An *Extended-Shadow-Code* Based Approach for Off-Line Signature Verification: Part -II- Evaluation Of Several Multi-Classifier Combination Strategies

Robert Sabourin & Ginette Genest

École de technologie supérieure, Département de génie de la production automatisée
4750 Avenue Henri-Julien, Montréal (Québec), H2T 2C8, CANADA
Tel: (514) 289-8932 Fax: (514) 289-9042 e_mail: sabourin@gpa.etsmtl.ca

Abstract

In a real situation, the choice of the best representation $R(\gamma)$ for the implementation of a signature verification system able to cope with all types of handwriting is a very difficult task. This study is original in that the design of the integrated classifiers $E(x)$ is based on a large number of individual classifiers $e_k(x)$ (or signature representations $R(\gamma)$) in an attempt to overcome in some way the need for feature selection. In this paper, the authors present a first systematical evaluation of a multi-classifier-based approach for off-line signature verification. Two types of integrated classifiers based on kNN or minimum distance classifiers and 15 types of representation related to the ESC used as a shape factor have been evaluated using a signature database of 800 images (20 writers x 40 signatures per writer) in the context of random forgeries.

1: Introduction

Shape factor definition is an ongoing problem in the field of pattern recognition. This problem also applies to handwritten signature verification, therefore [1], which mainly characterized by the following: the feature vectors have a high dimensionality, the number of reference signatures already available for training is normally very low (3 to 6 in practice) and the genuine signature shape is characterized by high intra-class variability over time. Consequently, the design of a signature verification system based on a single shape factor or a single shape representation is not a trivial task.

One solution is to design a class of shape factor tailor-made for the signature verification problem [2] and to build an integrated classifier permitting the cooperation of several classifiers. Combining classifiers is not new in the field of pattern recognition and has been investigated by several authors working in the field of character recognition. Several methods have been proposed and evaluated [3], but the *voting principle* seems more appropriate for the signature verification problem because we have to design one integrated classifier per writer enrolled in the verification system.

Combining classifiers adapted to a specific shape factor permits the design of a flexible integrated classifier in the sense that the choice of the best shape factor (e.g. feature selection) is overcome; that is to say, pertinent shape factors are chosen dynamically and consequently the verification system can be adapted to handwriting.

We propose in this paper the first systematical evaluation of a multi-classifier-based approach for off-line signature verification. A class of shape factors has been defined and experimental results presented in a companion paper [2]. The experimental protocol and signature database used for the experimentation phase of this work are the same as those used for the individual evaluation of the proposed 15 signature representations $R(\gamma)$.

2: Definition of integrated classifiers $E(x)$

Following Xu et al. [3], the case investigated in this work is related to a *Type 1* cooperation problem, that is to say, the output of each classifier e corresponds to abstract information like a label j related to class C_j. Let K represent the classifiers e_k, k=1,...,K which are responsible for assigning the label j_k to the unknown observation x, i.e. the production of the event $e_k(x)=j_k$. Here, the K classifiers are all of the same type: *kNN* [Figure 1] or *minimum distance* classifiers [Figure 2]. The problem can be stated as the combination of K events $e_k(x)=j_k$ for the definition of an integrated classifier E, which will be responsible for the attribution of a final label j to the unknown observation x, i.e. $E(x)=j,\ j\in\Lambda \cup \{M+1\}$. For the *kNN* classifiers depicted in Figure 1, label j_k corresponds to class C_1 if $k_1^{(K)} \succ k_2^{(K)}$ and inversely to class C_2 if $k_2^{(K)} \succ k_1^{(K)}$; otherwise the unknown observation x is rejected ($j_k=M+1$). The architecture of a *kNN*-based integrated classifier and that of a *minimum-distance*-based integrated classifier are depicted in Figures 1 and 2 respectively. Each block R_k represents a bar mask representation $R(\gamma)$ based on the *extended shadow code* as defined and evaluated in [2]. The same learning or comparison sets ($S^{(i)}_{Learn}$ or $S^{(i)}_{ref}$) are used for each individual classifier e_k in the proposed architectures. This greatly diminishes the number of reference signatures needed for the implementation of integrated classifiers $E^{(i)}(x)$. For the sake of clarity, the notation $E(x)$ will be

0-8186-7128-9/95 $4.00 © 1995 IEEE

used in the following paragraphs as an integrated classifier assigned to a specific writer *(i)*.

