The State of the Art in On-line Handwritten Signature Verification

G. K. Gupta
Faculty of Information Technology
Monash University, Building 75
Clayton, Victoria 3800, Australia

May 2006

Abstract

Using on-line handwritten signature verification (HSV) for authentication is attractive because of the long–standing tradition of its use in many common authentication tasks. HSV may be considered superior to many other biometric authentication techniques, for example fingerprints or retinal patterns, which are more reliable but also more intrusive. Furthermore, they require special and relatively expensive hardware to capture the image.

This paper presents a review of some on-line HSV techniques that have been reported in the literature. The paper also lists some commercial products that are currently available and discusses a number of applications of HSV.

1. Introduction

Our society is increasingly dependent on electronic storage and transmission of information and this has created a need to electronically verify a person's identity. Although handwritten signatures have often been used for identity verification, the situation has dramatically changed since the 9/11 terrorist attacks. Governments are now resorting to other biometric identity verification techniques, for example fingerprints and retinal patterns, which are more intrusive but are more reliable than handwritten signature verification (HSV). In spite of this trend there is considerable interest in authentication based on HSV because of the long–standing tradition of its use in a variety of authentication tasks.

Although there have been occasional disputes about the authorship of handwritten signatures [Osborn, 1929; Harrison, 1958; Hilton, 1956], verification of handwritten signatures has not been a major problem, since it appears that humans are generally very good at verifying genuine signatures.

[Miller 1994] and [Sherman 1992] discuss the importance of handwritten signature verification, also called signature dynamics, and note the availability of commercial products that use HSV (both note a product called Sign-On). Miller claims that more than 100 patents have been granted in the field of HSV (and dozens more have been granted since 1994), many of these however are for hardware to capture the signature. Miller and Sherman both note that although HSV is likely to become very important in the future, the technique will be widely accepted only if it is more reliable than the products that are currently on the market, in particular the technique needs to have lower false rejection rates (FRR or Type I Error).

The aims of authentication are likely to be different for different types of applications. For example, the primary concern of verification in a credit card environment (where the card holder presents a card to make a purchase and signs on an electronic device that automatically verifies the signature) must be to have a zero or near zero FRR so the genuine customer is not annoyed by unnecessary rejections. In this environment, fast verification is essential and, in addition, the information required for HSV should not require too much storage since it may need to be stored on a credit card strip or a smart card memory. A high level of security against forgeries may not be required and a false acceptance rate (FAR or Type II Error) of 10% might be acceptable since even that is likely to assist in reducing credit card fraud compared to the minimal checking that is done currently. On the other hand, in a security sensitive environment that was, for example, using HSV for granting an authenticated user access to sensitive information or other valuable resources, it would be necessary to have a high level of security against intruders and a zero or near zero FAR. An FRR of 10% would be a nuisance but might be acceptable. Of course, an ideal HSV system should have both the FRR and the FAR close to zero but no technique of HSV presently appears capable of performing consistently at this high level. It should be noted that FRR and FAR are closely related and an attempt to reduce one invariably increases the other.

A technique, which promises a small FRR when tested on an entire database, does not guarantee a small FRR for each individual. The performance figures reported in the literature are normally aggregate figures and it is common that some individuals have much larger error rates than the rest of the population in the test database. It is desirable that a technique not only have good aggregate performance but also good individual performance, but it has been found that some individuals have a very simple short signature which is relatively easy to forge or very large variation in their signatures, for example a person may have two different signatures, resulting in high individual FAR and/or FRR.

Most early work in HSV, in the early 1970s or before, focused on static (or off-line) HSV. Static HSV systems are those that require only an image of the signature. The advantage of these systems is that they do not need specialised hardware to capture signature information at the point of signing. Static HSV also has important areas of application, for example, in automatic cheque clearing; however there are disadvantages to the static approach. For example, static signature verification is sensitive to size and orientation of the signature. Also, static signature information is unlikely to be useful for storage on a credit card or smart card since generally significant storage is required to store a signature image. Also, since static techniques do not take advantage of the signature dynamics, the results generally are not as good as for on-line techniques. In this paper, the focus is on on-line HSV. For a more detailed discussion of static HSV, refer to the earlier surveys by [Plamondon and Lorette 1989] and [Leclerc and Plamondon 1994].

The present paper reviews some on-line HSV techniques that have been proposed in the literature with a focus on more recent developments. An attempt is made to describe important techniques and assess their performance based on published literature. The paper is organised as follows. Section 2 discusses the nature of handwritten signatures and how human signature experts verify handwritten signatures. A discussion of signature writing dynamics is presented in Section 3 which is followed in Section 4 by a discussion of some basic methodology of on-line HSV. Section 5 reviews some of the existing HSV literature using the parametric approach and Section 6 using the functional approach. Section 7 gives examples of some commercial products currently available and discusses how many of them are being used. Section 8 explains why it is so difficult to directly compare HSV systems and concludes the paper.

2. Nature of Handwritten Signatures

Handwritten signatures come in many different forms and there is a great deal of variability even in signatures of people who use the same language. Some people simply write their name while others may have signatures that are only vaguely related to their name and some signatures may be quite complex while others are simple. It is also interesting to note that the signature style of individuals relates to the environment in which the individual developed their signature. For example, people in the United States tend to use their names as their signature, whereas Europeans tend away from directly using their names.

It is well known that no two genuine signatures of a person are precisely the same and some signature experts note that if two signatures of the same person written on paper were identical they could be considered forgery by tracing. Successive signatures by the same person will differ, both globally and locally and may also differ in scale and orientation. In spite of these variations, it has been suggested that human experts are very good in identifying forgeries but perhaps not so good in verifying genuine signatures. For example, [Herbst and Liu 1977] cite references indicating that as high as 25% genuine signatures were either rejected or classified as no-opinion by trained document examiners, while no forgeries were accepted. Untrained personnel were found to accept up to 50% skilled forgeries.

[Osborn 1929] notes that handwriting shows great variation in speed and muscular dexterity. He claims that forgeries vary in perfection all the way from the clumsy effort which anyone can see is spurious, up to the finished work of the adept which no one can detect. Experience shows that the work of the forger is not usually well done and in many cases is very clumsy indeed. The process of forging a signature or simulating another person's writing, if it is to be successful, involves a double process requiring the forger to not only copy the features of the writing imitated but also to hide the writer's own personal writing characteristics. If the writing is free and rapid, it will almost certainly show, when carefully analysed, many of the characteristics of the natural writing of the writer, no matter what disguise may have been employed.

Osborn further notes that unusual conditions under which signatures are written may affect the signature. For example, hastily written, careless signatures, like those written in a delivery person's books, cannot always be used unless one has sample signatures that have been written under similar conditions. Furthermore, signatures written with a strange pen and in an unaccustomed place are likely to be different than the normal signatures of an individual. When a signature is being written to be used for comparison, this can also produce a self-conscious, unnatural signature.

Osborn discusses variations in signatures and notes that variations in handwriting are themselves habitual and this is clearly shown in any collection of genuine signatures produced at different times and under a great variety of conditions, which when carefully examined show running through them a marked, unmistakable individuality even in the manner in which the signatures vary as compared with one another. He presents as an example two sets of signatures (Figure 1) from a celebrated case of a contested will in New York. An estate worth more than six million dollars in the year 1900 was involved. The court accepted the will that had the five left hand side signatures as genuine in spite of significant variations in the signatures. For example, note the way the letters W and M have been written in the different signatures on the left compared with those on the right. The will with the four right hand side signatures was considered to have been signed by tracing and rejected as a forgery. Acceptance of the left hand side signatures led to the establishment of Rice University in Houston.

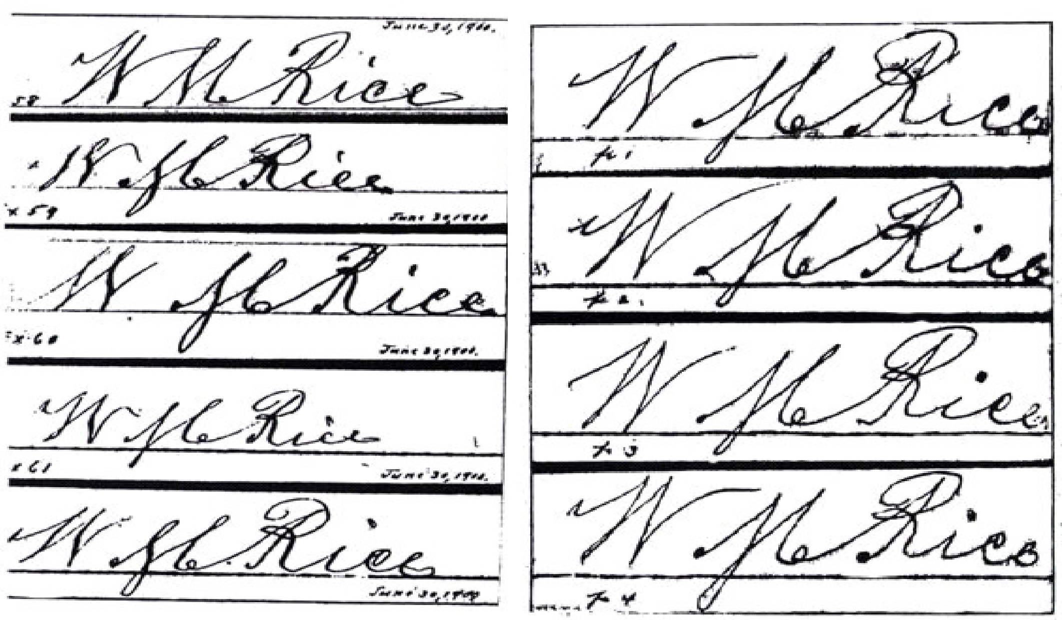

Figure 1. Genuine and Forged Signatures in a Disputed Will Case around 1900 [Osborn 1929]

To investigate signatures, Osborn recommends that several genuine signatures should always be obtained, if possible, with five signatures always a more satisfactory basis than one, and ten being better than five. To detect forgeries, Osborn gives a list of about 50 points that one needs to consider, including the following:

(a) Is the signature in a natural position?

(b) Does the signature touch other writing and was the signature written last?

(c) Is the signature shown in embossed form on the back of the sheet?

(d) Was the signature written before the paper was folded?

(e) Is the apparent age of the writing ink used consistent with the date of the document?

These points only show that HSV is far from trivial, but clearly most of these points cannot be applied to on-line HSV where a person's signature is collected on a graphics tablet rather than on paper.

[Hilton 1992] suggests that a signature has at least three attributes, form, movement and variation, and since signatures are produced by moving a pen on paper, movement perhaps is the most important part of a signature. The movement is produced by muscles of the fingers, hand, wrist, and for some writers the arm, and these muscles are controlled by nerve impulses. Once a person is used to signing his or her signature, these nerve impulses are controlled by the brain without any particular attention to detail.

Hilton further suggests that a person's signature does evolve over time and, with the vast majority of users, once the signature style has been established the modifications are usually slight. For users whose signatures have changed significantly over time, and such cases do occur although infrequently, the earlier version is almost always completely abandoned and the current version is the only one that is used. Only in some exceptional cases has it been found that a user may recall an old form of his or her signature, perhaps for signing special documents. [Liu, Herbst and Anthony 1979] found that in their experimentation with 248 users, three users continually varied between two signatures. This suggests that if an HSV system was to verify such exceptional cases of more than one signature by an individual, the system would need to maintain a list of reference signatures over time.

When a user's signature varies over time, should this variation be taken into account in HSV, assuming that the user might be using elements of a former signature in the current signature? Hilton comments that it is hard to answer this question since in the vast majority of cases the current signature is sufficient for verification purposes.

3. Dynamics of Signature Writing

The dynamics of the signature are captured by a graphics tablet in the data that is given by:

$$S(t) = [x(t), y(t), p(t)]^T \quad t = 0, 1, 2, \ldots, n$$

that is, it is a collection of x,y location values of the pen tip and the pen tip pressure values usually at equal time intervals. Some devices also capture azimuth and elevation. Many tablets sample at the rate of 200 times a second and the resolution of such devices is often 1000 pixels/inch although some have finer resolution. Typical American signatures are a writing of the person's name and therefore for American signatures the x-values typically grow linearly with time with small oscillations on the linear curve while the y-values show a more oscillatory variation with time, becoming positive and negative about the mean many times during a signature. An example is shown in Figure 2. Note the three pen-ups in the signature.

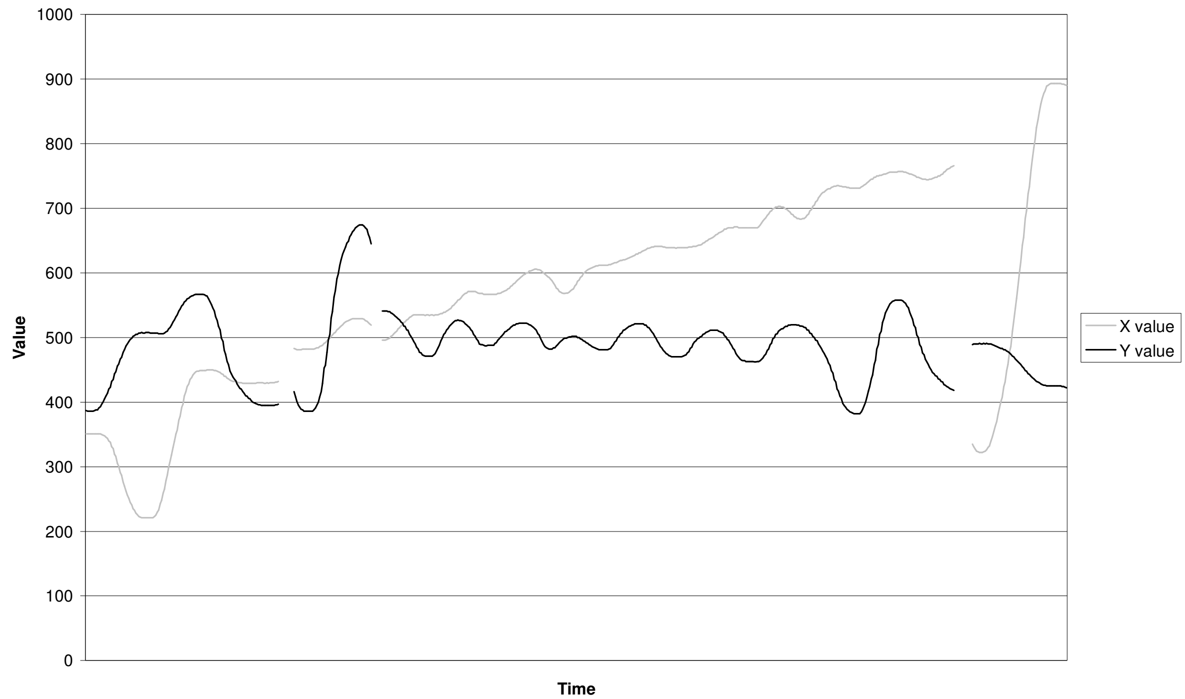

Figure 2. *x* and *y* Profiles of a Signature

The signature data, perhaps after appropriate smoothing, may be used to compute the derivatives of *x, y* and even *p* (assuming that the pressure data is not binary) if required. The first derivatives of *x* and *y* are the velocities in the two directions (which may be combined to compute total velocity, if required) and the second derivatives are the two accelerations. One may also wish to compute the third derivatives although these are not commonly used in HSV. The rate of change of acceleration is sometimes called *jerk*. Once these derivatives have been computed, the following signature data (assuming no pressure derivatives) is available[1]:

$$S(t) = [x(t), y(t), p(t), x_v(t), y_v(t), x_a(t), y_a(t), x_j(t), y_j(t)]^T \quad t = 0, 1, 2, \ldots, n$$

It is clear that every time a person signs his or her signature the number of samples obtained will be somewhat different. This variation in genuine signatures of the same individual makes it difficult to compare one set of values from one genuine signature with a set from another.

