Sample Selection for Optimising Signature Enrolment

Richard Guest, Mike Fairhurst

Sample Selection for Optimising Signature Enrolment

Richard Guest and Mike Fairhurst
Department of Electronics
University of Kent, UK
{r.m.guest},{m.c.fairhurst}@kent.ac.uk

Abstract

Accurate and representative template construction is critical for the successful and robust operation of dynamic signature verification systems. This paper presents a comparison of enrolment sample selection strategies using a large database of donated signature samples. Enrolment strategies evaluated include several that are based on sample variance within standard signature performance features, together with a series of established reference methods used for purposes of "baseline" comparison. Sample selection enrolment strategies are tested on two representative commercial signature verification systems by assessing the rate of successful enrolment and verification. The results show that using a prediction method based on the measurement of variance within samples for selection increases the ability to successfully enrol. The investigation is thus important in helping to develop possible best practice for implementing effective enrolment procedures.

Keywords: Biometrics, Automatic signature verification, Enrolment, Templates

1. Introduction

The handwritten signature is widely used as a form of personal biometric identification and verification [1]. All biometric systems require the enrolment of a user through a process of template creation, this template then being used for comparison against samples in future verification operations. The verification performance of a system depends on a number of factors but a key consideration is the quality of the data used for template construction at enrolment; indeed, most systems include an internal checking mechanism to assess whether enrolment samples satisfactorily capture representative user characteristics. Previous reported work on template construction has indicated performance enhancement in fingerprint [2] and face [3] biometric modalities by selecting a subset of samples from an individual's enrolment donation pool. The motivation behind the intelligent selection of a subset of samples is to improve the performance in terms of successful enrolment on a system as well as verification against formed templates; maximising the performance by the selection of available data.

This study describes a comparison of signature sample selection strategies for the creation of templates and, using two commercial signature verification systems, assesses the effectiveness of these strategies in terms of enrolment and verification success.

2. Enrolment Methodology

To assess enrolment and verification performance, two independent standard commercial signature verification systems were employed. Both systems use a three-sample *signature triplet* from each test subject to form an enrolment template. During the enrolment procedure each system checks the consistency of the triplet to assess if its suitability for the formation of a template. In this study, the commercial systems were treated as 'black-boxes' in that the internal methods for signature enrolment and verification were ignored – the study was merely concerned with the enrolment performance using the proprietary selection strategies.

The signature data used in this study were collected from members of the general public. 284 subjects (100 Male/184 Female with a mean age of 52.6, 27 left/257 right normal signing hand) were included in this study, giving a total of 6733 signatures in all. Subjects were asked to donate several signatures at their first visit and invited to donate further signatures during subsequent visits to a test centre. All subjects in this trial donated signature samples over two or more visits (or *sessions*) – the first session containing at least five signatures with a second session containing at least three signatures. A different session was defined as being more than 10 minutes after the previous signature was donated, with most sessions taking place more than a week apart. Data was captured at 100Hz sampling frequency from a conventional graphics tablet (304.8 x 304.8 mm) at a resolution of 500 lines per inch (19.56 lines per mm) using an inking pen on paper which was overlaid and secured on to the tablet surface.

In the experiments reported in this paper three method-groupings were devised to select the three signatures used for enrolment. Signatures could be selected from a single or from multiple collection sessions.

2.1. Method 1: Dynamic Variance Calculation

This method finds the triplet of enrolment samples with the lowest (and, for comparison, the highest) variance within a set of features *dynamically* selected from a subject's samples. A set of standard global features (measuring performance over the entire signature as opposed to a localized area of interest) were extracted from each signature. Global features were used as they are able to give an indication of general performance variation between signatures without the need for accurate and repeatable localization and segmentation processing. The extracted features are all commonly used in signature verification experimentation [4] [5]. This method was dynamic in that it allowed for a different selection of features for each subject and scenario. These features are listed in Table 1.