The event $e_k(x)=i$ could be represented by the characteristic function

$$T_k(x\in C_i) = \begin{cases} 1, & \text{when } e_k(x) = i \ \ AND \ \ i\in\Lambda \\ 0, & \text{otherwise.} \end{cases}$$

The general formulation of the decision rule based on the *voting principle* is [3]

$$E(x) = \begin{cases} j, & \text{if } [T_E(x\in C_j) = \max_{i\in\Lambda} T_E(x\in C_i)] \geq (\alpha * K) \\ M+1, & \text{otherwise.} \end{cases}$$

where

$$T_E(x\in C_i) = \sum_{k=1}^{K} T_k(x\in C_i), \qquad i=1, \ldots, M.$$

A value of $\alpha = 0.5$ corresponds to the simple majority rule and a value of $\alpha = 1.0$ states that a decision made by an integrated classifier $E(x)$ requires the unanimity of all individual classifiers $e_k(x)$. Here, we have $|\Lambda| = M = 2$ classes.

3: Experimentation

The proposed signature verification system has been tested using a standard signature database of 800 images (40 signatures written by 20 individuals). The handwritten signatures were written in a 3x12 cm rectangle, using the same type of writing tool (a Pilot Fineliner with flexible felt tip and black ink) and sheets of white paper. Let $\boldsymbol{R}$ be the reference database related to the first 20 signatures of each writer ($|\boldsymbol{R}|=400$), and $\boldsymbol{T}$ be the test database which is related to the last 20 signatures of each writer with $|\boldsymbol{T}|=400$.

Fifteen representations $R(\gamma)$ have been defined and evaluated using a *kNN classifier with vote* and a *minimum distance classifier* [2]. For each classifier tailor-made for a representation $R(\gamma)$, the performance of the verification system is reported in terms of Type I (ϵ_1, false rejection of genuine signatures) and Type II (ϵ_2, false acceptance of random forgeries) error rates evaluated for the 20 writers. The total error rate ϵ_t of a verification system is expressed in terms of ϵ_1, ϵ_2 and $P[\omega_i]$, the a priori probabilities of classes ω_i. This probability is set to 0.5 in our case: $\epsilon_t = ((\epsilon_1 \times P[\omega_1]) + (\epsilon_2 \times P[\omega_2]))$.

The previous experiments were therefore repeated 25 times for each signature verification system and the observations of class ω_2, that is to say, the subsets of $S^{(i)}_{learn}$ and $S^{(i)}_{gen}$, were redefined randomly following our standard protocol. In the case of the *minimum distance classifier*, the effect, on the global performance of the verification system, of the choice and number of reference signatures (class ω_1) in the definition of sets $S^{(i)}_{ref}$ were also taken into account.

The mean performances of the signature verification systems (e.g. the mean total error rate $\overline{\epsilon_t}$) resulting from the performance evaluation on 25 iterations of representations R(a) to R(o) used for both types of classifiers, for a total of 5625 individual experiments, are depicted in Figures 3 and 4 (see the individual labels) with parameters $k=\{1,3,5,7,9\}$ for the *kNN* classifiers, and with $1 \leq N_{ref} \leq 10$ for the minimum distance classifiers. At first glance, it is clear that verification systems built around *kNN* classifiers outperform those based on *minimum distance* classifiers for all representations $R(\gamma)$ under study. As an example, the mean total error rate $\overline{\epsilon_t}$ varies in the (0.010% - 2.156%) range of values for the *1NN* classifier [Figure 3] and in the (0.782% - 4.524%) range of values for the *minimum distance* classifier using $N_{ref}=6$ reference signatures [Figure 4].