3.1 Signature Velocities and Accelerations

A signature may be considered a sequence of strokes. [Dimauro, Impedovo and Pirlo 1994] define a stroke as a sequence of fundamental components, delimited by abrupt interruptions. Let us consider one such stroke in the *y*-direction (assume that the stroke is from a small *y* value to a large *y* value) and for the moment ignore any displacement in the *x*-direction since that can be treated in a similar way. If the *y*-velocity during the stroke is now examined, it is found that the stroke is characterised by a number of positive velocity values that are growing and reaching a peak somewhere perhaps mid-way through the stroke and then a number of

[1] We will use the symbols X, Y, P, XV, YV, XA, YA, XJ, YJ to denote this signature data. In addition, we will use TV, TA and TJ to denote total velocity, total acceleration and total jerk at each point respectively and AV, AA and AJ for averages of the signature velocity, acceleration and jerk.

positive values that are declining. This is clear if we look at the first major *y*-stroke that looks like an 'S' in Figure 2 and the corresponding y-velocity profile in Figure 4. Therefore the velocity profile of a stroke is a waveform, perhaps it looks like a bell-shape curve although the one in Figure 4 does not, starting with a zero velocity, reaching a peak and ending with a zero velocity. Furthermore, if the variation of the *y*-acceleration corresponding to a single peak velocity profile of a stroke is studied, it is found that the accelerations during the period of the single peak velocity profile will grow from zero value to reach a peak value perhaps mid-way during the velocity climb-up to the peak and then decline steadily to zero at the peak velocity. On the way down, the accelerations will grow in magnitude in the negative direction (as velocity reduces) and reach a negative peak around half-way down the curve and then decline steadily in magnitude (although still negative) to reach zero at the end of the curve. The *y*-acceleration in Figure 6 corresponding to the y-stroke in Figure 2 unfortunately does not look anything like what is expected due to jitter but it does show a sharp rise in positive acceleration followed by a sharp rise in negative acceleration. A stroke therefore will look like a single positive peak curve when the velocity profile is examined, and will look like a two peaks curve, one peak in the positive direction and another in the negative direction, when the acceleration profile is examined.

Theoretically the velocity profiles should therefore have double the peaks and valleys (call them extrema) than the *x* and *y* profiles would and the acceleration profiles would have four times. As we will see in practice it is not always so because of jitter. Furthermore the heights of the extrema of the velocity profiles will be smaller than the length of the corresponding strokes and the heights of the extrema of the accelerations will be smaller still. This is clearly illustrated by the velocity and acceleration profiles in Figures 4 and 6 which show that the majority of velocity values are between –5 and 5 while most acceleration values are between –1 and 1. These profiles are for the signature in Figure 2.

The shape of the velocity profile of a stroke has been studied by [Plamondon 1993] who states that the shape of velocity profile obtained when a rapid-aimed stroke (like those in a signature) is written, is approximately bell-shaped but the shape is asymmetric (although it is not so clearly evident in the velocity profiles in Figures 4 and 6) and it is claimed that this shape is almost preserved for movements that vary in duration, distance or peak velocity. It may be that this consistent shape is related to the way the central nervous system plans and controls movement. It is suggested that these profiles can be explained by log-normal curves. This is related to the Fitts' Law, a discussion of which can be found in [Fitts 1954].

Given that the position values are only accurate to one pixel, the velocity and acceleration approximations can have significant errors. For example, the velocity profile in Figure 3 is obtained by using the simple forward difference formula while the profile in Figure 4 is obtained using central differences. Much of the jitter, but not all, in Figure 3 is eliminated by the use of a more accurate formula in Figure 4. The differences are even more pronounced when we compare the two acceleration profiles in Figures 5 and 6. The large differences in the maximum values in Figures 5 and 6 should be noted. Jitter is still quite significant in the acceleration values in Figure 6.

A rough estimate for the height of the velocity profile may be obtained if the single peak curve is approximated with a triangle since the area under the curve is equal to the length of the stroke. Let the length of the stroke be d then the peak velocity v may be approximated by $v = 2d/t$ if t is the time taken to write the stroke. Clearly, high velocity peaks correspond to long strokes in the signature as illustrated by the velocity profiles in Figure 4.

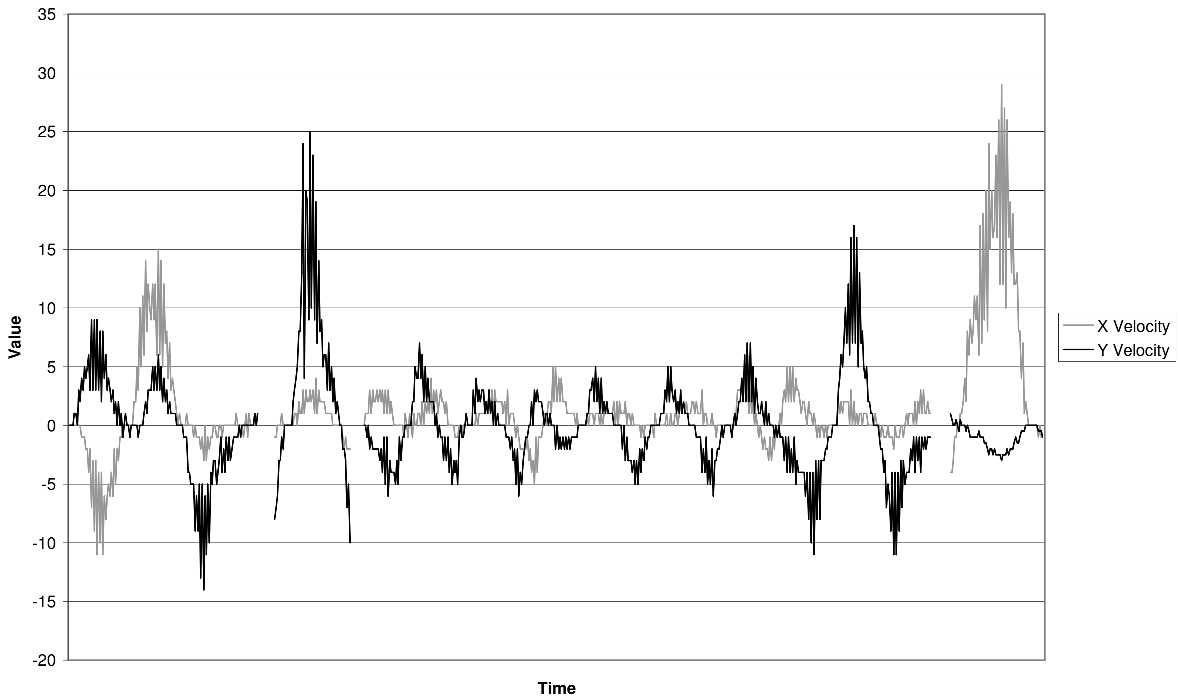

Figure 3. x and y Velocities of the Signature in Figure 2 using Forward Differences (y velocities are in darker colour)

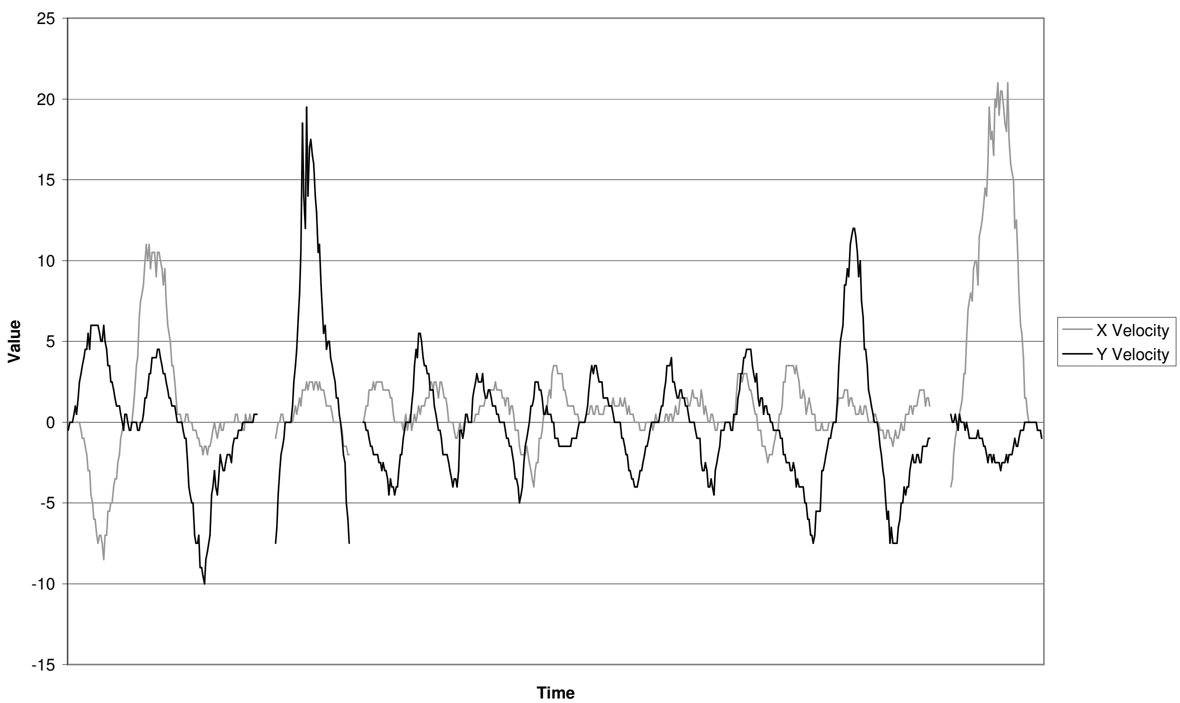

Figure 4. x and y Velocities of the Signature in Figure 2 using Central Differences (y velocities are in darker colour)

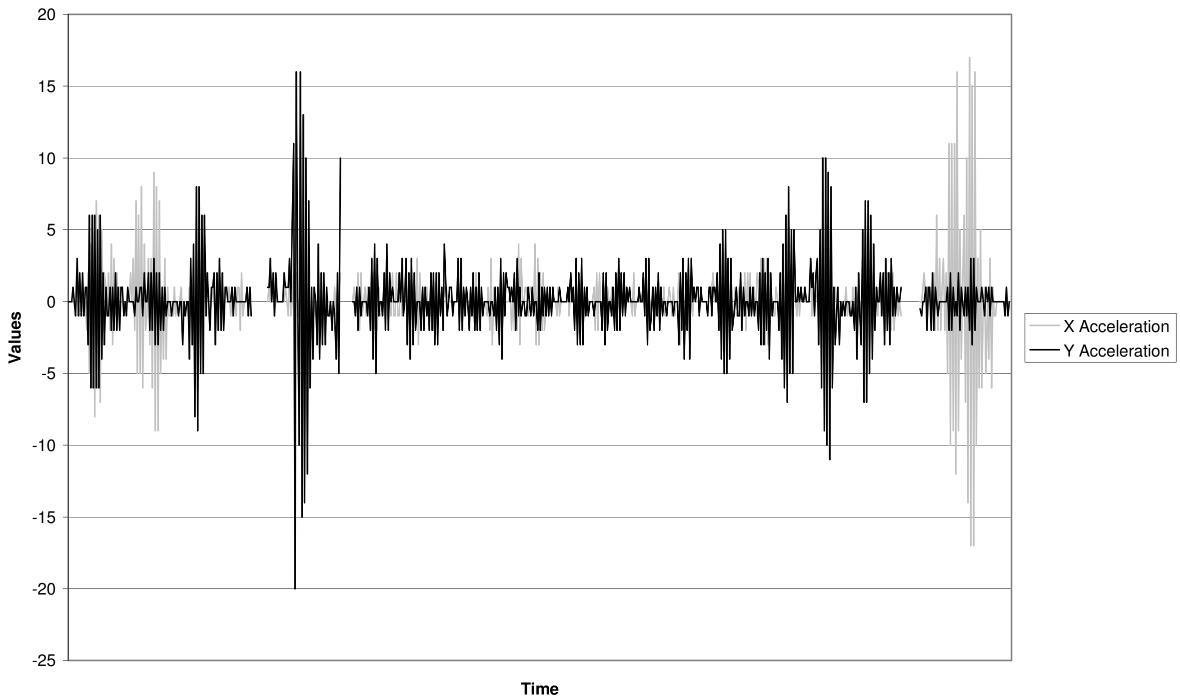

Figure 5. x and y Accelerations of the same Signature using Forward Differences (y accelerations are in darker colour)