Table 1. Feature Rankings

Rank	Feature
1	Total pen down duration / Total signing duration
2	Number of pen down sample points
3	Duration when velocity in the X plane < 0 / Total pen down duration
4	Duration when velocity in the Y plane < 0 / Total pen down duration
5	Total pen travel writing distance
6	Total pen down duration
7	Average velocity in the Y plane
8	Minimum velocity in the Y plane / Average velocity in the Y plane
9	X max coordinate - X min coordinate
10	Average velocity in the X plane
11	Minimum velocity in the X plane / Average velocity in the X plane
12	Duration when velocity in the Y plane > 0 / Total pen down duration
13	Duration when velocity in the X plane > 0 / Total pen down duration
14	Total pen travel writing distance / signature area
15	Y max coordinate - Y min coordinate
16	Time of min velocity in the X plane / Total pen down duration
17	Time of min velocity in the Y plane / Total pen down duration
18	(X max coordinate - X min coordinate) / (Y max coordinate - Y min coordinate)
19	Vertical midpoint - Y min / (Y max coordinate - Y min coordinate)
20	Total number of samples when velocity = 0 in the Y plane
21	Total number of samples when velocity = 0 in the X plane
22	Number of vertical midpoint crossings in signatures
23	Mean velocity / Maximum velocity
24	Average velocity / Maximum velocity in the Y plane
25	Number of pen ups within signature
26	Average velocity / Maximum velocity in the X plane
27	Time of 2nd pen down / Total signing duration
28	First instance of velocity ≠ 0
29	Correlation between velocities in the X and Y planes
30	Time of max velocity in the X plane / Total pen down duration
31	Time of max pen velocity / Total pen down duration
32	Time of max velocity in the Y plane / Total pen down duration
33	First order moment

The three stage process used for this method was as follows:

1. Separately assess the Coefficient of Variance (COV) for each feature across all samples for a particular subject in a selected signing session (all sessions, 1st or 2nd). A COV expresses the standard deviation of a dataset as a percentage of the mean value. In this way the magnitude of feature results do not prevent a direct comparison in variation. COV is calculated as:

$$COV = \frac{\sigma^2(x)}{\left|\bar{x}\right|} \times 100 \qquad (1)$$

where $\bar{x}$ is the mean of all values of a particular feature and $\sigma^2(x)$ is the standard deviation of all values of a particular feature.

2. Rank the features according to their COV values.

3. Find the sample triplet with the lowest (and, separately, highest) COV using the lowest (and highest) ranked *n* features (*n* was varied between 5 and 33 features).

A series of *scenarios* were devised to explore this method:

- Lowest COV (All Sessions, 1st Session and 2nd Session) - the three samples from all/1st/2nd signing sessions with the COV producing the minimum score using the *n* best features selected for enrolment.

- Highest COV (All Sessions, 1st Session and 2nd Session) - the three samples from all/1st/2nd signing sessions with the COV producing the maximum score using the *n* worst features selected for enrolment.

Comparison between these the High and Low COV methods provide an indication of whether the verification systems improve in enrolment performance in relation to degree of agreement or diversity within samples.

2.2. Method 2: A Priori Feature Selection

Using the same triplet selection scenarios as used in Method 1, 5 to 33 'best'/'worst' features were selected (from the same feature set) according to repeatability/stability rankings derived from a previous study(rather than dynamically selecting within each test subject as utilised in Method 1) [6]. These rankings were applied to all test subjects. Table 1 also shows the ranking of the features. Again, the signature triplet with the lowest and highest COV from the selected features determined the triplet selection used in enrolment.

2.3. Method 3: Reference baseline performance

For performance comparison against other standard reference systems, eight more strategies were used for sample selection:

- **First** 3 Samples (Session 1 and Session 2) - selects the first three samples from the first and second signing sessions.

- **Last** 3 Samples (All Sessions, Session 1 and Session 2) - selects the last three samples from the first and all signing sessions. This enables an analysis of enrolment performance change over prolonged use. The last three samples for all sessions are selected as data may be collected over more than two sessions. In this way, the performance of the three most recently donated signatures (when, theoretically, the user has become most familiar with the system) can be assessed.

- **Random** 3 Samples (All Sessions, Session 1 and Session 2) - selects three random signature samples from the first and all signing sessions.

In each case the selected sample triplets were used to independently enroll and form templates for each of the two target systems, each system determining whether or not the successful formation of an enrolment template had been achieved. The percentage of successful enrolments using each method was recorded.

3. Verification Assessment Methodology

Using the templates formed in a successful enrolment, verification performance was assessed using the two signature systems with the remaining signature samples for each test subject (i.e. signature samples that were not used during the enrolment phase). If enrolment was unsuccessful for a particular test subject then no verification was conducted. Verification performance was calculated by assessing the percentage of signatures successfully verified against formed templates.

