Combining Fisher Discriminant Analysis And Probabilistic Neural Network for Effective On-Line Signature Recognition

Souham Meshoul and Mohamed Batouche

Center Of Excellence in Information Assurance, Department of Software Engineering, CCIS, King Saud University, Riyadh, Kingdom of Saudi Arabia
{smeshoul, batouche}@ksu.edu.sa

ABSTRACT

The advent of new technologies enables capturing the dynamic of a signature. This has opened a new perspective for the possible use of signatures as a basis for an authentication system that is accurate and trustworthy enough to be integrated in practical applications. Automatic online signature recognition and verification is one of the biometric techniques being the subject of a growing and intensive research activity. In this paper, we address this problem and we propose a two-stage approach for personal identification. The first stage consists in the use of linear discriminant analysis to reduce the dimensionality of the feature space while maintaining discrimination between user classes. The second stage consists in tailoring a probabilistic neural network for effective classification purposes. Several experiments have been conducted using SVC2004 database. Very high classification rates have been achieved showing the effectiveness of the proposed approach.

1. INTRODUCTION

Biometric systems play an important role in the field of information security as they are intensively required for user authentication. They are commonly used to address the problem of person identification using human traits. Biometric systems can be first classified according to the type of human traits used to conduct the identification process. Human traits can be physiological like fingerprint, hand geometry and iris patterns or behavioral like Signature dynamics. Adoption of a particular biometric depends upon user preference and attitude, practicality issues, accuracy and level of security required and also technological issues. A comprehensive survey on biometrics can be found in [1].

Signature has been used to authenticate persons in financial and legal transactions since a long time ago. It is the most socially accepted and least controversial of all biometrics [1]. The need for signature analysis automation has entailed a proliferation of methods, tools and systems. Automatic signature recognition and verification (ASRV) can be applied to numerous applications ranging from banking applications, access control to e-commerce and e-government to name just a few.

Like other biometrics, automatic signature analysis addresses two main questions: "Is the person really who she claims to be?" and "Who is the owner of the signature?" The first one refers to signature verification in which the signature has to be identified as genuine or a forgery whereas the second one refers to signature recognition in which the identity of the signer has to be determined. Depending on the way signatures are represented and more precisely on the availability of time related information, methods for ASRV fall into two broad categories namely off-line methods and on-line methods [2-4]. Despite the wide acceptance of signature and its resistance to impostor attempts, ASRV remains a challenging task. This can be explained in one hand by the signature' variability. Signatures evolve by time and they are influenced by the physical and emotional conditions.

This work joins the numerous attempts to develop effective methods for ASRV. Our focus is on signature recognition. We describe in this paper a method that combines a linear discriminant analysis (LDA) and Probabilistic Neural Networks for effective identification. Basically, the problem to be solved is to identify the owner of a signature given an input signature and a set of reference signatures in database. Each signature is defined as an ordered list of points having several attributes like spatial coordinates, pen pressure, azimuth angle of the pen with the digitizing table and the altitude of the pen with the device.
In the following, sections 2 and 3 are devoted to the description of linear discriminant analysis and probabilistic neural networks. Section 4 describes the proposed framework for signature recognition. In section 5, we report on the conducted experiments and the obtained results. Finally, conclusions and perspectives are drawn.

2. FISHER DISCRIMINANT ANALYSIS

The goal of Fisher Discriminant Analysis also called Linear Discriminant Analysis (LDA) is to find an efficient way for maximum discrimination between classes in addition to dimensionality reduction. In the following, we explain its use in our context. LDA searches for Fisher discriminant vectors which group signatures of the same user and separates signatures of different users. Signatures are projected from N-dimensional space (N is the number of feature vector) to C-1 dimensional space (C corresponds to the number of users). More formally, the basic idea of LDA is to determine those vectors so that the Fisher Index defined as the ratio (det|Sb|/det|Sw|) is maximized.

Sw is the within class scatter matrix which measures the amount of scatter between signatures of the same user and is given by:

\[ S_w = \sum_{j=1}^{C} \sum_{i=1}^{N_j} (x_i - \mu_j)(x_i - \mu_j)^T \]  

(1)

where \( N_j \) is the number of signatures for user \( j \), \( x_i \) is a signature vector, \( i \) denotes the \( i \)-th signature vector for user \( j \), and \( \mu_j \) refers to the mean vector of class \( j \).

Sb is the between class scatter matrix which measures the amount of scatter between different users and is given by:

\[ S_b = \sum_{j=1}^{C} (\mu_j - \mu)(\mu_j - \mu)^T \]  

(2)

where \( \mu \) refers to the mean vector of all signatures.

LDA produces an optimal linear transformation which maps the input space to the projection space. Assuming that \( S_w \) is non-singular, the basis vectors of this function correspond to the first (C-1) eigenvectors with the largest eigenvalues of \( S_w^{-1}S_b \).