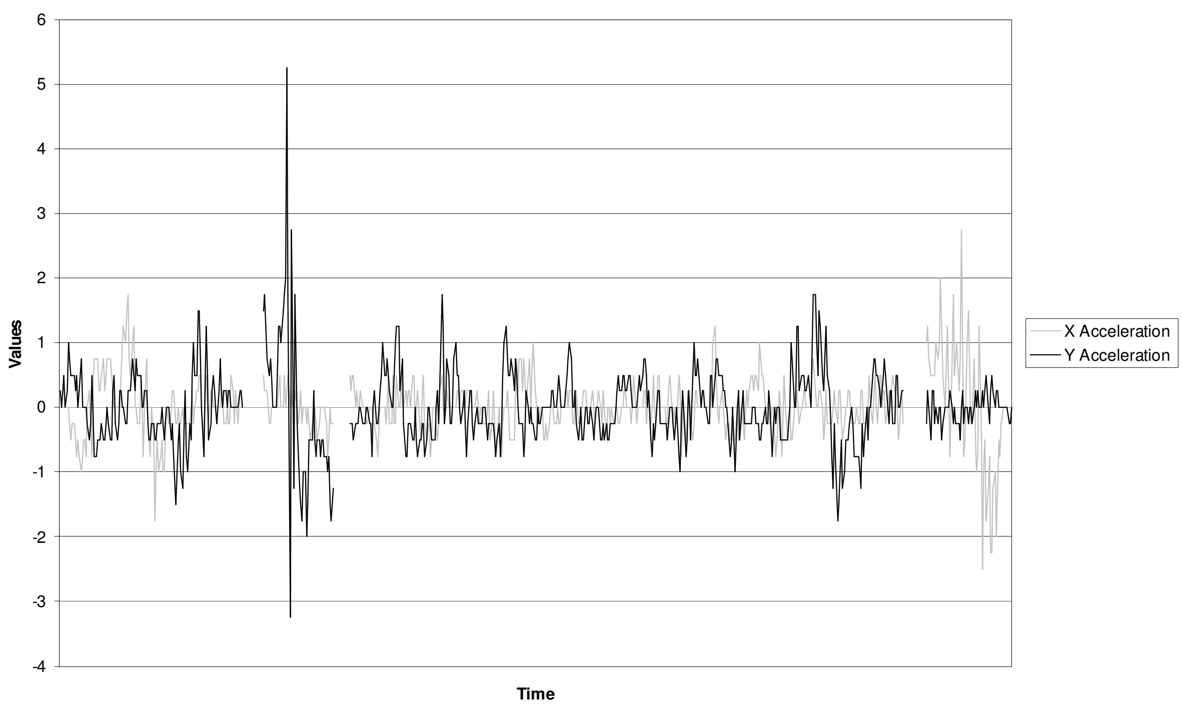

Figure 6. x and y Accelerations of the same Signature using Central Differences (y accelerations are in darker colour)

3.2 Example of a Forgery

Given the dynamics of signature writing, how does a forger forge a signature and how difficult is it for a forger to produce a good forgery? As noted earlier, it is believed that a forger cannot write another person's signature in a ballistic motion without a lot of practice and therefore producing a good forgery is never easy. Figures 7, 8 and 9 show position, velocity and acceleration profiles of a skilled forgery of the signature in Figure 2. These figures show that the profiles of this forgery are very different than those of the genuine signature in Figures 2, 4 and 6. The most obvious difference is the long pen-up time of the forgery. Also, most velocities and acceleration values for the forgery in this case are smaller than those for the genuine signature perhaps because the forger was being careful in forging the signature.

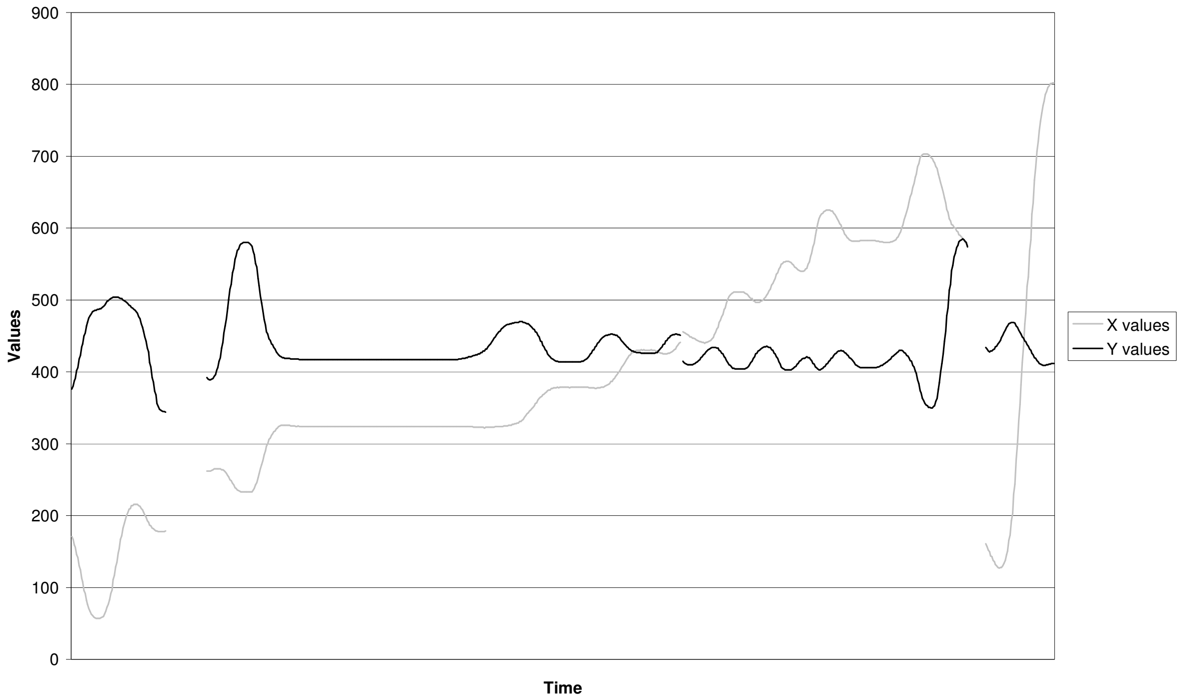

Figure 7. *x* and *y* Profiles of a Forgery for the Signature in Figure 2
(*y* velocities are in darker colour)

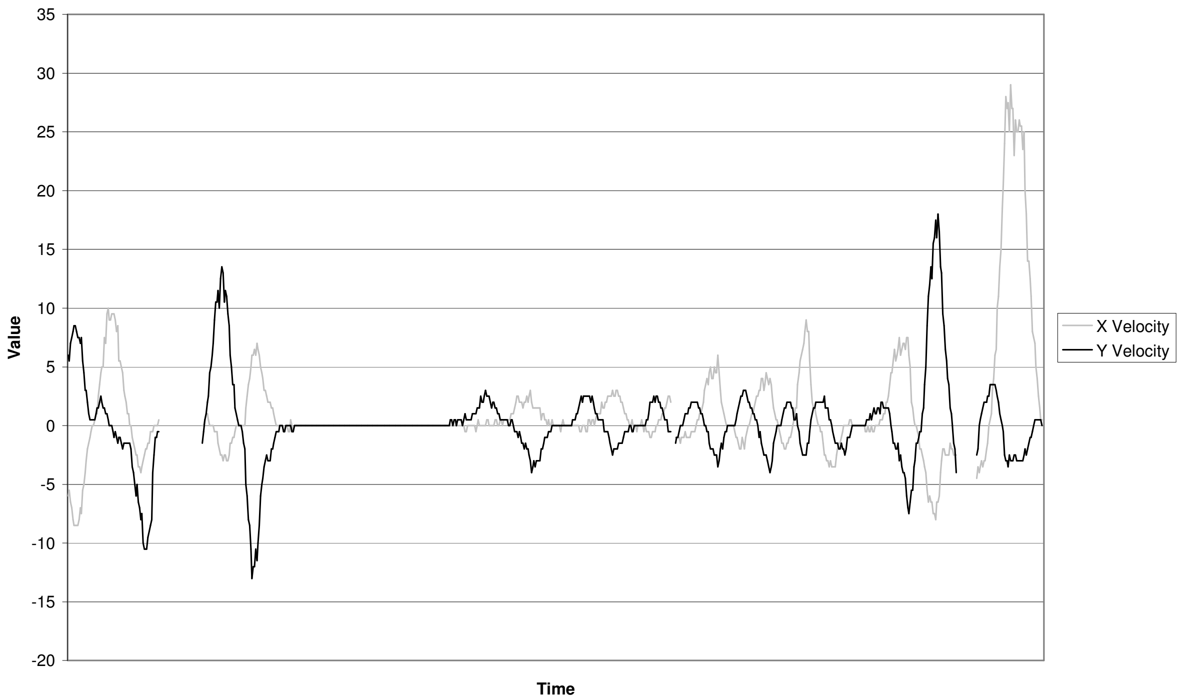

Figure 7. x and y Velocities of a Forgery for the Signature in Figure 2 using Central Differences (y velocities are in darker colour)

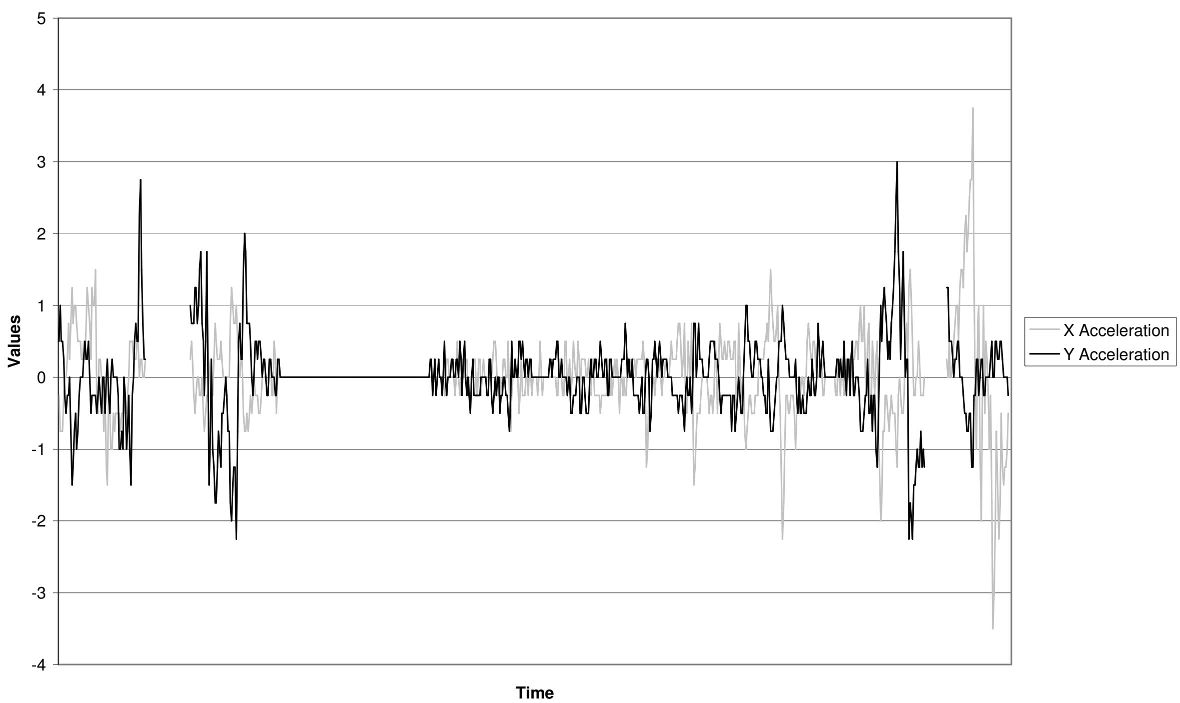

Figure 8. x and y Accelerations of the same Forgery using Central Differences (y accelerations are in darker colour)

As noted earlier, some signatures lend themselves more easily to forgery than do others. It is this complexity of forging of a signature that is studied by [Brault and Plamondon 1989, 1993] in an attempt to estimate the intrinsic risk of a signature being forged. They note that humans can only remember about seven variations of a pattern without error and so a forger cannot memorise all the details of a signature that is being forged. A process of minimisation or recoding of information that needs to be remembered takes place. For example, a forger might recode information about the signature of John Smith by first remembering the name and then remembering that John Smith's signature has a taller J than the forger would normally write himself and a rounder S and so on. Based on these observations, they derive an expression for an imitation difficulty coefficient of a signature. Roughly, the expression gives the difficulty of a given signature to be a function of the variation rates in length and direction of the strokes. It is claimed that people with small values of this coefficient and high variability between their signatures are the problematic signers.

4. Basic HSV Methodology

Most HSV techniques use the following procedure for performance evaluation:

1. Registration – Obtain a number of signatures for each individual at enrolment or registration time (these signatures are called *sample signatures* although other terms like *training signatures* have also been used).

2. Pre-processing and Building Reference Signature(s) – pre-process the sample signatures if required, compute the features required, and produce one or more reference signatures. Decide on what threshold will be used in verifying a signature.
3. Test Signature – when a user wishes to be authenticated, he/she presents the identification of the individual he/she claims to be and presents a signature (we call this signature the *test signature*). Pre-process as in Step 2 and compute the features of the signature.
4. Comparison Processing – the test signature is compared with the reference signature(s) based on the features' values and the difference between the two is then computed using one of the many existing (or specially developed) distance measures.
5. Performance Evaluation – for each signature that claims to be a genuine signature, compare the distance computed with the threshold decided in Step 2 above. If the distance is smaller, accept the signature, otherwise reject.
6. Repeat Steps 3-5 for the given set of genuine signatures and forgery attempts for this individual and Steps 1-6 for a new individual; compute the FRR, the skilled FAR and the random FAR.

Obtaining good estimates of FAR is very difficult, since actual forgeries are impossible to obtain. Performance evaluations therefore rely on two (some have defined more) types of forged signatures. A forgery may be *skilled* if it is produced when the forger has had access to one or more genuine signatures for viewing and/or practice. A forgery is called *random* when either another person's genuine signature is used as a forgery or the forger has no access to the genuine signature and is either only given the name of the person whose signature is to be forged or just asked to sign any signature without even knowing the name. Random forgeries generally lead to much smaller FAR than skilled forgeries.

Although performance evaluation as described above is essential, the evaluation is not always a true indicator of the performance of the technique since the signature databases used for testing often do not adequately represent the population at large. There is a great deal of variability in signatures according to country, age, time, habits, psychological or mental state, and physical and practical situations. Building a test database of signatures that is representative of real-world applications is a difficult task, since it is difficult enough to find people who will willingly sign 10 or 20 times. People are not always happy to have their signatures stored in a computer or given to others to practise forging them. Therefore most test databases are built using signatures from volunteers from the research laboratory where the HSV research is being carried out and as a result most test databases have very few signatures from people that are old, disabled, suffering from arthritis, or poor. Percentages of such people in the population are significant and these are the people whose signatures are likely to pose the greatest challenge for a HSV technique.

It has been reported that FAR and FRR are generally higher when the systems are used by a more representative group. The higher FRR by a more representative group is not surprising since the signing environment generally is less consistent than when signatures for a test database are collected. Furthermore, more signatures in this wider group are likely to be obtained from people whose signatures are variable either due to background or age compared to the relatively young research staff, students and professors whose signatures may comprise a test database. The reasons for a higher FAR in this group on the other hand are not clear. A more representative group could even lead to lower FAR than a test database since most test databases include skilled forgeries obtained after the forgers were allowed to practise their forgeries, which may not always be possible in the real world.