3.1: Experiment I

The goal of this experiment is to numerically evaluate the integrated classifiers $E(x)$ [Figures 1-2] using all 15 representations $R(\gamma)$ related to the *extended shadow code* [2]. In this first experiment, the cooperation of the $K=15$ classifiers is achieved with the help of the voting principle using the simple majority rule ($\alpha = 0.5$). Experimental results depicted in Figures 3 and 4 (solid line) show that the mean performance obtained with integrated classifiers $E(x)$ is as good as or better than the mean performance observed with the better individual classifier $e_k(x)$. For example, the mean total error rate $\overline{\epsilon_t}$ observed with representation R(n) and the *1NN* classifier is equal to 0.010% ($\sigma\pm0.015\%$) [Figure 3]. This result is almost the same as that obtained with the integrated classifier $E(x)$, which shows a mean total error of 0.023% ($\sigma\pm0.036\%$) if related dispersions in the data are taken into account. Consequently, we could say that the design of integrated classifiers $E(x)$ is a good method to employ in the implementation of signature verification systems, making it possible to overcome the difficult task of feature selection, provided that the individual performance obtained with representations $R(\gamma)$ is good enough.

3.2: Experiment II

Starting with the worst 3 representations $R(\gamma)$ related to both types of integrated classifiers under study, the experiment consists in implementing several signature verification systems with the sequential addition of the next 2 most powerful individual classifiers $e_k(x)$, that is to say, integrated classifiers $E(x)$ with K, the number of individual classifiers, taking one of the values in the set $K=\{3,5,7,9,11,13,15\}$. Using the voting principle with the simple majority rule ($\alpha = 0.5$), experiments made with both types of individual classifiers $e_k(x)$ confirm the hypothesis that the addition of better individual classifiers in the implementation of integrated classifiers $E(x)$ shows, an average, that this produces a better performance than that obtained with the best individual classifier $e_k(x)$ used in the implementation of $E(x)$ [Figures 5-6]. This is especially true when individual classifiers $e_k(x)$ are not correlated, and seems to be true when the *extended shadow code* is used as a class of shape factor. The used of integrated classifiers for the implementation of signature verification systems is particularly interesting when the performance of individual classifiers is not too high, as in the case of integrated classifiers based on *minimum distance* classifiers [Figure 6]. As an example, the experiment carried out with integrated classifiers based on representations R{a,f,g} shows a mean total error rate $\overline{\epsilon_t}$ = of 2.58% ($\sigma\pm0.37\%$), which is very powerful when compared to the mean total error rate obtained with the best individual *minimum distance* classifier based on representation R(g) which is equal to 3.83% ($\sigma\pm0.56\%$).

The second experiment consists in combining individual classifiers $e_k(x)$ in the reverse order, that is to say, starting with the best 3 classifiers, and then adding the next 2, and so on, with $K=\{3,5,7,9,11,13,15\}$. This is repeated for both types of classifiers, with the corresponding experimental results depicted in Figures 7 and 8. It is very interesting to note that the addition of less powerful classifiers did not significantly disturb the mean performance of verification systems based on both types of integrated classifiers $E(x)$. In the case of integrated classifiers $E(x)$ based on the combination of *1NN* classifiers [Figure 7], the variation in the mean total error rates $\overline{\epsilon_t}$ = could be explained by the fact that the best 2 individual classifiers based on representations R(n) and R(o) are highly correlated. In other cases, the addition of less powerful individual classifiers has always shown only a mild influence on the global performance of both types of integrated classifiers $E(x)$.

From these experiments, we can conclude that the addition of more powerful individual classifiers $e_k(x)$ (when implementing integrated classifiers) enhances the global performance of verification systems whenever they are not correlated. For the same reasons, the addition of less powerful classifiers does not destroy the global performance of integrated classifiers $E(x)$. This is very important in practice because the problem of feature selection might be overcome if the performance of individual classifiers is good enough.

3.3 Experiment III

Once the results of experiments **I** and **II** had been obtained, an attempt was made to characterize signature verification systems in terms of simple-rejection rule definitions. The numerical results obtained with integrated classifiers based on *1NN* classifiers and on the *minimum distance* classifiers using $N_{ref} = 6$ reference signatures are reported in Figures 9 and 10 respectively (see the dashed-lines labeled *a*). Using the majority rule with $\alpha = 13/15$, the performance evaluation of both verification systems corresponds to $[(\overline{\epsilon_t} = 0.004\%\ (\sigma\pm0.010\%), \overline{R_t} = 0.732\%\ (\sigma\pm0.186\%)]$ and $[\overline{\epsilon_t} = 0.13\%\ (\sigma\pm0.11\%), \overline{R_t} = 4.52\%\ (\sigma\pm0.54\%)]$ respectively. In practice, it is not always easy to select the best decision threshold α. Moreover, this scheme makes use of all 15 representations, which is a very costly solution for real implementations.

