Signature Verification: Benefits of multiple tries

R. S. Kashi
Avaya Labs Research, 233 Mt. Airy Road, Basking Ridge, NJ 07920, USA
E-mail: ramanuja@research.avayalabs.com

W. L. Nelson
Bell Labs, Lucent Technologies, 600 Mountain Ave, Murray Hill, NJ 07974, USA
E-mail: win@research.bell-labs.com

This method describes the advantages of a signature verification or any other biometric methodology having multiple tries. Traditional verification techniques acquire a test sample and compare with a model and a decision is made as to whether the test sample is genuine or a forgery. While it is clear that allowing a second try will reduce the false rejects, it will also increase the false accepts. However, from our experiments, the increase in false accepts was small compared to the dramatic reduction in false rejects. At the 1% false rejection (FR) rate, the addition of the second try reduced the false acceptance (FA) rate from 4.7% to 1.3%. So, in applications where very low FR is required, allowing the signer a second try appears to be a good option.

1 Introduction

Signature verification is a common behavioral biometric to identify human beings for purposes of establishing their authority to complete an automated transaction, gaining control of a computer, or gaining physical entry to a protected area. Signatures are particularly useful for identification because each person's signature is highly distinct from imposter's attempts, especially if the dynamic properties of the signature are considered in addition to the static shape of the signature. Even if skilled forgers can accurately reproduce the shape of signatures, it is unlikely that they can simultaneously reproduce the dynamic properties as well.

On-line signature verification schemes extract signature features that characterize spatial and temporal characteristics of a signature [1]. The feature statistics of a training set of genuine signatures are used to build a model or template for validating further test signatures [2]. Selecting a good model is the most important step in designing a signature verification system. Signature models are usually described by the set of parameters (features) which can be roughly divided into local and global feature subsets. Global features are calculated for sufficiently large segments of signatures (such as pen-down segments or the whole signature) while local features are calculated for smaller segments

(such as equally spaced sub-segments or even every signature sample). For example, total signature time and length of a signature are global features while slope tangent at each point is a local feature. Local features are more sensitive to handwriting variations than global features, but require more computer resources for processing and storage. The feature set selected should be a judicious choice of global and local features having maximum discriminative power while keeping the set cardinality small.

Another important component of the system is a distance measure between a signature and its model. A popular method is based on elastic matching by Dynamic Warping (DW) [3,4].Dynamic warping performs flexible matching of local features of a model and a signature sample. The closeness of the match is used as discriminator. An alternative strategy (commonly used in speech recognition) is using a Hidden Markov Model (HMM) [5]. An HMM performs stochastic matching of a model and a signature using a sequence of probability distributions of the features along the signature. During a signature verification, the probability that the signature is genuine is calculated. If this probability (which is called the likelihood function) is greater than a given decision threshold, the signature is accepted otherwise it is rejected. Alternatively, the difference between the likelihood score and the mean likelihood score obtained from training can be used as an error measure in isolation or in combination with other feature errors. This approach can be viewed as a statistical matching of the test signature with the signature HMM.

A statistical decision criterion for hypothesis testing (that the signature is genuine with the alternative hypothesis that it is a forgery) is usually evaluated in terms of two kinds of errors: Type-1 error in which the decision is "false" when in fact the hypothesis is true (false rejection), and a Type-2 error in which the decision is "true" when the hypothesis is false (false acceptance). If a feature space can be chosen such that the set of all possible outcomes when the the hypothesis is true is disjoint from the set of all possible outcomes when it is false, then the number of errors of both types can be reduced to zero by determining a decision boundary that separates the two sets. However, in real life situations the objective is to choose a feature space and decision boundary such that one type of error is minimized for a given value of the other. This error trade-off can be shown by plotting the Type-2 vs Type-1 errors as the decision boundary varies.

Traditional signature verification systems decide on a valid/forgery classification based on one test signature. In this paper we experiment with multiple (two & three) instances of the test signature. The verification results from each of the tries can then be combined in a variety of logical combinations to classify as a valid or forgery. In this paper we use the FR/FA trade-off curves

to compare the results of using one two and three tries.

The remainder of the paper is organized as follows: Section 2 presents the experimental procedure, Section 3 describes the results on the Murray Hill signature database [2] and the discussion of the results are presented in Section 4.