A number of other difficulties face a person who attempts to compare the results of many studies reported in the literature. Although there is now a public signature test database that could be used by all researchers for performance evaluation and comparison [Ortega-Garcia, Fierrez-Aguilar, Simon, *et al.* 2003], most researchers have their own test signature databases with a varying number of genuine signatures. Some have skilled forgeries while others do not. Some have screened the signature database to remove a number of signatures that for some reason were not acceptable while others have done no screening. The number of signatures used in building a reference signature often varies. Different tests and thresholds may be used. Some studies use a different threshold for each individual, while others use the same threshold for all. Clearly, state of the art in HSV would benefit if all research used the same testing protocols.

However, a first step towards overcoming these problems was taken in 2004 when the first International Signature Verification Competition (SVC2004) for on-line signatures was organised in Hong Kong. SVC2004 used two datasets of 100 signature sets each for testing. Each signature set consisted of 20 genuine signatures and 20 skilled forgeries from at least four forgers. The first 40 sets of each dataset were released to the participants before the competition. The first dataset only had information about pen coordinates while the second dataset also included additional information about pen orientation and pressure. The best system was found to have equal error rates (FRR = FAR) of between 2.5% and 3.0%. Several methods had equal error rate (EER) below 10%, while several others had EER above that.

Given that all on-line HSV techniques use signature features of one type or another, we define some terminology about them. A feature is called *global* if the feature is extracted from the whole signature (for example, signature path length, total time, average or RMS signature velocities, average or RMS signature accelerations and pen-up time). A feature is called *local* if it is extracted for each sample point (for example, position, velocity and acceleration values). Sometimes, *segmental* features are used when a signature is divided into segments and the feature is obtained for the whole segment. Based on these different features, there are at least two different HSV approaches commonly used although hybrid approaches combining both have also been proposed.

In the first approach, called the *parametric approach,* a set of values of global features or parameters are computed and compared. In this approach, once a set of features has been selected, only the values of the features from the reference signature(s) need to be stored. Also, when a test signature is presented, only the features' values are needed, not the signature. This often saves on storage which may be at a premium if, for example, the reference signature needs to be stored on a card, and that is why representing a signature by a set of features' values is sometimes called *compression* of the signature.

In the second approach, called the *functional approach*, all of the values of selected local features of a signature are assumed important and the test and reference signatures are compared point-to-point using sets of these values. In this approach much more information about a signature is often necessary. For example, given that a typical signature, if sampled 200 times a second, could produce 1000 samples or more, 25 features at each point would lead to 25,000 feature values for one signature which need to be compared with a similar number of values from another. The processing time then may become an issue. How the information about the reference signature(s) is to be stored also may be problematic if the system is to be used for a large number of individuals. Such large amounts of information may be difficult to store on a card.

4.1 Major Issues in the Parametric Approach

In the parametric approach, some major research issues are:

- How many global features are sufficient?
- Do the signatures require some transformation before computing the features, for example resizing, deslanting or smoothing?
- How many sample signatures will be used in computing the reference signature?
- Would the reference signatures be updated regularly? If yes, how would this be done?
- Should individual sets of features be used rather than the same features for all? Should individual thresholds be used rather than the same for all?
- How is the distance between a test signature and the reference signature going to be computed?

Regarding the number of features, there is often a temptation to include more and more features in a method in the hope of improving performance and some researchers have proposed as many as 100 global features. We believe that using a large number of features does not always lead to better performance and may create some difficulties. For example, if a method uses many features, the storage needed to store the values of those features for the reference signature is going to be relatively large and a credit card may lack sufficient capacity to store all the values. Also, when a test signature is compared to the reference signature, given that no two genuine signatures are identical, it is unlikely that a genuine signature will have all features' values close to the values for the reference signature. To ensure that all (or almost all) genuine test signatures are authenticated, a technique using a large number of features either must have a large threshold for the norm of the distance or use some criterion similar to the *majority classifier* used by [Lee 1992]. This majority classifier is not particularly satisfactory since it cannot be easily analysed theoretically. Also using it is in effect stating that although a large number of features are considered important, it is acceptable to ignore several of them in comparing the test signature to the reference signature. This appears contradictory.

A number of studies have investigated global features and we consider some of them. A large number of features were studied by [Crane and Ostrem 1983] who used an instrumented pen to sample three forces of the writing tip (viz. downward force and the *x* and *y* forces) and then initially computed 44 global features, some of them are given in Table 1. The method used simply removed one feature in turn from the current set of features (initially 44) and found the feature whose removal gave the lowest EER. The method continues until removing a feature does not reduce EER. 25 features were selected. To evaluate the proposed technique, a database of 5220 genuine signatures from 58 subjects was collected over a four-month period and 648 skilled forgeries from 12 forgers that were allowed to practise the signatures to be forged. EER as low as 1.5% was obtained although about half the genuine signatures were used for selecting the subset of features while the other half for testing.

[Lee 1992] and [Lee *et al.* 1996] describe a set of 42 features; many of them are listed in Table 1. A number of techniques for feature selection were used. In one of the techniques that uses only genuine signatures, for each feature, the mean for subject *i* is compared with the means of the same feature for all other subjects and the maximum of such distance is computed for each subject for each feature. Importance of each feature for an individual is then given by this maximum distance. This algorithm is used to find several subsets for each user; 34-features performed better than 42-features. When forgery data was available, distance was computed between the i^{th} feature of the subject and the corresponding feature of the forgeries for that individual and features with the largest distances were selected. It was shown that 23 or 24 features gave the best performance. Ten of them are listed in Table 1.

[Ketabdar *et al.* 2005] present a methodology for selecting the most discriminative global features. More than 150 global features from 60 papers were considered and an initial subset of 46, some of them listed in Table 1, was investigated. To perform feature selection, a near-optimal feature space search algorithm that avoids the exhaustive search was used. A cost function based on within-class variability being small and between–class variability being large was used. In order to take into account effects of correlation between feature vector components, the cost is computed on the whole feature vectors instead of individual features. The results of applying the method to a 25-users subset (25 genuine and 25 forgeries for each) of the MYCT database resulted in 12 features that are listed in Table 1.

[Fierrez-Aguilar, Nanni et al. 2005] evaluated 100 global features to find their discriminative power. This was done by computing the Mahalanobis distance between the mean of the training signatures of a person and the set of all training signatures from all users. Features were then ranked according to inter-user class separability. The top ten discriminative features are listed in Table 1.

A number of others researchers have studied global features. For example, [Richiardi *et al.* 2005] studied 46 global features and used MYCT database to select 12 best features including AV, AP, TT, pen down samples, DPVX, average and maximum pressure.

We summarise these studies in Table 1[2]. The summary shows that total time and average velocity are perhaps the best global features of a signature.

---

[2] Some more notation for global features is now presented: total signature time is TT, Maximum and minimum velocity are MV and NV (and in *x* and *y* directions MVX, MVY, NVX, NVY), number of pen-ups is NP, duration of positive velocity in *x* and *y* (DPVX and DPVY), duration of negative velocity in *x* and *y* (DNVX and DNVY), averages of positive and negative velocities in *x* and *y* (APVX, APVY, ANVX, ANVY), number of velocity zero crossings in *x* and *y* (NVZX, NVZY).

Reference	Features Studied	Best Features
[Crane and Ostrem 1983] 44 features (20 listed in the next column). 25 best features were selected, listed in the last column.	TT, For x, y and P features: scaled mean, SD, min, max, mean absolute, mean positive, numbers of positive and negative samples, number of zero-crossings, NP+1	TT, time up, time down, For x, y features: scaled mean, SD, min, mean positive, SD of P, max y, mean absolute P, mean positive P, numbers of negative P samples, number of zero crossings of y and P, max − scaled mean y and P, scaled mean − min y, NP, NP+1
[Lee *et al.*1996] 42 features (20 listed in the next column). Best 10 selected and listed in the last column.	AV, MV, TT, NVX, DPVX, DPVY, APVX, APVY, ANVX, ANVY, MVX-AVX, MVY− AVY, MVX− NVX, MVX− NVY, MVY− NVY, SD of X and Y, initial direction of first and second components, time at which MV occurs.	AV, TT, NVX, APVX, ANVX, APVY, ANVY, MVY− AVY, MVY− NVY, total pen down time
[Ketabdar *et al.* 2005] 46 features (20 listed in the next column). Top 12 best selected and listed in the last column.	TT, height, width and their ratio, time to width ratio, AV, MV and their ratio, average and SD of XV, SD of velocity and acceleration, MYV, AYA, AXA, DPVX, NP, average elevation and azimuth, AP	TT, DPVX, AV and average elevation, AV/MV, average P, SD of P, Max P, point of Max P, average elevation, DPVY/NP, number of tangent angles in the first half quadrant
[Fierrez-Aguilar, Nanni *et al.* 2005] 100 features (20 listed in the next column). Top 10 listed in the last column.	TT, NP, NVZX + NVZY, AJ, AJX, SDs of *x*, *y* acceleration and *x*, *y* velocity, AV/Max V, RMS V/max V, AV/MVX, AV/MVY, number of local maxima in *x* and *y*, RMS V/max V, RMS J, pen-down time/ TT	TT, NP, NVZX + NVZY, AJ, SD of AY, AX, VX and VY, SD of Y/(max Y – min Y), Number of local maxima in X

Table 1. Summary of Studies for Selecting Best Global Features

Is it necessary to transform a test signature before the features are computed? A number of transformational techniques have been proposed, the simplest only suggest smoothing the data but others suggest, for example, finding the bounding rectangle and scaling the test signature to the same size as the reference signature. While such transformations have been used, no systematic study of the benefits of using them has been reported. An example of a transformational technique is that of [Phelps 1982] who used a space-domain approach in which a close–fitting polygon was formed around the signature image. The signature area was then normalised and centred on a coordinate plane. Phelps showed that the area of overlap for a pair of valid signatures was consistently higher than for forgeries. [Kashi et al. 1997] use a cleaning process which included removal of irregular and redundant points, cusps were detected and marked, cubic B-spline smoothing was applied and the signature was then resampled at intervals of equal arc length.

Regarding the number of sample signatures, it was suggested earlier that at least five signatures are needed for satisfactory performance of a parametric technique. The reference signature is based on a set of sample signatures and for each element of the set of selected features the mean and standard deviation (SD) of the feature values has to be estimated. Clearly, to obtain good estimates of the mean and SDs of the features' values for the genuine signatures population, it is necessary to have several sample signatures. Experimentation by [Gupta and Joyce 1997a], [Nalwa 1997] and [Fierrez-Aguilar 2005] using different number of sample signatures show that performance improves with the number of sample signatures used and five or six sample signatures lead to acceptable performance without requiring too many sample signatures.

[Parks, Carr and Cox 1985] suggest that at least six sample signatures be used. They further note that if the six sample signatures are gathered under identical conditions, the SDs of the features might be too small to be an accurate estimate of the SDs of the person's signatures and a method for either increasing the SDs or in some cases discarding the previously obtained sample signatures was suggested. In extreme cases, the whole attempt to obtain a reference signature was aborted and a new set of sample signatures obtained on another occasion!

It is believed that a reference signature should be updated regularly, as the user's signature evolves over time. [Crane and Ostrem 1983] modified the reference signature when a signature was successfully verified by adding the new signature vector to the mean reference vector with a weight of 1/8. Adding such a weighting function does not have any effect on users whose signatures do not change over time and should have a positive effect on users whose signatures do. [Parks, Carr and Cox 1985] also suggest that the reference signature should be updated every time a signature is verified by applying a weighting of 10% to the new verified signature. Some HSV products listed in Section 7.1 also update reference signatures.

A number of researchers, for example [Fairhurst and Brittan 1994], advocate use of individual sets of features and individual thresholds. A careful study to evaluate the benefits has not been carried out. Problems arise because finding a set of individual features and individual thresholds without having access to a large number of training signatures of each individual is difficult. It may be that one could use the (say) five sample signatures and based on the mean and SD of each feature's values make a decision about which features and what threshold could be used for each individual. [Crane and Ostrem 1983] explore the possibility of using personalised feature sets for each person and show that personalised feature sets can improve the performance of a HSV system. [Parks, Carr and Cox 1985] suggest using different thresholds for different individuals and the possibility of basing the threshold in credit card verifications on the value of the merchandise being bought and the credit rating of the person. [Lee 1992] found that the performance of the best 10 common features was significantly worse than that of 10 individual features for each subject.

4.2 Computing Distance

Many different distance metrics to compute the distance between a test signature and the corresponding reference signature are available. We consider the following four approaches:

1. Euclidean and weighted Euclidean distance
2. Mahalanobis distance
3. Dynamic programming matching

4. Majority classifier

Euclidean and Weighted Euclidean Distance (ED and WED). Let *T* be the test signature with the feature set ($t_1, t_2, \ldots, t_n$) and let *R* be the reference signature denoted by two vectors, vector of means ($r_1, r_2, \ldots, r_n$) and standard deviations ($s_1, s_2, \ldots, s_n$) . Euclidean distance between T and the reference signature is then given by:

$$G(T) = \sqrt{\sum_{i=1}^{n} (t_i - r_i)^2}$$

It follows that by accepting all test signatures which are within a distance defined by some threshold; we are accepting all signatures within a spheroid around the reference signature. All features in this metric contribute equally. The major difficulty with this distance measure is that a test signature feature value that is far away from the corresponding reference signature value can have a large influence on the value of the distance. Also, the distance grows with the number of features.

A weighted or normalised Euclidean distance overcomes these difficulties by scaling the differences between the features' values by the standard deviation as well as scaling the root of the sum by the number of features as follows:

$$G(T) = \left(\frac{1}{n}\right)\sqrt{\sum_{i=1}^{n}\left(\frac{t_i - r_i}{s_i}\right)^2}$$

Distance measures based on the Euclidean metric are used quite widely, for example see [Crane and Ostrem 1983] and [Nelson, Turin and Hastie 1994].

Mahalanobis Distance (MD). MD differs from the Euclidean measure in that it takes into account the correlations of the data set and is scale-invariant. The distance is defined by the following expression if S is the covariance matrix of the features:

$$G(T) = \sqrt{(T - R)^T S^{-1} (T - R)}$$

Dynamic programming matching (DPM). DPM involves minimising the residual error between two functions by finding a warping function to rescale the time axis of one of the original functions. DPM is further explained by Lee and is investigated by a number of others, for example,]Parizeau and Plamondon 1990]. This measure is perhaps more suitable when the functional approach is used.

