illusory contours give the impression that they bound a square. The square looks like a 3-D form, ie. an identifiable visual object with a specifiable shape.

iii. **Illusory lightness differential:** The illusory square looks lighter than the surrounding white space. This is similar to the classic 2-D figure. However, the stereoscopic illusory contours tend not to fade when fixated upon.

iv. **Separation of surface layers:** The square appears to stand forward of the page. It looks like a surface layer that is separated from the depth plane of the pacmen. The Kanizsa square takes the appearance of an opaque white surface floating above the pacmen

v. **Perceptual asymmetry**. Fusion of half-images, in which disparity has the same magnitude but opposite sign, generates a very different 3-D perceptual outcome. Illusory contours form at different parts of the percept (cross fuse M-R in Fig 4.1a). The pacmen now look like portholes. A square looks like it is partly visible through those portholes. Each porthole appears to be bounded by illusory contours. These contours visibly continue the semicircular pacman boundaries to form complete portholes. Also note also that the space in between the portholes looks opaque and equidistant with the porthole boundaries. Anderson and Julesz described the difference between the percepts achieved at the two signs of disparity (crossed and uncrossed) as a perceptual *asymmetry*. Their terminology will be retained.

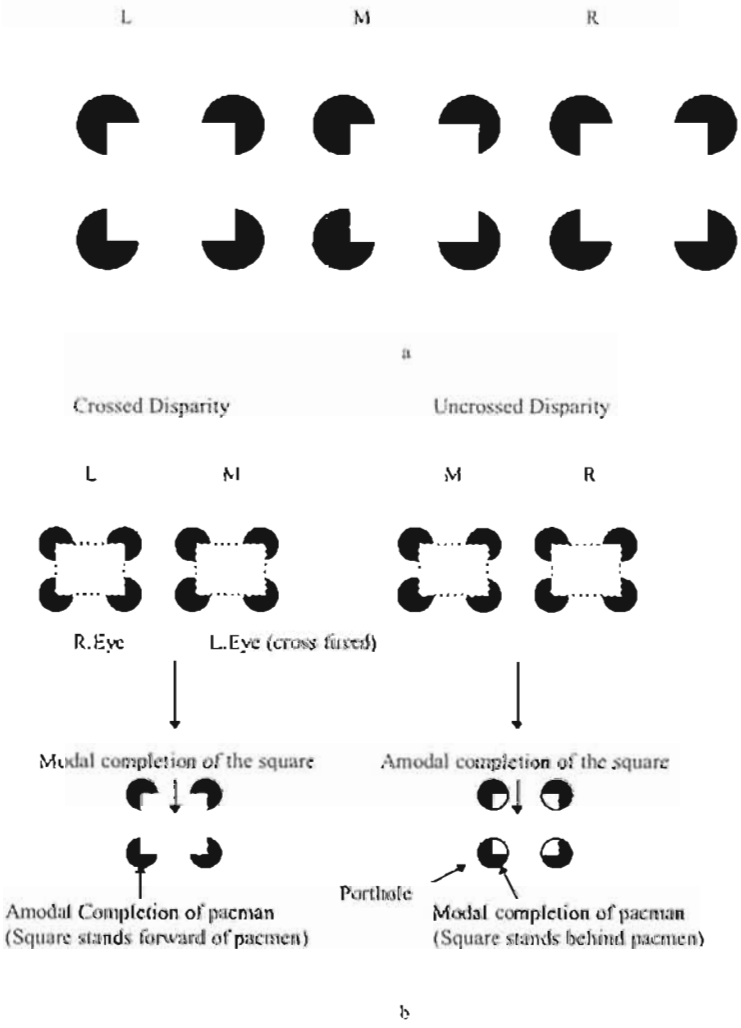



**Fig 4.1. A stereoscopic Kanizsa square**
In 4.1a crossed fusion of the left (L) and middle (M) half-images yields the perception of a distinctive Kanizsa square standing forward of the pacmen. The white region bounded by the illusory contours appears to stand forward of the pacmen. Crossed fusion of (M) and (R) demonstrates the reverse disparity sign. An amodal square results. Pacmen appear like 'portholes' through which a square is partly seen (The sequence will be reversed for divergent fusers). The perceptual effects of interest are pictured in (b). Note that there is a difference in perceptual organisation between the two signs of disparity.

The next section very briefly reviews recent insights into stereopsis in untextured stereograms of the type that generate stereoscopic illusory contours and / or the perception of separated of surface layers.

## 4.2 Recent explanations of 3-D illusory percepts in untextured stereograms

It was established in Chapter 3 that explanations of illusory phenomena associated with untextured stereograms vary considerably. What is clear is that assignment of matching

contours to different depths (in the classical sense of stereopsis) cannot account for the percept achieved when these images are fused. The intriguing phenomenology associated with the SKS suggests that stereopsis somehow modulates 3-D perceptual organisation. How stereopsis does this is not yet clear. The next sections briefly review two major recent proposals previously identified.

### 4.2.1 Form Computation Approach

Grossberg has proposed a computational theory (FACADE theory, 1994). The computational foundation of FACDE theory are neural networks whose function is extraction of 3-D form from the luminance patterns and retinal disparities available.

FACADE theory proposed that four functional components contribute. First, conventional disparity between corresponding luminance contours was computed, yielding depth. Second, the pre-visual closure of fragmented shapes was achieved by a Boundary-Contour-System. This mechanism connected contours such as those at the mouths of the pacmen, via boundary projections created by end-cut mechanisms. End-cut mechanisms are described as the output of *end-stopped* cortical cells that define contour termination. Finally, binocular FIDOs of the Feature Contour System fill-in extant spaces between these contours generating visible surface layers. Hence, neural networks underlying the perception of 3-D form drive the visual response. Grossberg (1994) states that these percepts are:

> ... manifestations of the mechanisms whereby the visual cortex generates informative 3-D representations of boundaries and surfaces with which to perceive the visual world (p. 48-49).

### 4.2.2 Surface Heuristic Approach

The other major theoretical approach maintains that key processes underpinning perception in untextured stereograms are inferential, since corresponding points are so scarce. As described in Chapter 3, an essential issue for this approach has been the perceived separation of surface layers and not extraction of form *per se*. The approach uses a modified Bayesian cue-integration framework. It sees perception as a statistical representation based on the probability that a certain image I, will be generated by a particular Scene S. Hence the manifest phenomenology is a product of the expression p(I|S). Two possible inferential mechanisms have been proposed: Nakayama and Shimojo's POGS and Anderson and Julesz luminance junction structures.

Nakayama and Shimojo's Principle of Generic Sampling (POGS) is the notion that the

system responds to ambiguous binocular stimulation (where retinal disparity is scarce) by inferring the most likely 3-D structure that would generate such a configuration of luminance. This inferential response was thought to be achieved by neural populations that had *learned* by ecological image sampling to anticipate certain three-dimensional organisations.

When faced with an under-determined pattern of retinal disparity, these networks responded by signalling learned 3-D perceptual structures (perceptual organisations). Monocular features were thought to cue a binocular neural inferential response with the result being a depth percept from minimal disparity information (Nakayama, 1996).

Anderson and Julesz (1995) took a different view. They were more concerned with how the system worked out which surface layer *owned* the luminance contours in the Kanizsa configuration. The argument was that the presence of non-matchable features at luminance junctions signals perception of an occluding edge. Mechanisms underpinning the inferential perceptual response they proposed were shown most strikingly by Anderson (1994). Anderson's l-junctions demonstrated that neither continuation nor completion were necessary to explain illusory contours and the separation of surface layers in stereopsis.

Anderson's extraordinary stimuli were constructed using thin line image pairs, even single line pairs, whose lengths were disparate. On fusion, subtle but distinctive illusory contours were generated in the absence of relatable features. Anderson and Julesz decided the system must decompose half-images into matchable and unmatchable elements. Unmatchable elements were interpreted by the system as monocular zones of a distant surface.

Anderson and Julesz placed particular importance on the so called *epi-polar constraint* which delegates all vertical image features unmatchable. Constraint of vertical disparities to monocular zones meant the system could identify particular sites at which the separation of surface layers must occur. Anderson (1997) has subsequently termed separation of surface layers at these sites *surface scission*, an inferential response peculiar to *a priori* 3-D luminance junctions (eg, L, X, T and l-junctions, see Chapter 3 for details).

### 4.3 Assignment of untextured *space* to depth

It is indeed a puzzling aspect of the SKS percepts that the system achieves separation of surface layers when point-disparity is so sparse. The two approaches summarised above do not address the possibility that phenomenology of the SKS relates to the manner in which the system

resolves the specifiable differences between the half-images presented to it. The next sections explore this issue.

### 4.3.1 Separation of surface layers in the Kanizsa square

In a random-dot stereogram, corresponding dots are assigned to a particular depth according to their unique retinal disparities. The depth of almost all texture features can be explained by local point-disparity computation (the exceptions being unpaired dots in the matrix). The question is, how is untextured space assigned to a depth plane in the absence of local point-disparity across texture gradients? In particular, how is surface separation achieved (perception of a depth step) in the absence of physical disparity discontinuity (a relative difference in point-disparities)?

Consider the RDS in Fig 4.2. The central matrix of disparate dots is bounded by a rectangle (for the purpose of subsequent demonstration). Crossed fusion of pair L-M sees the central matrix of dots (bounded by the small rectangle), stand forward, and in fusion of M-R the disparate matrix stands behind the non-disparate dots.

In each half-image there is a set of dots, adjacent to the disparate matrix, that is unmatchable. This region is represented schematically in 4.2b (for pair L-M) by a horizontal slice through the centre of the RDS. Unmatched dots are represented as cross-hatched squares, matched disparate texture is represented by grey squares and matched non-disparate texture is represented by white squares.

Figure 4.2c represents the resolution of inter-retinal correspondence in crossed fusion of pairs L-M. The figure on the right (4.2c) shows a stylised stereoscopic response. Matched pairs of disparate and non-disparate dots are fused. The result is that the disparate dot matrix (grey squares) stands forward of the projection plane (P Plane). Non-matched regions are resolved as monocular zones adjacent to the disparate matrix but on the same depth plane as non-disparate dots.

In Fig 4.2d crossed fusion of pair M-R is shown. The disparate region is seen to stand behind the non-disparate matrix at the P Plane. The unmatched regions are observed to stand on the same plane as the disparate dots. In these figures then, all dots in the matrix are fused and assigned to an appropriate depth plane except the un-paired dots. Hence, the stereoscopic depth of all regions of the matrix except the unpaired dots is differentiable by their relative positions in the overall matrix.

In fact the manner in which stereopsis achieves the depth of unpaired regions of the RDS has also been the subject of some debate. Most computational theories of stereopsis have imposed an *a priori* heuristic such as Julesz' *most distant surface rule* to explain how the system assigns depth to such features. As described in detail in Chapter 3, several theorists have recently accorded new importance to monocular regions of the RDS, as well as in untextured stimuli, in actually achieving the separation of surface layers.

Nakayama and Shimojo have argued that the system follows ecologically valid rules in achieving the perception of partial occlusion. This is plausible since, in natural vision, the unpaired regions always occur at the temporal projections adjacent to an occluding edge. The system may use this as a heuristic to resolve occlusion in untextured space. Their approach is that the responses of neural networks, neural populations, are established by learning.

Subsequently, Nakayama and his colleagues emphasise the importance of the eye of origin of non-matched monocular features as cues that trigger over-learned network outputs. These responses bias the interpretation of three-dimensional perceptual organisation where non-matched regions are ecologically valid, that is, where they occur as predicted by natural visual experience.

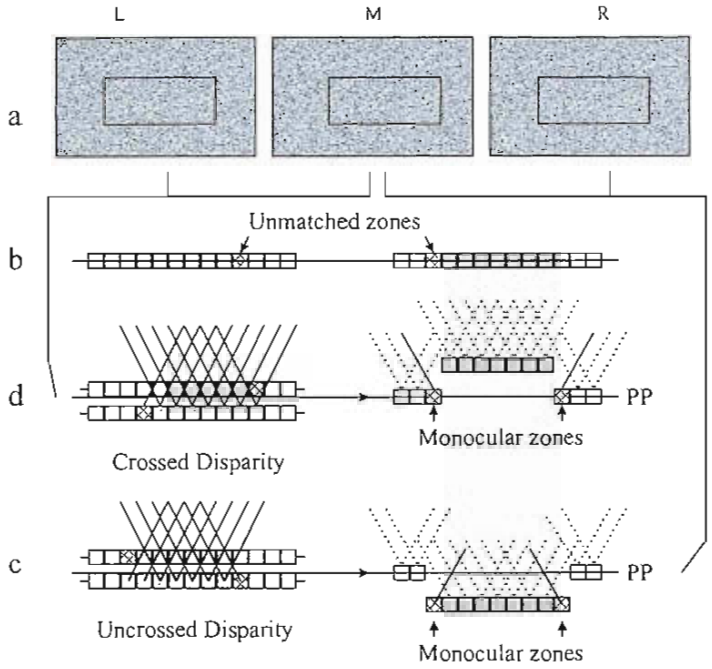



**Fig 4.2. Retinal correspondence and monocular zones in conventional RDS**
The set of half-images in (a) are meant to demonstrate the basic phenomenology associated with a typical RDS. A rectangle has been drawn to bound the disparate section of the random dot matrix. Crossed fusion of the pair L-M sees the disparate dots of the matrix stand forward of the non-disparate dots. Crossed fusion of pair M-R sees the disparate dots stand behind the non-disparate matrix. Between the two signs of disparity, the assignment of depth to surface layers is a product of local disparity resolution that involves matching corresponding points in the matrix then computing their distance according to the retinal disparity between them. This process is summarised in Fig b,c and d which represent the dots falling on an epi-polar slice through the stereograms in (a). Figure 4.2b shows the position of unpaired dots in the matrix. These unpaired dots are represented as cross hashed squares. Fig c represents the matching process in binocular fusion and the resulting stereoscopic phenomenology. Unpaired elements are assigned to the distant depth plane. Figure 4.2d represents the same process at the reverse sign of disparity. Again, the unpaired regions are assigned to the distant depth plane.

Anderson and Nakayama (1994) outlined a hypothetical binocular receptive field capable of detecting the breakdown of correspondence in the unmatched regions. However, as Anderson and Julesz have explained this is clearly not feasible in the case of untextured stereograms. In untextured stereograms, there is no breakdown of correspondence adjacent to an occluding edge. Yet, the separation of surface planes is still achieved in the complete absence of texture.

Anderson and Julesz later argued that non-matchable features of the untextured stereogram such as vertical disparity at luminance junctions could enable recovery of occlusion geometry (interpreted as monocular zones). The system inferred a depth step from this pattern of luminance resulting in propagation of illusory contours. A potential problem for this

approach is that perceptual asymmetry in the separation of surface layers is evident even in untextured stereograms with no unmatchable features at all.

Figure 4.3 shows a very simple untextured stereogram in which perceptual asymmetry occurs between the signs of disparity. Note that the disparate positions of the central white rectangle compared to the black rectangle are the same as the rectangle bounding the disparate dot matrix in Fig 4.2a above.

Crossed fusion of L-M yields the percept that the white rectangle stands forward of the P Plane. In this percept, assignment of depth to surface layers appears to be equivalent with pair L-M in the RDS. By this I mean that a central disparate region of the half-images stands forward of zero disparity at P. Reversing the sign of disparity, cross fuse pair M-R, sees the white rectangle standing behind the outer boundary of the black rectangle. *The contour at which surface layers separate appears to be different.* The question is, of course, *why*?

The perceptual difference between Figs 4.2 and 4.3 is that the separation of depth planes happens at a luminance step, a single contour, in the untextured figure, but at a discontinuity of disparity in the textured stereogram. The separation of depth planes in Fig 4.3 is, manifestly, not defined by a difference in disparity, a disparity discontinuity, between the near and distant depth planes. Separation of depth planes is achieved at a single luminance discontinuity, a single contour. The same contour differentiates the depth of both the near and the distant depth planes.

In Fig 4.3b, a central horizontal slice through the half-images in (a), shows that no unpaired regions exist. The only local retinal disparities present are at the vertical contours of the white rectangle. Figure 4.3c, which represents crossed fusion of L-M, suggests that the separation of surface layers to different depth planes might involve assignment of parts of the black rectangle, adjacent to the vertical boundaries of the white square, as *monocular*.

In Fig 4.3d, which represents crossed fusion of M-R, surface layers do not separate at the boundary of the white rectangle. Anecdotally at least, the contour at which separation of the depth planes occurs is at the outer boundary of the black rectangle, though the depth of the untextured black space is somewhat indeterminate. The black rectangle looks like a window.

The apparently straightforward untextured stereogram in Fig 4.3 is theoretically interesting because, in addition to the paucity of retinal disparity, there are no unpaired sectors or luminance junctions that might signal assignment of surfaces to depth planes. This simple

stereogram seems to undermine the idea that unpaired regions of untextured stereograms reveal surface separation and lead to induction of stereoscopic illusory contours.

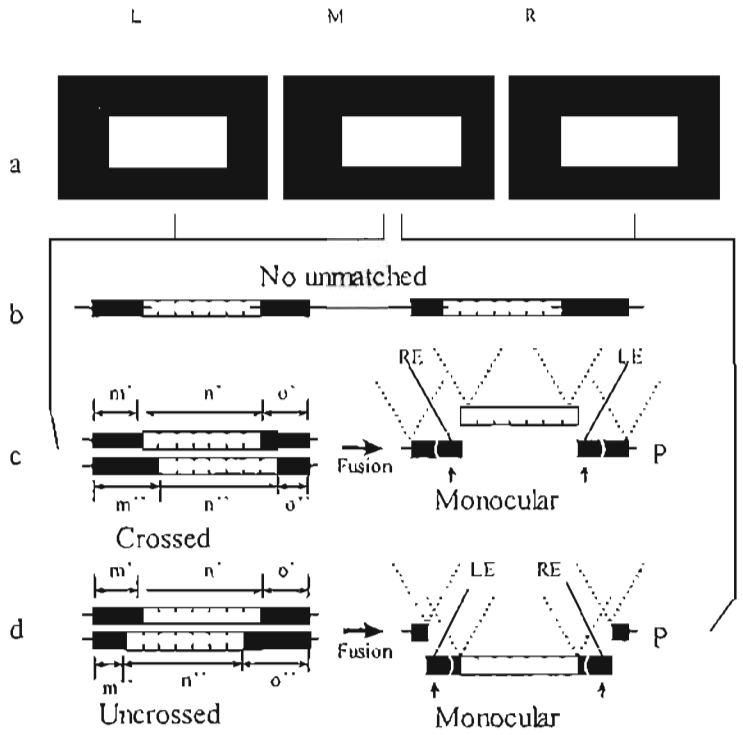



**Fig 4.3. Perceptual asymmetry in an untextured stereogram**
The set of half-images in (a) demonstrates a very simple untextured stereogram. Crossed fusion of the pair L-M sees the white rectangle stand forward of the black rectangle. Crossed fusion of pair M-R sees the white rectangle stand behind the outer boundary of the black rectangle and the black rectangle looks like a window. Assignment of depth to surface layers in these untextured pairs cannot be a function of local disparity resolution alone. This is because the same contour in each eye must somehow yield the percept of both the near and distance surface. As shown in (b) the half images define no *unpaired features* that might disambiguate the site of surface separation. Nevertheless, the system achieves just that. Figures (c) and (d) represent the fusion process. One aspect of binocular fusion evident here that has not been examined in detail previously, is the pattern of disparate binocular subtense of the half images. In (c) disparate subtense exists between region m' and m'', and o' and o''. No difference in subtense exists between n' or n''. In geometric terms, these differences in subtense are precisely reciprocal to the magnitude of disparity at the boundary of the white rectangle. Fusion at (c) and (d) must create a monocular zone where surface layers are separated. These zones are circled by a grey ellipse. Note that the magnitude of disparate subtense is the same in (d) but the sign of disparity is reversed.

An alternative to the Surface Heuristic approach is that there is more information available in the retinal image than has been recognised. Indeed, Fig 4.3 identifies another possibility that deserves consideration. Notice that fusion of half-images depicted in Fig 4.3c and 4.3d involves untextured regions that exhibit *disparate subtense* (relative horizontal disparity) at the black regions adjacent to the white rectangle. The magnitude of disparate

subtense must be precisely reciprocal to the disparity *between* the corresponding contours in each eye's view.

Since disparate binocular subtense is a geometric function of the magnitude of local retinal disparity in untextured stereograms, analysis of its role in 3-D perceptual organisation may help us to understand perception of stereoscopic illusory surfaces. In fact disparate subtense was identified by Grossberg as a component of FACADE theory.

### 4.3.2 Allelotropia in untextured space

Allelotropia is the deformation of each monocular image that is required to achieve the cyclopean, singular, binocular percept. The concept has been used to describe the manner in which the system interprets large scale half-images, such as human faces (Kaufman, 1974). However it stems from the 19th century propositions that the single vision achieved in binocular fusion is the average of the visual directions of corresponding contours in each eye (Hering, 1848, cited in Ono and Mapp, 1995). The textbook example of this effect is as follows. When one fuses the letters below, the "F" in the fused percept is seen at a central position between the non-disparate letters. Binocular fusion places the "F" at the average of its disparate position in both eyes (Grossberg, 1994).

EF G E FG

Grossberg has argued that allelotropia in untextured stereograms was achieved as binocular FIDOs filled-in unpaired regions in a Da Vinci configuration (in which unpaired elements existed) from appropriate *disparity pools*, after conventional stereopsis had resolved paired features. Grossberg claimed that this avoided the perception of 'gaps' in the binocular percept. However, Grossberg did not treat the concept in any great detail. The advantage of his analysis, is that it sets the problem in the context of an adaptive rather than a reflexive code of visual processing. A disadvantage for the psychology of binocular vision is that his thesis is not really testable in functional terms. It can only be falsified by an analysis of visual neural architectures.

Experiment 1 aimed to carefully manipulate the binocular information intrinsic to the SKS configuration and then to measure the subjective perceptual response of subjects to those manipulations. This was to begin to clarify the possible role of the stereoscopic response to

disparate subtense in the perceived separation of surface layers and modulation of the phenomenal character of the percepts, such as, illusory contours in the SKS.

Two possible components of the stereoscopic response in achieving the 3-D SKS are the distance of the illusory plane (seen depth) and the orientation of the plane (seen slant). To examine stereopsis in the initial experiment of the project, both a depth probe method and a slant estimation method based on Gillam (1968) were employed.

### 4.4 The stereoscopic slant response

The contribution of binocular vision to the perception of slant has also been presumed to be a stereoscopic response to the gradient of retinal disparity across an inclined textured surface. An empirical assessment of stereoscopic slant perception was first undertaken by Ogle (1950).

Ogle's geometric effect was produced using an afocal lens to magnify one eye's view of a fronto-parallel surface in the horizontal plane only. When observers fused the natural perspective view and the magnified perspective views of a surface they perceived the surface to be rotated about a central vertical axis. Two afocal lenses oriented obliquely in opposite directions generated the perception of slant about a horizontal axis. The second manipulation created disparate horizontal shear of each half-image of the surface. Recently, van Ee and Erkelens (1995) have shown that slant in any oblique axis can also be generated using appropriate components of horizontal magnification and shear.

Ogle's disparate magnification of a square is shown in Fig. 4.4a (vertical axis - horizontal magnification) and 4.4b (horizontal axis - horizontal shear). When presented with images manipulated in this way, observers report a slanted stereoscopic trapezoid where the distant edge of the trapezoid appears larger than the near edge. This is likely to arise because Ogle's geometric formula does not account for the change in the monocular images cast as a planar shape of fixed dimensions rotates. In fact Ogle's formula is constant with a rotated trapezoid. It appears that the perception of a trapezoidal shape points to very subtle size distance scaling of image proportions by stereopsis.

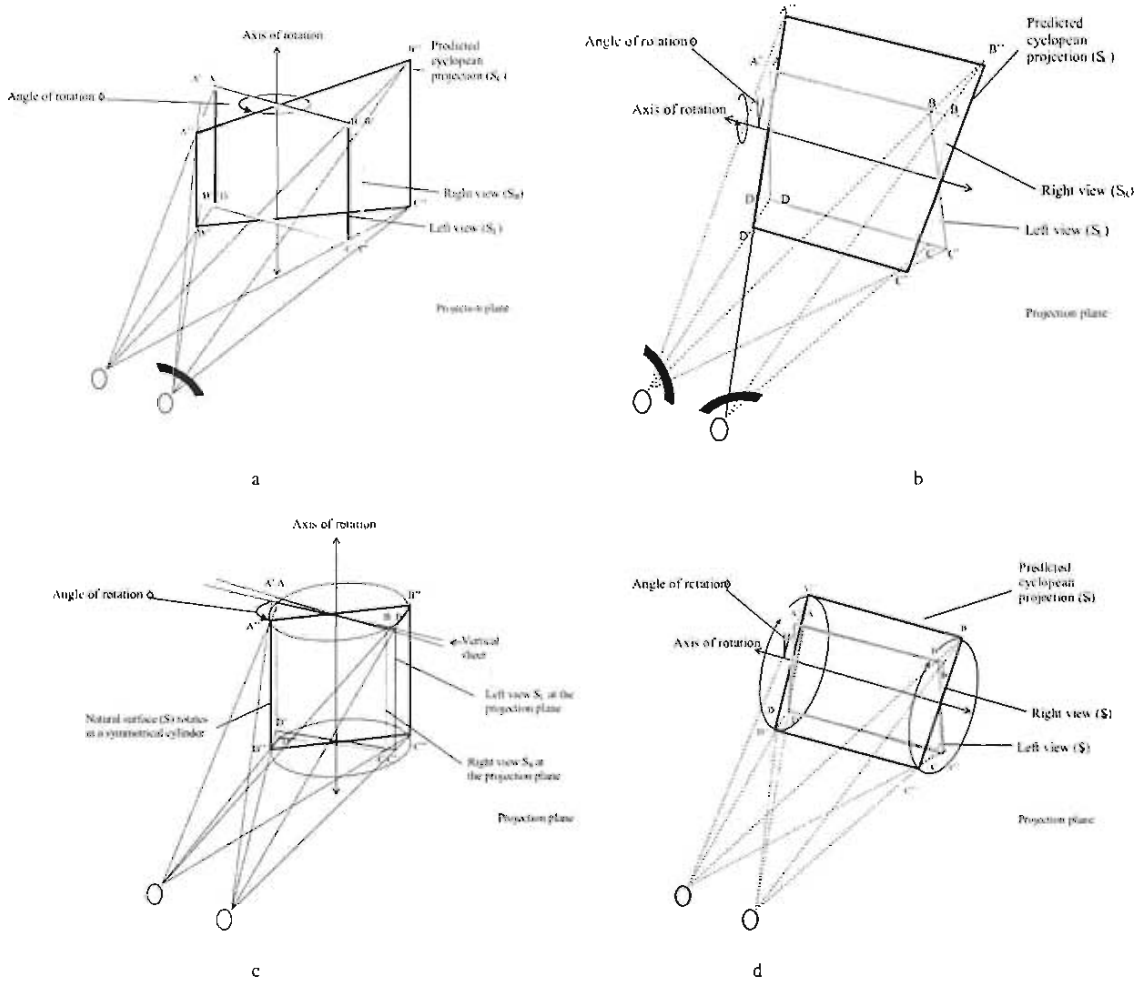



**Fig 4.4. Binocular parallax and stereoscopic rotation about the vertical and horizontal axes**

In (a) Ogle's magnification applied to a fronto-parallel plane, ABCD, gives the transformed magnitude A'B'C'D'. This predicts a trapezoidal stereoscopic surface A"B"C"D" on fusion – a rotation about the vertical axis (adapted from Gillam (1968)). Figure (b) shows Ogle's magnification applied to create relative horizontal shear of a fronto-parallel plane. This predicts a similar trapezoid rotated about the horizontal. In Figs (c) and (d) it is shown that a square surface ($S_C$) surface whose boundary prescribes the corner points A"B"C"D" rotates in a cylindrical volume of space. At a particular degree of rotation, each eye's view the shape cast by the surface at the distance of the P Plane will be a product of binocular parallax. In the right eye the image cast by the square ($S_R$) will subtend ABCD, and in the left view ($S_L$), its image will subtend A'B'C'D'.

Figures 4.4c and 4.4d demonstrate that a 2-D planar square rotated about its horizontal or vertical axis will describe a cylindrical arc in 3-D space. The disparate image shapes cast by the plane can be simulated at the plane of projection, that is, on a computer monitor as in my experiments. The basic technique is quite ancient. Indeed, the half-images created in this manner, equate to Wheatstone's perspective drawings (the first demonstrations of retinal disparity in the late 1830s). Each of the images projected at the computer monitor represents a perspective projection of the 2-D images of a 3-D object viewed at a specifiable distance. In other words the image at the retina simulates a natural visual perspective geometry.

The projection diagrams in Fig 4.4c and 4.4d essentially provide a geometric model of the silhouette of a square rotated in 3-D space. In the following experiment, this projection model was used to generate transformation of half images as described in Appendix A. Ogle's magnification could then be applied to the half-images to arrive at stereo-pairs that simulated natural binocular perspective seen from each eye's different view of a planar rotated surface. These image transformations enabled examination, with some precision, of the manner in which observers resolved the disparate size and shape of SKS images.

In the experiment to follow I also explored subjects' judgements of the overall depth of the SKS when fused. This was achieved by drawing the Kanizsa square in different positions in each half-image relative to the pacmen. These two manipulation were thought to provide a more thorough account of the stereoscopic response to SKS half-images than previous research. The next section outlines the general method developed to study the seen slant and seen depth responses all experiments.

## 4.5 Summary of the general method for slant and depth estimation

### 4.5.1 Apparatus

Stimulus generation was handled by a 4Mb Cambridge Research Systems (CRS), Vision Stimulus Generator (VSG 2/3) card. The VSG was mounted in an IBM compatible 486 D266 PC. Custom software routines were used to present the stimuli, manage the comparison stimulus and collect data. Images were presented on a 21in Vision Research (VR) monochrome monitor with a rapid phosphor decay, reducing cross-talk between images to a minimum. The VSG synchronised CRS Ferro-Electric Shutter-Valve Goggles with the monitor. Frame rate used was 230 frames per sec (115 per eye/sec).

### 4.5.2 Comparison stimulus

A planar stereoscopic comparison stimulus was employed both as a depth probe and in gauging subjects' slant estimations[4]. An octagonal shape with oblique lines drawn across its surface was chosen. This shape was thought to maximise orientation, angular and point-disparities. The comparator included an octagonal boundary (6 pixel line-weight) as a zero disparity reference frame. A central octagon (1 pixel weight) acted as the actual slant comparator. The comparator operated in two separate modes as a slant estimation device and as a depth probe. At the start of each trial, the comparator was presented at fronto-parallel. The comparator was activated by depressing the joystick trigger. Unless the trigger was pulled, movement made no change to the comparator.

*Seen Slant*

Stereoscopic rotation of the comparator was achieved by depressing the trigger then tilting the joystick forward or backward. In the case of rotation about the vertical axis, the joystick y-coordinate value was used to define the theoretical monocular transformation and retinal disparity appropriate to stereoscopic rotation.

In order to filter noisy output from the joystick and to maintain precision of comparator rotation, the y-value was only used to define a buffer threshold in movement of the joystick. The threshold meant that the joystick could be tilted forward thirty degrees or back thirty degrees with no effect. x-coordinate values were filtered out at this stage.

Tilting the joystick to yield a y-value greater or less than the defined thresholds controlled geometric transformation of the comparison stimulus. Stereoscopic rotation was incremented in steps of one degree as long as the joystick y-value was above threshold. This yielded stepwise stereoscopic rotation at a rate of about 10 degrees per second.

When the joystick was tilted to return y-values within the threshold limits, the incremental stereoscopic rotation ceased. The comparator maintained its position at that point and subsequent adjustments altered its rotation, until a button was pressed.

---

[4] The slant-estimation paradigm typically employs a real comparison stimulus. These are normally internally lit and operated mechanically by the subject. Such an arrangement proved impractical in this setting. Ambient illumination from the monitor and restricted field of view made it difficult to position a real comparator without confounding slant estimates with inappropriate shading and structural cues from the sides and the back of the comparator. An advantage of the virtual stimulus was that it was transformed by exactly the same geometric transformations as the test stimulus.

Moving the joystick to cross the opposite y-value threshold initiated slow incremental reduction of comparator rotation, or rotation in the opposite direction if disparity passed through zero degrees. This was to achieve a comparator that could rotated slowly and smoothly through about 170 degrees in either direction and passed smoothly through zero degrees from any position. When subjects had set the comparator to match the orientation of the test stimulus the joystick was tilted within the upper and lower y-value thresholds. Then subjects depressed either of two buttons on top of the stick.

Subjects were able to precisely manipulate the comparator, cease its incremental rotation, change the direction of its rotation, and make fine manipulations in either direction before finalising their judgement by pressing a button. When the button was pressed the current rotation of the comparator was recovered from the transformation formulae and stored.

For rotations about the vertical axis, the procedure was precisely the same except that the x-coordinate value of the joystick was used to define the movement thresholds. Each time the button was pressed the joystick was re-calibrated automatically and the comparator was reset to fronto-parallel.

*Seen Depth*

Use of the joystick as a depth probe was achieved in a similar fashion. The joystick y-coordinate value defined a movement threshold. Horizontal disparity was applied to each of the comparator half-images when the stick was tilted forward or back beyond threshold. When the y-value crossed the threshold horizontal disparity was incremented by 1 arcmin. This was the resolution defined by the pixel width of the monitor. A maximum disparity threshold was set at about the crossed and uncrossed fusion limits. Subjects were able to manipulate the comparator's stereoscopic depth behind the P Plane or above the P Plane within those limits. Once again, depth judgement was recovered from the disparity transformation of the stimulus when a button was pressed, and was the stored.

In summary, the comparator half-images were geometrically transformed as described by the perspective model appropriate to the rotation angle (see Appendix A). Disparity was then applied by transforming each image by Ogle's magnification factor according to the rotation angle. The comparator could be stereoscopically rotated through $340^0$ about either axis and could be set forward of, or behind, the P Plane as a depth probe. The comparator returned the theoretical

rotation angle in degrees, or the positional difference between the half images appropriate to standing disparity, in arcmins.

The comparator subtended $4.5^0$. Its centre was positioned $10^0$ below the centre of the test stimuli. Luminance values were the same as the test stimuli. Smooth rotation was achieved by redrawing the transformed images between frames. The software routine negated image flicker by redrawing the comparator only at $1^0$ intervals of rotation, and only redrawing while the comparison stimulus was actually meant to be in motion.

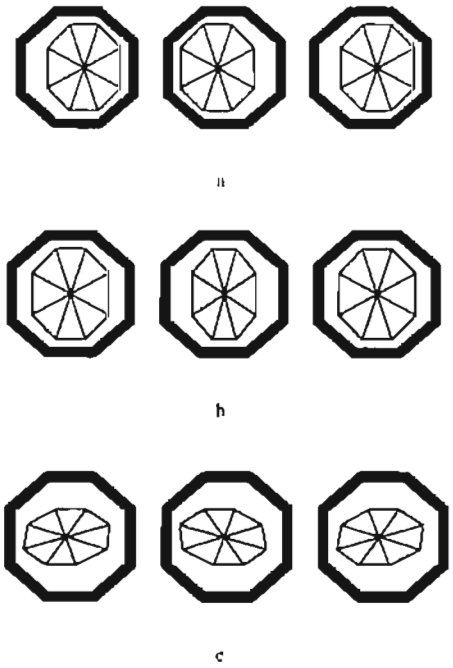


**Fig 4.5. Stereoscopic comparison stimulus**
Fusion of the half images in (a), demonstrates how the comparator could be used as a depth probe. In (b), the comparator is rotated about the vertical axis. In (c), it is rotated about the horizontal axis.

### 4.5.3 Participants

A panel of 96 undergraduate psychology students at James Cook University acted as participants in this research. Their ages ranged from 17 to 52 yrs, with a mean of 19.7 years and a Standard Deviation of 1.6yrs. First year students received course credit for participation. Other participants received no inducement. All participants were naive to the specific intention of the various studies. All were screened for normal stereo-acuity using the Randot test (20s arc at 40cm), and for normal or corrected-to-normal acuity on a standard Snellen chart.

### 4.5.4 Procedure

Subjects were seated at a desk in front of the monitor in a sound-attenuated testing room. Only dim ambient light from the monitor lit the room. Head movement was restricted using an adjustable chin rest.

Each subject was given a brief introduction to the concept of stereopsis. The operation of shutter-valve goggles was explained and they were comfortably fitted. A series of twenty practice trials were presented in random order. If subjects were not confident in manipulation of the apparatus after twenty trials, twenty more trials were completed. These involved 20 standing disparity judgements and 20 slant estimations about both axes. During practice the experimenter sat with the subject and responded to any questions that did not concern the nature of the experiments.

The comparison stimulus, all data collection, and progression through the trials was controlled by the subject, using a joystick. The subject navigated through the sequence of trials. There were no time limits in any trial.

## 4.6 Experiment 1 Perception of a SKS: seen slant, seen depth and seen lightness

The first experiment in the project was principally exploratory. It was designed to gain an understanding of the relationship between geometry of binocular parallax intrinsic to the stereograms projected and the phenomenology of the SKS observed when stereograms were fused. To do this, binocular disparity was defined mathematically then applied by shaping the "mouths" of Kanizsa square pacmen. The experiment aimed to capture the stereoscopic response and the phenomenal character of the resulting percepts along three dimensions:

1. Seen slant: this involved subjects setting a stereoscopic comparison stimulus to match the seen slant of an Kanizsa square
2. Seen depth: this involved subjects setting a stereoscopic depth probe to match the seen depth of a SKS
3. Seen lightness: this involved subjects rating the relative lightness of specified parts of the percept using a computerised Likert scale

### 4.6.1 Previous insights on slant estimation

Under laboratory conditions, observers tend to perceptually under-rotate surfaces defined by disparate textures. Under-rotation may be due to an information impoverished visual setting, that is,

a lack of depth reference frames used to scale retinal disparities (Stevens, 1991). Reference features such as disparity discontinuities, zero disparity frames, vertical shear, and filled monocular zones influence the interpretation of horizontal disparity gradients and the latency of slant judgements (Gillam and Borsting, 1988; Gillam, Findlay and Flagg, 1984, Gillam, 1994).

The context of retinal disparity can also impact upon slant judgements. For example, diagonal, horizontal and vertical markings have been shown to modulate the slant response in simple stereoscopic grids (Cagenello and Rogers, 1993; Gillam, 1968; Stevens and Brooks, 1988). There is also a tendency for slants about the vertical to be attenuated more than slants about the horizontal axis. This so-called *anisotropy* further demonstrates that seen slant is not predicted from point disparities alone:

> ...slant perception is not mediated by the fitting of surfaces to the depth value of point disparities. The slant elicited by a disparity ramp *[or a stereoscopically rotated surface]* depends on its context and orientation, which suggests higher order disparity arrangements, rather than point disparities, elicit a depth response ('act as stereoscopic primitives') *(Gillam and Rogers, 1991, p440).*

Two dominant explanations of slant-axis anisotropy have emerged in recent years. First, Cagenello and Rogers (1993) have argued that the degree of orientation disparity (relative differences in the orientation of matching contours between the eye's views) determines sensitivity to slant. They suggest that geometric asymmetry in orientation disparities between the axes explains the anisotropy. Secondly, Gillam has argued that slant-axis anisotropy is partly due to an asymmetric resistance to perspective conflicts between the axes.

Gillam and Ryan (1992) established that both slant-axis anisotropy and slant attenuation were substantially reduced in perspective-corrected configurations, that is, when perspective cues and disparity were congruent rather than in conflict. Nonetheless, Ryan and Gillam (1994) showed that perspective congruence between disparity and the configuration of surface markings still generated slant-axis anisotropy. Further, cyclopean edges in random-dot stereograms (Gillam and Ryan, 1992), in pseudo-random line grids (Ryan and Gillam, 1994) and in random-line cyclopean matrices (Gillam, Cook and Blackburn, 1995) appear to generate the same influences upon seen slant.

Clearly, the perception of stereoscopic slant in an SKS is a very different prospect for the system than resolving the orientation of a textured surface. Disparity information and boundary

information available is very sparse. In testing the judged position and orientation of an illusory figure I was essentially testing the capability of the visual system to deal with very sparse image difference information. Furthermore, it has been established that the percept achieved when fusing the SKS at opposite signs of disparity changes from modal (crossed disparity) to amodal (uncrossed disparity).

### 4.6.2 Measures

*Seen slant: Perceived orientation of an illusory surface*

The first manipulation of the SKS analysed the perception of slant of an illusory figure. Thus the dependent variable was termed Seen Slant. A square, equiluminant with its surrounds, was projected onto four black circles in each half-image. This square had been geometrically transformed to simulate the pattern of disparity created by rotating a square in natural perspective. Three independent variables were chosen to examine the sensitivity of the system to this manipulation of disparity: Theoretical Rotation, Slant Axis and Standing Disparity:

a. Theoretical Rotation: four levels of this variable were presented (0, 20, 30, 40, $50^0$). It was expected that seen slant (measured by a comparison stimulus) would be equivalent to the theoretical rotation applied to the SKS.
b. Slant-Axis: two levels of slant-axis were presented (horizontal and vertical). It was anticipated that a slant-axis anisotropy would emerge defined by greater attenuation of seen slant about the vertical axis relative to the horizontal.
c. Standing Disparity: three levels of standing disparity were presented (-20, 0, 20 arcmins). Standing disparity was achieved by shifting the Kanizsa square horizontally relative to the pacmen.

In summary, this experiment explored seen slant of a stereoscopic percept generated by projecting a pair of squares of homogeneous luminance to their surrounds on to non-disparate pairs of four black circles (a typical Kanizsa square configuration).

The square in each half-image was manipulated to simulate the horizontal image differences cast by a square rotated in natural perspective that stood at various depths above, equidistant to or behind the P Plane. This entailed manipulation of two components of retinal disparity defined by disparate binocular subtense (theoretical rotation and standing disparity). Figure 4.6 describes the geometric manipulations applied to the half-images of the Kanizsa square, diagrammatically.

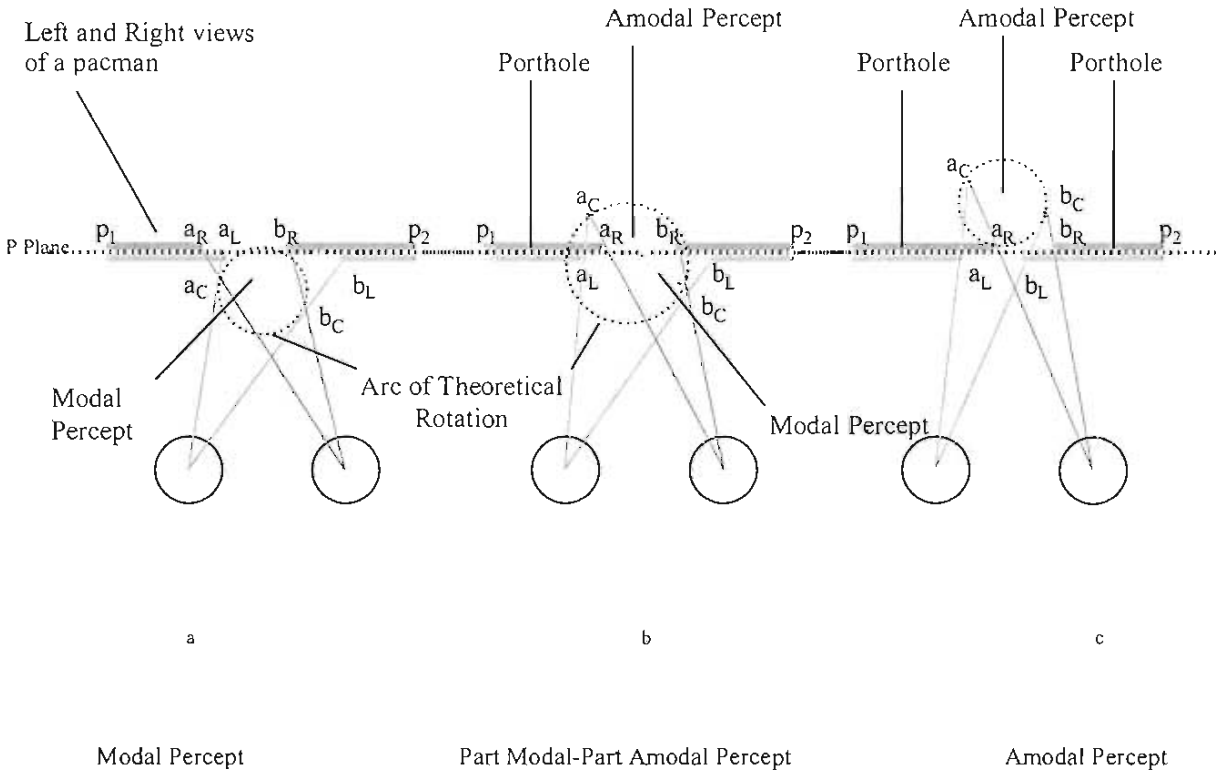



**Fig 4.6 Two components of disparity in a slanted SKS**

These projection diagrams show schematically the manipulations of binocular disparity achieved by projecting a white square to partly obscure four black circles to create disparate pacmen in each half-image. In (a) by appropriately manipulating disparate binocular subtense at the mouths of the pacmen (by standing disparity and relative magnification), it was predicted that fusion of the left half-image at $a_R$ and $a_L$ and $b_R$ and $b_L$ would yield a modal Kanizsa square whose right and left bounds stood at the cyclopean points $a_C$ and $b_C$ respectively. This was expected to result in a fully modal Kanizsa square rotated about the vertical axis. A similar manipulation at (b) but with no standing disparity predicts an illusory contour at $b_C$ with the surface slanted through the P Plane so that $a_C$ defined the depth of the amodal end of an illusory surface. In (c) the bounds of the square fused to yield the cyclopean points $a_C$ and $b_C$ would stand the figure behind the P Plane, predicting a stereoscopically rotated amodal surface.

*Seen Depth: Perceived distance of an SKS*

The second component of the depth percept manipulated was standing disparity. Thus, the experiment also measured the overall seen depth of the SKS at each level of standing disparity (-20, 0, 20 arcmins - see Fig 4.6). As explained above the standing disparity manipulation was achieved by a horizontal shift applied to the Kanizsa square in each of the half-images. This measure enabled an assessment of whether the overall stereoscopic depth of the illusory figure, theoretically a product of the relative horizontal disparity applied at the mouths of the pacmen, was discernible by observers. It was therefore anticipated that subjects' percepts would reflect the manipulation of this standing disparity (seen depth).

*Seen Lightness: Perceived lightness of portions of an SKS percept*

Since Kanizsa first demonstrated his remarkable illusory figures, lightness enhancement at illusory contours has been one of the most researched characteristics of the phenomenon. The cause of lightness enhancement is not yet resolved but is assumed to be at least partly related to brightness contrast (see for example Albert, 1993; Brigner and Gallagher 1974; Davi, Pinna and Sambin, 1992; Ehrenstein 1948; Frisby and Clatworthy 1975; Jory and Day 1979; Gregory 1977; Kanizsa 1975; Kennedy, 1975, 1976, 1981. 1987, 1988, 1991; Livingston and Hubel, Minguzzini, 1987; 1987; Purghe and Coren, 1992; Purge, 1991; Sambin, 1987).

A measure of lightness was used in this study to test the local character of the percepts. The hypothesis was that edges of the illusory figure standing forward of the P Plane (as in Fig 4.6a) would yield a rating of lightness suggesting lightness enhancement relative to the surrounding "white" homogeneous space - indicating a modal edge. Where the surface stood behind the P Plane (as in Fig 4.6) no lightness enhancement was expected.

## 4.5.3 Method

**Subjects**

Twelve available subjects were drawn from the pool of volunteers.

**Stereograms**

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a set of four black circles subtending $3^0$ in diameter (at 750mm viewing distance). The circles were positioned so that a square drawn through their centres

would subtend $7^0$ degrees. *Pacmen* were created by drawing a square, equiluminant with the background, so that it partly obscured the circles. The square symmetrically overlaid the circles (equal intrusion on all sides). The square intruded ½ the circles' radii creating Kanizsa square pacmen.

Mouths of the pacmen were geometrically transformed according to the monocular transformations of a square shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was then introduced by applying ½ Ogle's M to the overlaid square, symmetrically and in opposite signs (see Fig 4.7). Standing disparity was generated by shifting the overlaid square in equal and opposite directions in each half-image (see Fig 4.7).

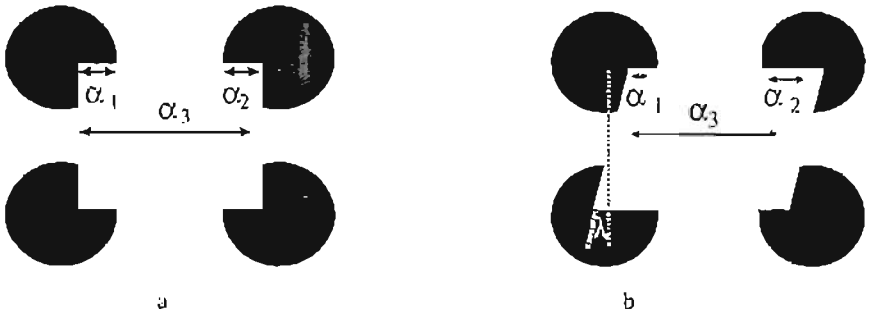



**Fig 4.7. Patterns of disparity in the SKS**
Fig 4.7 a shows one half-image supporting rotation about the vertical axis. Disparity appropriate to rotation about the vertical was applied by increasing the magnitude of the Kanizsa square ($\alpha_3$) in one eye relative to the other. This obviously changes the relative magnitudes of the mouths of the pacmen ($\alpha_1$) and ($\alpha_2$). Standing disparity was then manipulated by constraining ($\alpha_3$) while shifting the illusory figure in one eye relative to the other. This alters the relative magnitude of ($\alpha_1$) and ($\alpha_2$) in each eye. Fig (b) shows the horizontal differences applicable to rotation about the horizontal axis. To manipulate rotation, disparate shear ($\lambda$) was applied in eye while constraining ($\alpha_3$). Then standing disparity was introduced by again constraining ($\alpha_3$) while shifting the illusory figure in one eye relative to the other so adjusting the relative magnitudes of ($\alpha_1$) and ($\alpha_2$) in each eye.

Image pairs were presented on alternate frames at a rate of 200 frames per second. Through the shutter goggles, background luminance was 0.7 cd $m^{-2}$ with the black pacmen 0.09cd $m^{-2}$.

### Design and procedure

This study used a fully crossed, 5x3x2 repeated measures design to explore the effects of manipulation of the mouths of pairs of pacmen constant with theoretical rotation (0, 20, 30, 40, and $50^0$), slant-axis (Horizontal, Vertical) and standing disparity (+20, 0, and -20 minutes

of arc) of a projected surface. Three complete repetitions of this design were used, making 90 trials per subject in all. Details of the procedure were as described in section 4.4.

In the depth trials, subjects were asked to estimate the depth of the central axis of an illusory plane using the comparator as a depth probe. The depth of the comparator was manipulated as described in section 4.4. For the slant estimation trials subjects were asked to set the rotation of the comparison stimulus to match the apparent rotation of an illusory plane.

In this experiment only, subjects were also asked, on each trial, to judge the relative lightness of three specific regions of the SKS. Being a fully factorial design, lightness was judged in all the theoretical rotation, slant-axis and standing disparity manipulations. My interest here though was only to establish difference in lightness surface region (regions 1, 2, and 3). This yielded some 270 lightness scores. These were pooled across the three repetitions for each subject.

Subjects judged the lightness of specific regions of the configuration. These included the top edge of the plane rotated about the horizontal (region 1. below), then the central region of the plane (region 2. below) and finally the bottom of the plane (region 3. below). For the slanted surface, subjects judged the left edge (region 1) the centre (region 2), and the right edge (region 3) as shown in Fig 4.8.

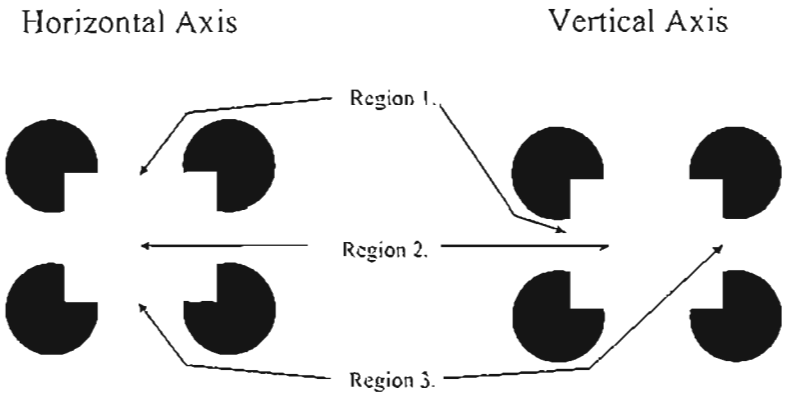



**Fig 4.8. Lightness judgement regions**
Subjects judged the relative brightness of the each of these regions in random order for each stimulus.

A pointer set 2 degrees from the test figure boundary (the outer perimeter of the disks) indicated the specific region to be judged. A sliding pointer (marked (a) in Fig 4.9) was presented at the centre mark of the scale for the start of each trial. Subjects were directed that

the extreme positive value of the scale indicated surface quality *much lighter* than the surrounding homogeneous regions and the negative end of the scale related to *much less light*[5].

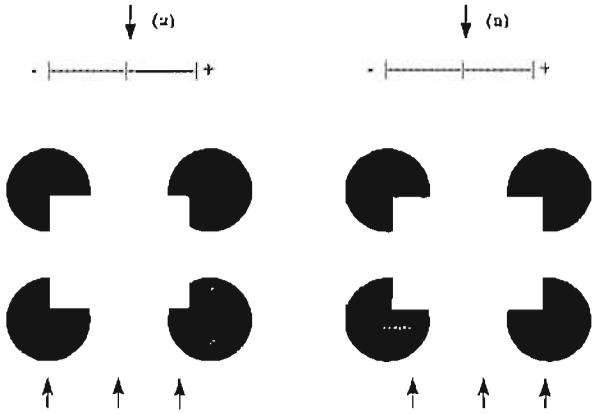



**Fig 4.9. Lightness judgement comparison stimulus**
Crossed fusion simulates the viewing situation faced by the subjects. At the top of the figure an arrow (a) was moved left or right to indicate a judgement of lightness at the regions indicated by one of the positions of the lower arrows. These were presented at a degree of disparity setting the figure well away from the depth of the illusory figure.

The lightness scale was operated by a joystick. Pressing the joystick trigger activated the scale. Movement of the joystick beyond a predefined threshold moved the pointer left if the stick was leaned left and right if the stick was leaned right. A lightness judgement was stored when the subject pressed either of two joystick buttons. This returned the pixel x-coordinate position of the pointer along the scale line when a joystick button was pressed. In turn, the x-coordinate position of the pointer was transformed into an eleven-point scale. The centre point of the scale represented 0 (no lightness difference), the far right represented 5 (lighter) and the far left, -5 (less light).

5 - Note that the subjects had the opportunity to practice lightness judgements on a select series of stimuli previous to the experimental trials. This was expected to give them an overall view of the relative lightness differences involved across the spectrum of the experimental stimuli.

### 4.6.3 Results

**Seen Slant**

To explore how binocular geometric factors impacted seen slant of the SKS percepts a three-way (5x2x3) repeated measures analysis of variance was used. The factors examined were theoretical rotation (0, 20, 30, 40, and $50^0$), slant-axis (horizontal, vertical) and standing disparity (+20, 0, and -20 arcmins). Obtained slant estimates were averaged across repetitions.

It was found that seen slant increased with theoretical rotation and there was a tendency for the larger slants (eg, $40^0$ and $50^0$) to be underestimated, as shown in Fig 4.10. The effect of theoretical rotation was significant $F_{(4,11)} = 147.298$, $p < 0.0001$.

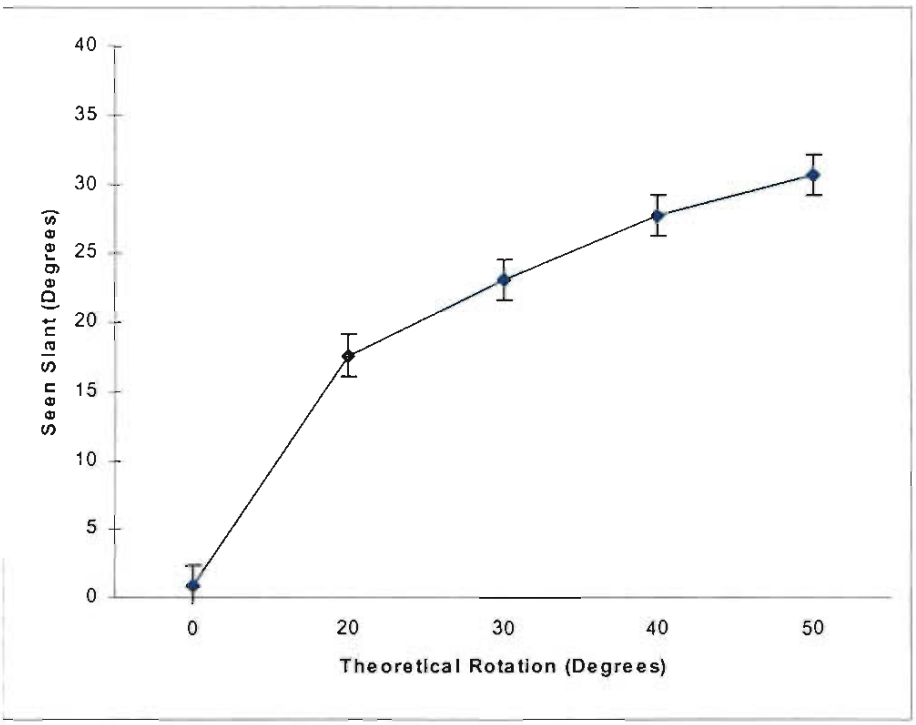



**Fig 4.10. Seen slant of a SKS across five levels of theoretical rotation**
This figure depicts the main effect of theoretical rotation (defined by relative magnification or shear) upon seen slant in SKS percepts. Standard error bars have been included.

The effect of slant-axis was also found to be significant: $F_{(1,11)} = 13.558$, $p < 0.05$, attesting to the slant-axis anisotropy illustrated in Fig 4.11. Seen slant about the horizontal axis was about 7 degrees greater than the vertical axis.

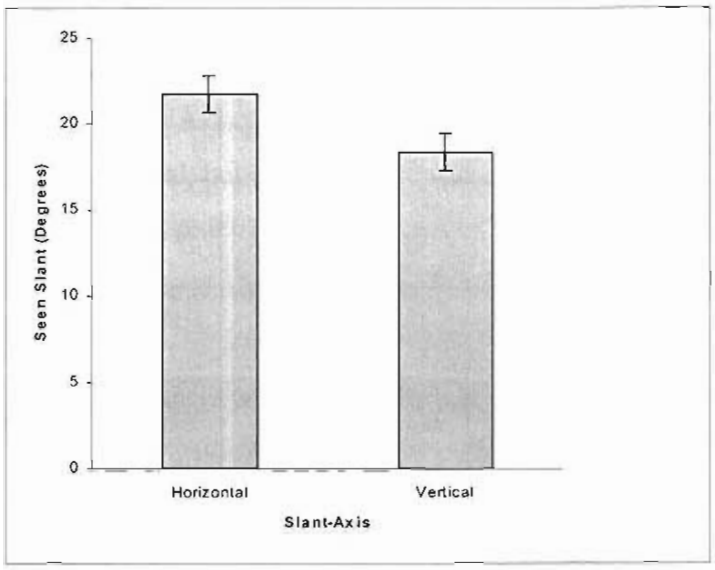



**Fig 4.11 Seen slant in a SKS rotated about the horizontal or vertical axis**
This figure graphs the main effect of Slant-Axis upon seen slant with standard error bars included.

Standing disparity also significantly affected seen slant: $F_{(2,11)} = 41.5$, $p < 0.001$, with surfaces standing forward of the P Plane (20 arcmins) judged about $5^0$ greater than rotations through the P plane (0 arcmins) and SKS standing behind the P Plane (-20 arcmins). Figure 4.12 demonstrates this effect.

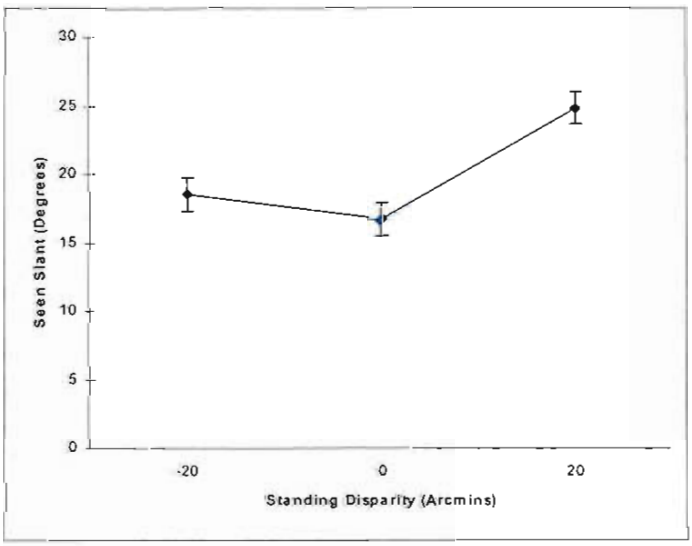



**Fig 4.12. Seen slant in a SKS across three levels of standing disparity**
This figure depicts the main effect of standing disparity upon seen slant in the SKS with standard error bars. At -20 arcmins the SKS stands fully behind the P Plane. The SKS has an amodal character and has the appearance of being viewed through four portholes. At 0 arcmins, the SKS rotates through the P Plane. It looks part modal and part amodal. At 20 arcmins the SKS stands fully forward of the P Plane. It has a modal character.

Two first order interactions between the main effects provide a further commentary on the effect of slant-axis and standing disparity on seen slant. First, Fig 4.13 depicts the interaction of Theoretical Rotation X Slant-Axis. Seen slant about the vertical was perhaps more attenuated than about the horizontal, but the difference was not clear cut. The interaction was significant however: $F_{(4,11)} = 4.67$, $p<0.01$.

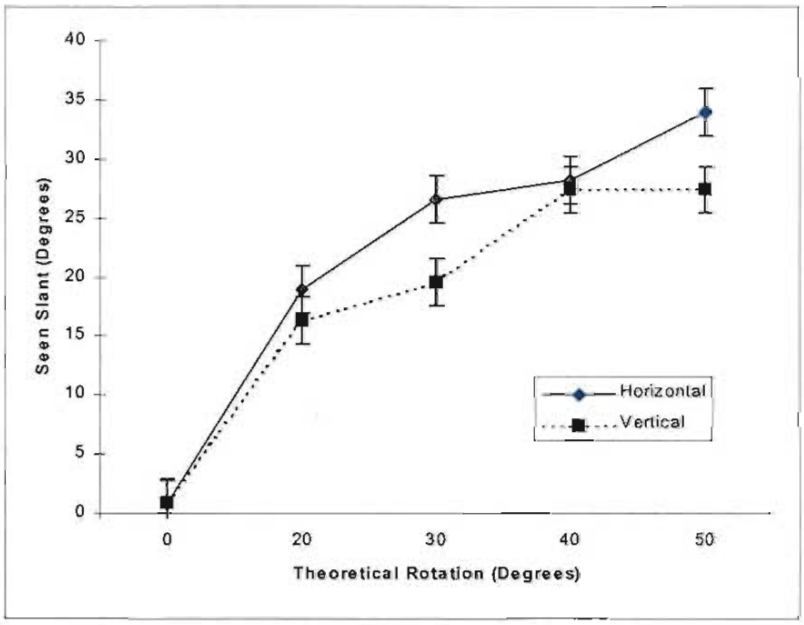



**Fig 4.13. Seen slant in a SKS: Interaction between theoretical rotation and axis**
This figure shows mean seen slant for the first interaction between slant-axis and theoretical rotation. Standard error bars have been included.

Figure 4.14 depicts the Theoretical Rotation X Standing Disparity interaction. The interaction was also significant $F_{(8,11)} = 4.53$, $p< 0.001$. Clearly, seen slant was substantially more consistent with theoretical rotation when the Kanizsa square stood forward of the P Plane.

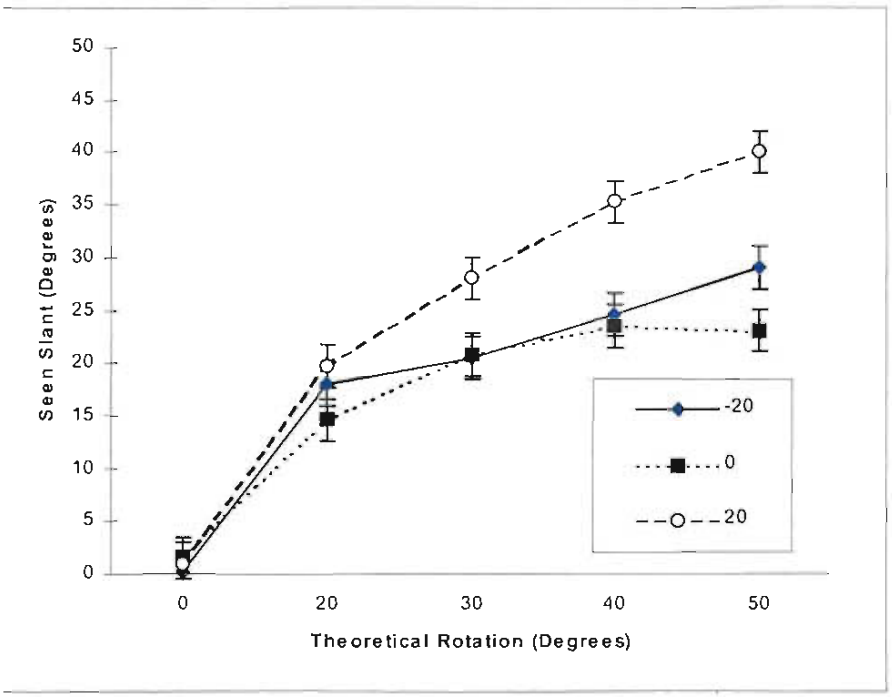



**Fig 4.14. Seen slant in a SKS: First order interaction between theoretical rotation X standing disparity**
This figure graphs mean seen slant for the interaction effect between theoretical rotation of an SKS and standing disparity. Standard error bars have been included.

These findings support the hypothesis that seen slant of a stereoscopic illusory surface would reflect the pattern of disparate subtense or shear applied to each half image. Typical of findings in seen slant for surfaces defined by disparate textures, larger rotations tended to be attenuated somewhat. Also consistent with previous research, anisotropy between the axes of rotation was evident. In addition, it was found that seen slant was attenuated substantially less with visible or modal completion of a SKS, that is, when the illusory surface stood in front of the P Plane (20 arcmins).

*Seen Depth*

This part of the experiment examined the ability of subjects to judge the depth of a SKS relative to the P Plane (the computer screen on which the disparate pacmen were drawn). A one-way repeated measures analysis of variance was conducted across three levels of standing disparity (+20, 0, and -20 arcmins). Results suggest that seen depth (defined by setting the comparator as a depth probe) varied predictably with standing disparity (see Fig 4.15). The

effect was significant, $F_{(2,11)} = 40.649$, $p < 0.001$.

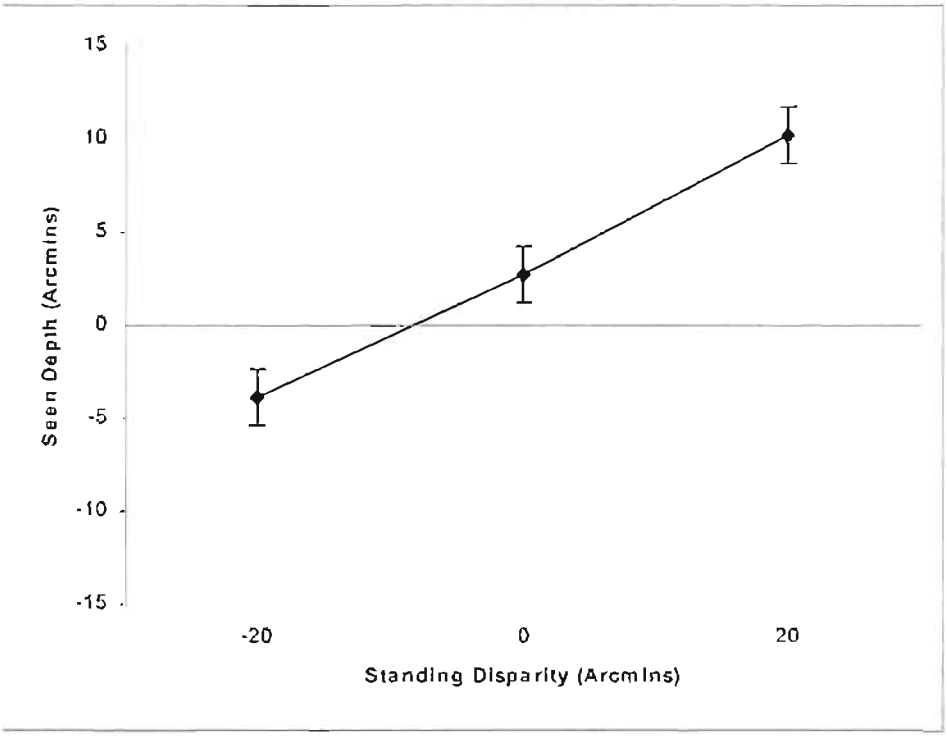



**Fig 4.15. Seen depth of a SKS**
This figure shows mean depth estimates of the central axis of SKS percepts at each level of standing disparity. Standard error bars have been included.

While observed depth judgements do not accurately match the theoretical disparity applied, recall that the task required only a global estimate. The task performed by subjects was to set the comparison stimulus to the depth that they saw the axis of a rotated illusory plane relative to the P Plane. The surfaces were typically at various degrees of slant and the surface being judged had no physical boundaries, perhaps making this task difficult. Nevertheless, the depth judgements observed are evidence that the system was able to access the standing disparity regardless of the actual presence of visible bounds (at +20 arcmins illusory bounding contours were present while at -20 arcmins they were not).

*Seen Lightness*

The objective of this measure was to test the perception of lightness of the SKS at different depths, that is, an SKS standing forward of the P Plane (+20 arcmins), standing at the P Plane (0 arcmins) and standing behind the P Plane (-20 arcmins). A one way repeated measures

analysis of variance was undertaken to analyse lightness judgements. The effect was significant: $F_{(2,11)} = 11.2$, $p < 0.001$ (see Fig 4.16). A slightly higher rating of seen lightness at + 20 arcmins compared to the mean judgements at 0 and - 20 arcmins supports the prediction that lightness enhancement would be associated with so called modal illusory SKS standing forward of the P Plane and to a lesser extent with the amodal SKS standing behind the P Plane.