2 Experimental Procedure

We conducted our experiments on the Murray Hill database and with algorithms described in our previous papers [2,6]. The test database consisted of 542 genuine signatures and 325 forgeries. Each reference set consisted of the first 6 signatures of every one of the 59 subjects. The digitizing tablet used for gathering the signatures had an 80 X 80 mm glass surface and a writing stylus that was electronically connected by a tether to the capacitive sensing system [7].

For the single try experiment the nine individual test signatures were solely used to determine whether the attempt was a valid or forgery. For the two-try procedure, if the signature/forgery passed the first try, it was accepted; if it failed, another signature/forgery sample was picked at random from the remaining signature/forgery set, and if that sample passed, it counted as an acceptance, otherwise a rejection. So we used an OR logic for deciding on a valid and an AND logic for a forgery. Therefore, if the signature was accepted in the first attempt, there was no need to go with the second try. The same logic combination was used with the three try case where an attempt is classified as valid if the first, second or third attempt passes and is considered a forgery if all the three attempts fail.

Assuming the OR logic for acceptance and the AND logic for rejection and further assuming that the individual attempts are independent we get the following results for the probability of false acceptance P_{FA} and the probability of false rejection P_{FR}:

$$P_{FA} = P_1(Accept|Forgery) + P_2(Accept|Forgery) \tag{1}$$

$$P_{FR} = P_1(Reject|Valid) * P_2(Reject|Valid) \tag{2}$$

where P_1 and P_2 represent conditional probabilities associated with the one and two tries respectively. If the probabilities are the same for the two independent tries then the the above two equations become

$$P_{FA} = 2 * P(Accept|Forgery) \tag{3}$$

$$P_{FR} = [P(Reject|Valid)]^2 \quad (4)$$

For n independent tries, Eq (3) has a multiplier of n, and Eq (4) has an exponent of n.

3 Performance results

The performance of signature verification methods tested on the Murray Hill database will be shown in terms of Type-1 false rejection(FR) errors vs. Type-2 false acceptance (FA) errors, as the decision threshold is increased in small increments.

In Figure 1, the FR/FA error trade-off curves are shown for the algorithm with global features only, for one try (solid line), two tries (dashed line) and three tries (dotted line). As seen from Figure 1, the second try improves the performance in all regions of the plot when compared to the single try. The equal error rate as seen from the figure decreases from about 4.5% to about 3% with the second try. At the 1% false rejection (FR) point, the addition of the second try reduced the false acceptance (FA) from about 13% to about 5%. Also seen in the figure, not much improvement is obtained with the three tries as compared to the two tries.

From the Equations (1) and (2) it is clear that while the false rejection rate decreases (since P_{FR} and P_{FA} are less than 1), the false acceptance rate increases, and hence would lead us believe that no significant gain could be achieved with multiple tries. However, as shown in Figure 2 the FR reduces significantly whereas the FA increases only slightly.

An even more dramatic performance improvement was observed with another algorithm using the combination of local and global features. The equal error rate as seen from the Figure 3 decreases from about 2.5% to about 1%. At the 1 % false rejection (FR) point, the two try technique reduced the false acceptance (FA) from about 5 % to about 1.2%.

Figure 4 shows the separate FR/FA vs decision threshold curves for the one-try, two-tries and the three-tries case, for the algorithm with global and local features. As in Figure 2 the FR curves shifts downwards with increasing number of tries (the false rejection is reduced), the FA curves shifts upwards (the false acceptances are increased). The disproportionate shifts in FR and FA leads to better performances as seen in Figures 1 & 3.

4 Discussion

The same principle of combining trials could be extended to mixed biometric modalities as well. Therefore, in a general case the P_1 and P_2 in equa-

tions (1 & 2) can represent the error probabilities of two different biometrics, eg., signature verification and finger-print verification. Such a combination in our opinion would significantly improve the performance curve when compared to the individual verification schemes. We also believe that such a scheme would be more robust if one of biometrics is chosen from the behavioral type, i.e, from signature, voice, or typing rhythms, and the other biometric is chosen from the physiological or the physical type, i.e, from fingerprints, iris-scans or hand geometry. This conclusion is drawn from the result shown in this study that the reduction in false rejection is far greater than the increase in false acceptance.