Majority classifier (MC). When many features are used, a single feature can unduly influence the decision through deviating far from the mean value, even if the other features have values close to their means for the genuine reference set. One way of alleviating this problem is to use MC, which is based on the "majority rules" principle. That is, it declares the signature being tested to be genuine if the number of feature values which pass a predetermined test is larger than half the total number of tested features. Using MC, [Lee 1992] achieved an EER of only 3.8% using 42 features. How the majority classifier performs on zero-effort forgeries is not clear. Nelson et al. also test the majority voting model and found it to give results that were worse than the Euclidean distance.

4.3 Some Theoretical Aspects of HSV

The theoretical basis of HSV using the feature-based approach is now briefly considered. This discussion follows the standard multivariate normal distribution approach which is discussed in most books on multivariate analysis (for example, refer to [Jobson 1992], Section 7.2). It is assumed that an infinitely large population of genuine signatures is available for an individual and that the signatures are represented by the values of their features. Let there be m features which are not always independent. Let f be a random variable which is a vector of feature values representing the genuine signature population. f is assumed to be normally distributed with mean vector μ and covariance matrix V.

Normally, of course, the mean and covariance matrix of the genuine signature population are not known and so the mean vector μ is replaced by the mean reference signature vector R and the covariance matrix V is replaced by a diagonal matrix that has the squares of the reference SDs (that is, S) and all correlations between the features are ignored. As discussed in Section 4, resulting simplified procedure therefore works as follows. The vector D is normalised by dividing each value in it by the corresponding SD in the vector S to obtain vector Z whose norm is then computed. The computed norm is now compared to a predefined threshold and the signature is authenticated only if the norm is smaller than the threshold.

Given that the population mean and the covariance matrix of the features' values is unknown and the correlations between the features have been ignored, the norm of the vector Z cannot be expected to closely follow the Chi-squared distribution. If the distribution were Chi-square, it would be possible to make accurate predictions for the threshold given the desired value for FRR. A simple example is now presented to show that larger threshold values need to be used in practice to achieve the same FRR as those which are predicted by the Chi-square distribution. Consider a single feature situation. An FRR of 1% corresponds to a Chi-square value of 6.63. Now if six features are used that have correlation coefficients of 1.0 between them, it is essentially using the single feature six times and therefore the correlated six features correspond to six times the value 6.63, that is 39.78 which is much higher than 16.8 if there was no correlation.

Of course, no prediction can be made about the expected FAR, since the forgeries would not have a normal distribution and there is no way to obtain even an estimate of the mean and covariance matrix of the forgeries.

4.4 Major Issues in the Functional Approach

In the functional approach, the following major research issues arise:

- How many local features should be used?
- How is the comparison to be carried out given that genuine signatures of an individual tend to have some variation amongst them?
- If the signature is divided into segments, how would segmentation take place?
- How would the distance between the test signature and the reference signature(s) be computed?

The question of how many local features are to be used is a difficult one. Some researchers use only a small number of features while others use up to 25 or even 50 local features. A number of studies have been reported that investigate which local features are important. We discuss some of them. [Plamondon and Parizeau 1988] compare the performance of three types of data: position, velocity and acceleration. To test the features, they used a database of 50 signatures each from 39 volunteers. On the average, 48 signatures remained for each signer after visual inspection of the data resulted in some signatures being rejected. Scaling was done before computing the velocity and acceleration values; the impact of scaling on these values is not clear. FAR with random forgeries varied from an average of 1.9% to 8.1%. The use of velocity data gave better results than displacement, which was better than acceleration.

The features studied by [Yanikoglu and Kholmatov 2003] and [Kholmatov and Yanikoglu 2004] are given in Table 2. The lowest error rates were obtained for the coordinate differences. They do not note that coordinate differences are essentially forward difference approximations to the velocities. [Lei and Govindaraju 2004] evaluate 22 commonly used local features, listed in Table 2. The results indicate that pressure, azimuth and altitude sequences do not have high consistency. The best local features were found to be the two velocities. Coordinate differences (these are approximate velocities, as noted earlier) and angle sequences also had high consistency. [Richiardi *et al.* 2005] select a set of 12 local features from an initial set of 39 consisting of 13 base local features and their first and second derivatives. Thus there are a number of duplicates in the set. Most of the 12 features were from the 13 base features, only two first derivatives (*x* velocity which is already in the base features and first derivative of acceleration was selected.

We summarise these studies in Table 2[3].

Reference	Features Considered	Best Features
[Plamondon and Parizeau 1988]	Position, Velocity and Acceleration	Velocity
[Yanikoglu and Kholmatov 2003] and [Kholmatov and Yanikoglu 2004]	Position relative to the first point, Δx, Δy Δc	Δx, Δy
[Lei and Govindaraju 2004] 22 features including four global features are listed in the next column. Eight best listed in the last column.	X, Y, [X, Y], V, XV, YV, P, A, EL, AZ, AV, centre of mass (*x*, *y*), torque, two curvature ellipses, AV, APVX, APVY, TT, curvature, cos(θ) and sin(θ), θ is the angle between the speed vector and the X-axis.	V, XV, YV, cos(θ) and sin(θ), X, Y, [X,Y]
[Richiardi et al. 2005] 39 features listed in the next column.12 best listed in the last column.	X, Y, P, AZ, EL, XV, YV, TV, XA, YA, TV, path tangent angle, log radius of curvature. First and second derivatives of all of the above 13.	X, Y, P, path tangent angle, TV, XV, TA, log radius of curvator, AZ, EL, YV, TJ

Table 2. Summary of Four Studies for Selecting Best Local Features

[3] Δx, Δy and Δc denote *x*, *y* coordinate differences and curvature differences between two consecutive points. AZ and EL denote pen azimuth and elevation.

To compare a test signature to its reference signature, one simple approach might be to compute correlations between the test signature values and the reference signature values. Such point-to-point comparison usually does not work well since some portions of any two genuine signatures of the same person can vary significantly and the correlation may also be seriously affected by translation, rotation or scaling of the signature. Another approach might be to segment the signatures being compared and then compare the corresponding segments using some alignment of segments if necessary. This appears to work better. Two other approaches are now being used successfully, namely dynamic time warping (DTW) and hidden Markov models (HMM). DTW performs flexible matching of local features of a test signature and its reference signature. HMM on the other hand performs stochastic matching of the two signatures using a sequence of probability distributions of the features along the signature.

[Parizeau and Plamondon 1990] consider three types of techniques for comparing a test signature with the reference signature: regional correlation, dynamic time warping and skeletal tree matching. Tree matching is a technique that involves building a tree of peaks and valleys of a signature. Once the trees for the test and reference signatures are available, the distance between them may be computed in terms of minimum number of operations needed to transform one tree to the other. It was found that no technique was globally superior to the other two although regional correlation was often better and much faster.

If the signature is to be divided into segments, a criterion for segmentation is required. It may be that any point at which the y velocity becomes zero is chosen as the end of one segment and the start of another. An alternate approach might be to identify extreme points (for example, maxima and minima of the signature's x and y values) and use them for segmenting the signature. [Plamondon and Parizeau 1988] suggest that segments be no longer than 0.7 seconds apiece as handwriting signals tend to fall out of phase beyond 0.7 seconds. [Hastie *et al.* 1991] on the other hand segment a signature at regions of low speed (15% of the mean speed) which they believe are also regions of high curvature.

Computing the distance between two set of signature features depends on the technique being used for comparison. General distance metrics have been discussed in Section 4.2.

5. The Parametric Approach

5.1 Introduction

As noted earlier, the parametric approach is based on using global features. A number of issues related to the parametric approach have been discussed already. We discuss pre-1993 developments only briefly since they are covered in [Plamondon and Lorette 1989] and [Leclerc and Plamondon 1994].

Some of the early work in HSV is not easily available but has been cited in the papers of [Herbst and Liu 1977] and [Lorette 1984]. The earliest cited work on HSV appears to be that of [Mauceri 1965]. Herbst and Liu report that Mauceri took 50 signatures from each of 40 subjects in evaluating his method using power spectral density and zero-crossing features. A FRR of 37% has been reported. Another study, [Farag and Chien 1972], used chain-encoded tablet data and tested the method on signatures from ten subjects. A FRR of 27% and a FAR of 27% have been reported. The work of [Sternberg 1975] who studied HSV using handwriting pressure has been cited by Herbst and Liu who report 6.8% FRR and 3.2% FAR for random forgeries and 17% FAR for skilled forgeries.

5.2 Pre-1993 Work

[Crane and Ostrem 1983] present an approach in which testing consisted of an enrolment phase in which 10 or 12 sample signatures were collected and a reference vector of features was formed by computing the mean and SD of each feature. The test signature vector was then compared to the reference signature vector and the Euclidean norm of the distance vector was computed which, if small enough, authenticated the signature, otherwise rejected it. Up to *three trials* were allowed and a false rejection occurred only if all three signatures failed the verification test. Several sets of results are presented but the definitions of FRR and FAR are not those normally used in the literature. Also, in some tests the experimenters removed some subjects that had very high variance in their signatures (3 out of 58 subjects failed enrolment!). FAR and FRR varying from 0.5% to about 3% are reported, but all experiments appear to have allowed three tries for verification.

[Lorette 1984] used seven dynamic features that are claimed to be invariant under rotation and magnification: number of connected components, number of loops, quantified cumulated phase for signatures on their whole, initial direction of track pen coded in four quadrants, total duration (it is not clear if this is invariant under magnification as claimed), duration of connected components (total time – pen-up time), and mean and maximum velocities in connected components. All the variables were normalised to have a mean of zero and SD of one and Euclidean distance was then computed. A database of 203 signatures from 14 volunteers (15 signatures from almost every volunteer) was used for evaluating the technique. The data was classified into 14 classes using hierarchical classification based on five signatures of each person. The remaining ten signatures were then assigned to nearest classes and this resulted in 8.3% FRR. An iterative process was then used to improve the classes and this resulted in 6.4% FRR. The details of improved classification are not provided and no forgeries were tested.

[De Bruyne 1985] proposes a set of 18 global features including six dynamic features and other static features. Dynamic features include: total time, number of pen lifts, writing time, and pen-up time (pen-up time + writing time = total time, so all three should not be used) as well as the maximum writing velocity and the time at which this velocity occurs. The static features include the following: area, proportion, SDs of x and y values, and ratio of total displacements in the x and y directions. Reference signature was computed using 10 sample signatures. The test signature was compared with the reference signature as well as with of forgeries and a maximum likelihood test is applied. Testing involved only 11 persons' signatures. 3% FRR and 2% FAR were obtained using 10 sample signatures.

[Hastie et al. 1991] describe a model in which a test signature is assumed to consist of a reference signature which is transformed from occasion to occasion. The authors describe the following five-step HSV method:

1. Smoothing – a cubic spline approximation is used to average out the measurement errors
2. Speed – speed is computed after smoothing
3. Time Warping – a time warp function is computed so that correspondence is found between the reference signature and the test signature
4. Segmentation – the signature is segmented using low speed regions (low speed is 15% of mean speed) into a sequence of segments called letters
5. Averaging – estimating the reference signature based on letters

The test signature is processed using steps 1 to 4 above. Distance between the test signature and the reference signature is found at the end of Step 3 and, if a decision is not made there, at the end of Step 4.

Results of using the method described above are presented in a companion paper [Nelson and Kishon 1991]. Ten samples of genuine signatures from each of 20 subjects were used for testing the technique as well as a number of forgeries for four of the 20 subjects. For those four subjects, FRR of 0% and skilled FAR of 18% was found. Nelson and Kishon make the point that shape and dynamics of a signature might play complementary roles in HSV since if a forger is trying to get the shape right he is unlikely to get the dynamics right, and vice versa. They describe computing pen velocities and accelerations in the x and y directions using spline-smoothing and further computations of path velocities, path accelerations, path tangent angles, tangential accelerations, centripetal accelerations, jerks and curvatures. They note that both the pen pressure and speed have highly repeatable patterns in valid signatures of a person although others have claimed that pressure is not very useful. They use the features: signature time, path length, root mean square (RMS) speed, RMS centripetal, tangential and total acceleration, RMS jerk, average horizontal speed and integral of centripetal acceleration magnitude.

A two-stage verification scheme is described in which a first screening stage is used to reject those signatures that are obviously not close to the reference signature. The four best measures from the above list were selected for each user on the basis of smallest SD with respect to their mean and these were used for the screening test which appears to have worked well for the given data. Applying the first stage to signatures of only four individuals, FRR of zero and skilled FAR of about 17% was obtained. The second stage is not discussed.

[Lee 1992] in his thesis aims to design an HSV system yielding good performance for point-of-sale applications. He developed a database of 5603 genuine signatures from 105 human subjects and 4762 forgeries for all of them. The forgeries consisted of random, skilled and timing forgeries. Each genuine signature was forged by two forgers, each forger contributing to all three types of forgeries and six samples of each type of forgery being collected from each forger. This process yielded 3744 forgeries (however $105 \times 2 \times 18 = 3780$). Further forgeries were collected for 22 subjects randomly selected from the 105 individuals. Eight different individuals provided six skilled forgeries of each of the 22. These provided 792 forgeries (although $22 \times 8 \times 6 = 1056$). It therefore appears that some of the forgeries were rejected. A subset of the database consisting of 11 genuine signatures each (six for each were used for the reference signature and five for testing) for 22 individuals and 704 forgeries from eight forgers for each individual, each providing 4 forgeries for each person. It is not quite clear how this small subset was selected. An EER of 3.8% was reported.

5.3 1993-2005 Work

[Chang, Wang and Suen 1993] present a technique based on Bayesian neural networks for on-line HSV of Chinese signatures. A set of 16 features is used which include the following: total time, average velocity, number of segments, average length in the eight directions of the signature, width/height ratio, left-part/right-part density ratio, and upper-part/lower-part density ratio. Using a database of 800 genuine signatures from 80 people and 200 simple and 200 skilled forgeries by 10 forgers, about 2% FRR, 2.5% skilled forgeries FAR and 0.1% zero-effort FAR were obtained.

[Nelson, Turin and Hastie 1994] propose a set of 25 features which include two time-related features, six features related to velocities and accelerations, four shape-related features, eight features giving the distribution density of the path tangent angles and four giving angle-sector densities of the angular changes and a feature relating to the correlation between the two components of pen velocity. They discuss the statistical basis of HSV and then use three different methods for computing the distance between the reference signature and the test signature, namely the Euclidean distance method, Mahalanobis distance method and the quadratic discriminant method. A simple method of feature selection is described which essentially consists of computing the ratios of the SD to the mean for each feature and rank-ordering the features according to this ratio. It is not explained why a feature with small normalised SD would provide good discrimination between the genuine signatures and forgeries. A variety of schemes, for example using individual best 8, 10, 12 or 14 of the 25 features, are evaluated. The performance of all these sets is similar although the individual best 8 and 10 seem to perform the best with FRR near zero. Using a Euclidean distance method, they identify best 10 of the 25 features and obtain 0.5% FRR and 14% FAR.

