

Handwritten Signature Verification Using Complementary Statistical Models

Thesis submitted by
Alan McCabe
November, 2003

for the Degree of Doctor of Philosophy
in the School of Information Technology at
James Cook University of North Queensland.

Supervisor:
Doctor Bruce Litow

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institute of tertiary education. Information derived from the published and unpublished work of others has been acknowledged in the text, and a list of references is given.

Alan McCabe

November 10, 2003

Statement on Access to this Thesis

I, the under-signed, the author of this thesis, understand that James Cook University of North Queensland will make it available for use within the University Library and, via the Australian Digital Thesis network, for use elsewhere. 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 written consent of the author; and to make proper written acknowledgement for any assistance I have obtained from it.

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

Alan McCabe

November 10, 2003

ELECTRONIC COPY

I, the undersigned, the author of this work, declare that the electronic copy of this thesis provided to the James Cook University Library, is an accurate copy of the print thesis submitted, within the limits of the technology available.

Signature

Date

Acknowledgements

I would like to thank a number of people who have made it possible for me to complete this project. Firstly I would like to thank the School of Information Technology and in particular my supervisor Dr Bruce Litow. I appreciate Dr Litow's time, constructive advice and his words of wisdom, many of which I remain unable to comprehend.

I also need to thank a number of friends and work colleagues for providing an outlet for the many frustrations over the past few years, as well as always being a willing source of distraction, whether it was required or not. I would especially like to thank Gerry O'Reilly who has continually remained patient, supportive and understanding, regardless of the pressures or the situation. She has always managed to put a smile on my face.

Finally and most significantly, I would like to thank both of my parents, Frank and Vicky, whose generosity quite simply knows no bounds. None of this would have been possible without the unlimited support, in all its forms, offered by them both.

Abstract

There is considerable interest in computerised personal identification and in particular in *biometrics*, a branch of identification that deals with verifying physical or behavioural characteristics of human beings. This thesis is concerned with the development of the particular biometric of handwritten signature verification, which is superior in many ways to other biometric authentication techniques that may be reliable but are much more intrusive.

Specifically, this project involves the use of two complementary artificially intelligent systems in the form of neural networks and hidden Markov models. Five sample signatures are used to build a reference in each of the independent models and experimentation and testing is done using an extensive database of almost 4,000 genuine signatures and forgeries. The confidence levels from each model are then combined and tested on unseen signatures resulting in an equal error rate of 1.1%. Further experimentation is performed and includes analysis of different verification scenarios, error contribution and the importance of visual feedback when signing. Finally, experiments are conducted exploring the possibility of “signing” handwritten passwords, with the developed system resulting in an equal error rate of 0.7% in the worst case.

Major Contributions

- A new method of signature segmentation, detailed in Appendix A, which dramatically reduces “false” segments;
- The development of successful methods for selecting effective training forgeries from a body of other users’ reference signatures;
- A method for comparing signatures via string edit distance, outlined in Appendix B;
- Examination and development of methods for combining the output of neural networks and hidden Markov models;
- The introduction of several new features not previously used in on-line handwritten signature verification;
- A study on the need for a signer to have visual feedback when performing their signature.

Contents

1	Introduction	1
1.1	Handwriting	3
1.2	Handwritten Signatures	3
1.3	Design Overview	6
1.4	Thesis Structure	8
2	Literature Review	9
2.1	Applications of Handwritten Signature Verification	10
2.2	Automated Handwritten Signature Verification	11
2.2.1	Dynamics of Signature Production	11
2.2.2	The Basic Methodology	14
2.3	Review of Earlier Work	21
2.3.1	Combining Local and Global Features	28
3	Neural Networks	32
3.1	The Theory of Neural Networks	33
3.1.1	Learning in Neural Networks	35
3.1.2	Linear Networks	44
3.1.3	Multi-Layer Perceptrons	47
3.1.4	Radial Basis Functions	53
3.1.5	Bayesian Networks	54
3.1.6	Kohonen Self-Organising Maps	56
3.1.7	Autoassociative Networks	57
3.2	Applications to Handwritten Signature Verification	57
3.3	Previous Work	58
3.4	Methodology	62
3.4.1	Pre-processing	62
3.4.2	Signature Database	63
3.4.3	Extracted Features	66