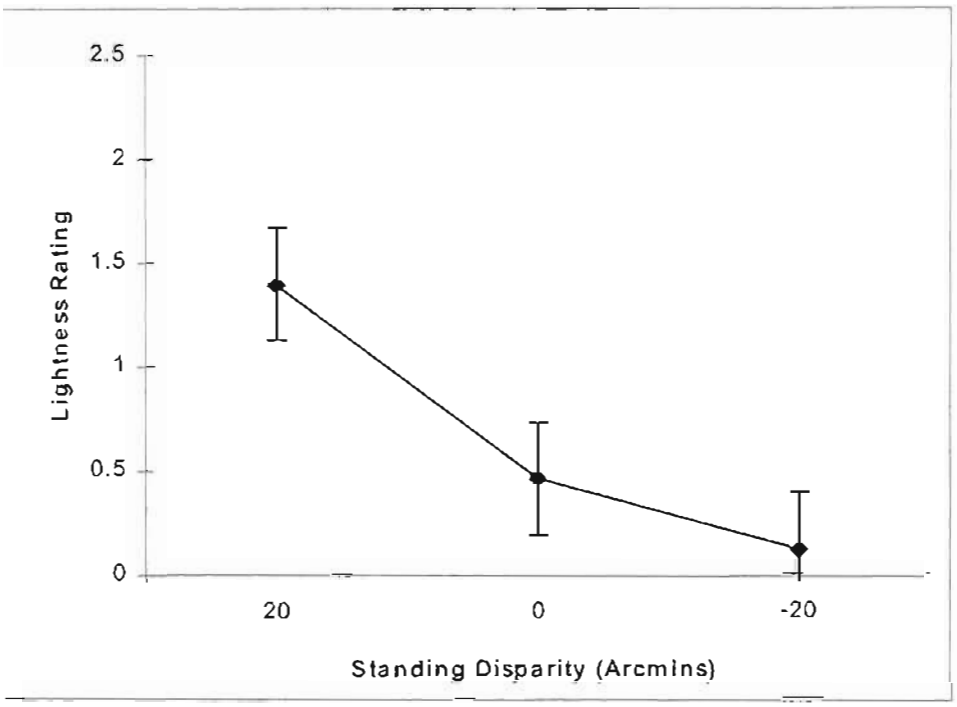



**Fig 4.16. Seen lightness in a SKS**
This figure plots the mean lightness rating (along an eleven point Likert scale with 0 as the centre point) for each level of standing disparity that were applied to the SKS. Standing error bars have been included.

### 4.6.4 Discussion

These results provide initial descriptive evidence that stereopsis is able to access and integrate substantial stereoscopic information in SKS half images. As seen in Fig 4.10, seen slant approximated theoretical rotation and in Fig 4.16 depth judgements varied according to the manipulation of standing disparity. However, the slant response appears to have been substantially attenuated for the amodal and part-modal SKS. In other words, in the case of an amodal SKS, subjects were aware of a slanted square seen through portholes (see Fig 4.14 at -

20 arcmins standing disparity). In the case of the part-modal SKS, subjects were aware of a slanted square where half the SKS was seen through portholes (see Fig 4.14 at 0 arcmins standing disparity). All of this suggests that basic binocular image processing mechanisms at least partly underpin the SKS percepts. Nevertheless, there does appear to be an impact of visible spreading or completion mechanism upon the integration of the image difference information. Perhaps modal connections linking disparity signals are stronger than amodal connections.

For free fusers, the perceptual outcome that the three metrics describe may be seen when the stereogram below is fused. In Fig 4.17a the white square drawn to overlay black disks to create the Kanizsa pacmen has been enlarged in half-image (L), in the horizontal plane, relative to half image (R). When these two images are fused (crossed or uncrossed free fusion) observers report an Kanizsa square that is rotated in the third dimension, that is, in the z plane. The Kanizsa square looks modal (exhibits a visible illusory contour) at the near-end and amodal at the other. This effect might be described as *apparent change in modality of completion* between the near and distant ends of the Kanizsa square. Somehow, lightness enhancement arises at the modal (visible edge between the pacmen) of the Kanizsa square.

There is also a difference in the perceptual quality of the pacmen. The two pacmen that appear partly occluded by the Kanizsa square look opaque while the other pacmen look like portholes. These effects are described pictorially in Fig 4.17b.

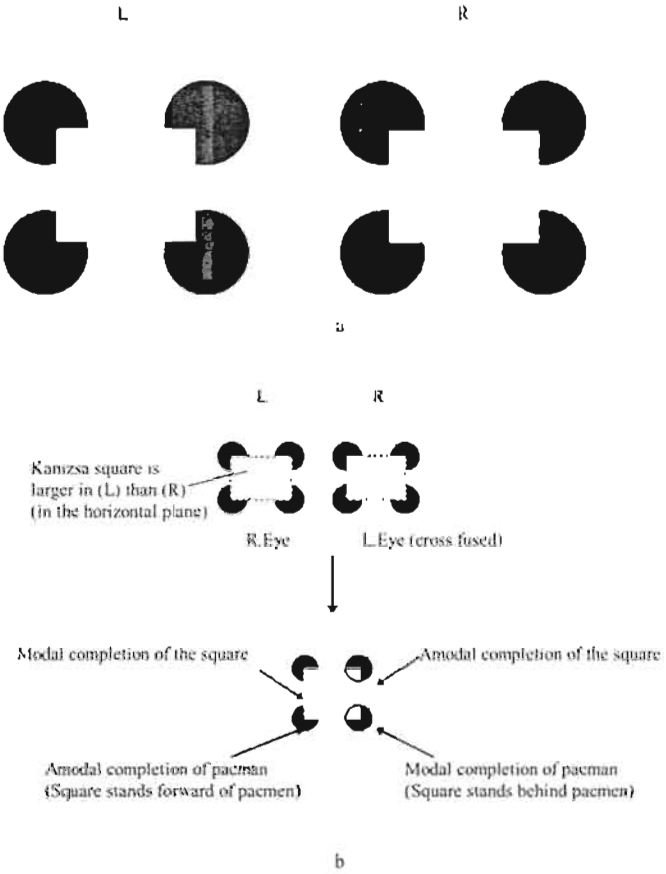



**Fig 4.17. A slanted SKS**
In fig 4.17a, crossed or uncrossed fusion of the half-images (L) and (R) yields the percept that an Kanizsa square appears rotated or slanted in three-dimensional space. The phenomenology associated with this stereoscopic rotation is shown in 4.17b. One end of the SKS looks modal, that is, exhibits an illusory contour. The other end of the square looks amodal. The pacmen standing at back of the modal illusory edge appear opaque and at the other end of the square they look like portholes.

In summary the results of this experiments reveal a close relationship between basic stereoscopic processes and the phenomenology of the SKS. Depth judgements and seen slant approximated the theoretical disparities applied and subjects' lightness judgements suggested that illusory edges were only evident at edges of the Kanizsa square standing forward of the plane of projection. It appears that there was no induction of illusory contours bounding the Kanizsa square unless stereopsis generated interpolation of surface layers.

Separation of surface layers at the P Plane may have somehow modulated the modality of completion or in Grossbergs terms, filling-in by binocular FIDOs. Parts of the illusory surface standing forward of the P Plane appear to filled-in. Parts of the illusory surface standing behind the P Plane appear to be occluded by an opaque surface at the depth of the P

Plane. This was most dramatically evident when the illusory figure slanted through the P Plane.

Figure 4.18 demonstrates a Kanizsa square rotated about the horizontal axis. The finding that seen slant was predicted by theoretical rotation of an illusory square, suggests that *resolution* of binocular geometry has a role in the perceptual organisation of the images presented. Whether modal or amodal, or whether part modal/amodal the slant response was still evident. However, a slant axis anisotropy suggests that the integration of disparity measures must be subtle.

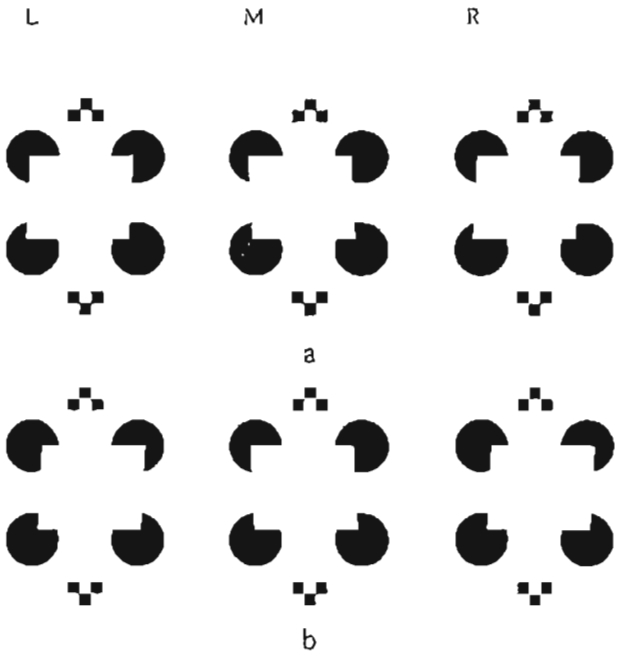



**Fig 4.18. A Kanizsa square rotated about the horizontal axes**
Figure 4.18a shows a stereoscopic illusory figure rotated through the plane of projection (the page). Crossed fusion yields a surface rotated about the horizontal axis. In both percepts (L-M and M-R), the near edge of the illusory surface looks modal and the distant edge appears amodal. At the same time, the occluded pacmen look opaque and the occluding pacmen look like portholes. The surface separation here seems to involve filling-in of a surface at the P Plane that peels forward to the modal illusory edge. In (c), a set of stereograms are presented to show the manipulation of standing disparity. Crossed fusion of L-M sees an illusory surface rotating about the horizontal axis but partly occluded by an opaque surface between the pacmen. Fusion of M-R sees an opaque modal illusory plane standing forward of the pacmen. Separation of surface planes appear to be generated at the illusory contours.

## 4.7 On the relationship between projection geometry and the SKS percepts: A summary description

When the pattern of perspective projection at the mouths of the SKS pacmen predicted that an edge of the square (the overlaid white square used to generate disparate pacmen) would stand forward of the pacmen (local crossed disparity), a modal edge formed. A visible illusory

contour was seen between adjacent pacmen. The illusory contour was not an isolated perceptual event but marked the separation of a near slanted surface separated from a distant frontal surface (homogeneous surrounds). The distant surface occurred at the depth of the pacmen.

Likewise, when an edge of the SKS was predicted to stand behind the P Plane it was not visible. Subjects still saw the partly visible shape slant. Traditionally such an edge was called amodal. Only the actual mouths of the pacmen were assigned to a distant depth plane. The untextured region between the pacmen was assigned to the depth of the bounding arc of the portholes. An illusory contour closed the bounding arc of the porthole. That contour marked a depth step between the fronto-parallel surface at the P Plane and the mouths of the pacmen. It was concluded that the separation of surface layers must depend on the sign of local disparities.

Figure 4.19 is used to point out aspects of how the 3-D illusory SKS percepts might relate to binocular image processing. The solid black lines represent the disparate pacmen presented on the monitor screen. Figure 4.19a depicts the percepts in question.

The figure represents visual projections at a horizontal slice through pacmen in the SKS half-images. In turn, those visual projections define the shape and size of images captured at the retinae. Clearly, the results of Experiment 1 suggest that the system responds to the SKS almost as if a textured square, was rotated through the P Plane. In the SKS there is no surface texture. There are only four luminance steps (contours) that can possibly be fused by stereopsis. Two contours occur at the outer boundaries of the pacmen. These are non-disparate. They are evidently assigned equidistant with the P Plane (the monitor screen).

Experiment 1 has established that fusion of uncrossed disparity to produce the cyclopean point $a_c$ is associated with the perception of a porthole. Fusion of the crossed disparity to create the cyclopean point $b_c$ is associated with the perception of an illusory edge and the separation of a white surface layer from the depth plane of the pacmen. The white layer was demonstrated to slant from $b_c$ toward the other end of the Kanizsa square.

Such a percept appears to involve fitting regions of the disparate images to appropriate depth and orientation by separating surface layers at the pacmen. The implication is that the three-dimensional percept induced by fusion of the SKS must closely reflect constraints upon stereopsis existing within the distributed pattern of 2-D differences in the retinal half-images, that is, disparate subtense between contours. The next chapters will explore binocular image

processing mechanisms other than point-disparity computation that might account for the intriguing 3-D illusory SKS percepts.

In conclusion, an exploratory study of the phenomenology associated with the stereoscopic response to disparate Kanizsa squares showed that the illusory squares' perceived depth and orientation was much like a surface defined by point-disparities. It was found that at P Plane some mechanism appeared to modulate the mode of the Kanizsa square completion. This modulatory effect appeared to be related to the manner in which the system separated the various untextured components in the stereograms to particular depth planes. The percepts quantified appear to point to an image processing mechanism that involves resolution of disparate subtense. Later chapters examine these issues in detail. At this stage it appears that there is more information for stereopsis in the Kanizsa square half-images than previously acknowledged by the Surface Heuristic approach to 3-D illusory percepts in particular.

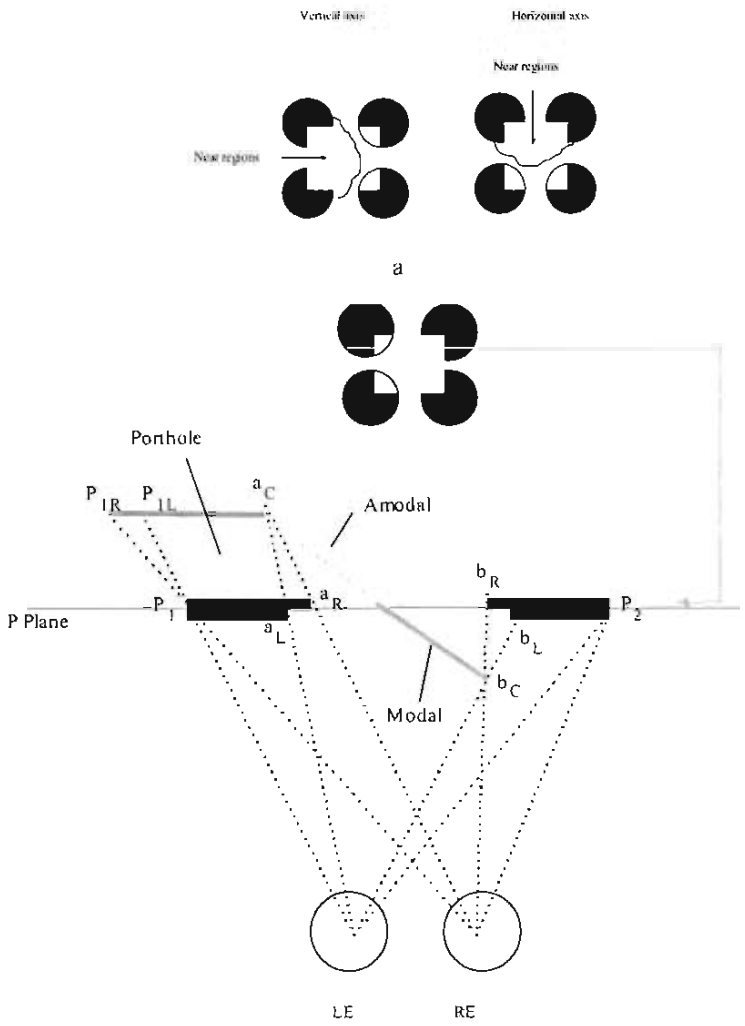



**Fig 4. 19. Visual projections and binocular fusion of corresponding contours in a SKS rotating through the P Plane**

There are just four luminance steps (contours) that can be fused in this horizontal slice through the disparate pacmen of an SKS. These are the steps at point P1 and P2 at the outer edges of the pacmen. There is no disparity between these points so I can assume that they are perceived to stand at the depth of the plane of projection. An uncrossed disparity exists between aL and aR . This disparity predicts a cyclopean point at the depth of aC the edge of the amodal portion of a SKS – but visible to both eyes. Similarly, there is a typical crossed disparity between point b L and b R whose fusion predicts the cyclopean point bC. Note however that the luminance step that predicts this crossed disparity also is the luminance step at which surface layers separate—the inner boundary of the pacman separates from the near white surface layer. This means that the boundaries of the near and distant surfaces are predicted by fusion of the same luminance contour. Note also that the magnitude of bR_bL is precisely the magnitude of the monocular zone that is visible to the right eye and occluded from the left. This means that the perception of a near surface edge at bC appears to resolve the depth of both the near and distant surface boundaries. In turn this means that the disparate binocular subtense between P2_bR and P2_bL is also resolved. If this is so, then resolution of the disparate subtense between P1_aR and P1_aL might be achieved by a similar configuration of projections that sees the separation of surface layers at P2. This suggests that a monocular zone of magnitude P1R_P1 might contribute to perception of a porthole. It might be feasible then that the separation of surface layers at these two different luminance steps would support the perceptual asymmetry evident in the SKS rotated through the P Plane. This would achieve allelotropia across each pacman, that is, the deformation required to fit each eye's image to the other to generate the cyclopean percept.

# 5. Toward a functional model of the perception of a SKS

***Summary:*** *This chapter discusses the detailed 2-D layout of image differences in a single SKS pacman. A functional model is developed that might account for the phenomenological properties of the slanted Kanizsa square described in Experiment 1.*

## 5.1 Binocular mechanisms and perception of a SKS

Subjects' slant and depth estimates in Experiment 1 equate well with the seen depth and orientation recovered from textured surfaces. On the face of it, it seems that from few local disparities the visual system somehow assigned an illusory Kanizsa square to a predictable depth and orientation - despite only very sparse point-disparities. The question is: what kinds of visual processes lead to such remarkable phenomenology?

The Surface Heuristic and Form Computation approaches emphasise ambiguity of such untextured stereograms. They have assumed that the 3-D illusory percepts are underdetermined by retinal disparity, in a narrow sense, and so have proposed mechanisms by which the system might compensate for sparse disparity.

Another way of interpreting the Experiment 1 findings is that the system indeed accesses image difference information other than local point-disparities and integrates this into the 3-D percept. This should not be surprising because the utility of relative disparity (disparate binocular subtense) for stereopsis is well established (see Chapter 1 for an extensive review). Theorists previously concerned with the 3-D illusory percepts have perhaps overlooked this issue. In this chapter then, I attempt to develop a possible alternative explanation of the SKS percepts.

## 5.2 Stereoscopic response to pacman half-images

### 5.2.1 Perspective projection and the SKS half-images

Recall that in Experiment 1 the shape of the SKS pacmen mouths simulated the silhouette of a surface at specifiable depth and orientation (partly obscuring four black circles). The "square" projected at each half-image was shaped to model a disparate view of the pacman at each eye (see fig 5.1).

This model geometric surface was then rotated to varying degrees about each axis. Hence, the shapes of the mouths of the pacmen simulated perspective projection of the corners of a surface onto each retinae. Perspective projection creates a rather complex pattern of

image differences that results at each pacman and between the pacmen. The 2-D layout of half-image differences is the product of binocular parallax (in the natural setting).

Indeed Wheatsone's original experimentation in stereopsis literally used sketches of the each eye's slightly different view of simple 3-D objects as half-images. My manipulations have attempted to do the same thing. The difference being, of course, that local disparity at surface contours and texture points is almost entirely missing in the SKS. It is clear that the perceptual organisation of surface layers in the SKS reflects the stereoscopic response to disparate shapes of the pacman half-images.

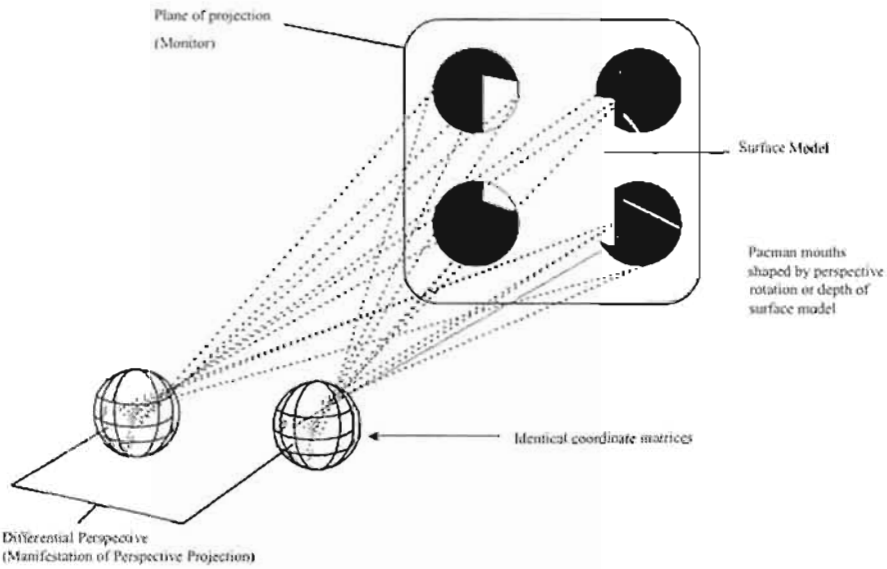



**Fig. 5.1. Perspective projection in construction of SKS half-images**
The image differences in the shape of the pacmen simulated perspective differential cast by a model surface (a square) silhouetted against the circles (pacmen).

## 5.2.2 Perceptual asymmetry at a single stereoscopic pacman

Consider Fig 5.2. In physical terms, each pacman is bounded by a single continuous steep luminance contour that changes direction sharply at several locations. Points at which the contour changes direction most sharply are labelled A B and C. The remainder of the contour describes a uniform arc through D.

The image pairs in Fig 5.2 demonstrate two *signs* of disparity[4]. With crossed fusion of L-M, a white surface stands forward of the pacman. Fusion of M-R generates the porthole

[4] This figure was created by drawing a small square, equiluminant with the page to partly obscure a single black

effect[7]. Clearly, fusion of disparate pacmen yields the perception of two surface layers even when removed from the SKS configuration. This suggests a basic stereoscopic mechanism separate from the actual generation of illusory contours is partly responsible for the 3-D percept.

Note that a perceptual asymmetry still occurs. Fusion of L-M, generates a depth step at the mouth of the stereoscopic pacman through ABC. In fusing M-R, a depth step is also generated but not at the mouth of the pacman. Instead, a depth step occurs along the entire arc of the pacman (along ADC). An illusory contour forms that continues the arc of the pacman through CA. A white region at the mouth of the pacman (ABCA), stands behind the arc ADC. The pacman now looks like a porthole.

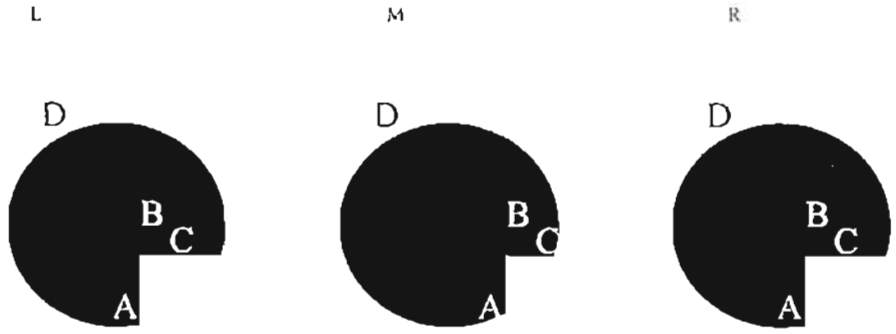



**Fig 5.2. Fusion of disparate pacmen**
Crossed fusion of L-M generates two separate surface layers. A near edge at ABC appears to bound a layer assigned to a near depth plane that partly occludes a black circle - assigned to a distant surface. Fusion of M-R yields a porthole. The boundary of the porthole looks to be completed by an illusory contour between A and C (the percepts will be reversed for divergent free fusers). Through the porthole, a small white sector between ABC is somehow assigned to a distant depth plane. This region looks like the corner of, say, a square that is partly occluded by a surface with a porthole cut through it. Separation of surface layers arises along the arc ADC (the porthole boundary) - not at ABC.

If we can understand how separation of surface layers is achieved then perhaps mechanisms underpinning the 3-D illusory SKS percepts may be illuminated. The first step is to define exactly what image differences exist in these simple pacman half-images.

### 5.2.3 The 2-D layout of image differences at a single SKS pacman

Figure 5.3 schematically describes binocular fusion of a pacman by drawing the right image over the left. This enables direct comparison of the projected images for both signs of

---

circle. This arrangement was repeated three times to create half images L,M and R. The square I M was the shifted about two mm to create retinal disparities.

disparity. In effect, this drawing equates to motor alignment (by vergence eye movements) of the paired retinal coordinate matrices. When vergence is locked the origin of each matrix is the centre of each line of site (grey lines represent aligned binocular coordinate lines).

Note that in Fig 5.3, regions shaded grey, are actually of opposite contrast in each eye. For crossed disparity, the region is white in the right eye, and black in the left. For uncrossed disparity, the region is black in the right eye and white in the left.

For crossed disparity, horizontal point-disparity exists at the mouth of the pacman, since $B_R\_A_R$, in the right eye occupies a disparate position to $B_L\_A_L$ in the left. There are also disparities in the magnitude of the regions $B_R\_C$ and $B_L\_C$; and, between $B_R\_A_R$ and $B_L\_A_L$. These are differences in the visual angles subtended at each retina (disparate subtense). Uncrossed disparity reverses the eye of origin of these differences.

There is also a pattern of binocular subtense across the pacman that is reciprocal to point-disparity (for example between $D\_B_R$ and $D\_B_L$). The visual angle subtended by these regions in each eye is disparate. Moreover, there is a difference in the actual length of the arc of the pacman boundary between $A_R\_D\_C$ and $A_L\_D\_C$.

---

, Uncrossed fusion will reverse these effects.

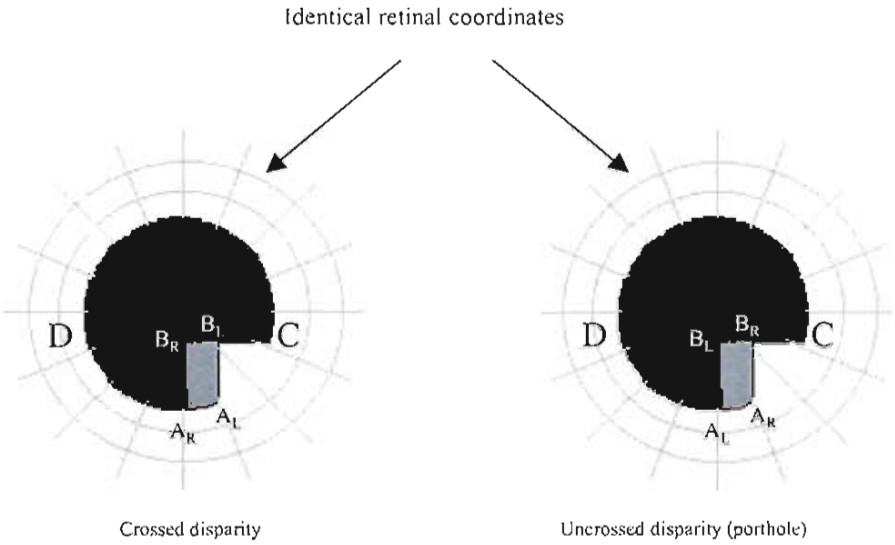



**Fig 5.3. Binocular disparity of a stereoscopic pacman**
Assuming that half-images of a pacman are fused by aligning the non-disparate arc of the half-images for both crossed and uncrossed disparity, it possible to identify the pattern of retinal differences that exist when the mouths of the pacmen are manipulated. A retinal disparity exists between the pacman boundary at $B_R$-$A_R$ and $B_L$-$A_L$. The disparate region of the fused pacman is shaded grey. For crossed disparity, the disparate region is black in the left eye and white in the right eye. For uncrossed disparity this region is white in the left eye and black in the right. Note that there is a vertical disparity present between $A_R$-$A_L$ at each sign of disparity. Further, note that disparate binocular subtense exists between $B_R$-$B_C$ and $B_L$-$B_C$, again at both signs of disparity. This disparate subtense is reciprocated exactly, between $B_R$-D and $B_L$-D (between the arc and the mouth of the pacman). Finally note that a vertical disparate subtense occurs between $B_R$-$A_R$ and $B_L$-$A_L$ at both signs of disparity.

Clearly, substantive image differences in the shape of each half-image are present even in a single stereoscopic pacman.

### 5.2.4 Resolving 2-D image differences: Crossed disparity

The stereoscopic response to the 2-D layout of images differences at a disparate pacman can be described using a simple projection diagram (see Fig 5.4). The figure depicts just the horizontal projection geometry subtended at the two retinae by each pacman half image (along an epi-polar line D_C). This is the percept where it looks as though a white surface layer at the mouth of the pacman is perceived to float about the pacman. The dotted line D_C represents the P Plane (eg. the page, or in Exp 1, the computer monitor).

At either end of the pacman, luminance steps at D and C project to identical points in each retinal matrix. In fact, the same can be said of all points along the arc of the pacman. However, along D_C, disparate luminance steps occur at $B_R$ and $B_L$. These luminance steps occur at the mouth of the pacman. They project to disparate points in each retinal coordinate

matrix. The system fuses these signals and achieves the sensation (cyclopean) that a point $B_C$ exists at a depth at which visual projections through $B_R$ and $B_L$ intersect.

Clearly, $B_C$ is not an isolated visual point, but is seen as a near edge at which two surface layers are somehow separated. The 3-D impression is that a near surface layer partly occludes a pacman. The whole mouth of the pacman is assigned to stand forward.

Fusion of the luminance steps at $B_R$ and $B_L$ , the contours that yield the near "white" edge, also leaves a black "edge" standing on the P Plane with ($B_R$ _ $B_L$) a monocular zone marked visible only to the left eye.

In theory, either a depth step at $B_C$ or perception of a slanted surface might be signalled by the disparate magnitudes of D_ $B_R$ and D_ $B_L$. However, observers do not report that the pacmen look twisted or bent - as this disparate subtense might suggest. Assigning the white surface segment $B_C$_C' to a near depth plane is, in manner of speaking, a particular *resolution* of the 2-D layout of image differences in achieving the singular or cyclopean 3-D percept.

The term *resolution* is meant to pose the possibility that the system is faced with a 2-D data set in which the size and shape of a black luminance term (the pacman) is manifestly different in each retina. Most of the border of the black luminance term falls on identical retinal points along a non-disparate arc.

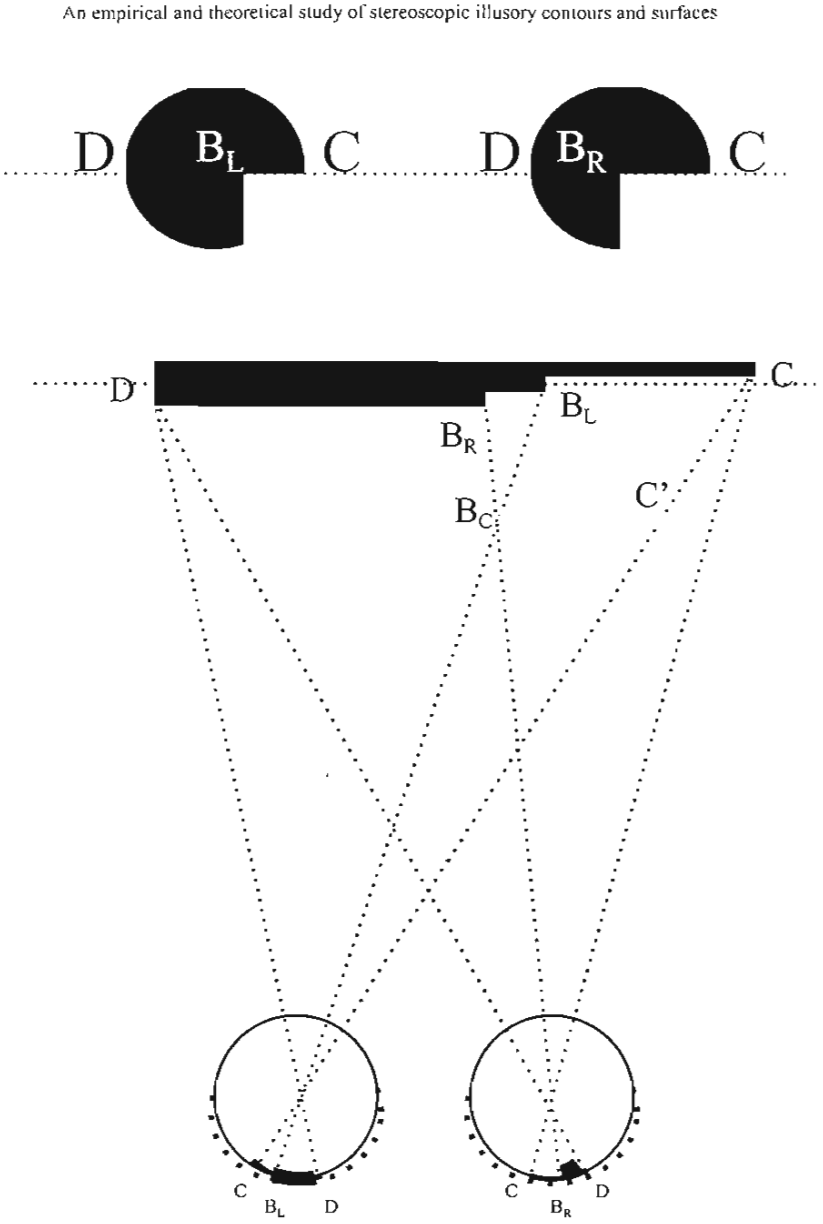



**Fig 5.4 Crossed fusion of disparate pacman half-images**
This figure represents the projection array associated with crossed fusion of a single pacman, its relation to the shape of the retinal images and perceived depth. D_C represents the magnitude of the pacman. The mouth of the pacman cuts into $B_L$ in the left eye and $B_R$ in the right. Fusion of $B_R$ and $B_L$ predicts the cyclopean point $B_C$. Note that point $B_C$ stands at the intersection of crossed visual projections that correspond both with $B_R$ and with $B_L$. Assignment of the edge $B_C$ yields a monocular zone visible to the left eye. The disparate magnitude of $B_R$_D - $B_L$_D are resolved in the separation of surface layers to $B_C$. Interestingly this has implication for the white surface layer assigned to the depth of $B_C$. The magnitude of disparate subtense ($B_R$_C - $B_L$_C) would either project beyond the edge of the pacman in the left eye (as in $B_C$_C'), or be seen as slanted through $B_C$_C.

In the absence of a disparity discontinuity (a step in the point disparity field available in textured surfaces) a crossed configuration of visual projections ($B_R$-$B_C$-$B_L$) yields

separation of surface layers. Is it possible that the system actually splits the percept into two layers at a single contour, where $B_C$ aligns $B_R$ in the right eye and $B_L$ in the left?

### 5.2.5 Surface separation in functional terms: Crossed disparity

The analysis above suggests that perceived surface separation in untextured stereograms involves processing a pattern of disparate subtense in addition to conventional point disparity. Figure 5.5 depicts likely component processes in fusing a single disparate pacman.

First, images of the pacman impinge on the sensor matrix at the retina. This is traditionally termed *Image Registration*. Next, the system coordinates *Vergence Lock* using vergence eye movements. Achieving vergence lock aligns two identical retinal coordinate matrices (see Fig 5.5a).

Conceptually this enables precise *Image Comparison* (Fig 5.5b). At the arc of the pacman, say points D and C, image comparison will detect no disparity. All possible points along that part of the contour register at identical coordinate points in each retina except at the mouth of the pacman. Hence, image comparison (see 5.5b) will return positional *Parity* between the retinal images as well as positional *Disparity* for contours at the mouth of each pacman image.

In the pacman half-images, just two contours will register (actually, this is just one curved contour). Between those contours there is a homogenous luminance data term. Retinal space between contours is a different size (ie. exhibits disparate subtense). Note that the difference in subtense is precisely reciprocal to the magnitude of conventional point-disparity. Indeed, it is a geometric fact of binocular parallax that the two retinal measures are immutable and reciprocal. For example, increasing disparity at the mouth of the pacman ($B_L$-$B_R$) necessarily decreases subtense across the pacman (D_ $B_L$- D_$B_R$) and the reverse is true.

Not surprisingly then, it appears that a binocular *Surface Separation* mechanism must be sensitive to disparate subtense between contours as well as conventional point disparity since, in untextured space, they are essentially the same thing. Figure 5.5c proposes how the system might assign depth according to positional disparity and disparate subtense (and in order to resolve projected image differences).

First, note that subtense across the whole pacman (D_C) is constrained to be interpreted at fronto-parallel because the arc of the pacman is non-disparate (equal curvature). Given that the disparate subtense at D_$B_L$- D_$B_R$ is therefore constrained, the system assigns

the disparate region of the field *region* $B_L\_C$ - $B_R\_C$ (the mouth of the pacman) to a near depth plane.

The system seems to split the percept into two layers where the visual projections cross. Something unusual must happen at this contour – if binocular subtense is to be resolved. The white side of the luminance step must be assigned to the edge of the near surface layer $B_C$ (partly occluding the black side of the step, the pacman itself). The position of $B_C$ is predicted by the intersection of crossed visual projections through $B_L$ and $B_R$ assigned to the depth plane of the pacman. But the black side of the same contour must be assigned to the distant plane. This means that a single contour is shared between the near and the distant surface layers. For convenience, I will term this perceptual outcome a *split-projection* configuration (since two layers share the same visual projections).

This response raises the possibility that the 3-D SKS percept is due to *Surface Spreading* from the region separated by binocular image processing at individual pacmen.

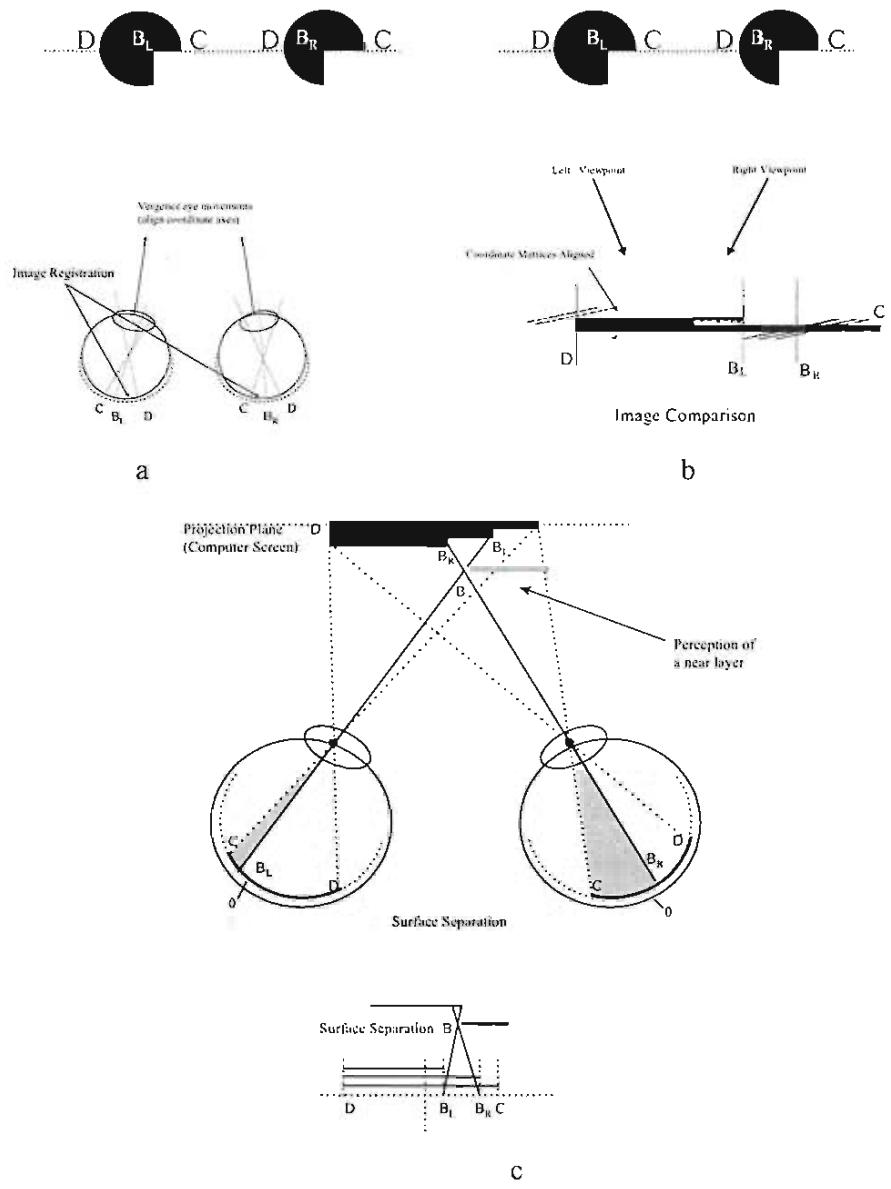



**Fig 5.5. Binocular image processing at an SKS pacman: Crossed disparity**

This figure depicts conceptually components of a binocular image processing scheme that might achieve surface separation at a pacman. The figure examines a horizontal slice through the pacman at C_D. In (a) vergence eye movements align the coordinate matrices of the retina. The images are registered at the sensory array. C and D will register identical retinal points within the matrices. The mouth of the pacman $B_R$ and $B_L$ will register at different positions. In (b) image comparison will return the relative coordinate positions of points D, $B_L$,$B_R$ and C. Despite the sparse point disparity, image comparison yields measures of subtense between contours as well as the actual disparate position of contours since two are immutably related. Fig (c) describes the observation that the cyclopean edge at $B_C$ might be achieved by separation of surface layers in a crossed fusion configuration of projections. Subtense D_$B_R$- D_$B_L$ is constrained to a frontal plane. Otherwise the pacmen would look slant or bent. Separation of surface layers in such a manner may underpin spreading of a near separated surface layer.

### 5.2.6 Resolving the 2-D layout of image differences: Uncrossed disparity

A simple projection diagram can also be used to describe the stereoscopic response to a 2-D layout of half-images (that is different perspective views) created at the opposite sign of disparity in the SKS pacman (the percept here, looks like a porthole; see Fig 5.6). The figure depicts a horizontal slice of the visual array of projections subtended at each retina by the pacmen half-images along D_C (standing on the P Plane).

In aligning the pacman half-images along D and C, the arcs of the pacman boundaries are non-disparate (posses precise retinal parity). Fusion of retinal points projected from $B_R$ and $B_L$ yield the cyclopean point $B_C$. The perceived depth of $B_C$ is predicted by the intersection of visual projections through $B_R$ and $B_L$. This looks like a conventional uncrossed fusion of disparate contours.

Clearly, $B_C$ is seen as a distant edge standing behind the bounds of the porthole. It is the edge of the porthole that now appears to partly occlude the distant surface. In theory, either a depth step at $B_C$ or perception of a slanted surface might be signalled by the disparate magnitudes of $D\_B_R$ and $D\_B_L$, that is, disparate subtense across the pacman. Assigning the white surface segment $B_C\_C$ to a distant near depth plane is once again a resolution of the disparate subtense at specifiable parts of the pacman in each eye's view.

Moreover, in the absence of a disparity discontinuity (a step in the point disparity field available in textured surfaces) a crossed pattern of visual projections can yield separation of surface layers not at $B_C$ but at C and D. Therefore, an uncrossed fusion configuration at $B_C$ seems to require that the system splits the percept into two layers *along the arc of the pacman.* This amounts to surface separation in a split-projection configuration as for crossed fusion, *but at a different contour* (hence the perceptual asymmetry between signs of disparity).

In perceiving the porthole, the disparate subtense $B_R\_C - B_L\_C$ is interpreted as a monocular zone visible in the left eye ($C_R\_C_L$). Moreover, separation of surfaces at D means that the region $D_R\_D_L$ is also interpreted as a monocular zone. By assigning the black space between $B_C$ and D to a distant depth plane, the system achieves a percept that resolves the subtense $D\_B_R - D\_B_L$.

In summary, the porthole percept seems to involve a particular stereoscopic response to patterns of disparate binocular subtense, that is, surface separation. Surface separation ostensibly occurs at D and C. But these are non-disparate points. A particular stereoscopic response that I have termed a split-projection configuration appears to be associated with

resolution of disparate subtense in separation of surface layers at the bounding arc of the pacmen (a porthole is seen).

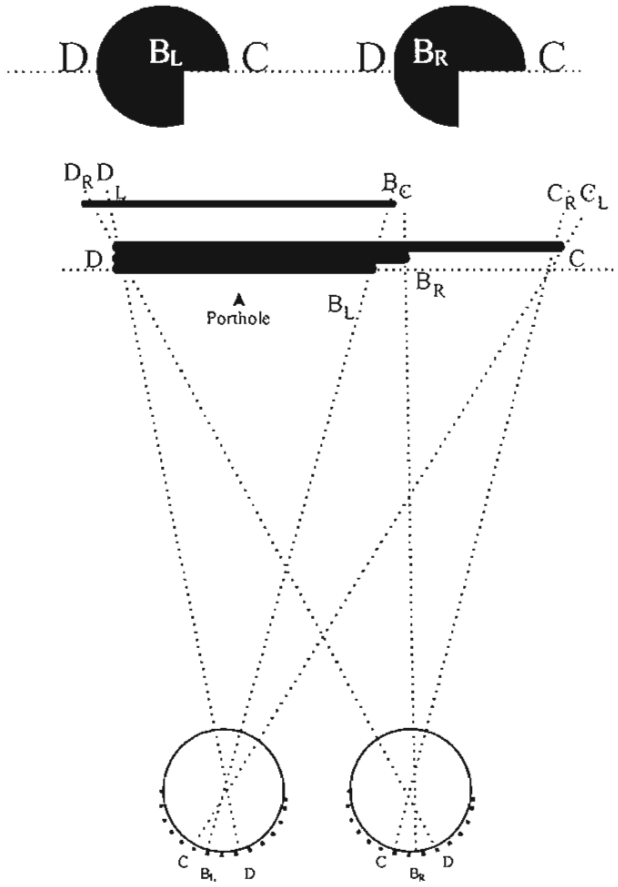



**Fig 5.6 Projection geometry and the perception of a porthole**
This figure is meant to represent the resolution of visual angles (subtense) and retinal disparities associated with fusion of uncrossed disparity in a single pacman taken out of the global context of the Kanizsa square. D_C represents the subtense of the pacman. The mouth of the pacman cuts into $B_L$ in the left eye and $B_R$ in the right. Fusion of $B_R$ and $B_L$ predicts the cyclopean point $B_C$. This point stands at the a distant depth plane. Assignment of that luminance step to a distant depth plane means that the magnitude of binocular subtense on either side of the point $B_C$ might be resolved as a slanted surface from $B_C$ to D or $B_C$ to C. However this is not the percept reported in the perception of a Porthole. Another possible resolution of these disparate angles, is that the magnitude of their differences be hidden from one eye as a monocular zone, that is, behind the arc of the porthole - behind D and C. These image differences are C_ $B_L$ - C $B_R$ (the horizontal luminance step and D_ $B_L$ - D_$B_R$ (the magnitude of the pacman that is reciprocal to disparity between $B_L$ and $B_R$). This means that points D and C represent apices of two triangular occlusion configurations - the sites at which surface layers are separated. This diagram predicts that C, on the near surface layer, aligns points $C_L$ in the right eye and $C_R$ in the left eye, with the magnitude of the differences representing a monocular zone. Similarly, D aligns points $D_L$ in the right eye and $D_R$ in the left eye.

### 5.2.7 Surface separation in functional terms: Uncrossed disparity

Perception of surface separation to achieve a porthole percept might involve processing disparate subtense in addition to conventional point disparity computation. Figure 5.7 depicts likely component processes in fusing a single disparate pacman to achieve the porthole.

Images of the pacman impinge on the sensor array at the retina - image registration. Next, the system coordinates vergence eye movements to lock vergence angle. Achieving vergence lock aligns two identical retinal coordinate matrices (see Fig 5.7a).

Image comparison will reveal reversed disparity and differences in subtense (Fig 5.7b). At the arc of the pacman points D and C, image comparison will detect no disparity. Points along that part of the contour register identical coordinate points in each retina except at the mouth of the pacman. Hence, image comparison (see 5.7b) will return positional *Parity* between the retinal images as well as positional *Disparity* for contours at the mouth of each pacman image.

Intriguingly, surface separation (Fig 5.7c) must split the percept into two layers at two non-disparate points that fall on the arc of the pacman. At both D and C, the white side of a black-white luminance step is assigned to the edge of a near surface layer partly occluding the black side of the same configuration, the pacman. The position of $B_C$ is predicted by the intersection of uncrossed visual projections through $B_L$ and $B_R$. However no depth step is perceived at this contour. This raises the possibility that the porthole percept achieves a separation of surface layers in a split-projection configuration that leads to the spreading of surface layers at the P Plane. Distant and near surface layers share the luminance step at the arc of the porthole bounds.

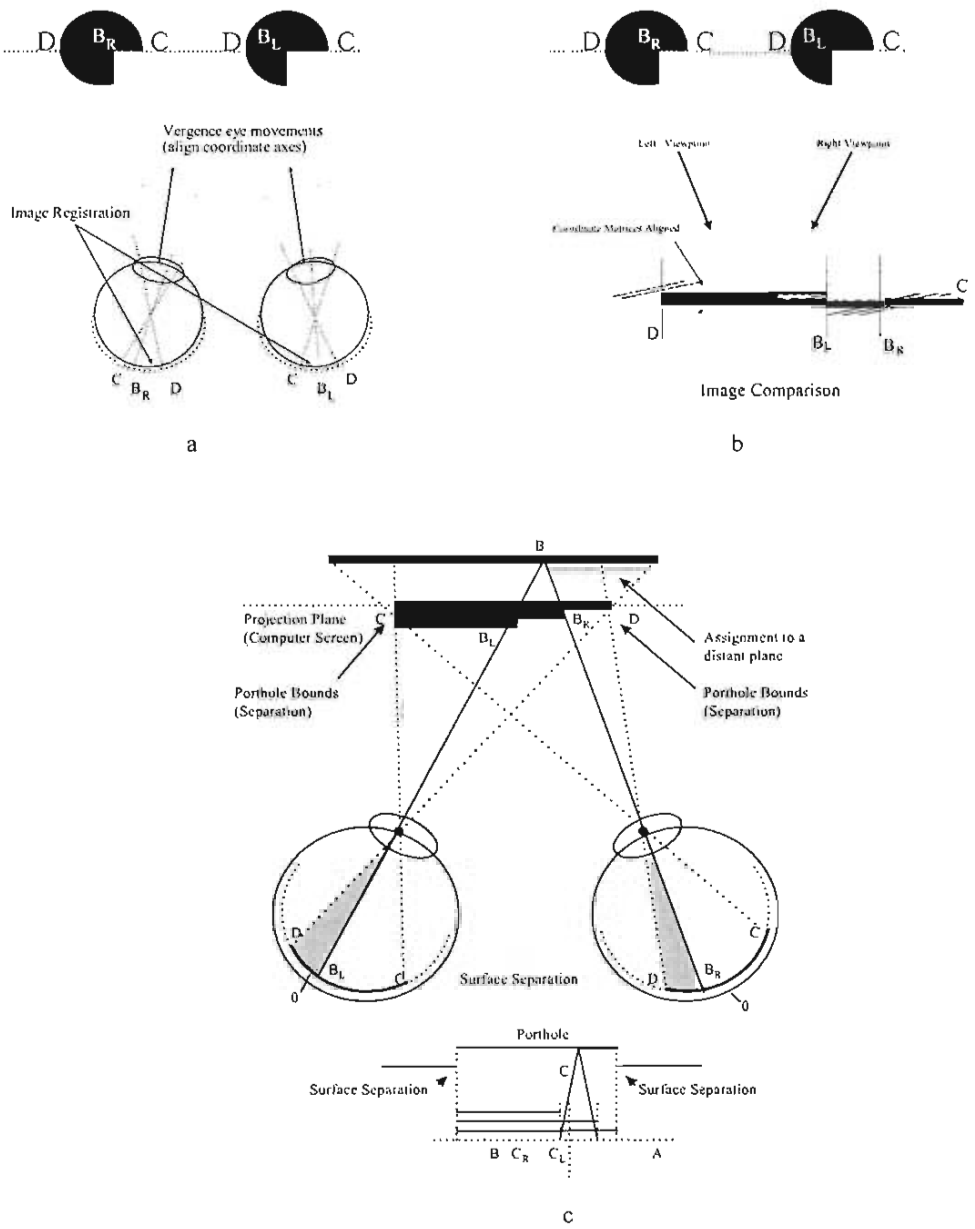



**Fig 5.7. Binocular image processing at an SKS pacman: Uncrossed disparity**

This figure depicts conceptually components of a binocular image processing scheme that might achieve surface separation at a pacman. The figure examines a horizontal slice through the pacman at C_D. In (a) vergence eye movements align the coordinate matrices of the retina. The images are registered at the sensory array. C and D will register identical retinal points within the matrices. The mouth of the pacman $B_R$ and $B_L$ will register at different positions. In (b) image comparison will return the relative coordinate positions of points D, $B_L$,$B_R$ and C. Despite the sparse point disparity, image comparison yields measures of subtense between contours as well as the actual disparate position of contours since two are immutably related. Fig (c) describes the observation that the cyclopean edge at $B_C$ might be achieved by separation of surface layers in a crossed fusion configuration of projections. Subtense D_$B_R$- D_$B_L$ is eonstrained to a frontal plane. Otherwise the pacmen would look slant or bent. Separation of surface layers in such a manner may underpin spreading of a near separated surface layer.

## 5.3 A functional model of a pacman and the SKS percepts

It is possible then to develop a hypothetical and functional model of the binocular perception mechanisms that appear to underpin the SKS percepts. Figure 5.6 is a diagram that summarises the proposed relationship between binocular image processing and the phenomenology evident when disparate pacmen are fused. It also suggests that this local site of binocular image processing, a single pacman, may be a key to understanding the 3-D illusory percept in the SKS. Surface separation at the individual pacmen may initiate surface spreading toward adjacent pacmen. The process is as follows:

I. **Image Registration** Pacmen at 5.8a are registered at the retina, appropriate to crossed or uncrossed disparity. The grey regions represent the area of the mouths – obviously these regions are not bounded in the Kanizsa square.

II. **Vergence Lock** Vergence eye movements align the retinal coordinate matrices, that is, they align identical retinal coordinate points.

III. **Image Comparison** The retinal sensory arrays register both retinal disparity and retinal parity (Fig 5.8b). In the right eye, the mouth of the pacman intrudes farther into the pacman than the left for crossed disparity (and vice verse for uncrossed). The arc of the pacman is non-disparate (retinal parity).

IV. **Surface Separation**

a. Crossed disparity

The system assigns the disparate black pacman to the depth of retinal coordinate parity. The white space at the mouth of the pacman (shown as grey) is assigned to a near depth plane in a manner that resolves disparate subtense (Fig 5.8c). This appears to be achieved in a split-projection configuration at the mouth of the pacmen.

b. Uncrossed disparity

The system assigns the entire space enclosed by the pacman, and the mouth of the pacman to a distant depth plane in a manner that resolves disparate subtense (Fig 5.8d). This appears to be achieved in a split-projection configuration at the bounds of the pacmen.

V. **Surface Spreading**

a. Crossed Disparity

The near surface, separated at the bounds of the mouth of the pacman may spread outward across homogenous space (toward adjacent pacmen for example) as shown in Fig 5.8c.

b. Uncrossed Disparity

The near surface spreads outward across homogenous space from the perceived depth step at the pacman bounding arc and across the mouth of the pacman which takes the appearance of a porthole as shown in Fig 5.8d.

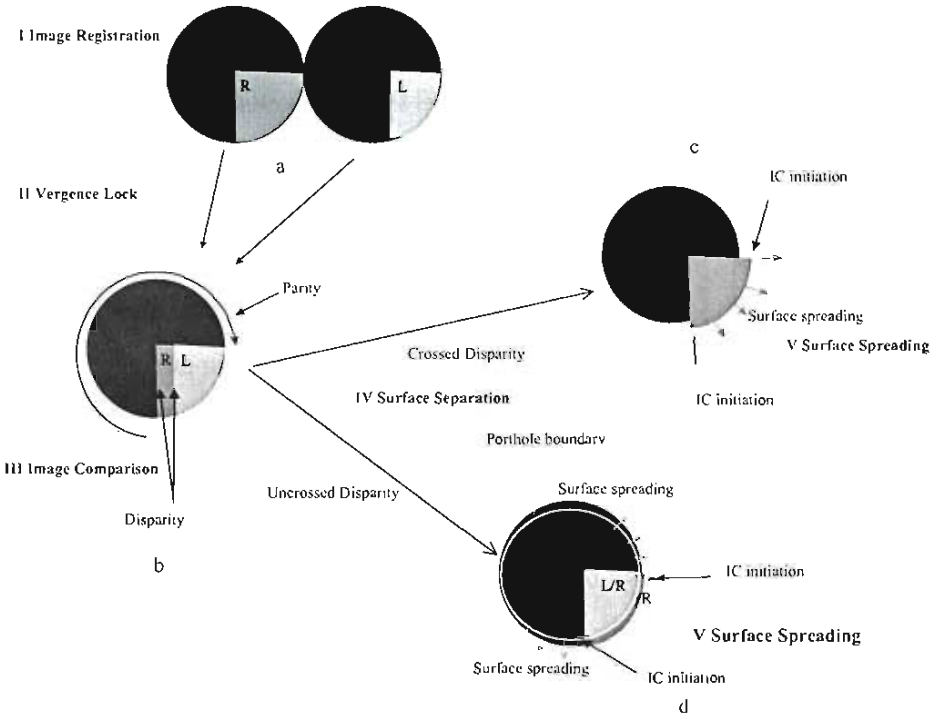



**Fig 5.8. A basic functional map of perceptual processes underpinning perception of a single pacman**
See text for details

In summary, this section has examined binocular processing of the horizontal disparities across a single pacman of the SKS in some detail for both signs of disparity. Separation of surface layers has been accounted for as the system resolves typical positional disparities and the reciprocal differences in subtense between contours that impinge upon the retina. Collectively, these disparities amount to a set of 2-D image differences measures

resulting from perspective projection. The 3-D illusory percepts appear to be closely associated with mechanisms by which the binocular vision system resolves the 2-D layout of inter-retinal differences (created from perspective projection). In that sense, the system behaves in the manner of a Binocular Image Processor where image differences are interpreted in terms of the fit of 2-D image measures (structured by the visual array) into a 3-D perceptual space.

### 5.4 Concluding remarks: Binocular image processing and the SKS.

In the above sections, I have described how binocular image processing might underpin the separation of surface planes in a single disparate pacman. This analysis was based on disparities (at a single pacman) that simulated two perspective views of a black circle partly obscured by a white surface.

From this point, the proposal presented in this chapter will be termed a Binocular Image Processing And Surface Spreading (BIPASS) model of the SKS percepts. The model is descriptive of possible surface perception processes. It emphasises an association between 3-D percepts and the perspective projection of images registered at the retina. What this association means is that 3-D perceptual structure, even in untextured stereograms is immutably related to resolution of the differences in 2-D image structure. In turn, this suggests that so called *illusory* 3-D SKS percepts established in Experiment 1 may not be underdetermined by image difference information as assumed by the Surface Heuristic and Form Computation accounts of illusory 3-D percepts.

In conclusion, it appears that the 3-D SKS percepts may be described in functional terms that encompass surface separation (a binocular image processing mechanism) and surface spreading (an outcome of surface separation). In a phenomenological sense, the 3-D SKS percepts may be products of binocular image processing and surface spreading. Figure 5.9 summarises my BIPASS model for stereopsis in the SKS percepts.

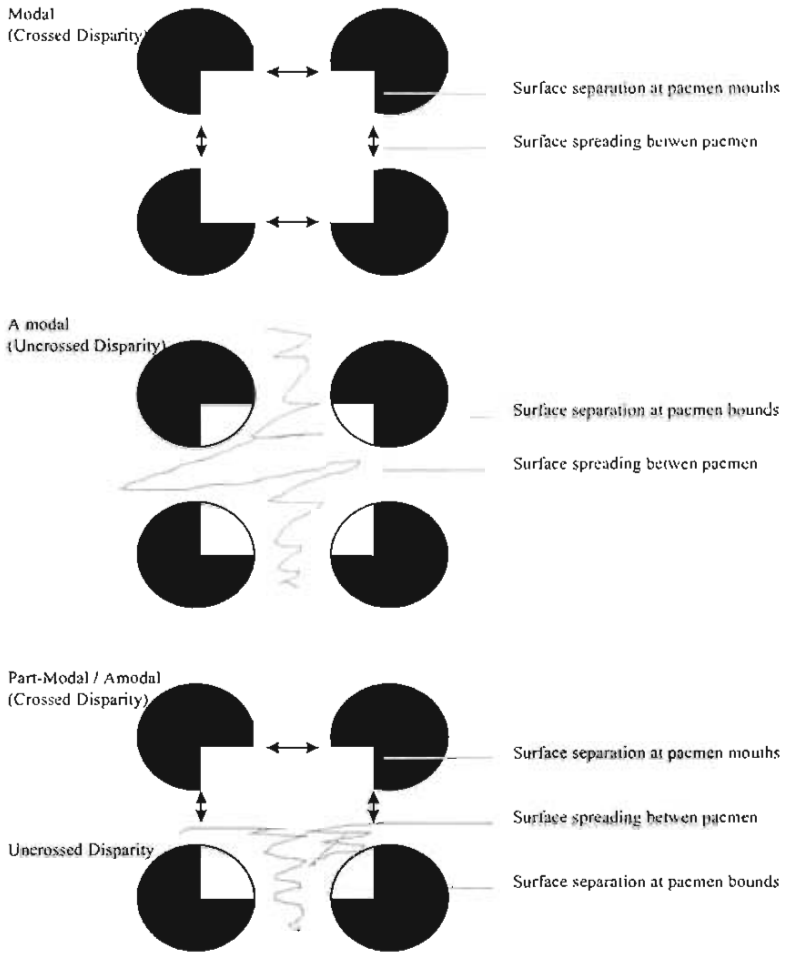



**Fig 5.9 A BIPASS model and phenomenological properties of the SKS percept**
Crossed disparity applied to a model surface drawn to create the pacmen yields surface separation at the pacman mouths. The near surface layer spreads between the pacmen. Crossed disparity yields surface separation at the pacman bounds and this yields surface spreading from the pacman bounds. The part-modal / amodal SKS results when crossed disparity is applied to one pair of pacmen and uncrossed at the other.

# 6. Binocular image processing and surface spreading in the SKS percepts

***Summary:*** *The objective of this chapter was to assess the nature of surface spreading and the integrative mechanisms (such as Grossberg's FCS activities) associated with the SKS percepts. To do this, two pacmen from the SKS were replaced with a luminance and depth referent (a line, Experiment 2 or dot, Experiment 3). Seen slant was used to test whether such a configuration would invoke the sensation of surface spreading between pacmen and referent. An unusual ambiguity was identified for the SKS half-images supporting orientation in the vertical slant-axis. Lastly, Experiment 4 compared seen slant of a Kanizsa diamond against an SKS to examine the nature of this ambiguity.*

## 6.1 Experiment 2 Surface spreading: A configuration effect on disparity interpretation

Chapter 5 examined stereopsis at a single SKS pacman. Separation of surface layers was observed at the mouth of the pacman (crossed disparity) or at the arc of the pacman (uncrossed disparity) when half images were fused. These percepts suggest that separation of surface layers in the SKS need not be associated with completion mechanisms. My argument is that a basic stereoscopic response underpins surface separation, associated with perception of the 3-D illusory figures, that is not well documented. In framing the BIPASS model, I suggested that the response might be explained in functional terms that include a mechanism that leads to the apparent spreading of surface layers separated in depth.

The BIPASS model recognised that the stereoscopic response accesses image difference measures including but not limited to retinal point-disparities across disparate pacmen. I suggested that the perception of separate surface layers at the pacman might be related to the manner in which the system *resolved* physical differences in shape of each eye's pacman image.

It was proposed that the phenomenological properties of the SKS percepts were, in part, a product of surface separation at specifiable parts of the SKS (the pacmen mouths or bounding arc). Depending on the sign and magnitude of disparity, the BIPASS model predicts modal, amodal and part-modal SKS percepts in terms of surface separation and spreading as follows. This is summarised in Fig 6.1.

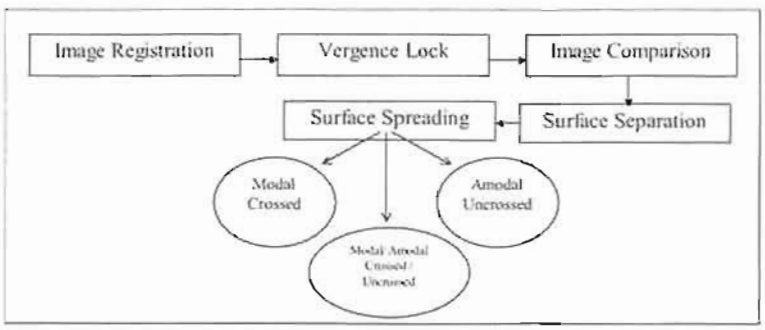



Fig 6.1. A BIPASS model and the SKS percepts

### 6.1.1 Phenomenal character of surface spreading

Perception of 3-D surface spreading is observed when stereograms, constructed using two disparate pacmen set adjacent to each other, are fused (see Fig 6.2, cross fuse L-M). An illusory contour is generated between A and D. The illusory contour looks like a classic modal illusory connection across homogeneous ground. For convenience in later discussion, I will term these stimuli (those constructed from paired adjacent pacmen) the stereoscopic Kanizsa square-pair (SKS-P).

The BIPASS model suggests that the visible contour between A and D (cross fuse L-M in Fig 6.2) is not just an illusory connection but in fact bounds a near surface layer (and therefore a distant surface layer). In other words the illusory contour *is* a fold or depth step in perceived space (constructed as the system resolves the disparate shape of pacmen according to the BIPASS model). Moreover, the near surface layer appears to merge with the homogeneous surrounds of the pacmen (at C and F). Illusory contours here merge with the surrounding space. The subjective sense of depth displacement may dissipate and the illusory contours fade into homogeneous surrounds. A near separated surface layer spreads from the mouths of the SKS-P pacmen and gradually fades into ground.

Finally, when disparity in the SKS-P is reversed, portholes are seen (cross fuse M-R in Fig 6.2). No illusory contour forms between A and D; instead, an illusory contour *completes* the pacmen at A_C and D_F in a smooth arc. This illusory contour *is* a depth step. Further, white corners seen through the portholes look like they belong to the same partially seen object. This looks like a classic amodal connection between partly occluded features aligned along a smooth 2-D trajectory across and behind the space between the portholes (after Kellman and Shipley, 1991).

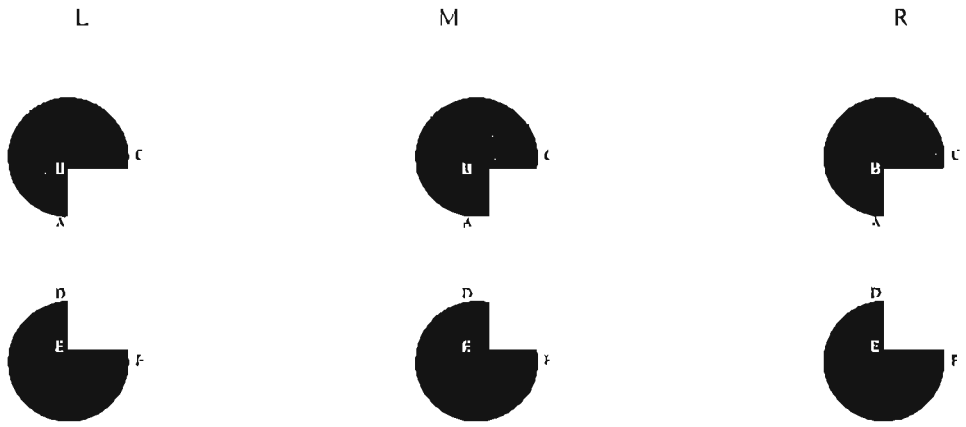



**Fig 6.2. Surface spreading between adjacent stereoscopic Kanizsa pacmen**
Crossed fusion of the pair L-M yields separation of surface planes at the mouths of the pacmen. A strong illusory contour appears to continue the near surface layer between A and D. This contour demarcates a depth step across a physically homogeneous plane of projection. Thus the separation of surface layers represents the spreading of illusory near surface layer in between the pacmen. The layer seems to fade into the plane of projection beyond C and F. Fusion of M-R sees the classic porthole effect. Two white corners within the mouths of the pacmen appear to stand at a distant depth plane. There is no illusory contour formed between A and D. An illusory contour does emerge along CA and DF giving the impression that the pacmen are completed. These contours again demarcate a depth step between the near and distant surfaces.

In light of these observations, it does not seem feasible to understand generation of the SKS percepts without exploring surface spreading in between the pacmen in some detail. The question is how to reconcile observations of surface layers in a SKS-P configuration with the classic concept of modal or amodal completion. Indeed, the mechanism of illusory contour induction has often been thought to involve integrative connections across homogeneous regions of the visual field between physically discontinuous luminance contours (binding together physically discontinuous image features).

For example, Kanizsa's original explanation of the illusory square was that the perceptual system completed a partially bounded *form* leading to the impression of contours where none existed, that is, modal completion[x] (see section 2.1). Since then it has been demonstrated that a partially bounded form is not necessary for the perception of illusory contours. Day (1987) argued that the critical factor was that configurations yielding 2-D illusory contours were a cue to edge.

Kanizsa (1987), and Kennedy, (1975, 1976, 1981, 1987, 1988, 1991) have systematically shown that a *completion process* alone cannot account for the perception of

---

[x]. He argued that in fact the amodal completion of the pacmen was the central factor

illusory contours in 2-D stimuli. Kellman and Shipley's (1991, 1992) theory of illusory contour continuation between relatable contours attempted to account for this[9]. Moreover, Anderson (1994) and Anderson and Julesz (1995) have also shown that no completion or even continuation process *per se* is necessary for induction of illusory contours in stereopsis. However, it was shown in Chapter 5 that a single disparate pacman taken out of the SKS context (crossed disparity) yields no illusory contours.

Grossberg has maintained that visible (modal) and invisible (amodal) connections underpin the percepts identified in Fig 6.2 (between adjacent pacmen). The connections link physically uncorrelated luminance contours across a homogeneous field of the image. In Grossberg's view a pre-visual Boundary Contour System creates both connections in precisely the same manner. This system is a network of neural activity that links the activities of hyper-complex cells in the visual cortex. Its role is to integrate the end-cut responses of the cells.