[Gupta and Joyce 1997a] propose an algorithm with the aim of using a small set of global features that are easy to compute and invariant under most two-dimensional transformations (e.g. rotation, slant, size). They use six features in the initial experiments: total time, number of velocity sign changes in the *x* and *y* directions, number of acceleration sign changes in the *x* and y directions, and total pen-up time. The HSV algorithm described in Section 4 is now used based on Euclidean distance. It is shown that time by itself is the best single discriminator and pen-up time is also a good discriminator. Including path length in the set of attributes improves the performance of the technique and good results were obtained when path length was included and the reference signature built using 10 sample signatures. An FRR of about 0.5% was obtained with FAR of little more than 10%. The authors note that:

1. A surprising number of forged signatures have feature values that are more than 20 SDs away from the reference signature mean, many are more than 50 SDs away. This would not have been surprising for random forgeries but all forgeries examined were produced by volunteers who had practised forging the signatures.

2. The values of many forged signatures were usually far away from the reference signature means for the following features: total time, acceleration sign changes, pen-up time, path length, and the *x*-acceleration zero count. For other attributes, namely the two velocity sign changes, *y*-acceleration zero count and the segment count, many more forged signatures had feature values closer to the corresponding reference signature mean.

An earlier version of this technique was implemented by Texas Instruments using the TMS320 DSP family [Dullink *et al.*, 1995].

5.4 Summary of Parametric Techniques

In the last few years, it is becoming increasingly apparent that a purely parametric approach is unlikely to lead to much better performance than that reported above. Most recent developments therefore are based either on using the functional approach or using a combination of parametric and functional approaches.

6. The Functional Approach

6.1 Introduction

In the functional approach, all the values of selected local features of a signature are assumed important and the test and reference signatures are compared point-to-point using sets of these values. Therefore much more information about the signatures is often necessary.

6.2 Pre–1993 Work

As discussed earlier, the functional approach is based loosely on the idea of comparing a test signature with a reference signature by comparing the different parts of the signature separately and combining these comparisons to achieve an overall similarity measure. As noted earlier, difficulties arise since even the genuine signatures of one person usually have a certain level of variation, making a direct point-to-point comparison almost impossible. In order to make more effective comparisons, a system must perform some type of alignment of the test and reference signatures in an attempt to "line-up" corresponding parts of the signatures. This alignment may in turn create problems of its own, since forgeries undergo alignment as well as genuine signatures.

The techniques described in this section rely mostly on the functional approach but many use global features as well.

[Herbst and Liu 1977] used two orthogonal accelerometers mounted on an experimental pen to sample a signature at the rate of 200 times per second. They found that most signatures were of 2-10 seconds in duration with an average time of about 5 seconds. Each signature was heuristically partitioned into segments and corresponding segments were cross-correlated after the segments were modified (aligned) based on the duration of the interval and discrepancies between the test signature and the reference signature. It was found that segments in the range of 1-2 seconds performed the best. Longer segments were more difficult to align due to minor variations within the segments. Often segmentation was based on pen-down and pen-up occurrences within the signature. Another approach to segmentation was to use equal time components of the signature.

Before applying the correlation test, the algorithm rejected a signature if it took time that was more than 20% different than the reference signature. This resulted in almost 60% of forgeries being rejected. Once this initial test was passed, the correlation test was applied after the segments were shifted if necessary by up to 20 per cent of the reference signature's segment duration, extra pen lifts were eliminated and correlation results were weighted to penalise excessive shifting, and weighted by the reference signature length of the segment since larger segments were believed to be harder to forge. The technique was evaluated by first collecting five sample signatures each from 70 users. One or two reference signatures were selected such that the distance between the selected signatures and the remaining signatures was at least equal to a prespecified value. These were considered to be the "best" reference signatures. Another 695 genuine test signatures and 287 forged signatures were used for testing. FRR of more than 20% was obtained with an FAR of around 1%, although, not surprisingly, a much lower FRR was obtained if a user was able to undergo three trials to get verification.

[Liu, Herbst and Anthony 1979] propose using writing pressure in addition to the two acceleration measurements used by [Herbst and Liu 1977]. It was found that correlation between pressure waveforms shows little discrimination since the gross form of the pressure waveform dominates the correlation values, but they found that it was more effective to remove low frequency paper contact components of the pressure waveform. Using the acceleration and pressure correlations separately as well as together, they carried out some experiments using signatures from 24 subjects and obtained results close to 16% FRR and below 1% FAR, better than earlier results of [Herbst and Liu 1977].

To overcome the limitations of the regional correlation technique of Herbst and Liu, [Yasuhara and Oka 1977] suggest that the reference signature and the test signature be compared using non-linear time alignment, a technique that has been used in speech recognition. The technique involves building what they call *time registration paths* by plotting the reference signature pen force values against the test signature values. An algorithm is designed which then searches for a path for which the Euclidean distance between the reference signature and the test signature is minimised. This distance is tested against a threshold value and if below the threshold the signature is verified, otherwise it is rejected.

In a US patent, [Bechet 1990] proposes a method in which *x* and *y* velocities are used to compare a reference signature with the test signature. The method involves dividing the velocity signals into discrete time segments and then comparing the corresponding velocity segments. Bechet notes that it was observed that each signature consists of a number of segments where the number is predetermined for each person's signature. Random signals were found to occur between the segments including variations in position of the segments and their duration and some parts of the segments were found to be missing. They used 600 ms segments that are compared after a complex standardisation process involving shifting and scaling so that the difference between the reference segment and the test segment is minimum. Total distance between the test signature and reference signature is then obtained using the segments, and a decision is made regarding acceptance of the test signature based on the number of segments that pass a threshold test.

6.3 1993–2005 Work

[Plamondon 1994] presents a multilevel HSV system that uses global features as well as point-to-point comparison using personalised thresholds. Global features used include total pen-down time, the percentage of pen-up time and the percentage of time when the angular velocity is positive. These features are used for the first stage of verification. The signature is normalised using rotation and scaling and local correlations are computed between portions of the test signature velocity values and the corresponding values of the reference signature using segments alignment based on elastic matching. This second stage is followed by a third stage involving calculation of variations between the normalised coordinate values of the test signature and the reference signature using local elastic pattern matching. For evaluation, three signatures from each of eight individuals were used and eight other people provided three forgeries for each of the eight genuine signers in 64 sessions after having access to genuine signatures and with information on the dynamics of these signatures having been provided as a sound sequence. Another set of genuine signatures was obtained from six other subjects, each providing nine signatures, three of which were used as reference signatures. Tests were performed with the two databases for adjusting the discriminating function to minimise errors. It therefore appears that test signatures and reference signatures were used in deciding individual thresholds that minimised errors. Therefore the results obtained, FAR of 0.5% and FRR of 0.0%, cannot be considered reliable. Also, the testing was very limited and the number of signatures tested was small.

As noted earlier, [Dimauro, Impedovo and Pirlo 1994] believe a signature consists of a sequence of fundamental components called strokes which occur in positions that are constant in the signature of each individual, although the number of components in one signature may be different than in another signature of the same subject. Their method involves finding the components of the test signature and then checking that the components appear in the sequence derived from the reference signatures. If the sequence does not fit, the test signature is rejected, otherwise the components are compared with those of the reference signatures. The technique was tested on signatures of twenty persons who provided 50 reference signatures and 40 test signatures each and forgeries from ten people who watched all genuine signatures being signed and provided four forgeries for each of the twenty genuine signers (thus each forger provided 80 forgeries). An overall FRR of 1.7% and FAR of 1.2% was obtained. The major problem with the evaluation is the use of 50 reference signatures, which is quite unrealistic.

[Nalwa 1997] challenges the notion that the success of on-line HSV hinges on capturing velocities or forces during signature production. His proposed approach is based on using jitter, aspect normalisation, parameterisation over normalised length, sliding computation window, centre of mass, torque, moments of inertia, moving coordinate frame and saturation, weighted cross-correlation and warping.

Jitter is proposed since it is claimed that a person forging a signature is constantly correcting the pen trajectory to conform to an a priori curve. Aspect normalisation is based on the observation that individuals do not scale their signatures equally along both x and y dimensions. Parameterisation of a signature over its normalised arc-length is recommended as opposed to parameterisation over time. Once parameterised, a sliding window is used to compute the following five characteristics of the signature over each window: centre of mass (x and y coordinates of the centre in the window), torque (twice the signed area, negative if clockwise, swept with respect to the origin by the portion of the signature within the window), and moments of inertia about the x-axis and y-axis within the window. Warping is now used in comparing the reference signatures with the test signature so that an overall error measure is minimised. The algorithm presented consists of normalisation (to make the algorithm largely independent of the orientation and aspect of the signature), description (to generate the five characteristics) followed by comparison (to compute the error between the signature and the reference signatures). It is recommended that errors from different models be combined using the harmonic mean so that if one of the errors is low then the mean is low and the signature is verified.

Three signature databases were used to test the proposed algorithm. The first consisted of 904 genuine signatures from 59 different individuals and 325 skilled forgeries. Signatures of one of the signers were removed before testing. The second set consisted of 982 signatures from 102 individuals, collected in a single session. There were 401 skilled forgeries. Some genuine signatures and forgeries from this dataset were also removed. The third dataset had 790 genuine signatures from 43 signers and 424 skilled forgeries. The results from the three test databases and one that included all three, using 4, 5 and 6 reference signatures, are presented. The EER varies between 2 and 5. As noted by the author, it is very difficult to compare results of one study with those of another, in particular if other researchers have no access to the test database that has been used.

[Kashi *et al.* 1997] describe a method using both local and global features. The signature is first normalized using its Fourier coefficients. The global features presented by Nelson, [Turin and Hastie 1994] are then used. Local features of inclination angle and difference between adjacent angles are used to build a left-to-right HMM model. As noted earlier, signatures were pre-processed to remove irregular and redundant points, cusps were detected and marked, cubic B-spline smoothing was applied and the signature was resampled at intervals of equal arc length. Using the Murray Hill signature database of 542 genuine signatures and 325 forgeries, it is shown that the HMM performs worse than using only the global features and a technique combining both is judged the best, resulting in EER of 2.5%. Six sample signatures were used.

[Gupta and Joyce 1997b] describe a technique which uses x and y profile extrema values to capture a signature's shape. The technique is best described by using an example. Consider Figure 9 which shows the x and y profiles of a signature fragment. In the figure local maxima and minima of the x profile are labelled A and B respectively. Corresponding labels for the y

profile are C and D. The extrema of both profiles together may be represented by the string ADBCABDPBCDCABDPBCADBC. The letter P has been inserted for pen-up.

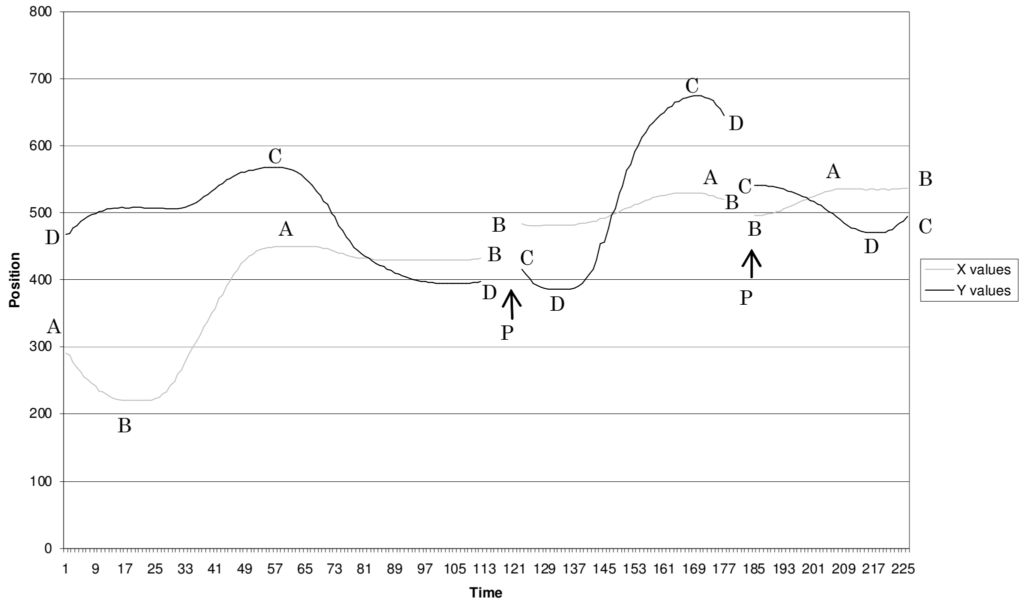

Figure 9. *x* and *y* Profiles of a Signature Fragment

To compare two signature shapes, the string representations of the extrema for each are found and the two strings are then compared and distance between them is calculated. Gupta and Joyce use a number of sample signatures as reference and compare the test signature to each of these to find either the mean distance or the smallest distance. The basis of using the smallest distance is that the sample signatures provide a collection of signatures that show the habitual variations in a person's signature and the test signature should be compared with the reference signature closest to it. The distance so computed is now compared with the threshold and if smaller the test signature is accepted.

The above representation of *x* and *y* extrema was found to be not sufficiently effective and further work was carried out by [McCabe 1997] and [Gupta and Joyce 2001]. Essentially, the representation was modified to include information about long time gaps between successive extrema by inserting a symbol T for every four samples. A typical string representation of a signature is more than 100 symbols long, including 2-3 P symbols and perhaps 20 T symbols in addition to the A, B, C and D. Using a database of signatures of 60 people, the improved implementation yields skilled forgery EER of 4.8%. Five sample signatures were used. [McCabe 1997] refined the technique further by controlling the variability in reference signatures so that the ratio of SD to reference mean is at least 0.175 and the ratio of reference mean to average reference string length is no more than 0.25. These changes result in significant improvement in performance although these parameters were tuned using the test database. EER of close to 2% was obtained. Use of fewer reference signatures was also evaluated. Three reference signatures result in EER of 4.3% while only one in about 6.2%.

[Dolfing *et al.* 1998] outline a method of using hidden Markov models in which each signature is modelled with a single HMM without any preprocessing. The signature samples are the usual five parameters captured by a tablet. The samples were blocked into segments by the condition that *y* velocity be zero. A feature vector of 32 features (13 spatial features, 13 dynamic features and 6 contextual features) was computed for each segment. The test database consisted of 1530 genuine signatures, 3000 amateur forgeries by 51 people and 240 professional forgeries. Each individual contributed 30 signatures which were divided into training, validation and testing set of 15, 5 and 10 signatures respectively. Using individual thresholds, EER of 1.9-2.9% were obtained for different types of forgeries. The error rate was reduced to about half if writers with the shortest signatures were removed from the database. It is not quite clear how the individual thresholds were computed.