3.4.4	Experimental Setup	90
4	Hidden Markov Models	105
4.1	The Theory of Hidden Markov Models	106
4.1.1	Markov Models	107
4.1.2	The Hidden Layer	109
4.1.3	Bakis Models	117
4.1.4	Learning in Hidden Markov Models	118
4.2	Applications to Handwritten Signature Verification	121
4.3	Previous Work	121
4.4	Methodology	125
4.4.1	Signature Segmentation	125
4.4.2	Extracted Features	126
4.4.3	Experimental Setup	132
5	Combining Multiple Models	137
5.1	Previous Work	138
5.2	Methodology	145
5.2.1	Experimental Setup	146
5.3	Further Results	152
5.3.1	Removal of “Short Signatures”	153
5.3.2	Contribution to Overall Error	153
5.3.3	Allowing Users Another Chance When Rejected	155
5.3.4	Varying the Size of the Reference Set	156
5.3.5	Zero-Effort False Acceptance Rate	157
5.3.6	The Importance of Visual Feedback When Signing	158
5.3.7	Manually Adjusted Personal Thresholds	159
5.3.8	Signing a Password	160
6	Conclusion	162
	Appendices	165
A	The Extremum Consistency Algorithm	165
A.1	Introduction	165
A.2	The Problem Domain - Motivation	166
A.3	The Algorithm	168
A.4	Successful Applications	171
A.4.1	Direction Based Handwritten Signature Verification	172

A.4.2	Velocity Based Handwritten Word Verification	175
A.4.3	Physiology Research - Tracking Fluctuations in Infant Face Temperature	176
A.5	Future Work	176
A.6	Conclusion	176
B	Signature Similarity Via Edit Distance	178

List of Figures

1.1	<i>A typical handwritten word, sampled at 205 points per second. Each dot in the handwriting represents the position of the pen tip for one of these samples.</i>	2
1.2	<i>The various forces at work in generation of handwriting. A is the “pen pressure” exerted by the writer perpendicular to the axis of the writing instrument; B is the “point load”, which is the component of pressure exerted perpendicular to the writing surface, responsible for indentations, line thickness etc.; and C is the “travel action”, the pressure component exerted in two dimensions (forward/sideward for upstrokes or crossings and drag/backward for downstrokes) across the writing surface, responsible for line generation. Together these forces contribute to features like velocity, acceleration, shape etc.</i>	4
2.1	<i>The equal error rate is the error value at which the false acceptance and false rejection rates are the same.</i>	15
3.1	<i>The general structure of a neuron.</i>	36
3.2	<i>The sigmoid activation function.</i>	37
3.3	<i>The hyperbolic activation function.</i>	37
3.4	<i>The ramp activation function.</i>	38
3.5	<i>The step activation function.</i>	39
3.6	<i>The sign activation function.</i>	39
3.7	<i>Various feed-forward network topologies. (a) A simple two-input, one output network with no hidden layers. (b) A two-layer network with two inputs, two hidden nodes and one output node. (c) A more complicated network consisting of eight input nodes each connected to four nodes in the hidden layer and a single output node. (d) A network similar to that in (c) except with two hidden layers.</i>	41

3.8	<i>A linearly-separable feature space.</i>	45
3.9	<i>A linear network.</i>	46
3.10	<i>Linear separability of Boolean functions - the axes represent the input values and the dots represent the output (a solid dot is a 1 and a hollow dot is a 0). (a) The AND function which is linearly separable. (b) The OR function which is linearly separable. (c) The XOR function - it is not possible to draw a single line to separate the classes.</i>	46
3.11	<i>The sampled coordinates captured from the handwritten word "farley".</i>	65
3.12	<i>Interpolation of the sampled coordinates produces the off-line, or static, image of the word.</i>	65
3.13	<i>Contributors to the database grouped according to (a) nationalities, (b) handedness (left or right), (c) age and (d) gender.</i>	67
3.14	<i>This is an illustration of the difficulty that a potential forger has in trying to identify the pen-down ratio. The sample in (a) is a genuine signature and (b) is an attempted forgery based on the forger having seen an off-line version of the signature (both taken from the signature database used in this project). The pen-down ratio for the genuine signature is 0.992 and is 0.879 for the forgery (forgeries were typically found to have much lower pen-down ratios, presumably because of the extra attention to detail).</i>	70
3.15	<i>Horizontal length of a typical handwritten word is a simple feature to comprehend and calculate. The horizontal length of this sample is 1,345 pixels.</i>	70
3.16	<i>Cursivity varies widely between different authors while tending to remain similar for different samples produced by the same author. For example consider parts (a) and (b) above that contain words written by different authors with very different cursivity values of 16.0 and 3.6 respectively. Part (c) is another sample of the same written word as (b), by the same author, and has a very similar cursivity value of 3.8.</i>	72