Grossberg suggests that the visual or phenomenological character of the modal contours (that they are visible) is created by the Feature Contour System's binocular FIDOS (filling-in domains). These mechanisms fill-in the features at particular depth planes from the activity of various disparity pools. All of these activities, Grossberg argues, are modulated by feedback from a higher-level Object Recognition System. The ORS operates like a hypothesis tester. If the parts recovered from the BCS and FCS don't fit a recognisable object mould then the networks re-iterate their activity.

FACADE theory is a powerful explanatory device. FACADE theory is functionally very similar to Kanizsa's conceptual approach to the classic 2-D stimuli that emphasised completion. It is also similar to Kellman and Shipley's arguments in that it poses modal and amodal connections as two faces of the same coin. In fact all three of these theories predicts integrative projections cast across homogeneous spaces of the visual field between relatable luminance contours that can be either visible or invisible depending on the visual context.

The BIPASS model differs somewhat from both of these ideas but is intimately related to them. The BIPASS proposal is that in resolving differences between the two retinal images (the SKS) the system splits apart two surface layers. It does this at the mouths of the pacmen (crossed standing disparity) or at the curved bounds of the pacman (uncrossed standing disparity). Then, surface spreading can emerge from the contours at which the system splits

[9] These issues were addressed in detail in chapter 3. See section 3.1.

the percept into layers. Evidently surface spreading can arise along oblique depth planes *ie.* can look slanted. The mechanism can create part-modal 3-D SKS percepts. This suggests that previously proposed filling-in mechanisms (e.g. Grossbergs's binocular FIDOs) must compute filling-in along a depth gradient. The BIPASS model need not rule out an integrative or completion mechanism. It does, however, emphasise the extensive array of retinal measures available from image comparison.

### 6.1.2 Exploiting subtleties in the layout of disparity to explore 3-D separation and spreading of untextured layers

A basic tenet of the BIPASS model is that surface separation might be a stereoscopic response to disparity in the size and shape of large scale visual features. The resulting percept is the sense of one surface partly occluding another at the pacmen. Surface spreading then might be thought of as continuation of the perceived separation of depth layers between large scale features.

Subtleties in the layout of disparities, in Fig 6.2, demonstrate something interesting about the nature of surface spreading in the SKS percepts. In particular, horizontal contours at the mouths of the pacmen in the SKS-P possess a peculiar ambiguity for stereopsis (at B_C and E_F). Luminance contours horizontal to the line of sight (particularly untextured contours), yield limited disparity information (Anderson, 1997; Gulick and Lawson, 1976). It was determined to exploit these subtleties in the next two experiments to examine the nature of possible modal (visible) and amodal (invisible) completion responses that are at the heart of theories such as Grossberg's.

The horizontal luminance contours in question are those at B_C and E_F in the SKS-P half-images presented in Fig 6.2[10]. Note that in the right eye (image L for crossed fusion) the contour (B_C) is longer than the corresponding contour (B_C) in the left eye's view (image M for crossed fusion). The system could feasibly interpret horizontally aligned contours in several different ways. Modal and amodal connections may impact on interpretation of the orientation of these contours. The obvious way to test this possibility was to manipulate the SKS-P configuration by inserting a referent some distance from the pacmen to see if a surface spread to that referent.

---

[10] Note that these "contours" are actually segments of the single pacman bounding contour.

Figure 6.3 demonstrates the source of potential ambiguity. A luminance and depth referent that is physically uncorrelated with these horizontal contours, such as an adjacent pacman (in the SKS), may affect a change in interpretation of horizontal disparities. There are potentially several different feature scales at which disparity could be interpreted. One interpretation is based on the disparate angles subtended by the luminance contour ($A_R$_B – $A_L$_B) itself. Another interpretation is disparate subtense between the luminance contour and the referent, that is, the disparate subtense between C and E in each eye's view ($A_R$_E – $A_L$_E). A third possible interpretation is of a surface standing fronto parallel ($D_R$_C – $D_L$_C).

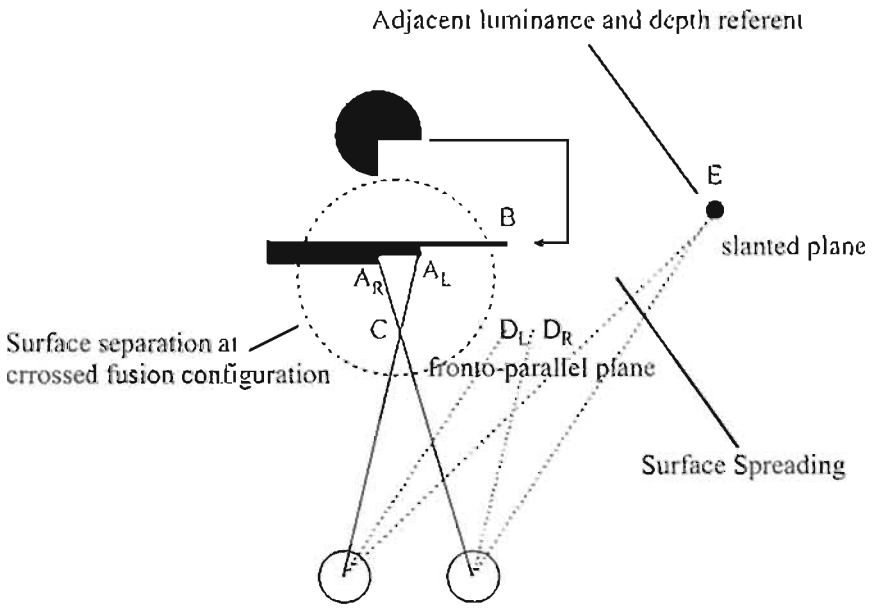



**Fig 6.3. Interpretations of disparities present at horizontal contours in fusing the SKS pacmen**

When crossed (or uncrossed) disparity is introduced to the mouths of a pair of pacmen then they are fused, surface separation occurs where $A_R$ and $A_L$ are fused. The result is a depth step at point C. The disparate subtense between $A_R$_B and $A_L$_B can theoretically be resolved in several ways. Three possibilities are shown here. First, resolution to a frontal plane means that the disparate subtense remains. Points $D_R$ and $D_L$ mark the ends of the depth plane. The difference in subtense at $D_R$_C and $D_L$_C extends beyond the arc of the pacman at B. Second, the disparity can be interpreted as part of a difference in subtense. The result is surface spreading toward point E. This would yield a slanted surface plane. Third, the same image difference could theoretically resolved in a steep slant toward point B, the arc of the pacman.

In perceiving slant and depth in SKS percepts (in Experiment 1), the orientation of horizontal contours must have been interpreted in terms of disparate subtense *between* pacmen and not local image differences, that is, according to a large feature scale. Experiment 1 demonstrated that the system quite precisely achieved this. Modal, amodal and part modal slanted percepts were produced. This means that the remarkably accurate impression of the slanted amodal and

even the part-modal SKS involved a disparity measure (an image difference) other than the actual physical dimensions of the horizontal contours. The system seems to be responsive to the disparate subtense of spaces *between* luminance contours.

It is important to note that interpretation of disparity in the horizontal slant-axis, in a SKS pacman, is not affected in the same way as the vertical slant-axis. Introducing horizontal shear at the mouth of the pacmen creates an equivalent difference in subtense at B_C and E_F (see Fig 6.4). However, it is the disparate orientation of B_C and E_F that defines the perceived stereoscopic (or cyclopean) orientation of the contour. This means that local and large feature scale disparities are congruent with each other.

The magnitude of the disparate length of these contours (C_B and E_F) has no bearing upon their orientation when the images are fused[11]. Their stereoscopic orientation is defined by fusion to a singular cyclopean contour. This is not the case for horizontal contours. Experiment 2 was designed to test the difference between the physical layout of image differences for shear (rotation of a surface in a SKS-P about the horizontal) and simple relative magnification (rotation of a surface in a SKS-P about the vertical).

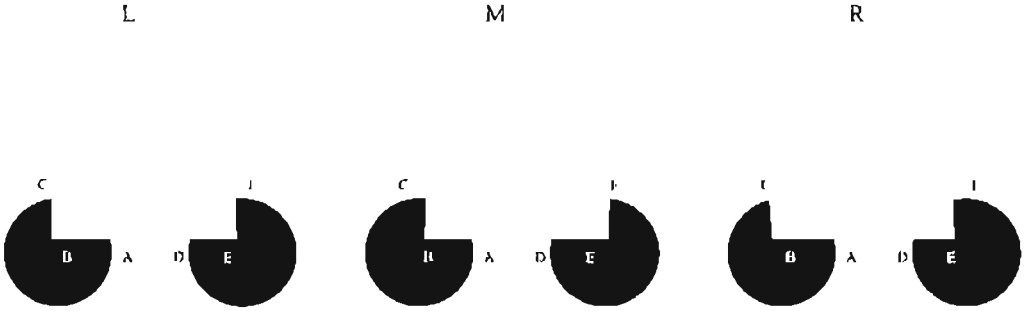



**Fig 6.4. Stereoscopic rotation about the horizontal axis in an SKS-P**
Crossed fusion of the L-M pair yields the percept of a surface layer separated from the pacman and assigned to a near surface layer (modal surface). Orientation of the near surface layer is defined by the disparate orientation of contours at B_C and E_F. However, the resolution of horizontal contours of disparate length at B_A and D_E appears to be interpreted as disparate subtense across homogeneous space since the near surface spreads between the pacmen at AD. Fusing pair M and R sees the orientation of the sheared contours maintained in the opposite sign of disparities (amodal surface) viewed through portholes.

Experiment 2 was designed to explore the effect of configuration differences upon surface spreading. A line was introduced to the SKS-P configuration. The line was intended to represent a possible luminance and depth referent. It was positioned so that its ends aligned

[11] Resolution of disparate subtense at this local level of the pacman bounds was discussed in detail in Chapter 5

with the bounds of the mouths of the pacmen. I wondered if such a referent would be sufficient to influence surface spreading (and therefore disparity interpretation). The basic stimuli are shown in Fig 6.5. These figures will be called the stereoscopic Kanizsa square-line configuration (SKS-L).

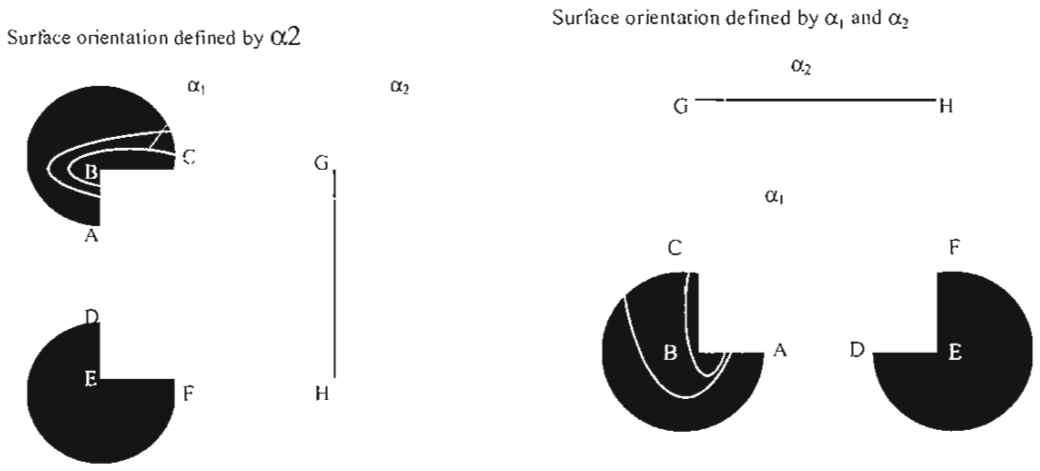
In the SKS-L half-images, faint illusory contours appear to continue from the pacmen toward the ends of the line G_H. Experiment 2 tested whether a 3-D surface layer separated at the pacmen by stereopsis (ie. at the pacmen mouths), would spread in between the pacmen and the line. This was to be determined by measuring seen slant. So seen slant in the SKS-L configuration was to be compared to that in the SKS-P configuration, discussed above, while maintaining constant disparities at the pacmen. Seen slant was used as a metric to examine the conditions required for surface spreading.

In the SKS-L configuration, the actual position of the line in each half-image could be defined by the same model of relative magnification or shear applied to the SKS configuration in Experiment 1. Hence, for theoretical rotation about the vertical slant-axis (Fig 6.5a), disparity was achieved by manipulating the relative magnitude of the square set of points describing EBGHE in each eye's view. In constructing stimuli appropriate to the horizontal slant-axis (Fig 6.5b), disparate shear was applied to EBGHE. Standing disparity was also introduced to this square EBGHE. In manipulating standing disparity, the square EBGHE was shifted horizontally relative to that in the left eye's view.

In order to generate the perception of a slanted modal or amodal surface layer the visual system must interpret disparate subtense of horizontal contours (B_C and E_F) in relation to the magnitude of disparate subtense between B-G and E-H. In the horizontal slant-axis shear at the pacman itself (BC and EF) and the position of the line itself (BG and EH) both prescribe the same retinal disparity, that is, disparity appropriate to rotation in the horizontal slant-axis.

---

(section 5.3).




a - Vertical slant-axis

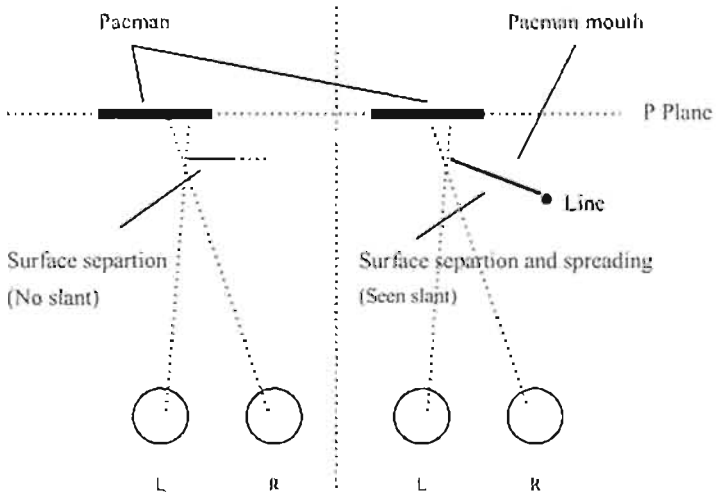
b- Horizontal slant-axis

**Fig 6.5. Continuation and surface spreading in an SKS-L**
These figures show the basic stimulus used in Experiment 3. In (a) a line has been placed adjacent to two pacmen. The line represents a luminance and depth referent that may induce, or support, spreading of illusory contours towards its ends. The ends of line (H and G) are continuous with the horizontal contours B_C and E_F in (a) and with the vertical contours B_C and E_F in (b). Disparities appropriate to the surface standing forward or behind the projection plane was applied by shifting the entire square set of points EBGHE slightly in eaeh eye. Disparities appropriate to vertical slant-axis were introduced by varying the horizontal magnitude of EBGHE in each eye. The same applied to rotations about the horizontal axis but with horizontal shear applied to EBGHE. This was meant to test a possible difference in orientation of an illusory surface layer due to alternative resolutions of disparities in the vertical slant-axis (a). No such difference exists in (b).

Experiment 2 therefore compared the spreading of a separated surface layer in the SKS-P and SKS-L configurations. Given that precisely the same disparities were applied at the pacmen for these stimuli, it was thought the comparison would yield differences in surface spreading due to the effect of configuration. These differences should provide insight into 3-D surface spreading and its relationship to the stereoscopic response. Comparison of seen slant in the SKS-P and SKS-L configurations was expected to highlight *interpretation* of disparity layout that could be associated with some kind of completion mechanism.

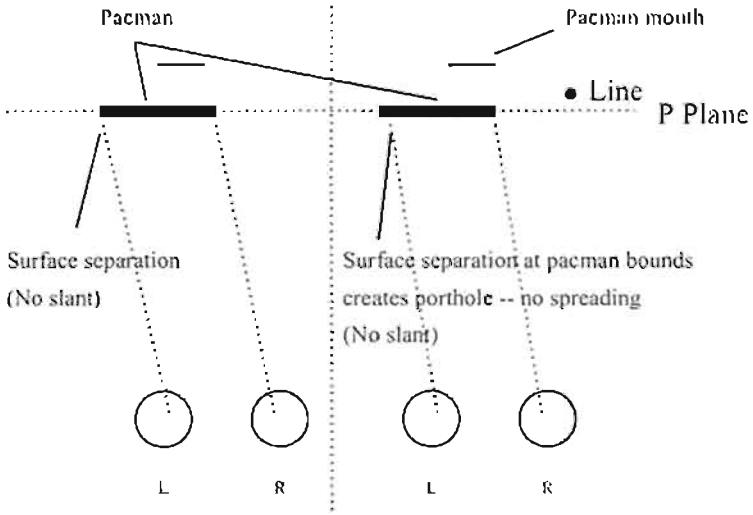
As described in section 6.2.2. a slant-axis effect was predicted because of the ambiguous nature of horizontal contours in rotation about the vertical axis. It was expected that a slanted stereoscopic surface would be seen in between pacmen in the SKS-P only for the horizontal axis of theoretical rotation (generated by horizontal shear and therefore not influenced by configuration). It would be expected that a slanted stereoscopic surface would be seen in the SKS-L for both axes of rotation. Hence, a comparison between the two axes of slant was included.

It has been shown in Chapter 4 that standing a SKS behind the P Plane "switched off" the modal (visible) surface appearance between the pacmen. This has been associated with an amodal completion mechanism. Such a mechanism clearly should not impact the SKS-P, but may play a role in organisation in the SKS-L. Hence, it was anticipated that no slant would be seen when the SKS-P was stood behind the P Plane (ie. uncrossed standing disparity). A standing disparity component was therefore added to disparities at the pacmen for both SKS-P and SKS-L configurations. This was to compare spreading between the modal SKS-L percepts (standing forward of the P Plane) and the porthole percept (standing behind the P Plane). These phenomenological predictions are summarised in Fig 6.6.




a. SKS-P b. SKS-L

Standing forward of P Plane (standing disparity = 20 arcmins)




c. SKS-P d. SKS-L

Standing behind P Plane (standing disparity = -20 arcmins)

**Fig 6.6. Surface separation and spreading in SKS-L and SKS-P configurations rotated about the vertical axis**

The disparities applied to these configurations are the same. In a, fusion of the SKS-P pairs, with crossed disparity (20 arcmins), should lead to surface separation, but limited surface spreading and hence no seen slant at the mouths of the pacmen. In b, fusion of the SKS-L pairs, with crossed disparity (20 arcmins), should lead to surface separation associated with surface spreading toward the line. Stereoscopic slant should result. In c, fusion of the SKS-P pairs, with uncrossed disparity (-20 arcmins) should lead to surface separation (porthole), but no surface spreading and hence no seen slant at the mouths of the pacmen. In d, fusion of the SKS-L pairs, with uncrossed disparity (-20 arcmins) should lead to surface separation (porthole), but no surface spreading and hence no seen slant at the mouths of the pacmen. Seen slant should be evident in all rotations for both levels of standing disparity and for both SKS-L and SKS-P configurations.

*Measures: seen slant in the SKS-P and SKS-L percepts*

In summary, an experimental comparison of the SKS-P and SKS-L stereograms was devised to analyse seen slant of an illusory surface layer that forms between adjacent stereoscopic pacmen. A square, equiluminant with its surrounds, was projected onto two black circles in each half-image. This square had been geometrically transformed to simulate the pattern of disparity created by rotating a square in natural perspective then drawing it to partly obscure two Kanizsa square pacmen. A fully factorial repeated measures design compared the impact upon seen slant of four independent variables: Theoretical Rotation (0, 40$^0$), Slant-Axis (horizontal , vertical), Standing Disparity (-20, 20$^0$) and Configuration (the SKS-P, SKS-L comparison).

## 6.1.3 Method

*Subjects*

Ten subjects were drawn from the pool of volunteers. Screening and practice sessions were the same.

*Stereograms*

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a pair of black circles subtending 3$^0$ in diameter (at 750mm viewing distance). The circles were positioned so that a square drawn through their centres would subtend 7$^0$. *Pacmen* were created by drawing a square, equiluminant with the background, so that it partly obscured the circles. The square symmetrically overlaid the circles (equal intrusion on all sides). The square intruded 1/3 the circles' radii creating a pair of SKS type pacmen.

Mouths of the pacmen were transformed according to the monocular transformations of a square shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was then introduced by applying ½ Ogle's M to the overlaid square, symmetrically and in opposite signs (see Fig 6.7). Standing disparity was generated by shifting the overlaid square in equal and opposite directions in each half-image (see Fig 6.7). A one pixel weight black line was drawn at the boundary of the transformed overlaid square adjacent to the pacmen to create the SKS-L stereograms and was absent in the SKS-P configurations.


**Fig 6.7. Patterns of disparity in the SKS-L**
The figure at (a) shows one half-image supporting vertical slant-axis. Disparity appropriate to vertical slant-axis was applied by increasing the magnitude of the Kanizsa line-square ($\alpha_2$) in one eye relative to the other. This changes the relative magnitudes of the mouths of the pacmen ($\alpha_1$). Standing disparity was then manipulated by constraining ($\alpha_2$) while shifting the line-square in one eye relative to the other. The transformed overlaid square used to generate appropriate disparities is shown as a dashed line. Fig (b) shows the horizontal differences applicable to horizontal slant-axis in the line-square stimuli. Disparate shear ($\lambda$) was applied in each eye while constraining ($\alpha_2$). Then standing disparity was introduced by again constraining ($\alpha_2$) while shifting the illusory figure in one eye relative to the other so adjusting the relative magnitude of ($\alpha_1$) in each eye.

Image pairs were presented on alternate frames at 200 frames per second. Through the shutter goggles, background luminance was 0.7 cd $m^{-2}$ with the black pacmen 0.09cd $m^{-2}$.

*Design and Procedure*

The study used a fully crossed, four-way (2x2x2x2) repeated measures design to explore the effects on seen slant of theoretical rotation (0, $40^0$), in conjunction with standing disparity (+20 arcmins [standing above the P Plane], -20 arcmins [standing behind the P Plane]) slant-axis (Horizontal/Vertical) and configuration (SKS-P, SKS-L). Six complete repetitions of this design were used, making a total of 72 slant estimation scores.

Subjects were asked to judge the orientation, if any, of a surface layer at the mouths of the pacmen. This was achieved by rotating the comparison stimulus to match the orientation of that region of the stimulus. The comparison stimulus returned the degree of disparity applied to the comparison stimulus.

### 6.1.4 Results and discussion

Results for each subject were averaged across repetitions. Each independent variable was treated separately.

A one-way (1x2) repeated measures analysis of variance was conducted to examine the effect of theoretical rotation (0, $40^0$) upon the seen slant of an illusory surface observed when half-images were fused. Seen slant at $40^0$ theoretical rotation was some $6^0$ greater than at $0^0$ (see Fig 6.8) and the effect was significant: $F(1,7) = 24.53$, $p < 0.01$. A substantial residual seen slant was found at $0^0$. This was somewhat surprising. The residual was probably due to perception of slant in the SKS-P (where a surface appears to spread from the mouths of the pacmen standing forward of the P Plane into the homogeneous surrounding regions).

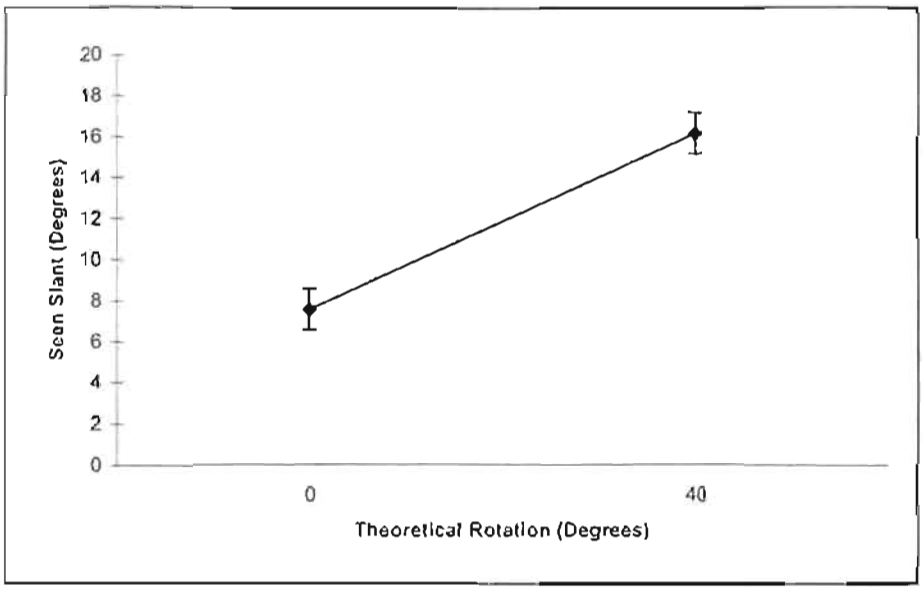



**Fig 6.8. The impact of theoretical rotation upon seen slant in the SKS-P and SKS-L configurations**
This figure graphs mean slant estimates for the main effect upon seen slant of $0^0$ and $40^0$ theoretical rotation. Standard error bars have been included.

Similarly, a one-way (1x2) repeated measures analysis of variance was conducted to examine the effect of Standing Disparity (+20, -20 arcmins) upon the seen slant of an illusory surface observed when half-images were fused. It was found that for surfaces standing above the P Plane (20 arcmins), seen slant was about $6^0$ greater than those standing behind the P Plane (-20 arcmins) and that effect was significant: $F_{(1,7)} = 21.55$, $p < 0.001$ (see Fig. 6.9).

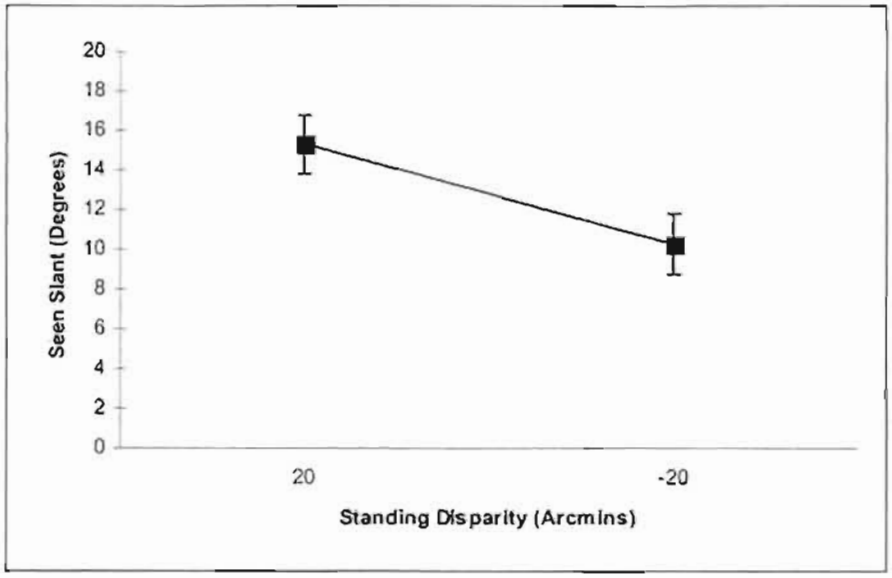



**Fig 6.9. Effect of standing disparity upon seen slant in SKS-L and SKS-P configurations.** This figure compares the mean slant estimates for the two levels of standing disparity. 20 arcmins standing disparity is associated with modal 3-D illusory surface and -20 arcmins was associated with amodal percepts (seen through portholes). Standard error bars have been included.

Stimulus configuration impact upon seen slant was also analysed using a one-way (1x2) repeated measures ANOVA that compared SKS-P and SKS-L configurations. Seen slant in the SKS-P configuration was some $4^0$ less than the SKS-L. The effect was significant: $F_{(1,7)} = 16.90$, $p < 0.01$ (see Fig 6.10). A similar comparison of slant-axes however, yielded no significant difference.

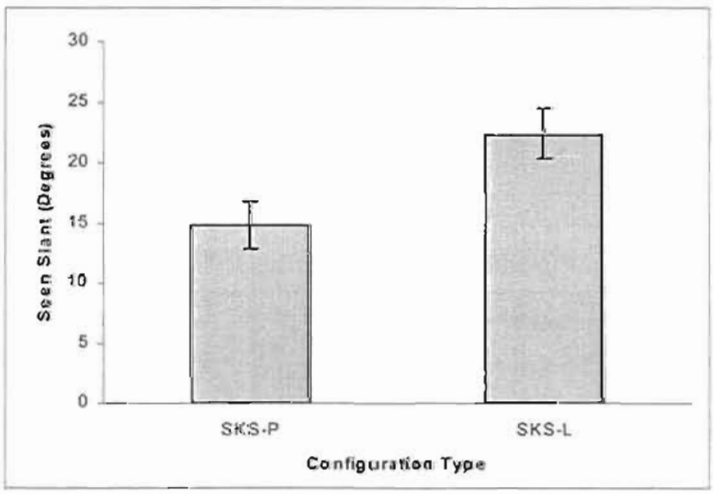



**Fig 6.10. Impact of seen slant in the SKS-P and SKS-L configurations**
This figure graphs mean slant estimates (seen slant) for each of the configurations presented (SKS-P and SKS-L). Standard error bars have been included.

*One way ANOVA at $40^0$ Theoretical Rotation*

To examine the impact of standing disparity, slant-axis and configuration upon seen slant, a one-way (1x8) repeated measures ANOVA was conducted upon the seen slant data returned for theoretical rotation of $40^0$. This comparison enabled a direct assessment of the subtle phenomenology that had been anticipated (the spreading of a surface layer from surface separation at the mouth of the pacmen toward a luminance and depth referent). This effect was significant: $F(7,8) = 6.9$, $p < 0.0001$. Figure 6.11 graphs the mean slant estimates at each configuration (SKS-P / SKS-L), each level of standing disparity (20 / -20 arcmins) and each slant-axis (horizontal / vertical – slant axes are differentiated by shading). A Least Squared Means comparison was conducted across all means to directly contrast independent variables with theoretical rotation constant at $40^0$ (see Table 1).

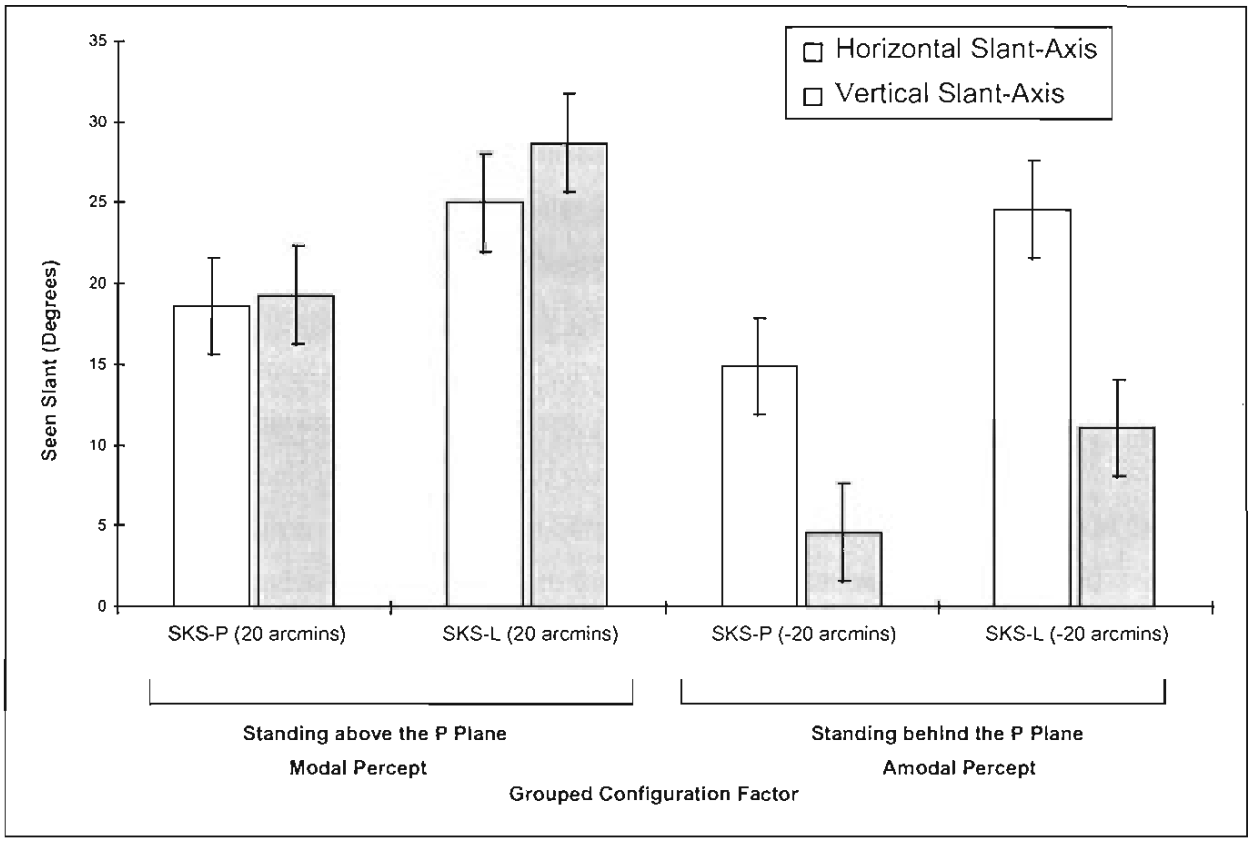



**Fig 6.11. One -way comparison of standing disparity, slant-axis and configuration**
This figure plots the mean slant estimates for the SKS-P and SKS-L configurations at the two level of standing disparity and for each separate slant-axis (these have been shaded to indicate horizontal and vertical slant-axis). Note that theoretical rotation for all of these presentations was $40^0$. Standard error bars have been included.

Interestingly, there was no significant difference between the slant-axes for either SKS-L or SKS-P configurations with +20 arcmins standing disparity (standing above the P Plane, surface separation at the moths of the pacmen). Seen slant was evident even in the SKS-P with no adjacent referent. This suggests that a surface layer spread from the mouth of the pacmen oriented toward the P Plane and gradually merging with it. The sense of slant was substantially stronger in the SKS-L however (as expected). From Table 1, about the vertical, the SKS-P(+) x SKS-L(+) comparison $t = -2.178$, $p < 0.05$ and about the horizontal there was no significant difference between configurations. There were also no significant differences between the two axes when standing disparity stood the surface layer forward of the P Plane.

The anticipated impact of the –ve sign of standing disparity was supported. From Table 1, the SKS-P(-) configuration showed a difference between slant-axis of $10^0$ and the difference was significant $t = 2.382$, $p<0.05$. Moreover, the SKS-L(-) configuration yielded a difference in seen slant between each axis (slant-axis anisotropy) of some $13^0$, and this was a significant difference: $t=3.137$, $p< 0.05$.

In a general sense, these findings mean that for the SKS-P configuration standing above the P Plane, observers' slant estimates are consistent with perception of a surface rotated at about $18^0$ and spreading from the pacmen into homogeneous surrounds. For the SKS-L configuration, seen slant was about $25\text{-}30^0$. This is consistent with perception of a surface layer separated at the mouths of the pacmen and spreading from the near to the distant depth plane or luminance and depth referent.

In the porthole percept (-ve sign of standing disparity), there is evidence that subjects experienced only a very weak association between the adjacent referent and the mouth of the pacman and this region was seen as near to fronto-parallel in rotation about the vertical, whereas about the horizontal axis, local disparity was interpreted as slanted. These findings are therefore not supportive of the computational approach to amodal (invisible) connections.

| | Vs. | Diff. | Std. Error | t-Test | P-Value |
|---|---|---|---|---|---|
| SKSP+H | SKSP+V | -.750 | 4.304 | -.174 | .8624 |
| | SKSL+H | -6.375 | 4.304 | -1.481 | .1450 |
| | SKSL+V | -10.125 | 4.304 | -2.352 | .0227 |
| | SKSP-H | 3.750 | 4.304 | .871 | .3878 |
| | SKSP-V | 14.000 | 4.304 | 3.253 | .0021 |
| | SKSL-H | -6.000 | 4.304 | -1.394 | .1696 |
| | SKSL-V | 7.500 | 4.304 | 1.743 | .0877 |
| SKSP+V | SKSL+H | -5.625 | 4.304 | -1.307 | .1973 |
| | SKSL+V | -9.375 | 4.304 | -2.178 | .0342 |
| | SKSP-H | 4.500 | 4.304 | 1.046 | .3009 |
| | SKSP-V | 14.750 | 4.304 | 3.427 | .0012 |
| | SKSL-H | -5.250 | 4.304 | -1.220 | .2284 |
| | SKSL-V | 8.250 | 4.304 | 1.917 | .0611 |
| SKSL+H | SKSL+V | -3.750 | 4.304 | -.871 | .3878 |
| | SKSP-H | 10.125 | 4.304 | 2.352 | .0227 |
| | SKSP-V | 20.375 | 4.304 | 4.734 | .0001 |
| | SKSL-H | .375 | 4.304 | .087 | .9309 |
| | SKSL-V | 13.875 | 4.304 | 3.224 | .0023 |
| SKSL+V | SKSP-H | 13.875 | 4.304 | 3.224 | .0023 |
| | SKSP-V | 24.125 | 4.304 | 5.605 | .0001 |
| | SKSL-H | 4.125 | 4.304 | .958 | .3426 |
| | SKSL-V | 17.625 | 4.304 | 4.095 | .0002 |
| SKSP-H | SKSP-V | 10.250 | 4.304 | 2.382 | .0212 |
| | SKSL-H | -9.750 | 4.304 | -2.265 | .0279 |
| | SKSL-V | 3.750 | 4.304 | .871 | .3878 |
| SKSP-V | SKSL-H | -20.000 | 4.304 | -4.647 | .0001 |
| | SKSL-V | -6.500 | 4.304 | -1.510 | .1374 |
| SKSL-H | SKSL-V | 13.500 | 4.304 | 3.137 | .0029 |

Table 1. Least Squared Means pairwise comparison (all possible pairs)

The table lists pairwise comparison for the configuration and Standing Disparity effects (20 arcmins = "+" and – 20 arcmins = "-") upon seen slant in the SKS-P and SKS-L for rotations at $40^0$ for the horizontal (H) and vertical (V) slant-axis.

These findings suggest that part of the explanation of surface spreading is some kind of linkage between the pacmen and an adjacent luminance contour but that an adjacent referent is not necessary to yield a slanted spreading layer. The line in the SKS-L appears to act as a depth anchor by which the system interprets an ambiguous disparity in the arrangement of horizontal contours (rotation about the vertical axis). The line seems to lead to assignment of the space between it and the adjacent pacmen to an oriented depth plane in the modal percept but to a lesser extent in the amodal percept.

Slant estimates in the modal form of the SKS-L and SKS-P (stands above the P Plane) suggests that the presence of modal illusory contours generated a somewhat stronger connection between the pacmen and the adjacent luminance and depth referent in the SKS-L configuration. This finding is consistent with the findings in Experiment 1 that seen slant in SKS configuration was substantially greater for percepts in which standing disparity positioned the 3-D illusory plane fully forward of the P Plane.

Overall, Experiment 2 suggests that large-scale configuration effects play a role in 3-D perceptual organisation of the 3-D Kanizsa line-square stimulus. The observation that slant estimates even for the horizontal slant axis were substantially attenuated suggests that the long distance connections, between pacmen and line, were weaker in the SKS-L than those in the SKS.

The findings do not strongly support a mechanism such as Grossberg's pre-visual Boundary Contour System. If the modal and amodal form were underpinned by an identical BCS activity one would expect seen slant in the SKS-L to be no different regardless of its subsequent appearance (modal or amodal). Still, the findings point to a subtle effect of configuration.

It is perhaps not clear whether the results above were influenced by subjects attempting to make sense of an unlikely set of visual cues or that demand characteristics of the design also influenced slant judgements. Nevertheless, the findings suggest that modal surface spreading involves some kind of confirmatory mechanism (perhaps similar to Grossberg's Object Recognition System). The overall configuration of the stimuli appeared to mediate interpretation of local retinal disparity values, particularly for the modal SKS-L percepts.

Recall that the images created were a product of a 2-D layout of image differences modelled on perspective projection. The configuration effects in this experiment suggest that a set of image difference measures supplementary to conventional point disparities can impact 3-D perceptual organisation. This argument is summarised in Fig 6.12.

In terms of the BIPASS model, the kind of pre-visual connections proposed by Grossberg are not necessary for the induction of surface separation and this is in agreement with Anderson's arguments about illusory contour induction. Perhaps a higher confirmatory process that reconciles retinal differences at different feature scales influences the 3-D trajectory of surface spreading.

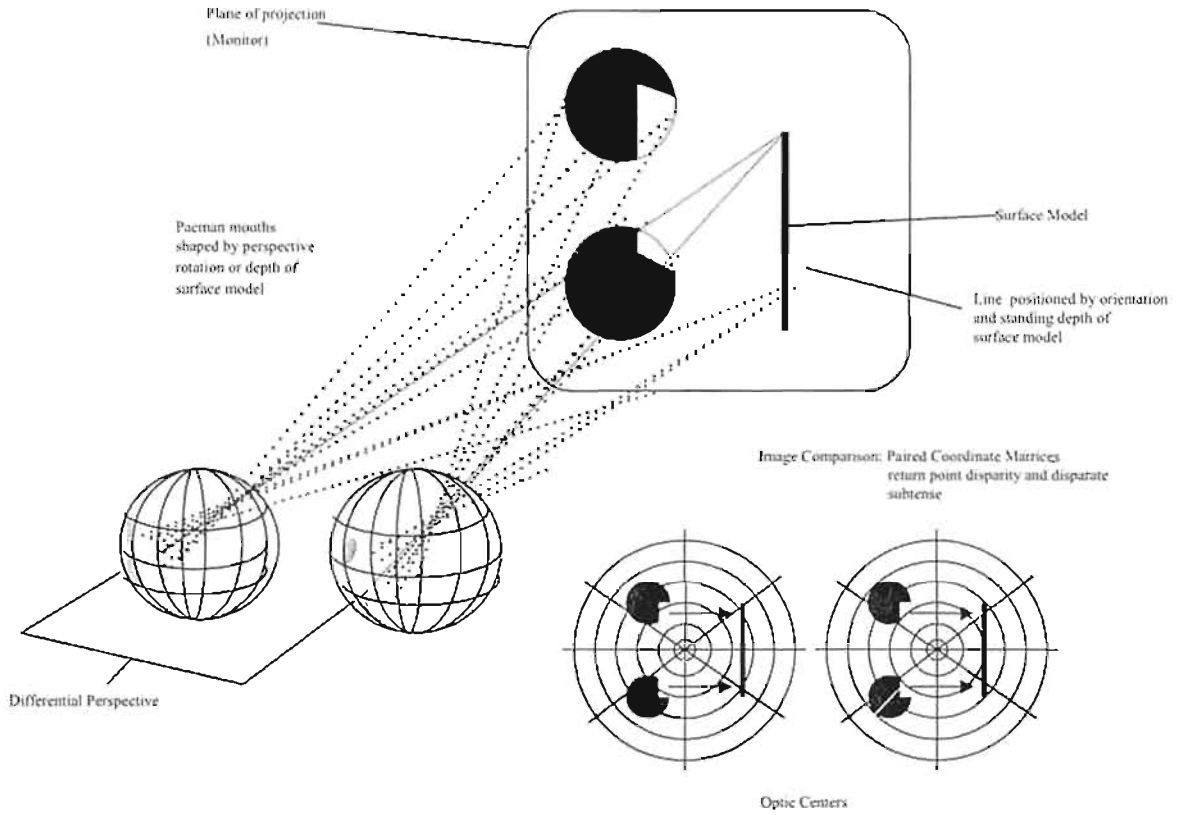



**Fig 6.12. 2-D layout of image differences and image comparison in the SKS-L**
Image comparison appears to be responsive to both the point disparity at contours and disparate subtense. Disparate subtense can occur in two respects: as a difference in magnitude of luminance contours or as the magnitude of untextured "space" between contours. The system seems responsive to both feature scales of image differences.

For free fusers the stimuli presented in this experiment are shown in Fig 6.13. The question remains as to whether or not surface spreading that appears in these stereograms can emerge in the absence of a completion process.

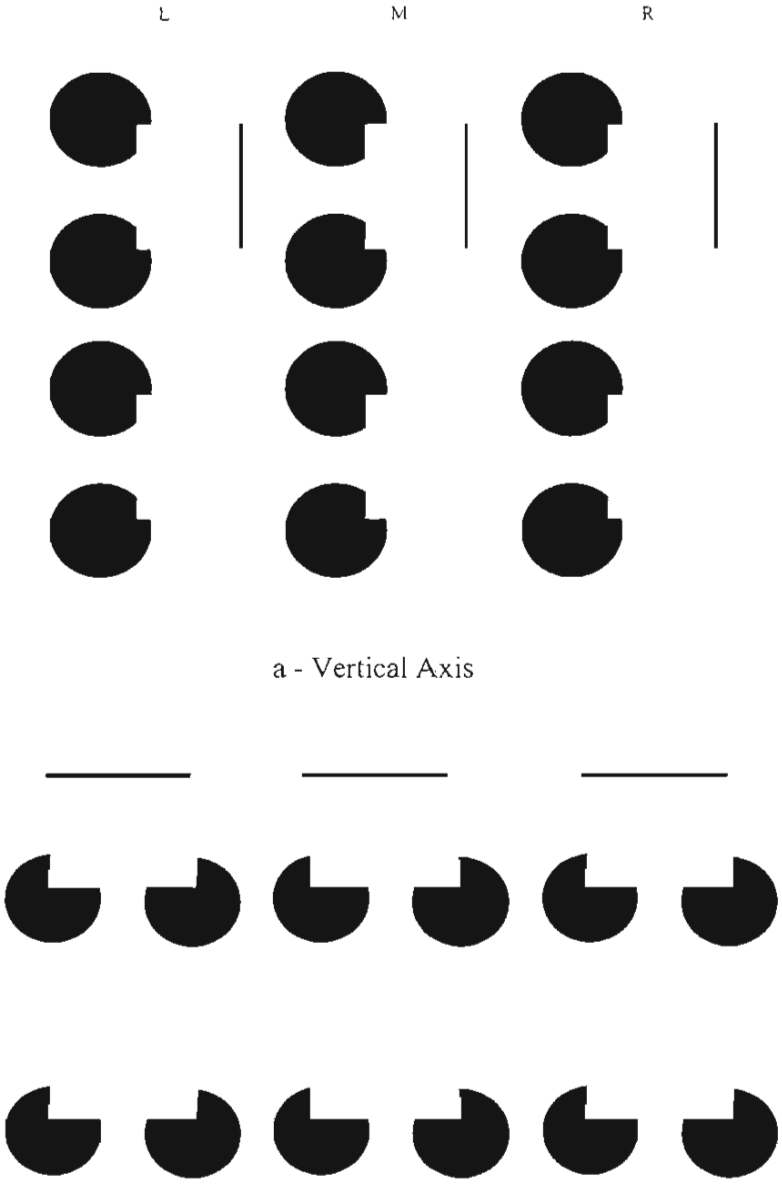



**Fig 6.13. Surface spreading in the SKS- and SKS-P configurations**
In (a) crossed fusion of the L-M pairs (top with bottom) compares the presence and absence of a line for vertical slant-axis with uncrossed disparity. Fusion of the M-R pairs (top and bottom) compares the same for crossed disparity. Fig (b) shows the same effects for horizontal slant-axis. As experiment 2 has shown, there is a difference between the perceptual manifestations of retinal disparity that means horizontal contours in vertical slant-axis are ambiguous.

The next experiment attempted to repeat these findings in a more direct manner and in a manner thought to reduce the likelihood of completion mechanisms being involved. The line

in the SKS-L configuration was replaced with a small dot to see if spreading would still emerge. Surface spreading alone might impact upon the interpretation of horizontal disparities.

## 6.2 Experiment 3 Surface spreading without completion?

Experiment 3 tested induction of surface spreading independently to contour continuation or completion. Figure 6.14 shows once again, that the surface layer separated from the depth plane of pacmen stands at a near depth plane (cross fuse L-M). The effect here is that the surface layer at the mouths of the disparate pacmen (B_C and E_F) appears to merge or fade out into the homogeneous space surrounding the pacmen. Note that the previous experiment showed subjects tended to see the separated surface layers as fronto-parallel when standing behind the P Plane but not when standing above the P Plane.

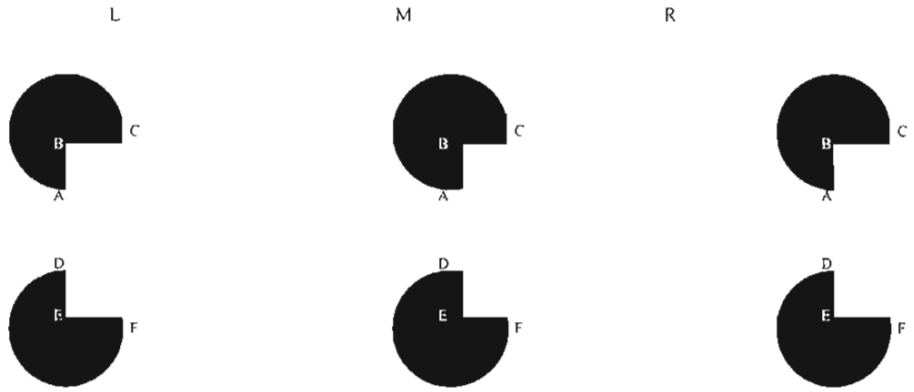



**Fig 6.14. Surface spreading**
Crossed fusion of the L-M pair sees a surface layer between the pacmen standing at a near depth plane. The layer appears to merge with the homogeneous surrounds from the mouths of the pacmen along the contours A_C and E_F. Fusion of the M-R pairs sees the porthole effect, where an illusory contour marks the depth step between a surface layer assigned to a distant depth plane (behind the portholes) across C_C and D_F.

Since introduction of a luminance and depth referent (a line) could influence the orientation of this apparently spreading surface layer I wondered if positioning a small circle adjacent to disparate pacmen would have an equivalent effect. Such a stimulus means that the luminance and depth referent set adjacent to stereoscopic pacmen does not generate any obvious completed form. Hence, the slant estimation technique was again used as a metric of the influence of stimulus configuration. In this case a small circle - a dot - was placed adjacent to the stereoscopic pacmen. This configuration will be called the stereoscopic Kanizsa square - dot (SKS-D). Standing disparity was manipulated to see if this referent might constrain the

interpretation of disparate horizontal contours at the mouths of the pacmen relative to the SKS-P configuration.

The actual position of the dot in each half-image was defined by the same perspective corrected model of relative magnification applied to the SKS configuration in Experiment 1 and to the SKS-L, in Experiment 2. To do this the dot was literally *attached* to the edge of an overlaid square. The square was of homogeneous luminance with the white background. The dot was positioned at the midpoint of the edge of the overlaid square adjacent to the pacmen (see Fig 6.15).

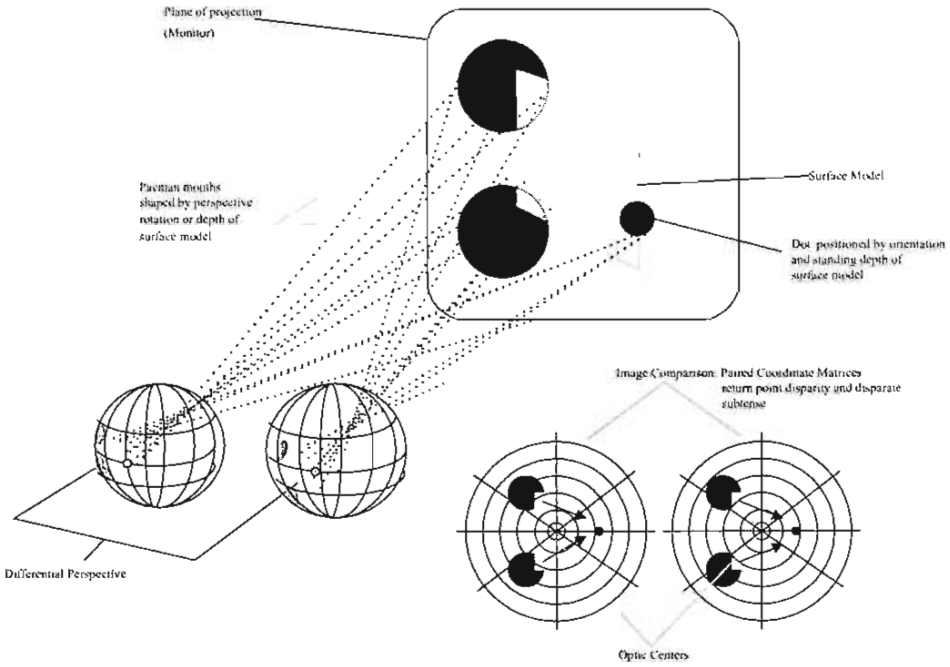



**Fig 6.15. 2-D layout of image differences and image comparison in the SKS-D**

Image comparison appears to be responsive to both the point disparity at contours and disparate subtense. Disparate subtense can occur in two respects: as a difference in magnitude of luminance contours or as the magnitude of untextured "space" between contours.

### *Measures: seen slant in the SKS-P and SKS-D percepts*

Manipulation of disparity in this experiment simulated rotation of a square in the vertical slant-axis only (see Fig 6.16). This created pacmen whose mouths were formed by horizontal magnification only. Disparate pacmen were achieved by manipulating the relative magnitude of the square describing BDFHB in each eye's view. Hence, manipulating the square positioned the small dot to precisely the same extent as adjacent pacmen in Experiment 1, and the adjacent line in Experiment 2.

Standing disparity was also introduced to the square BDFHB (see Fig 6.14). In manipulating standing disparity, the square BDFHB was shifted horizontally relative to that in the left eye's view. Again, positioning the dot. In order to generate the perception of a slanted modal or amodal surface layer the system must interpret the disparate magnitude of horizontal contours at BC and HG in terms of the (horizontal) magnitude of BE and HE.

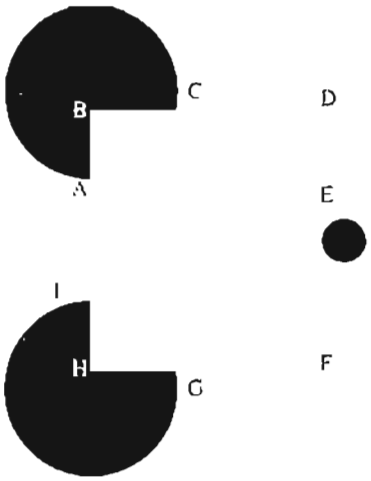



**Fig 6.16. Separation of surface layers and *spreading* toward an adjacent dot**
Crossed fusion of the L-M pair sees a subtle surface layer spreading toward the dot opposite. Free fusion using small images seems to reduce this effect somewhat compared to the same effect on a large computer monitor. If one compares the top and bottom set of L-M pairs, the lower pair sees the mouths slanted toward the dot. This is because no disparity is actually available in this figure that defines orientation of the mouths.

A fully factorial repeated measures design compared the impact upon seen slant of three independent variables: Theoretical Rotation (20, 40$^0$), Standing Disparity (-20, 20$^0$) and Configuration (the SKS-P, SKS-D comparison). It was anticipated seen slant would be evident in conditions where the dot was present in the SKS-D configuration, but in the absence of this depth referent (SKS-P configuration), seen slant would tend toward zero. A significant effect of theoretical rotation was also expected. Moreover, it was predicted that seen slant would be greater when the mouth of the pacmen stood in front of the P Plane than when standing disparity stood the mouth of the pacmen behind the projection plane.

### 6.2.1 Method

*Subjects*

Eight available subjects were drawn from the department pool of volunteers.

*Stereograms*

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a pair of black circles subtending $3^0$ in diameter (at 750mm viewing distance). The circles were positioned so that a square drawn through their centres would subtend $7^0$ degrees. *Pacmen* were created by drawing a square, equiluminant with the background, so that it partly obscured the circles. The square symmetrically overlaid the circles (equal intrusion on all sides). The square intruded about 1/3 the circles' radii creating SKS pacmen.

Mouths of the pacmen were geometrically transformed according to the monocular transformations of a square shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was then introduced by applying ½ Ogle's M to the overlaid square, symmetrically and in opposite signs (see Fig 6.17). Standing disparity was generated by shifting the overlaid square in equal and opposite directions in each half-image (see Fig 6.17). A six arcmin black circle was drawn at the boundary of the transformed overlaid square adjacent to the pacmen. Its position in each half-image was defined automatically by its relationship to the overlaid square.

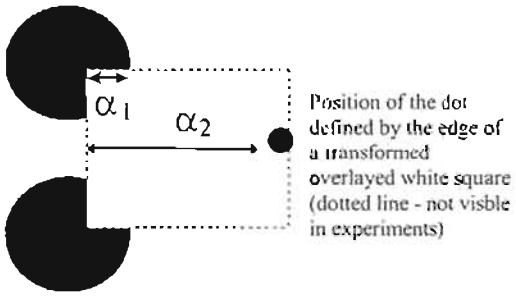



Fig 6.17. Patterns of disparity in the SKS-D stimuli
This figure shows one half-image supporting vertical slant-axis. Disparity appropriate to vertical slant-axis was applied by increasing the magnitude of the Kanizsa dot-square ($\alpha_2$) in one eye relative to the other. This changes the relative magnitudes of the mouths of the pacmen ($\alpha_1$). Standing disparity was then manipulated by constraining ($\alpha_2$) while shifting the dot-square in one eye relative to the other. This manipulates the relative magnitude of ($\alpha_1$).

Image pairs were presented on alternate frames at a rate of about 200 frames per second. Through the shutter goggles, background luminance was 0.7 cd $m^{-2}$ with the black pacmen and the small circle 0.09cd $m^{-2}$. In half of the stimuli, a the luminance-depth referent was absent.

*Design and procedure*

A fully crossed (2x2x2) repeated measures design was used to establish the effect of theoretical rotation ($20^0$, $40^0$), standing disparity (+20arcmins [standing above the projection plane], -20arcmins [standing behind the projection plane]) and configuration (SKS-P / SKS-D) upon seen slant. Five complete repetitions were carried out, making a total of forty slant estimation scores per subject.

Subjects were once again asked to judge the orientation of the pacman mouths by manipulating the comparison stimulus, as before.

### 6.2.2 Results and discussion

A one-way repeated measures analysis of variance was used to examine the effect of configuration (SKS-P / SKS-D) on seen slant. The configuration presented was found to affect seen slant (see Fig 6.18). The effect was significant $F_{(1,7)} = 24.6$, $p < 0.001$. In the SKS-D presence of the dot referent impacted upon the interpretation of disparities. I contend that this occurred as a surface layer spread from the mouths of the pacmen toward the luminance and depth referent.

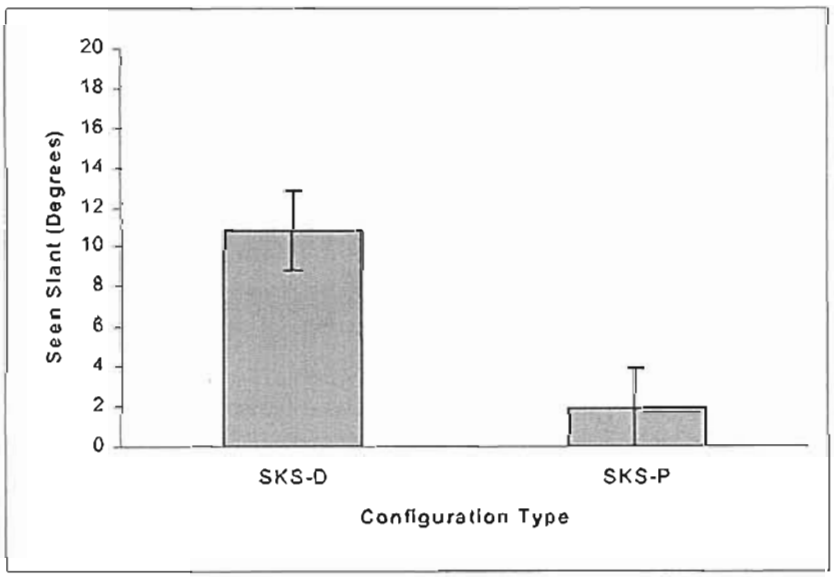



**Fig 6.18. The effect of configuration upon seen slant in SKS-D and SKS-P percepts**
This figure plots the mean seen slant estimates for each configuration SKS-D and SKS-P. Standard error bars have been included.

However, a one-way repeated measures analysis of variance examining the impact of theoretical rotation, that is, the relative horizontal magnitude of the overlaid homogeneous

square between the pacmen and the small circle, generated no significant effect on seen slant: $F_{(1,7)} = 0.0018$, $p > 0.9$.

In contrast to Experiment 2 Subjects tended to perceive the partly occluding surface layer as fronto-parallel in the SKS-P. This may have been due to the reduced magnitude of disparity at the mouths of the pacmen in Experiment 3. The orientation of the layer was, however, modulated by the presence of a single dot, in no obvious sense at least, part of a 'closed' form. Moreover, the finding that theoretical rotation had no impact upon seen slant is important. The presence of the referent dot and not theoretical rotation (local retinal disparities) influenced seen slant. This is not actually surprising since, in the absence of an adjacent referent there is no physical measure upon which to differentiate standing disparities from relative horizontal magnification in these horizontally aligned contours. That is, arguably, the nature of the ambiguity of horizontal contours in these abstract pacmen.

A one-way repeated measures ANOVA was also conducted to examine the effect of standing disparity upon seen slant in this configuration. Once again, that effect was significant; $F_{(1,7)} = 8.087$, $p < 0.05$ as shown in Fig 6.19. Recall that the sign of standing disparity dictates the phenomenological asymmetry between the porthole (-20 arc mins) and the separation of a near surface layer (+20 arc mins). The finding that possible amodal connections are weak between pacmen and adjacent referent (in terms of the seen slant metric) is now repeated.

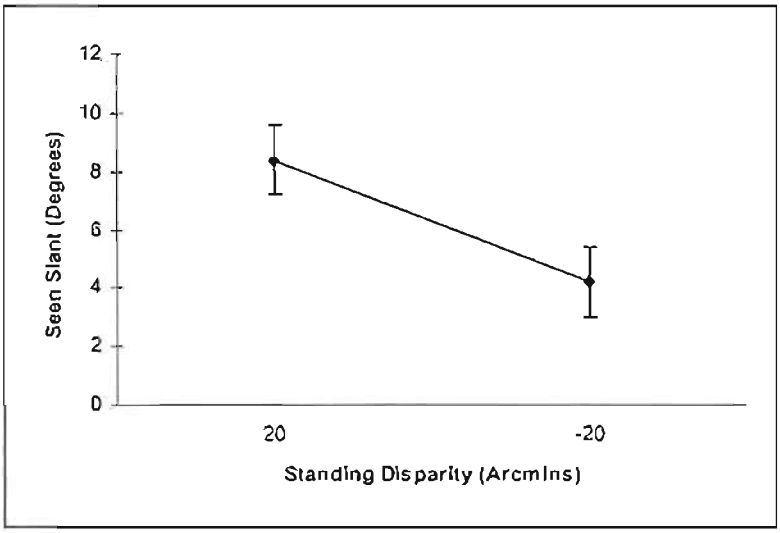



**Fig 6.19. Effect of standing disparity upon seen slant in SKS-P and SKS-D configurations**
This figure graphs seen slant means for each level of standing disparity presented in the SKS-P / SKS-D comparison. Standard error bars have been included.

The findings here are quite straightforward. There was a tendency for both a partly occluding (+ve sign of standing disparity) and a partly occluded (-ve sign of standing disparity) surface layer to appear fronto-parallel in the absence of a luminance and depth referent. This is in agreement with Experiment 2. Nonetheless, the presence of an adjacent luminance and depth referent - a dot - did impact upon the seen slant of a modal partly occluding (+ 20 arcmins standing disparity) surface layer.

To observe a spreading surface layers as looking oriented in depth, a disparity averaging mechanism may be in place. Grossberg's arguments that layers are filled from appropriate disparity pools by binocular FIDOs will need therefore to account for orientation. That is, the FIDOs need to be able to predict a gradient of disparity pooling. Not merely "near" and "far" pools.

For free fusers the stimuli used in this experiment are presented in Fig 6.20. The texture of the page seems to inhibit perception of surface spreading compared to the computer generated version of the configuration. In the simple experiments above, I used a polished monochrome monitor screen with luminance set to a fairly dim level. Crossed fusion of the pair at (a) may see the surface layer that is assigned to the near depth plane slanted toward the dot and fusion of the pair in (b) sees the surface layer merge with the surrounds. Subjects judged the later figure to generate a surface layer oriented at about fronto-parallel.

In summary, a surface layer appeared to spread toward the luminance and depth referent. But, only when standing disparity meant that fusion invoked a near surface layer standing above - partly occluding - the pacmen (+20 arc mins). The magnitude of seen slant was greatly attenuated. Indeed, theoretical differences in the magnitude of seen slant were not above significance at $\alpha = 0.05$.

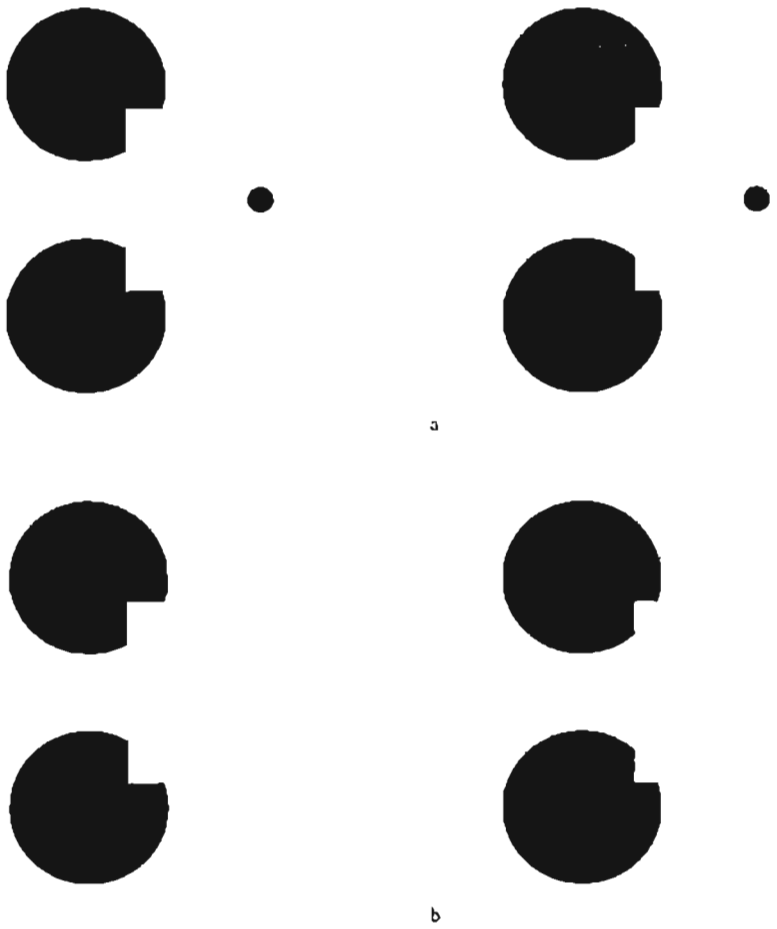


**Fig 6.20. Surface spreading in the SKS-D**
Experiment 3 found that a dot (a luminance and depth referent) set adjacent to disparate pacmen affected the perceived orientation of the Surface layer separated at the mouths of the pacmen (see Fig a). When the pairs in (a) are cross fused a surface layer may appear to spread toward the dot (can be compared to the SKS-P in (b)).

## 6.3 Experiment 4 Slant-axis anisotropy and "image feature disparity scale" in part modal / amodal Kanizsa figures

Experiment 4 directly compares seen slant about the two slant-axes using Kanizsa configurations of different shapes. This was to study the impact on seen slant of ambiguity inherent in the separation and orientation of illusory surface layers in the vertical slant-axis. Ambiguity of orientation seems to occur when part of the surface layer separated from the pacmen is defined by horizontal contours. This potentially means that results established previously for the vertical slant-axis were at least partly a product of a conflict in the image feature scale used to interpret image differences (including the slanted Kanizsa figures seen in Experiment 1).

Experiment 2 and Experiment 3 showed that perception of surface spreading in SKS-L and SKS-D configurations involved a possible conflict between interpretations of image differences at two levels of scale. Since the physical measure of image differences are identical, "levels of scale" here refers to the relative size of disparate features (spaces between pacmen and referent) and not to the actual scale (magnitude) of local disparities. The finding that seen slant was dramatically attenuated in the SKS-L and SKS-D configurations arguably underscores the ambiguous nature of image differences underpinning seen slant in the vertical slant-axis (for a square shape).

The fact that slant was seen at all is evidence for the presence of some confirmatory component of surface spreading that in a sense *binds* the adjacent white segments at the mouths of the pacmen in the SKS. In other words, where a conflict between levels of scale exists some confirmatory process appears to drive the interpretation of that disparity.

Manipulations in Experiment 2 and Experiment 3 appear to have reduced the impact of a luminance and depth referent stepwise. This has proportionally increased the relative weight of physical disparity ambiguities upon the global metric of the percept - seen slant. Moreover, the effect was particularly strong in the vertical slant-axis.

My argument is that horizontal contour ambiguities can represent, perceptually, a conflict between scales of disparity interpretation. This seems to mean that, all else being equal (viewing distance, vergence, interocular distance, theoretical rotation), seen slant in the vertical slant-axis will be attenuated relative to seen slant in the horizontal slant-axis. This is because there is less conflict, between the possible scale of disparity interpretations in the horizontal slant-axis (for a square shape).

There are many studies, using seen slant as a metric, showing that the context of retinal disparity can impact upon slant judgements. Diagonal, horizontal and vertical markings across a line grid modulate the slant response (Cagenello and Rogers, 1993; Gillam, 1968; Stevens and Brooks, 1988). Seen slant about the vertical tends to be attenuated more than slants about the horizontal axis (Gillam and Ryan, 1992). The difference is possibly due to the manner in which the vertical slant-axis is manifest using horizontal inter-retinal differences alone.

The difference in seen slant about these two arbitrary axes, in pseudo-randomly dotted line grids, has been called slant-axis anisotropy. Slant-axis anisotropy has been attributed to two possible causes. Cagenello and Rogers (1993) said that orientation disparity (relative

differences in the orientation of matching contours between the eye's views) determines sensitivity to slant. Hence, slant-axis anisotropy arises because there is asymmetry in the manifestation of orientation disparities between the axes.

Gillam has argued that slant-axis anisotropy is at least partly due to an asymmetric resistance to perspective conflicts. And, Ryan and Gillam (1994) have shown that slant-axis anisotropy and slant attenuation were substantially reduced when "perspective cues" were not in conflict. Gillam and her colleagues have not defined perspective cues in stereoscopic terms but in monocular terms. So we will not use the term perspective cue here.