The scheme which we have described in this paper aimed at reducing the false rejection rate which is required in banking and point of service applications, where the business does not want to offend its customers and hence would want a very low FR. However, in a different scenario, e.g., logging in as a super-user, the logic would be reversed in the sense that the transaction is accepted iff (if and only if) both the tries (possibly with different biometric modalities) passes and is rejected if either one is rejected. In such a case the false acceptance is multiplicative and the false rejection is additive.

The other main conclusion of this study is that significant performance improvement was achieved without any change in the underlying algorithms. The only change made was of a procedural nature, i.e., allowing for a multiple try mode.

References

1. G. Lorette and R. Plamondon. Dynamic approaches to handwritten signature verification. In *Computer Processing of Handwriting*, pages 65–85, 1990.
2. W. Nelson, W. Turin, and T. Hastie. Statistical methods for on-line signature verification. *International Journal of Pattern Recognition and Artificial Intellingence*, 8, 1994.
3. V. S. Nalwa. Automatic on-line signature verification. *Proceedings of the IEEE*, 85:215–239, 1997.
4. Y. Sato and K. Kogure. Online signature verification based on shape, motion, and writing. In *Proc. 6th of ICPR*, pages 823–826, 1982.
5. L. R. Rabiner and B. H. Juang. An introduction to hidden markov models. *IEEE ASSP Magazine*, 3(1):4–16, 1983.
6. R. S. Kashi, J. Hu, W. Turin, and W. Nelson. On-line handwritten signature verification using hidden markov model features. In *Proceedings of the fourth international conference on document analysis and recogni-*

tion, pages 253–257, 1997.

7. G. E. Blonder, R. A. Boie, L. W. Ruedisueli, E. R. Wagner, and Y. H. Wong. Capacitive moment sensing for electronic paper. AT&T Bell Laboratories Internal Document, 1991.

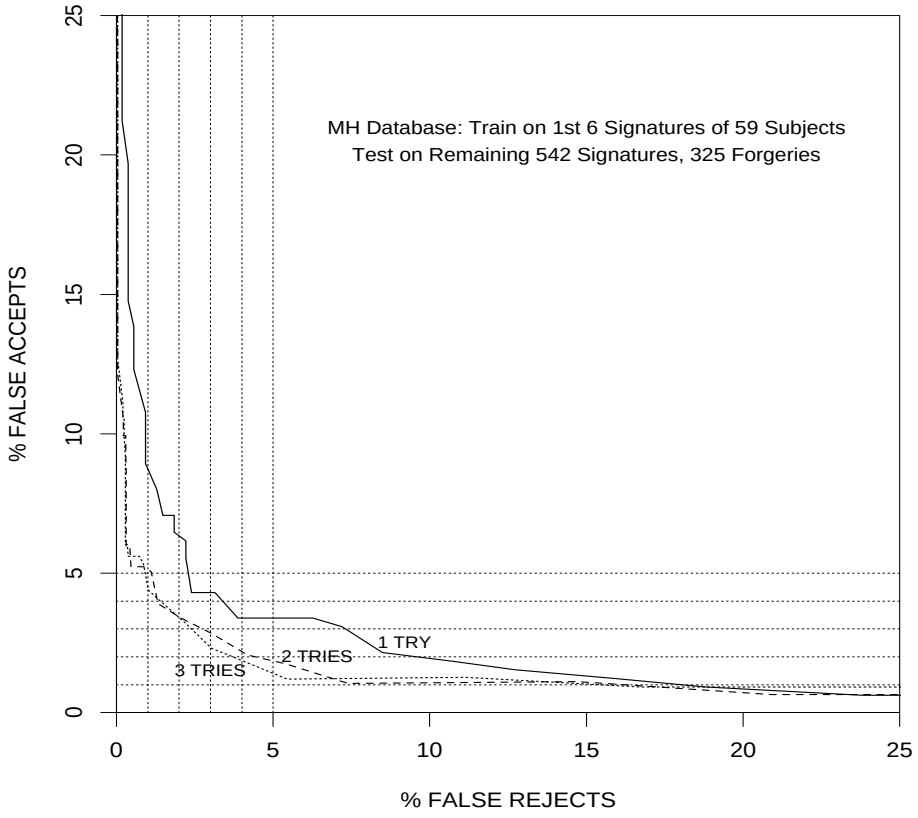

Figure 1: Error trade-off performance curves for the algorithm with global features only