Now, the question is: Is it advantageous to use all 15 representations and an arbitrary threshold α, or to build an integrated classifier $E(x)$ based on a subset of individual classifiers $e_k(x)$ and the simple unanimity rule for the combination of classifiers? In this way, starting with the best individual classifier $e_k(x)$, an attempt was made to characterize both signature verification systems. These experimental results are represented by the solid lines (b) in Figures 9 and 10. Using the unanimity rule and the best K=5 individual classifiers, the performance evaluation of both verification systems corresponds to $[(\overline{\epsilon_t} = 0.004\%\ (0.010\%), \overline{R_t} = 0.183\%\ (0.058\%)]$ and $[\overline{\epsilon_t} = 0.05\%\ (0.06\%), \overline{R_t} = 2.88\%\ (0.51\%)]$ respectively. The same experiment was conducted, but in the reverse order, that is to say, the worst individual classifier was used first, and the next best classifiers were added sequentially. Here, we can observe that the global performance of both systems is enhanced by the addition of better classifiers, at the expense of higher rejection rates $\overline{R_t}$ (see the solid lines *c* in Figures 9 and 10).

We can therefore conclude from these experiments that the unanimity rule requires some sort of feature selection for the performance optimization of integrated classifiers. Consequently, it seems preferable to take into account a large number of individual classifiers when implementing verification systems based on the concept of integrated classifiers, and to settle decision threshold α properly.

4: Conclusion

The *extended shadow code* used as a shape factor seems discriminant enough for the signature verification problem in the elimination of random forgeries when the resolution of the bar mask array is high enough (see R(n) and R(o) in [2]). Moreover, the use of integrated classifiers permits the implementation of signature verification systems without the a priori feature selection that results in a single shape factor.This scheme also permits the design of more general verification systems tailor-made for all types of handwriting. Future work will be directed towards the evaluation of this concept on a very large signature database.

5: References

[1] Sabourin, R., Plamondon, R. and Lorette, G., "*Off-line Identification with Handwritten Signature Images: Survey and Perspectives*", in Structured Document Image Analysis, Baird,

H.S., Bunke, H. and Yamamoto, K., editors, Springer-Verlag, pp 219-234, October 1992.
[2] Sabourin, R. and Genest, G., *"An Extended-Shadow-Code Based Approach for Off-Line Signature Verification: Part -I- Evaluation of The Bar Mask Definition"*, in the 12th ICPR, Jerusalem, Israel, October 9-13, 1994, pp450-453.
[3] Xu, L., Krzyzak, A. and Suen, C.Y., *"Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition"*, IEEE Trans. on Systems, Man, and Cybernetics, Vol. 22, No. 3, May/June 1992, pp. 418-435.