[Jain, Griess and Connell 2002] propose a system in which certain critical points, for example, start and end points of a stroke and points of trajectory change, are extracted for each signature. The number of strokes is used as a global feature. Two types of local features, spatial and temporal, are extracted from the x and y coordinates. The spatial features are static features that relate to the shape of the signature: the x and y coordinate differences between two consecutive points, the absolute y coordinates with reference to the centre of the signature, the sine and cosine of the angles with the x-axis, the curvature, and the grey values in a 9 x 9 pixel neighbourhood. Signature samples are then resampled uniformly with equidistant spacing. Each signature is then transformed into one long stroke by concatenating all the strokes followed by smoothing.

The temporal feature was the speed at local points. The absolute and relative speed at each resampled point and average speeds between two critical points were calculated and their importance was investigated. Many experiments were carried out to select features as well as to compare a common threshold with writer-dependent thresholds. The speed feature was found to be very useful and writer-dependent thresholds were found to be better than a common threshold.

Local features of each signature are represented as a string at each sampling position and a modified string matching is used to find dissimilarity values. A penalty for differences in the number of strokes is included. In the verification process, a test signature is compared with each reference signature and the dissimilarity values are combined into one value. The proposed technique was tested using two datasets. The first dataset contained 520 signatures, ten signatures each from 52 writers, collected in one session. The second dataset was a superset of this dataset and contained a total of 1,232 signatures collected from 102 writers, seventeen of which contributed more than ten signatures in multiple sessions over a period of up to one year. Twenty writers provided three skilled forgeries each (a total of only 60) after viewing an original signature. The best error rates using a common threshold were 3.3% FRR and 2.7% FAR and the best error rates using writer-dependent thresholds were 2.8% FRR and 1.6% FAR. The FAR rates appear to be based on random forgeries. No FAR for skilled forgeries was reported.

[Feng and Wah 2003] use dynamic time warping (DTW) to match two signatures so that the match with the least global cost is found. The process changes the input such that the end of the input waveform is aligned with that of the reference and the peaks and valleys are aligned with those of the reference. Both x and y signals are independently warped. The DTW technique has two drawbacks, first the technique is expensive in computation time and, secondly, warping aligns not only genuine signatures but also forgeries. This paper presents an approach called extreme points warping (EPW) which involves warping only selected important points (the peaks and valleys) of a signature. EPW involves first finding these points (the algorithm ignores small peaks and valleys), matching them and then warping the segments between them. EPW and DTW are compared on a database of 1000 signatures of 25 users and the results show that EPW is much faster than DTW and leads to slightly better EER although the error rates are still quite high.

[Ortega-Garcia, Fierrez-Aquilar, Martin-Rello and Gonzalez-Rodriguez 2003] present results of using the five time sequences, *x* and *y* coordinates, pressure, inclination and attitude as well as three derived sequences, path tangent angle, path velocity and log curvature radius. If the first and second derivative of each of these sequences is computed, the total time sequences are 24. A signature sample that has say 1000 samples would generate 24,000 values. The functional values are normalised to obtain zero mean and unit standard deviation. Signatures are modelled using hidden Markov models (HMM) based on the sequences. The technique was tested using a signature database of 15 genuine signatures and 15 forgeries each from 50 people. The tests, using the same threshold for all, resulted in 4.83% EER which reduced to 0.98% by using user-specific thresholds.

[Quan and Ji 2005] define sixteen types of extrema points including eight maxima and minima in the *x* and *y* directions (two different points for clockwise and anticlockwise motion of the pen) and eight different combinations of maxima and minima (for example, two different *x* maxima and *y* minima). These extrema are identified in the signature and some that are too close to other extrema are removed. The pattern for the test signature is then compared with that of the reference signature using derivative dynamic time warping. The distance between the two is computed using a simplified Mahalanobis distance. The testing procedure is somewhat unclear but it appears six samples for each signer were used to find a reference signature. It involved comparing each of the six samples with the other five and counting the number of matching points. The signature that has the largest total matching points was then selected as the reference for the individual. Using random forgeries, an EER of 3.8% was obtained.

[Fierrez-Aguilar, Nanni *et al.* 2005] use a non-parametric statistical recognition strategy based on 40 features selected from a list of 100 global features. Parzen Gaussian windows are then used. The signature database, 16,500 signatures from 330 signers was divided into a training set and a test set. Skilled forgeries EER of between 5% and 7% were obtained with 5 training signatures 1–2% with 20 training signatures. The random forgeries EER was 1–1.5% for 5 training signatures and around 0.5% for 20 training signatures. Employment of user-dependent thresholds was shown to improve the performance.

[Fierrez-Aguilar, Krawczyk *et al.* 2005] compare two function-based approaches to HSV. The two approaches are local and regional approaches. The local approach using dynamic time warping (DTW) is that presented by Jain et al. (2002) to which modifications were made in preprocessing (resampling done in the original paper was not done), feature extraction (features Δx, Δy and p were now extracted at each point), and matching (gap penalty was changed to be based on the distance between two feature vectors). The system was further modified not to employ user-dependent thresholds. It was found that the DTW system outperforms the HMM system in almost all test conditions if user-independent thresholds are used. It was shown that a system combining the DTW and HMM systems employing user-dependent thresholds performs significantly better than individual DTW or HMM systems.

6.4 Summary of Functional Techniques

In the last few years, it has become apparent that functional techniques can generally outperform purely parametric techniques. Most recent developments therefore are based either on using the functional approach or using a combination of parametric and functional approaches.

7. HSV Products and Applications

It has been estimated that the HSV market was worth almost $6m in 2003 and was expected to grow to above $100m in five years. A number of commercial products in HSV are being advertised and some of them are listed here. The list is not comprehensive because of fast growth in this field and there is no simple mechanism to find a list of all products. In some cases it is even difficult to discover if a company that was offering HSV products still exists, some have gone out of business or have been taken over by other companies.

7.1 HSV Products

Communication Intelligence Corporation (CIC, www.cic.com) appears to have a number of signature verification products. A product called Sign-On, it is claimed, allows HSV to be built-in to a variety of widely used software enabling the system to use a handwritten signature instead of a password. It uses not only the signature image, but acceleration, stroke angles, start and stop pressures (if available) and other factors. The signature information can be updated each time a successful verification occurs. The product uses six signatures plus a final verifying signature to build a reference signature. The test signature is then compared with the reference signature resulting in one of three judgments: true, forgery or ambiguous. The product is claimed to have a 2.5% ERR but details of its performance evaluation are not available. Sign-On can also be used to capture, verify and encrypt signatures directly on a PDA and may be used to secure the device which then requires a valid signature to unlock it.

SOFTPRO, a German company, claims that more than 200 financial institutions are using the modules of its SignPlus system (www.signplus.com). SignPlus allows capturing and verification of both static and on-line signatures. JustSign is another on-line HSV software (www.justsign.com) that uses *x*, *y*, pressure, sequence, time and pen strokes information.

SiSign from SiVault Systems (www.sivault.com/solutions/sisign.html) uses the standard HSV procedure that requires registration. It is suggested that registration could be done at a point-of-sale terminal if the customer has another reliable form of identification (for example, a driver’s licence). A signature captured by a pressure-sensitive mobile phone may also be used.

IBM Israel (www.haifa.il.ibm/projects/image/sv) has an HSV product that claims a total error rate of 1.5%. It follows the usual HSV procedure in requiring user enrolment. It is claimed that the reference signature can be stored using only 100-150 bytes of storage. WonderNet, an Israeli company, has a number of products that use HSV techniques for including handwritten signatures in documents. The HSV procedure uses three sample signatures to build a reference signature that is used for verifying test signatures. The system modifies the reference signature each time a signature is authenticated.

Although details are not available, SMARTpen Biometric Authentication System (BiAS) uses HSV to authenticate individuals by the biometric characteristics of their signatures (www.smartpen.com). Another HSV product, DioPass, recognises and identifies on-line signatures on the screens of devices such as PDAs, smartphones and PCs with digitizer. The speed, order of making strokes, number of strokes and handwriting pressure as well as the shape of the signature are used for HSV.

BioSig from Vector Intelligence is an on-line HSV system developed by Lucent Bell Labs. It is based on three patents. The software uses a technique that divides the signature into a sequence of line segments which are ordered according to the time sequence of the signature. Each segment is assigned a stroke direction value that depends upon the orientation of the segment. It also captures 23 features including speed and stroke direction. Six sample signatures are recommended for building a reference signature. A verification score is then found and the signature is accepted or rejected based on that score. An error rate of 1.2% is claimed.

Cyber-Sign (www.cybersign.com) is a dynamic HSV system that analyses velocities in the two dimensions, shape, stroke order, pen pressure and timing information of signature writing. The technique adapts to natural signature changes over time. This product was earlier called ID-007. No details of how it performs are available.

The company Silanis Technology Inc. has a product called ApproveIT, which adds signatures to documents directly from a pen-based input or from a previously captured signature on a file which is safe from tampering due to a password protection feature. If a document is signed using ApproveIT, it is verified to ensure the contents have remained unchanged since the prior approval. If a document is modified after approval, the signature will not print or display itself on the altered document.

SignatureOne Profile Server provides a framework for maintaining electronic signature profiles for HSV. Profile Server works in conjunction with CIC's Sign-it suite of client products. Crypto-Sign also offers a HSV system that uses on-line verification based on capturing x, y, p and t coordinates. The system can be used on a PDA either with an inkless stylus or with an ink-based stylus writing upon a paper on the digitizer.

7.2 HSV Applications

HSV has been reported to have many applications. Web sites of vendors of HSV products often list where and how their products are being used. Applications in the legal, medical, and banking sectors are common. The applications involve access control, data privacy, personalised document approval and on-line shopping.

Credit cards deal with perhaps as much as $10 billion worth of purchases every day. It has been reported that credit card issuers lose more than $1 billion each year due to credit card fraud. Although this sum is substantial, it is small when expressed as a percentage of total credit card purchases, certainly well below 1%. Credit card issuers however earn only a small commission on the purchase price and therefore the total losses due to credit card fraud are significant for the credit card issuers.

Credit card issuers, mostly banks, have taken a number of steps in the last few years to reduce credit card fraud. These have included requiring the credit card holder on receipt of a new card to phone the credit card issuer to activate the card in an attempt to reduce fraud involving lost or stolen cards. This has resulted in some decline in credit card fraud, which was rising before this action was taken.

A reliable HSV technique could have applications in reducing credit card fraud although there appear to be some hurdles that must be overcome if the technology is to be useful. A stolen card has the owner's signature on the back and this makes it particularly easy to forge the signature, given minimal checking of signatures at the place of purchase. This problem of forging the signature does not even arise if a credit card has been intercepted in the mail, since any signature could be then put on the back of the card. One possible approach to reducing credit card fraud would be to require the owner of a new credit card to visit the bank that issued the card and supply sample signatures so that information from the signatures could be put on the card electronically making it unnecessary to have the owner's signature appear on the card, a signature that may be viewed and forged by someone who has stolen the card.

The above scheme of course has a problem. If one credit card issuer requires that the owner of a new card to go to the bank and produce identification and give sample signatures while other credit card issuers continue with the present procedure of requiring no such visit, a credit card requiring signatures would not be very popular. The suggested approach is therefore not convenient for the customer and is unlikely to be adopted in a very competitive marketplace where customers are being offered new credit cards almost every day.

Another approach might be possible. When the owner of a new card makes the first purchase using the card, the sales staff could be asked (by the credit card terminal) to check the customer's identification (for example, a driver's licence) and provide the identification number, as is done when cashing a cheque. This scheme does have merit in that it only creates a minor nuisance when the customer is asked the first time for an identification, but using this approach the HSV technique may not work very well since the system has only one signature to serve as its reference signature. Further signatures will of course become available as the customer makes more purchases, but then it is not possible to confirm that those signatures belong to the genuine card owner.

Yet another approach might be possible. In this approach, the customer uses the credit card normally, but his or her signatures are captured electronically and compared with a signature profile that has been built over the last few weeks or months. When the result of comparison shows a significant mismatch, a suitable action is taken which may include either rejecting the purchase being charged to the card or, preferably, bringing the mismatch information to the attention of a human operator who can take appropriate action, for example contacting the card owner.

HSV might also have uses in computer user authentication if an HSV technique could be designed that provided a high level of security against intruders through a zero or near zero FAR. It might be suitable for user authentication not only at login time but also for accessing sensitive applications such as sensitive databases or exclusive software. A typical on-line HSV system will of course require that a signature input device like a graphics tablet be connected to each workstation to capture signature details. This technique has the potential to replace the password mechanisms for accessing computer systems in some situations. The major disadvantage of this approach of course is the requirement that a graphics tablet be attached to each workstation.

Reliable HSV could well have other applications. For example, it might assist in reducing the forging of passports. Although the US Government is now requiring that biometric information be included in passports, there is considerable opposition to such a move in many countries. HSV could be a more acceptable approach since an application for a passport normally requires that the applicant go to some authorised office to file an application form and that signatures be certified in the presence of an authorised officer. It is thus not unreasonable that the office where a passport application is filed might require the applicant to provide a set of sample signatures, which would be captured electronically and used for building a reference signature. That reference signature could then be placed on a magnetic strip on the passport (passports are likely to feature magnetic strips for faster processing anyway). At the port of entry, the person entering a country would then be required to sign his or her name on a graphics tablet and the signature would be compared with the reference signature on the passport strip. Forging of passports would then be almost impossible.

Conclusions

On-line HSV is a very active area of research and a large number of research papers have been published in the last few years (for example, refer to the bibliography http://iris.usc.edu/Vision-Notes/bibliography/char1012.html). We have attempted to present a snapshot of developments in the field with focus on recent developments. In doing so we have not been able to cover all the techniques, for example we have not discussed techniques that use neural networks (e.g. [Pender 1991] and [Lee 1996] or wavelets (e.g. [Lejtman and George 2001].

Our study suggests that any technique that uses a purely parametric approach is unlikely to provide an EER of less than 5% if a reasonably large signature database is used. Therefore most recent developments use functional approaches and some of such studies have achieved good performance, a number of them claiming well below 5% EER. On-line HSV is an active research area and we believe further improvements are likely.

As noted earlier, the state of the art in HSV makes it difficult to draw definitive conclusions about which techniques are the best. The following factors frustrate a person wishing to compare different techniques:

1. Performance of an HSV system that uses different features for different individuals is better than a system that uses the same features for all.

2. Performance of an HSV system that uses different threshold values for different individuals is better than a system that uses the same threshold value for all.