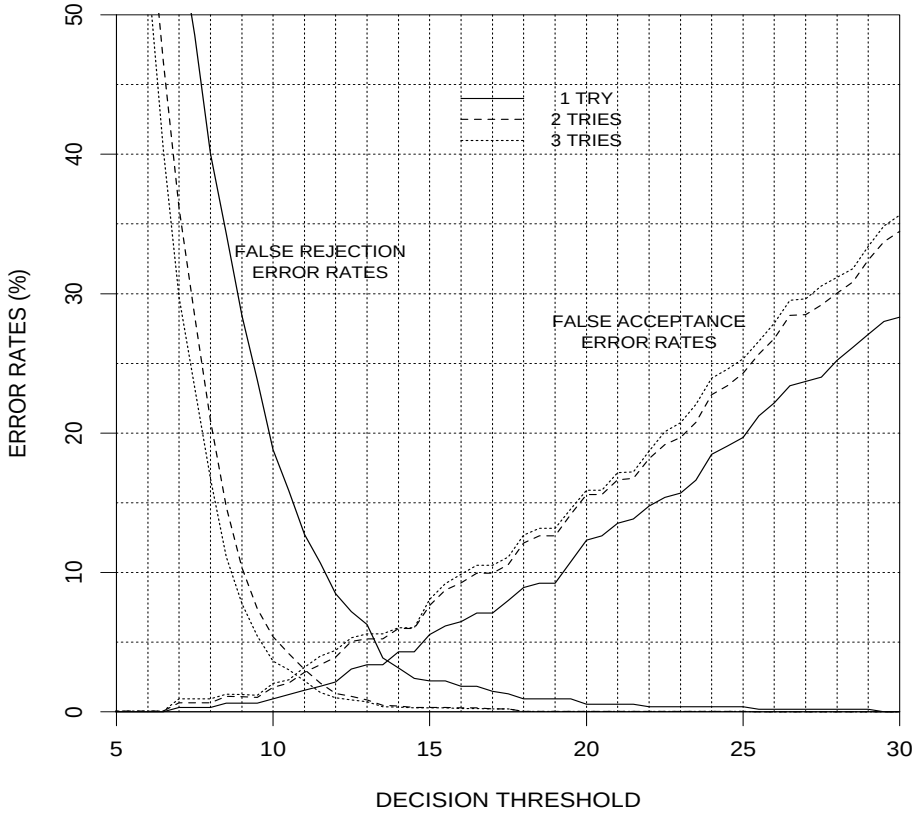

Figure 2: Separate FA & FR curves for global features only

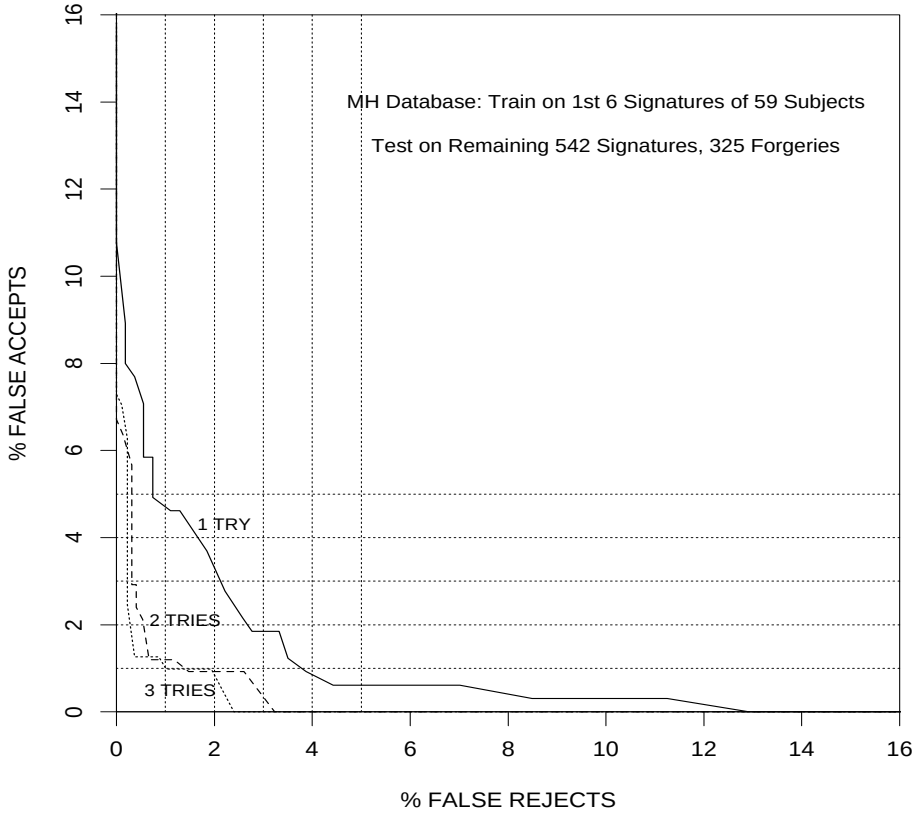