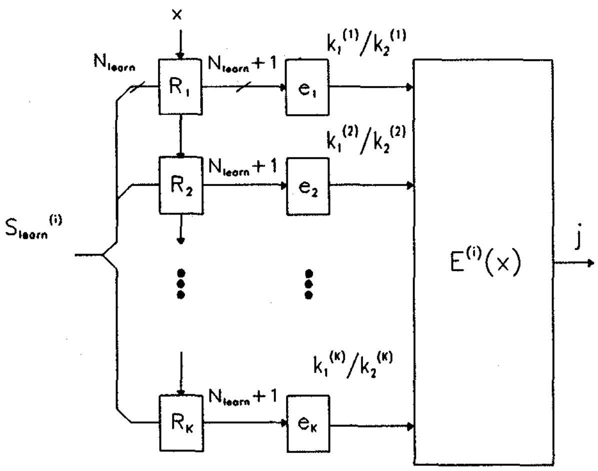

Figure 1

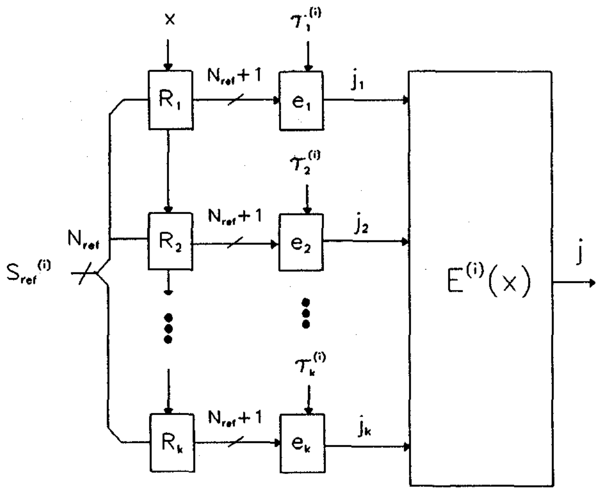

Figure 2

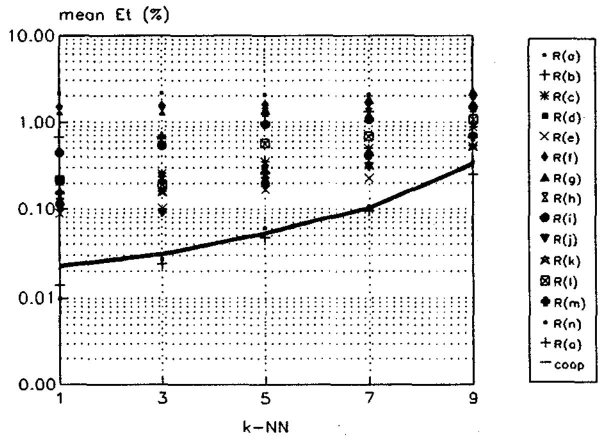

Figure 3

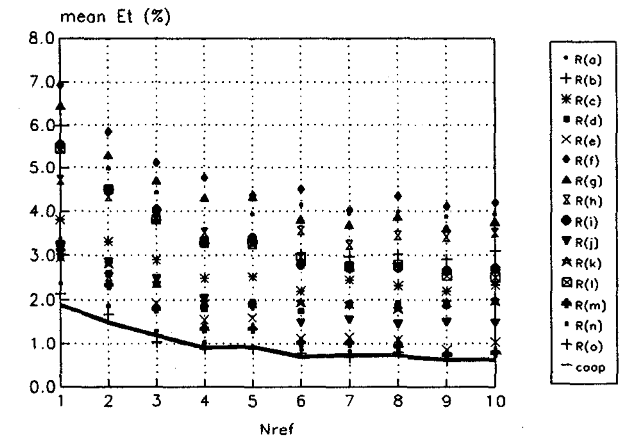

Figure 4

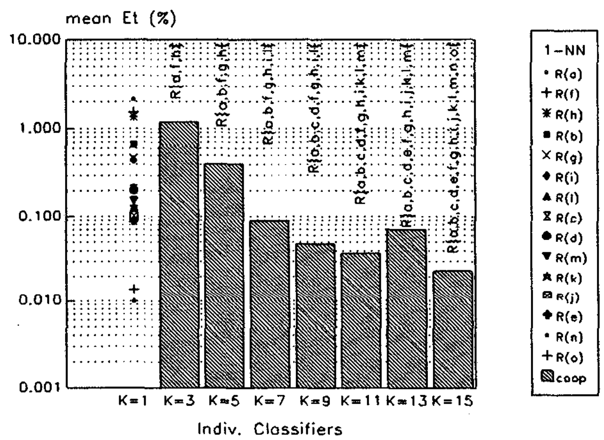

Figure 5