Experiment 1 showed that a slant-axis anisotropy existed where seen slant in the horizontal slant-axis tended to be significantly greater than the vertical slant-axis. Experiments 2 and 3 showed that seen slant in the vertical slant-axis required the system to somehow override disparities. Rotations about the vertical axis in the SKS may therefore require interpretation. In the horizontal slant-axis disparity in the contours defining orientation are largely congruent with disparate subtense across the homogenous spaces.

To test the logic of these arguments a simple stimulus was developed in which the components of physical retinal disparity in a Kanizsa configuration were thought to be very similar for either arbitrary slant-axis - vertical and horizontal. A stereoscopic Kanizsa diamond (SKD) was constructed. In the Kanizsa diamond, Ogle's relative horizontal magnification (vertical slant-axis) and horizontal shear (horizontal slant-axis) yield a pattern of disparity in which the components of seen slant at the mouth of each pacman and between pacmen are congruent. That is, disparities at and between pacmen are not in conflict, or at least, any residual conflict will be equivalent for both slant-axes.

Figure 6.21 shows that half-images of a diamond configuration can be defined in precisely the same way as previously discussed for the square. The diagram shows a view of the rectilinear projection geometry that shapes the retinal images of a diamond. In this case the diamond is rotated about the vertical slant-axis.

The dimensions of the two retinal images of a diamond are predicted geometrically; by the dimensions of the 2-D diamond, its angle of rotation, the system's interocular distance and the viewing distance. Figuratively speaking, to create stereoscopic half-images we take a horizontal-vertical slice through the perspective view of a rotated diamond at the distance of the plane of projection - the computer display. The formal transformations used to achieve the half-images are presented in Appendix A.

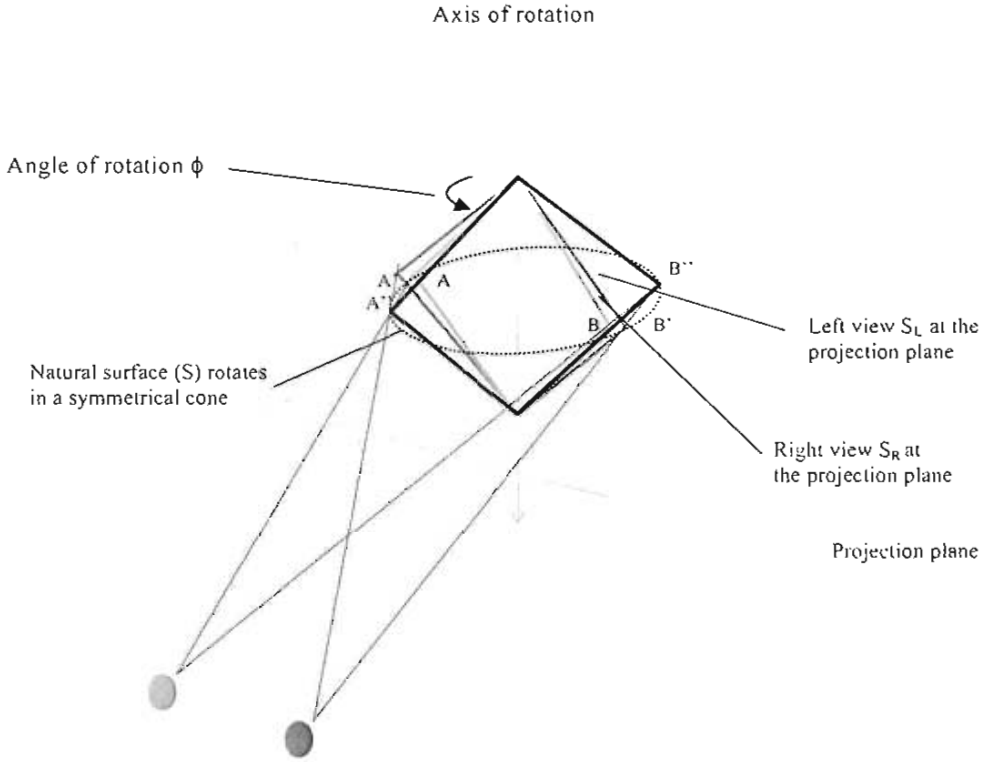



**Fig 6.21. Projection geometry underpinning 2-D layout of image differences captured at the retina in the case of a rotating planar diamond shaped surface**
A diamond shaped surface (S) rotates symmetrically through a conical prism. For any angle of rotation ($\phi$), the images at each retina will prescribe a geometric retinal disparity. This 2-D layout of image differences can be simulated by generating the images that project to the eye's from the distance of a plane of projection (eomputer monitor). At that distance, the rotated diamond shape A''B''C''D'', projects the images ABCD in the left eye's view ($S_L$) and A'B'C'D' in the right eye's view ($S_R$). Horizontal slant-axis can be generated similarly by applying horizontal shear to the retinal images.

The actual shape defined by the transformation of a diamond was achieved by attaching the four corners of the diamond to a square of the same vertical and horizontal dimensions. Hence we could draw a square shape and a diamond shape in which the horizontal and vertical and vertical dimensions in each half image derived from precisely the same transformations. Possible monocular perspective cues were identical.

If a diamond shape the same luminance as the background is drawn to partly obscure four pacmen, the shape of the diamond can be transformed as shown to create stereo half-images. The central prediction in this experiment was that the way that the pattern of disparity manifest in a SKS and a SKD diamond defined by exactly the same transformation of horizontal magnitude or shear will have different perceptual consequences. These *asymmetries* are shown in Fig 6.22.

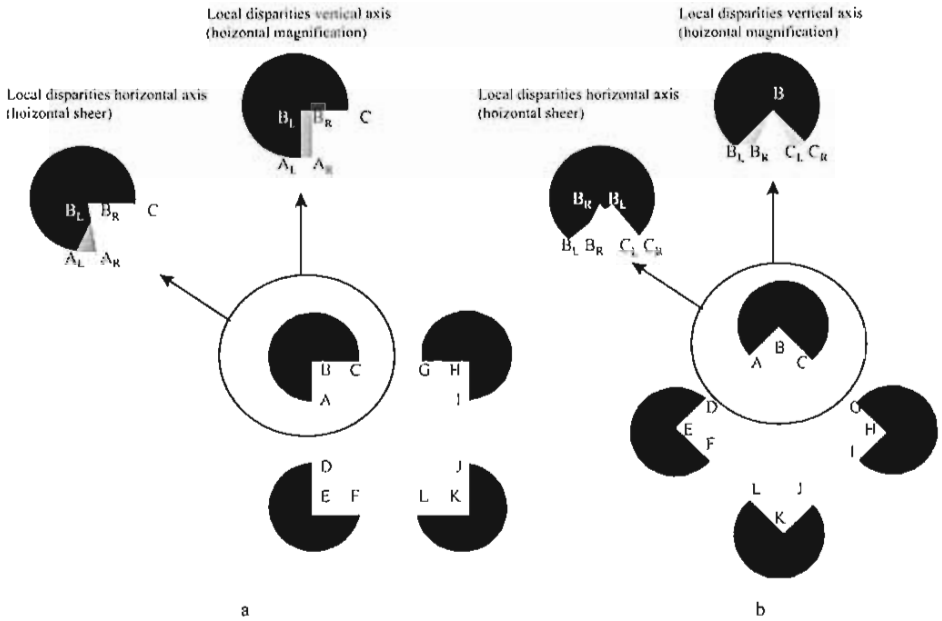



**Fig 6.22. Shape of disparities in two stereoscopic Kanizsa figures.**
This figure shows the differences in disparities created when a diamond and square are transformed then drawn to obscure four black circles. The grey segments are meant to show the bounds of the right eye's retinal image of the pacman. But, the right image has simply been draw to overlay the left - so the left view is partly obscured. Each shape is transformed by the same magnitude in the same direction (horizontal magnification or shear). In (a) relative horizontal magnification of an overlaid square creates disparities in the pacman that includes horizontal contours at $B_L$_C - $B_R$_C. For the horizontal axis, disparity exists in the disparate shear along $B_L$_ $A_L$ and $B_R$_ $A_R$. The horizontal contour at $B_L$_C - $B_R$_C has no bearing on orientation. In (b) for both axes of rotation orientation disparities exist at all contours of the mouths of the pacmen.

*Measures: seen slant in the SKS and SKD percepts*

Experiment 4 tested the argument that a slant-axis anisotropy in the SKS was partly due to the manner in which disparities in the SKS were physically manifest. Disparities in the vertical slant-axis are, theoretically, ambiguous and may conflict against disparate subtense between pacmen. Hence, in a SKD in which disparities that are, theoretically, not ambiguous, no slant-axis anisotropy should result.

This experiment employed a fully factorial repeated measures design to examine seen slant across theoretical rotation in both the horizontal and vertical slant-axes in for the SKS and SKD configurations. It was anticipated that no slant-axis anisotropy would be evident in the Kanizsa diamond. Moreover, it was expected that seen slant in the vertical slant-axis for a SKD would be greater than seen slant in the vertical slant-axis of a SKS. The significance of these predictions is that they will further illuminate the relationship between retinal disparities

in Kanizsa figures. Planned comparisons were prepared to contrast the two configurations at each slant-axis in the first –order interaction between main effects.

### 6.3.1 Method

*Subjects*

Six subjects were drawn from the pool of volunteers for the purposes of this experiment. These subjects were drawn from volunteers who had partaken in previous experiments. They were considered to be experienced in manipulation of my experimental instrumentation.

*Stereograms*

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a set of four black circles, each subtending $3^0$ in diameter (at 750mm viewing distance).

**SKS**: Four black circles were positioned so that a square drawn through their centres would subtend $7^0$ degrees. *Pacmen* were created by drawing a square, equiluminant with the background, so that it symmetrically partly obscured the circles. The square, prior to transformation, intruded ½ the circles' radii creating SKS pacmen.

Mouths of the pacmen were geometrically transformed according to the monocular transformations of a square shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was then introduced by applying ½ Ogle's M to the overlaid square, symmetrically and in opposite signs (see Fig 6.23a and b).

**SKD:** The four circles were positioned so that a diamond, based on the same dimensions as the square above, drawn through their centres, would subtend $7^0$ degrees. The mouths of the pacmen were geometrically transformed by drawing a drawing a diamond, equiluminant with the background, so that it partly obscured the circles. The diamond shape was created by attaching its corners to an identical square to that above. The diamond symmetrically overlaid the circles (equal intrusion on all sides). The diamond intruded ½ the circles' radii creating SKS pacmen.

Mouths of the pacmen were therefore geometrically transformed according to the monocular transformations of a diamond shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was introduced by applying ½ Ogle's M or relative shear to the overlaid diamond, symmetrically and in opposite signs (see Fig 6.23c and d).

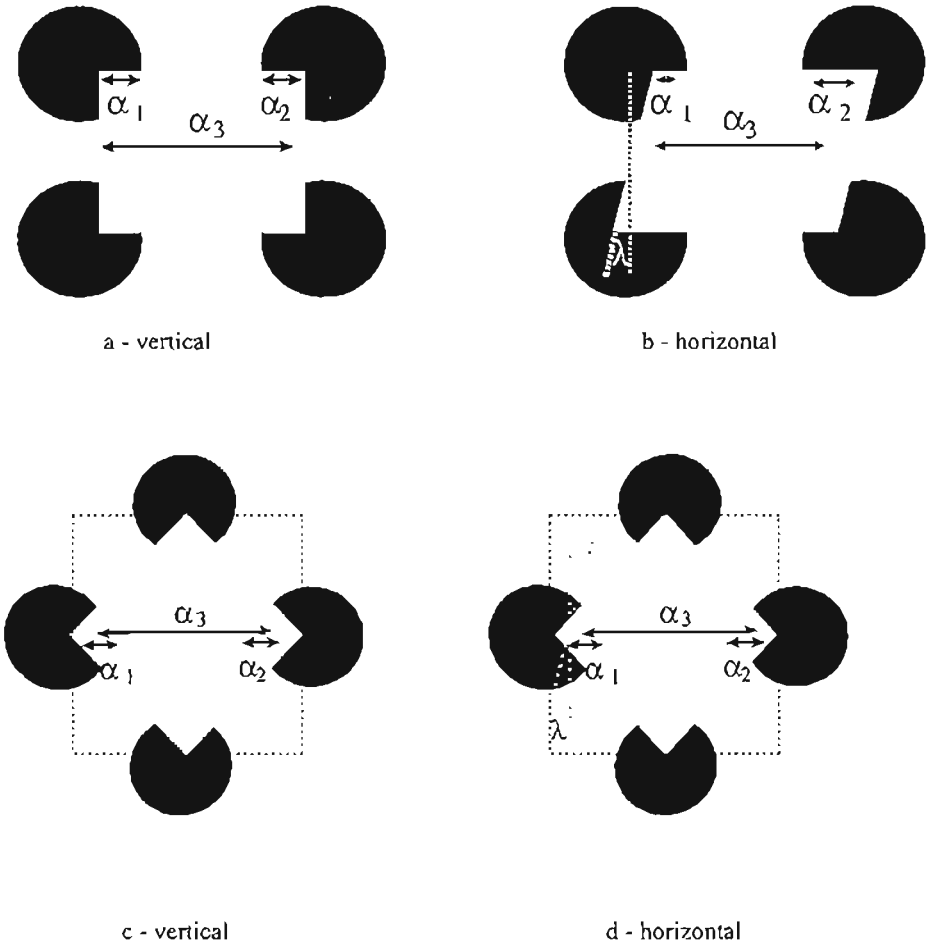



**Fig 6.23. Patterns of disparity in the SKS and SKD**

Fig 6.23a shows one half-image supporting vertical slant-axis. Disparity appropriate to vertical slant-axis was applied by increasing the magnitude of the SKS ($\alpha_3$) in one eye relative to the other. This obviously changes the relative magnitudes of the mouths of the pacmen ($\alpha_1$) and ($\alpha_2$). Fig (b) shows the horizontal differences applicable to horizontal slant-axis. To manipulate rotation, disparate shear ($\lambda$) was applied in eye while constraining ($\alpha_3$). Then standing disparity was introduced by again constraining ($\alpha_3$) while shifting the illusory figure in one eye relative to the other so adjusting the relative magnitudes of ($\alpha_1$) and ($\alpha_2$) in each eye. Fig 6.21c shows one half-image supporting vertical slant-axis for a diamond. Disparity appropriate to vertical slant-axis was applied by increasing the magnitude of the diamond ($\alpha_3$) in one eye relative to the other. This changes the relative magnitudes of the mouths of the pacmen ($\alpha_1$) and ($\alpha_2$). Fig (d) shows the horizontal differences applicable to horizontal slant-axis. To manipulate rotation, disparate shear ($\lambda$) was applied in each eye while constraining ($\alpha_3$).

Image pairs were presented on alternate frames at a rate of about 200 frames per second. Through the shutter goggles, background luminance was 0.7 cd $m^{-2}$ with the black pacmen 0.09cd $m^{-2}$.

*Design and procedure*

The study used a fully crossed, three-way (2x2x4) repeated measures design to explore the effects of configuration (SKS, SKD), slant-axis (horizontal/vertical) and theoretical rotation (20, 30, 40, $50^0$), respectively, on estimated slant of a Kanizsa configuration. Three complete repetitions of this design were used, making 64 trials per subject in all. Details of the procedure were as described in section 4.5. Subjects were asked to set the rotation of the comparison stimulus to match the apparent rotation of an illusory plane.

### 6.3.2 Results and discussion

A three-way (2x2x4) repeated measures analysis of variance examined the effects of configuration (SKS / SKD), slant-axis (horizontal / vertical) and theoretical rotation (20, 30, 40, $50^0$), respectively, on seen slant of a Kanizsa configuration. Results were averaged across the repetitions.

Theoretical rotation was found to impact predictably upon seen slant and the effect was significant: $F(3,5) = 336.97$, $p < 0.001$. Figure. 6.24 shows the mean slant estimates for theoretical rotation. Slant-axis was also found to affect seen slant. A slant-axis anisotropy of about $4^0$ was found and that effect was also significant: $F(1,5) = 292.547$, $p < 0.041$.

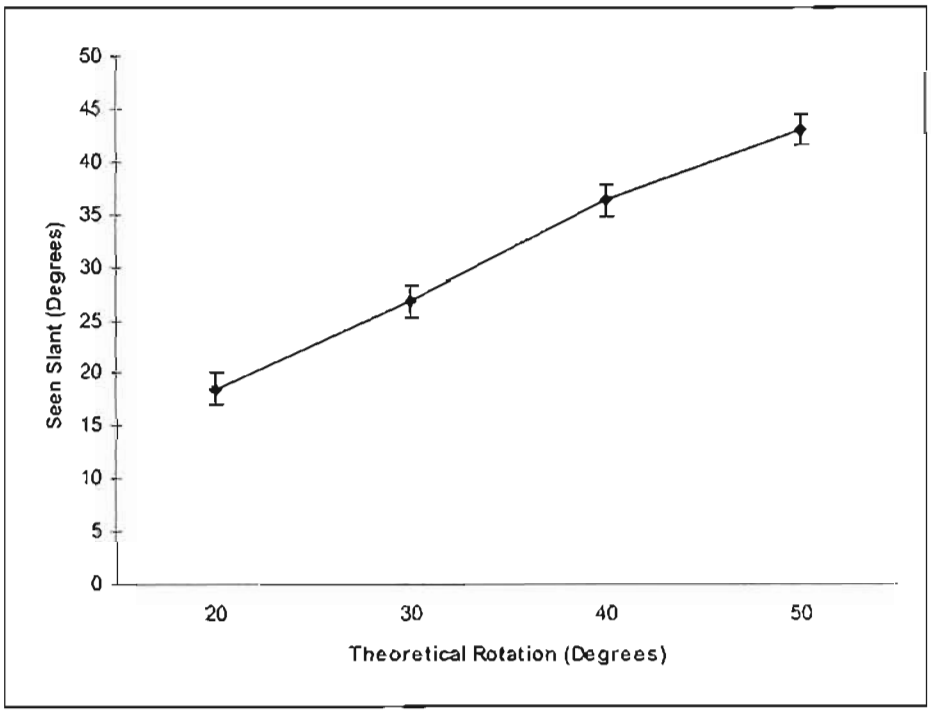



Fig. 6.24. The effect of theoretical rotation upon seen slant in the SKS and SKD configurations

This figure graphs mean slant estimates for the main effect of theoretical rotation upon seen slant. Standard error bars have included.

The shape of the Kanizsa configuration (SKS / SKD) was also found to effect seen slant. And, that effect was significant: $F(1,5) = 14.6$, $p = 0.01$. Figure 6.25 shows this comparison. A small difference between these global means of about $2^0$ was evident.

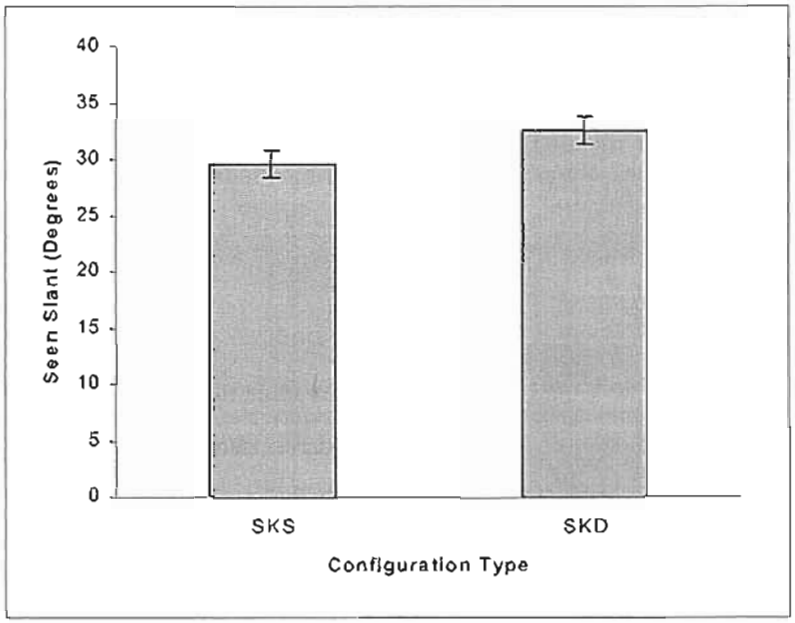



Fig.6.25. The effect of Configuration upon seen slant in stereoscopic Kanizsa figures
This figure graphs mean slant estimates for the main effect of configuration (SKS / SKD). Standard error bars have been included.

Two interactions, between key factors of interest describe differences between the illusory shapes. Firstly, an interaction between configuration and slant-axis affected seen slant as depicted in Fig. 6.26. That interaction was significant: $F(1,5) = 10.467$, $p < 0.05$. Planned comparisons revealed that the difference between means at each slant-axis was significant for the SKS: $F(1, 4) = 17.862$, $p < 0.01$ attesting to the slant-axis anisotropy seen in Fig 6.26 for the SKS configuration only.

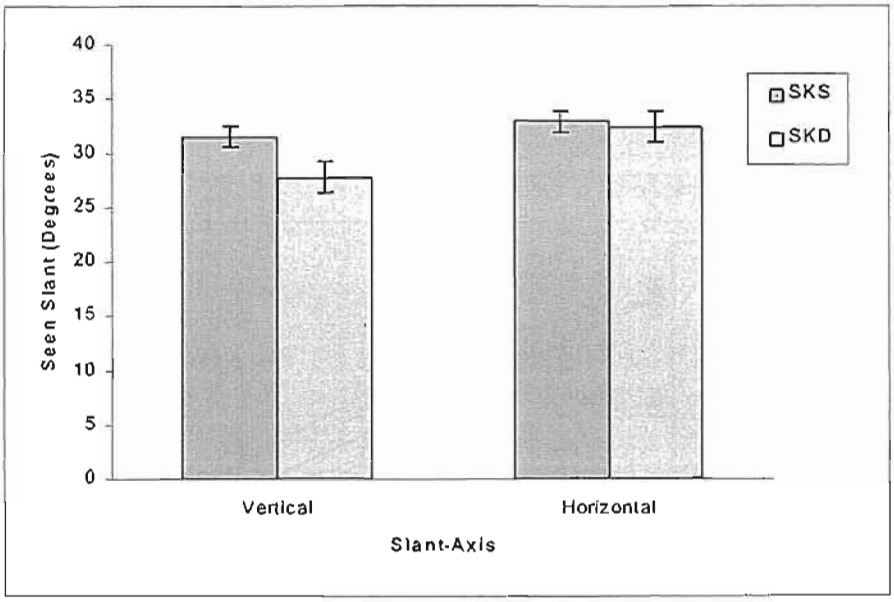



Fig 6.26. An interaction shape and slant-axis in two stereoscopic Kanizsa figures
This figure plots the first order interaction between slant-axis and configuration (SKS / SKD) upon seen slant. Standard error bars have been included.

Figure 6.27 graphs means for the three-way interaction between shape, slant-axis and theoretical rotation. A difference between the shapes occurs at each degree of rotation about the vertical slant-axis. The three way interaction between shape, slant-axis and theoretical rotation was found to effect seen slant. The interaction was also significant: $F(3,6) = 11.495$, $p < 0.001$. However, the impact of shape was clearly evident only at the larger theoretical rotations. Planned comparisons revealed a slant-axis anisotropy for the square at $40^0$: $F(1,4) = 11.052$, $p < 0.01$ and at $50^0$ degrees $F(1, 4) = 87.20$, $p < 0.001$. These findings support the hypotheses:

a. that no slant-axis anisotropy would be evident in the SKD
b. that seen slant in the vertical slant-axis for the SKD would be greater than seen slant in the vertical slant-axis in the SKS.

In Fig 6.27 the means are remarkably close to the theoretical rotation for both shapes. However, slant attenuation, particularly at the larger magnitudes of theoretical rotation is still evident and seen slant for square shapes about the vertical slant-axis were most attenuated. The SKD shape yielded seen slant very near to the theoretical - at each rotation - about each slant-axis. However, only at thirty and forty degrees was a slant-shape anisotropy in seen slant

evident for the horizontal slant-axis. The effects observed, appear to be greater at greater theoretical rotation. Planned contrasts showed that slant-axis anisotropy occurred in the square shape at forty and fifty degrees theoretical rotation only.

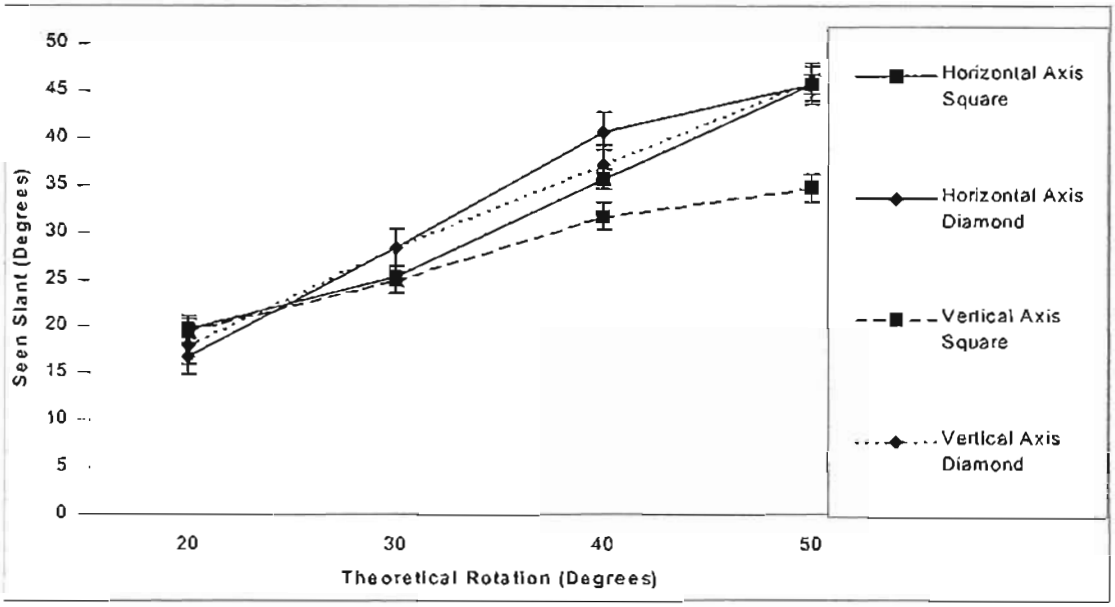



**Fig 6.27. Effect of theoretical rotation x shape x slant-axis upon seen slant in the SKS and SKD configurations**

The figure plots mean slant estimates at four levels of theoretical rotation for each slant-axis and for each shape. Standard error bars have been included.

In summary, slant estimates associated with the SKD configuration suggest that when geometric asymmetries between slant-axes are minimal, slant-axis anisotropy and seen slant attenuation, even in the 3-D illusory percepts, are also minimal. This suggests that the system seems to be vulnerable to a conflict between the scales of retinal features underpinning disparity measures specific to horizontal contours underpinning stereoscopic rotation about the vertical axis.

Vulnerability of the binocular vision system to horizontal contours in the SKS percepts and the seemingly precise interpretation of SKD percepts reveals some important characteristics of disparity processing in the 3-D illusory SKS percepts. It suggests that image comparison mechanisms contemporaneously return disparities at multiple scales. Moreover, disparity processing in the SKS percepts is not a one off computation but a dynamic and

relative process. Some kind of confirmatory processing seems to modulate the aggregate image difference data term.

Seen slant appears to be an appropriate metric of the relative impact of physically separated retinal image features in stereopsis. Given the theoretical ambiguity of a square shape (in manifestation of retinal disparity in the vertical slant-axis), there is a clear influence of a higher confirmatory or completion processes in the estimate of seen slant of an untextured surface. Moreover, the three experiments in this chapter have demonstrated that it is possible to systematically manipulate interaction between the disparities at the pacmen and between the pacmen in stereopsis.

These findings support the argument that an intimate relationship exists between the SKS percepts and 2-D layout of image differences projected to the retina. The BIPASS model attempted to demonstrate mechanisms that underpin the percepts. It appears that another step is required in the BIPASS model relating to confirmatory processing such as object recognition (Grossberg) or smoothing constraints (Kellman and Shipley) that impact the precise trajectory of illusory contours and disparity interpretation. Fig 6.28 summarises the modification of the BIPASS model.

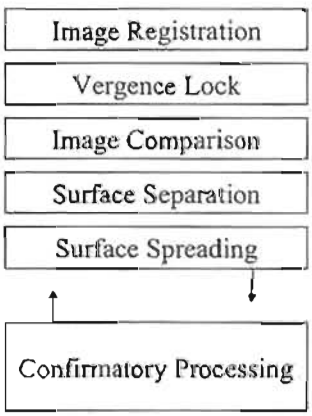



**Fig 6.28. A revised BIPASS model of the SKS percepts**

For free fusers examples of the stimuli in question are provided in Fig 6.29. Crossed fusion of the L-M or M-R pairs will yield a SKD or SKS oriented in the vertical slant-axis.

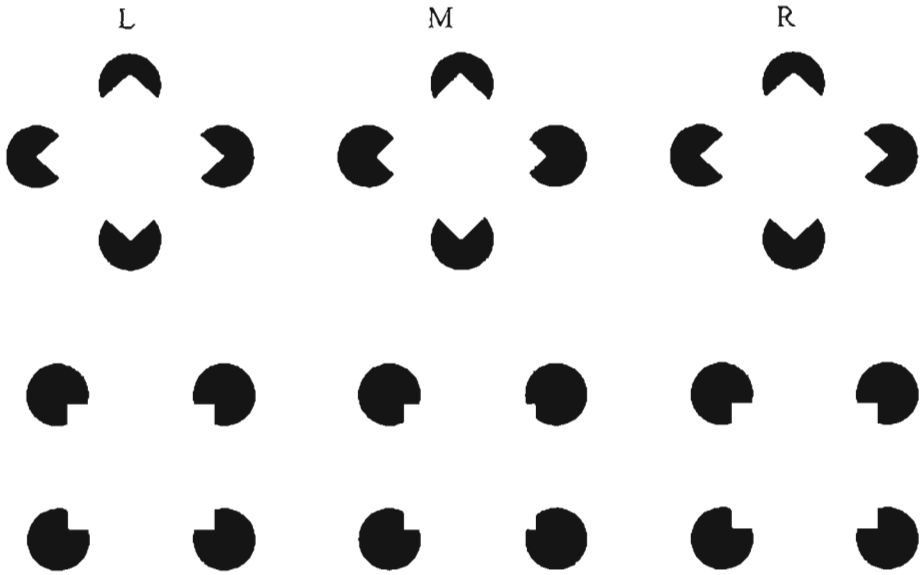



Fig 6.29. Perception of slanted part modal / amodal Kanizsa squares and diamonds

Finally, the phenomenological properties of these 3-D illusory percepts, particularly the diamond shapes, are quite remarkable. The SKDs look very stable. Yet they appear to literally pass *through* the projection plane (the page) when they are fused (see Fig 6.30). Note that the projection plane is the zero-disparity (or zero-crossing) point in terms of the sign of disparity. As such it demarcates modulation between a *modal* and *amodal* looking parts of the SKD.

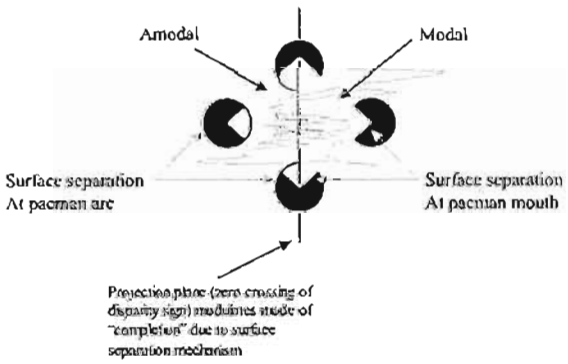



Fig 6.30. Phenomenological properties of a stereoscopic Kanizsa diamond

The BIPASS model explains these remarkable percepts in terms of the quantitative separation of surface layers at specifiable contours. An important aspect of these percepts appears to be the image difference measures available across the two retina constraints upon image

comparison. Some kind of confirmatory processing such as a completion mechanism but includes integration of disparity values (that may conflict) appears to impact upon the perceived orientation of separated surface layers in these configurations.

# 7. Surface separation in the absence of typical point disparities

**Summary:** This chapter examines 3-D illusory SKS percepts formed in the absence of conventional disparity. Two experiments used seen slant to explore the stereoscopic response to these stimuli. It is concluded that the percepts can be explained in the terms of binocular image processing mechanisms underpinning surface separation and surface spreading.

A case has been made that the stereoscopic response to the SKS is two fold. It involves response to conventional point-disparities at luminance contours and disparate subtense across large-scale parts of the retinal images (such as pacmen and spaces between them). In natural vision, these two aspects of disparity are immutably related. That is because they are direct products of the projection of light from objects in the world onto each retina - binocular parallax. The phenomenological properties of the SKS percepts appear to reflect the manner in which binocular vision processes those disparities; but in a highly unusual context. A BIPASS model has been proposed that identifies likely functional processes.

Recall that the Surface Heuristic approach emphasises an inferential response (eg inverse ecological optics) triggered by particular local features (such as unpaired regions). On the other hand, the Form Computation approach emphasises integration of physically uncorrelated luminance contours. Grossberg's view is that integrative connections bind together "end-cut" mechanisms creating to a visual entity as a neural syncytium.

Both of these approaches propose means by which the system overcomes the lack of retinal point-disparity in untextured stereograms. In contrast, the BIPASS model suggests that the perceived separation of surface layers in the SKS percepts is partly attributable to *resolution* of disparate subtense constrained by vergence locking. The percepts derived from stimuli presented in this chapter seem to strongly support that view.

Experiments 5 and 6 examine evidence that inter-retinal differences in the SKS percepts need not include conventional retinal point disparity processing at all. Resolution of disparate subtense may be sufficient to generate surface separation. Surface separation in turn initiates surface spreading between depth and luminance referents such as adjacent pacmen.

## 7.1 Experiment 5 A Kanizsa square visible to stereopsis but with pacman mouths unpaired

This section examines a SKS in which the mouths of inducing pacmen are unpaired. No retinal point-disparity exists in such half-images (in the conventional understanding of interocular correspondence). For convenience, the configuration will be termed the stereoscopic Kanizsa square-unpaired (SKS-U). The next section demonstrates the SKS-U percept.

### 7.1.1 Phenomenology of a SKS-U percept

In Fig 7.1 an illusory surface is clearly visible (cross fuse L-M) despite the fact that the pacman mouths are unpaired[12]. A set of four portholes is perceived (cross fuse M-R) in the reverse sign of *disparity*, that is, where the eye of origin of the pacman mouth is reversed. An amodal white square seems partly visible behind the depth plane of the pacmen - through portholes. The difference between percepts at the two signs of "disparity" has been termed perceptual asymmetry (after Anderson and Julesz, 1995).

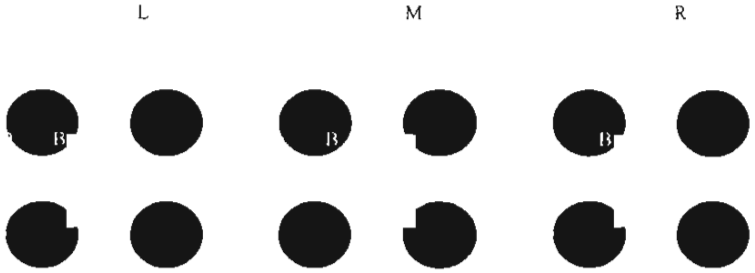



**Fig 7.1. An Unpaired SKS**
In half-images L, M and R pacmen have been paired with a full black circle. Crossed fusion of L-M yields the percept of a Kanizsa square standing forward of the projection plane. Crossed fusion of M-R yields the percept of an amodal square standing behind the four pacmen which now appear as portholes. Uncrossed fusion will reverse these effects. The labels ABCD identify corresponding regions of a pacman-eircle pair.

The literature in this area is presently dominated by the idea that faced with uncertain images the system gets by on the learned rules of monocular occlusion. Subsequently, there has been particular interest in the detection of unpaired regions in untextured stereograms.

---

[12] These half-images were formed by drawing four black circles, then drawing a white square homogenous with its surrounds to partly obscure the circles in the typical Kanizsa configuration. This image was then repeated to generate the three image pairs. In half-image L, the right hand pair of circles has been brought forward to overlay the square so that only the left pair of circles have "mouths". In M, the same process was used to generate mouths only in the right hand pacmen. The same process was applied to R, a replica of M.

### 7.1.2 Monocular occlusion zones and the SKS-U

The percepts achieved when free fusing the half-images in Fig 7.1 are theoretically interesting. Not only do they represent stereopsis in the absence of point-disparity, they also break, or at least *bend*, ecological geometric *rules* of monocular occlusion zones (see Chapter 3 for a detailed review). The significance of this, is that several theorists have based their arguments upon the ecological validity of monocular features at partial occlusions.

Howard and Rogers (1995) state four geometric principles of monocular occlusion zones; as they arise in natural vision. Figure 7.2, represents linear projection that typically creates monocular zones in stereopsis. The diagram depicts a near surface ($S_1$) partly occluding a distant one ($S_2$). A monocular zone of magnitude $\phi$ occurs where a section of the distant surface is visible only to the right eye.

Howard and Rogers's rules are as follows:

1. The monocular regions in each eye are on the temporal side of the occluding surface S1.
2. A monocular zone due to occlusion is more distant than the binocular object that creates it (surface S1).
3. Eye movements affect the physical size of the monocular zone only slightly.
4. The angle of subtense of a monocular zone is inversely proportional to the distance of the occluded object. For a binocular object at a given distance, the angle $\phi$ increases with the distance between surfaces S1 and S2 increases.

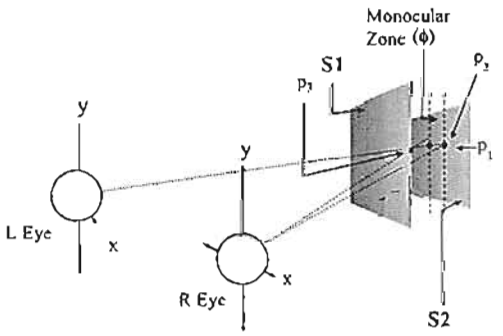



**Fig 7.2. Interposition and monocular occlusion**
Surface $S_1$ here occludes $S_2$ resulting in a monocular zone on $S_2$. The points $P_1$ on the distant surface and $P_3$ on the near surface are visible to both eyes while $P_2$ is visible only to the right eye. The result is a monocular zone of magnitude $\phi$. Note that the BIPASS model emphasises that surface separation involved the edge of the surface layer $P_3$ will correspond with point $P_1$ in the right eye and $P_2$ in the left eye. This crossed disparity configuration is an analogue of the split-projection configuration where a near surface is separated from a distant layer where visual projections cross.

The rules that define the shape and size of monocular zones are a direct geometric function of perspective projection of light to the retina (binocular parallax). Monocular zones have often been seen as redundant to stereoscopic information processing, however, that account is under revision. Julesz's *most distant surface rule* is the classic treatment of monocular features by previous computational approaches to stereopsis (see also Poggio and Poggio, 1984). Monocular features are more recently treated as particular *cues* to the presence of an occlusion relationship (Nakayama, 1996).

The SKS-U in Fig 7.1 bends the first two rules expressed by Howard and Rogers. In fusing the L-M pair, in Fig 7.1, a monocular feature occurs on the nasal side of the perceived *occluding* edge - on the near surface layer. In fusing M-R they are on the temporal side (of the occluding feature). In fact monocular pacman mouth can represent either occluding, and/or occluded monocular zones. Further, the SKS-U *bends* ecological rules because there is of course an unpaired region at the arc of the black circle fused with the pacman. The percepts at both signs of disparity necessarily involve monocular regions assigned to both depth planes simultaneously.

The SKS-U therefore seems to represent a problem for theories based on detection and response to monocular features as discrete primitives (see Nakayama, 1996 for example). Nakayama has claimed that the significance of monocular features for stereopsis was dependent on the eye of origin. He argues that monocular features trigger particular binocular responses that invoke the perception of occlusion. Anderson and Nakayama (1994) developed a hypothetical binocular receptive field construct, capable of detecting the breakdown of correspondence at occlusion in textured surfaces. Anderson and Julesz's also described decomposition of the half-images into matchable and unmatchable features, where unmatchable features yielded the perception an occluded surface layer.

Even Grossberg's notion that binocular FIDOs fill-in unpaired regions using signals from near or far disparity pools also seems undermined by the SKS-U. This is because the unpaired Kanizsa square configuration has no actual disparities that define the presence of a near surface edge. There are no disparity pools from which to fill-in the "gaps" of the stereoscopic percept. This means that the notion of allelotropia presented by Grossberg requires an alternative local mechanism to achieve surface separation. Here in lies the advantage of the BIPASS model.

The BIPASS model suggests that resolution of binocular differences in the size and shape of spaces between contours is integral to the percept achieved when the SKS-U half-images are fused. Consider, for example, the two retinal images cast by linear perspective projection in Fig 7.2. The horizontal subtense (relative size) of the distant surface S2 in each image will be smaller in the left eye than the right. If the left eye's image of S2 (S2L) equals $\phi$ S2L then the angle subtended by S2 in the right eye will equal $\phi$ S2L plus $\phi$ S2R, (the subtense of the monocular zone). This is a purely geometric observation based on linear projection. Indeed it is plausible to pose another rule of monocular occlusion as follows:

> The magnitude of the difference in angle subtended by the distant surface (S2) at each retina, is directly proportional to the distance of separation of the occluding (S1) and occluded surface (S2) (such disparity will also be minimally affected by eye movements).

There are two possible *resolutions* of disparate subtense (in Fig 7.2). It is possible that the distant surface, S2, look slanted. This is unlikely: if the bounds of S2 are unambiguously oriented, that is, they are not defined by contours horizontal to the line of sight (as in most natural edges); and/or, if the surfaces are textured (since each texture point potentially yields a local depth computation). The second possibility is the that S2 is perceived to stand behind S1. In other words, by splitting the percept into two layers at, say, P3-P1-P2 the system achieves a single cyclopean view of surfaces separated in depth.

### 7.1.3 Binocular Image Processing and the SKS-U

The SKS percepts generally, are interesting because they appear to have revealed something of the nature of mechanistic binocular information processing beyond point-disparity computation. They demonstrate (I argue) how binocular image processing can achieve surface separation in the absence of disparity discontinuity. The SKS-U percepts in particular, are interesting because the perceived depth step appears to be underpinned by the resolution of disparate subtense in the absence of point-disparity computation.

This experiment examined the stereoscopic processes involved by manipulating the size of visible parts of the SKS-U pacmen relative to the black circle with which they are paired. This was done in a way that simulated 2-D layout of image differences cast by a partly occluding object. A simplified drawing of the binocular array of projections is described in Fig 7.3. The pacmen are partly occluded in one eye's view by drawing a white square (shaped

according to the perspective model in Appendix then with disparity applied using Ogle's magnification factor).

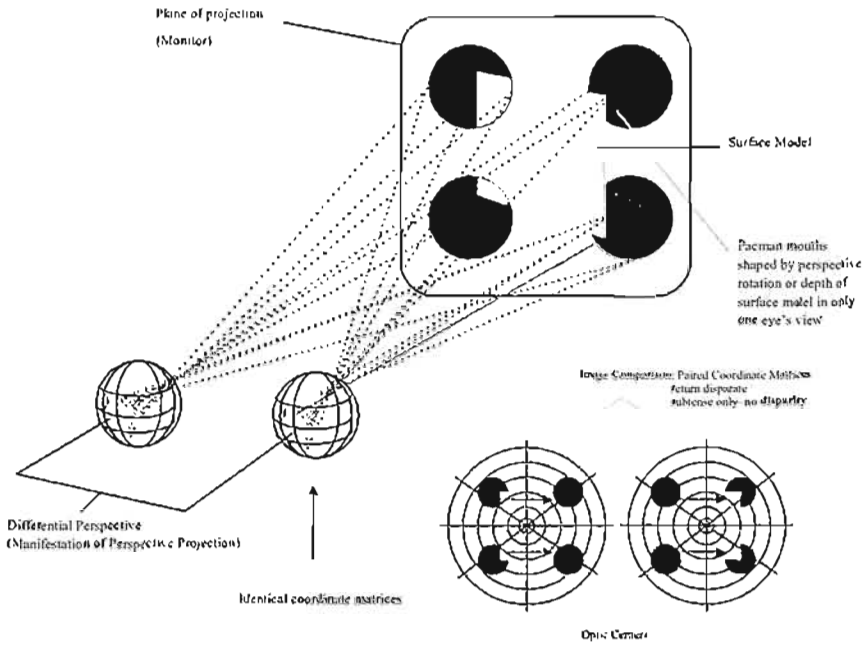



**Fig 7.3. 2-D layout of image differences in the SKS-U half-images**
This figure demonstrates the creation of SKS-U half-images using a square surface model to shape the pacmen. The configuration will reflect the 2-D layout of image differences views of a surface overlaying four circles at the distance of the projection plane. Subsequently, an image comparison process will yield substantial parity between the half-images. While disparity is in the form of disparate subtense indicated by the arrows.

The reader may find it useful to keep the concept of 2-D layout of image differences in mind as the discussion next turns to a detailed BIPASS model of the SKS-U percepts (see Fig 7.4).

*A BIPASS model of the SKS-U percepts*

Figure 7.4a I, describes the Image Registration and Image Comparison components of the BIPASS model. The images projected at the monitor (projection plane) are shown. Each image contains two pacmen and two circles. Useful sites of these image pairs have been labelled. A small monocular pacman mouth is present at $B_R$ in the right image and $C_L$ in the left (crossed fusion). The projection diagram in Fig 7.4a I examines the image comparison process when vergence is locked. The thick black lines represent each eye's view of the half-images at the at the projection plane (monitor).

*Vergence Lock and Image Comparison: Modal SKS-U percept*

Assuming that the arcs of the pacmen bounds are fused to yield single vision at points A, B, C and D. The greater part of the luminance contour bounding each pacman-circle is a non-disparate arc in each eye's view (ie. fused at fronto-parallel). Vergence angle, locked when these arcs are fused, will constrain the set of differences in subtense of large scale features of uniform luminance. There are few actual luminance contours. The figure outlines the manner in which these disparate features project onto the retina and the image measures achieved by alignment of the retinal coordinate matrices. Retinal parity between the images is extensive (equal subtense in each eye). For example, coordinate measures (D-C), (D-A), (B-A), (D-B), (A-C) and (C-B) will yield identical retinal measures.

A region of overlap exists at each eye's view at of the pacman mouths $C_L$_C (black in right eye and white in left) and at $B_R$_B (black in left eye and white in the right). This overlap yields a series of possible retinal difference measures, for example, $C_L$-D, $C_L$-B, $C_L$-A etc will yield disparate subtense. The luminance overlap at $C_L$_C and $B_R$_B is therefore constrained within a set of retinal parity and disparity measures, that is, measures of magnitude of angles subtended at the retina.

*Surface separation and spreading: Modal SKS-U Percept*

Figure 7.4a II describes how the set of retinal measures relates to the percept achieved by fusing the half-images. The percept is a white square standing forward of the four pacmen.

My argument here is that the 3-D illusory percept is achieved as the system resolves image differences at the retina by surface separation. The concept of surface separation is that the system splits (in functional terms) the percept into two layers at specifiable luminance contours (luminance steps). The luminance steps in question correspond to B and $B_R$ and C and $C_L$. These contours each demarcate a white-black luminance step.

In surface separation, the black side of the luminance step (D_C and B_A) remains assigned to the projection plane since A, B, C,D are constrained by the uniform (non-disparate) curvature of the pacmen-circle pairing. However, the white side of the luminance step (regions C_$C_L$ and B_$B_R$) are displaced to a near depth plane (to B' and C'). Displacement of the white surface layer achieves resolution of the disparate subtense across the pacmen (D_$C_L$) - (D_C) and (A_$B_R$) - (A_B) the non-disparate subtense between them.

Interestingly, surface separation involves a particular configuration of visual projections at B- $B_R$-B' and C-$C_L$-C'. This configuration means that the white surface edge displaced to B' and C' represents the intersection of the line of sight of each eye through the luminance step. These cyclopean edges represent a Split-projection configuration of double fusion in each eye. Hence, the magnitude of separation of surface layers as described by the BIPASS model is predictable from the magnitude of the disparate subtense across the pacmen-circle pair. The next two experiments test that prospect using seen slant.

The projection diagrams examine just one horizontal (epi-polar plane) of image differences. It should be noted that what is actually displaced onto a near disparity plane are the entire mouths of the pacmen. The surface spreading component of the process involves continuation of the white surface layer between displaced sectors of the retinal fields. The perception is that the entire white region partly bounded by these sectors is assigned to the near depth plane. Illusory contours bound the displaced white layer. So the modal illusory contours represent the spreading of a depth displacement between the mouths of the pacmen.

The BIPASS model proposes that a very similar sequence of functional stereoscopic processes underpins the perception of an amodal SKS-U. This occurs when the relative direction of intrusion of the mouths of the pacmen is reversed, that is, when the eye of origin of the pacmen-circle pairs are reversed. Figure 7.4b demonstrates this process.

*Vergence Lock and Image Comparison: Amodal SKS-U percept (portholes)*

Once again, assume that the arc of the pacmen bounds are fused to yield single vision at A, B C and D as described in Fig 7.4b I. Figure 7.4b I represents vergence locked at those points. This yields a set of differences in subtense of large scale features. These disparate features project onto the retina. Vergence lock achieves alignment of the retinal coordinate matrices, enabling image comparison.

In Image Comparison, the pacman mouths project to $B_L$_B (black in left eye and white in right) and $C_R$_C (black in right eye and white in the left). Luminance at $C_L$_C and $B_R$_B is constrained within a set of retinal point-parity and disparate binocular subtense at the two retinae.

*Surface separation and spreading: Amodal SKS-U Percept (portholes)*

The BIPASS model suggests that the 3-D illusory percept (an amodal SKS-U) is achieved as the system resolves image differences at the retina by surface separation and this results in

surface spreading at the depth plane of the pacman arcs (the projection plane). Surface separation splits the percept into two layers at specifiable luminance contours (luminance steps).

Figure 7.4a II once again describes how the set of retinal measures relates to the amodal SKS-U percept (that looks like portholes) achieved from fusion of the half-images (according the BIPASS model). The projection diagram is used to describe the percept. White corner segments (a white amodal square) stand behind the portholes. The point of this projection diagram is to outline how the 3-D percept is related to fit the binocular projection geometry subtended at the retinae.

Perceptual asymmetry arises because the system splits the percept into two layers at the arc of the pacmen-circle pairs (porthole percept) rather than at the pacman mouths (modal percept). In the projection diagram, the luminance steps in question in this case correspond to A, B, C and D. These contours each demarcate a white-black luminance step.

First consider surface separation at the mouth of the pacman. The mouths of the pacmen C_$C_R$ and B_$B_R$ are displaced to a distant depth plane. Regions of disparate subtense across the pacmen-circle pairs are also assigned to the distant depth plane.

In surface separation at the porthole bounds, the white side of the luminance step (D_C and B_A) remains assigned to the projection plane since A, B, C,D are constrained by the uniform (non-disparate) curvature of the pacmen-circle pairing. However, the black side of the luminance step within the entire arc of the pacmen is displaced to the displaced to a distant depth plane. Displacement of the black surface layer achieves resolution of the disparate subtense across the pacmen (D_C) - (D_$C_R$) and (A_B) - (A_$B_L$). There are no illusory contours generated at these portholes since the unpaired black circle contour, in each pacman-circle pair stands on the projection plane (or rather, the percept is split along this contour).

Surface separation involves the Panum analogue at B-B'$_R$-B'$_L$ and C-$C_L$'-C'$_R$ (at the pacmen mouths). The same configuration occurs at A-A'$_R$-A'$_L$ and D-$D_L$'-D'$_R$. This configuration means that a white near surface edge (porthole) at A, B, C and D bounds the displaced region within the porthole at the intersection of the line of sight of each eye onto the displaced surface. Once again these cyclopean edges represent a split-projection configuration. In sum, the magnitude of separation of surface layers as described by the BIPASS model is predictable from the magnitude of the disparate subtense across the pacmen-circle pairs.

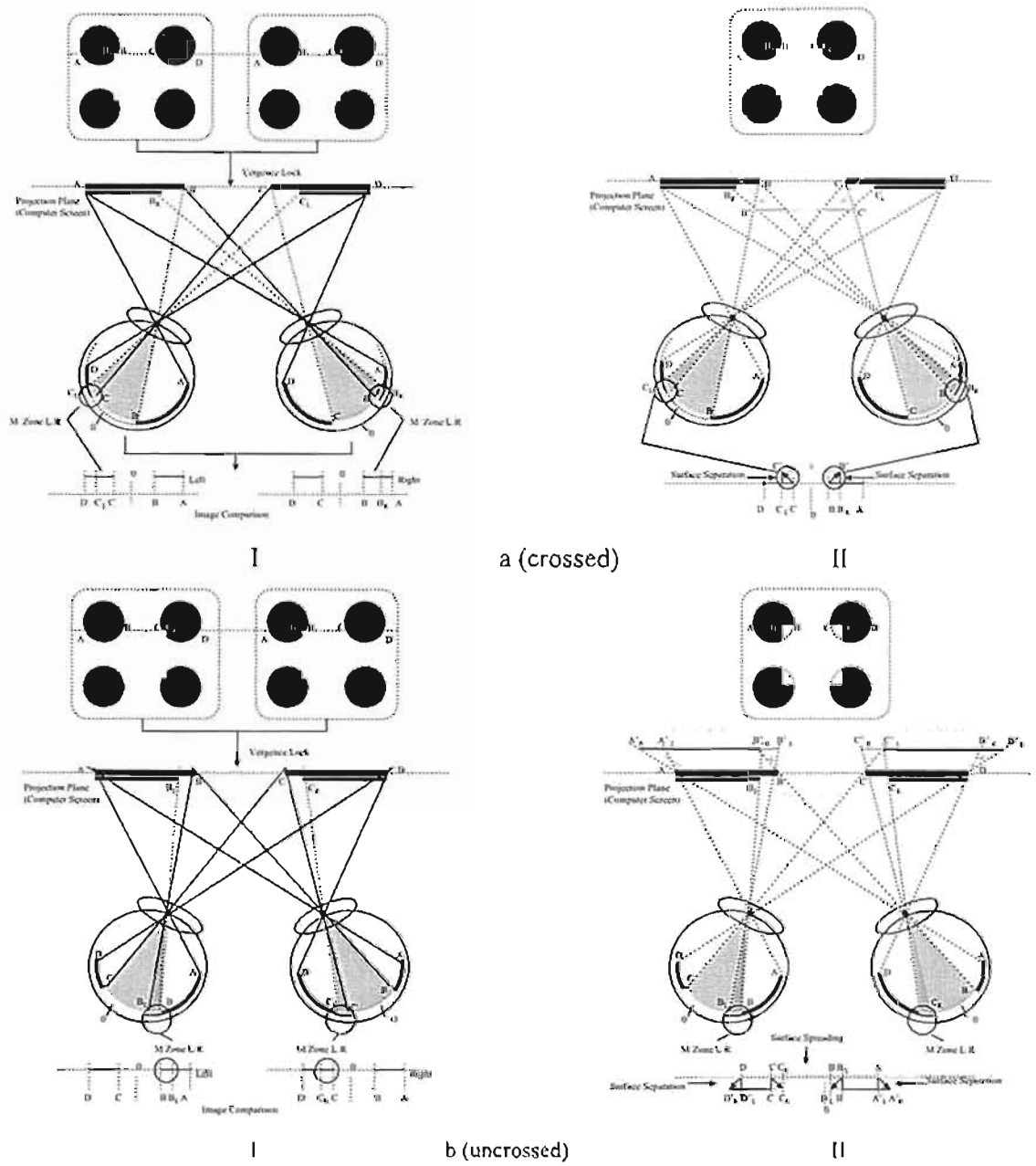



Fig 7.4. Binocular Image Processing underpinning the SKS-U percepts

This figure outlines a BIPASS model of stereopsis that may underpin the SKS-U percepts. In 7.4aI, Image Registration, and Vergence Lock are achieved in aligning the arc of the pacmen-circle pairs. Vergence Lock aligns the retinal coordinate matrices (whose origin are the optic centres). Image Comparison will return retinal parity at ABCD but no traditional point-disparity. It will however return disparate subtense across the pacmen-circle pairs $(A\_B_L)_{left} - A\_B)_{right}$ and $(D\_C_R)_{right} - (D\_C)_{left}$ ; and between the pairs at $(C\_B_L)_{left} - (C\_B)_{right}$ and $(B\_C_R)_{right} - (B\_C)_{left}$. In 7.4aII, surface separation and spreading are shown. The system separates surface layers at C' and B', the edges of a near white surface standing forward of the pacmen. The layer must spread across the space between the pacmen-circle pairs. In Fig. 7.4bI and II, the sequence is precisely the same. However, the system separates surface layers at the ABC and D, that is, the bounds of the pacmen-circle pairs. This yields the percept of set of portholes through which is seen an amodal square shape. Hence the percepts achieved are in a very real sense a resolution of disparate subtense measures.

Note that the projection diagrams examine just one horizontal (epi-polar plane) of image differences. Surface separation displaces the entire field bounded by the black circles to a distant depth plane in the porthole percept. The surface spreading component of the BIPASS model, at this sign of disparity, involves spreading of the white surface layer between from the bounds of the portholes.

*Measures*

In previous experiments, it has been demonstrated that disparate shaped pacmen set into a Kanizsa square configuration induced stereoscopic slant in an illusory surface when the half-images were fused. In this experiment we addressed the question as to whether or not the same effect would be evident in the SKS-U by manipulating subtense across the pacmen-circle pairs. It was predicted that surface spreading would arise between these separated surface layers. A very simple direct test of this prospect was designed.

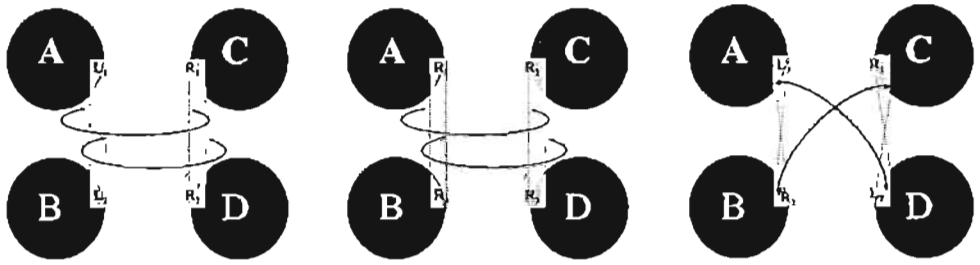
Figure 7.5 shows the manipulations carried out in Experiment 5. These diagrams (Fig 7.5a, b and c), show the two half-images aligned in the way that binocular fusion must align the luminance boundaries - so that the position of the unpaired pacman mouths is constrained. Each fused pacman has been labelled A, B, C or D.

In Fig 7.5a the unpaired mouths are presented in the left eye's view of pacmen A and B, that is, $L_1$ and $L_2$. The right eye's view of A and B are full circles. The unpaired mouths $R_1$ ad $R_2$ are in the right eye's image of pacmen C and D. Here the left view of C and D are full circles. The hypothesis was that this configuration (as in Fig 7.1) would induce no seen slant.

Figure 7.5b shows the arrangement predicted to induce seen slant in the vertical slant-axis. The unpaired mouths of the pacmen are positioned in the right eye's half-image of A,B,C and D at $R_1$, $R_2$, $R_3$ and $R_4$. In that case the left eye's view of A,B,C and D is of full circles. With crossed fusion of this configuration, it is predicted that A and B produce point disparity surface separation that sees assignment of the unpaired pacmen mouths to a ear surface plane. At C and D, a porthole will be invoked as the sectors $R_2$ and $R_3$ are assigned to a distant depth plane behind the depth plane off the pacmen. The hypothesis was that this configuration would yield seen slant in the vertical slant-axis.

Finally, in Fig 7.5c, the unpaired pacman mouths were introduced at A and D in the left eye's half-image ($L_1$ and $L_2$). The left eye's view of B and C was of full circles. Mouths

were positioned in the right eye's view of C and D, that is, at $R_1$ and $R_2$. The right eye's view of A and D was of full circles. The hypothesis was that this configuration would yield seen slant in the horizontal slant-axis.




a. Fronto-parallel b. Vertical slant-axis c. Horizontal slant-axis

Fig 7.5. Half-images underpinning an SKS-U
These figures show the position of unpaired pacmen mouths once the motor component of stereopsis led to fusion of the luminance boundaries to produce the circles or pacmen labelled A,B,C, D. In (a) the left pair of pacmen (A and B) in the right eye and right pair of pacmen in the left eye's view (C and D) have "mouths" (at $L_1$, $L_2$, $R_1$ and $R_2$). This predicts a Kanizsa square standing forward or behind the pacmen depending on the sign of binocular fusion. Seen slant of the surface induced should be zero. In (b) all pacmen in the right eye's half-image (A,B,C and D) have mouths (at $R_1$, $R_2$, $R_3$ and $R_4$). This predicts seen slant in the vertical slant-axis. In (c), diagonally opposite pacmen in each eye's half image as shown (at $L_1$, $L_2$, $R_1$ and $R_2$).

This experiment therefore tested the general hypothesis that manipulating the eye in which pacmen, that is, circles with mouths, were presented would generate predictable stereoscopic rotation of a separated surface layer.

### 7.1.4 Method

*Subjects*

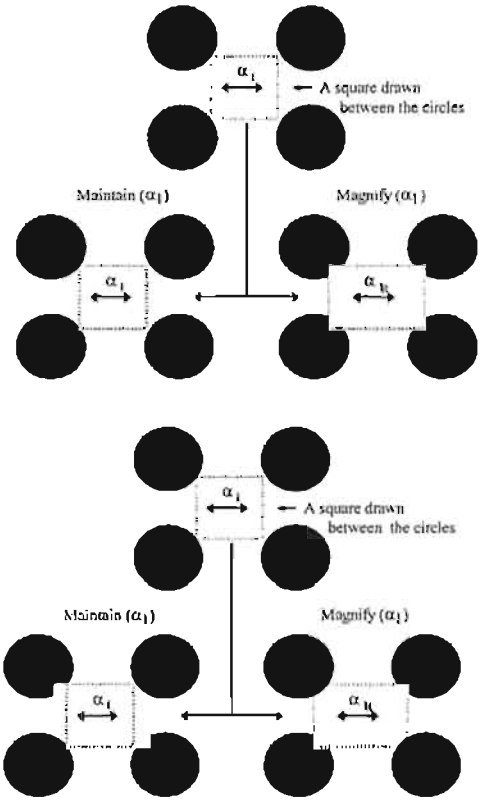
Ten participants were drawn from the pool of volunteers. They were subjected to precisely the same practice sessions and were screened for normal stereo-acuity as before.

*Stereograms*

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a set of four black circles subtending $3^0$ in diameter (at 750mm viewing distance). The circles were positioned so that a square drawn through their centres would subtend $7^0$ degrees. *Pacmen* were created by drawing a square, equiluminant with the

background, so that it just touched the boundary of the circles but did not obscure them. The square symmetrically intersected the arc of the circles (equal all sides).

Mouths of the pacmen were geometrically created according to the monocular transformations of a rotated square shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was then introduced by applying ½ Ogle's M to the overlayed square, symmetrically and in opposite signs (see Fig 7.5). The circles were then stood forward or behind this square to create a pacman in the appropriate eye's half-image. A magnification of this overlayed square was introduced. The magnification was appropriate to a stereoscopic rotation of about $40^0$ in both slant-axes was applied by manipulating the image pairs as shown in Fig 7.6.




a. Vertical slant-axis b. Horizontal slant-axis

**Fig 7.6. Creating inter-retinal differences in SKS-U half-images**
The half-images used to test the SKS-U were constructed from the same basic image. In (a) a set of four black circles was drawn then a white square was drawn between those circles. To generate slant in the vertical slant-axis the left and right half-images were then created by magnifying one eye's view of the square horizontally. In (b) images were constructed by magnifying both of the overlain white squares (in the left and right views). Then diagonally opposite circles were then stood forward of the square as shown.

Image pairs were presented on alternate frames at a rate of about 200 frames per second. Through the shutter goggles, background luminance was 0.7 cd $m^{-2}$ with the black pacmen 0.09cd $m^{-2}$.

*Design and procedure*

This study used a one way (1x3) repeated measures design to explore the effects of manipulation of the pacmen upon seen slant in the unpaired Kanizsa square. The design compared the effects of theoretical rotation (zero$^0$ rotation, 40$^0$ in the vertical slant-axis and 40$^0$ in the horizontal slant-axis). Ten complete repetitions of this design were used, making 30 trials per subject in all. Details of the procedure were as described in section 4.5.

### 7.1.5 Results and discussion

A one-way (1x3) analysis of variance was used to examine the effect of manipulation of mouths of the pacmen (zero, 40$^0$ in the vertical slant-axis, 40$^0$ in the horizontal slant-axis) upon seen slant. The effect was found to be significant: $F(9,2) = 14147.82$, $p < 0.001$. Results were pooled across the repetitions.

A slant underestimation of approximately 10 degrees was evident for both axes. Figure 7.7 compares the means for the effect. The difference between the slant-axes was within standard error.

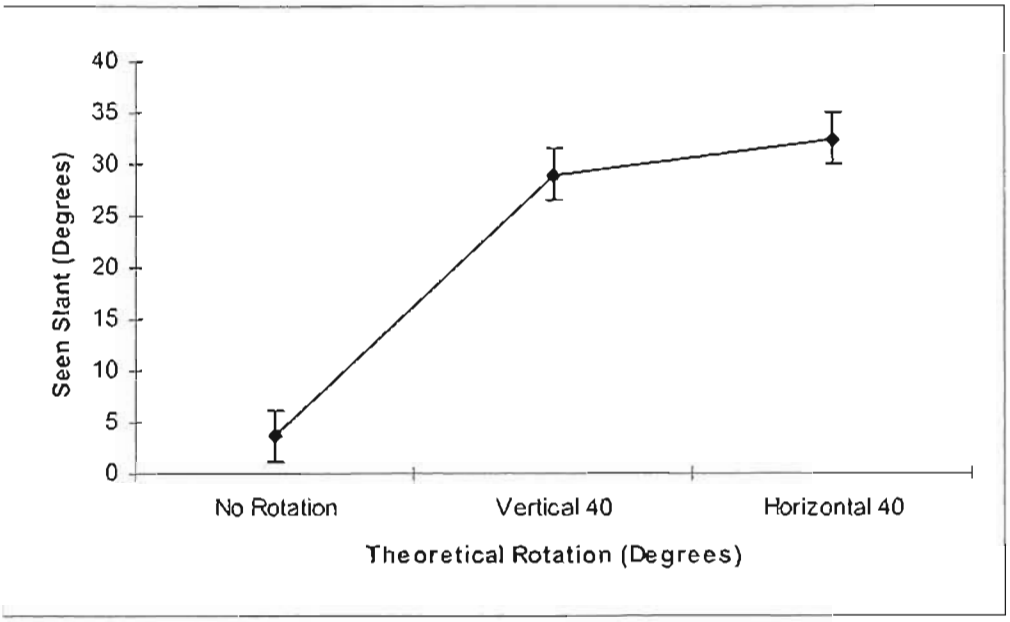



**Fig 7.7. The impact of theoretical rotation upon seen slant in a SKS-U**
This figure plots mean slant estimates for three levels of theoretical rotation in the SKS-U with standard error bars shown.

Results support the hypothesis that manipulation of the eye in which a monocular feature arose would impact predictably upon seen slant. All three specific hypotheses were supported. This is a remarkable finding since there is no disparity in the traditional sense in the half images presented. Separation of a surface plane can be associated with the displacement in depth of monocular mouths of pacmen in an unpaired Kanizsa square configuration. Further, the orientation of the surface layer separated from the depth of the pacmen appears to be governed by disparate subtense inter-retinal magnitudes on the space between the pacmen.

The eye of origin of unpaired features in untextured stereograms has been seen as a crucial cue to the perceptual organisation of untextured stereograms (Nakayama, 1996). The findings here concur. The difference in theoretical development though is what separates these findings from Nakayama's. The BIPASS model is the argument that monocular features are *resolved* not because of their eye of origin, but because the eye of origin is a function of the disparate subtense geometric relationships in binocular array. Moreover, sector corresponding to the pacman mouths of the black circles is also "monocular". In other words, I argue that Nakayama's emphasis of the *eye of origin of unpaired features* is only a partial description of the processes involved. Nakayama ignores the disparate binocular subtense that is concomitant to the presence of unpaired features.

The percepts generated by fusing the SKS-U were seen as part modal and part amodal slanted figure. Figure 7.8 depicts a selection of half-images used in Experiment 5 - for free fusers[13].

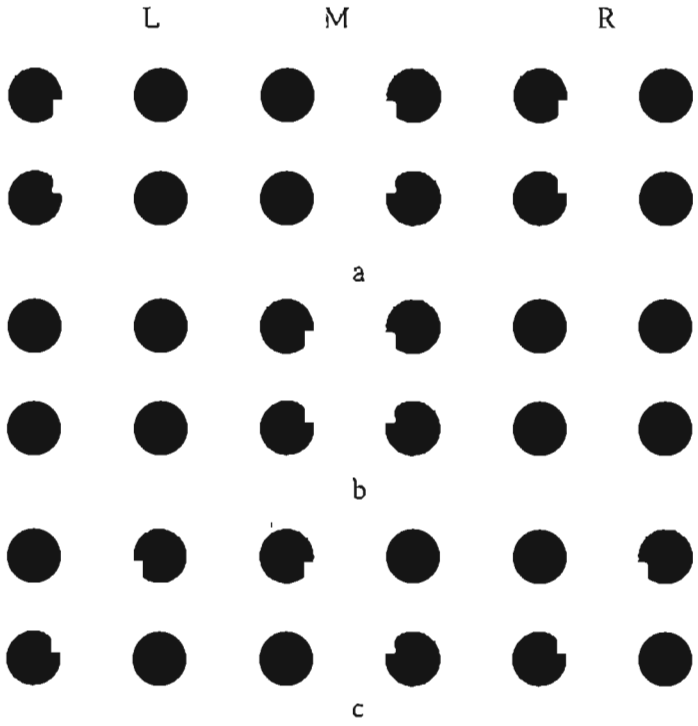



**Figure 7.8. Orientation in an SKS-U**
In (a) crossed fusion of pairs L-M yields a Kanizsa square standing forward of the pacmen standing on a distant surface plane. Fusion of M-R sees the Kanizsa square standing behind the pacmen. When the unpaired pacmen mouths are evident, appropriately contrived (see Fig 7.3) the percept generated is one of a part modal- part amodal form slanting through the projection plane. Crossed fusion of L-M in (b) sees a form slanted in the vertical slant-axis, with the right edge standing forward and the left edge standing behind the pacmen. Finally, crossed fusion of the L-M pair in (c) sees a part modal - part amodal form oriented in the horizontal slant-axis. The bottom edge stands forward and the top edge is seen through portholes. Fusing M-R sees this relationship reversed[14].