Figure 3: Error trade-off performance curves with global and local features

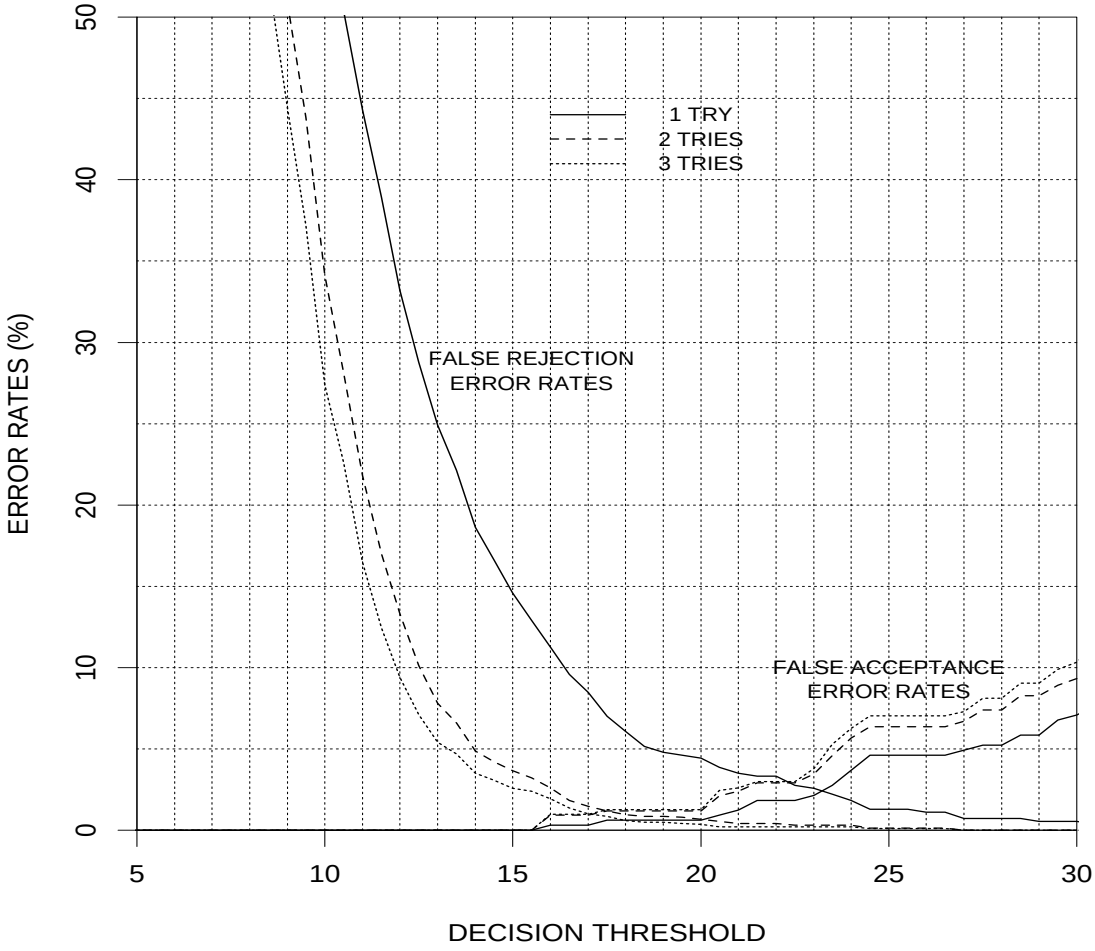

Figure 4: Separate FA & FR curves for global and local features