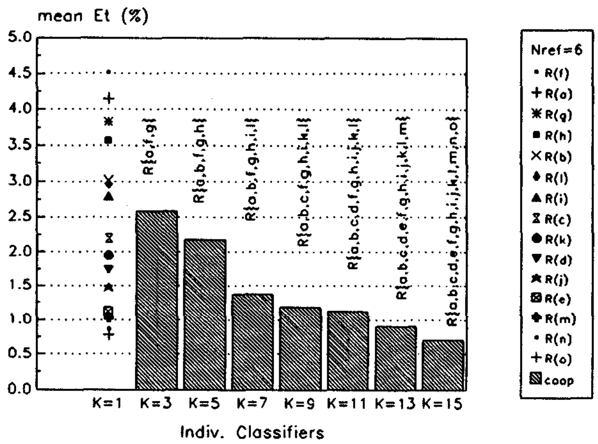

Figure 6

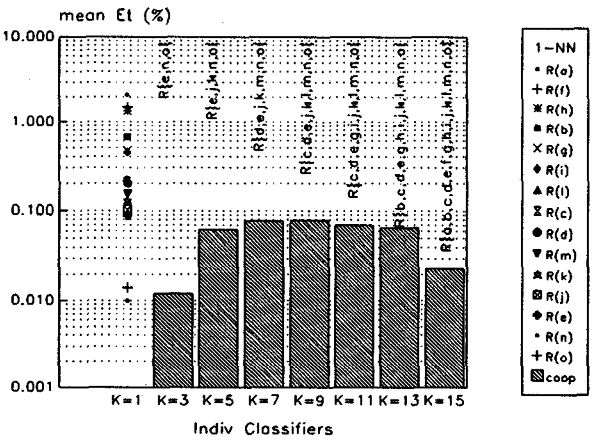

Figure 7

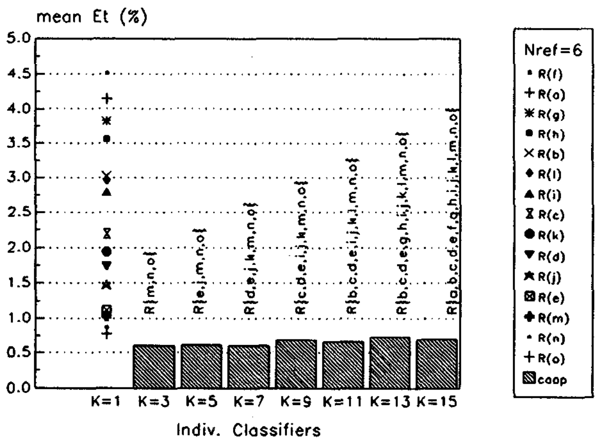

Figure 8

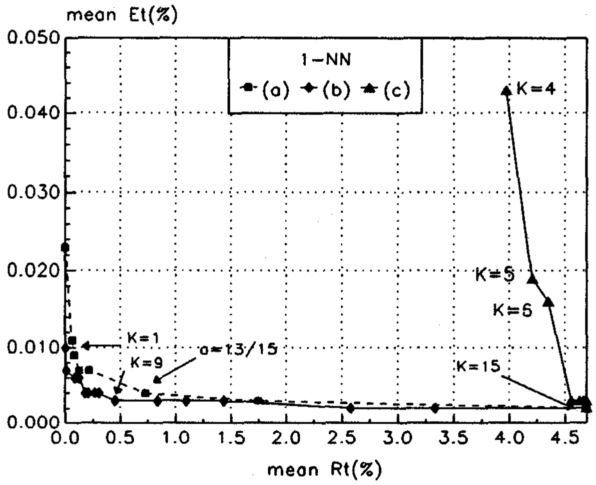

Figure 9

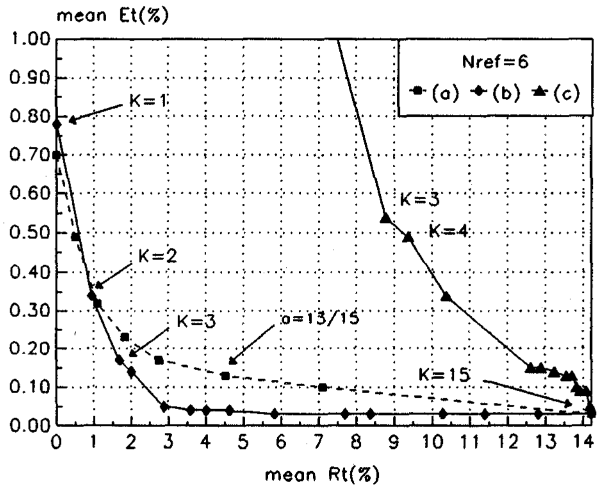

Figure 10