3.17	<i>Cursiveness varies somewhat between authors and is a feature that is highly indicative of natural handwriting style. (a) shows a signature with a seemingly high cursiveness, but the actual value for this is 12 which is significantly lower than the signature in (b) at 125. These signatures are examples of how visual inspection can be quite deceptive in estimation of cursiveness.</i>	73
3.18	<i>This is an illustration of how the different measures of central tendency can give different midpoints for the calculation of top-heaviness. The figure has three horizontal lines drawn to illustrate the location of the calculated midpoint using the different central tendency measures.</i>	74
3.19	<i>Different handwriting samples can result in quite different curvature values. For example, (a) shows a sample in which the writing is quite flat and not well-formed, resulting in a curvature value of 3.96. Conversely (b) shows a sample with a much more pronounced forming of the handwritten characters resulting in the higher value of 5.22.</i>	75
3.20	<i>The process of calculating the average curvature per stroke. (a) shows the entire handwritten word and (b) shows an isolated view of one of the extracted strokes.</i>	75
3.21	<i>Two signatures sections produced by the same author illustrating the consistency in the number of strokes. The crosses on the handwriting represent the stroke boundaries. Both of these samples have 21 strokes and as can be seen the segmentation is quite consistent.</i>	76
3.22	<i>A typical handwriting sample with labels indicating the ascenders, descenders, mean vertical displacement, ascender height and descender depth.</i>	77
3.23	<i>The maximum height of a signature or handwritten word is defined as the distance from the top of the highest ascender to the bottom of the lowest descender. The vertical line seen here to the right of the writing sample is the maximum height, and in this case is calculated as 1,005 pixels.</i>	78

3.24	<i>The gradient of the line between each pair of consecutive points is determined (a sample of which is shown in part (a)), and the mean of those values found - this mean is the slant of the handwriting. Part (b) illustrates the computed slant value, drawn as a series of dotted lines laid over the handwriting sample.</i>	79
3.25	<i>“Long strokes” extracted from a typical handwriting sample. The long stroke is represented as bolded handwriting with the remainder of the handwriting appearing as a broken line in the background.</i>	80
3.26	<i>Calculation of handwriting slant through regression of “long strokes”. (a) shows one of the long strokes extracted from a typical handwriting sample. (b) shows a close-up view of that same stroke with the straight line being the line-of-best-fit as produced by simple linear regression. The gradient of this line is taken as the handwriting slant. (c) shows the same handwriting sample used in (a) and is overlayed with a series of straight lines parallel to the calculated slant using the regression of long strokes.</i>	81
3.27	<i>The dotted line represents a single vertical projection, one of many used in the calculation of vertical overlap. The crosses are the points of intersection between the vertical projection and the handwriting stream (there are five in this instance). The average number of intersections is then a measure of the degree of handwriting slant (the higher the slant the higher the number of intersections).</i>	82
3.28	<i>Stroke concavity is depicted in this figure, showing a closeup of a stroke segment with a line-of-best-fit (found by simple linear regression) drawn through four points. The concavity is then found by taking the square root of the sum of squares of the minimum distance from each point in the stroke to the line-of-best-fit.</i>	83
3.29	<i>Depending on the feature used, it may be necessary to remove certain pixels from calculation. Typically, fragmented information such as the dotting of ‘i’s and the crossing of ‘t’s are removed.</i>	84