## 7.2 Experiment 6 Stereoscopic rotation of the unpaired Kanizsa square

A second study was conducted to ascertain whether the magnitude of surface separation was quantifiable and predictable from the disparate subtense separating pacmen. In the absence of point disparities, the assignment of a surface layer to depth may involve resolution of disparate subtense in the 2-D layout of image differences. Figure 7.9 shows the BIPASS model for the SKS-U stimuli. The perception of a part-modal part-amodal form, slanted in the

[13] Note that the texture of the age seems to disrupt the continuity of these forms compared to the experimental stimuli.
[14] Uncrossed fusion of these pairs will precisely reverse these percepts.

vertical slant-axis is predicted by separation of surface layers to create a triangular pattern of visual projections at the near surface edge. At the other end of the plane separation of surface layers happens at the pacman arc. In these projection drawings, the diameter of the corresponding pacmen and circles has been scaled so that in (a) the magnitude of seen slant is greater than in (b).

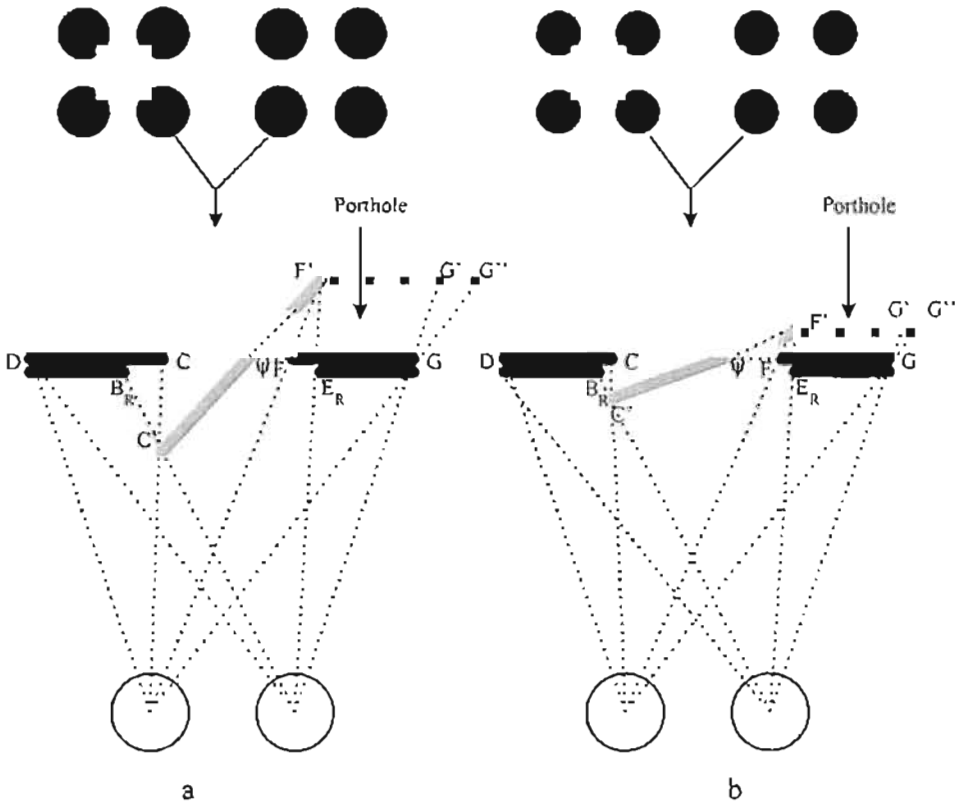



**Fig 7.9. Manipulating seen slant in the SKS-U**

In (a), fusion of an SKS-U is shown. Both sets of black lines represent the fusion of a circle (left eye) and a pacman (right eye). The grey line represents a part modal / part amodal illusory form slanted in the vertical slant-axis. The near surface layer separated from the pacmen spreads toward the projection plane at about ψ. The distant edge at F' is seen through a porthole.

Separation of surface layers in (a) occurs at a triangular occlusion configuration which leaves the $B_R$_C visible to the left eye. C'' represents the edge of the near surface layer. This arrangement of surface layers resolves the disparate magnitude D_$B_R$ - D_C. Adjacent to the configuration yielding the near surface layer a porthole occurs between F and G. This means that F and G represent the apices of triangular configurations where the near opaque surface is separated from the indeterminate black porthole. F' therefore is assigned to depth in a way that resolves the disparate subtense F_G (left eye) and $E_R$_G (left eye). Disparity in the space between the pacmen, that is, $B_R$_$E_R$ (right eye) - C_F (left eye) predicts the orientation of C'_F' an illusory form that is modal above the projection plane and amodal behind it.

In (b) precisely the same configurations arise. With the magnitude of the differences between D_C - D_$B_R$ and F_G - $E_R$ _G reduced by reducing the scale of the pacmen diameters. This manipulation predicts a lesser degree of orientation at C'_F'.

*Measures*

Experiment 6 manipulated the relative differences in subtense between the pacmen in the SKS-U to generate disparate subtense inter-retinal differences that predicted seen slant in a part modal / part amodal illusory form. This was achieved by maintaining the magnitude of the white square overlaying the pacmen but scaling the diameters of the pacmen. The hypothesis was that seen slant of about 20, 30, and 40 degrees about each axis could be achieved by precisely scaling the pacmen.

## 7.2.1 Method

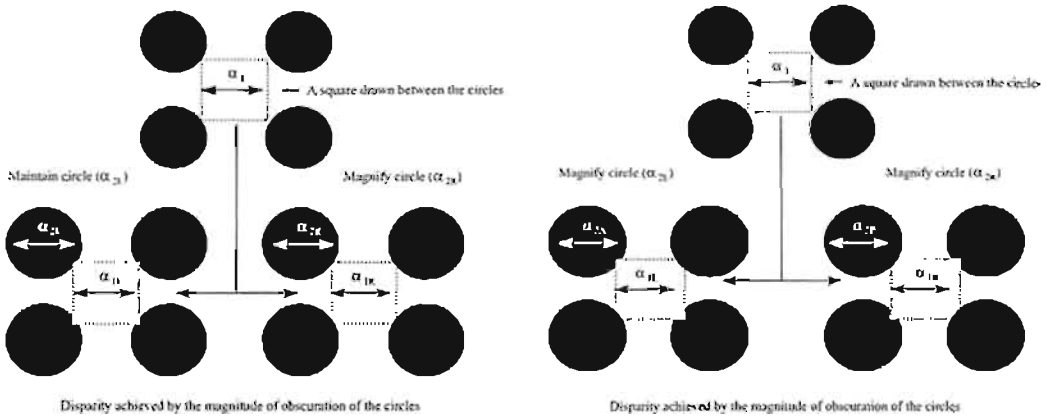
*Subjects*

Ten participants were drawn from the pool of volunteers. They were subjected to the same practice sessions and screening as before.

*Stereograms*

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a set of four black circles subtending $3^0$ in diameter (at 750mm viewing distance). The circles were positioned so that a square drawn through their centres would subtend $7^0$ degrees. *Pacmen* were created by drawing a square, equiluminant with the background, so that it just touched the boundary of the circles but did not partly obscure them. The square symmetrically aligned the arc of the circles (equal all sides).

Mouths of the pacmen were geometrically created by transformation of the diameter of the circles by scaling them symmetrically in proportions that equated to +/- ½ Ogle's M to each eye's view. Finally, the appropriate circles were stood behind the overlain square (see Fig 7.9).




a b

**Fig 7.9. Patterns of disparity interocular differences in an unpaired-SKS**
Fig 7.9a shows one half-image supporting vertical slant-axis. Disparity appropriate to vertical slant-axis was applied by increasing the magnitude of the Kanizsa square ($\alpha_3$) in one eye relative to the other. This obviously changes the relative magnitudes of the mouths of the pacmen ($\alpha_1$) and ($\alpha_2$). Standing disparity was then manipulated by constraining ($\alpha_3$) while shifting the illusory figure in one eye relative to the other. This alters the relative magnitude of ($\alpha_1$) and ($\alpha_2$) in each eye. Fig (b) shows the horizontal differences applicable to horizontal slant-axis. To manipulate rotation, disparate shear ($\lambda$) was applied in eye while constraining ($\alpha_3$). Then standing disparity was introduced by again constraining ($\alpha_3$) while shifting the illusory figure in one eye relative to the other so adjusting the relative magnitudes of ($\alpha_1$) and ($\alpha_2$) in each eye.

Image pairs were presented on alternate frames at a rate of about 200 frames per second. Through the shutter goggles, background luminance was 0.7 cd $m^{-2}$ with the black pacmen 0.09cd $m^{-2}$.

*Design and procedure*

This study used a fully crossed, two-way (3x2) repeated measures design to explore the effects of manipulation of the diameters pacmen (or circles), to simulate theoretical rotation and axis of slant respectively, on seen slant of an SKS-U. The design examined the effect of theoretical rotation (20, 30 and $40^0$) and slant-axis (horizontal, vertical). Six complete repetitions of this design were used, making 36 trials per subject in all. Details of the procedure were as described in section 4.5.

## 7.3.2 Results and discussion

A two-way (3x2) repeated measures analysis of variance was used to explore the effects of theoretical rotation (20, 30, $40^0$) and slant-axis (horizontal/vertical) on seen slant in the SKS-

U. Results were averaged across trials prior to analysis.

Seen slant varied predictably with theoretical rotation, as shown in Fig 7.10, and that effect was significant: $F_{(2,9)} = 259.908$, $p < 0.001$. A typical slant underestimation of about $5^0$ was established. There was no evidence of slant-axis anisotropy in these data. The difference between the disparate subtense means for slant-axis manipulation was not-significant $F_{(1,9)} = 0.148$, $p > 0.8$. This is not surprising since there could be no asymmetry between the point disparity and disparate subtense in the manifestation of 2-D layout of image differences in these configurations (because there was no point-disparity).

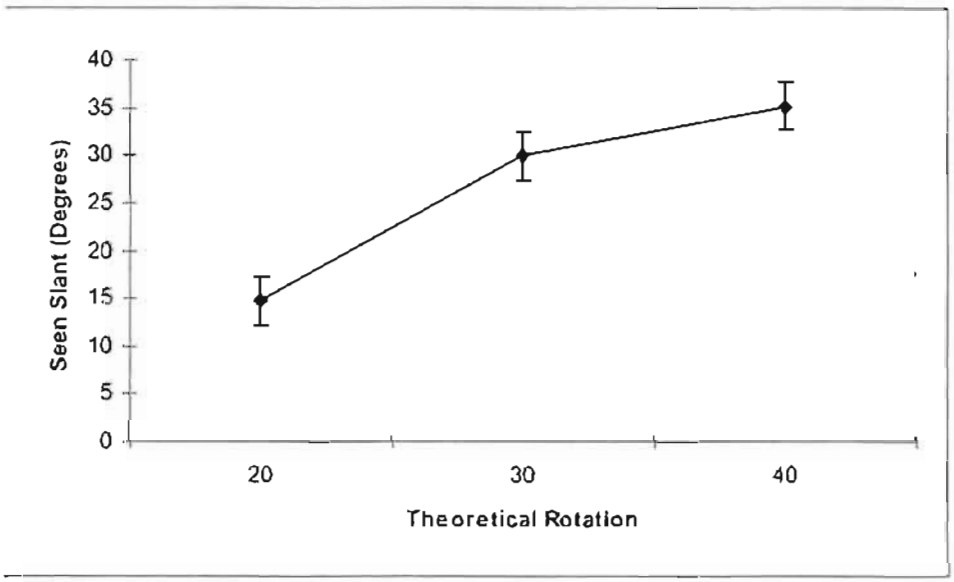



**Fig. 7.10. The effect of theoretical rotation upon seen slant in the SKS-U**
This figure graphs mean slant estimates for three levels of theoretical rotation in the SKS-U configuration where theoretical rotation was achieved by manipulating the diameter of pacmen.

These findings and the earlier demonstrations suggest that the system is able to access interocular difference information in a way that is very much more abstract than conventional point-matching. That is not to say that retinal correspondence is redundant in the untextured stereograms. The important stereoscopic aspects appear to be:

1. vergence lock at corresponding pacman / circle boundaries
2. separation of surface planes where a monocular displacement was generated
3. resolution of disparate binocular subtense in assigning surfaces to depth planes

4. spreading of the near surface layer

An approximation[15] of the stimuli presented in Experiment 6 is given in Fig 7.11 for free fusers.

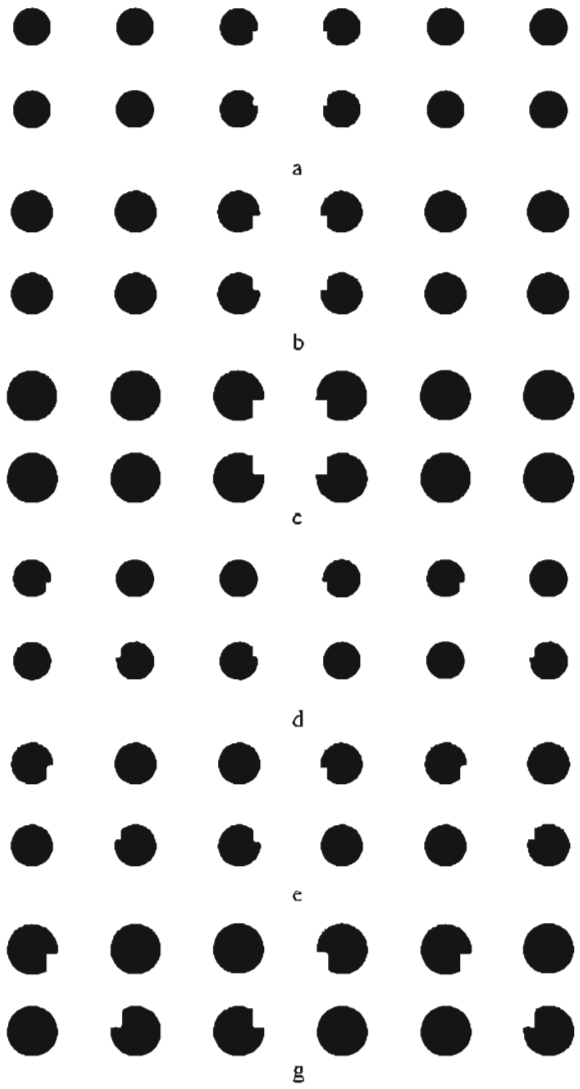



Fig 7.11. Seen slant in an SKS-U
Crossed fusion of the L-M or M-R pairs in (a to c) yield the percept of a part-modal form oriented to slant in the vertical axis. Seen slant increases from (a) to (c). The increase in slant, as shown in Experiment 6 is due to scaling the diameter of the circles used to construct the pairs. The same applies to pairs (d through g), but orientation is in the horizontal slant-axis.

[15] The stimuli here posses only an approximation of the differences applied in the experiment. They have been created to demonstrate the relative nature of the effects only.

In summary, at least in the context of the SKS-U percepts, monocular features can contribute to stereopsis and can occur as either occluded or occluding features in separation of surface planes. The evidence is that it is not the processing of the monocular features *per se* that is important, but the processing of disparate binocular subtense. This should not be taken to suggest that images created using unpaired features generate stereopsis in the same manner as point-disparity computation. The crucial factor that these stimuli indicate about stereopsis is that the concept of retinal disparity needs to be broadened.

This is of course the argument of Nakayama and Shimojo and Anderson and Julesz. I argue however, that resolution of disparate subtense across the image pairs partly defines the system's response to monocular zones - not inference. Gillam, Blackburn and Nakayama, (1999) have recently shown that slant is experienced when black panels in which one eye views an unpaired white line. It appears that the degree of slant observed is similarly related to subtense separating the ends of the panel pairs.

An example of the capacity of stereopsis to achieve a 3-D percept by surface separation (in resolving the 2-D layout of image differences) is shown in Fig 7.12. The half-images contain disparate subtense in oblique axes that is putatively resolved by surface separation. The perception is of a surface plane oriented in oblique slant-axes. For example, in Fig 7.2 the modal and amodal versions of the SKS-U can be rotated through about $45^0$ about the z plane, and still be readily fused to generate a strong sense of surface separation. This suggests that the system may not be constrained by epi-polar lines in the same manner as described for retinal point matching (see for example Howard and Rogers, 1995; Anderson and Julesz, 1995).

In figure 7.12a, fusion of half-images comprising unpaired pacman mouths that are obliquely adjacent in the cyclopean percept generates complete 3-D illusory percepts. By manipulating the disparate subtense in these half-images subtense percepts in which surfaces stand forward of the pacmen (cross fuse L-M), or behind them (cross fuse M-R), can be perceived.

Further, fusion of the pairs 7.12b (monocularly adjacent) and 7.12c (diagonally adjacent) demonstrate that it is possible to generate percepts of obliquely orientated part modal / amodal surfaces using the SKS-U configuration. Quite stable 3-D, slanted percepts arise.

Stereopsis is usually thought of as a mechanism that compares horizontal differences in the position of matchable pairs of texture points. The stereograms in Fig 7.12 have no matchable disparate texture-points where the surfaces separate, and the image differences present are not, it seems horizontal at all. Indeed Anderson claims that vertical disparity is a cue to surface separation. However, the rotation of the illusory figures suggests that it is the resolution of disparate subtense not an not inferential response to vertical disparity that is the important issue.

Nonetheless, there is undoubtedly some instability in these figures. It is not clear whether this is rivalry *per se* or the perceptual fading of the illusory contours. Some almost vertical image differences in these configurations that appear to be quite stable. It is hoped that future research into these kinds of non-epipolar patterns of disparate subtense will enable a more detailed account of binocular image processing that the description posed in this thesis.

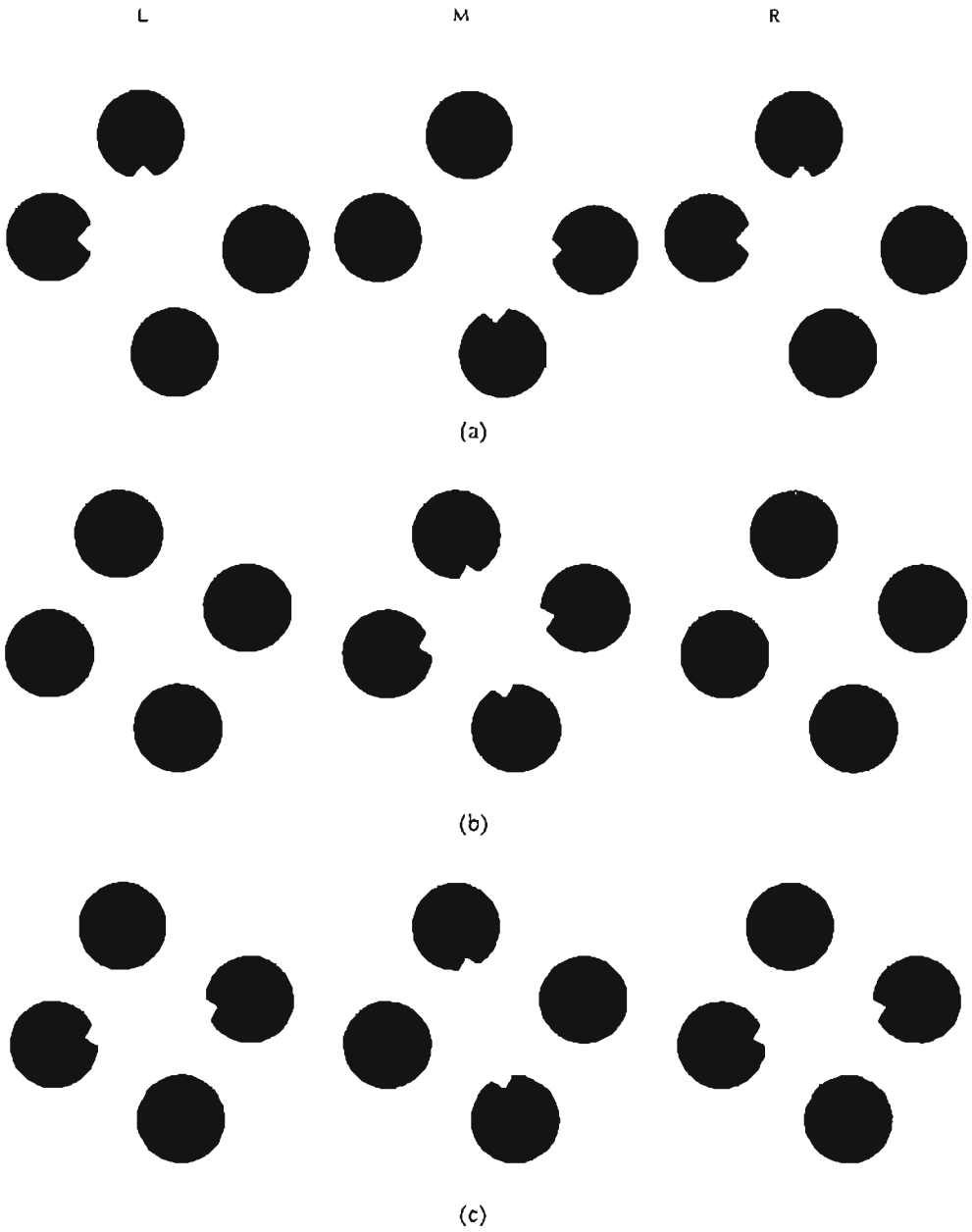



Fig 7.12. Orientation of an SKS-U in oblique slant-axes
Fusion of these figures yields a percept equivalent to the patterns of occlusion resolved by surface separation. These percepts suggest that though vertical disparity may be an important local aspect of partial occlusion, the resolution of disparate subtense in oblique directions relative to the line of site dramatically impacts perceived orientation of perceived surface layers. My argument is that disparate subtense across large scale spaces between pacmen may be the crucial factor - not vertical disparity as a special cue *per se*.

In conclusion two quite simple experiments in this chapter tested the prediction of the "BIPASS model" that the depth assigned to surface layers separated in fusing the SKS-U half-

images was predictable from the magnitude of disparate subtense across the pacman-circle pairs. This suggests strongly that stereoscopic mechanisms are fundamentally involved in generation of these 3-D SKS-U percepts. It is likely that such mechanisms may be common to many other stereograms yielding the 3-D illusory percepts. The next chapter assesses the Ehrenstein grid figure in terms of binocular image processing.

# 8. Contrast spreading in a stereoscopic Ehrenstein square

*Summary: This chapter examines contrast spreading in a stereoscopic Ehrenstein square (SES). The chapter firstly describes the SES percepts and outlines the Surface Heuristic and Form computation approaches to explaining the phenomena. The possibility that the SES percepts might be related to binocular image processing similar the SKS percepts is explored. Three experiments are presented concerning seen slant, ambiguous disparity and completion-type mechanisms. Results suggest that contrast spreading in stereopsis is indeed related to the SKS percepts. In particular, the manner by which the system resolves an extensive structure of 2-D image differences (in image pairs) seems to be very similar.*

## 8.1 Stereopsis and contrast spreading

### 8.1.1 A Stereoscopic Ehrenstein Square

Contrast spreading is reported when observers fuse image-pairs constructed from an Ehrenstein Grid with a grey cross at the centre (see for example Nakayama, Silverman and Shimojo, 1998). For convenience the stimuli will be called the stereoscopic Ehrenstein square (SES). Each SES half-image is composed of four black panels. A cross that has intermediate luminance is drawn between the squares. Disparity is introduced by changing the position of the central cross in each eye relative to the position of the black panels. Figure 8.1 presents image pairs appropriate to crossed and uncrossed disparities. Parts of the half images have been labelled in Fig 8.1b to help describe the 3-D percept[16].

When the half images L-M are cross-fused (Fig 8.1a) there is the impression that a thin transparent grey surface or film stands forward of the black panels[17] (this phenomenology is well documented, see for example Grossberg, 1994). Observers report illusory contrast spreading where the grey surface crosses the corners of the black panels between A to E, F to C and; D to H, G to B (in Fig 8.1b). The grey layer has a roughly oblong shape.

---

[16] This has been done because the letter labels can themselves impact upon organisation of surface layers.

[17] These half-images were created by drawing four black panels 1,2,3 and 4. A grey square was drawn to overlay the black panels. The center of the grey square was positioned in the middle of the set of four black panels. The grey square was then stood behind the black panels. This configuration was copied and re-drawn to create L, M and N. Disparity was then applied by shifting the grey square in half-image M to the right about 4mm. Pairs L and R are identical.

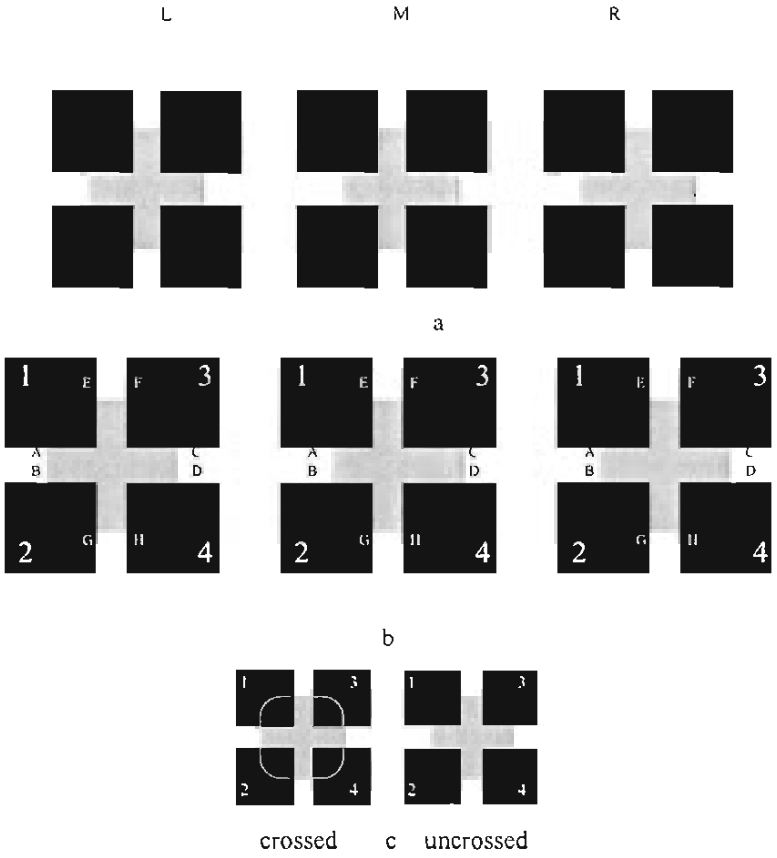



**Fig 8.1. Stereoscopic contrast-spreading**
In (a), crossed fusion of L-M yields the percept of a transparent grey figure standing forward of opaque black panels. Fig (b) is the same stimulus but has been labelled to help explain the 3-D illusory percept. In fusion of the L-M pair illusory contrast spreading occurs across the corners of the squares at FC, DH, GB and AE. Crossed fusion of M-R will see a similarly shaped object (amodal closure) standing behind opaque black panels. The 3-D percepts are summarised in fig (e). In the crossed disparity case (the L-M pairs) "modal" contrast spreading appears to complete a roughly square shape (represented by the grey line). In uncrossed disparity, the white and black parts of the image are opaque. The grey cross now stands behind the black panels and gives the impression of an amodal "squarish" shape.

A sketch of the established 3-D illusory percept is included in Fig 8.1c. Contrast spreading describes the perception of illusory change in contrast that seems to involve completion of an interposed grey layer as it cuts across the corners of the black panels.

Crossed fusion of pair M-R, (in Fig. 8.1a ) yields the impression that an opaque grey figure stands behind the depth plane of the black panels. Disparities are reversed[18]. No contrast spreading is perceived. Also, the white area between the black panels looks transparent and

[18] Reverse disparity (pair M-R) is simply produced by shifting the grey square that stands behind the black panels in the opposite direction.

not like a surface layer. The cross now looks like an opaque plane partly obscured by black panels (sketched in Fig 8.1d).

At both signs of disparity the whole cross is stood forward or behind the black panels. The perceptual difference experienced by observers (fusing the two signs of disparity) cannot be explained just by assignment of "near" or "distant" depth values to particular contours. There is a change in the quality of the grey layer between the two signs of disparity. Point-disparity occurs at the contours A_B and C_D in the SES half images. Mechanisms that cause surface spreading in the SES appear to be switched on or off depending on the sign of disparity at just two contours.

### 8.1.2 Recent accounts of stereoscopic contrast spreading

The Ehrenstein figures presented in this chapter stem from an effect called *Neon Colour Spreading* first demonstrated by van Tuijl (1975). Stimuli inducing 2-D Neon Colour Spreading were refined by Redies and Spillman (1981) who showed that when a red cross was inserted in the centre of an Ehrenstein grid the red colour spread into the region bounded by illusory contours (see Chapter 2 for a review). The spreading effect is common to achromatic versions of the stimuli (Anderson, 1997).

A Surface Heuristic account (Anderson and Julesz, 1995; Anderson, 1997; Nakayama, Shimojo and Ramachandran, 1989; Nakayama, 1996) has argued that the transparency perceived is due to contrast ordering and detection of monocular features. Contrast ordering is the sequence of change in contrast along contours separating white, black and grey regions. Mertelli (1974) defined the relationship between luminance order and transparency, explaining that that to appear transparent, a surface must have intermediate luminance. This is the case in the contrast-spreading stimulus in Fig 8.1.

Anderson (1997) suggests that luminance order along the black-white and black-grey contours (say at A, square 1 in Fig 8.1) facilitates *scission* of a particular intermediate contrast region into two *causal layers*. He argues that when two contours are aligned, a change in contrast that does not disrupt *contrast polarity* can result in the lower contrast region being *decomposed* into two separate layers. Hence, transparency results when two surface layers can represent a change in contrast along aligned contours. A shift in contrast polarity means a shift from high contrast to low contrast, or *vice versa*.

Watanabe and Sato (1989); Watanabe, Takeichi and Shimojo (1990); Watanabe and Shimojo (1990); Watanabe (1995) have argued that these remarkable contrast spreading

effects are due to the parallel operation of Grossberg's (1987) Boundary and Feature Contour mechanisms, but in the stereoscopic domain. This is a Form Computation account.

Grossberg's explanation has been that contrast spreading is an emergent output of the image segmentation process he describes in FACADE theory. The Boundary Contour System integrates vertical and horizontal end-cuts at the vertical and horizontal extents of the grey cross (End-cuts are the output of hyper complex receptive fields). Integration of these end-cut signals creates a rounded pre-visual boundary that unites disparity signals. The Feature Contour System's binocular FIDOs then fill the area within the boundary contour from a zero-disparity pool. In addition, contrast spreading is the output of the filling-in domains. Figure 8.2 illustrates this process[19].

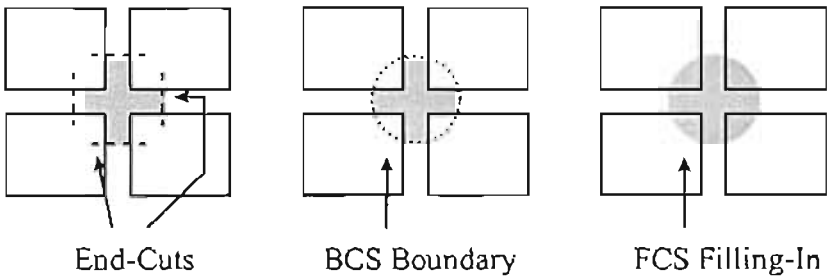



**Fig 8.2. Grossberg's FACADE theory of contrast spreading**
In Grossberg's view, the contrast spreading percept is the result of a sequence of end-cut, Boundary integration by the BCS and finally, filling in by the FCS's binocular FIDO.

Grossberg's (1994) explanation of the reversed disparity sign percept, where no contrast spreading occurs, is that black square boundaries inhibit the filling-in process but not the integrative BCS mechanisms. However, Anderson has shown that stereoscopic contrast spreading can occur in configurations where it appears unlikely that the percept is caused by integration of form *per se*. Anderson's "I-Junction" demonstrations have established that integrative perceptual mechanisms like Grossberg's BCS are not necessary conditions for the induction of illusory contours and contrast spreading in stereopsis.

### 8.1.3 Stereoscopic T-Junctions in the SES: A surface heuristic

T-Junctions are seen as a particularly important surface heuristic[20] (Anderson and Julesz, 1995; Nakayama and Shimojo, 1992; and Anderson, 1997). T-Junctions are highlighted in Fig 8.3. The label "T" describes the three-way junction of luminance values white, black and grey. The stem of the T-junction (T-stem) is the contour that bounds white and grey. The top of the

[19] This figure was shown earlier in Chapter 3.

T-junction (T-top) is made up of the white-black and grey-black edge of the black square (say, square 1).

In the SES half-images, there is a change in contrast along the contour defining the T-top (highlighted in Fig 8.3a). Anderson's argument about the separation of surface layers along the T-top is this:

> When two aligned contours undergo a discontinuous change in the magnitude of contrast *[along the T-top]*, but preserve contrast polarity, the lower contrast region is decomposed into two causal layers *(Anderson, 1997, p2)*.

Anderson's concept essentially comes from Mertelli's (1974) *surface scission*. Anderson argues that stereopsis enhances and amplifies surface scission. His view is that the part played by stereopsis is disambiguation of occlusion geometry, for example, decomposing the half-images into matchable and non-matchable features. Unmatchable features signal partial occlusion (in the manner of an *a priori* heuristic). The system separates the percept into two causal layers *attaching surface quality to contours*.

In a traditional account of stereopsis the SES percepts are under-determined by local point-disparities. Nevertheless there has been no detailed analysis of image difference information available in the SES stereograms or the manner in which the 3-D illusory percepts are related to inter-retinal differences that do exist. This chapter attempts to do so.

### 8.1.4 Binocular image processing and the SES percepts: Crossed disparity

This section examines the 2-D layout of luminance contours making up SES half-images. It describes how binocular image processing might be related to perceptual asymmetry in the SES percepts.

It is probably not controversial that SES half-images register at each retina and vergence angle is locked by coordinated eye movements. This aligns coordinate matrices enabling the system to exploit inter-retinal parity and disparity to achieve a single cyclopean percept. Comparison of the two images between coordinate matrices will return point-disparities and disparate subtense. Indeed, in untextured stereograms, the two are immutable. It is not possible to generate point-disparity without creating a reciprocal disparate subtense.

*Layout of disparity in SES half-images*

---

[20] Note that this is my term - not the authors'. It is a something of a simplification of their idea.

The layout of 2-D image differences can be systematically described by mapping one half-image onto the other (see Fig 8.3a). The T-stems, that is, the grey-white contours $A_R\_B_R$ in the right and eye and $A_L\_B_L$ in the left eye, are disparate. T-stem $C_R\_D_R$ in the right eye is disparate to $C_L\_D_L$ in the left. Two upside down T-Junctions are adjacent the ones highlighted. Non-disparate T-Junctions also occur, at E and F and G and H. The black-white luminance contour bounding the greater part of the black panels is also non-disparate.

There is also disparate subtense along both sides of the T-tops. For example the grey-black luminance contours at the T-top (in Fig 8.3a), $O\_A_R$ and $O\_A_L$ subtend disparate visual angles. Further, there is a reciprocal disparate subtense along the grey-white contours at the T-tops, for example, $M\_A_R$ - $M\_A_L$. The total image magnitude along the T-top is not disparate, that is, M_O subtends the same magnitude in each eye.

In sum, image comparison will return a set of retinal parity and disparity measures including but not limited to point-disparity at the T-stems.

*Resolution of 2-D image differences*

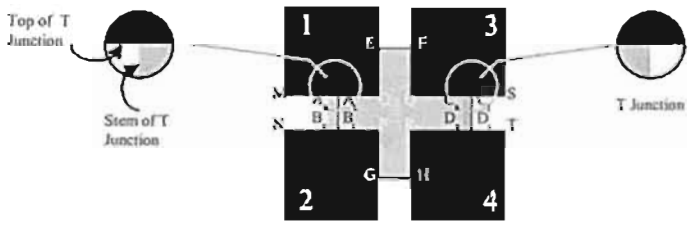
My argument is that in functional terms the SES percepts amount to a resolution of the layout of 2-D images falling upon the retinal coordinate matrices. The system responds like an image processing device designed to recover a single head-centric cyclopean view. Figure 8.3b (crossed disparity) is a projection drawing that shows how image differences appear to relate to the SES percept. The drawing examines a single epi-polar plane between points M and S (the bottom of squares 1 and 3 in Figs 8.3a). The black lines represent black square boundaries, ie. T-tops. The grey lines represent the grey-black portion of those boundaries (and cross the space between the squares). The disparate T-stems are at $A_R$ (right eye) $A_L$ (left eye) and $C_R$, $C_L$.

Fusion of disparate contours $A_R$ to $A_L$ and $C_R$ to $C_L$ looks like a conventional crossed disparity. Conventional stereopsis will assign these contours to the depth of A' and C'. These points represent the two vertical bounding edges of a near grey surface layer. To resolve disparate subtense adjacent to these contours the system must simultaneously assign the white-black portion at the T-top ($M\_A_R$, $M\_A_L$) to the zero disparity plane. The system must also assign the grey-black portion at the T-top ($A_R\_O$, $A_L\_O$) to the zero disparity plane. Hence the near edge is separated from a distant surface still visible when correspondence is resolved.

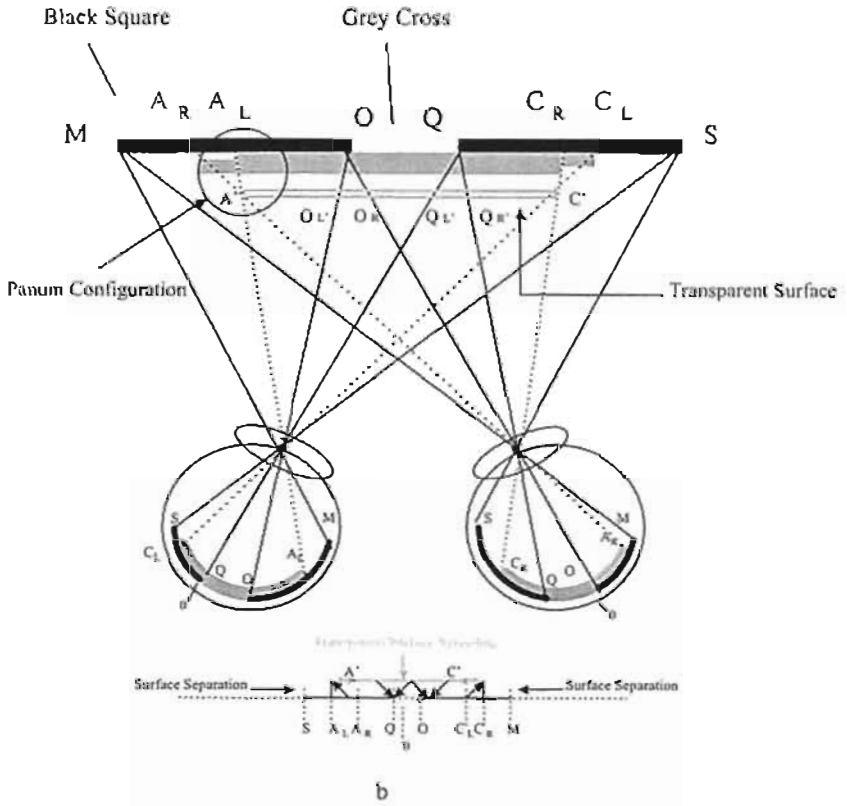
Surface separation (the perceived depth step at the T-stems) resolves point-disparity in both the position and disparate subtense of contours along the T-stem and the T-top. This could only be achieved at the split-projection configuration $A\_A_R\_A_L$ with A at the apex of a split-projection configuration. The system seems to split the percept into two layers along the T-stems. This leaves the white side of T-stems at the projection plane but stands the grey side of the T-stem forward.

Fusion of the T-stem at $C_R$-$C_L$ (Square 2 in 8.3a), precisely mirrors the Split-projection configuration at A. Between these two disparate T-Junctions there is disparate subtense along the T-tops ($A_R\_O$ - $A_L\_O$ and $C_R\_Q$ - $C_L\_Q$). There is no disparity at all between O and Q. It seems reasonable that the system might interpret the T-tops at A_O and C_Q as slanted. This does not seem to happen. When the near grey layer is transparent, O and Q must stand on the projection plane. Visual projection to these contours cross at the projection plane, through the near layer (at $O_R$' and $O_L$' and at $Q_R$' and $Q_L$').

Thus binocular subtense constrains separation of surface layers to a near depth plane (between A and C) and fusion of contours at the distant depth plane (along the T-tops) simultaneously. The only way that this can happen is when the near surface layer is transparent. Since contours are fused at and through a near surface, then it is transparent. The system does not have to infer causal layers but separates surface layers to different depth planes where both layers are visible.




a




b

**Fig 8.3. Surface separation at crossed disparity in the SES**

Figure 8.3a shows the half-images from Fig 8.1 aligned by vergence eye movements. Two T-Junctions are highlighted. Disparity exists at the stem of these T-Junctions. So contours $A_R\_B_R$ - $A_L\_B_L$ and contours $C_R\_D_R$ - $C_L\_D_L$ are disparate. In Fig 8.3a disparate subtense exists *along* the black-white contours $M\_A_R$ - $M\_A_L$, $N\_B_R$ - $N\_B_L$, $C_R\_S$ – $C_L\_S$, and $D_R\_T$ – $D_L\_T$. Disparate subtense also occurs along the black-grey contours $A_R\_O$ – $A_L\_O$, $B_R\_P$ -- $B_L\_P$, $C_R\_Q$ – $C_L\_Q$, $D_R\_R$ -- $D_L\_R$. There is also disparate subtense across the image at $M\_C_R$ – $M\_C_L$, $N\_D_R$ – $N\_D_L$, $S\_A_R$ – $S\_A_L$ and $T\_B_R$ – $T\_B_L$. These disparities are precisely reciprocal in magnitude to local disparity at the T-stems.Figure 8.3b is a perspective drawing of a single epi-polar line through M and S. The black lines at M_O and Q_S are the black square bounds (T-top). These "boundaries" are composed of a black–white contour and a black-grey contour. The grey lines represent the grey-black contour along the T-top (and cross the space between them). The disparate T-stems occur at $A_R$ / $A_L$ and $C_R$ / $C_L$. The near surface layer stands at A_C.Fusion of disparities at $A_R$-$A_L$ and $C_R$-$C_L$ recovers the near surface. The grey side of each T-stem is assigned to A-C. But, the white side of the T-stem, say $M\_A_R$ and $M\_A_L$, is assigned to the zero disparity plane at the T-top. One way this could be achieved is a Split-projection configuration where the system splits the percept into two layers along a projection line that is single at the near surface but projects to disparate points on a distant surface plane. Also note that the system separates surface layers but still assigns $O\_A_R$ – $O\_A_L$ and $Q\_C_R$ – $Q\_C_L$ to the distant black (along the T-top). To achieve this A_C must be transparent.

*Binocular image processing in the SES percepts: uncrossed disparity*

Figure 8.4 examines binocular image processing at the opposite sign of disparity. Vergence eye movements will align the retinal images along, say, the black square bounds. The diagram in Fig 8.4a maps one half-image onto the other. The image differences are exactly the same magnitude as those outlined above but the eye of origin of those magnitudes has been reversed.

The manner in which the system detects and responds to the half-images can be analysed using Fig 8.4b (uncrossed disparity), a projection drawing along an epi-polar line between M and S (the bottom of squares 1 and 3 in Figs 8.4a). Figure 8.4b poses the opposite sign of disparity (compared to Fig 8.3a) at T-stems: $A_R$-$A_L$ and $C_R$-$C_L$. The percept is a grey surface between A' and C' standing behind the black panels. No separation of surface layers occurs along the T-stems. Instead, the system seems to split the percept into two layers at the inner bounds of the black panels, that is, at Q and O and along the T-tops. Uncrossed disparity at the T-stem means that a depth step occurs along the T-top and the grey region between the two black panels is assigned to the depth of the T-stems. The grey surface looks opaque.

The grey surface is separated from the inner boundaries of the black panels along crossed visual projections where Q and O are at the apex of a Split-projection configuration. In that way, the system resolves disparate subtense at $A_L\_O - A_R\_O$ and $C_L\_Q - C_R\_Q$. If this were not the case, these contours would appear slanted from O and Q toward A and C respectively.

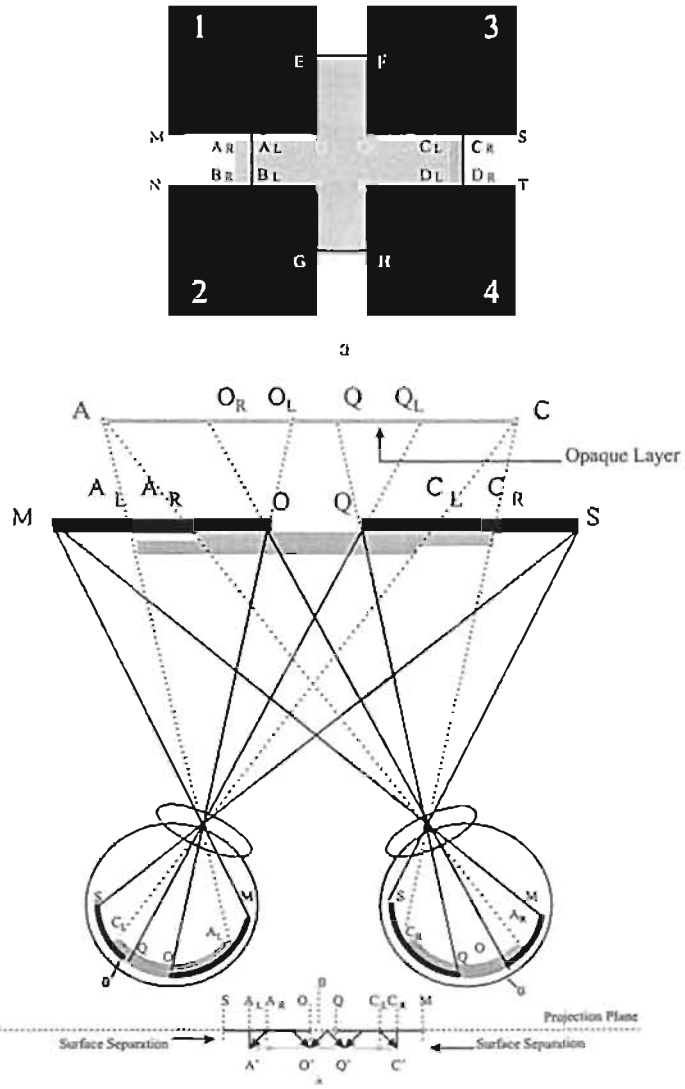



**Fig 8.4. Uncrossed disparity and stereoscopic surface separation in the SES**
Figure 8.4a compares half-images from Fig 8.1 aligned by vergence eye movements. Uncrossed disparity exists at the stem of the T-Junctions. So contours $A_R\_B_R$ - $A_L\_B_L$ and contours $C_R\_D_R$ - $C_L\_D_L$ are disparate.

In Fig 8.4a disparate subtense exists *along* the black-white contours $M\_A_R$ - $M\_A_L$, $N\_B_R$ - $N\_B_L$, $C_R\_S$ – $C_L\_S$, and $D_R\_T$ – $D_L\_T$. Disparate subtense also occurs along the black-grey contours $A_R\_O$ – $A_L\_O$, $B_R\_P$ – $B_L\_P$, $C_R\_Q$ – $C_L\_Q$, $D_R\_R$ – $D_L\_R$. There is also disparate subtense across the image at $M\_C_R$ – $M\_C_L$, $N\_D_R$ – $N\_D_L$, $S\_A_R$ – $S\_A_L$ and $T\_B_R$ – $T\_B_L$. Disparate subtense is precisely reciprocal in magnitude to local disparity at the T-stems.

Figure 8.4b is a perspective drawing of a single epi-polar line through M and S. The black lines at M_O and Q_S are the black square bounds (T-top). The disparate T-stems occur at $A_R$ / $A_L$ and $C_R$ / $C_L$. The near surface layer stands at A_C. At this sign of disparity $A_R$ / $A_L$ and $C_R$ / $C_L$ are in an uncrossed relation to the T-top. No surface separation is achieved at the T-stem but the contours are fused and assigned to the depth of C. This means that the disparate subtense at $O\_A_R$ – $O\_A_L$ and $Q\_C_R$ –$Q\_C_L$ might be resolved in a Split-projection configuration at O and Q respectively. This is a requirement for the surface A_C to appear fronto-parallel. Otherwise the percept would be of a surface slanted A-O and one slanted Q-C.

In sum, the stereoscopic response to the SES half-images seems to involve resolution of disparate subtense, that is, disparity in the size and shape of large-scale image features. The system splits the percept into two depth planes where visual projections cross (I have previously called this a split-projection configuration). Correspondence here looks similar to a Panum configuration where a single corresponding contour (luminance step) visible to both eyes (a near edge) aligns the two disparate visible edges of a distant plane (each eye's view of the edge occupying a different line of sight).

### 8.1.5 Surface spreading in the SES

This section examines the relationship between surface separation and surface spreading in the SES percepts. Sketches that describe surface separation at a single T-junction in the SES are presented in Fig 8.5. A three-term luminance T-junction is represented with important parts of the junctions labelled.

In Fig 8.5a the white homogeneous region bounded by the T-junction is labelled A, the black region D and the grey region B. Disparity at the T-stem means that the region at C is white in the left eye and grey in the right for crossed disparity. Region C is white in the left eye and grey in the right for uncrossed disparity.

Figure 8.5b describes the position of the perceived depth step, for crossed disparity. Contrast spreading seems to be generated in crossed fusion of the white-grey contour at the T-stem. Separation of surface layers happens when the grey region at C is displaced in depth to a near surface leaving C as a white monocular zone visible only to the left eye.

The system achieves surface separation by displacing the grey region to the depth at which the two lines of sight cross (ie. it stands the grey surface forward). This suggests that illusory contrast spreads from the *site* of surface separation across the T-top. The T-top is locked at a zero disparity depth plane. Hence the boundary of the illusory contrast generated by crossed disparity is a near surface layer separated from a distant plane, that is, a fold in perceived space that spreads from the grey side of the T-stem.

Figure 8.4c describes image processing where contrast spreading does not occur (uncrossed disparity). In fusing the disparate T-stem, both grey and white regions are assigned to a distant depth plane. There is no separation of surface layers along the T-stem. Instead, a depth step is achieved along both the black-white and the black-grey contour at the T-top.

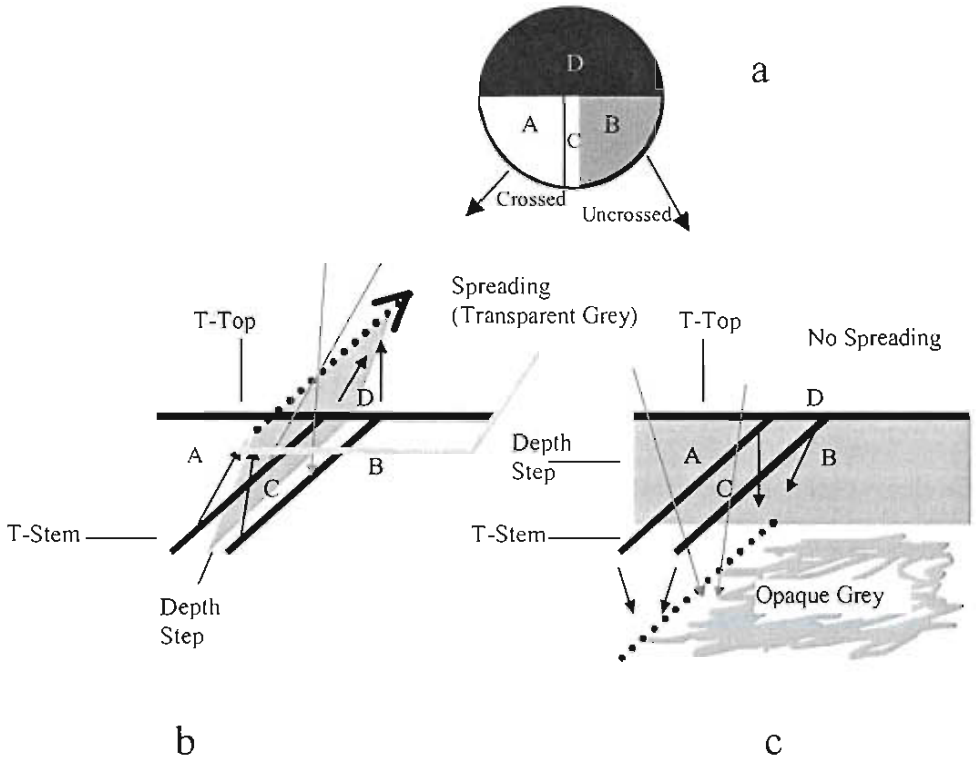



**Fig 8.5. Surface separation and perceptual asymmetry in the SES percepts**
This figure shows the way the system may assign surface layers at particular contours in the contrast spreading configuration. In (a) the local disparities at the T-junction are shown. For crossed fusion the region at C is white in the left eye and grey in the right. For uncrossed disparity C is white in the right eye and grey in the left. Fig (b) shows how surface separation is achieved in crossed fusion. The region C (grey in the right eye) is displaced to a near depth as the T-stem bounding region A and D is fused in a triangular configuration of projections. This yields a depth step along the T-stem. In Fig (c) a depth is created along the T-top since an uncrossed fusion configuration assigns both the white region A and the grey region D to a distant depth plane.

Three experiments, 7, 8 and 9 addressed the validity of this argument using the seen slant metric. The metric enables manipulation the relative size and shape of the grey cross in the SES half-images by drawing a grey square behind the four black panels. The shape and size of the grey square can be manipulated according to the theoretical rotation of a surface plane. In Experiment 7 the basic predictions of the BIPASS model were tested using SES half-images. Experiment 8 compared two differently shaped Ehrenstein figures to examine surface separation and spreading where all four T-stems generated binocular disparity. Finally, in Experiment 8 the mechanism of surface spreading and its relationship to confirmatory processes such as completion was explored.

## 8.2 Experiment 7 seen slant in SES percepts

Chapter 5 developed a descriptive and functional model (BIPASS) to help explain intriguing phenomenological aspects of the SKS percepts. The next section summarises a similar explanation of illusory contrast spreading and perceptual asymmetry in the SES percepts. Figure 8.6 revisits the basic BIPASS model applied to a SES and demonstrates the phenomenal outcomes that the model predicts.

Image registration and vergence lock enables image comparison, returning both parity and disparity values from the 2-D layout of the retinal images. Surface separation resolves disparity in the size and shape of large-scale image features and initiates surface spreading. Subsequently, confirmatory processing such as completion, continuation or Grossberg's Object Recognition System, impacts on the actual trajectory of the illusory 3-D contrast spreading. The model is summarised in Fig 8.6a.

Figure 8.6b shows the phenomenology predicted, in functional terms, by the BIPASS model for crossed disparity. Surfaces are separated into two layers at the T-stems (A_B and C_D). This generates surface spreading across the T-tops (eg. M_A and C_D). The near transparent grey layer spreads towards the adjacent non-disparate T-stems (E_F and G_H). There is no disparity at E_F and G_H however. This may make surface spreading somewhat unstable.

Figure 8.6c shows the phenomenology predicted by the BIPASS model for uncrossed disparity. Surfaces are separated into two layers at the T-tops (ie. along M_A_AC_S and N_B_D_T). no surface spreading emerges. The grey layer is assigned to the distant depth plane.

One way to initially explore the validity of the BIPASS model was to manipulate the subtense of the grey cross in the SES half-images in the same manner as in the SKS related experiments. A case has been made that contrast spreading arises, in the SES percepts, because of surface separation at the T-stems. Separation of surface layers along the T-top yields no contrast spreading. An exploratory study was designed in which these two effects might arise in the same 3-D percept in perception of a slanted SES. That is, by manipulating the binocular subtense between T-stems perception of a part modal/amodal Ehrenstein figure whose orientation reflected the magnitude of retinal disparities applied, was anticipated.

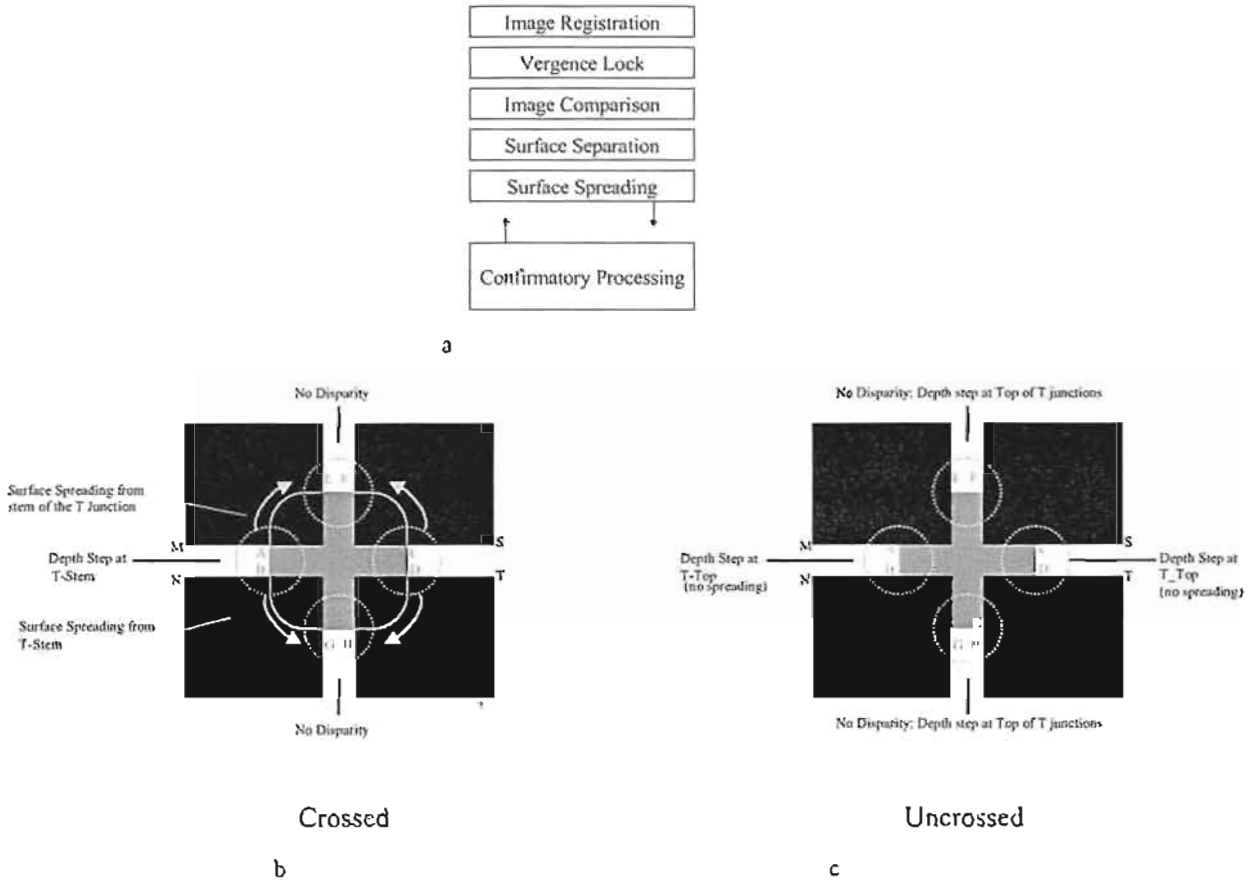



**Fig 8.6. A BIPASS model of the perception of contrast spreading effects in the SES**
The BIPASS model is described in (a). Applied to the SES percepts, the BIPASS model predicts the phenomenal outcomes when half-images are fused. For crossed disparity, shown in (b), surface separation at the disparate T-Junctions initiates spreading of a near surface layer across the homogeneous luminance toward non-disparate T-Junctions. In (c) disparities at T-Junctions indicated are uncrossed. This means that the surface layers separate at the T-tops no contrast spreading arises.

The BIPASS model is the proposal that, in functional terms, the system generates surface spreading at T-stems in which crossed point-disparity is present between the image pairs. It also predicts that when T-stems possess uncrossed point-disparity there will be no surface spreading generated. When the subtense of the arms of the grey cross is disparate (but yielding no standing disparity) surface spreading will be generated at one end of the cross and not the other. This is because the grey surface layer will appear to slant through the P Plane.

Figure 8.7 is used to help explain the logic of the BIPASS model for a stereoscopically rotated SES percept. The diagrams demonstrate how half-images in Experiment 7 were constructed and shows the predicted phenomenology arising from binocular fusion of the transformed image pairs.

Figure 8.7a is a projection diagram along a single epi-polar line across the top of the T-Junctions in the SES. This diagram predicts that disparate subtense between the T-stems will yield a stereoscopically rotated SES percept. This figure shows seen slant in the vertical slant-axis.

In Figs 8.7b and c, the position of the T-stems has been manipulated by drawing a grey square behind the four black panels. A magnification factor appropriate to an arbitrary degree of theoretical rotation was achieved by transforming the size or shape of the grey square in each image (manipulating the subtense or shear at T-stems).

Figure 8.7b shows the half-images simulating vergence lock, ie. with the left image mapped over the right. This highlights the relative magnification of the horizontal arms of the central grey cross to induce rotation in the vertical slant-axis (Just two T-stems carry any point-disparity information). It was predicted that when these pairs were fused, the horizontal arms of the cross would stand forward of, say, the left side of the figure when local disparities at the T-stems were crossed and behind the black panels when uncrossed. It was anticipated the subjects judgement of seen slant would reflect the degree of relative magnification applied to the grey cross. The predicted phenomenology is shown in Fig 8.7d.

Figure 8.7c demonstrates manipulation of the SES that we expected would yield seen slant in the horizontal slant-axis (again the left image overlays the right). To achieve these manipulations the central arms of the grey cross were transformed by applying horizontal shear to the T-stems. Figure 8.7f describes the phenomenology predicted in that case.

Figures 8.7e and 8.7f show that the SES half-images for each slant-axis will yield disparity at only two of the four arms of the grey cross. This suggests that spreading of the near surface layer will be in conflict with the slant response. In particular, T-stems at the vertical arms of the cross in Fig 8.7e and 8.7f are non-disparate. The BIPASS model predicts that stereopsis will not directly contribute toward surface separation and surface spreading at these T-stems.

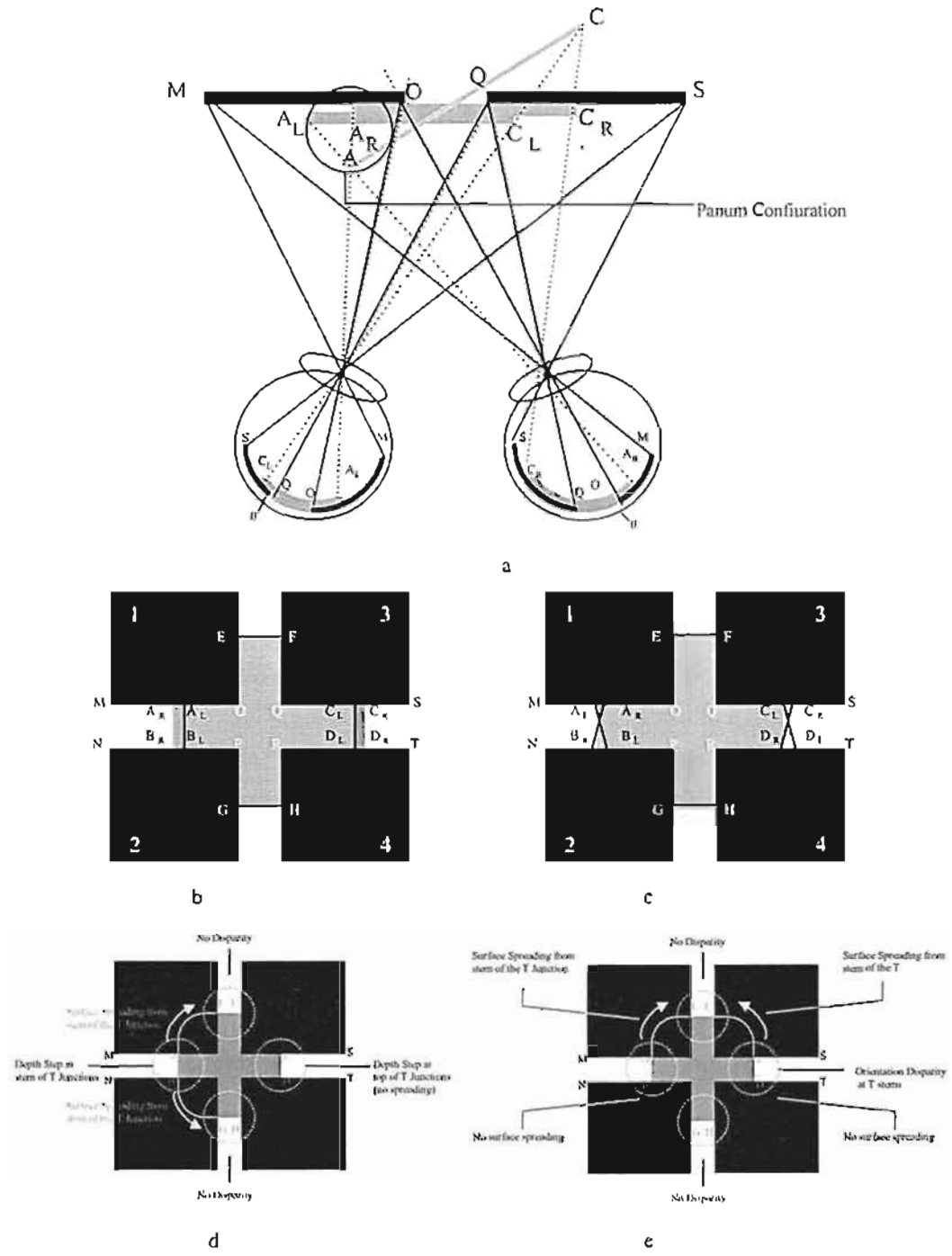



Fig 8.7. Making half-images to produce a slanted stereoscopic Ehrensteln percept
In (a) a projection drawing demonstrates that when local disparities at the T-stem $A_R$-$A_L$ are crossed and the local disparities at the T-stem $C_R$–$C_L$ are uncrossed the grey surface layer will look slanted along A-C. Note that the subtense between T-stems $A_R$_$C_R$-$A_L$_$C_L$ and the subtense along M_$C_R$ and M_$C_L$ are immutably reciprocal to those local disparities. Fig (b) shows the half-images where the right eye's view of the horizontal cross arms has been magnified. Fusion of these half-images should yield stereoscopic rotation of the cross arm. This predicted the phenomenology described in (d). That is, stereoscopic rotation in the vertical slant-axis and spreading of the surface from the near layer. Fig (c) shows the half-images where the each eye's view of the horizontal cross arms has been sheared. Fusion of these half-images should yield stereoscopic rotation of the cross arm. This predicted the phenomenology described in (e). That is, stereoscopic rotation in the horizontal slant-axis and spreading of the surface from the near layer.

*Measures*

Experiment 7 was an exploratory study of seen slant generated by fusing disparate SES half-images leading to a part-modal / part-amodal SES percept. In this experiment, disparities were applied to the SES to simulate the retinal images cast by a rotating grey square partly obscured by four lack panels. Hence, Ogle's magnification factor was applied to the grey square then transformed to yield theoretical rotation of the square at 0, 20, 30, 40 and $50^0$ about both slant-axes. The hypothesis was simply that the seen slant of a part-modal/amodal SES percept would reflect the magnitude of theoretical rotation.

## 8.2.1 Method.

*Subjects*

Ten subjects from the participant pool took part. All were screened and subjected to the same practice sessions as before.

*Stereograms*

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a set of four black panels subtending about $3^0$ (at 750mm viewing distance). The black panels (0.09cd $m^{-2}$) were positioned so that a square drawn through their centres would subtend $7^0$ degrees. A central grey cross was created by drawing a grey square (0.4cd $M^{-2}$) so that it obscured the four black panels. The grey square overlaid the black panels symmetrically (equal intrusion on all sides) intruding ½ the width of the black panels. The grey square was then stood behind the black panels against a background luminance of 0.7 cd $m^{-2}$.

The shape and size of the grey square could be manipulated to generate appropriate disparity at the grey cross seen by observers in the SES. So the grey square was geometrically transformed according to the monocular transformations of a square shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was then introduced by applying ½ Ogle's M to the grey square, symmetrically and in opposite signs in each eye's view of the grey cross (see Fig 8.8).


**Fig 8.8. Patterns of disparity in the SES**
Fig 8.8a shows one half-image supporting rotation about the vertical axis. Disparity was applied by increasing the magnitude of the grey square (α) in one eye relative to the other. Fig (b) shows one half-image supporting rotation about the horizontal axis. Disparate shear (λ) was applied to the grey square in an opposite direction for each eye's view while constraining (α).

*Design and procedure*

A two-way (5x2) repeated measures design was used to explore the effects of theoretical rotation ($0,20,30,40,50^0$) and slant-axis (Horizontal / Vertical) on the estimated slant of a grey cross in the Ehrenstein figure. Six complete repetitions of this design were used, making a total of 60 trials per subject in all. Details of the procedure were as described in section 4.5 above.

## 8.2.2 Results and discussion

A two-way (5x2), repeated measures analysis of variance examined the effects of theoretical rotation (0, 20, 30, 40, and $50^0$) and slant-axis (vertical, horizontal) on seen slant. Obtained slant estimates were averaged across repetitions.

Seen slant varied predictably with theoretical rotation, as shown in Fig 8.9. The effect was significant: $F(4,9) = 31.44$, $p < 0.001$. The slant-axis comparison did not approach significance however. Figure 8.9 shows that slant estimates in the SES were largely consistent with theoretical rotation but were substantially attenuated. Subjects' slant estimates were not accurate, or at least, the slant estimates did not precisely reflect the theoretical rotation of the Ehrenstein square. Clearly, the response of the system involved processing disparate subtense and/or orientation of the central grey cross. Subjects tended to underestimate the slant of the Ehrenstein figure at larger disparities (compared to the theoretical rotation of $0,20,30,40,50^0$). This result provided an interesting exploratory insight into the mechanisms that underpin perception of the SES.