3. Performance of an HSV system that uses more signatures for building the reference signature is better than a system that uses a smaller number of signatures.

4. FRR of an HSV system that uses more than one test signature to make a judgement about whether the subject is genuine or not is better than a system that uses only one signature.

5. Performance of an HSV system that uses the genuine signatures as well as some or all of the forgeries that are used in performance evaluation in building the reference signature or in deciding which features to choose and/or what threshold to select is better than a system that does not use any test signatures in building the reference signature.

6. Performance of an HSV system that is tested on a database of test signatures that have been screened to eliminate some subjects with problem signatures is likely to be better than a system that has not carried out any such screening.

However, more careful testing by researchers and availability of public test databases is resulting in welcome progress in the area. Continuing progress in performance evaluation and comparison of HSV techniques is needed.

We have briefly described some HSV products available in the market. Interest in these products has obviously grown since 9/11 terrorist attacks. The products are being used in a wide variety of applications.

References

BECHET, L. Method of Comparing a Handwriting with a Reference Writing. US Patent 4, 901, (1990), 358.

BRAULT, J., and PLAMONDON, R. How to Detect Problematic Signers for Automatic Signature Verification. Int Carnahan Conf on Security Technology, (1989), 127-132.

BRAULT, J., and PLAMONDON, R. A Complexity Measure of Handwritten Curves: Modeling of Dynamic Signature Forgery. *IEEE Trans on Systems, Man, and Cybernetics*, 23, 2, (1993), 400-413.

DE BRUYNE, P. Signature Verification using Holistic Measures, *Comp Security*, Vol 4, (1985), 309-315.

CHANG, H., WANG, J., and SUEN, H. Dynamic Handwritten Chinese Signature Verification. Proc Second Int Conf on Document Analysis and Recognition, (1993), 258-261.

CRANE, H. and OSTREM, J. Automatic Signature Verification using a Three-axis Force-Sensitive Pen. *IEEE Trans on Systems, Man, and Cybernetics*, SMC-13, 3, (1983), 329-337.

DIMAURO, G. IMPEDOVO, S. and PIRLO, G. Component-Oriented Algorithms for Signature Verification, *Int Journal of Pat Rec and Art Int*, 8, 3, (1994), 771-793.

DOLFING, J. G. A., AARTS, E. H. L. and VAN OOSTERHOUT, J. J. G. M. On-line signature verification with Hidden Markov Models. In Proc. International Conf on Pattern Recognition, (1998), 1309-1312.

DULLINK, H., VAN DAALEN, B., NIJHUIS, J., SPAANENBURG, L. and ZUIDHOF, H. Implementing a DSP Kernel for Online Dynamic Handwritten Signature Verification using the TMS320 DSP Family, Texas Instruments, (1995).

FAIRHURST, M. and BRITTAN, P. An Evaluation of Parallel Strategies for Feature Vector Construction in Automatic Signature Verification Systems, *Int Journal of Patt Rec and Art Int*, 8, 3, (1994), 661-678.

FARAG, R. and CHIEN, Y. On-line Signature Verification. Proc Int Conf on Online Interactive Computing, Brunel University, London, (1972), 403-424.

FENG, H. and WAH, C. C. 2003, Online Signature Verification using a new Extreme Points Warping Technique, *Pattern Recognition Letters*, Vol 24, No 16, pp 2943-2951.

FIERREZ-AGUILAR, J., KRAWCZYK, S., ORTEGA-GARCIA, J. and JAIN, A.-K. Fusion of Local and Regional Approaches for On-line Signature Verification", in Proc. Intl. Workshop on Biometric Recognition Systems, IWBRS, Springer LNCS-3781, Beijing, China, (2005), 188-196.

FIERREZ-AGUILAR, J., NANNI, L., LOPEZ-PENALBA, J., ORTEGA-GARCIA, J. and MALTONI, D. An On-line Signature Verification System Based on Fusion of Local and Global Information, in Proc. 5th IAPR Intl. Conf. on Audio- and Video-based Biometric Person Authentication, AVBPA, Springer LNCS-3546, New York, USA, (2005), 523-532.

FIERREZ-AGUILAR, J., ORTEGA-GARCIA, J. and GONZALEZ-RODRIGUEZ, J. Target Dependent Score Normalization Techniques and their Application to Signature Verification, in Proc. Intl. Conf. on Biometric Authentication, ICBA, Springer LNCS-3072, Hong Kong, (2004), 498-504. (Selected for IEEE Trans. SMC-C Special Issue on Biometric Systems)

FITTS, P. The Information Capacity of the Human Motor System. In Controlling the Amplitude of Movement, *J Exp Psych*, 47, (1954), p 381-391.

GUPTA, G. and JOYCE, R. A Study of Some Pen Motion Features in Dynamic Handwritten Signature Verification. Unpublished Report, Department of Computer Science, James Cook University, Townsville, Australia, (1997a).

GUPTA, G. and JOYCE, R. A Study of Shape in Dynamic Handwritten Signature Verification. Technical Report, Computer Science Dept, James Cook University, Townsville, Australia, (1997b). http://www.cs.jcu.edu.au/Research/techreports/ftp97.html#97-4

GUPTA, G. and JOYCE, R. A New Technique of Capturing Shape in Online Handwritten Signature Verification. Proc of XXXVI Annual Convention, CSI2001, S. Sen and S. K. Bamdopadhyay (eds.), (2001), C-41 – C48.

GUPTA, G. and MCCABE, A. A Review of Dynamic Handwritten Signature Verification, Department of Computer Science, James Cook University, Townsville, Australia, (1997). http://www.cs.jcu.edu.au/~alan/Work/HSV-Lit_rev.html

HARRISON, W. *Suspect Documents*. Praeger Publishers, New York, (1958).

HASTIE, T., KISHON, E. CLARK, M. and FAN, J. A Model for Signature Verification. Proc IEEE Int Conf on Systems, Man, and Cybernetics, Vol 1, Charlottesville, (1991), 191-196.

HERBST, N. and LIU, C. Automatic Signature Verification Based on Accelerometry. *IBM J Res Dev*, (1977), 245-253.

HILTON, O., *Scientific Examination of Documents*. Callaghan and Co, Chicago, (1956).

HILTON, O., Signatures — Review and a New View. *J of Forensic Sciences*, JFSCA, 37, 1, (1992), 125-129.

HONGO, Y., MURAMATSU, D., and MATSUMOTO T. Modification of Intersession Variability in On-line Signature Verifier, AVBPA 2005, Lecture Notes in Computer Science 3546, Springer Verlag, (2005), 455-463.

IGARZA, J. J., GOIRIZELAIA, I., ESPINOSA, K., HERNÁEZ, I., MÉNDEZ, R., and SÁNCHEZ, J. Online Handwritten Signature Verification Using Hidden Markov Models. CIARP 2003, (2003), 391-399.

IGARZA, J. J., GÓMEZ, L., HERNÁEZ, I., and GOIRIZELAIA, I. Searching for an Optimal Reference System for On-Line Signature Verification Based on (x, y) Alignment. ICBA 2004, 519-525.

JAIN, A. K., GRIESS, F. D., and CONNELL, S. D. Online Signature Verification. *Pattern Recognition*, 35, 12, (2002), 2963-2972.

JOBSON, J. *Applied Multivariate Data Analysis*, Volume II: Categorical and Multivariate Methods. Springer-Verlag, New York, 1992).

KASHI, R. S., HU, J., NELSON, W. L., and TURIN, W. On-line Handwritten Signature Verification using Hidden Markov Model Features, Proc of the 4th Int Conf on Document Analysis and Recognition, 1, (1997), 253-257.

KETABDAR, H., RICHIARDI, J., and DRYGAJLO, A. Global Feature Selection for On-line Signature Verification, Proc. 12th Conf of the Int Graphonomics Society, Solerno, Italy, (2005).

KHOLMATOV, A. and YANIKOGLU, B. Biometric Authentication using Online Signatures (extended version of the above), Lecture Notes in Computer Science 3280 ISCIS 2004, pp 373-380

LECLERC, F. and PLAMONDON, R. Automatic Signature Verification: The State of the Art –1989-1993, *Int Journal of Patt Rec and Art Int*, 8, 3, (1994), 643-660.

LEE, L. On-line Systems for Human Signature Verification. Ph.D. Thesis, Cornell University, 1992.

LEE, L., BERGER, T., and AVICZER, E. Reliable On-line Human Signature Verification Systems. *IEEE Trans on Pattern Recognition and Machine Intelligence*, 18, 6, (1996), 643-647.

LEE, L.L. Neural approaches for human signature verification, 3rd International Conference on Signal Processing, 2, (1996), 1346-1349.

LEI, H. and GOVINDARAJU, V. A Comparative Study on the Consistency of Features in On-line Signature Verification, Proceedings of the Structural, Syntactic, and Statistical Pattern Recognition, in A. Fred, T. Caelli, R. P. W. Duin, A. Campilho and D. de Ridder (eds.), Lisbon, Portugal, (August, 2004), Lecture Notes in Computer Science, Springer Verlag, Volume 3138, p 444.

LEJTMAN, D. Z. and GEORGE, S. E. 2001, On-line Handwritten Signature Verification Using Wavelets and Back-propagation Neural Networks, Sixth International Conference on Document Analysis and Recognition (ICDAR'01), (2001), p 992.

LIU, C., HERBST, N. and ANTHONY, N. Automatic Signature Verification: System Description and Field Test Results. *IEEE Trans on Systems, Man, and Cybernetics*, SMC-9, 1, (1979), 35-38.

LORETTE, G. On-line Handwritten Signature Recognition based on Data Analysis and Clustering. Proc of 7th International Conf on Pattern Recognition, Montreal, (1984) 1284-1287.

MAUCERI, A. Feasibility Studies of Personal Identification by Signature Verification. Report No. SID 65 24 RADC TR 65 33, Space and Information System Division, North American Aviation Co, Anaheim, California, (1965).

MCCABE, A. Implementation and Analysis of Handwritten Signature Verification Technique, Honours Thesis, Department of Computer Science, James Cook University, Townsville, Australia, (1997). http://www.cs.jcu.edu.au/Research/topics/theses/1997/Alan.html

MILLER, B. Vital Signs of Identity. *IEEE Spectrum*, (1994), 22-30.

NALWA, V. Automatic On-line Signature Verification. *Proc IEEE*, 85, 2, (1997), 215-239.

NELSON, W. and KISHON, E. Use of Dynamic Features for Signature Verification. Proc IEEE Int Conf on Systems, Man, and Cybernetics, 1, Charlottesville, (1991), 201-205.

NELSON, W., TURIN, W., and HASTIE, T. Statistical Methods for On-line Signature Verification. *Int Journal of Patt Rec and Art Int*, 8, 3, (1994), 749-770.

ORTEGA-GARCIA, J., FIERREZ-AQUILAR, J., MARTIN-RELLO, J. and GONZALEZ-RODRIGUEZ, J. Complete Signal Modeling and Score Normalization for Function-Based

Dynamic Signature Verification. In J. Kittler and M. S. Nixon (eds.), AVBPA 2003, Lecture Notes in Computer Science 2688, (2003), 658-667.

ORTEGA-GARCIA, J., FIERREZ-AQUILAR, J., SIMON, D., GONZALEZ, J., FAUNDEZ, M., ESPINOSA, V., SATUE, A., HERNAEZ, I., IGARZA, J.-J., VIVARACHO, C., ESCUDERO, D., and MORO, Q.-I. MCYT Baseline Corpus: A Bimodal Biometric Database, IEE Proceedings on Vision, Image and Signal Processing, Special Issue on Biometrics on the Internet, 150, 6, (2003), 395-401.

OSBORN, A. *Questioned Documents*. Boyd Printing Co, Albany, NY, 2nd Edition, (1929).

PARIZEAU, M. and PLAMONDON, R., A comparative Analysis of Regional Correlation, Dynamic Time Warping, and Skeletal Tree Matching for Signature Verification. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 12, 7, (1990), 710-717.

PARKS, J., CARR, D. and FOX, P. Apparatus for Signature Verification. US Patent Number 4,495, 644, (1985).

PENDER, D. Neural Networks and Handwritten Signature Verification. PhD Thesis, Department of Electrical Engineering, Stanford University, (1991).

PHELPS, R. A Holistic Approach to Signature Verification. Proc of 6th International Conf on Pattern Recognition, 2, (1982), 1187.

PLAMONDON, R. Looking at Handwriting Generation from a Velocity Control Perspective. *Acta Psychologica*, 82, (1993), 89-101.

PLAMONDON, R. The Design of an On-line Signature Verification System: From Theory to Practice, *Int Journal of Patt Rec and Art Int*, 8, 3, (1994), 795-811.

PLAMONDON, R., ALIMI, A., YERGEAU, P. and LECLERC, F. Modelling Velocity Profiles of Rapid Movements: A Comparative Study. *Biological Cybernetics*, 69, (1993), 119-128.

PLAMONDON, R., and LORETTE, G. Automatic Signature Verification and Writer Identification — The State of the Art. *Pattern Recognition*, 22, (1989), 107-131.

PLAMONDON, R., and PARIZEAU, M. Signature Verification from Position, Velocity and Acceleration Signals: A Comparative Study. Proc of 9th International Conf on Pattern Recognition, 1, (1988), 260-265.

QUAN, Z. and JI, H. Aligning and Segmenting Signatures at Their Crucial Points Through DTW. In D. Huang, X. Zhang and G. Huang (eds.), Advances in Intelligent Computing: International Conference on Intelligent Computing, ICIC 2005, Hefei, China, Proceedings, Part I, Lecture Notes in Computer Science, Volume 3644, Springer Verlag, (2005), 49-58.

RICHIARDI, J., KETABDAR, H., and DRYGAJLO. A. Local and Global Feature Selection for On-line Signature Verification. Proc Eighth International Conference on Document Analysis and Recognition (ICDAR'05), (2005) 625-629.

SHERMAN, R. Biometric Futures. *Computers & Security*, 11, (1992), 128-133.

STERNBERG, J. Automated Signature Verification using Handwriting Pressure. 1975 WESCON Technical Papers, No 31/4, Los Angeles, (1975).

YANIKOGLU, B., and KHOLMATOV, A. An Improved Decision Criterion for Genuine/Forgery Classification in On-line Signature Verification. Presented at ICANN, (June 2003). http://students.sabanciuniv.edu/~alisher/abstracts/online-sig-icann2003.pdf

YASUHARA, M., and OKA, M. Signature Verification Experiment Based on Nonlinear Time Alignment: A Feasibility Study. IEEE Trans on Systems, Man, and Cybernetics, SMC-7, 3, (1977), 212-216.