4. Results

Table 2 shows the percentage of subjects that successfully enrolled on the two commercial systems using the three sample selection strategies. It is interesting to note the variability in enrolment performance from the different reference strategies (and also the variation in enrolment performance between the two systems given a common set of signature samples). Although the within-system differences are small between the reference methods, these results show that the choice of enrolment samples can affect enrolment performance. While the small difference may not provide substantive evidence, the last three samples of a session typically provide optimum performance for both systems, outperforming strategies based on either the first three samples or a random selection, possibly indicating that as the enrollee becomes more familiar with the system, the more ‘stable’ the signature becomes.

Table 2. % of subjects successfully enrolled using the reference strategies

Triplet	Session	System 1	System 2
First	1	84.9	96.6
	2	85.9	96.9
Last	All	86.9	96.7
	1	85.5	96.9
	2	85.5	96.9
Random	All	82.4	93.6
	1	84.4	93.8
	2	85.9	96.9

To further investigate this effect, the sum of the COVs across the first three samples in the first and second sessions were analysed. The mean COV for the first session was 242.6 compared to 229.1 in the second session indicating a greater stability of samples in the second session. It can be argued therefore that using reference methodologies, if enrolment data from a single session is required, stability between samples improves within later sessions and templates could be formed from these data if feasible within an enrolment scenario. Optimum reference strategy results will, however, be achieved by using samples from more than one session.

Figures 1 to 4 show the enrolment results across the range of selected features using the standard commercial systems tested (Systems 1 and 2) and Enrolment Methods 1 and 2 as described in Section 2. In these Figures (and Figures 5 to 8 which show verification results) the following legend applies:

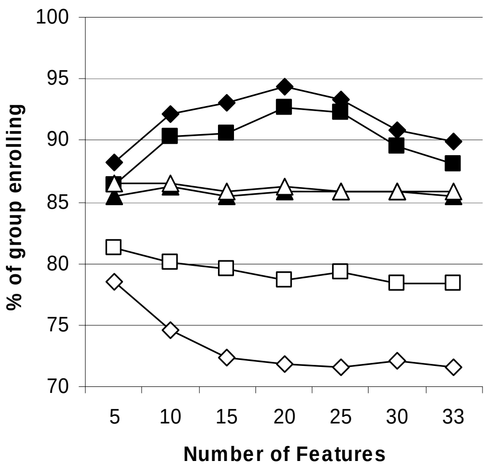

Figure 1. System 1, Method 1 Enrolment

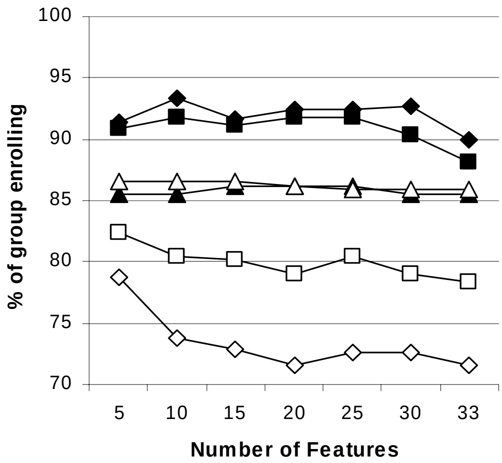

Figure 2. System 1, Method 2 Enrolment

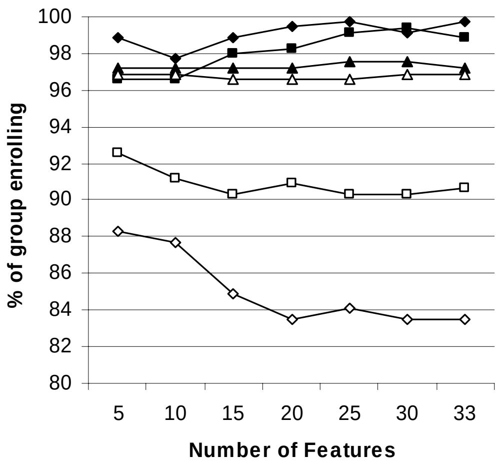

Figure 3. System 2, Method 1 Enrolment

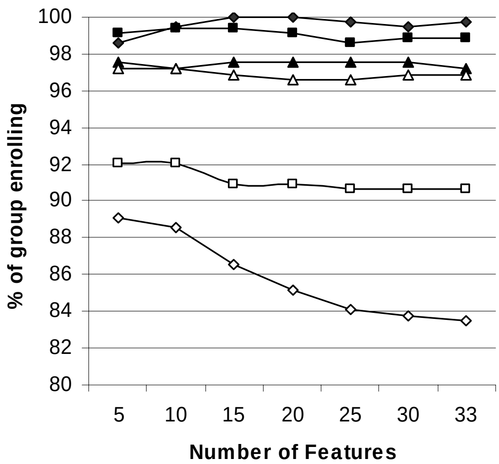

Figure 4. System 2, Method 2 Enrolment

Both systems return optimum performance when using a triplet with Low COV indicating that both best exploit similarity rather than diversity of samples when forming templates. Taking the optimum reference results for the two systems (86.87% for System 1 and 96.89%

for System 2) the results show that improvement in enrolment rates can be achieved by judicious strategy selection. In particular, taking the 'best' 15 to 20 samples with Low COV from more than one session results in higher enrolment rates.

For System 1, the dynamic selection of 20 features (Method 1) produced the highest enrolment rates. Again, the selection of 20 features produced the highest enrolment rates for System 2; however, optimum performance was achieved from selection of ranked features (Method 2). In all cases, as with the reference data, samples selected across all sessions produce the best enrolment rates. When using COV selection methods with a single session, data from the first session produced the best result although this can be attributed to the fact that a greater number of signature samples were captured in this session, thereby generating more combinations from which to select an optimum subset for enrolment. None of the High COV results produced an improvement on reference performance.

Table 3 shows the verification results in terms of the percentage of signatures successfully verified against formed templates (from a successful subject enrolment) across the entire dataset. It is evident that there is a trade-off between enrolment and verification performance, with the strategies that result in low enrolment rates producing high verification rates. Hence, System 1, which had lower enrolment rates than System 2 (as shown in Table 3), now has higher verification rates due to a lower number of verification samples and more 'tightly' controlled templates. Indeed, for System 1, the best verification performance is obtained from a random selection of samples across all sessions - the method which resulted in the lowest enrolment rate for System 2 (again see Table 3). For System 2 the picture is less clear, with the best verification results being produced by the strategy with the best enrolment rate, perhaps highlighting the robust nature of templates created by the second system.

Table 3. % of signatures successfully verified using templates formed by standard enrolment strategies.

Triplet	Session	System 1	System 2
First	1	80.0	60.8
	2	84.0	62.7
Last	All	84.8	60.2
	1	83.0	59.1
	2	84.0	61.4
Random	All	89.8	58.6
	1	83.9	59.3
	2	84.0	62.7

Figures 5 to 8 show the successful verification rates using the remaining signature samples verified against the corresponding enrolment templates.