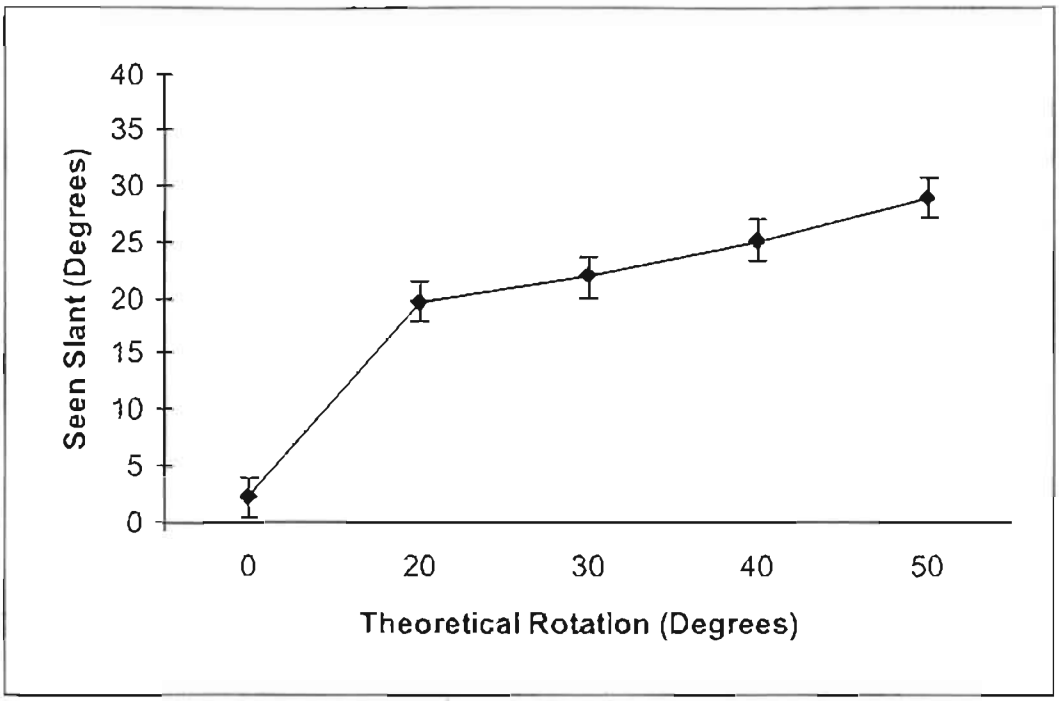
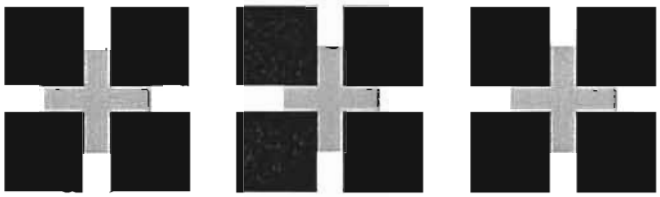



**Fig 8.9. The impact of theoretical rotation upon seen slant in a SES**
This figure plots the mean slant estimates for levels of theoretical rotation applied to a grey square drawn behind four black panels in a SES configuration.

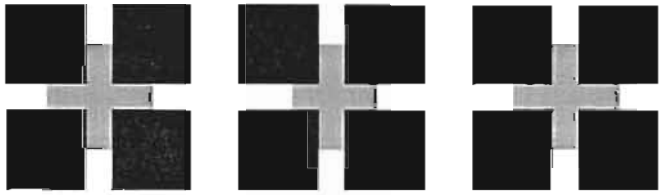
Predictable seen slant in a part-modal / part-amodal SES percept demonstrates the simple but important theoretical point that the stereoscopic response to SES half-images is related to resolution of the 2-D layout of the half-image configurations (ie. the stereoscopic response seems not to be merely a heuristic response to local occlusion cues). Moreover, the results suggest that a computation of 3-D form must be more complex than mere filling-in of computed disparity values. The simple BIPASS model anticipates this. Image comparison mechanisms return a set of retinal parity as well as disparity measures and conflicts can emerge between measures occurring at different feature-scales. As we have seen, this is particularly the case at horizontally aligned luminance contours (where there is ambiguity between allocation of image differences to either surface separation or surface orientation). The slanted percept is likely to be a resolution of image disparity and parity information over time and proceeds in such a way that contrast spreading can cross depth planes.

Some examples of the stimuli presented to subjects are shown in Fig 8.10. The shapes that emerge at the grey cross may look "deformed". Similarly, subject's slant estimates suggest that seen slant is an aggregate of disparity processing and surface perception. This

seems particularly strongly evident in the percepts derived from fusion of the pair at Fig (c). The figure presents disparities appropriate seen slant in the horizontal slant-axis[21].



a - 10% relative magnification



b - 20% relative magnification

c - 20% relative horizontal shear

**Fig 8.10. Stereoscopic rotation of an Ehrenstein figure**
Crossed fusion of L-M and M-R in Fig 8.10a yields the percept of a grey layer rotated in the vertical slant-axis. At the near end of the grey layer, the layer looks transparent and appears to spread across the grey squares. The shape of the grey layer will appear uneven - not a uniform plane. Deformation of the shape is likely to be due to interaction between local disparity processing, a tendency to complete the figure and the resulting size-distance scaling. In (b) seen slant is increased by increasing the magnification factor. In (C), crossed fusion of L-M and M-R yields stereoscopic rotation in the horizontal slant-axis is shown. In all of these pairs there is no disparity at the ends of the vertical arms of the cross.

The BIPASS model was developed to explain some of the remarkable phenomenology of the SKS percepts. The model hinges on the likelihood that fusion of untextured half-images recovers a more thorough metric of inter-retinal differences than local point disparities alone. The system appears to utilise disparate subtense across large feature scales. Disparate subtense appears to constrain the manner in which the system organises the 3-D illusory percepts as seen in this experiment. That is, it contributes to surface separation and perception of the rotation of an illusory 3-D surface layer through the P Plane.

---

[21] Note that when these stereograms are printed on paper, the texture of the paper seems to weaken the effects that we observed in the laboratory setting, under dim light and with quite large scale figures on a high-quality (and polished) monitor and viewed with shutter valve goggles.

This analysis differs from Anderson, Nakayama, and their colleagues who take the Surface Heuristic approach. These authors' theories are grounded in perceptual inference. Their theories are designed to overcome the paucity of local matches in untextured stereograms. However, the 3-D percepts achieved from untextured stereograms are not so underdetermined by 2-D image differences as Anderson, Nakayama, and their colleagues have assumed. Consider this excerpt from Anderson (1997, p419):

> One of the most challenging problems facing theories of perceptual organisation is in determining the surface properties that generate image structure ... the term 'surface properties' refers to the underlying causes of the image formation process, such as illumination conditions surface reflectance opacity and shape. Recovering these surface properties is difficult because the mapping of the 3-D scenes onto 2-D images is many-to-one, which implies that there is no simple means of inverting the image formation process. Yet, we rarely have phenomenological access to this ambiguity; the properties of the world usually seem unambiguous and stable. This apparent stability in the face of noninvertability suggests that the system imposes heuristics, rules, or constraints to recover surface properties. If this logic is correct, then a fundamental problem in understanding perceptual processing is to determine what these heuristics are, and how they shape our perceptual experience of surface structure.

Anderson argues against Nakayama's view that stereopsis in ambiguous conditions is driven by *inverse ecological optics*. Recall Nakayama and Shimojo's account of the Principle of Generic Sampling (see Chapter 3 for a review of the POGS). This was the idea that in the face of sparse matching primitives interpretation of untextured stereograms was biased toward a generic 3-D view. That is, critical features of the stimuli triggered learned responses in particular neural populations. The results of this experiment suggest that a quantitative relationship may exist between surface separation and disparate subtense at just the horizontal arms of the cross. In the next experiment stimuli were prepared in which surface separation was predicted at all T-stems.

In summary, the observation that seen slant was substantially attenuated relative to theoretical rotation suggests that the presence of non-disparate T-stems at the vertical arms of

the cross may conflict against the perceived rotation of the cross. In other words, the degree of seen slant may have been impacted by conflict between two interpretations of disparity (large scale => slanted verses local-scale => fronto-parallel). This experiment provides an initial insight into the relationship between the 2-D layout of the retinal images and the 3-D perceptions developed. The next experiment more directly tests this reasoning.

## 8.3 Experiment 8 Contrast-spreading and seen slant in a diamond and a square

This experiment compared seen slant in two differently shaped Ehrenstein figures. Seen slant in a SES was compared to a stereoscopic Ehrenstein diamond (SED). Differences between the two shapes are theoretically important for understanding the way stereopsis accesses and utilises inter-image differences in untextured stereograms. In particular, the SED raises the question of vertical and oblique images differences and their importance in the separation of surface planes in the 3-D illusory contrast spreading percepts.

### 8.3.1 Vertical inter-retinal differences and separation of surface layers in a SES

The Surface Heuristic approach to the 3-D illusory percepts emphasises the importance of unpaired features of the two half-images. Anderson, in particular, has placed importance on vertical image differences. He claims that unmatchable parts of contours at contour junctions reveal partial occlusion geometry. Since vertical image differences at luminance junctions are unmatchable, they might disambiguate occlusion geometry and so facilitate surface scission (separation of surface layers).

In Anderson's arguments, the central problem the system faces is determining which contours, at a contour junction, fit to which depth layer (in the eventual 3-D percept). The system can then work out which luminance value belongs to which contour and which surface quality belongs to which luminance value. That is, the system interprets a particular region of a particular luminance as opaque or transparent. In order to do this the system has to have *a priori* knowledge of how contours and unpaired features (half-occluded features) fit together in the natural world.

I concur with this view of a half-occluded region as a geometric principle, but, I have misgivings about Anderson's thinking on perceptual inference mechanisms. To discuss the issue in some detail an anecdote from Anderson is presented below (Anderson, 1997, p437):

One geometric fact of binocular contour junctions generated by one surface occluding another is that they generate both horizontal and vertical displacements in the two eyes (except for occluding contours that are perfectly horizontal relative to the line of sight ...). This can be appreciated immediately by performing the following exercise. Holding your head upright , place your left hand in front of your head and point your index finger 45 degrees upwards (ie roughly towards 2 0'clock). Now place your right hand behind the left , and point your right index finger upwards and to the left - 45 degrees. Arrange your hands so that your fingers appear to project an X-shaped image to each eye, while keeping the two fingers separated in depth with the right finger behind the left. Now alternatively open and close your left and right eye, and observe how the V-shaped junctions appear to shift horizontally and vertically as the left and right eyes alternately open and close. Note also that the vertical shifts in these junctions alternatively occlude and reveal portions of the far finger ... we describe unmatched features on a partly occluded surface as half-occluded. If the visual system was capable of detecting these half-occluded features, then they could potentially provide information about the geometry of occlusion.

One simple method for recovering unmatchable features is to restrict matching to epi-polar lines, ie. One dimensional, horizontal "slices" of the two image planes ...Specifically, we have shown that the visual system appears to use an epi-polar matching constraint to determine which features at contour junctions are half-occluded ... which in turn are used to recover the geometry of the occluding surfaces (through the formation of illusory contours).

Let us clarify some of these issues with reference to the present thesis.

*Binocular geometry and vertical image differences*

The unpaired regions Anderson talks about are indeed a geometric product of binocular parallax. Nevertheless, binocular parallax generates a great deal more information that Anderson does not mention (when he talks about untextured stereograms). These unpaired features are important as they are part of the overall 2-D layout of image differences. However, they are not the whole story.

*Local and non-local components of stereopsis in a finger X*

Consider this: suppose that you were holding your fingers in an "X" as Anderson describes. Imagine that some object obscured just the "finger junctions" (so you couldn't see the fingers actually cross). Would the perceived depth relations at the fingers look any different? Obviously not. Also, if you rotate the crossed fingers, move one finger relative to the other, or move your eyes and then your head the unpaired region on the partly occluded finger will not necessarily remain vertical. Depth relations will be maintained.

An inter-image property available, and stable, across all of these contexts is the relative proportions of parts of the occluding and partly occluded finger, that is, disparate subtense. It is also purely a geometric fact that the difference in subtense in each eye's view of a partly occluded region must equal the magnitude of the half-occluded region on that surface, vertical or otherwise.

Contours that are horizontal to the line sight are a rare case in which the system will be vulnerable to ambiguity in the partial occlusion geometry. In natural vision, horizontal contours will be the exception not the rule. Contours that are oriented in any manner other than perfectly horizontal to the line of sight, seen with two eyes, will have disparate subtense at partial occlusion that is also a product of binocular parallax due to their orientation.

The BIPASS model has suggested that point-disparity computation should not be considered to be an isolated visual event but occurs in the context of reciprocal patterns of disparate subtense. It has been suggested that the system may not need to literally detect half-occluded regions to recover occlusion geometry. The linkage between disparate subtense and point-disparity processing, at partial occlusions, seems to be the mechanism labelled surface separation.

*Surface (finger) separation and disparate subtense where vertical image differences arise*

In each 2-D retinal half-image, the contour bounding the occluding finger (in Anderson's example), as it crosses the occluded finger, occupies a different position because of binocular parallax. The most basic principle of stereoscopic vision is that fusion of those disparate contours yields a singular boundary of the near finger.

An interesting problem for the system arises here. Anderson and Julesz (1995) asked how the system could differentiate contours on a surface from contours at boundaries? To Anderson and Julesz this was another example of the uncertainty created by a paucity of

binocular matches in untextured stereograms. In the finger-cross example, the question applies to, say, disparate contours at the occluding finger joints as compared to contours at the occluding finger bounds. They argued that detection of unpaired vertical image differences by decomposition of the images into matchable and unmatchable features was a solution to this uncertainty.

This is possible in principle, of course, but something else is different about boundary contours (where the fingers cross in Anderson's example). In fusing Anderson's finger X, the boundary of the occluding finger *must* align the visible boundary of the partly occluded finger along the same line of sight in each eye. This is also an immutable geometric principle. Because, in the 2-D images, the partly occluded finger starts precisely when the occluding finger ends. The two surface objects must *share* the contour.

So, while Anderson is concerned with which surface *owns* contours at partial occlusion, what differentiates boundary contours is that boundary contours are *shared* between the two depth planes. Arguably, no surface exclusively owns the contour. It appears that vergence lock can constrain cyclofusion to achieve contour sharing. It seems surface separation can be achieved, in untextured stereograms when the system assigns a depth step along a contour that is shared between an occluding and occluded depth plane.

This was precisely the point of *surface separation*, at a split-projection configuration, in the SES. In crossed fusion, the contour (the luminance step) bounding the near finger (in Anderson's example) is fused at the intersection of a pair of visual projections (lines of sight) between which, half-occluded parts of the distant finger occur.

Since, the differences in binocular subtense at the distant depth plane[22] are monocular zones, it was shown in the unpaired SKS that disparate subtense may itself be sufficient to yield stereoscopic depth. These figures also showed that monocular features can occur on partly occluding surface layers in those unusual circumstances. In that case the requirement of a region being unpaired was not a valid heuristic to guide surface separation.

There are certainly vertical image differences present when a SED is constructed using a diamond drawn behind the four black panels to shape the grey cross. Figure 8.11 demonstrates a SED and identifies the vertical image differences that are apparent in that

---

[22] By resolution of disparate subtense we mean that in achieving a singular (cyclopean) viewpoint combination of the retinal images must yield a view in which the sum of the subtense at each depth plane must fit into the singular view. A difference in subtense between two singular contours can either define a surface oriented across several depth planes or a depth step. A depth step resolves disparate subtense by separation of surface layers into

shape, but not in the SES. The vertical image differences arise in the SED because of the oblique orientation of the T-stems. In the SES only two T-stems generate disparity while in the SED all four T-stems generate disparity. This is just an accident of the background shape used to draw the Ehrenstein cross in both stimuli.

Fusion of the oblique T-stems in the SED means that the system must achieve cyclofusion where the actual lengths of the T-stems are different. Their orientation, however, must be constrained by their relative orientation. This is not the case in the SED. The perception of partial occlusion at both signs of disparity must therefore be intimately related to resolution of the disparate subtense of the grey, black and white regions that are constrained within the 2-D structure of luminance contours.

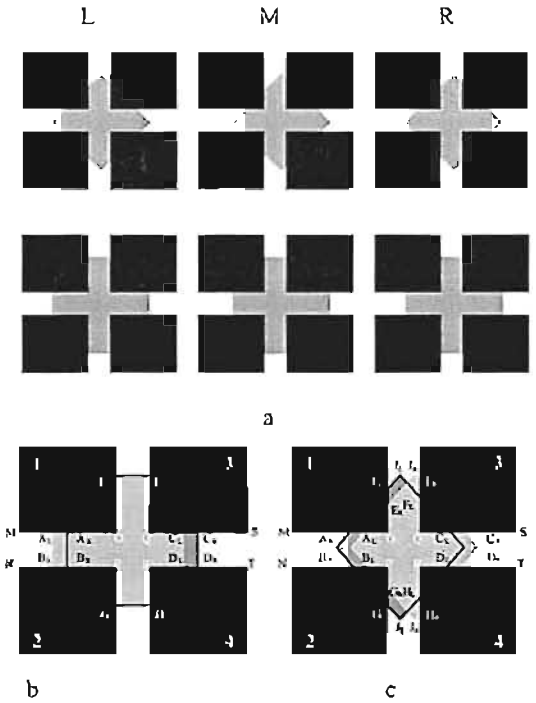



**Fig 8.11. Local disparities in Ehrenstein figures**
In (a), stereograms representing a square shaped and a diamond shaped Ehrenstein figure are shown. Both have precisely the same horizontal disparity. Crossed fusion of the pairs at L-M sees a layer standing forward of the black panels with contrast spreading across the corners of the black panels. Crossed fusion of the M-R pairs poses the reversed disparity sign. A grey layer stands behind the black panels. No contrast spreading is seen. In (b) the half-images comprising the square shape are mapped over each other. There is no disparity at EF or GH. While in (c) the half-images comprising the diamond are mapped onto each other. There are vertical image differences exist at the side-on T-Junctions (along the T-top) $E_L$_$E_R$, $F_L$_$F_R$, $H_L$-_$H_R$, and $G_L$_$G_R$. Along the T-stems vertical image differences occur at the bottom pair of T-Junctions: $J_L$_$G_L$ - $J_R$_$G_R$ and $J_L$_$H_L$ - $J_R$_$H_R$; and at the top of the grey cross at $I_L$_$E_L$ - $I_R$_$E_R$ and $I_L$_$F_L$ - $I_R$_$F_R$. Note that there are no vertical image differences at the T-junction in the horizontal arms of the cross.

two depth planes.

In summary, I have three misgivings about Anderson's ideas in relation to the stereoscopic Ehrenstein figures:

1. They do not explain the perception of a depth step and illusory contrast spreading where no vertical image differences arise. That is, only horizontal image differences are present in the SES along the horizontal arms of the grey cross.
2. Focus on an epi-polar constraint does not predict the quantitative nature of stereopsis that exists in separation of transparent surface layers when the system must fuse contours at two depth-planes simultaneously (as in Experiment 7).
3. By emphasising disambiguation of presumably ambiguous point-disparities in generation of surface separation, Anderson does not allow for one very simple non-local aspect of stereopsis that appears important, namely, disparate subtense.

The next experiment was designed to test the ambiguity of the 3-D illusory percepts using the seen slant metric.

### 8.3.2 Seen slant as a metric of interaction between local disparities and disparate subtense

In the absence of texture, as Anderson explains, *contours horizontally aligned relative to the line of sight* can be a source of ambiguity in stereopsis (see also Anderson and Julesz, 1995). I would argue that this is because it is difficult for the system to differentiate whether a difference in subtense along some horizontal contours is due to the contour's orientation or another surface partly occluding the contour.

In the retinal images, at any 2-D orientation, other than horizontal, binocular fusion of contours can potentially signal 3-D orientation and/or relative depth. This is true of oblique contours bearing zero orientation disparity. The same is true of horizontally aligned contours in the SES. Horizontally aligned T-stems in the SES (E_F and G_H in Fig 8.12c) cannot be informative about depth. There is simply no disparity created at those T-stems. The SES half-images may be ambiguous because two T-stems signal disparity and two do not.

An SED possesses no disjunction between disparities at the T-stems. All four T-stems manifest point-disparity consistent with rotation of a surface. Therefore the depth and orientation of each T-stem in the SED relative to the black panels would be expected to be recovered accurately by the surface separation mechanism.

To explore the implications of these two shapes upon disparity processing the grey cross at the centre of SES and SED half-images was transformed using precisely the same model of perspective projection. Each eye's view simulated the viewpoint arising if a planar

square or diamond was rotated in the vertical slant-axis. This was achieved by altering the horizontal magnitude of the square and diamond in each eye's view to a degree defined by Ogle's relative magnification corrected for perspective projection[23]. The image differences are described in Fig 8.12a and Fig 8.12b. An example of these manipulations as they are physically manifested in the central grey cross is given for a square in Fig 8.12c and diamond in Fig 8.12d.

[23] We only manipulated horizontal image magnitudes

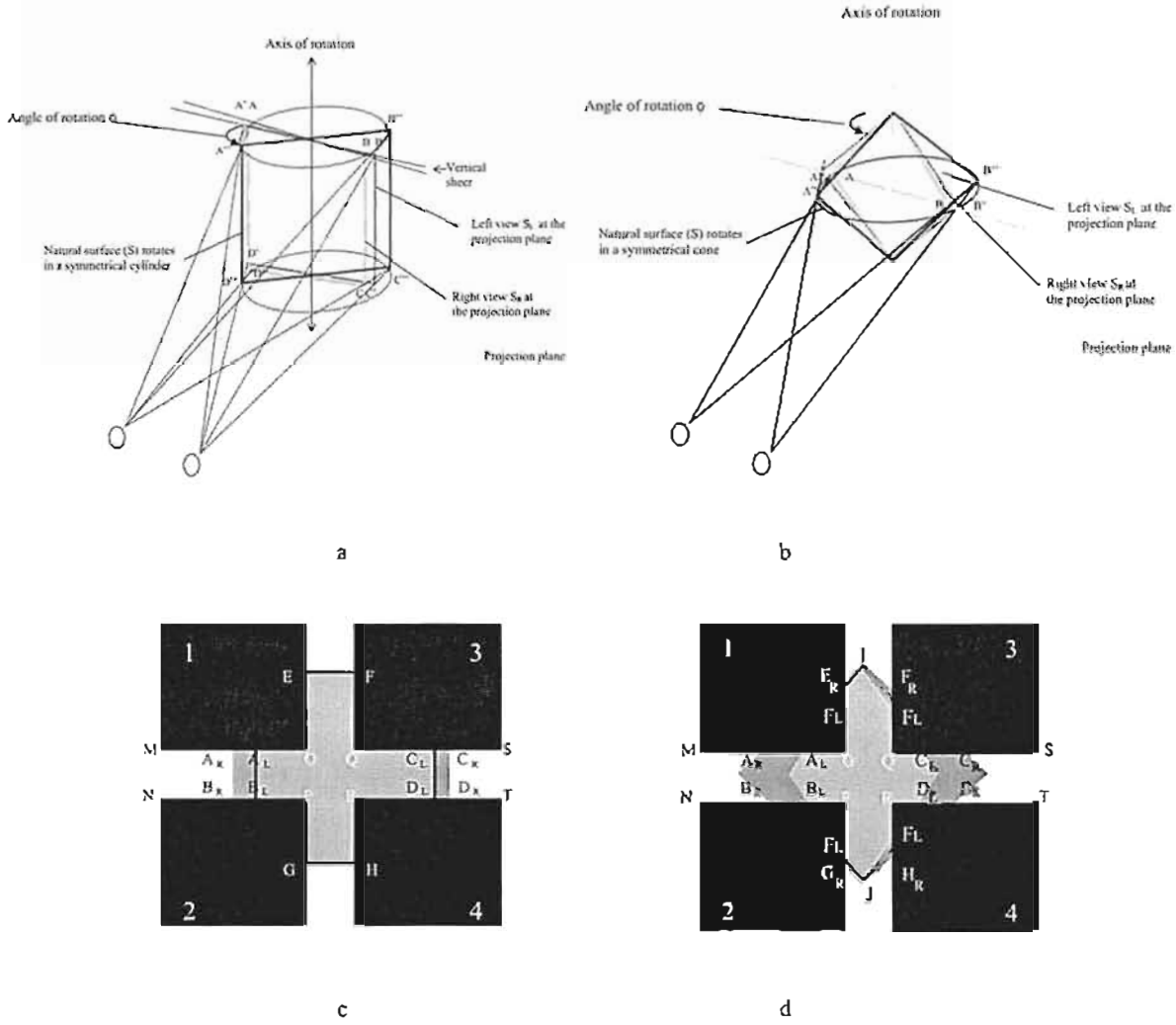



**Fig 8.12. Binocular parallax and stereoscopic slant in a diamond and square**

In (a) a square surface ($S_C$) whose boundary corner points A"B"C"D" will describe a cylindrical arc as it rotates in the vertical slant-axis. At a particular degree of rotation, each eye's view the shape cast by the surface at the distance of the projection plane will be a product of binocular parallax. In the right eye the image cast by the square ($S_R$) will subtend ABCD, and in the left view ($S_L$), its image will subtend A'B'C'D'. In (b) the same is the case for a diamond. But it rotates in a bi-conical cylinder. Half image that simulate each eye's viewpoint can be generated by taking transverse slices through these abstract figures (a) and (b) at the distance of the projection plane. When the black panels of the Ehrenstein figure are drawn to obscure the disparities described by (a) and (b) a pattern of 2-D image differences are achieved. An example of the manipulations is shown for both the Ehrenstein square in (c) and Ehrenstein diamond in (d). In (c) relative magnification results only in local disparity at the central grey cross occurs at the vertical T-stems $A_R\_B_R$ - $A_L\_B_L$ and $C_R\_D_R$ - $C_L\_D_L$. In (d) relative magnification results in orientation *and* disparate subtense. At T-stems: $I\_F_R$ - $I\_F_L$, $I\_E_R$ - $I\_E_L$, $J\_G_R$ - $J\_G_L$, $J\_H_R$ - $J\_H_L$. Disparity and disparate subtense arises at the T-stems at the horizontal arms of cross as well.

*Measures*

In summary, this experiment explored seen slant in a stereoscopic percept generated by projecting a pair of squares or diamonds of homogeneous intermediate luminance to their surrounds (black squares on a white background) behind four non-disparate black panels (a typical Ehrenstein configuration).

The square or diamond in each half-image was transformed to simulate the horizontal image differences cast by a square or diamond rotated in natural perspective that stood at various depths above, equidistant to or behind the projection plane (on which stand black panels). This was achieved the central cross in the Ehrenstein figures appropriate to the Keplerian projection arrays in Fig 8.13. Figure 8.13 describes the geometric manipulations applied to the half-images diagrammatically.

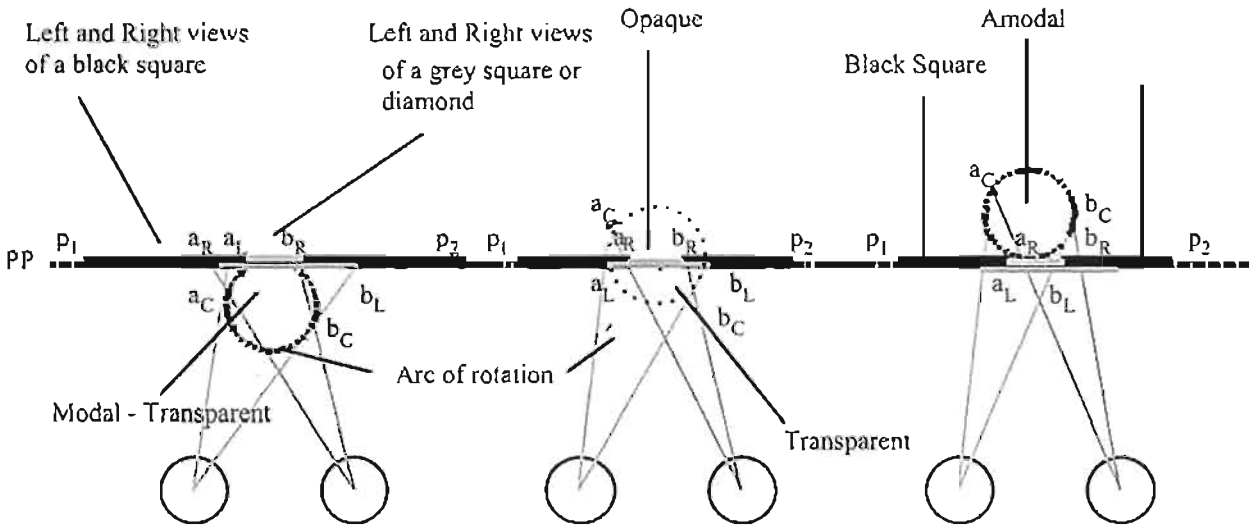



**Fig 8.13. Two components of disparity in a slanted Ehrenstein figure**

These projection diagrams show schematically the manipulations of binocular disparity achieved by projecting a grey square or diamond behind four black panels to create a disparate cross in each half-image. In (a) by appropriately manipulating disparate subtense at the cross (by standing disparity and relative magnification), it was predicted that fusion of the left half-image at $a_R$ and $a_L$ and $b_R$ and $b_L$ would yield a modal transparent Ehrenstein square or diamond whose right and left bounds stood at the cyclopean points $a_C$ and $b_C$ respectively. This was expected to result in a fully modal stereoscopic Ehrenstein shape rotated about in the vertical slant-axis. A similar relative magnification at (b) but with no standing disparity predicts a illusory contrast spreading at $b_C$ with the grey surface slanted through the projection plane so that $a_C$ defined the depth of the opaque end of the surface. In (c) the bounds of the square or diamond fused to yield the cyclopean points $a_C$ and $b_C$ would stand the figure behind the projection plane, predicting a stereoscopically rotated opaque surface.

In the SES, the central cross contained certain T-stems that were not informative of stereoscopic rotation. Therefore the hypothesis was that seen slant in the SED would be attenuated less than in the SES given that precisely the same large-scale disparity transformation was applied to both the square and diamond shapes. The key comparison in this experiment was therefore a first order interaction between the background shape and theoretical disparity. Planned comparisons were to provide a direct comparison of mean seen slant differences between each shape at each level of theoretical rotation for the square and diamond.

### 8.3.3 Method

*Subjects*

Ten subjects were drawn from the department pool of volunteers. They were asked to complete the same practice trials as in previous experiments.

*Stereograms*

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a set of four black panels subtending about $3^0$ (at 750mm viewing distance). The black panels (0.09cd $m^{-2}$) were positioned so that a square drawn through their centres would subtend $7^0$ degrees. A central grey cross was created by drawing a grey square (0.4cd $M^{-2}$) or a grey diamond so that it obscured the four black panels. The grey square or diamond overlaid the black panels symmetrically (equal intrusion on all sides) intruding about ½ the width of the black panels. The grey square or diamond was then stood behind the black panels against a background luminance of 0.7 cd $m^{-2}$.

The shape and size of the grey square or diamond could be manipulated to generate appropriate disparity in the SES. The grey square and diamond were geometrically transformed according to the monocular transformations of a square shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was then introduced by applying ½ Ogle's M to the grey square, symmetrically and in opposite signs in each eye's view of the grey cross (see Fig 8.14).

Standing disparity was then introduced by shifting the square or diamond shape in opposite directions sufficient to generate +20, 0 and -20 arcmins of retinal disparity.


a SES b SED

**Fig 8.14. Patterns of disparity in the SES**
Fig 8.14 shows one half-image supporting rotation about the vertical axis for the SES (a) and SED (b). Relative magnification was applied by increasing the magnitude of the grey square ($\alpha$) in one eye relative to the other. Relative magnification was applied to the diamond in the same way - by increasing the magnitude of the grey square ($\alpha$) in one eye relative to the other. Standing disparity was applied by shifting the square or diamond horizontally relative to the black panels while maintaining $\alpha_1$.

*Design and procedure*

A three-way (2x5x3) repeated measures design was used to explore the effects of configuration shape (diamond, square) theoretical rotation ($0,20,30,40,50^0$) and standing disparity (+20, 0, - 20 arcmins) on the estimated slant of a grey cross in the Ehrenstein figure. These disparities were applied appropriate to stereoscopic rotation about the vertical slant-axis only (i.e. only horizontal magnification was applied to the grey cross). Three complete repetitions of this design were used, making a total of 72 trials per subject in all. Details of the procedure were as described in section 4.5 above.

## 8.3.4 Results and discussion

A three-way (2x5x3) repeated measures analysis of variance was used to explore the effects shape (diamond, square) and theoretical rotation (0, 20, 30, 40, and $50^0$) and standing disparity (+20, 0, - 20 arcmins) on seen slant. Obtained slant estimates were averaged across repetitions.

It was fond that seen slant varied with theoretical rotation and the effect was significant: $F_{(4,9)} = 107.884$, $p < 0.0001$ (see Fig 8.15).

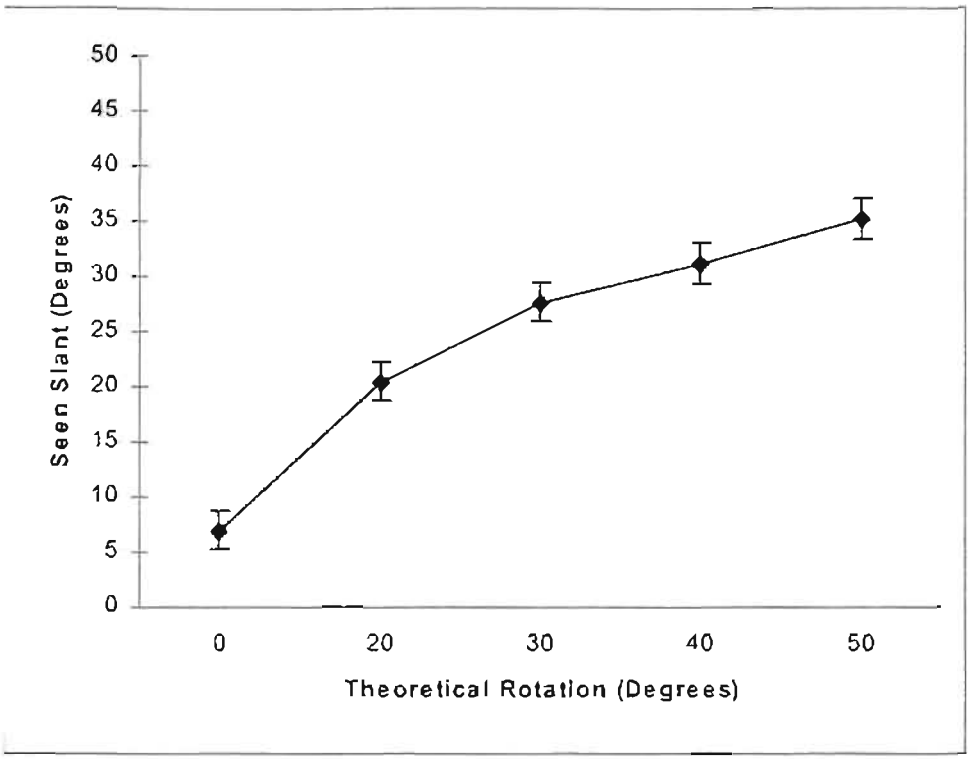



**Fig 8.15. The effect of theoretical rotation upon seen slant in stereoscopic Ehrenstein configurations**
This figure plots mean seen slant for levels of theoretical rotation applied to the Ehrenstein configurations. Standard error bars have been included.

It was also found that standing disparity affected seen slant (see Fig 8.16) and that effect was significant $F_{(2,9)}= 22.597$, $p < 0.001$. At a standing disparity of +20 arcmins (modal form of the SES or SED), seen slant averaged about $7^0$ greater than standing disparity at -20 arcmins (amodal form of the SES or SED). Across the different percepts substantial seen slant was evident however.

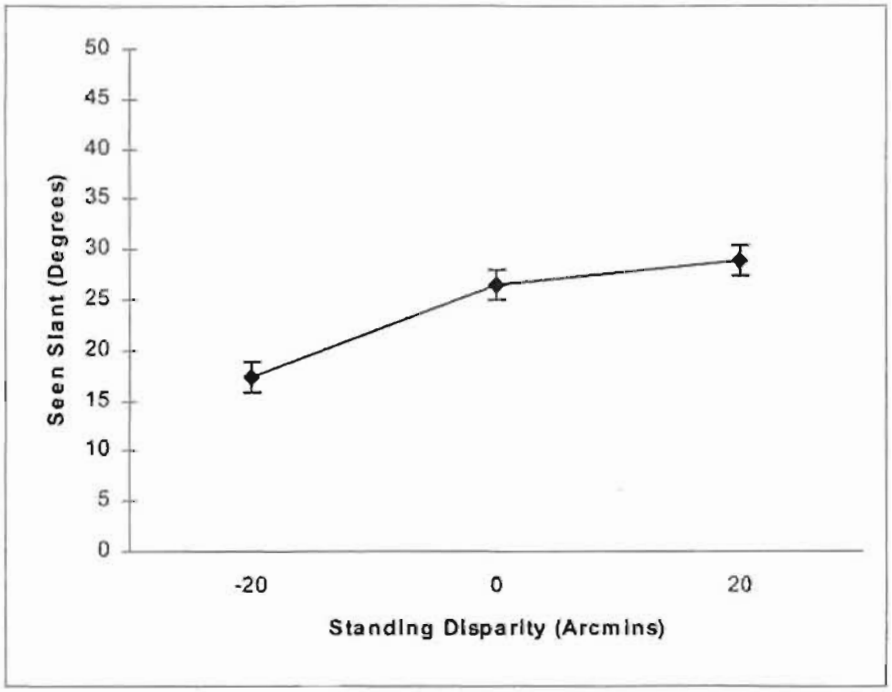



**Fig 8.16. The effect of standing disparity upon seen slant in stereoscopic Ehrenstein configurations**

This figure plots mean seen slant for levels of standing disparity applied to the Ehrenstein configurations. -20 arcmins standing disparity is associated with an amodal appearance while 20 arcmins is associated with the modal percept. 0 arcmins standing disparity is associated with a part modal and part amodal appearance.

As predicted, seen slant in the SED configuration was somewhat greater than the square and the effect was significant, $F_{(1,9)} = 7.43$, $p < 0.05$ (see Fig 8.17). Seen slant in the SED was about $7^{\circ}$ greater than the mean seen slant for the SES.

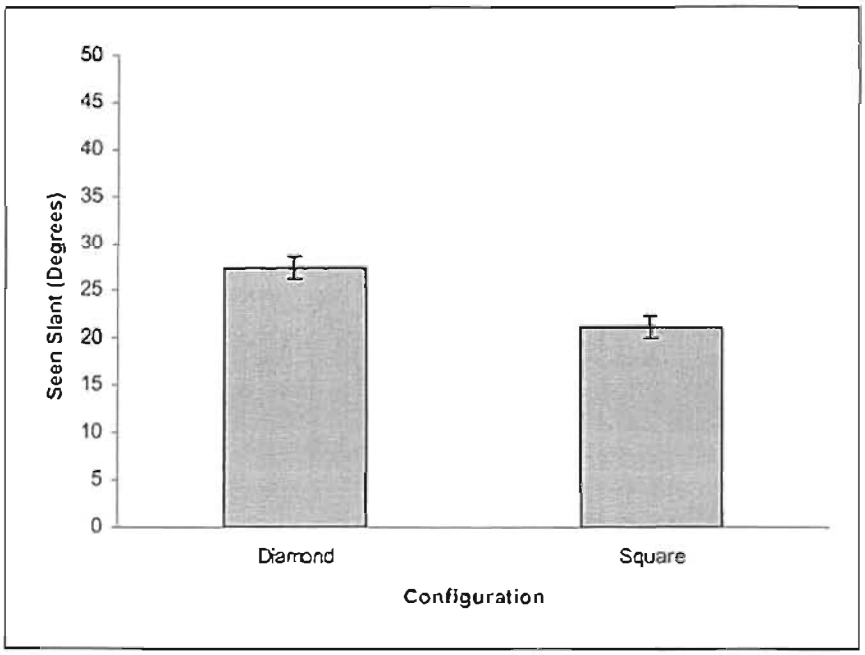
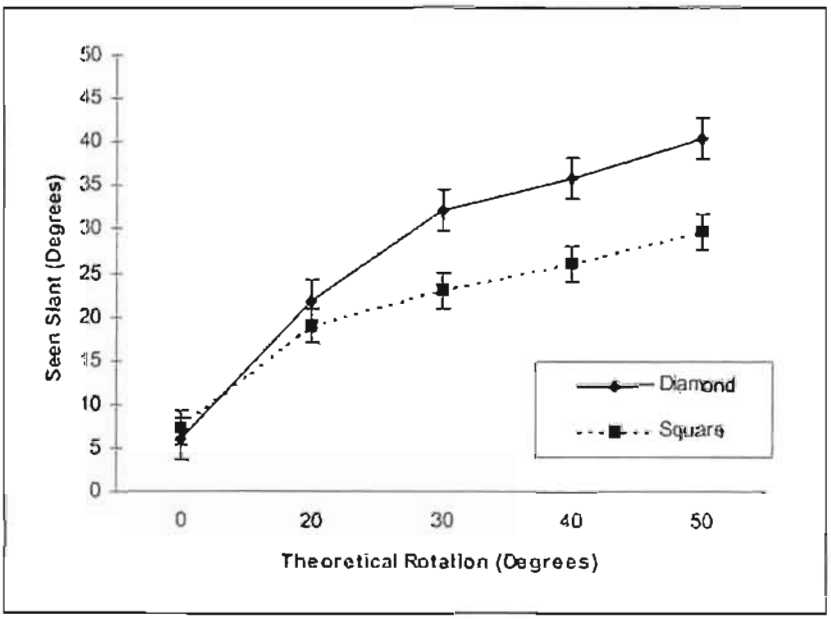



**Fig 8.17. The effect of shape upon seen slant in stereoscopic Ehrenstein configurations**
This figure plots mean seen slant for a square (SES) and diamond (SED) shaped grey figure partly obscured by four black panels in Ehrenstein configuration.

A first order interaction between shape and theoretical rotation also yielded a significant effect upon seen slant: $F_{(3,5)} = 8.47.52$, $p < 0.001$. This is the key finding for the experiment. The finding supports the hypothesis that the system had access to a more thorough metric of retinal disparity in the SED compared to the SES (by accident of the shape of the configuration of T-stems). Planned comparisons revealed significant differences between the two shapes at $30^0$, $F(1,4) = 25.05$, $p < 0.001$; at $40^0$, $F = (1,4) = 29.25$, $p < 0.001$; and at $50^0$, $F = (1,4) = 34.85$, $p < 0.001$.

These finding suggest that slant attenuation was lesser, for the Ehrenstein diamond, at each magnitude of stereoscopic rotation. (see Fig 8.18). Hence, seen slant in the SED is closer to the theoretical rotation applied to the central grey diamond shape.




**8.18. A shape x theoretical rotation interaction effect upon seen slant in SES and SED configurations**
This figure plots mean slant estimates for a first order interaction between shape (square / diamond) x theoretical rotation in stereoscopic Ehrenstein configurations.

The SED, whose T-stems were all consistent with theoretical rotation, yielded more accurate seen slant than in the SES where only two T-stems were consistent with slant. What these findings suggest is that all else being equal (eccentricity, viewing distance and relative magnification of the image pairs) that stereopsis integrates diverse disparity measures across image pairs. Seen slant and the perception of slanted contrast spreading layer must be associated with integration of disparity values.

I have explained that these disparity values can arise at different feature scales, and that those difference measures can conflict. Hence, the results point to a relatively ambiguous horizontal 2-D alignment of the T-stems in the SES. Their effect on the overall slant response may represent what Gillam and Ryan (1994) have called a *cue conflict* against stereoscopic rotation of the grey surface layer perceived. In a sense, it is surprising is that stereoscopic slant is seen at all in the SES because so few contours are actually disparate. To achieve seen slant the system must assign large feature scale *spaces* in between contours to an oblique orientation. Moreover, as indicated in the previous discussions, several of the disparities that do exist in the SES configuration are in conflict with the eventual perceived orientation of the SES percept (upon fusion).

Seen slant has previously been shown to be sensitive to various configurational, that is, non-local aspects of stereopsis. The finding, now repeated, that a SES could be seen to

slant suggests interaction between image differences at different feature scales, that is, between disparities at the horizontal T-tops and disparate subtense between the vertical T-stems. The conflict appears to exist between disparate subtense and point-disparities at specifiable contours in the SES. This suggests that some confirmatory integrative mechanism underpins interpretation of the horizontal T-stems stems with vertical T-stems and so on.

Experiment 8 suggests that the integrative mechanism appears to be operating in the fully modal, the fully amodal, and the part-modal/amodal percepts (positive, negative and zero standing disparity respectively). But, because seen slant was generally greater for positive standing disparity, it may be that there is an asymmetry in the strengths of visible-non-visible integration. The next experiment will address that prospect in detail.

An SED yielded relatively accurate slant judgements. The shape of the SED means that all T-stems are precisely aligned across the black panels in each half-image. Since the T-stems all have the same magnitude of orientation disparity, when disparate half-images are fused the T-stems will also be aligned across three dimensional space.

Alignment and precisely equal orientation disparities at the T-stems of the SED have two theoretical implications. Firstly, the system would not actually require an integrative completion mechanism to achieve stereoscopic rotation at all. Perceived spreading of the illusory grey layer could be considered to simply require filling-in between luminance and depth referents. Secondly, fusion of the T-stems and their assignment to a near or distant depth layer seems to arise in the manner described with the BIPASS model invoking surface separation and spreading.

Despite the presence of vertical image differences, surface spreading in the SED was generated across the T-tops when the T-stems stood forward of the projection planes. No surface spreading was generated when the T-stems stood behind the T-tops. Moreover, the relative accuracy of seen slant in the SED suggests that the mechanism underpinning surface separation must be tightly constrained by cyclofusion of the oblique T-stems. Resolution of disparate subtense between bounding contours is also likely to be constrained in that manner because the two are immutably related to transformation of the diamond shape and / or its relative position in the half-images.

The likely relationship between vertical image differences and surface separation is addressed in Fig 8.19 which depicts a pair of T-Junctions in the SED. Various luminance

values in this figure have been labelled. A is black, B is white and C is grey. The disparate T-junction images projected to the eyes are shown.

A vertical difference of the type that Anderson has referred to is at O_E along the stem of the T-junction in Image 2 (Fig 8.19). This region is not matched in each eye, according to the epi-polar constraint. Anderson's claim is that the system infers that the unmatched region must be a half-occluded feature.

It is possible to sketch the surface separation process using simple 2-D diagrams. When this is done the physical 2-D constraints upon 3-D perceptual organisation at surface separation may be crystallised.

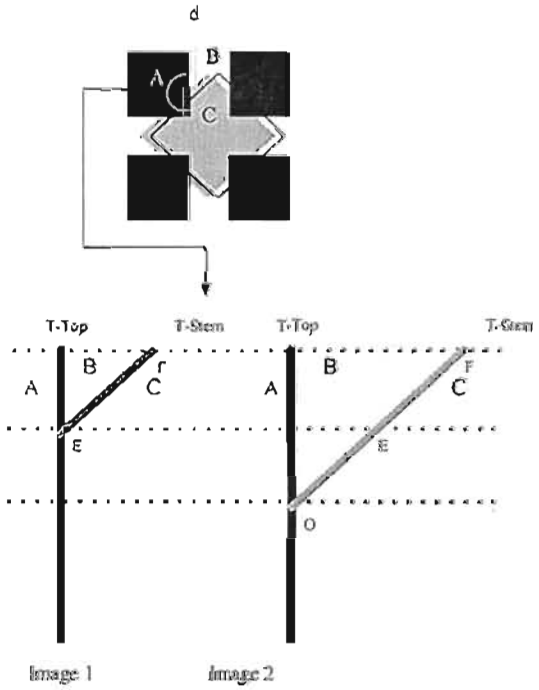



**Fig 8.19. Image comparison and the epi-polar constraint on binocular matches at T-Junctions**
This figure describes oblique T-Junctions in the SED. Image 1 and Image 2 represent the two retinal images of this T-Junction. When disparity is crossed in the Ehrenstein figure Image 1 will be seen by the right eye and Image 2 by the left and the reverse is true for uncrossed disparity. The Long black vertical line represents the T-top. Luminance values surrounding the T-Junctions are labelled: A is black, B is white and C is grey. The T-stem is represented by the line E_F in Image 1 and O_E_F in image 2 and the dotted lines represent epi-polar lines at critical points along the T-top.

A BIPASS model based interpretation of surface separation at the above T-junction is shown in Fig 8.20. The overriding principle of the BIPASS model is that the components of the T-Junctions, the luminance steps (not the lines meant to represent them), are not isolated events in the retinal image (as seen slant has suggested). It is possible to map one eye's image

on to the other to simulate vergence eye-movements achieving vergence lock. Then, it seems feasible to speculate on what the patterns of luminance reveal to the system.

In 8.20a the two eye's images are mapped over one another simulating image comparison. In the coordinate matrices the two T-tops occupy identical positions, that is, zero disparity. There is a vertically located region of luminance that has a different value in each eye, namely, region D (Recall that the lines in these figures represent contrast change along the T-top and T-stem). This means that D is grey in one eye and white in the other. A typical horizontal crossed fusion of the T-stem (see figure 8.19b and 8.19c) shows how this non-epi-polar difference in binocular subtense is resolved in surface separation.

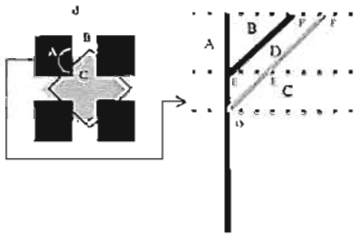
In Fig 8.20a a depth step is achieved in crossed fusion of the T-stems. The T-top remains assigned to the zero disparity plane. To achieve this the system must fuse the region E_F in each eye (the grey-white contour at the T-stem). This suggests that a sector of the near surface, O_E projects partially across the T-top. At this sign of disparity surface spreading emerges. Why does this disparate subtense not create stereoscopic slant? The answer is that the orientation of the contour, the luminance step between grey and white is constrained to fronto-parallel.

In uncrossed disparity, surface separation yields a depth step along the T-top and not along the T-stem. This is shown in Fig 8.20c. The T-stem is assigned to the distant depth plane, but no separation of surface layers occurs at the T-stem because of the uncrossed configuration. Surface separation seems to involve splitting the images along the T-top with the subtense (O_E) a white monocular zone visible only in the right eye. A Split-projection configuration at the T-top means that the near and distant planes share the luminance step at the T-top. Orientation of the contour along the T-stem, the luminance step between grey and white, is constrained to fronto-parallel.

A sketch of the perceived depth relationships is shown in Fig 8.20d. In crossed fusion of the T-junction a depth step is achieved at the T-stem and in uncrossed fusion it is achieved along the T-top. One way to explore the constraints imposed by oblique contours was to introduce orientation disparities at the T-stems. A simple experiment was designed to examine seen slant in a SED.

Essentially, the argument is that with vergence locked, fusion of T-Junctions are not isolated events but constrained by a complex set of 2-D differences returned by image comparison. Those differences occur in the position of luminance contours and in the size and

shape of spaces between those contours. Surface separation is a resolution of those interocular differences in yielding a singular cyclopean view. At partial occlusion, where a depth step is perceived, near and distant layers can share bounding contours.



Vergence Lock and Image Comparison at T-Junctions

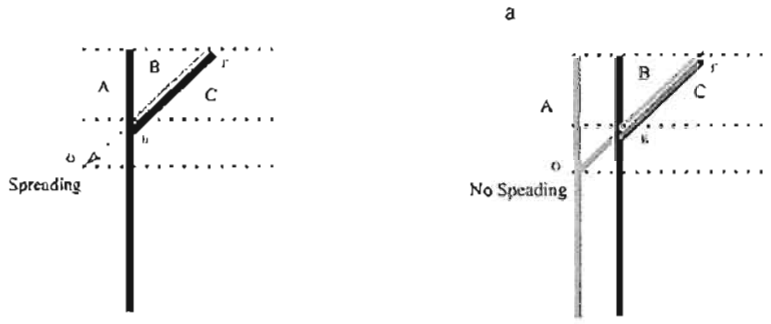



**Fig 8.20. Fusion of T-Junctions in a SED**

In fusing the stereoscopic T-Junctions described in Fig 8.19, the system must align the grey-black portion of the T-top and the white-black portion of the T-Junction. This is represented in (a). Crossed disparity at the T-stem is shown in (b). A simple view of crossed fusion is that it assigns the T-stem to a near depth plane creating a depth step along the T-stem (E_F). The system must assign the T-top to a distant depth plane and the T-stem is the depth step between the grey and white layers. A triangular crossed fusion means that the system separates surface layers along the T-stem but maintains the white-black-white-grey region B on the distant depth plane. Hence, the step yields a transparent grey layer by displacing the region d (grey in the left eye) to a near depth plane. The region O_E (in the left eye for crossed fusion), may actually overhang the T-top in the left eye. Fusion of uncrossed disparity is shown in (c). Alignment of the T-stem and the T-top requires uncrossed fusion along the T-stem E_F. The section of the contour E_O is seen only in the right eye as the system cross-fuses E. The result is a depth step along the T-top.

These ideas are summarised as well in (d) for crossed and uncrossed fusion. In crossed fusion a depth step occurs at the T-stem and in uncrossed fusion orientation disparity occurs at the T-top. In crossed fusion there may be a monocular region of the grey portion of the grey white T-stem assigned to the near depth plane. This was identified as a possible contributor to spreading of an illusory surface layer in the stereoscopic Kanizsa square. It may also contribute to spreading of an illusory contrast layer across the SES. Though we can't make too much of this aspect since at the horizontal arms of the cross the same kind of near monocular region does not happen.

It is a reasonably simple matter to demonstrate that vertical image differences don't appear to be critical in surface separation mechanism except that they are a product of disparate subtense in the proportion of near and distant surface layers (in the SED at least). A demonstration is included below that negates all vertical differences in the SED. It seems that surface spreading still occurs.

In that crossed configuration, the grey-white T-stem contour is shared by the grey (near) and white (distant) depth planes. At the same time, the T-tops (grey-black and white-black contours along the black panels) at the zero disparity depth plane are visible *through* the interposed grey layer. Spreading is associated or invoked by the separation. It should be noted that in the Split-projection configuration of projections at the T-stem, the white and grey luminance regions still *share* the T-stem contour at the same depth plane.

For free fusers, a selection of stereograms similar to those that subjects observed is presented in Fig 8.21a and Fig 8.21b. Crossed fusion of L-M in (a) and (b) yields a slanted SES where the central grey cross stands at a distant depth plane as if partly occluded by the black panels. This requires two components of disparity in the diamond shape - orientation and standing disparity. Crossed fusion of M-R in both (a) and (b) gives the impression that a near grey transparent and slanted surface layer floats above the black panels. Crossed fusion of the pairs at Fig 8.21 (c) should yield the perception of a part-modal/amodal figure slanted through the projection plane (depth layer of the black panels).

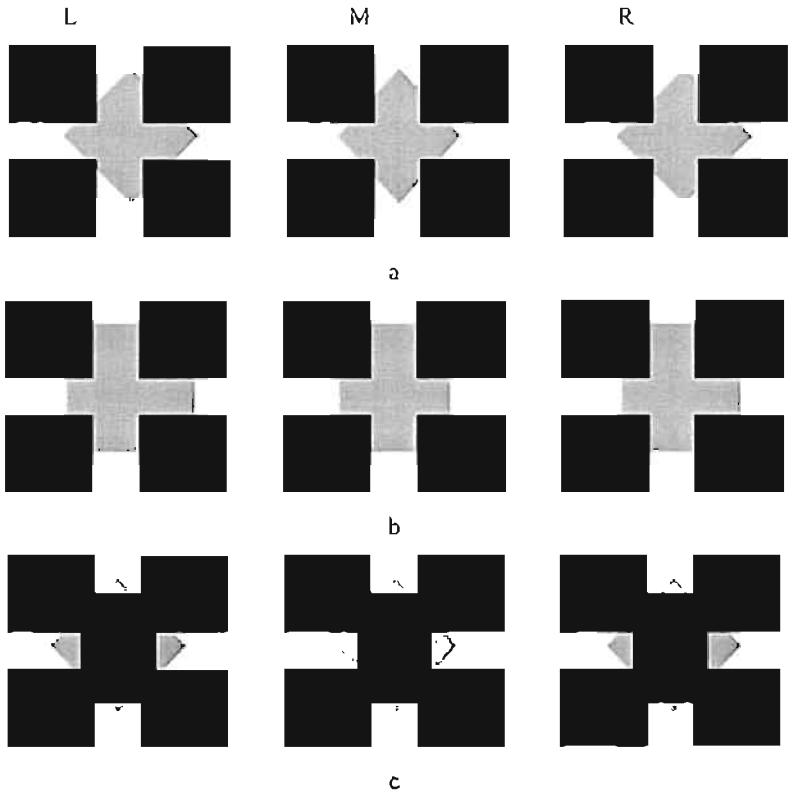



Fig 8.21. Slant and standing disparity in SESs
Crossed fusion of the L-M pairs in fig 8.20a (diamond in back) and 8.20b (square in back) gives the impression that an opaque grey surface layer stands behind the black panels and is oriented in the vertical slant-axis. Crossed fusion of the M-R pairs in (a) and (b) sees a near transparent grey surface layer floating above the black panels. Contrast spreading occurs where the grey surface appears interposed over the black panels. In (a) and (b) subjects judgements of seen slant were less attenuated for the square shape. In Fig (c) the vertical inter-retinal differences have been negated to demonstrate that a very clearly defined Ehrenstein diamond shape can be seen slanting *through* projection plane – transparent at one end (near) – opaque at the other (far). There are only horizontal disparities in this figure.

The next experiment addresses in some detail the nature of integrative process such as completion, in some detail.

## 8.4 Experiment 9 Seen slant and integrative mechanisms in contrast spreading

Previous experimentation in this chapter has suggested that resolution of both point-disparity and larger feature scale disparities, such as disparate subtense, underpin the stereoscopic response in 3-D SES percepts. Seen slant in the SED showed that when these two inter-retinal image measures were consistent with stereoscopic slant, the stereoscopic response did not appear to be underdetermined by disparity. These two contingencies of inter-retinal

differences appeared to constrain uncertainty even in untextured stereograms involving partial occlusions. Clearly, across a textured surface, a very large number of possible disparity feature-scales could exist (ie. defining disparity curvature across the surface in relation to its visible edges).

Another way of phrasing this ability of the system is that it appears to be responsive to disparity between large-scale image features (eg, a partly obscured diamond) as well as small scale image features (contour positions). The BIPASS model has proposed that resolution of disparate subtense constrains 3-D perceptual organisation in untextured stereograms. Surface separation and spreading result. Constraint stems from the recovery and resolution of the detailed 2-D layout of the half-images.

*Seen slant as metric of integrative processes in the SES percepts*

The final experiment presented in the thesis involves another simple slant judgement task. The seen slant metric offered an opportunity to exploit ambiguous retinal disparities to examine the nature of confirmatory integrative visual processing in the SES configuration.

To test for a *tendency* to complete the background square in the Ehrenstein figures an SES stimulus was employed (in which only horizontal disparities at two of the four T-stems exists). The logic of the experimental manipulations carried out was this. If the tendency to complete the Ehrenstein figure were a third component of the 3-D perceptual organisation, then, manipulation of that tendency would impact on the seen slant metric. Seen slant appears to be sensitive to different manifestations of disparity. Then, by constraining these aspects and changing the magnitude of the black panels the tendency to complete the background grey square might emerge as a separate component of the seen slant metric.

Evidence for these propositions stems from Experiments 7 and 8. These experiments established that a SES whose disparities are created by drawing a grey square behind the black panels yielded seen slant. Seen slant was about proportional to the relative horizontal magnitude of the grey square drawn in each eye's view. Seen slant was quite dramatically attenuated. Attenuation increased as retinal disparities were increased.

These contingencies suggested that seen slant in the Ehrenstein figure was not purely a function of the magnitude of retinal disparities between the grey cross half-images. This is not surprising since there is no actual disparity at the vertical arms of the central grey cross. That prospect brings with it an important theoretical implication, the local zero disparity measure at

the vertical arms of the grey cross seems to have impacted upon the interpretation of disparity at the horizontal arms of the grey cross.

The point-disparity signals at the four luminance contours at the T-stems of the SES are separated by substantial regions of homogeneous luminance. There is also disparate subtense between the T-stems. There is no disparity at or along or between the T-stems at the vertical arms of the SES. Therefore some kind of long-range interaction between them seems to have emerged. Attenuation of seen slant in the stereoscopic Ehrenstein square might be due to some kind of feature integration mechanism, that is, completion.

*Confirmatory processing in the SES: Completion, Confliction or both?*

Figural completion is a somewhat controversial topic in vision science. Figural Completion is associated with the perceptual judgement that fragmentary visual information is integrated to create some kind of emergent visual *form* (Kanizsa 1975). One mechanism of completion in the SES most recently advocated by Grossberg and his colleagues is a pre-visual boundary constructed by a BCS.

In construction of this pre-visual boundary the system is blind to depth, that is, disparity signals and to specific luminance values. The FCS then fills the integrated boundary from disparity pools. It is the activity of the FCS that creates the perception of an illusory contrast change manifest as contrast spreading. In Grossberg's view the outcome of these parallel integrative mechanisms is compared to *a priori* templates in an Object recognition System (ORS).

Grossberg's approach makes no reference to the possibility of conflict between the various depth and luminance components of his BCS-FCS-ORS feed-back loops. It is not difficult to imagine that conflicting relative weights in these systems could co-exist. Resulting perceptual outcomes could reflect those weights and might even change over time depending upon the context of interpretation (vision).

Grossberg's work concerns integrative visual mechanisms. His theory reaches across 2-D and 3-D domains of perception. The main concern of this project was binocular visual processes and stereoscopic information. A valuable notion derived from Grossberg's approach was allelotropia. The BIPASS model re-caste allelotropia in the terms of constraints imposed on 3-D organisation generated by disparate subtense.

Nonetheless, in reference to the SES, Grossberg's theory makes one important prediction about the part of T-Junctions. That is, the illusory contour that bounds of the grey

contrast spreading layer *continues* from the T-stems across the T-top and meets with adjacent T-stems that mirror luminance order.

Kellman and Shipley also argued that continuation of the T-stem contour was a visible (modal) or invisible (amodal) projection across the luminance discontinuity at the black panels. Kellman and Shipley suggest that this process is part of figure ground splitting between geometrically relatable contours.

I concur with these arguments to a point. The BIPASS model is slightly different than both Grossberg's and Kellman and Shipley's. The theoretical difference is subtle but important. In particular, it is important because the BIPASS model may account for Anderson's effects where an illusory contour can be seen in the absence of any obvious continuation process.

The BIPASS model proposes that spreading involves continuation of a near surface layer, separated from the depth plane of the black panels. Moreover, it is not a "contour" that continues across the black panels but the depth step or *separation* of two depth layers. Surface separation can happen in isolation to adjacent luminance and depth referents (as established in the stereoscopic Kanizsa square-pair, stereoscopic Kanizsa square-line and stereoscopic Kanizsa square-dot manipulations – see Chapter 6).

It has been established previously that Anderson has shown that completion is not the *cause* of illusory phenomena in stereopsis, just and Jory and Day (1979) and others have shown in 2-D stimuli. Regarding continuation mechanisms, Anderson (1997, p18) says this:

> ...contour continuation processes are not necessary for the formation of illusory contours in binocular vision, and hence, any theory that requires the presence of contours along the direction that the illusory contour forms must be incorrect ...That is not to say, however, that continuation processes do not play a role in illusory contour generation when such image features are present However, there is currently no evidence that uniquely identifies these features as either necessary or sufficient for the formation of illusory contours.

Again, I concur with Anderson in that respect. It is equally likely, however, that a satisfying explanation of the 3-D illusory percepts will, in the end, need to account for confirmatory processes such as continuation or completion or confliction mechanisms.

Kellman and Shipley (1992) discovered something important about the mechanism of illusory contours that appeared to offer a useful metric for the present experiment. Kellman and Shipley had subjects rate the intensity of illusory contours. They found that the subjective intensity of illusory contours was roughly a ratio of the physical length of the edge separating contour inducers (eg. pacmen) and the physical length of the contour inducing the illusory contour. For example, in a Kanizsa square the intensity of the illusory contours would be strong with big pacmen close together and weak with small pacmen a long way apart.

This experiment employed the concept in the SES, the logic being that incrementally reducing the relative size of the black panels impacted the proportion of inducing contour (the stems of the T-Junctions) reducing in proportion to the separation of the T-stems (the magnitude of the black panels). In doing so, it was predicted that the relative weight of the *completion* component of stereoscopic slant would increase as the size of the black panels was decreased. Subsequently, it was predicted that the relative weight of the local disparity component of seen slant would decrease. Therefore, seen slant might be expected to increase as the size of the black panels decreased.

Therefore, this experiment examined evidence from Experiments 7 and 8 that some completion/confirmatory process compatible with my understanding of surface separation was active in the 3-D percept evoked by fusing the SES. Manipulations of the Ehrenstein figure are described using Fig 8.22.

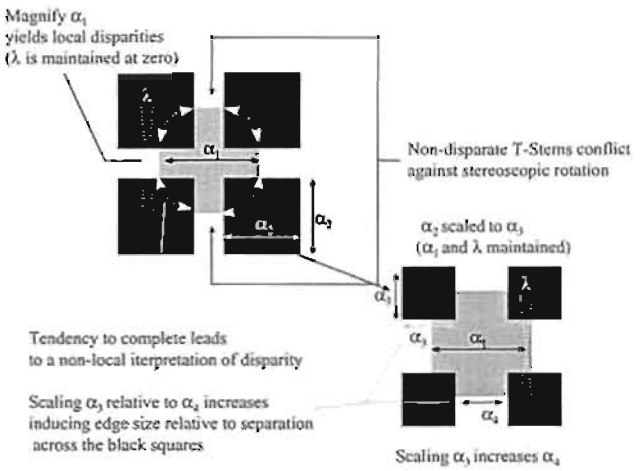
The objective of the experiment was to compare configurations in which interaction between contours signalling the same zero disparity and some disparate subtense across the configuration in the manner of completion would impact the seen slant metric. An Ehrenstein figure was constructed by drawing a grey square behind the four black panels. Both the magnitude of the grey cross and the local disparity at the grey cross were constrained while transforming the size of the black panels. Subjects then judged the slant of the illusory stereoscopic surface layer.

Disparities at the T-stems of the horizontal arms of the SES were applied by drawing a grey square behind the black panels. That grey square was manipulated appropriate to a perceptive corrected transformation of Ogle's magnification factor between the half-images. The perspective corrected relative magnification of the grey square defined a constant theoretical rotation of forty degrees in either the vertical or horizontal - slant axis. No standing disparity was applied.

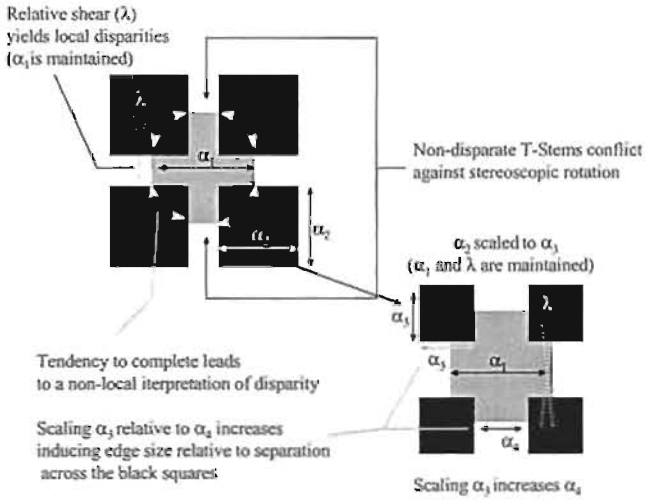
*Measures*

In summary, the effects of two independent variables upon seen slant were employed in Experiment 9: the relative size of the black panels and slant-axis. This was achieved by scaling the black panels symmetrically by an arbitrary percentage[24]. Scaling factors were: 0, -5%, -10%, -15%, and -20%. A scale factor of 0 degrees meant that the size of the black panels was the same as in the previous two experiments (see procedure for details) and a -20% scale factor reduced the size of the black panels relative to the 0% condition by 20%. The centre point of the black panels was precisely maintained. The hypothesis was that incremental reduction of the magnitude of the squares; all else being equal, would decrease slant attenuation in the SES and so yield an incremental increase in seen slant. It was also anticipated that the horizontal axis of rotation would be little affected by changing the relative size of the black panels relative to the vertical axis. Hence, an interaction between slant-axis and relative scale of the black panels was anticipated. Planned comparisons were employed to establish differences between slant-axes at each level of relative black-panel scale.

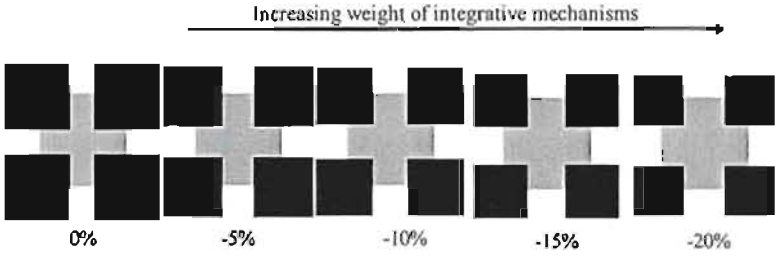
[24] We say *arbitrarily* but pilot testing was used to establish a reasonable *arbitrary* range.




a - Vertical slant-axis




b - horizontal slant axis




c

**Fig 8.22. Re-scaling black panels in the SES**

Manipulation of the Ehrenstein square appropriate to slant in the vertical slant-axis and scaling of the black panels is shown in (a). This involves relative magnification of the grey square drawn behind the black panels (ie. changes $\alpha_1$). Manipulation of the Ehrenstein square appropriate to slant in the horizontal slant-axis and scaling of the black panels is shown in (b). This involves relative shear of the grey square drawn behind the black panels (ie. changes $\lambda$). In both (a) and (b), re-scaling the black panels ($\alpha_2$ to $\alpha_3$) changes their proportion of the inducing length of the T-stems ($\alpha_4$).

### 8.4.1 Method

*Subjects*

8 subjects were drawn from the department pool. They performed the same practice trials and were screened in the same way as previously.

*Stereograms*

Half-images were presented at the centre of the monitor, at eye-level in the mid-sagittal plane. Each half-image consisted of a set of four black panels subtending about $3^0$ (at 750mm viewing distance). The black panels (0.09cd $m^{-2}$) were positioned so that a square drawn through their centres would subtend $7^0$ degrees. A central grey cross was created by drawing a grey square (0.4cd $M^{-2}$) or a grey diamond so that it obscured the four black panels. The grey square or diamond overlaid the black panels symmetrically (equal intrusion on all sides) intruding about ½ the width of the black panels. The grey square or diamond was then stood behind the black panels against a background luminance of 0.7 cd $m^{-2}$. black panels were re-scaled by simply reducing their magnitude (0%, -5%, -10%, -15%, and -20%).