3.30	<i>Different degrees of parallelism. (a) and (b) are two sections of signatures taken from different authors with different values for parallelism. The sample in part (a) has a parallelism value of 0.10, whereas (b) has 0.34.</i>	86
3.31	<i>The various stages in the calculation of baseline consistency. (a) shows the original handwritten word, (b) shows the extracted minima for non-descender characters and (c) shows the line-of-best-fit calculated for these points using linear regression. The baseline consistency is then the square root of the sum of the squares of the distances between the extracted minima and the line. The baseline consistency of this handwriting sample is 25.2.</i>	87
3.32	<i>The area of a signature. (a) shows the original sample and (b) shows the calculated area. In (b) the black lines represent the vertical extremities (the maximum and minimum intersections with vertical projections) and the shading shows the area of the signature segment. The area of signatures with multiple components is found by summing each of the independently calculated component areas.</i>	89
3.33	<i>This figure illustrates “middle-heaviness”, which is defined as the percentage of the bounding box of a signature that is interior to the sample itself. The bounding box is shown in the figure and all shaded pixels are points interior to (or part of) the sample. The area of the shaded pixels is then divided by the area of the bounding box to give middle-heaviness.</i>	90
3.34	<i>The physical spacing between components is a measure of the average distance between the last point sampled in a component and the first point sampled in the immediately following component (if any). This distance is illustrated in the figure and defaults to zero if there is only one component.</i>	91
3.35	<i>Convergence of training and verification errors. (a) In a linear network. (b) In a multi-layer perceptron with a single hidden layer.</i>	93
3.36	<i>Error rates resulting from varying the number of hidden nodes in a MLP with one hidden layer.</i>	94
3.37	<i>Convergence of error rate using back-propagation in a typical MLP with no negative examples.</i>	99

3.38	<i>The performance of the optimal network structure using different threshold values.</i>	104
4.1	<i>A Markov model of the weather.</i>	108
4.2	<i>A HMM that models coin tosses using (a) 2 and (b) 3 states.</i>	111
4.3	<i>A summary of Expectation Maximization (EM) training in hidden Markov models.</i>	117
4.4	<i>A 5-state left-right, or Bakis, model. This example features loop (for example, 1 to 1), forward (1 to 2) and skip (1 to 3) transitions.</i>	118
4.5	<i>This figure presents a signature from the signature database. Using the velocity based stroke (VBS) technique for segmentation results in 290 strokes, whereas the extremum consistency approach results in 217 strokes.</i>	126
4.6	<i>This figure represents the horizontal length of a typical stroke. (a) contains the original handwritten sample with an extracted stroke in bold. (b) shows an expanded view of that same stroke with the horizontal length marked.</i>	127
4.7	<i>Different strokes can result in quite different curvature values. For example, (a) shows a sample stroke that is quite flat, resulting in a curvature value of 0.01. Conversely (b) shows a sample with a much more pronounced curve that results in the higher curvature value of 0.40.</i>	128
4.8	<i>Handwriting slant calculated using stroke end-points. This is the same stroke as shown in Figure 4.7, depicted here as the series of sampled points. The solid line to the immediate right is the calculated slant.</i>	129
4.9	<i>An illustration of handwriting slant calculated through regression. (a) shows the original word as a series of sampled points with the extracted stroke in bold and (b) shows slant calculated via regression.</i>	129
4.10	<i>A graphical representation of the beginning and ending gradient values within the stroke.</i>	131
4.11	<i>A plot of the HSV results using the Segmental K-Means learning algorithm and different threshold values.</i>	135
4.12	<i>A plot of the HSV results using the Baum-Welch learning algorithm and different threshold values.</i>	136
5.1	<i>The combination of models described in previous chapters.</i>	146