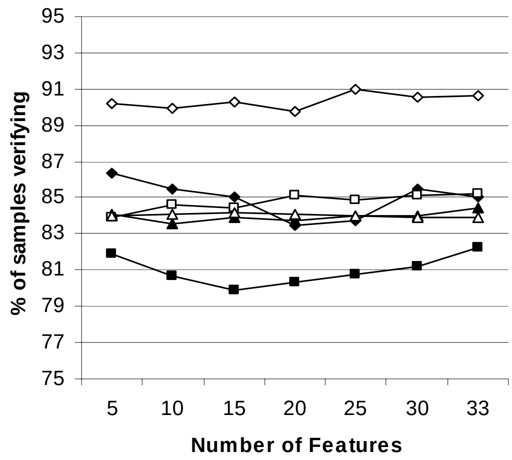

Figure 5. System 1, Method 1 Verification

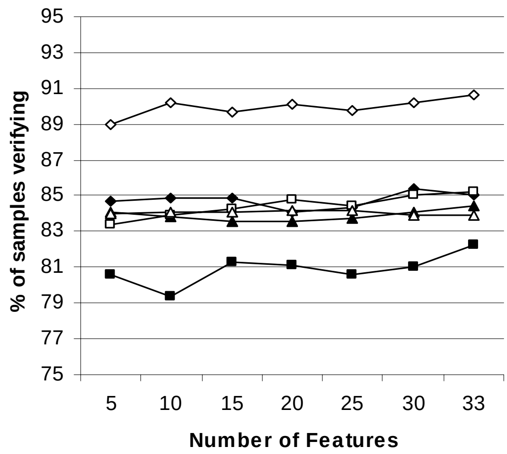

Figure 6. System 1, Method 2 Verification

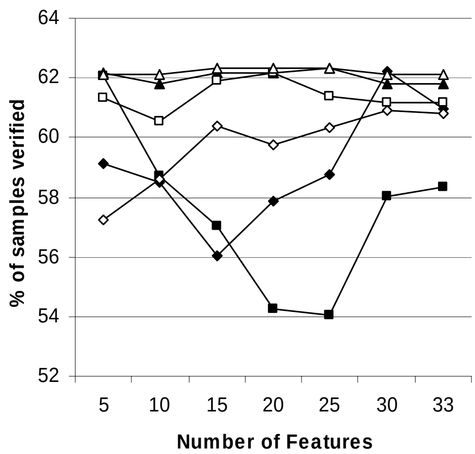

Figure 7. System 2, Method 1 Verification

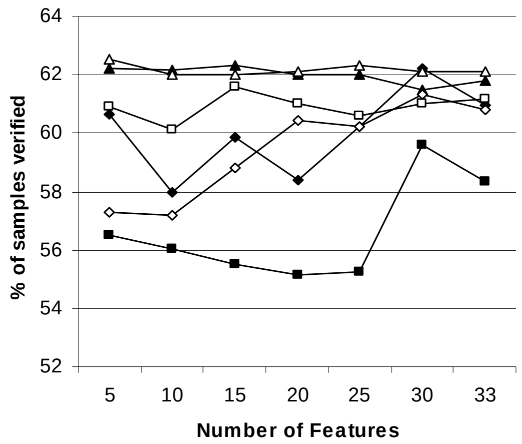

Figure 8. System 2, Method 2 Verification

Observing the verification results from both systems it is evident that, as with the reference results, a low enrolment rate typically results in a higher verification rate (for example with all the High COV enrolment scenarios). Assessing the performance of System 1, however, we can note that some methods (for example the Low COV across all sessions) show improved performance over the reference methods for both enrolment and verification. The findings are slightly different for System 2, with excellent verification results being obtained from the Low COV scenario with enrolment samples taken from the second session. Both this second session Low COV and the second session High COV scenarios have increased performance over the reference methods.

5. Conclusion

In this paper we have shown that two methods based on the dynamic selection and a priori knowledge of feature variation within signature samples can lead to improved template construction during enrolment. Using two commercial signature verification systems for testing purposes, both variance-based enrolment methods also result in higher verification rates based on the templates they produce when compared with some standard reference strategies.

It is clear from this investigation that familiarity and practice in sample donation for any particular system is likely to be important in ensuring subsequent performance optimization. Of particular importance, however, is that our experimentation has been carried out using established commercial systems and with data from a cross-section of the general public in a typical everyday transaction environment, and hence the strategies investigated are assessed in conditions which reflect likely performance characteristics achievable in practical applications.

Collectively, the results of the initial study reported here represent a first step towards developing some viable guidelines for best practice in the application of signature verification systems, and future work in this area will aim to develop this theme in a more objective and generic way.

Acknowledgment

The authors gratefully acknowledge the support of the UK Engineering and Physical Sciences Research Council (EPSRC) for part of the work reported here.

References

[1] R. Plamondon, S.N. Srihari, "On-line and off-line handwriting recognition: A comprehensive survey", IEEE *Trans PAMI*, 22, January 2000, pp. 63-84.

[2] X. Jiang, W. Ser, "Online fingerprint template improvement", *IEEE Trans PAMI*, 24, August 2002, pp. 1121–1126

[3] X. Liu, T. Chen, S. M. Thornton, "Eigenspace updating for non-stationary process and its application to face recognition", *Pattern Recognition*, 36 (9), September 2003, pp. 1945-1959

[4] B. Kashi, J. Hu, W.L. Nelson, W. Turin, "Hidden Markov Model approach to online handwritten signature verification", IJDAR, 1(2), 1998, pp. 102-109

[5] L.L. Lee, T. Berger, E. Aviczer, "Reliable on-line signature verification systems", *IEEE Trans PAMI*, 18, June 1996, pp 643-647

[6] R.M. Guest, "Age stability in handwritten dynamic signature verification systems", accepted for publication in *Pattern Recognition Letters*, 2006.