The shape and size of the grey square was manipulated to generate appropriate theoretical rotation. The grey square was geometrically transformed according to the monocular transformations of a square shape as described in Appendix A. Disparity appropriate to stereoscopic rotation was then introduced by applying ½ Ogle's M to the grey square, symmetrically and in opposite signs in each eye's view of the grey cross (see Fig 8.23).

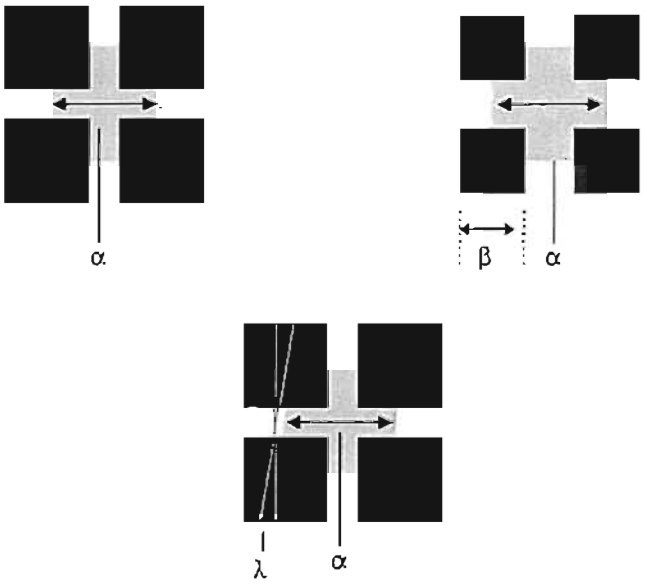



**Fig 8.23. Patterns of disparity in the SES**
Fig 8.23a shows one half-image supporting rotation about the vertical axis for the square. Relative magnification was applied by increasing the magnitude of the grey square (α) in one eye relative to the other. Similarly relative magnification was applied by increasing the magnitude of the grey square (λ) in one eye relative to the other. The black panels were re-scaled by adjusting their magnitude $\beta_1$.

*Design and procedure*

A two-way (2x5) repeated measures design was used to explore the effects of slant-axis (horizontal, vertical) and relative scale of the black panels (0, -5, -10, -15, -20%) on the estimated slant of the stereoscopic Ehrenstein square whose theoretical rotation was $40^0$. Three complete repetitions of this design were used, making a total of 30 trials per subject in all. Details of the procedure were as described in section 4.5. Subjects were asked to rotate a comparison stimulus defined by luminance contours to match the degree to which they judged a contrast-spreading figure to be rotated.

## 8.4.2 Results and discussion

A two-way (2x5), repeated measures analysis of variance examined the effects of slant-axis (horizontal, vertical) and the relative scale of the black panels (0%, -5%, -10%, -15%, -20%) on seen slant. Obtained slant estimates were averaged across repetitions.

The relative scale of the black panels was found to impact upon seen slant and the effect was significant: $F_{(4,2)} = 12.13$, $p = 0.001$. Figure 8.24 compares the mean slant estimates across the incremental changes in relative scale of the black panels. The findings

here were that the relative scale of the black panels impacted on seen slant but not entirely as predicted. While it was anticipated that seen slant would increase as the scale of the black panels decreased, this appears to have been the case up to point. At -20% relative scale, seen slant was found to be attenuated more than at -10% and -15% relative scale.

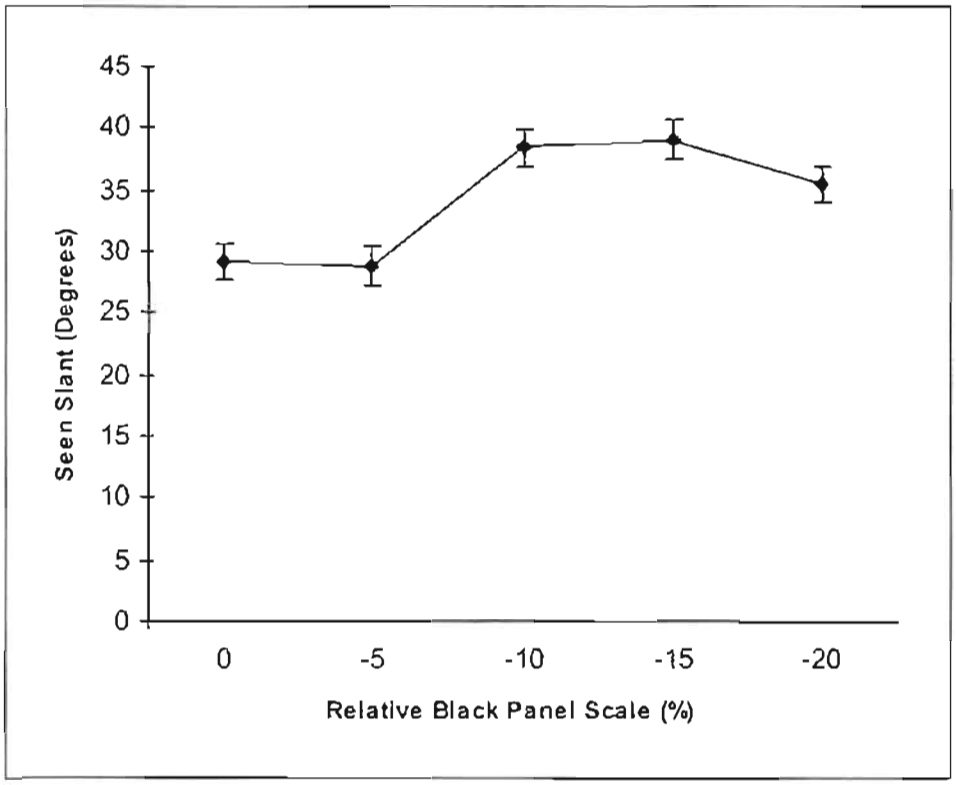



**Fig 8.24. The impact of relative scale of the black panels upon seen slant in a SES**
This figure plots mean slant estimates (seen slant) for the levels of relative scale of the four black panels in the SES.

Slant-axis had no significant impact on seen slant. The key result is the first order interaction between factors. The relative scale x slant-axis interaction was found to impact upon seen slant and that interaction was also significant: $F_{(4,7)} = 11.93$, $p = 0.001$.

Mean seen slant estimates for each level of black panel scale and each slant-axis are plotted in Fig 8.25. Clearly, scaling of the black panels had an axis-asymmetric effect upon seen slant. For rotation about the vertical axis reduction in the scale of the black panels yielded increased seen slant until -10% then a reduction of seen slant has emerged. In contrast, for the horizontal slant-axis seen slant increased from the -5% manipulation to slightly greater than veridical at -20% scale in the manner predicted. At the -20% relative scale a typical slant-axis anisotropy (about 10 degrees) is present where seen slant about the Horizontal axis tends to be slightly greater than the vertical axis.

Planned comparisons revealed that the difference between means at the 0 % scale was not significant while at the -5% scale the difference between seen slant at each axis was significant $F(1,4) = 16.4$, $p < 0.001$. At 10% seen slant was not significantly different, while at the -15% scale, $F(1,4) = 5.4$, $p < 0.05$; and the -20 % scale, $F(1,4) = 25.8$, $p < 0.001$ differences in seen slant were significant.

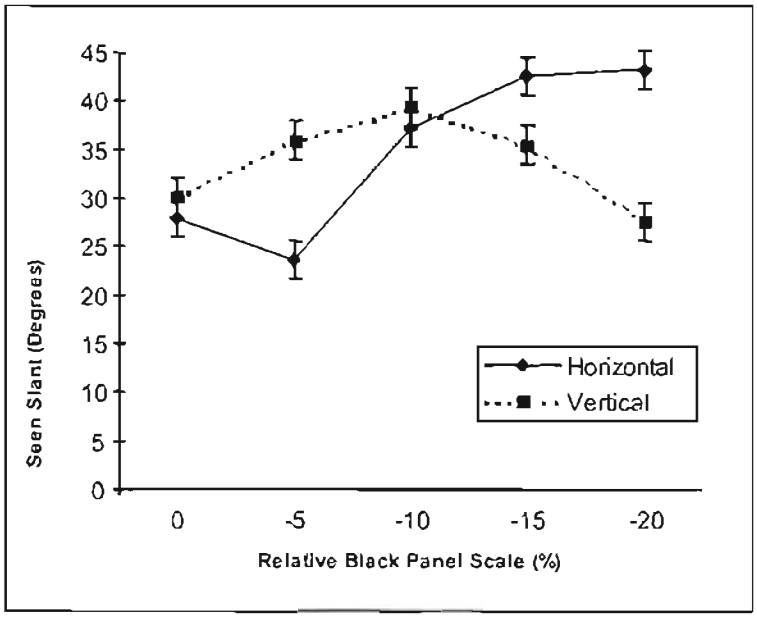



**Fig 8.25. Interaction of relative scale of the black panels x slant-axis**
This figure graphs mean seen slant for a first order interaction effect between relative scale of four black panels x slant-axis. The black panels were drawn to obscure a grey square theoretically rotated at $40^0$. Standard error bars have been included.

In rotation about the horizontal axis the relative scale of the black panels had a quite dramatic impact upon seen slant. Between the -5 and -20% levels, seen slant approximately doubled from about 24 to $40^0$. Recall that the large feature scale disparity (disparate subtense or shear across the horizontal arms of the central grey cross) was precisely the same magnitude in all stimuli presented in this experiment, that is, appropriate to theoretical rotation of $40^0$. Results therefore suggest that when the SES half-images are fused, the slant response appears to reflect integration of disparity values. Evidently, integration of those disparities by the system depends upon integration of the overall configuration. Reducing the size of the black panels was predicted to increase the strength of illusory connections across those panels.

In previous discussions, I have interpreted this as evidence of some kind of confirmatory visual processing such as completion that contributes to the perception of

spreading surface layer by driving the interpretation of ambiguous local disparities. Disparity values are not isolated image difference measures but are interconnected in the overall slant response. In that manner they appear to carry relative weights.

I would argue that the increase in seen slant for the SES rotated about the horizontal axis shows the increased weight of those local disparities consistent with theoretical rotation. However, in the SES rotated about the vertical axis, reducing the scale of the black panels below -10%, appears to have increased the relative weight of local disparities (non-disparate T-stems at vertical cross arms) that were not consistent with theoretical rotation (after -10% relative scale). Hence, seen slant was increasingly attenuated in the vertical slant-axis.

In summary, the relative sizes of the T-tops and T-stems in the SES were reduced as the scale of the black panels reduced. Disparity values evidently occupy different weights in the overall slant response depending on the relative magnitude of the black panels. Moreover, it was found that the impact of relative weights of disparity values was axis-asymmetric.

This experiment incrementally reduced the relative size of the black panels in the SES, impacting the proportion of T-stems (by changing the relative magnitude of the black panels). It was found that a *completion* component of stereoscopic slant increased as the size of the black panels was decreased but only for rotation about the Horizontal axis.

Results suggest that the relationship between the tendency to complete the stereoscopic Ehrenstein square as defined by the relative proportions of black panels compared to the magnitude of the T-stems was more complex than anticipated. The evidence for a completion mechanism or confirmatory mechanism appears to be quite strong in the SES, however. The next chapter addresses the significance of these findings in relation to a number of stereograms that have been produced in the last decade or so.

# Part 3 On the generality of binocular vision processes underlying stereoscopic illusory contours and surfaces

***Summary:*** *Part 3 of the thesis reviews the possible generality of processes identified in the BIPASS model for perception of a range of stimuli generating 3-D illusory percepts. It then summarises experimentation and makes concluding remarks.*

*In* ***Chapter 9*** *the BIPASS model is reviewed and used to offer an alternative functional account of percepts from several key stimuli. These stimuli have been principally proposed by the Surface Heuristic approach. The chapter attempts to describe how the 2-D structure of half-images is related to the 3-D illusory percepts. Grossberg's approach is also reviewed. It is suggested that a gap in his theory appears to be the apparently mechanistic separation of surface layers in stereopsis.*

*Finally,* ***Chapter 10*** *summarises the project and proposes some possible implications for natural binocular vision. Several stereo photographs are used to demonstrate the possible relevance of the BIPASS model for natural vision. It is suggested the outcome of binocular vision may be a complex product of 2-D relationships between point-disparities, disparate subtense and the topographic architecture of the binocular sensory array.*

# 9. Generality of binocular vision processes in untextured stereograms

*Summary: Previous chapters have suggested that even when point-disparities were scarce, in untextured stereograms, stereoscopic mechanisms still played an important role in depth perception. A functional BIPASS model was framed to describe the perception of an SKS. The model was also applicable to the SES percepts. Using a series of demonstrations, that to some extent are speculative, this chapter applies the model to several other types of stimulus configuration. These include a simple black cruciform, an oblique cruciform, Anderson's l-junctions and stereo capture effects.*

## 9.1 Binocular image processing and 3-D Illusory Percepts

### 9.1.1 A brief review

This project has developed and tested a functional BIPASS model to describe stereoscopic mechanisms underpinning the SKS percepts. It is a model that stems from a sensory systems approach to the perception of the 3-D Illusory percepts (see Fig 9.1). Though still immature, the model appears to be applicable to the SES (contrast spreading) as well.

In summary, my proposal is that the 3-D illusory percepts are the products of several system processes. Inter-retinal measures of parity and disparity are detected across the two retinal coordinate matrices at multiple feature scales (image comparison). Retinal disparities are then resolved in separation of surface layers and this can yield surface spreading toward a luminance and depth referent. Higher-order confirmatory processing appears to influence the precise trajectory of surface spreading and integration of disparity values.

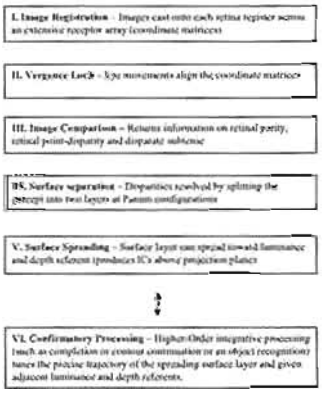


Fig 9.1. A BIPASS model of visual system processes underpinning the SKS and SES percepts

The system's response to untextured stereograms appears to involve resolution of inter-retinal differences in image structure at various scales. A key concept is that the system can resolve image differences by separating the singular percept into surface layers in the absence of disparity discontinuity. A case has been made that surface separation occurs where resolution of image differences involves a "split-projection configuration".

The term "split-projection configuration" is simply a description of the binocular vision projection geometry that is associated with perception of partly interposed surface layers. Surface separation arises at contours where projection lines cross. This happens when near and distant surface layers share the same luminance contour. In such a configuration a near edge is fused along a line of sight to each eye that also aligns the more distant disparate edge of a partly occluded surface. In this way, separation of surface layers can be achieved when no texture is available. Moreover, local disparity is not actually necessary.

A key clue to mechanisms involved, I have argued, is the difference between percepts at the two directions of "disparity". The BIPASS model predicts asymmetry even in the simplest possible untextured stereograms. It also provides a parsimonious construct with which to examine the 3-D illusory percepts (if in limited functional terms at this stage). The next sections attempt to show this.

### 9.1.2 A step-wise description of surface separation and perceptual asymmetry in some related untextured stereograms

Discussion in this section refers to Fig 9.2. The stereograms are intended to demonstrate, step-wise, the manner in which resolution of the 2-D layout of inter-retinal disparity and parity, seems to relate to the phenomenal properties of the percepts (such as illusory contours, portholes and transparency).

First, note that in this set of stereograms vergence is locked at the same angle and disparities in all sets of pairs are virtually identical. Image comparison will return only very subtle differences in retinal parity and disparity between the half-images. All of the stereograms exhibit perceptual asymmetry between the signs of disparity.

*A simple black and white untextured stereogram*

Crossed fusion of the L-M pair in Fig 9.2a sees a white disk (bounded at ABCD) floating above a black square. In the untextured retinal images, just one luminance step (contour)

demarcates the circle from the square in each image. My BIPASS model would suggest that both the near surface (disk) and distant surface (square) must share the black-white contour ABCD (in a split-projection configuration) in crossed fusion of the L-M pair, in Fig 9.2a. This was called surface separation. The mechanism resolves disparate subtense between the outer black-white square bounds and the inner black-white circle bounds. By splitting the percept into two layers along the curved contour (ABCD) the system resolves disparate subtense and point-disparity.

There are no detectable unpaired zones in the half-images. A geometric fact of binocular fusion of this untextured stereogram is that the magnitude of the portions of the distant surface (the black square) allocated to monocular zones is precisely the magnitude of disparate-subtense at ABCD. The magnitude of this difference in subtense "fits" behind the near surface (the white square), that is, it is occluded from the view of one eye in a split-projection configuration.

Crossed fusion of the M-R pair sees a white disk visible through a square porthole. Surface separation occurs at different contours. Uncrossed disparity at ABCD means that the curved luminance contour cannot be shared in a split-projection configuration. The system separates surface layers at the bounds of the black square instead. This resolves disparate subtense between the square bounds and the circle bounds. The entire area bounded by the square-boundary is assigned to a distant depth plane by splitting the percept into two surface layers at the luminance step at the square boundary.

*A BIPASS model and stereoscopic illusory contours*

By manipulating the simple untextured stereogram (in Fig 9.2a) it is possible to summarise how the BIPASS model might account for perception of 3-D illusory contours.

In Fig 9.2b, a white cross has been drawn to overlay the black square and white circle. The result is a set of physically discontinuous white-black contours at ABCD. Disparities at those contours have not changed. Crossed fusion of the L-M pair sees a white surface (an illusory disk) stand forward of four black "squares". Stereoscopic illusory contours seem to complete the white disk across EFG and H. The illusory contours look smoothly rounded.

Therefore, surface separation occurs in the same manner as for Fig 9.2a, that is, at the curved luminance contours at ABC and D. Non-disparate curvature of the contours at ABC and D must constrain the perceived orientation of the near layer. Separation of surface layers

in this specific context leads to surface spreading between contours at which the split-projection configuration arises (at ABC and D). Illusory contours represent a depth step at the bounds of a near surface layer spreading toward adjacent luminance and depth referents.

Some kind of higher-order processing seems to be involved in actually shaping the perceived spreading of the near layer. This has been termed confirmatory processing. There does not seem to be a satisfying functional explanation of the appearance of *completion* in the literature.

Perceptual asymmetry is evident when pair M-R (Fig 9.2b) are cross-fused. The percept is a white surface (disk) seen through four windows at ABCD. Uncrossed disparity at the curved black-white contours at ABC and D precludes a split-projection configuration (contour shared between two layers). Instead, surface separation happens at the *window bounds* (EFG and H). A split-projection configuration will resolve disparate subtense at non-curved black-white contour segments. Surface separation at the inner *window bounds* yields near surface spreading and hence illusory contours. Surface spreading seems to complete the *window frames*. Surface spreading is visible as stereoscopic illusory contours.

*Surface separation, contrast spreading and transparency*

Next, consider the stimuli in Fig 9.2c. A transparent grey disk appears to float above four black squares (L-M pair is cross fused). Point disparities at ABCD and disparate subtense remain precisely the same. What is different is that there are now non-disparate grey-white contours present (bounding the four grey sectors). Surface separation occurs along the curved black-grey contour at ABC and D. The mechanism is just the same as in Fig 9.2b. However, the non-disparate grey-white contours stand at the same depth as the black squares. These grey contours are visible through the separated surface layer making it transparent.

Reversing the sign of disparity in Fig 9.2c (fuse M-R) yields the percept that an opaque grey disk stands behind four windows. The BIPASS model predicts that the percept results as the system separates surface layers at all grey-white and black-white contours at ABCDEFG and H. At these contours a split-projection configuration is established.

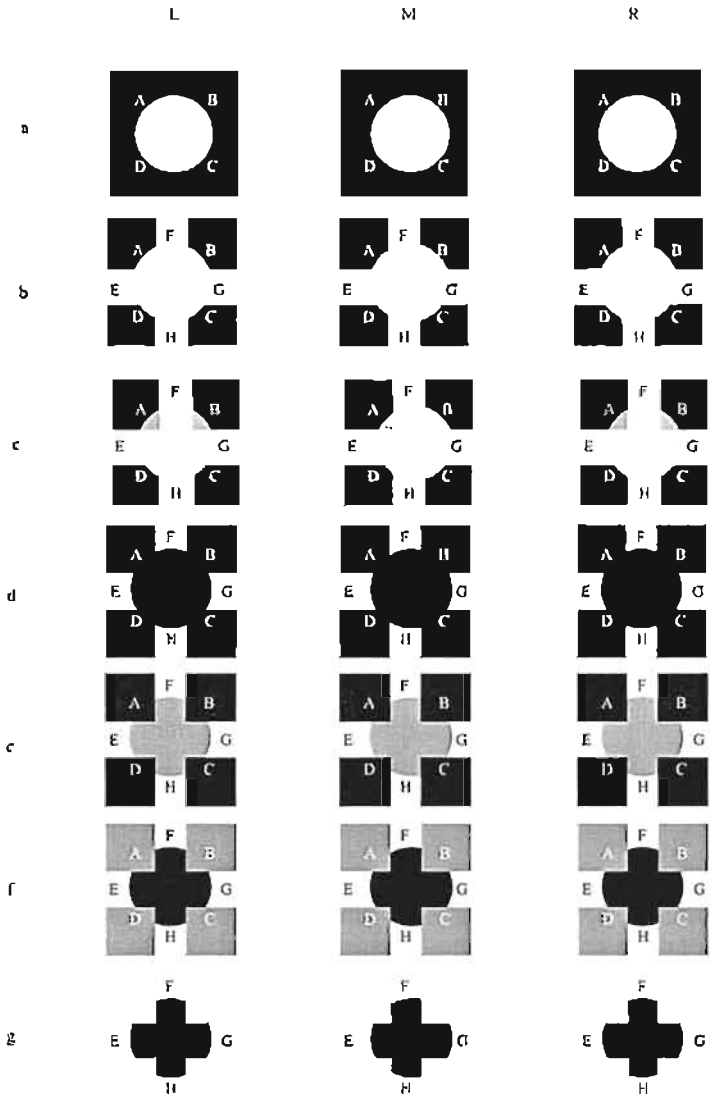



Fig 9.2. Surface separation and spreading

In (a) crossed fusion of L-M sees an opaque white disk standing above a black square. Separation of surface layers occurs along the bounds of the white disk in a split-projection configuration that also resolves disparate subtense. Fusing M-R sees the disk visible through a square window. Surface separation happens at the bounds of the window (straight black-white contours). In (b) the same disparities are present but contours are discontinuous. Surface separation occurs at the curved contours at ABCD in fusing L-M. Illusory contours complete the white disk in between the contours at which surface separation occurs. However, fusing M-R sees illusory contours spread across the near surface at the bounds of four windows. Fig (c) shows the contrast spreading effects. Fusing L-M sees a transparent circle stand forward of the black squares. The surface separation mechanism is almost the same but the system must split the percept at the curved grey-black contours (ABCD) while standing the straight grey contours on the projection plane. This is what must cause transparency. Fusing the M-R pairs sees a grey circle visible through four square windows. The system separates layers at the straight black and grey contours. Fusing L-M in (d) shows that when there is no intermediate contrast the circular shape is seen as opaque and separation now occurs at the curved black-white contours. Fusing M-R however sees the black sectors at ABCD look opaque and stand on the projection plane separated from a distant opaque black layer. Fig (e) sees a near transparent layer spread across the black squares (when L-M are fused). Separation occurs along the curved grey-white contours in fusing L-M and at the black-white and black-grey contours EFGH when M-R are fused. In (f) very similar percepts arise with luminance values manipulated and in (g) only two luminance values are present. Separation occurs at the curved black-white contours (EFGH) in fusing L-M and at the straight contours in the fusing M-R.

*Other manipulations of luminance values and stereopsis in untextured stereograms*

Further subtle manipulations of the luminance terms in Fig 9.2 yields predictable 3-D illusory percepts. In the remaining stereograms in Fig 9.2, surface separation occurs at EFG and H with crossed fusion. Manipulation of the luminance values adjacent to EFG yields predictable change in the percepts yielded when the image pairs are fused.

Crossed fusion of the L-M pair in Fig 9.2d yields a percept where an opaque black disk stands forward of four partly occluded squares. Illusory contours spread from surface separation at the black-white contour in EFG and H. Surface spreading is across black spaces at ABC and D.

Fusing the M-R pair sees the disk stand behind the black sectors. The sectors appear to complete as mis-shaped opaque squares. Surface separation arises at the black-white contours bounding those four squares.

In Fig 9.2e a grey transparent disk stands forward when the L-M pair is fused. Surface separation occurs at EFGH, that is, at the white-grey contours. A grey illusory surface spreads across the black squares (at ABC and D). The near grey surface is transparent since all grey-black contours are visible through the near layer. Their position and orientation are constrained at zero disparity. Crossed fusion of the M-R pair yields the percept of an opaque grey disk behind the four black squares. Surface separation occurs at the bounds of the squares.

By manipulating luminance values in (9.2f), surface separation is precisely equivalent to (e) (at EFG and H) when the L-M pair is fused. No illusory contrast spreading occurs. However, illusory contours still cross the grey squares at ABC and D. The effect is a transparent glass like appearance. Fusion of the M-R pair leads to surface separation at the grey-white contours. A black disk stands behind the grey squares.

Finally, similar point disparities exist in Fig 9.1g. Surface separation still occurs at EFG and H resolving disparate subtense and point-disparities. The percept has the appearance of a near "porthole" through which are seen four white sectors on the distant depth plane. This seems counter-intuitive until one recognises that that disparate subtense is precisely the same in Figs (d, e, f and g). Fusion of the M-R pair sees four white sectors standing forward of an opaque black circle. The percepts are constrained by subtense and the sign of those binocular differences.

In sum, this section has attempted to demonstrate how some diverse 3-D illusory percepts appear to be related to some basic stereoscopic mechanisms that collectively can be described as binocular image processing. Those mechanisms are Image Registration, Vergence Lock, Image Comparison, Surface separation and Spreading, and a higher-order Confirmatory process.

## 9.2 Binocular image processing and Nakayama's "bent" cruciform

This section addresses a percept introduced by Nakayama and Shimojo (1992) in the terms of binocular image processing. It gives a possible alternative process based on the BIPASS model. Nakayama and Shimojo showed a simple disparate cruciform to some 200 students. They found a universally repeated asymmetry in 3-D perceptual organization between directions of disparity involving perception of illusory contours at crossed disparity but not at uncrossed disparity.

Nakayama and Shimojo explained the difference in percept by the Principle Of Generic Sampling (POGS). Recall that POGS is the idea that the system interprets untextured half-images according to the likelihood of the arrangement of surface layers in the natural world. Nakayama and Shimojo called this, *inverse ecological optics*. By this they mean that the perceptual response to ambiguous arrangement of luminance is a product of soft-wired neural populations whose associations are developed through image sampling , that is, neural network training.

The stimulus in question is reintroduced in Fig 9.3a. Vertical arms of the cross (half-image M in this case) are disparate, that is, shifted left or right. Crossed fusion of L-M sees the vertical arms of the cross stand forward of the horizontal arms. Illusory contours traverse the horizontal arms. Crossed fusion of the M-R pair (or uncrossed fusion of the L-M pair) sees the horizontal arms of the cruciform bend forward. No illusory contours form and there is no separation of surface planes in a depth step.

A BIPASS model can account for the difference between percepts in terms of stereoscopic mechanisms. Figure 9.3b, depicts a crossed disparity configuration at the white-black contour at the vertical arms of the cruciform. That is, at the intersection of the cross arms. Disparate points on those contours are labelled Q and O.

Split-projection configurations at $Q\_Q_R\_Q_L$, and $O\_O_R\_O_L$ resolve disparate subtense at the vertical arms. The split-projection configurations resolve disparate subtense along each horizontal arm in separating surface layers. The illusory contours generated in cross fusing

the L-M pair represent the bounds of a separated surface layer spreading between split-projection configurations.

Figure 9.3c is a sketch of the percept achieved at the opposite sign of disparity. The system assigns points Q and O to a more distant depth plane than the ends of the horizontal arms of the cross (points $P_1$ and $P_2$). The uncrossed visual projections from Q and O preclude surface separation. Along the vertical arms of the cross both the black and white regions (surfaces) share the bounding contours but at the same depth pane.

Being horizontal to the line of sight the 3-D orientation of the horizontal cross arms is unconstrained (can be interpreted as slanted). Disparate subtense of the horizontal arms can be interpreted as seen slant. It does not seem to be necessary to invoke *inverse ecological optics* to explain these percepts.

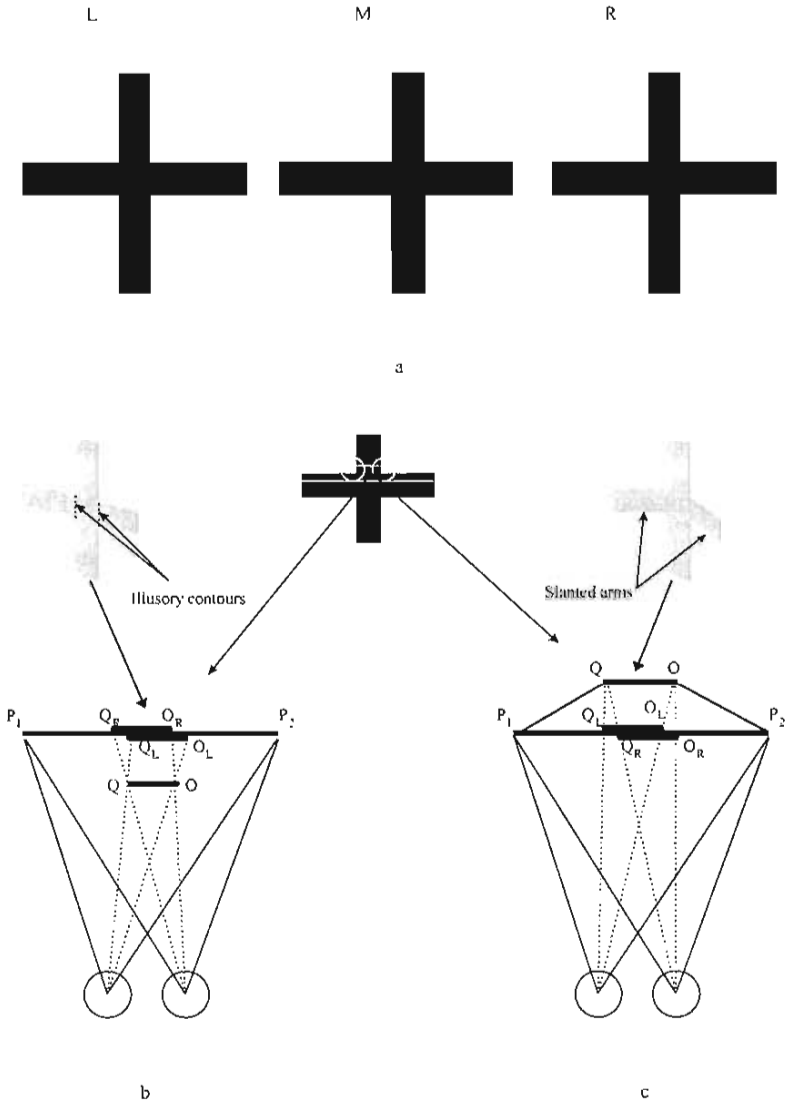



**Fig 9.3. Surface separation and perceptual asymmetry in a simple cruciform**
Crossed fusion of L-M in 9.3a (Anderson and Julesz, 1995) generates the perception of a vertical arm visibly occluding the horizontal arm. Illusory contours form between contours at which stimuli are assigned into two layers. Crossed fusion of the M-R pair sees the horizontal arms of the cross bent toward the observer. There is no separation of surface layers. Figures 9.3b and c are sketches of the array subtending the half-images. In (b) contours marking the junction between horizontal and vertical arms of the cross, are labelled Q and O. Crossed fusion of these junctions relative to $P_1$ and $P_2$ creates surface separation. Points Q and O form the apex of a split-projection configuration. For point Q, say, this resolves disparate subtense $[Q_L\_P_1 - Q_R\_P_1]$. The same is not possible when disparity is reversed. In 9.3c, Q and O represent uncrossed disparities relative to $P_1$ and $P_2$. No surface separation arises. No illusory contours result. Slant in the horizontal arms of the cross is generated by the binocular subtense $[Q_L\_P_1 - Q_R\_P_1]$ and $[O_L\_P_2 - O_R\_P_2]$ (Nakayama and Shimojo, 1992).

This description of surface separation and spreading in Nakayama's cross is simply one possible account of perception. Nevertheless, the BIPASS model predicts some quite counter-intuitive effects from simple manipulations of the stereogram. For example, in Fig 9.4

introduces just two vertical lines at the horizontal arms of the cruciform. The inner luminance contour bounding these vertical lines become contours along which the system separates surface layers, that is, resolution of disparate subtense by splitting the percept into two layers (at a split-projection configuration).

In phenomenological terms, crossed fusion of the L-M pair in Fig 9.4 sees the vertical arms stand forward of the horizontal arms (and the vertical lines) as before. However, in the M-R pair, crossed fusion does not yield the percept of a folded cruciform. The horizontal arms of the cross do not slant. Instead, surface separation occurs along the vertical lines. Hence, observers see a depth step at the inner black-white contour bounding the vertical line. Fusion assigns white regions *between* disparate contours to the distant depth plane.

Introduction of the vertical lines leads to a depth percept that is equivalent to 3-D perceptual organization in the *porthole* effect. Important visual projections (that projections from luminance contours) are sketched in 9.4b and c. The drawing at (c) highlights the observation that surface separation resolves disparate subtense along the horizontal arms of the cross.

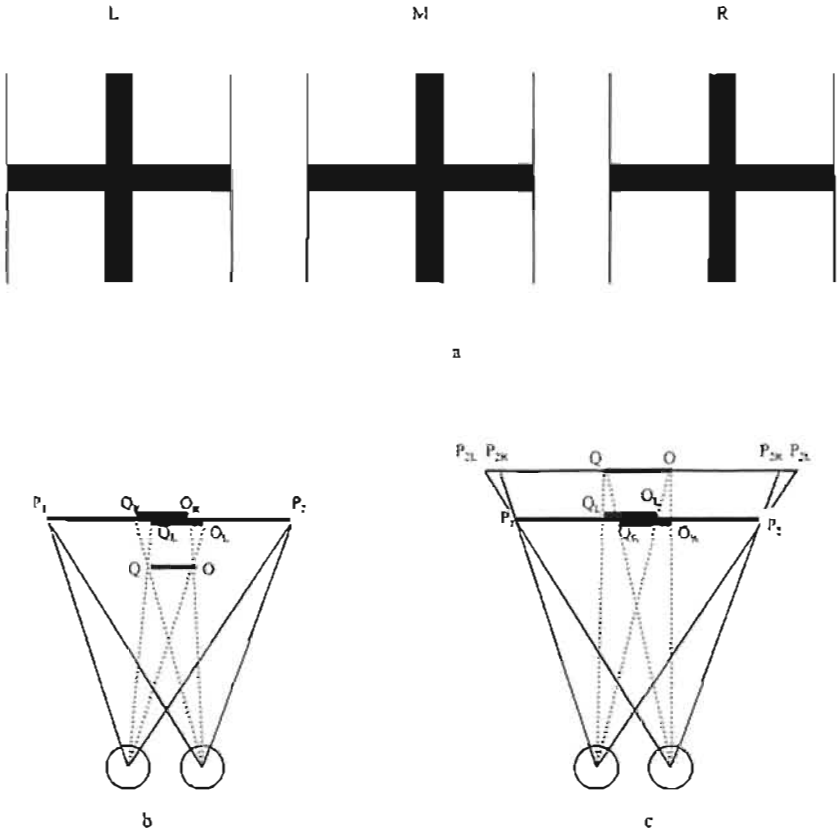



**Fig 9.4 Surface separation in a stereoscopic cruciform**
In 9.4a, crossed fusion of L-M yields a percept where the vertical arms of a central cross stand forward of the horizontal arms. Crossed fusion of M-R sees all contours except the two thin vertical lines assigned to the distant depth plane. These percepts are described using a sketch of visual projections in (b) and (c). Fig 9.4b shows that addition of two thin vertical lines to the cross has little impact upon the perceived organisation of surface layers. However, in 9.4c, a split-projection configuration of projections is supported at both $P_1$ and $P_2$. Thus, all of the space between those contours is assigned to the distant depth plane resolving disparate subtense.

## 9.3 A BIPASS model for the Anderson and Julesz oblique cross

Anderson and Julesz (1995) also addressed Nakayama and Shimojo's simple figure. To Anderson and Julesz the percepts showed that no unmatchable monocular features *disambiguate* occlusion geometry in Nakayama's cruciform. They described the luminance arrangements as *ambiguous L-Junctions*. To show the ambiguity they simply rotated Nakayama's cruciform $45^0$ before applying disparity. In this section, I briefly revisit the Anderson and Julesz explanation in reference to what will be termed the *oblique* cross.

Figure 9.5 shows the oblique cross with one arm shifted horizontally (in M). Upon fusing the L-M and M-R pairs, there is no phenomenal asymmetry between the two directions

of disparity. At each sign of disparity, illusory contours cross the more distant arm in an equivalent position.

Anderson and Julesz claimed that this was because the vertical image differences present were unmatchable due to an epi-polar constraint on vertically disparate elements of the retinal images. They argued that the unpaired components of the oblique cross arms were interpreted according to inference about *causal* geometric structure. Anderson and Julesz may have overlooked disparate subtense in the *length of the oblique cross arms*, disparate subtense *between the cross arms* (white space) and *orientation of the oblique arms*. The orientation of the oblique contours means they are highly unlikely to look *bent*, that is to look slanted.

In the functional terms of the BIPASS model, a split-projection configuration along the near oblique arms of the cross will resolve disparate subtense between the arms of the cross by surface separation (at both signs of disparity). Spreading of the near surface layer between the split-projection configurations will yield stereoscopic illusory contours across the distant arm of the cross. Higher order confirmatory process such as completion may (but not necessarily) facilitate continuation of that separated surface layer.

It is not difficult to show that non-epi-polar features are not necessary to generate illusory contours or surface separation the oblique cross. For example, in Fig 9.5b the oblique cross has been overlaid with a rectangle that negates vertical disparities. Now, no vertical disparities exist in the half-images. Yet, fusion maintains symmetric organisation of surface layers between the two signs of disparity. That is, stereoscopic surface separation occurs at equivalent contours at each sign of disparity.

I would argue that the reason for the perceived similarity between percepts in the oblique cross is not a matching constraint but a constrained relationship between subtense and orientation. The oblique orientation of the cross arms means that orientation "parity" in each eye renders them unambiguously fronto-parallel. Resolution of disparate subtense in the split-projection configuration will be therefore be stable. The percepts generated appear to be logical outcomes of stereoscopic mechanisms outlined in the BIPASS model.

Explanation of the oblique cross percepts from a BIPASS account is quite simple. The key is to recognise the importance of the magnitude of white spaces between the arms of the cross in the retinal images. At both signs of disparity a split-projection configuration along the near arm of the cross will resolve that disparate subtense. Figures 9.5c and 9.5d demonstrate this possibility.

Several non-disparate arrows have been inserted in Figs 9.5c and 9.5d that in each stereogram point to the contour along which surface separation occurs. In (c) surface separation occurs along the arm BC. Note that the gap between the arrowheads and arm BC, in the half-images, is the magnitude of the disparate subtense in the white space between the arms. The same effect is evident in Fig 9.5d. Here surface separation arises along the white-black contours at each side of the AD arm. The non-disparate arrows are intended to demonstrate the magnitude of disparate subtense.

Since separation of surface layers is achieved along the near arms of the cross, it is spreading of the near surface layer between these black-white contours, that generates illusory contours. It does not seem to be necessary to invoke *a priori* constraints or *causal* relationships to explain the percepts in the oblique cross. Substantial constraints exist in the 2-D patterns of image differences contained in the half-images.

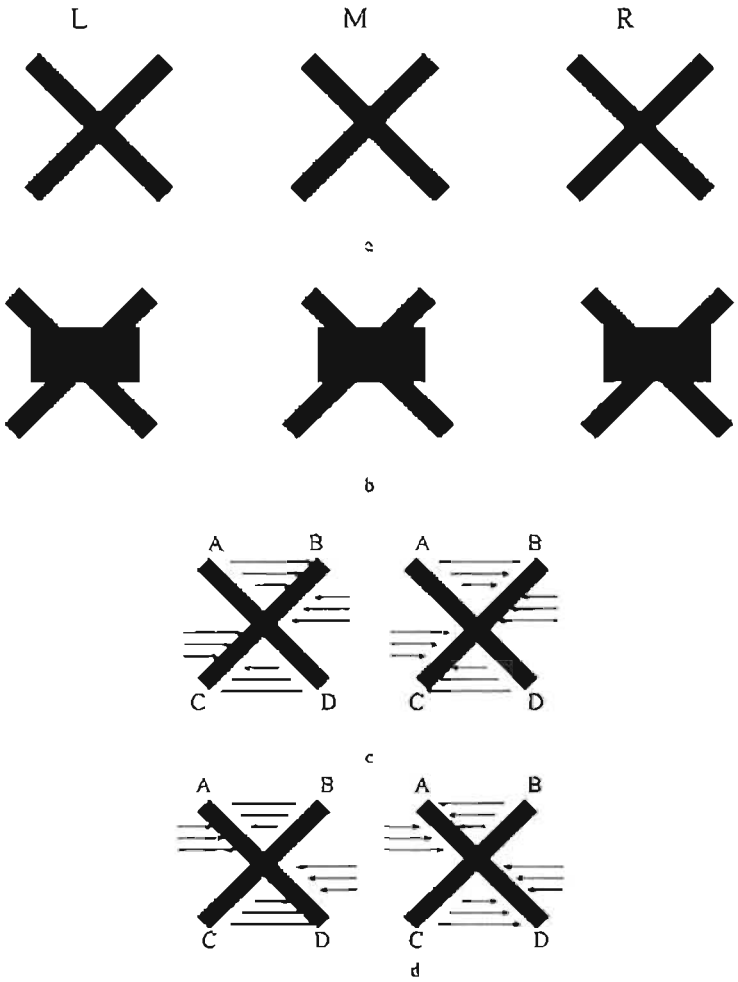



**Fig 9.5. Anderson and Julesz Oblique Cross**
In (a), crossed fusion of the L-M and M-R pairs yields no perceptual asymmetry between the signs of disparity when the half-images are fused. Illusory contours continue cross the distant rectangle in both instances. The near arm appears to be *modally* completed. In Fig (b), a similar effect can be generated where a rectangle overlays all vertical disparities. Figure (c) shows the contour along which the separation of surface layers arises (arm BC). Fig (d), the reversed sign of disparity yields surface separation along the arm AD with crossed fusion. In both percepts the across are non-disparate. The gap between the arrowheads and the appropriate arm is equal to the magnitude of disparate of disparate subtense in the space between the arms. It is surface separation in a split-projection configuration of visual projections that yields the depth step in these stereograms. It is spreading of the near surface layer across between the two sides of the near arm of the cross that produces the illusory contours.

## 9.4 Binocular image processing and Anderson's stereoscopic I-junctions

This section attempts to address a stereoscopic stimulus first demonstrated by Anderson (1994). The percepts produced are crucial for understanding stereoscopic illusory contours. Generation of the percept clearly involves no completion process and no continuation process

*per se*. Nevertheless, a BIPASS model of stereoscopic mechanisms may account for Anderson's I-junctions.

*Phenomenology associated with Anderson's I-junction*

Anderson and Julesz argued that I-junctions, where a vertical disparity exists, were evidence that the system decomposes the retinal images into matchable and non-matchable features. Unmatchable portions of the vertical lines reveal occlusion geometry to the system. The system assigns these portions of the line to a distant depth plane. They are interpreted as partly occluded.

Anderson's images are re-introduced in Fig 9.6a. Crossed fusion of the L-M pair yields the percept of an illusory contour at the tips of the vertical line. Anderson (1994) initially described the contours in terms of Bayesian inference based on the notion that conditional priors (luminance junction patterns) existed. Those priors dictated the response of the system. However, Anderson (1997) sees the illusory contour as evidence of an end-cut mechanism. An end-cut mechanism is the output of orientation sensitive hypercomplex (also called end-stopped) receptive fields (after Grossberg, 1994). The end-cut mechanism proposed requires that the contours are orthogonal, *or nearly* orthogonal to the orientation of the contour inducing the *end-stopped* signal.

These intriguing *illusory* contours clearly do arise at a different depth plane from the fused vertical lines. They seem to fall on a near depth plane. Moreover, they appear to mark separation between near and distant surfaces, that is, a depth step.

It is interesting to note that in all of Anderson's demonstrations using vertical lines, the illusory contours formed are not orthogonal to the vertical lines. In the configuration shown, for example, the illusory contours don't appear orthogonal but are substantially rotated about the z-axis of projection. Only when the disparate lines are aligned horizontally (by rotating the entire stimulus $90^0$ see Anderson, 1994) do the illusory contours appear orthogonal to the horizontal lines. Figure 9.6b shows the perceived orientation of the illusory contours. This appears related to the eye (the half-image) in which the longer vertical line is presented. The small squares are *fusion locks* used to constrain vergence eye movements.

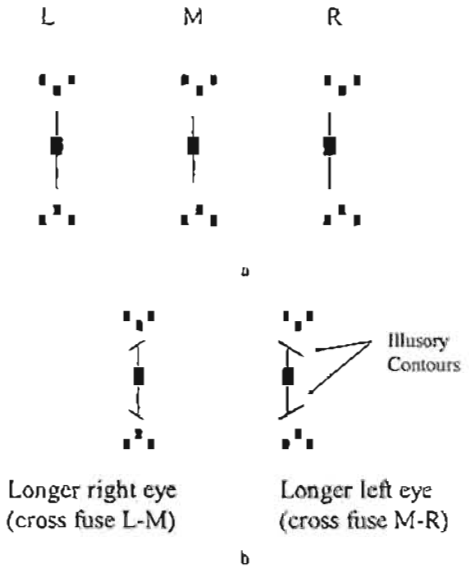



**Fig 9.6. Illusory contours induced at I-junctions**
In (a) crossed fusion of the L-M pair sees an illusory contour generated at the tip of the vertical lines (Anderson, 1994). With the vertical line in right half-image longer than the left, the contours is rotated about the Z-axis by about -45 degrees as shown in (b). Cross fusion of the M-R pair sees the contour rotated in the other direction (45 degrees), as shown in (b).

*Toward a binocular image processing account of the I-junction percept*

In natural vision, an I-junction might happen if we looked through a round window at some vertical object, say a tree or the corner of a building. Because of binocular parallax, the 2-D layout of the half-images may contain a vertical difference in the magnitude of the object in each eye's view within the frame of the window. One eye may see a longer distant edge than the other (disparate binocular subtense). This can be simulated using a very simple stereogram. In Fig 9.7a an ellipse has been drawn and within it a vertical line. The line has a disparate position and disparate length in each half-image just the same as Anderson's I-junction. The percepts generated in fusing the ellipses can help to explain the generation of illusory contours at the tips of the I-junctions above.

Crossed fusion of half-images L-M in Fig9.7a sees the line stand at a distant depth plane relative to the ellipse and fusion of Fig 9.7b yields the percept that the line stands forward. This appears reasonably straightforward. However, careful comparison of the two percepts reveals a subtle perceptual asymmetry between the signs of disparity that has been common to all the untextured stereograms so far examined.

When pair L-M in Fig 9.7a is cross fused the line looks like it is visible though a porthole. All of the white region within the ellipse has been *captured* to the depth of the line.

The BIPASS model has suggested that such effects emerge when the system splits the percept in two in a way that resolves disparate subtense returned from image comparison. So in the ellipse in Fig 9.7a (L-M) uncrossed disparity in the position of the line precludes a split-projection configuration along the line. Instead the system splits the percept into two layers at the bounds of the ellipse. Where the line is separated from the ellipse at points AB, the visual projections from the line to each eye must cross in a split-projection configuration. The disparate subtense of the line is resolved as the line stands behind the near bounds of the ellipse.

When pair M-R are cross fused in 9.7a the line stands forward of the ellipse. In order to resolve disparate subtense created by the disparate position of the line the system must split the percept into two layers along the white-black contours on either side of the vertical line.

Stereograms at Fig 9.7b are intended to show how the I-junction is related to the separation of surface layers in the ellipse. It was suggested that to resolve uncrossed disparity in the position of the line relative to the ellipse that visual projections at the line ends must cross at AB. This is the case as well for Anderson's I-junctions. Hence binocular image processing splits the percept into two layers at the projection plane in both the ellipse and the I-junction stimuli. The illusory contours created in the I-junction are created as the system separates surface layers at an oblique split-projection configuration where the distant surface shares a near surface edge.

Crossed fusion of the M-R pair yields no separation of surface layers at the I-junction and hence no illusory contours. This is because disparate subtense is resolved in a split-projection configuration along the vertical line.

In summary, the projection drawing in Fig 9.7c is intended to help explain how resolution of disparate vertical subtense in the vertical line is resolved by splitting the percept into two layers at AB (the very tip white-black contour) where visual projections to the distant layer must cross, as the line is fused. The oblique orientation of the line is predicted by the degree of separation of the eyes.

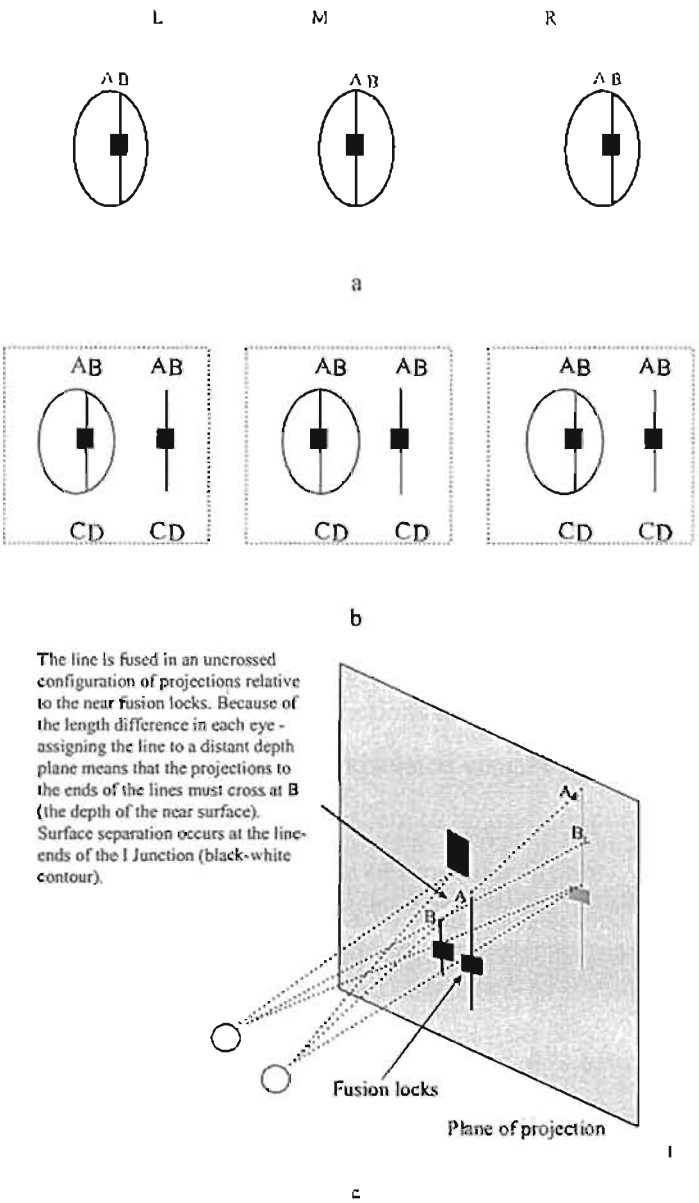



**Fig 9.7. Stereoscopic surface separation at I-junctions**
In (a) crossed fusion of the L-M pair yields the percept that the vertical line stands behind the ellipse. The white space within the ellipse is assigned to the distant depth plane. This is because disparate subtense between the line and the ellipse bounds is resolved in a split-projection configuration at the inner bounds of the ellipse (see text for details). Crossed fusion of the M-R pair sees the line stand forward of the ellipse. In (b) equivalent percepts are generated fro the ellipse and the eye junction. The degree of disparity is identical at both sings. The letters A and B mark sites along the contour bounding the ellipse and at the end of the disparate lines. To fuse these lines along the same *line of sight* in cyclofusion visual projections must cross at AB the luminance step must be shared between the near and distant surfaces. In this manner, binocular geometry predicts that the orientation of the illusory contour in the I-junctions is a product of the separation of the eyes - binocular parallax.

Evidence for the proposed binocular image processing applied to these I-junctions can be demonstrated using some simple images that introduce step-wise manipulation of the disparities present in Anderson's I-junction stimuli.

First, in 9.8a, in the absence of fusion locks, lines of different lengths do not invoke strong illusory contours. This may be because the fusion locks actually constrain cyclofusion. The unconstrained lines may induce binocular rivalry - illusory contours will therefore be unstable - since fusion is unstable.

Figure 9.8b shows that thickening the lines causes quite strong rivalry at their ends. Illusory contours will not arise. In Fig 9.8c, vertical differences exist but in the *context* of obliquely oriented contours. Strong illusory contours are evident, but only in the L-M pair. In the M-R pair (uncrossed disparity), there is a change in the site of surface separation to the vertical bounds of the thickened line. This is an intriguing effect noticed by Howard and Rogers (1995).

Next, in 9.8d and 9.8e, a series of these thickened lines joined together can form a triangular shape that takes the appearance of an aperture when fused. Vertical disparities exist at all vertical black-white contours. Disparate I-junctions arise at every vertical luminance intersection. Resolution of the I-junctions is not an isolated visual event though. Disparate subtense and the mechanism of surface separation seem to be more important than a matching constraint in constraining the phenomenal outcome of binocular fusion.

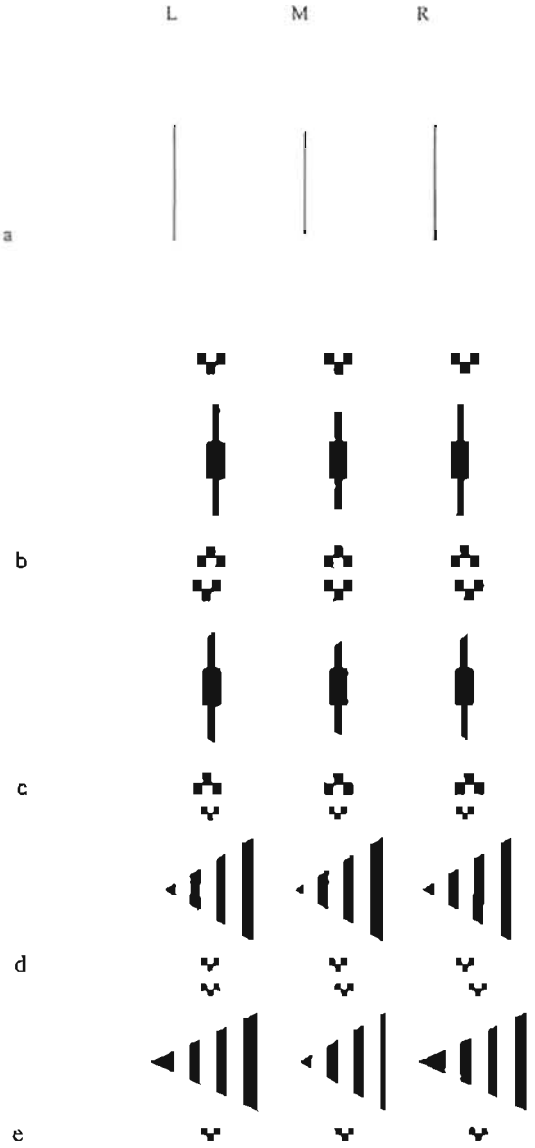



**Fig 9.8. Vertical image differences and stereoscopic I-junctions**

In (a) the absence of fusion locks lessens or negates the tendency to perceive illusory contours emanating from vertical image differences. When the I-junction lines are increased in thickness rivalry occurs. This rivalry is resolved if the vertical differences are combined with orientated contours at the tips of the I-junctions as in (c). Here crossed fusion of the pairs in L-M reinstates separation of surface planes which invokes the illusory contour but in the opposite sign of disparity (fuse M-R), the site of surface separation is now along the vertical boundaries of the lines. Finally in (d) and (e) these thickened I-junctions are arranged in a manner congruent with the perception of a triangle. In (d) the opposite signs of disparity show the feasibility of the two alternative sites of surface separation (at the outer perimeter of the triangle). In (e) surface separation occurs according to the contours at which crossed configurations of projections are created in cyclofusion yielding a stereoscopically slanted triangle.

In summary, a BIPASS model appears to contribute toward understanding Anderson's intriguing stimuli, in a way that is parsimonious with previous discussions.

## 9.5 Binocular image processing and "Stereo Capture"

*Stereo Capture* is another intriguing perceptual outcome of binocular fusion in untextured stereograms. Stereo Capture is a *wallpaper effect* where, the perceived depth of an illusory form captures a periodic pattern of luminance (see for example, Ramachandran and Cavagnah, 1985). Ramachandran for example argues that stereo capture occurs as the brain takes a short cut by inferring depth where disparities are sparse. I believe that binocular image processing mechanisms may help define the depth of the "captured" layer after all[25].

The basic phenomenon is re-introduced in Fig 9.9. Figure 9.9a are stereograms that invoke stereo capture when fused. The Kanizsa square in (M) has been shifted to the right about 2mm. This creates disparate shaped pacmen. A set of horizontal lines overlays the Kanizsa squares.

Crossed fusion of the Kanizsa squares in the L-M pair generates the percept that a near illusory surface stands forward of the pacmen. The bounds of the Kanizsa square, seems to capture the horizontal lines to the depth plane of a near surface. This is the classic stereo capture effect.

At the opposite sign of disparity, fusion of the M-R pair, a typical porthole effect arises. The system assigns four segments, within the bounds of the pacmen, to a distant depth plane seen through the porthole. The four sectors within the bounds of the portholes capture the horizontal lines to the distant depth plane a well.

I would argue that the stereo capture effect seems to support a BIPASS model of the 3-D illusory percepts. In fusing L-M, the system assigns horizontal lines to the near depth plane. My proposal is that the system splits the percept into two layers at a split-projection configuration at the mouths of the pacmen. The separated near layer spreads between pacmen and across the sites at which the horizontal lines abut the pacmen.

In fusing the M-R pair, no surface separation occurs at the pacman mouths. Instead, it occurs at the white-black contour at the bounds of the pacmen. Again, surface spreading cuts

---

[25] It turns out that at least one of four readers of this project prior to submission did not "see" the "stereo capture" percepts generated as I will describe them. There are clearly more complex issues of ambiguity present than my simple descriptive model addresses at this stage.

across the horizontal lines. Stereoscopic surface separation at the pacman bounds therefore assigns the portions of the horizontal lines within the pacmen to the distant depth plane.

Manipulating the horizontal lines within the Kanizsa squares can support these quite simple ideas. Horizontal lines are not actually "captured" by the illusory surface layer but assigned according to displacement of surface layers depth in surface separation.

Figure 9.8b shows such a manipulation. Now, the horizontal lines are not disparate in position or length. This leaves a gap between the lines and the vertical contour at the pacman mouths. The gap is precisely reciprocal to the magnitude of disparate subtense at the pacman mouths. Fusing pair L-M, yields the percept of a transparent illusory Kanizsa square. The horizontal lines stand at the depth plane of the disks. Surface separation still occurs at the mouths of the pacmen. No "capture" arises.

Fusing the M-R pair sees the same effect. The system does not assign the lines to depth. Kanizsa square *portholes* still emerge. Horizontal lines stand on the depth plane of the pacmen bounds and the portholes do not *complete*. A split-projection configuration at the bounds of the horizontal lines is the site of surface separation. That is, surface separation must here occur at the white-black contour at the line-ends (as in Anderson's so called I-junctions).

Another way to demonstrate the source of capture is to use a non-disparate Kanizsa figure and then manipulate the depth of the horizontal lines. In Figure 9.8c there is no disparity in the Kanizsa squares themselves. The position of the horizontal lines relative to the Kanizsa square have been manipulated so that now the horizontal lines are disparate.

Fusing the L-M pair (9.8c) sees the horizontal lines stand forward of the Kanizsa square. Notice that the system also assigns the space between the lines to the near disparity plane. This is because the gaps between the line ends and the pacman mouth boundaries are resolved within a split-projection configuration at those lines ends yielding surface separation. The near surface layer spreads between the line-ends. The percepts looks like an opaque striped surface standing forward of the pacmen.

Disparities are the same magnitude but opposite direction in the M-R pair (in c). Since no disparity exists at the mouths of the pacmen - no porthole effect emerges. The system assigns the horizontal lines (an the space between them) to a distant depth plane. Separation of surface layers occurs at the mouths of the pacmen.

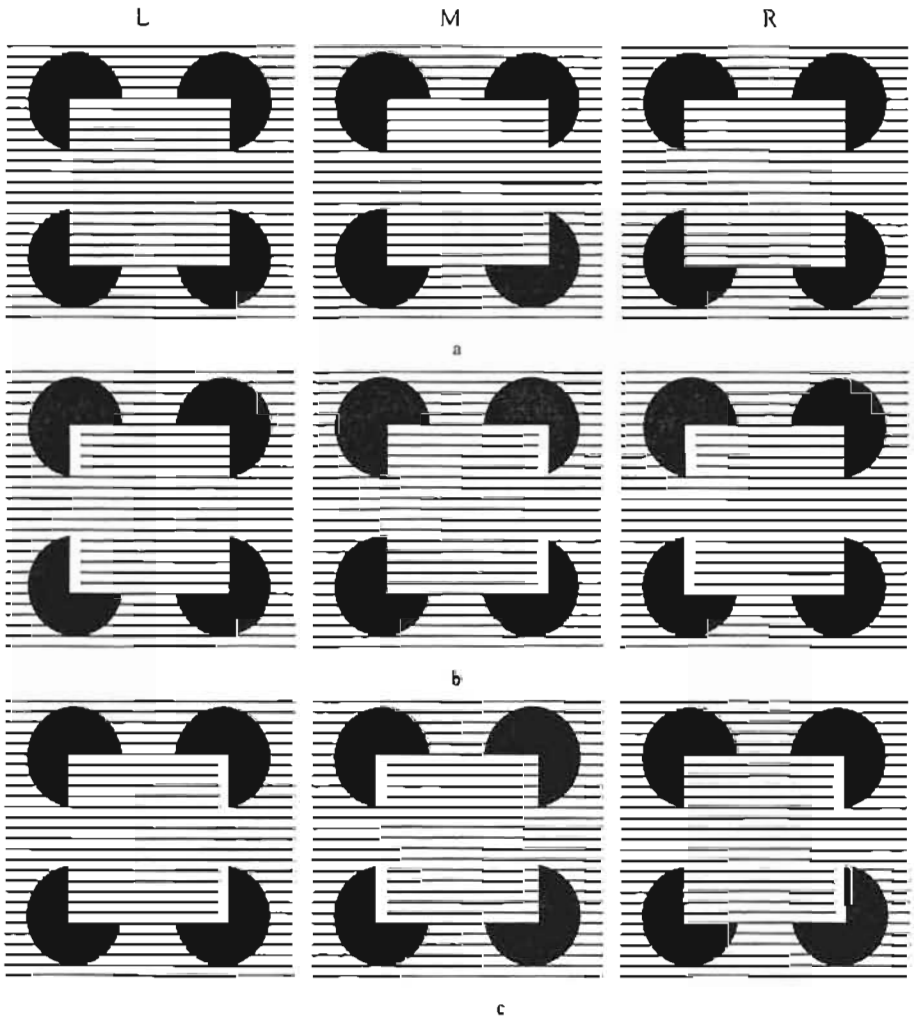



**Fig 9.9. Separation and stereo capture**
In (a) fusion of the L-M pair demonstrates the classic stereo capture effect. The horizontal lines bounded by the Kanizsa square are assigned to a near depth plane. Fusion of the M-R pair in (a) sees the porthole effect where horizontal contour segments bounded by the horizontal contours are captured to the distant plane. The porthole appears to be completed across near horizontal line assigned to the depth of the projection plane. In (b) the horizontal lines have been manipulated so that they are non-disparate. In fusing the L-M pair a transparent illusory square is seen. Fusing the M-R pair sees the porthole effect arise. Horizontal contours are assigned to the depth of the pacmen - as are the white spaces between the lines. In (c) disparity in the position of the horizontal lines has been introduce while the Kanizsa squares are non-disparate. Now fusion of the L-M pair sees the horizontal lines stand forward. Crossed fusion sees the horizontal lines stand behind the pacmen (see text for details).

One final demonstration in this section shows that oriented random-lines such as those chosen by Gillam (1995) *resist* capture. Obliquely oriented contours bounding these lines

provide unambiguous information regarding distance and orientation. Both the orientation of the lines themselves, their bounding luminance contours and the spaces between them are all tightly constrained to zero disparity (or at least to the same disparity plane).

In the top set of half-images in Fig 9.10, a glass like surface film stands forward of the projection plane, above the random line surface (cross fuse L-M). There is no surface capture, yet the separation of surface planes is maintained as previously described in the Kanizsa Square. In the reversed disparity case (cross fuse M-R), the lines remain on the near surface while the typical porthole effect is seen.

In the bottom, set of half-images, disparity applied to random line elements within the Kanizsa Square itself generates a separation of the surface layers. Those lines stand on the near occluding surface, in the case of crossed disparity (cross fuse L-M), and on the distant surface in the uncrossed example (cross fuse M-R).

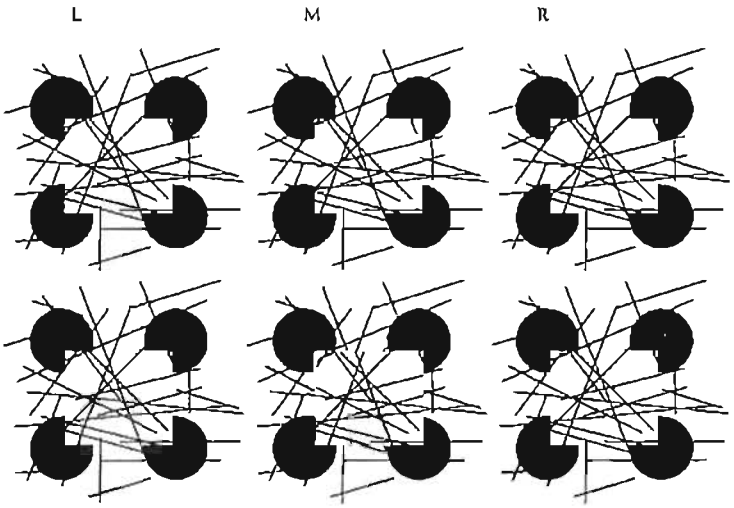



**Fig 9.10 Surface features and failure of stereo capture**

In the top set of half-images (Gillam, 1995), crossed fusion of L-M yields the percept of a transparent or glass like illusory surface standing forward of the projection plane with the unambiguously oriented random features lying at the level of the pacmen. Both the curvatures of the pacmen, the lines, and the spaces between the lines constrain this surface separation. Fusing M-R sees the porthole effect developed, but with the random lines not captured at the pacman mouths. In cross fusing L-M in the lower set, disparity in the position of the random line matrix has been applied. This figure has the appearance of a Kanizsa square with random lines standing forward of the disks or behind the pacmen in fusing M-R.

In the next chapter, I conclude that the mechanisms underpinning an BIPASS model may indeed apply to textured surfaces and to stereopsis in a natural setting.

# 10 Concluding remarks on binocular image processing and surface spreading

*Summary: This chapter reviews experimentation and the development of a BIPASS model of SKS and SES percepts. It then explores some possible implications of the model for understanding untextured stereograms and for stereopsis in the natural setting.*

## 10.1 Research summary and development of the BIPASS model

This project concludes that some basic stereoscopic mechanisms underpin perception of the 3-D illusory percepts. The binocular vision system has access to sources of inter-retinal differences other than point disparity and this has been overlooked in previous understanding of stereoscopic illusory contours and surfaces. My experimental findings suggest that the Surface Heuristic account of stereoscopic illusory contours and surfaces, based on inferential perceptual schemes, has over emphasized the ambiguity of untextured stereograms for stereopsis. Similarly, the Form Computation approach appears to have over-emphasised the importance of integrative visual mechanisms. My evidence, from experiment and demonstration, is that the relative physical dimensions of untextured features in the half-images comprising the stimuli, constrain the stereoscopic response considerably. Stereoscopic illusory contours and surfaces appear to be intimately related to a stereoscopic response described in functional terms by the BIPASS model. This model is a functional description of processes that appear to involved.

This thesis presented nine simple experiments that, together with a number of demonstrations, suggest something quite important about stereopsis. Six experiments studied the SKS percepts. Three experiments examined the SES percepts. All experiments used seen slant as a metric of the stereoscopic response.

Seen slant allowed comparison of the percepts experienced, when untextured stereograms are fused, against a stereoscopic comparison image that simulated 3-D surface orientation. This enabled analysis of the way that the system must resolve large-scale, disparate, untextured features and regions in the half-images. Moreover, phenomenological characteristics of the 3-D illusory percepts, such as opacity and transparency, also appeared to be related to the manner in which the system resolved disparate feature-scale inter-retinal image differences.

I have concluded that the findings in this set of nine experiments cannot be solely explained by the inferential visual responses, as suggested by theorists such Anderson, Nakayama and their collaborators. These authors emphasise local factors in surface separation. Nor can they be solely explained in terms of feature integration processes described in Grossberg's FACADE theory. My view of the stereoscopic response is that it is sensitive to disparity at multiple feature-scales simultaneously (parity and disparity: in the relative size, position and separation of features in the field of view). The 3-D illusory percepts seem to emerge as the system resolves the 2-D layout of those disparities by separating surface layers at particular luminance contours. The BIPASS model is a description of the processes that appear important.

## 10.2 Possible implications of the BIPASS model for understanding stereoscopic illusory contours

The experimental findings and analyses mostly apply to stereoscopic illusory contours and surfaces. We should not under-emphasise the stereoscopic information available even in completely untextured stereograms. However, there is a larger context for this discussion as well. The basic principles that underpin the BIPASS model might represent a guide toward understanding a range of stereoscopic percepts. Recall the work of Gulick and Lawson (1976). Gulick and Lawson addressed the question of texture density in stereopsis. In fact, they discovered that "stereoscopic contours" were seen in stereograms containing only sparse textures. Gulick and Lawson argued that texture density was a continuum that linked sparsely textured stereograms with Julesz's Random-Dot-Stereograms.

In Gulick and Lawson's stereograms (see, for example, Fig 10.1a) disparity was introduced by shifting the position of just one set of black squares in a uniform matrix. Crossed fusion of the L-M pair in 10.1a sees a white near surface layer floating above the

matrix. Gulick and Lawson called the edges of the near layer "stereoscopic contours". As you fuse these images notice that there are two rows of squares adjacent to the depth step - but only the outer set of squares is actually binocular.