5.2	<i>The MLP structure that produced the lowest overall error rate when combining the constituent systems. Each of the weight values W_i is optimised via a learning algorithm.</i>	150
5.3	<i>The average signature duration (in seconds) per signer.</i>	154
5.4	<i>The overall error rate versus the duration threshold. Signers with an average signature duration less than t seconds were removed from consideration. As can be seen, error rates generally improve as signature duration increases.</i>	154
5.5	<i>A plot of individual contributions to overall error rate, sorted in order of increasing contribution.</i>	155
5.6	<i>The overall error rate versus the number of reference signatures used.</i>	157
5.7	<i>(a) A signature sample captured using a stylus to provide visual feedback to the signer. (b) A signature sample from the same author captured without the use of the stylus.</i>	158
A.1	<i>This figure represents some local minima situations which are typically encountered in processing the input stream. The horizontal axis represents increasing time and the vertical axis can represent various stream types such as velocity, direction and temperature. Specifically, (a) contains a valid minimum, (b) contains only a single valid minimum (there are actually two minima, but the second is the result of the local maximum in the centre, which should be ignored in this environment) and all others contain no “true” minima. An effective algorithm should reflect this.</i>	167
A.2	<i>Situations like this are the result of the limited resolution of the hardware used to capture a stream. The black line represents the actual value of the stream and the black dots represent the recorded value. Time is represented on the horizontal axis. This situation typically arises when the hardware is a graphics tablet which rounds the position of the pen tip to the nearest pixel, but also comes up with (say) temperature observations when the actual temperature is rounded to the nearest tenth of a degree for recording.</i>	168

A.3	<i>A finite state machine expressing the EC algorithm. The movements between vertices (states) are defined by the comparison between points in the input stream and the comparisons are included on the edges in the diagram. Additionally there are actions to be performed when vertices are reached - these are also included in the diagram.</i>	169
A.4	<i>An illustration of step and width calculation. Valid steps occur between time points 0 and 1, 1 and 2, 3 and 4 and 5 and 6. Backward steps occur between time points 6 and 7, as well as between 9 and 10.</i>	170
A.5	<i>An illustration of a large downslope with a small upslope. The minimum indicated by (1) is more likely the result of noise than a genuine valley and should be ignored (that is, it is necessary to take the minimum of the upslope and the downslope, rather than the sum or average). The better minimum would be that indicated by (2).</i>	171
B.1	<i>A portion of a signature that has been segmented. The crosses on the diagram represent stroke start and/or end points. . . .</i>	179
B.2	<i>The start-point of each stroke is placed at the origin and the quadrant in which the end-point lies is recorded as the observation. The line appearing in the figure in quadrant A represents a stroke extracted from a handwritten word.</i>	179

List of Tables

3.1	<i>A summary of the main database of signatures used in experimentation.</i>	68
3.2	<i>The training performance of three network structures.</i>	96
3.3	<i>The classification accuracy of the three implemented neural network training algorithms.</i>	97
3.4	<i>The convergence speed of the three implemented neural network training algorithms.</i>	97
3.5	<i>The performance of the HSV system using different approaches to obtaining negative examples.</i>	102
3.6	<i>The performance of the HSV system using different MLP structures. Unless otherwise stated, the structure made use of a single hidden layer.</i>	103
4.1	<i>The training performance of the two HMM training algorithms used in experimentation. The “Number of Epochs” is the mean number of epochs required for convergence to occur (with standard deviation in brackets). The relative time compares the elapsed time prior to convergence.</i>	134
4.2	<i>A comparison of the modelling accuracy of the Segmental K-Means and Baum-Welch HMM training algorithms. All significant HSV-related error rates are reported in this table.</i>	135
4.3	<i>A comparison of the modelling accuracy of the Segmental K-Means and Baum-Welch HMM training algorithms when ten reference signatures are used to train the model.</i>	136
5.1	<i>The results using the two different voting mechanisms to combine the classifiers.</i>	147
5.2	<i>The resulting error rates using the different confidence-based approaches to combining the classifiers.</i>	152

5.3	<i>The most successful results for each of the different model scenarios used during development.</i>	152
5.4	<i>A breakdown of the error rates for various reference set sizes, optimised to give the lowest overall error rate.</i>	156
5.5	<i>The resulting error rates using different types of forgeries in the “signing passwords” variant of the HSV system.</i>	161
A.1	<i>Error rates using various methods of overcoming local extrema in a specific signature verification environment. If there are parameters involved in the operation (such as convolution window size) then the parameters which produced the lowest overall error rates were used to generate the results.</i>	174