Next, cross-fuse the L-M pair in the RDS (10.1b). You will see a near surface layer. A depth step occurs at its bounds. Gulick and Lawson said that the boundaries of the surface layers were *stereoscopic contours* derived from the same process of form extraction as in the sparse matrix. If the figure is analysed carefully, monocular regions occur in precisely the same parts of the dense matrix as in 10.1a.

Gulick and Lawson thought that the mechanisms by which a depth step was achieved in the sparse matrix must be precisely the same as in the RDS. Evidence, from experimenting with the SKS percepts is in agreement with them. Gulick and Lawson claimed that the difference between their matrices and an RDS was that texture features were smaller and denser in an RDS. They said that the small dense texture features camouflage any obvious difference in the position of corresponding luminance contours in an RDS.

It may also be useful to think of the phenomenal outcome of fusing Gulick and Lawson's stimuli in terms of the BIPASS model. In Gulick and Lawson's matrices, the system separates surface layers at the inner most luminance boundary of the monocular set of black squares. This separated near surface layer then spreads between the squares bounded by stereoscopic illusory contours.

The surface separation mechanism resolves disparate subtense across the sets of black squares by supporting an overlap of surface layers. My argument is that the disparate matrix squares contain two components of retinal disparity, local disparity (at disparate luminance contours) and disparate binocular subtense (between luminance contours). Disparate subtense constrains binocular fusion in a sparse matrix in precisely the same manner that it does in completely untextured stereograms such as a Kanizsa square.

In fusing the Julesz, RDS, a similar separation of surface layers might arise. That is, surface separation occurs at local crossed fused luminance contours, adjacent to monocular parts of the matrix. This suggests that the actual visible edge of the near surface layer involves the spreading of a near surface layer between dots in the RDS that support crossed visual projections –resolving disparate subtense across the matrix by surface separation.

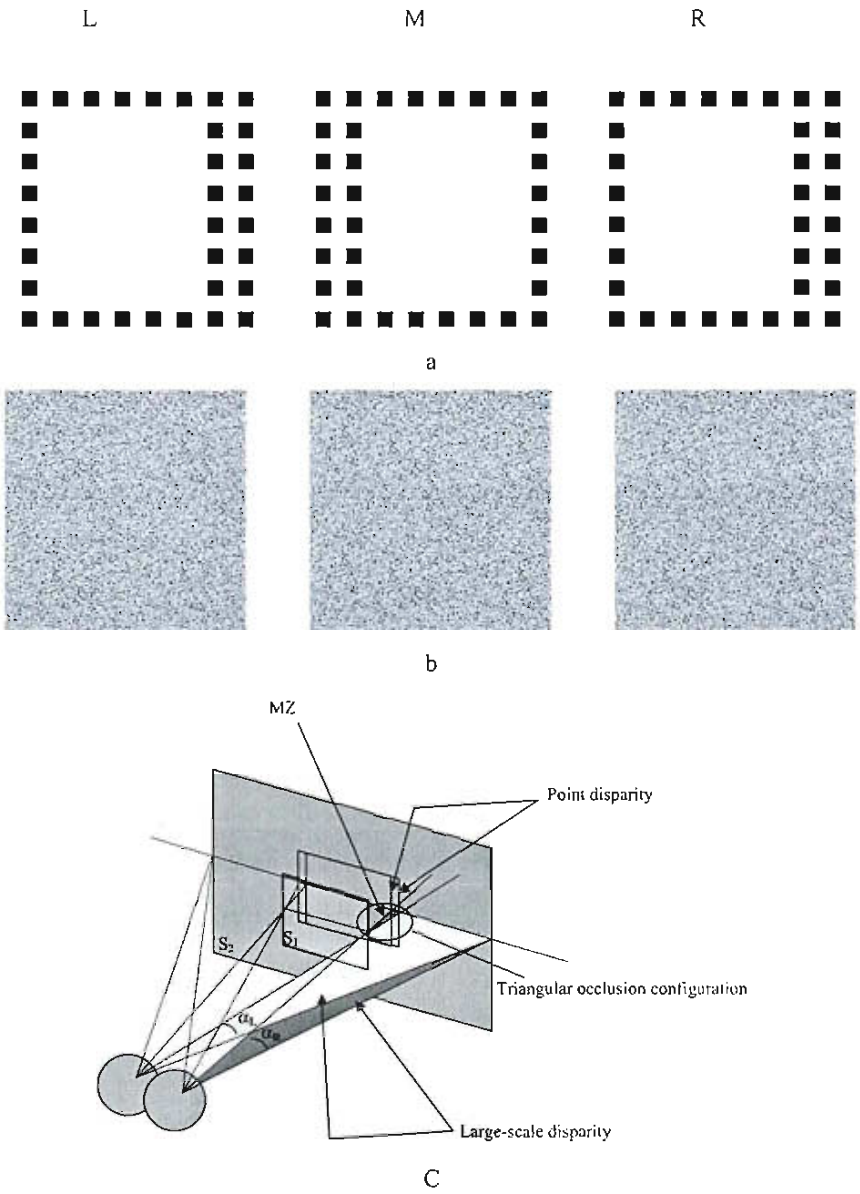



**Fig 10.1. Stereoscopic illusory contours and Random-Dot-Stereograms**
Crossed-fusion of the L-M pair in (a) or uncrossed fusion of pair M-R, sees a near surface layer stand forward of the sparse matrix. In (b), a near textured surface stands forward. Both stereograms have the same disparity – but clearly texture in the RDS is more dense. In both of these percepts (crossed fusion of L-M or uncrossed fusion in M-R for both (a) and (b)) disparate subtense between the edges of the stereograms and the near surface layer may constrain the perception of stereoscopic depth. A projection drawing in (c) demonstrates the subtense in question ($\alpha_L$ and $\alpha_R$) that must constrain the combination of half-images images. The percept is of a near surface S2 standing in front of a distant surface S1. Surface separation involves an overlap of surface layers. An MZ on surface S1.

It appears that stereoscopic illusory contours and surfaces may represent one end of a continuum of stereoscopic phenomena. The continuum involves separation of surface layers in depth. Densely textured random-dot-stereograms fall at one end of the continuum while very sparsely textured stereograms and untextured stereograms such as the stereoscopic Kanizsa figures (or even Anderson's I-Junctions) fall at the other end. But the processes underpinning the percepts are the same. A mid-point connection between the poles of the continuum might be Gulick and Lawson's sparse matrices.

## 10.3 Possible implications of the BIPASS model for understanding binocular vision in more natural contexts

In this section, I discuss a very simple contrived context of what I will call pseudo-natural binocular vision. I briefly venture to discuss the possible relevance of the BIPASS model for a more general understanding of stereopsis. Consider Fig 10.2.

The stereograms in Fig 10.2 have been constructed using a pair of stereo photographs taken of my desk. A digital camera on a tripod was used. Two photographs were taken at about 3 meters from the background wall. Two photos were taken from viewpoints about 65mm apart (commonly cited as inter-ocular distance). So, I have reproduced a pseudo-natural context of retinal disparity. I have then overlayed the photos with a white frame. The photos have been shifted relative to (behind) the frame to make a very basic point about binocular vision and the importance of binocular subtense.

Crossed fusion of these photos in (a), sees the objects on the desk visible through a window. Why do I see a window – a square porthole? Arguably, it is because the system separates a near surface layer (the page) from the background in resolving disparate subtense. By adjusting the frame in (b), the objects in view stand further forward (note that the contrived manipulation constrains vergence angle). Also, the local disparity gradients across the objects and disparities between the objects are identical. Only the magnitude of background subtense has been changed. In (b) the mug, for example stands about level with the page when the images are fused. In (c) it stands forward of the page. The point is, if all local disparities in these photos are the same, then the depth perceived stems from the resolution of relative disparate subtense across the background. Resolution is achieved by separation of surface layers to different depth planes.

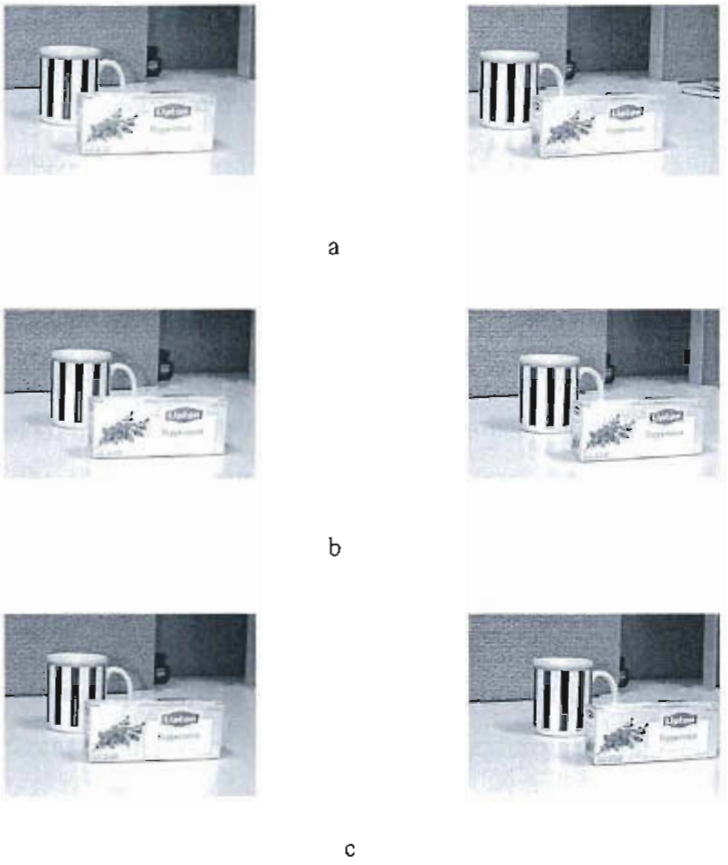



**Fig. 10.2. Disparate binocular subtense and surface separation in stereo photographs**
In fig (a) crossed fusion yields a scene viewed through a square aperture. This is an analogue of the porthole effect in untextured stereograms. Surface separation at the aperture boundary resolves the magnitude of large-scale background differences. In (b) the disparities between near objects are maintained. The disparities in back are manipulated by shifting the photo behind white frames. In (c) the background differences are negated by the same means. The porthole is therefore negated, and near objects stand forward of the plane of projection.

Also note that in each image in Fig 10.2, the contour that separates the background from, say the mug, is actually the same physical contour. Where the teabag packet partly occludes the mug – the same pattern of projections occurs. This is the mechanism of surface separation at particular contours where binocular lines of sight cross at a near surface edge (the split-projection configuration of contours). Separation occurs at the actual image boundary itself – hence the appearance of a window – in (a) and (b). In (c) there is no

disparate subtense in the background frame. In achieving surface separation, I argue, the system must resolve near surface layers in the split-projection configuration, and so resolves disparate subtense. In a manner of speaking stereoscopic depth may be a product of the resolution of both local and non-local inter-retinal differences.

I can represent this functional understanding of binocular processes by drawing a set of visual projections through the object bounds as in Fig 10.3. One way of looking at this arrangement of objects at different relative depths is as a series of surface layers. Across those layers there are local disparity gradients that define curvature and orientation. The system must separate surface layers at contours, which correspond in each image. At object boundaries, the visual projections cross (split-projection). The system seems to separate layers by achieving image overlap where disparate subtense is resolved behind the near surface layer.

For example, the edge of the mug in (a) is deduced from the position of the contours $A_R$ and $A_L$ in each respective retinal image. A crossed configuration of visual projects separates the mug from its background and in so doing resolves $B\_A_R - B\_A_L$. Hence, removing all texture from objects as in the stereograms in (b), separation of surface layers is still readily achieved. If my argument is valid, then the position of objects in space, and separation of the different depth planes (in the absence of texture) is recovered not only from local but is at least partly revealed to stereopsis from a process that resolves background or foreground subtense. This equates to simultaneous processing of several different disparity feature-scales.

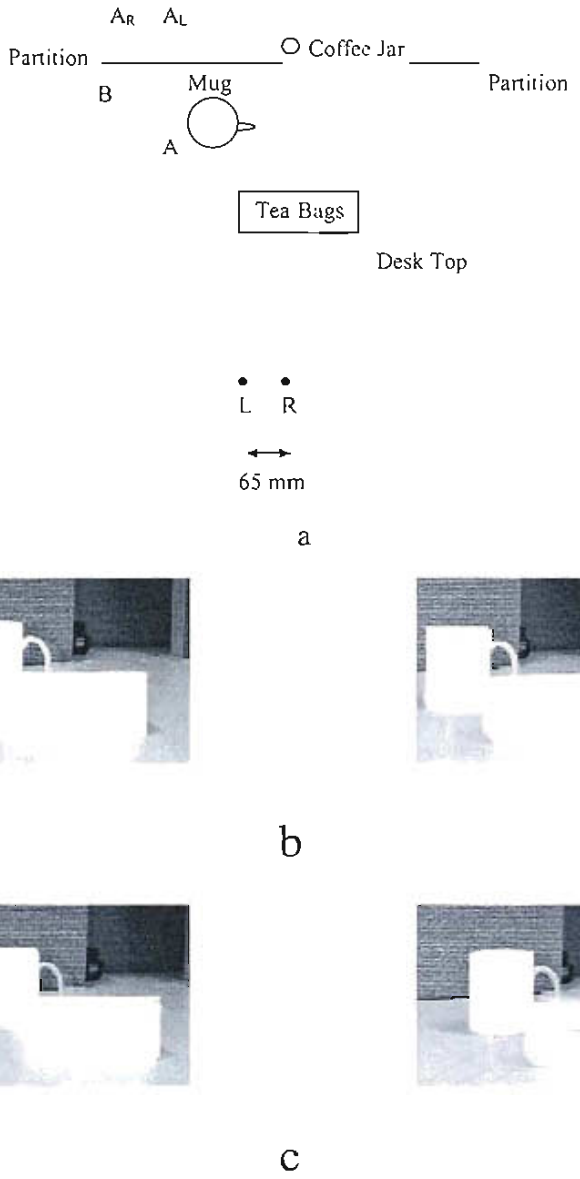



**Fig. 10.3. Disparate subtense and surface separation in the absence of texture**
Figure (a) reconstructs the geometry subtended at each eye by a small set of mundane stationary objects. In the 2-D layout of the retinal images, object boundaries and backgrounds are *the same* luminance contours. The system must achieve separation at these contours so that a near edge is separated from distant depth plane at those contours (split-projection). This interpretation is a resolution of disparate subtense. Crossed fusion at (b) positions all objects behind a square window, an analogue of the porthole effect. Point-disparities across the object surfaces have been removed. The disparities at the background can now be manipulated by shifting the photo behind white edge boundaries. In (c), the objects now stand at a nearer depth despite a lack of surface texture. Resolution of disparate subtense to yield depth therefore appears to be an important and repeatable stereoscopic response.

## 10.4 Possible implications of the structural organization of binocular vision for 3-D perception

One purely physiological constraint that may have implications for the way in which the images addressed are resolved, is a very basic aspect of all the perceptual systems. That is, the structure of the system – its topographic and retinotopic organisation (see for example Kandel, Schwartz and Jessel, 1991).

A pair of parallel arrays of luminance sensors at the retinae converging to a single array of binocular receptive fields at the striate cortex, may itself constrain resolution of local disparities. Figure 10.4 presents a highly stylised representation (eg. the optic chiasm is ignored) of the massively parallel organisation of retinal coordinate geometry. Resolution of a cyclopean 3-D percept from the combination of two retinal images may involve restrictions on perceived contour alignments and surface separation because each local region of each retinal image is physically constrained by its position in the visual *coordinate matrix*. Hence the spaces between contours and image features will be constrained in achieving the percept. My evidence is that disparity values at different scales can conflict however.

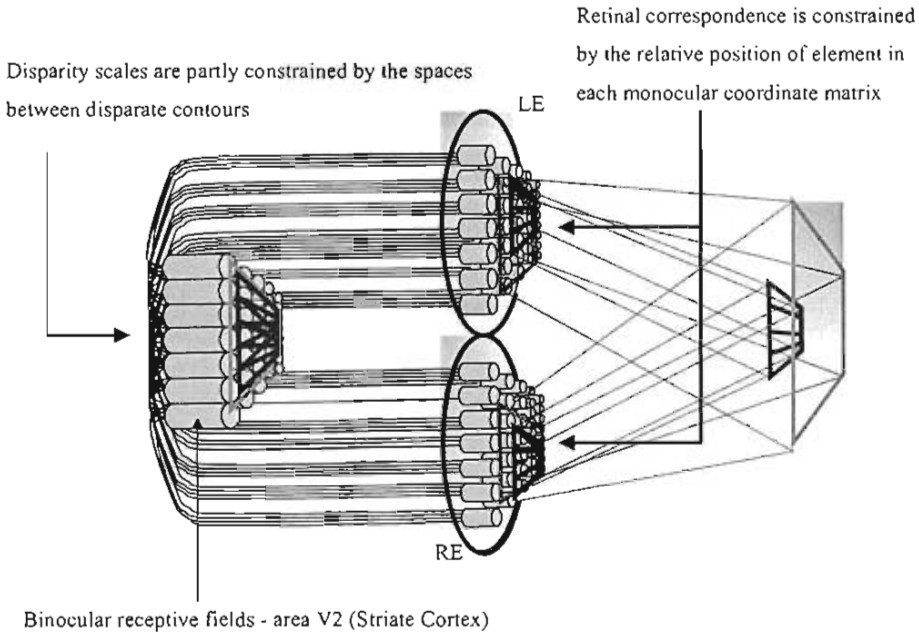



**Fig. 10.4 A system-structural representation of parallel retinal and cortical topography**
Authors such as Howard and Rogers (1995) explain that in binocular fusion vergence eye movements align the retinal coordinate matrices. Given the topographic organisation of the *system* this might mean that not only local positional differences but also disparate spaces between features constrain the eventual 3-D percept.

In conclusion, a fruitful avenue of future research in this area may be to examine how binocular subtense might constrain the correspondence problem in textured images. Historically, correspondence represents a major computational problem both for theoretical and practical applications of stereoscopic image processing. In David Marr's (1982) seminal work in stereopsis, for example, he employed a range of disparity filters to compute coarse-to-fine disparities and so constrain correspondence by progressive data reduction - a "levels of analysis" approach. A dramatic saving in computational effort might well be achieved simply a two-stage processing scheme: resolving disparate subtense then local correspondence.

In other work I have also begun to examine the patterns of binocular optic flow that surround occlusion, in the perception of spinning and looming illusory surfaces. The spatio-temporal patterns of non-local binocular correspondence appear to offer another interesting area in which to pursue the separation of perceived space into surfaces and shapes. A recent finding for example has suggested that stereopsis enhances the maximum threshold at which stroboscopic motion is detected for Kanizsa figures[26].

Another possible avenue of research may be the use of skeletal contours and partial occlusion patterns in the generation of stereoscopic 3-D displays. It may be possible to use the propensity for binocular vision to access sparse disparities to build effective low-fidelity virtual environments using very much reduced computational power compared to those based on densely textured point disparities. According to this research, the system can quite effectively yield predictable depth percepts in the complete absence of textures.

---

[26] Preliminary results were presented at the 23rd Australasian Experimental Psychology Conference (Huf and Ryan, 1997).

## References

Agostini, T., & Proffitt, D. R. (1993). Perceptual organisation evokes simultaneous lightness contrast, *Perception* , *22*, 263-272.

Albert, M. K. (1993). Parallelism and the perception of illusory contours, *Perception* , *22*, 589-595.

Albright, T. (1995). My most true mind thus makes mine eye untrue, *Trends in neurosciences* , *18*, 331-333.

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception* , *26,* 419-454.

Anderson, B. L. (1994). The role of partial occlusion in stereopsis, *Nature, 367,* 365-368.

Anderson, B. L. & Julsez, B. (1995). A theoretical analysis of illusory contour formation in stereopsis, *Psychological Review* , *102*, 705-743.

Anderson, B. L. & Nakayama, K., (1994). Toward a general theory of stereopsis: Binocular matching, occluding contours, and fusion, *Psychological Review, 101,* 414-445.

Becker, M. F., & Knopp, J. (1978). Processing of visual illusions in the frequency and spatial domains, *Perception & Psychophysics, 23,* 521-526.

Berkley, M. A., Debruyn, B., & Orban, G. (1994). Illusory, motion, and luminance-defined contours interact in the human visual system, *Vision Research, 34,* 209-216.

Bloomfield, S. (1973). Implicit feature and stereoscopy, *Nature,* 256

Bonaiuto, P. Giannini, A. M., & Bonaiuto, M., (1991). Visual productions with or without amodal completion, *Perception, 20*, 243-257.

Bradely, D. R. (1987). Cognitive contours and perceptual organisation, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 53-61. New York: Springer.

Bradley, D., & Lee, K. (1982). Animated subjective contours, *Perception and Psychophysics* , *32,* 393-395.

Brigner, W. L., & Gallagher, M. B. (1974). Subjective contour: Apparent depth, or simultaneous brightness contrast, *Perceptual and Motor Skills, 38,* 1047-1053.

Bruce, V., Green, P. R., & Georgeson, M. A. (1996). *Visual Perception*, Psychology Press, London: Taylor and Francis.

Bulthoff, H., & Mallot, H. (1988). Integration of depth modules: Stereo and shading *Journal of the optical society of America*, *5*,1749-1758

Bulthoff, H., Fahle, M., & Wegmann, M. (1991). Perceived depth scales with disparity gradient, *Perception, 20,* 145-153.

Burt. P., & Julesz, B. (1980). Modifications of the classical notion of Panum's fusional area. *Perception, 9,* 671-143.

Cagenello, R., & Rogers, B. J. (1988). Local orientation differences affect the perceived slant of stereoscopic surfaces, *Investigative ophthalmology and visual science, Supplement, 29,* 399.

Cagenello, R., and Rogers, B. J. (1993). Anisotropies in the perception of stereoscopic surfaces: The role of orientation disparities, *Vision Research, 33,* 2189-2201.

Carmen, G.J., and Welch, L. (1993). Three dimensional illusory contours and surfaces, *Nature, 360,* 585-587.

Chao-Yi, L., & Kun, G. (1995). Measurements of geometric illusions, illusory contours and stereo-depth at luminance and colour contrast, *Vision research, 35,* 1713-1720.

Cohen, M.A., & Grossberg, S. (1984). Neural dynamics of brightness perception: features, boundaries, diffusion and resonance, *Perception and Psychophysics, 36,* 428-456.

Coren, S. (1991). Retinal mechanisms in the perception of subjective contours: The contribution of lateral inhibition, *Perception , 20,* 181-191

Coren, S., & Porac, C. (1991). Subjective contours and apparent depth: A direct test, Perception and Psychophysics , *33,* 197-200.

Coren, S. (1972). Subjective contours and apparent depth, *Psychological Review, 79,* 359-367.

Coren, S., Porac, C., & Theodor, L.H. (1986). The effects of perceptual set on the shape and apparent depth of subjective contours, *Perception and Psychophysics , 39,* 327-333.

Davi, M. Pinna, B., & Sambin, M. (1992). Amodal completion versus induced inhomogeneities in the organisation of illusory figures, *Perception , 21,* 627-636.

Day, R.H. (1987). Cues for edge and the origin illusory contours: An alternative approach, in in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 83-90. New York: Springer.

Day, R.H., & Jory, M.K. (1980). A note on a second stage in the formation of illusory contours, *Perception and Psychophysics, 27,* 89-91.

Day, R.H., & Kasperczyk, R.T. (1983). Amodal Completion as a basis for illusory contours, *Perception and Psychophysics, 33,* 355-364.

Dresp, B. & Bonnet, C. (1991). Psychophysical evidence for low-level processing of illusory contours and surfaces in the Kanizsa square, *Vision Research , 31,* 1813-1817.

Dresp, B. Lorenceau, J. & Bonnet, C. (1990). Apparent brightness enhancement in the Kanizsa square with and without illusory contour formation, *Perception , 19*, 483-489.

Dresp, B., & Bonnet, C. (1995), Subthreshold summation with illusory contours, *Vision Research , 35*, 1071-1078.

Dresp, B., Salvano-Pardieu, V., & Bonnet, C. (1996). Illusory form with inducers of opposite contrast polarity, *Perception and Psychophysics , 58*, 111-124.

Durgin, F., Tripathy, S. P., & Levi, D. M. (1995). On the filling in of the visual blind spot: some rules of thumb, *Perception , 24*, 827-840.

Francis, G., Grossberg, G., & Mingolla, E. (1994). Cortical dynamics of feature binding and reset: Control of visual persistence, *Vision Research , 34*, 1089-1104.

Francis, G., & Grossberg, G. (1996). Cortical dynamics of form and motion integration: Persistence, apparent motion, and illusory contours, *Vision Research , 36*, 149-173.

Frisby, J.P., & Clatworthy, J.L. (1975). Illusory contours: Curious case of simultaneous brightness contrast? *Perception , 4*, 349-357.

Fujita, N. (1993). An occluded contour becomes visible with reversal of disparity, *Perceptual and Motor Skills , 77*, 271-274.

Garding, J., Porrill, J., Mayhew, J. E. W., & Frisby, J. P. (1995). Stereopsis, vertical disparity and relief transformations, *Vision Research, 35*, 703-722.

Gegenfurtner, K. R., Brown, J.E., & Rieger, J., (1996). *Technical Report. No. 28: Integration processes in the perception of real and illusory contours.* Max Plank Institut fur Biologische Kybernetik: Germany.

Gillam, B. (1968). Perception when stereopsis and perspective conflict: Experiments with aniseikonic lenses, *Journal of experimental psychology: human perception and performance, 78*, 299-305.

Gillam, B., & Borsting, E., (1988). The role of monocular regions in stereoscopic displays, *Perception, 17*, 603-608.

Gillam, B., & Lawergren, B. (1983). The induced effect, vertical disparity, and stereoscopic theory, *Perception and Psychophysics, 34*, 121-130.

Gillam, B., & Rogers, B.J. (1989). Orientation disparity, deformation, and stereoscopic slant perception, *Perception, 20*, 441-448.

Gillam, B., & Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in the perception of stereoscopic slant, *Perception, 21*, 427-439.

Gillam, B. (1987). Perceptual grouping and subjective contours, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 35-46. New York: Springer.

Gillam, B. (1995). Quantitative depth without binocular correspondence not yet demonstrated, *Nature, 373*, 202-203.

Gillam, B. (1995). Paper presented at the 22th Annual Australian Experimental Psychology Conference, Brisbane, April, 1995.

Gillam, B., Blackburn, S., & Cook, M. (1995). Panum's limiting case: double fusion, convergence error or da Vinci stereopsis, *Perception, 24*, 333-346.

Gillam, B., Blackburn, S., & Nakayama, K. (1999). Stereopsis based on monocular gaps: Metrical encoding of depth and slant without matching contours, *Vision Research, 39*, 493-502.

Gillam, B., Chambers, D., & Russo, T. (1988). Postfusional latency in stereoscopic slant perception and the primitives of stereopsis, *Journal of experimental psychology: human perception and performance, 14*, 163-175.

Gillam, B., Flag, T., & Finlay, D. (1984). Evidence for disparity change as the primary stimulus for stereoscopic processing, *Perception and Psychophysics, 36*, 559-564.

Ginsburg, A. P. (1975). Is the illusory triangle physical or imaginary? Nature , 257, 219-220.

Ginsburg, A. P. (1987). The relationship between spatial filtering and subjective contours, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 61-68. New York: Springer.

Goldstein, E. B. (1996). *Sensation and Perception*, Pacific Grove: Brooks/Cole.

Gove, G., Grossberg, S., & Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback, *Visual Neuroscience , 12*, 1027-1052.

Greggory, R. L. (1972). Cognitive contours, *Nature* , 238, 51-52.

Gregory, R. L. (1987). Illusory contours and occlusion surfaces, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 27-34. New York: Springer.

Gregory, R. L. (1977). Vision with isoluminant colour contrast: 1. A projection technique and observation, *Perception, 6*, 113-119.

Gregory, R. L. & Harris, J.P., (1974). Illusory contours and stereo depth, *Perception and Psychophysics, 15*, 411-416.

Gregory, R. L. (1973). The confounded eye, in Gregory, R.L., and Gombrich, (Eds). *Illusions in nature and art*, Gerald Duckworth: London.

Gregory, R. L. (1974). *Concepts and Mechanisms in Perception,* Shcribner: New York

Gregory, R. L. (1986). Illusory surfaces as perceptual postulates, *Perception and Psychophysics, 39,* 220.

Gregory, R. L. (1998). *Eye and Brain: The psychology of seeing*, Oxford: Oxford University Press.

Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons can signal illusory contours. *Nature, 365,* 550-552.

Grossberg, G. (1994). 3-D vision and figure-ground separation by the visual cortex, *Perception and Psychophysics , 55,* 48-120.

Grossberg, S., & Mingolla, E. (1987). The role of illusory contours visual segmentation, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 53-61. New York: Springer.

Grossberg, S., & Todorovic, D. (1988). neural Dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena, *Perception and Psychophysics 43,* 723-742.

Grossberg, S. (1987). Cortical dynamics of three-dimensional form, colour, and brightness perception: I Monocular theory, *Perception and Psychophysics , 41,* 87-116.

Gulick, W. L., & Lawson, R. B. (1976). *Human stereopsis: A psychophysical analysis*, New York: University Vermont.

Guzman, A. (1968). Decomposition of a visual scene into three dimensional bodies, in A. Grasseli, (Ed.), *Automatic interpretation and classification of images*, New York: Academic Press.

Halpern, D. F., (1987). The functional equivalence of objective and illusory brightness enhancement, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 53-61. New York: Springer.

Howard, I. P., & Ohmi, M. (1992). A new interpretation of the role of dichoptic occlusion in stereopsis, *Investigative ophthalmology and visual science. supplement, 33,* 1370.

Howard, I.P., & Rogers, B.J. (1995). *Binocular vision and stereopsis*, Oxford: Oxford University Press.

Hubel, D. (1980). A big step along the visual pathway, *Nature, 380,* 197-198.

Jory, M., & Day, R. H., (1979). The relationship between brightness contrast and illusory contours, *Perception, 8,* 3-9.

Julesz, B. (1964). Binocular depth perception without familiarity cues, *Science, 145,* 356-362.

Julesz, B. (1960). Binocular depth perception of computer generated patterns, *Bell Systems Technical Journal, 39,* 1125-1162.

Julesz, B. (1971). Foundations of cyclopean perception, Chicago: Chicago press.

Julesz, B., (1986) Stereoscopic vision, *Vision Research, 26,* 2, 1601-1612.

Kandel, E.R., Schwartz, J.H., & Jessel, T. M. (1991). *Principles of neural science*, London: Prentice Hall.

Kanizsa, G. (1976). Subjective contours, *Scientific American* , 234, 48-52.

Kanizsa, G. (1979). *Organisation in vision: Essays on gestalt perception* , Praeger: New York.

Kanizsa, G. (1955). Marzini quasi-percetti in campi con stimolozione omogenea. *Revista di psicologia, 49,* 7-30.

Kaufman, L. (1974). *Sight and mind*, New York: Liverton Publishing.

Kellman, P. J. & Loukides, M. G. (1987). in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 53-61. New York: Springer.

Kellman, P. J., and Cohen, M. H., (1984). Kinetic subjective contours, *Perception and Psychophysics, 35,* 237-244.

Kellman, P. J., and Shipley, T. F., (1991). A theory of Visual Interpolation in Object Perception, *Cognitive Psychology, 23,* 141-221.

Kellman, P. J., and Shipley, T. F. (1992). Perceiving objects across space and time, *Current Directions in Psychological Science, 1*, 193-199.

Kennedy, J. M. (1988). Line endings and subjective contours, *Spatial Vision, 3*, 151-158.

Kennedy, J. M. (1975). Depth at an edge, co planarity, slant depth, change in direction and change in brightness in the production of subjective contours, *Italian journal of psychology,* 2, 107-123.

Kennedy, J. M. (1976). Sun figure: An illusory diffuse contour resulting from an arrangement of dots, *Perception, 5*, 479-481.

Kennedy, J. M. (1981). Illusory brightness and the ends of petals: Change in brightness without aid of stratification or assimilation effects, *Perception, 10*, 583-585.

Kennedy, J. M. (1987). in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 53-61. New York: Springer.

Kojo, I., Liinasuo, M., & Rovamo, J. (1993). Spatial and temporal properties of illusory figures, *Vision Research, 7.* 897-901.

Laurie, C. Warm, J., Dember, W., & Frank, R. (1994). Determinants of stability in the perception of subjective contours, *Perception and Psychophysics* . *55*, 394-398.

Lawson, R.B., & Gulick, W.L. (1967). Stereopsis and anomalous contour, *Vision research*. 7, 271-297.

Lesher, G. W. (1995). Illusory contours: Toward a neurally based perceptual theory, *Psychonomic Bulletin and Review* , *2* , 279-321.

Lesher, G. W., & Mingolla, E. (1993). The role of edges and line ends in illusory contour formation, *Vision Research* , *33*, 2253-2270.

Li, C., & Kun, G. (1995). Measurement of geometric Illusions, illusory contours and stereo-depth at luminance and colour contrast. *Vision Research 35*, 1713 - 1720.

Liu, L., Stenvenson, S. B., & Schore, C. M. (1994). Quantitative stereoscopic depth without binocular correspondence, *Nature, 367*, 66-69.

Livingston, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, colour, movement and depth, *Journal of Neuroscience. 7*, 3416 - 3468.

Mallot, H, A., Arndt, P. A., & Bulthoff, H. H. (1995). *The mechanism of intensity based stereo*, Technical Report No. 18, Max Plank Institut fur Biologishe Kybernetik.

Marr, D., & Pogio, T. (1979). A computational theory of human stereo vision, *Proceedings of the Royal Society of London*, 204, 301-328.

Marr, D. (1982). *Vision, A computational investigation into the human representation and processing of visual information,* San Francisco: W.H. Freeman.

Masini, R., Ferraro, M., & Costa, T. (1994). Stereokinetic effects with sharp and fuzzy illusory contours, *Perception* . *23*, 1437-1445.

Mayhew, J. E. W., & Frisby, J. P. (1980). The computation of binocular edges, *Perception, 9*, 69-86.

Mayhew, J. E. W., & Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception, *Nature, 297*, 376.

McKee, S. P., Bravo, M. J., Smallman, H. S., & Legge, G. E., (1995). The uniqueness constraint, and binocular masking, *Perception, 24*, 49-65.

Mcourt, M. E., & Paulsen, K. (1994). The influence of illusory contours on the perception of luminance increments and decrements, *Visual Research* , *34*, 2469-2475.

Merteli, F. (1974). The perception of transparency, *Scientific American, 230,* 90-98.

Meyer, G. E., & Ming, C. Y. (1988). The visible persistence of illusory contours, *Canadian Journal of Psychology , 42*, 479-488.

Meyer, G.E., & Petry, S. (1987). Top-down and bottom up: The illusory contour as a microcosm of issues in perception, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 1-8. New York: Springer.

Minguzzini, G. F. (1987). Anomalous figures and the tendency toward continuation, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 53-61. New York: Springer.

Mitcheson, G. J., & McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception, *Vision research. 30*, 1781-1791.

Mitcheson, G. J., & Westheimer, G. (1984). The perception of depth in simple figures, *Vision Research. 24, 9*, 1063-1073.

Nakayama, K., and Silverman, G. H. (1988). The aperture problem: I. Perception of nonrigidity and motion direction in translating sinusoidal lines, *Vision Research , 28*, 739-1990.

Nakayama, K. (1996). Binocular surface perception, *Proceedings of the National Academy of Science. USA, 93*, 634-639.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects, *Perception , 18*, 55-68.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects, *Perception , 18*, 55-68.

Nakayama, K., & Shimojo, S. (1990). Da Vinci stereopsis: depth and subjective occluding contours from unpaired image points, *Vision Research, 30,* 1811-1825.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces, *Science , 257*, 1357-1363.

Nakayama, K., Shimojo, S., & Ramachandran, V.S. (1989). Transparency: Relation to depth, subjective contours, luminance, and neon color spreading, *Perception , 19*, 497-513.

Neutrali, A. N., & Haskell, B.G., (1988). *Digital pictures – representation and compression*, Plenum Press.

Ninio, J., (1985). Orientational vs horizontal disparity in the stereoscopic appreciation of slant, *Perception. 14*, 305-314.

Ogle, K. N. (1950). *Researches in binocular vision,* New York: Hafner.

Ogle, K. N. (1959). Theory of stereoscopic vision, in Koch, S., (Ed) *Psychology: A study of a science.* New York: McGraw-Hill.

Ono, H., & Mapp, P. (1995). A restatement and modification of Wells-Hering's laws of visual direction, *Perception, 33,* 2685-2696.

Ono, H., Shimono, K., & Shibuta, K. (1992). Occlusion as a depth cue in Panum's limiting case, *Perception and Psychophysics, 51,* 3-13.

Parks, T. E. (1994). On the microgenesis of illusory figures: A failure to replicate, *Perception , 23,* 857-862.

Parks, T. E., & Pendergrass, L. (1982). On the filtered-components approach to illusory visual contours, *Perception and Psychophysics, 32,* 491-493.

Parks, T. E. (1983). Further difficulties for the filtered-components approach to illusory visual contours, *Perception and Psychophysics , 34,* 190-192.

Parks, T. E. (1986). Illusory figures, illusory objects, and real objects, *Psychological Review , 93,* 207-215.

Parks, T. E. (1987). Illusory figures and pictorial objects, in Petry, S., and Meyer, G.E, (Eds), *The Perception of Illusory Contours ,* New York: Springer.

Peterhans, E., and von der Heydt, R. (1991). Bridging the gap between psychophysics and physiology, *Trends in Neurosciences , 14,* 112-119.

Peterhans, E., von der Heydt, R., & Baumgartner, G. (1986). Neuronal responses to illusory stimuli reveal stages of visual cortical processing, in J. D. Pettigrew, K. J. Sanderson, & W. R. Levick, (Eds.), *Visual Neuroscience ,* pp. 42-65. Cambridge: Cambridge University Press.

Peterson, M. A., & Gibson, B. S. (1994). Object recognition contributions to figure-ground organisation: Operations on outlines and subjective contours, *Perception and Psychophysics , 56,* 551-564.

Petry, S. & Meyer, G. (1987). (Eds.). *The perception of illusory contours,* Springer-Verlag.

Pizzlo, Z., Rosenfield, A., and Weiss, I. (1997). The geometry of visual space: About the incompatibility between science and mathematics, *Computer Vision and Image Understanding, 65,* 613-634.

Pogio, G.F., & Pogio, T. (1984). The analysis of stereopsis, *Annual review of neuroscience, 7,* 379-412.

Pradiso, M. A., Shimojo, S., & Nakayama, K. (1989). Subjective contours, tilt after-effects, and visual cortical organisation, *Vision Research , 29,* 1205-12

Prazdny, K. (1985). On the nature of inducing forms generating perceptions of illusory contours, *Perception and Psychophysics, 37,* 237-242.

Prazdny, K. (1986). Illusory contours from inducers defined solely by spatiotemporal correlation, *Perception and Psychophysics , 39,* 179-186.

Prazdny, K. (1986). 'Capture' of stereopsis by illusory contours, *Nature, 324,* 393-394.

Purghe, F. & Coren, S. (1992). Amodal completion, depth stratification, and illusory figures: A test of Kanizsa's explanation, *Perception . 21,* 325-335.

Purghe, F. and Coren, S. (1992a). Subjective contours 1900-1990: Research trends and bibliography, *Perception , 51,* 291-304.

Purghe, F. (1993). Illusory contours from pictorially three dimensional inducing elements: counter evidence for Parks and Rock's example, *Perception . 22,* 809-818.

Purghe, F., & Katsaras, P. (1991). Figural conditions affecting the formation of anomalous surfaces: overall configuration versus single stimulus part, *Perception , 20,* 193-206.

Purghe, F. (1991). Is amodal completion necessary for the formation of illusory figures?, *Perception , 20,* 623-636.

Ramachandran, V., Ruskin, D., Cobb., & Rogers-Ramachandran, D. (1994). On the perception of illusory contours, *Vision Research , 34,* 3145-3152.

Ramachandran, V. S., & Cavanagh, P. (1985). Subjective contours capture stereopsis, *Nature, 317,* 527-530.

Ramachandran, V. S. (1985). Apparent motion of subjective surfaces, *Perception , 14,* 127-134.

Ramachandran, V. S. (1986). Capture of stereopsis and apparent motion by illusory contours, *Perception and Psychophysics , 39,* 361-373.

Ramachandran, V. S. (1987). Visual Perception of surfaces: A biological theory in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours,* pp. 105-115. New York: Springer.

Redies, C., & Spillman, L. (1981). The neon colour effect in the Ehrenstein illusion, *Perception, 10,* 667-6681.

Reynolds, R. (1981). Perception of an illusory contour as a function of processing time, *Perception, 10,* 107-115.

Ritter, M. (1977). Effects of disparity and viewing distance on perceived depth. *Perception and Psychophysics, 5*, 317-320.

Ritter, M. (1979). Perception of depth. Processing of simple position disparity as a function of viewing distance. *Perception and Psychophysics, 25*, 360-380.

Rock, I. (1987). A problem solving approach to illusory contours, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 90-96. New York: Springer.

Rock, I., & Anson, R. (1979). Illusory contours as the solutions to a problem, *Perception . 8,* 665-681.

Rogers, B., & Bradshaw, M. (1993). Vertical disparities, differential perspective and binocular stereopsis, *Nature, 361*, 253-255.

Rogers, B., & Bradshaw, M. (1995). Disparity scaling and the perception of frontoparallel surfaces, *Perception, 24*, 155-179.

Rogers, B., & Cagenello, R. (1989). Disparity curvatures and the perception of three dimensional surfaces, *Nature, 339*, 135-137.

Rogers, B., & Graham, M.E. (1983). Anisotropies in the perception of three-dimensional surfaces, *Science, 221*, 1409-1411.

Rogers. B. J. (1995). Vision is getting easier every day, *Perception, 25*, 1227-1232.

Ryan, C., & Gillam, B. (1994). Cue conflict and stereoscopic surface slant about horizontal and vertical axes, *Perception, 23*, 645-658.

Sambin, M. (1987). A dynamic model of anomalous figures, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 53-61. New York: Springer.

Saye, A., & Frisby, J. P. (1975). The role of monocularly conspicuous features in facilitating stereopsis from random-dot stereograms, *Perception, 4*, 159-171.

Sekular, A. B., Palmer, S. E., & Flynn, C. (1994). Local and Global processes in visual completion, *Psychological Science , 5*, 260-267.

Shapely, R., & Gordon, J. (1987). The existence of interpolated illusory contours depends on contrast and spatial separation, in S. Petry, & G. E. Meyer, (Eds.), *The Perception of Illusory Contours*, pp. 53-61. New York: Springer.

Shipley, T. (1975). Visual contours in homogeneous space, *Science,* 150, 348-350.

Shipley, T. F., & Kellman, P. J. (1990). The role of discontinuities in the perception of subjective figures, *Perception and Psychophysics, 48*, 159-270.

Shipley, T. F., & Kellman, P. J. (1991). Perception of partly occluded objects and illusory figures: Evidence for an identity hypothesis, Journal of experimental psychology: *Human Perception and Performance, 18*, 106-120.

Shipley, T.F., & Kellman, P.J. (1992). Strength of visual interpolation depends on the ratio of physically specified to total edge length, *Perception and Psychophysics, 52*, 97-106.

Siegal, S., & Petry, S. (1991). Evidence for independent processing of subjective contour brightness and sharpness, *Perception , 20*, 233-241.

Simmonds, M. B. (1975). Stereopsis and subjective contours, *Perception and Psychophysics , 15*, 401-404.

Smith, A., & Over, R. (1975). Tilt after effects with subjective contours, *Nature , 257*, 581-582.

Smith, A., & Over, R. (1976). Colour selective tilt after effects with subjective contours, *Perception and Psychophysics , 20*, 305-308.

Spillman, L., & Dresp, B. (1995). Phenomena of illusory form: Can we bridge the gap between levels of explanation? *Perception, 24*, 1333-1364.

Spillman, L., (1977). Contrast and brightness illusions, in H. Spekreijse, L. H. van der Tweel, (Eds.), *Spatial contrast*, pp. 38-46, Amsterdam: North-Holland.

Stevens, K. A., & Brooks, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces, *Vision Research*, 28, 371-386.

Stevens, K.A., (1983). Evidence relating subjective contours and interpretations involving interposition, *Perception , 12*, 491-500.

Stevens, K.A. (1991). The perception of three-dimensionality across continuous surfaces, in J. R. Ellis, M. K. Kaiser, & A.C. Grunwald, *Pictorial communication in virtual and real environments*, pp. 88-105, London: Taylor and Francis.

Takeichi, H. Shimojo, S & Watanabe, T. (1992). Neon flank and illusory contour: Interaction between the two processes leads to colour filling in, *Perception , 21*, 313-324.

Takeichi, H., Watanabe, T., & Shimojo, S. (1992). Illusory occluding contours and surface formation by depth propagation, *Perception , 21*, 313-324.

Tsal, Y., Meiran, N., & Lavie, N. (1994). The role of attention in illusory conjunctions, *Perception and Psychophysics , 55*, 350-358.

Tyler C. W. (1977). Is the illusory triangle physical or imaginary, *Perception*, 603-604.

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity, *Vision research, 15*, 583-590.

Ullman, S. (1976). Filling in the gaps: The shape of subjective contours and a model for their generation, *Biological cybernetics, 25,* 1-6.

Vallortigara, G., & Bressan, P. (1994). Occlusion, transparency and stereopsis: a new explanation for stereo capture, *Vision Research , 34,* 2891-2896.

van der Zwan, R., & Wenderoth, P. (1994). Psychophysical evidence for area V2 involvement in the reduction of subjective contour tilt after effects by binocular rivalry, *Visual Neuroscience, 11,* 823-830.

van der Zwan, R., & Wenderoth, P. (1995). Mechanisms of purely subjective contour tilt after effects, *Vision Research, 35,* 2547-2557.

Van Ee, R., & Erkelens, C. J. (1995). Binocular perception of slant about oblique axes relative to a visual frame of reference, *Perception. 24,* 299-314

Van Tuijl. R. (1975). A new visual illusion: Neonlike colour spreading and complementary color induction between subjective contours, *Acta Psychologica 39,* 441-445.

von der heydt, R., & Peterhans, E. (1989). Ehrenstein and Zollner illusions in a neural theory of contour processing, in J. J. Kulikowski, C. M. Dickinson, & I. J. Murray, (Eds.). *Seeing contour and colour*, pp. 27-39, Oxford: Pergamon Press.

Wallach, H., & Zuckerman, C. (1963). The constancy of stereoscopic depth. *American journal of psychology, 76.* 404-412.

Watanabe, T., & Cavanagh, P. (1992). Depth capture and transparency of regions bounded by illusory and chromatic contours, *Vision Research , 32,* 527-532.

Watanabe, T., & Cavanagh, P. (1993). Transparent surfaces defined by implicit X junctions, *Vision Research , 33,* 239-2346.

Watanabe, T., Nanez, J., & Moreno, M., (1995). Depth release of illusory contour shape in the Ehrenstein grid, *Vision Research , 35,* 2845-2851.

Watanabe, T., & Sato, T. (1989). Effects of luminance contrast on colour spreading and illusory contour in the neon colour spreading effect, *Perception and Psychophysics , 65,* 427-430.

Watanabe, T., & Takeichi, H. (1990). The relation between colour spreading and illusory contours, *Perception and Psychophysics , 47,* 457-467.

Watanabe, T. (1995). Orientation and colour processing for partially occluded objects, *Vision Research , 35,* 657-655.

Yuille, A. L., & Bulthoff, H., (1993). *Bayesian decision theory and psychophysics*, Technical

Report No. 2, Max Plank Institut fur Biologishe Kybernetik.

# Appendix A A simple binocular perspective projection model

**Perspective correction: Rotation about the vertical axis**

Two diagrams are presented below to show perspective transformation of the angles of subtense of a square as it rotates about the vertical axis. Conceptually, there are two components of this process, a transformation in azimuth ($\alpha$) and a transformation in elevation ($\beta$) relative to the axis of rotation (zero disparity).

*Horizontal Compression*

$\alpha$ transformation is addressed first, as described in Fig. A1, where y is the distance to the picture plane; $\theta$ is the angle of rotation; $\Delta y$ is the distance of the distant or near edge of the rotated rectangle from the picture plane; $w_1$ is half the horizontal magnitude of the fronto-parallel square; and $w_2$ is the perspective corrected (projected) magnitude of $w_1$ as it passes through the arc of rotation. So, $\alpha_F$ is the angle subtended by $w_2$ at distance $y + \Delta y$, and $\alpha_N$ is the angle subtended by $w_2$ at distance $y - \Delta y$.

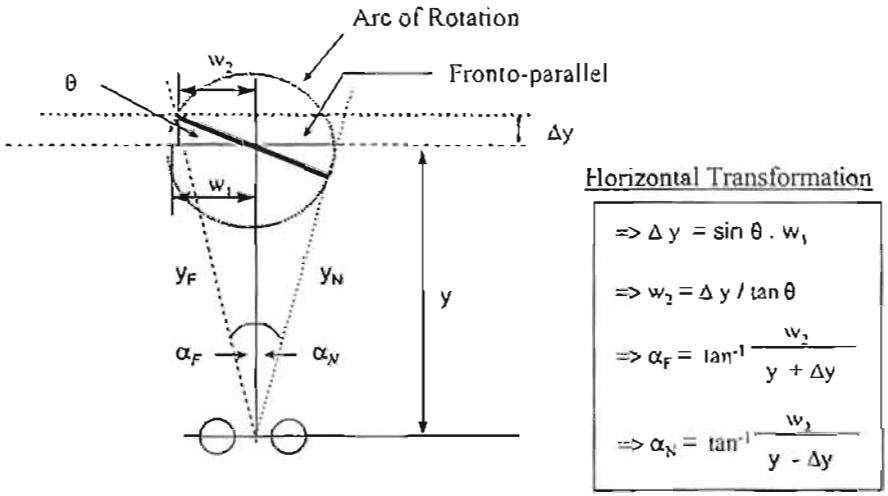



**Fig A1 $\alpha$ Transformation: Vertical axis of rotation**

*Vertical Shear*

A β transformation can be calculated by determining the distance of the far edge ($y_F$) according to $\alpha_F$ described above. The angle of elevation subtending the distant edge ($\beta_F$) is therefore the angle of subtense of the far edge at distance $y_F$, and the angle of elevation subtending the near edge ($\beta_N$) is the angle of subtense of the near edge at distance $y_N$. This is shown in Fig A2.

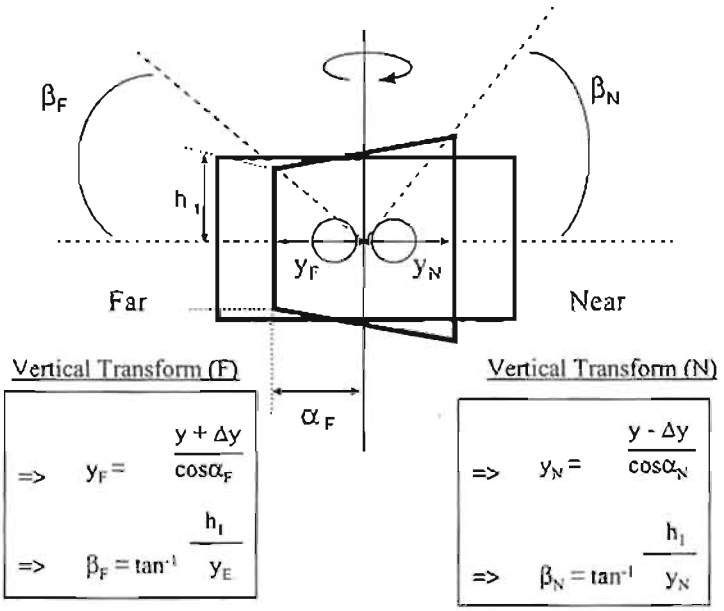



**Fig A2 β transformation: Vertical axis of rotation**

In summary, these transformations generate compression of horizontal dimensions and vertical shear in the square relative to the angle of its rotation. A symmetrical stereoscopic rotation of this shape was achieved by applying ½ the required magnification factor in equal but opposite measure to each half image. In other words, as one image was compressed, the other was expanded —another example of reciprocal disparity patterns at occlusion.

**Perspective correction: Rotation about the horizontal axis**

A similar transformation was required to generate stereoscopic rotation about the horizontal axis. The major components of the transformation are compression of the vertical dimensions and shear in the horizontal dimensions of the figure according to the angle of rotation.

*Vertical Compression*

A β transformation in Fig A4 describes perspective rotation about the horizontal axis, where $h_1$ is half the vertical magnitude of the fronto-parallel square; and $h_2$ is the perspective corrected magnitude of $h_1$ as it passes through an arc of rotation about the horizontal axis. So $\beta_F$ is the angle subtended by $h_2$ at distance $y + \Delta y$, and $\beta_N$ is the angle subtended by $h_2$ at distance $y - \Delta y$.

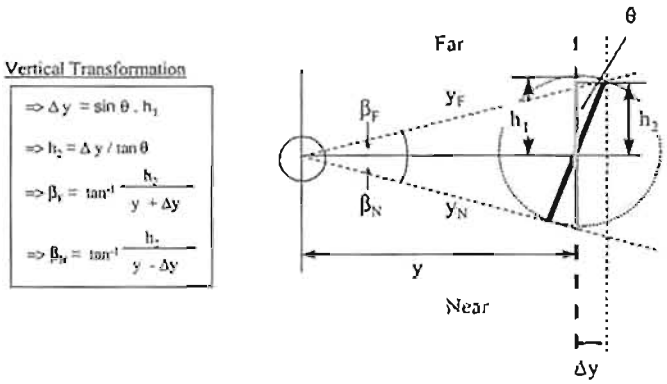



**Fig A3 β Transformation: Horizontal axis of rotation**

*Horizontal Shear*

Finally, an α transformation yielding horizontal shear can be calculated by determining the distance of the near edge ($y_N$) according to $\beta_N$ as described above. As Fig A4 demonstrates: the angle of azimuth subtending the distant edge ($\alpha_F$) is the angle of subtense of the far edge at distance ($y_F$); and the angle of azimuth subtending the near edge ($\alpha_N$) is the angle of subtense of the near edge ($w_1$) at distance ($y_N$).

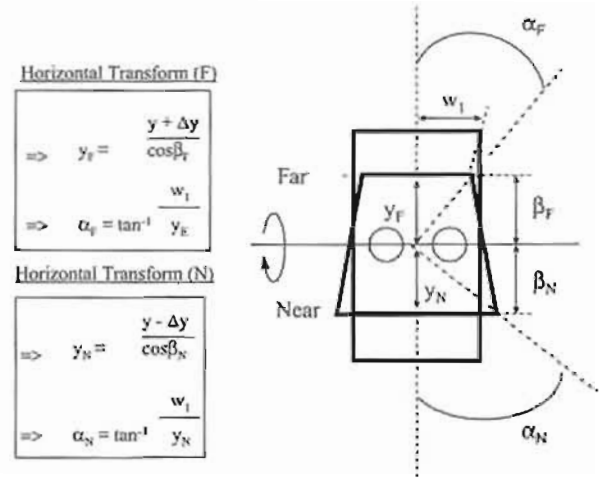



Fig A4 $\alpha$ Transformation: Horizontal axis of rotation

# Appendix B Sample operating software

## Run-Time Stimulus Presentation (Borland Objects Oriented Pascal 7.0)

```
program EXPERIMENTATION (infile, outfile);
{This program designed
to run a series of experiments based arround manipulations of a Kanizsa
Square}
uses
 Crt, Graph, dispobj, bgidrv, vsgdrv, gfxobj, exp_stim, exp_imag, tools;
const
 fillstyle = solidfill;
 NumImages = 32;                { 32 images in an experimental set }
 NumSets = 1;                   { 1 Set in an experiment }
type
 SetType = array[1..NumImages] of PTImage;   { the type of image displayed }
 OrderType = Array[1..NumImages] of integer; { Order array used in RandomiseSet }

var
 display : tvsgdisplay;          { the display object }
 image : SetType;                { image array  }
 i : integer;
 marker: tmarker;
 comparator : tcomparator;       { angular comparator }
 sets : integer;                 { experimental set index }
 txtbuff : string;               { buffer for constructing output text }
 numbuff : string;               { buffer for converting nums to text }
 outfile : text;
 datafile: string;
 { draws the comparator line on the display and waits until the user
 has lined the comparator up with the image.  returns when return
 key is pressed.  When this procedure returns the angle of the
 comparator can be accessed by using the comparator.deviation
 function. }
procedure trial(var comparator : tcomparator; const display : tdisplay);
var

 keypress: char; { contents of last key pressed }
 extended, exit : boolean; { flags }
 angle : angletype;
begin
 exit := false; { not ready to exit }

 repeat
  keypress := ' ';
  if(KeyPressed) then
  begin

     { First we need to decide if this is a normal character
```

```
      or an extended scan code. If the readkey call returns
      a value other than 0 then we have read a normal character
      otherwise we have received a scan code which will be
      aquired on the next call to readkey }

    keypress := readkey;
    if(keypress <> chr(0)) then extended := false
    else
    begin
      extended := true;    { set extended flag }
      keypress := readkey;  { and read actual key value }
    end;

    { Now extended will be true if an extended scan key was received
      and the actual scan code will be in keypress.  If extended
      is false then keypress contains an ascii char }

    if(extended = true) then
    begin
        if(keypress = 'M') then { -> }
        begin
            comparator.PointNeedle(display, comparator.deviation + 1);
        end
        else if(keypress = 'K') then { <- }
        begin
            comparator.PointNeedle(display, comparator.deviation - 1);
        end
        else if (keypress = 'H') then { ^ }
        begin
        end
        else if (keypress = 'P') then { v }
        begin
        end
    end  { extended scan code }
    else { we have a normal ascii character }
    begin
      if (keypress = chr(13)) then exit := true { exit true on return key pressed }
      else if(keypress = 'L') then display.SyncLensNow(cc_LeftOnRightOff)
      else if(keypress = 'R') then display.SyncLensNow(cc_LeftOffRightOn);
      end;
  end;
  until exit = true; { return key pressed }
end;
{ Used by RandomiseSet - Checks to see if x already exists in the ordertype
  array. If it does then x is not a valid integer. }
function valid(x : integer; ptr : integer; const order : OrderType) : boolean;
var i : integer;
begin
  valid := true;
  for i := 1 to ptr do
```

```
  begin
     if(x = order[i]) then valid := false;
  end;
end;
{ This procedure randomises the order in which the stimuli will be presented to the
  subject.  It does this by creating a randomly ordered set of the numbers 1..N where
  N is the number of stimuli being presented.  The randomly ordered set is created by
  drawing a random number in the range 1..N (without replacement) so the domain of the
  random set decreases as each number is drawn.

  When the randomly ordered set has been created it is used as an index to move the
  pointer to the image objects from the image set into a temporary set.  It is this
  process which randomises the order of presentation.  When this process is complete
  the pointers are copied back into the image set in their new order. }

procedure RandomiseSet;
var
   order : OrderType;
   temp : Settype;
   ptr : integer;

var i : integer;
begin
    ptr := 1;
    repeat
        i := random(NumImages) + 1;    { number in range 1..NumImages }
        if(valid(i, ptr, order)) then
        begin
            order[ptr] := i;        { save index number in order array }
            ptr := ptr + 1;         { increment index }
        end;
    until ptr = NumImages + 1;

    { now shuffle the pointers to the images in the Image pointer array
      using the order array as the shuffling key }
    for i := 1 to NumImages do temp[i] := Image[order[i]];
    for i := 1 to NumImages do Image[i] := temp[i];
end;
{ writes one set of trial data to the output file }
Procedure DumpData(s : integer; angle : AngleType; image : PTImage);
begin
    write(outfile, i, ' ');                    { experimental set number }
    If(Image^.LeftStimulus = Type_HRotKSq) then write (outfile,'H ');
    if(Image^.LeftStimulus = Type_VRotKSq) then write (outfile,'V ');

    write(outfile, image^.Identify, ' ');    { left stimulus object type }
    write(outfile, image^.LeftParm  :3, ' ');          { left stim parameter }
    write(outfile, image^.RightParm :3, ' ');          { right stim parameter }
    writeln(outfile, angle:3:0, ' ');                  { User defined comparator }
    If(Image^.LeftStimulus = Type_HRotKSq) then write ('H ');
```

```
    if(Image^.LeftStimulus = Type_VRotKSq) then write ('V ');
    write(' ',image^.Identify, ' ');    { left stimulus object type }
    write( image^.LeftParm :3, ' ');          { left stim parameter }
    write( image^.RightParm :3, ' ');         { right stim parameter }
    writeln( angle:3:0, ' ');                 { User defined comparator }

end;

{// End of support functions ////////////////////////////////////////////}

{ this is the start of the experiment proper. }

begin
    writeln;
    writeln;
    writeln;
    writeln('         ILLUSORY CONTOURS IN 3D');
    writeln('         exp1');
    WRITE('Subject Filename? : ');
    readln(datafile);
    display.init; { initialise the display }
    marker.init(cc_fground,cc_solid);

    { This is the simple compass comparator }
    comparator.init(cc_fground, cc_outline, 50);
    comparator.moveto(display.MidX, display.Y - 50);
    comparator.calibrate(90);

    { These are our stimulus images - they will be randomised
      by a call to RandomiseSet before each set of trials so
      they will be presented to the user in a random fashion.
      when it is constructed.  When we have finished with these
      images we need to destruct them to release the image object
      AND any resources it may have aquired.  Note the call to the
      image destructors in the final few lines of the program }

  {KSq Horizontal Rotation}
  image[1] := new(PTImage, init(1,
                new(PTHRotKSq, init(50, 200, 100, 20)),
                new(PTHRotKSq, init(50, 200, 100, 20))));
  image[2] := new(PTImage, init(1,
                new(PTHRotKSq, init(50, 200, 100, 40)),
                new(PTHRotKSq, init(50, 200, 100, 40))));
  image[3] := new(PTImage, init(1,
                new(PTHRotKSq, init(50, 200, 100, -20)),
                new(PTHRotKSq, init(50, 200, 100, -20))));
  image[4] := new(PTImage, init(1,
                new(PTHRotKSq, init(50, 200, 100, -40)),
                new(PTHRotKSq, init(50, 200, 100, -40))));
```

```
{KSq Vertical Rotation}
image[5] := new(PTImage, init(2,
            new(PTVRotKSq, Init(50, 200, 100, 20)),
            new(PTVRotKSq, Init(50, 200, 100, 20))));
image[6] := new(PTImage, init(2,
            new(PTVRotKSq, Init(50, 200, 100, 40)),
            new(PTVRotKSq, Init(50, 200, 100, 40))));
image[7] := new(PTImage, init(2,
            new(PTVRotKSq, Init(50, 200, 100, -20)),
            new(PTVRotKSq, Init(50, 200, 100, -20))));
image[8] := new(PTImage, init(2,
            new(PTVRotKSq, Init(50, 200, 100, -40)),
            new(PTVRotKSq, Init(50, 200, 100, -40))));

{KSq Horizontal:  Random Monocular Noise}

image[9] := new(PTImage, init(3,
            new(PTHRotKSqRandDots, Init(50, 200, 100, 20)),
            new(PTHRotKSqRandDots, Init(50, 200, 100, 20))));
image[10] := new(PTImage, init(3,
            new(PTHRotKSqRandDots, Init(50, 200, 100, 40)),
            new(PTHRotKSqRandDots, Init(50, 200, 100, 40))));
image[11] := new(PTImage, init(3,
            new(PTHRotKSqRandDots, Init(50, 200, 100, -20)),
            new(PTHRotKSqRandDots, Init(50, 200, 100, -20))));

image[12] := new(PTImage, init(3,
            new(PTVRotKSqRandDots, Init(50, 200, 100, -40)),
            new(PTVRotKSqRandDots, Init(50, 200, 100, -40))));

{KSq Vertical:  Random Monocular Noise}

image[13] := new(PTImage, init(4,
            new(PTVRotKSqRandDots, Init(50, 200, 100, 20)),
            new(PTVRotKSqRandDots, Init(50, 200, 100, 20))));
image[14] := new(PTImage, init(4,
            new(PTVRotKSqRandDots, Init(50, 200, 100, 40)),
            new(PTVRotKSqRandDots, Init(50, 200, 100, 40))));
image[15] := new(PTImage, init(4,
            new(PTVRotKSqRandDots, Init(50, 200, 100, -20)),
            new(PTVRotKSqRandDots, Init(50, 200, 100, -20))));
image[16] := new(PTImage, init(4,
            new(PTVRotKSqRandDots, Init(50, 200, 100, -40)),
            new(PTVRotKSqRandDots, Init(50, 200, 100, -40))));

{KSq Horizontal: Random Noise at Random Depth}

image[17] := new(PTImage, init(5,
            new(PTHRotKSqRandRanDots, Init(50, 200, 100, 20)),
            new(PTHRotKSqRandDots, Init(50, 200, 100, 20))));
```

```
image[18] := new(PTImage, init(5,
        new(PTHRotKSqRandRanDots, Init(50, 200, 100, 40)),
        new(PTHRotKSqRandDots, Init(50, 200, 100, 40))));

image[19] := new(PTImage, init(5,
        new(PTHRotKSqRandRanDots, Init(50, 200, 100, -20)),
        new(PTHRotKSqRandDots, Init(50, 200, 100, -20))));

image[20] := new(PTImage, init(5,
        new(PTHRotKSqRandRanDots, Init(50, 200, 100, -40)),
        new(PTHRotKSqRandDots, Init(50, 200, 100, -40))));

{KSq Vertical: Random Noise at Random Depth}

image[21] := new(PTImage, init(6,
        new(PTVRotKSqRandRanDots, Init(50, 200, 100, 20)),
        new(PTVRotKSqRandDots, Init(50, 200, 100, 20))));


image[22] := new(PTImage, init(6,
        new(PTVRotKSqRandRanDots, Init(50, 200, 100, 40)),
        new(PTVRotKSqRandDots, Init(50, 200, 100, 40))));

image[23] := new(PTImage, init(6,
        new(PTVRotKSqRandRanDots, Init(50, 200, 100, -20)),
        new(PTVRotKSqRandDots, Init(50, 200, 100, -20))));

image[24] := new(PTImage, init(6,
        new(PTVRotKSqRandRanDots, Init(50, 200, 100, -40)),
        new(PTVRotKSqRandDots, Init(50, 200, 100, -40))));

{KSq Horizontal: Random Inducers}

image[25] := new(PTImage, init(7,
        new(PTHRotKSqRandInducers, Init(50, 200, 100, 20)),
        new(PTHRotKSqRandInducers, Init(50, 200, 100, 20))));

image[26] := new(PTImage, init(7,
        new(PTHRotKSqRandInducers, Init(50, 200, 100, 40)),
        new(PTHRotKSqRandInducers, Init(50, 200, 100, 40))));

image[27] := new(PTImage, init(7,
        new(PTHRotKSqRandInducers, Init(50, 200, 100, -20)),
        new(PTHRotKSqRandInducers, Init(50, 200, 100, -20))));

image[28] := new(PTImage, init(7,
        new(PTHRotKSqRandInducers, Init(50, 200, 100, -40)),
        new(PTHRotKSqRandInducers, Init(50, 200, 100, -40))));
```

```
{KSq Vertical: Random Inducers}
image[29] := new(PTImage, init(8,
            new(PTVRotKSqRandInducers, Init(50, 200, 100, 20)),
            new(PTVRotKSqRandInducers, Init(50, 200, 100, 20))));

image[30] := new(PTImage, init(8,
            new(PTVRotKSqRandInducers, Init(50, 200, 100, 40)),
            new(PTVRotKSqRandInducers, Init(50, 200, 100, 40))));

image[31] := new(PTImage, init(8,
            new(PTVRotKSqRandInducers, Init(50, 200, 100, -20)),
            new(PTVRotKSqRandInducers, Init(50, 200, 100, -20))));

image[32] := new(PTImage, init(8,
            new(PTVRotKSqRandInducers, Init(50, 200, 100, -40)),
            new(PTVRotKSqRandInducers, Init(50, 200, 100, -40))));

for i := 1 to NumImages do Image[i]^.scale (1);

{ text management - uses string buffer }
txtbuff := 'Press <Return> to start session'; display.centertext(txtbuff);
readln;
sound(440);
delay(500);
NoSound;
display.clear;
assign (outfile, datafile);
rewrite (outfile);
{this section of code randomises the stimuli within an experimental
 set - presents each stimulus to the subject and records their response
 to the stimuli.  It does this for each set in the session pausing
 briefly between each set. }

for sets := 1 to NumSets do
begin
  RandomiseSet;
    for I := 1 to NumImages do
    begin
        Write('       Trial  ',i,' of ',NumImages,' in iteration ', NumSets,'  :  ');

        { Set the display intensity required for the experimenal
          stimulus, draw the image, the reset the intensity to
          its original value }

        Image[i]^.Draw(display);


        { manage the orientation of the comparator - horizontal
          line for horizontal rotation, vertical line for
          vertical rotation and horizontal line for no
```

```
        rotation - maybe we should alter the code so that
        no rotation randomly presents a vertical or horizontal
        line }
      if(Image[i]^.LeftStimulus = Type_HRotKSq) then
      begin
       comparator.Moveto(Display.MidX - 280, Display.MidY);
       marker.Moveto(Display.MidX - 200, Display.MidY);
       marker.draw(Display);
       comparator.calibrate(0);  { Set line vertical }
       comparator.Draw(display); { Normalise and draw the comparator }
      end
      else
      begin { This code handles the case where there is no rotation }
       comparator.calibrate(90);
       comparator.Moveto(Display.MidX, Display.MidY + 180);
       marker.Moveto(Display.MidX, Display.MidY + 100);
       marker.draw(Display);
       comparator.Draw(display);
      end;

      { Now we must allow the subject to adjust the comparator or
        comparators to represent the desired angle.
        On return the comparator(s) angle can be aquired by using
        the comparator.deviation function to read the current value
        of the comparator.
        using trial as a procedure rather than a function allows for the
        use of more than one type of comparator in a single trial.}

      trial(comparator, display);

      DumpData(sets, comparator.deviation, Image[i]);
      display.clear;
      delay(500);
      sound(800);
      delay(100);
      NoSound;

    end;
    sound(50);
    delay(1000);
    NoSound;

    str(sets, numbuff); txtbuff := 'End of Session !';
    display.centertext(txtbuff);
    readln;
end;
Writeln('End of Experiment');
readln;
{ At the end of the session we need to clean up the objects
  we created at the start.  The image destructors release the
```

```
    memory aquired by the image objects and by objects within the
    image object.  The display destructor takes care of releasing
    any resources that the display constructors may have aquired
    (such as font memory in the VSG card )
  for i := 1 to NumImages do dispose(image[i], done);
  display.done; { destroy the display object }
  close (outfile);
end.
```