3. PROBABILISTIC NEURAL NETWORK

Probabilistic Neural Networks (PNN) introduced by Donald Specht in 1988 [5, 6] are a kind of radial basis network suitable for classification problems. Contrary to backpropagation, the PNN training process is very fast and converges to a global optimum.

PNNs are based upon the Bayes strategy for decision making and Parzen window estimation [6]. Using the latter technique, the probability density functions (PDF) required by Bayes’ theory can be easily determined.

Suppose \( n \) is the number of training samples, \( m \) is the feature space dimension, and \( x_i \) is the \( i \)-th training sample for a certain class (user 1 for example), then the Parzen estimate of the PDF for class 1 is [7]

\[ F_1(x) = \frac{1}{(2\pi)^{m/2}\sigma^m} \sum_{i=1}^{n} \exp\left[-\frac{(x-x_i)^T(x-x_i)}{2\sigma^2}\right] \]  

(3)

The \( \sigma \) is the “smoothing parameter” which represents the single free parameter for this algorithm.

As shown in figure 1, PNN architecture is like multilayered feedforward network with four layers: an input layer, a pattern layer, a summation layer, and an output layer.

![Probabilistic neural network architecture](image)

\( \text{Figure 1. Probabilistic neural network architecture.} \)

The input layer merely distributes features of input vector to the pattern layer. The latter, with one neuron for each training sample, is fully connected to the input layer. The weights \( w_i \) for these connections are set equal to the different training patterns. After the input pattern \( x \) and the weights \( w_i \) are normalized to unit, the pattern layer computes distances from the input vector to the training samples and produces a vector whose elements indicate how close the input is to a training sample. Each neuron \( j \) performs a radial transfer function which can be simplified as follows:

\[ \exp\left( \frac{x^Tw_j - 1}{\sigma^2} \right) \]  

(4)

The pattern layer neurons representing the same class are connected to the same summation unit. The summation layer, with one neuron for each class, sums the outputs of pattern layer for each class and produces outputs which represent the probabilities that the vector belongs to each class. Finally, the output layer picks the maximum of these probabilities, and provides the target class for the input vector given by:

\[ \text{targetclass}(x) = \arg \max_{j} \left( \frac{1}{n} \sum_{i} \exp((x^Tw_i - 1)/\sigma^2) \right) \]  

(5)
4. SIGNATURE RECOGNITION USING THE PROPOSED FRAMEWORK

In this section, we describe the framework used for online signature recognition. First, an offline step is performed as a training phase to construct the PNN model. Then the recognition process is applied in an online step.

4.1. Training phase

The training process acts in two main stages: linear discriminant analysis and PNN model construction. For this purpose, first input signatures need to be processed in order to meet PNN construction requirements like getting signatures with the same size. After that LDA is applied to reduce the dimensionality of the feature space while maintaining discrimination between classes. Then all training signatures are projected on Fisher space and used for PNN construction.

Feature processing and extraction is one of the critical parts of any ASRV system. In our work, we make use of dataset available at SVC2004 online signature database [8]. For each signature, the provided dynamic features are x-coordinate, y-coordinate, Time stamp, Button status, Azimuth, Altitude, and Pressure. Figure 2 shows one of the existing signatures.

For every signature, the invariant dynamic parameters after sampling and normalization: G (distances to the center of gravity), pressure, azimuth and altitude are recorded as the result of the preprocessing step. Then the whole dataset is used as input to linear discriminant analysis (LDA) component.

Every signature is represented by the vector of all selected features. Operating in the manner described in section 2, LDA provides Fisher discriminants which are linear combination of original signatures features. These ones allow to best discriminate between users by grouping signatures of the same user and separating signatures of different users. The training signatures are then projected onto Fisher space leading to a reduced set of feature vectors (Fisher signatures) which are used to train PNN in order to derive its model.

The architecture of our PNN model for signature classification is defined as follows. The input layer corresponds to the features of Fisher signatures. The pattern layer comprises all training signatures. The summation layer contains as many neurons as number of users (classes). The output layer is a competitive layer which identifies the target user based on probabilities given by summation layer. The PNN model evolves according to the dynamic described in section 3 in order to set connections weights.

4.2. Recognition phase

Once the PNN model is constructed, it can be used for recognition purposes. During an online step, the user identification is carried out as follows. The input signature is preprocessed to get its invariant dynamic parameters as described before, and then it is projected on Fisher space. The resulting feature vector is used to feed the PNN model in order to identify the corresponding user.

5. EXPERIMENTAL RESULTS

Several experiments have been conducted to assess the performance of the proposed approach. The 40 sets of signature data provided in SVC2004 have been used. Every set corresponds to a user and includes 20 signatures. Therefore, the dataset includes 800 signatures. The experimental protocol has been carried out in the following manner.

A first set of experiments has been conducted to determine the suitable signature size during sampling step. Best results were obtained for a sampling with 70 points. Therefore, each signature is defined by 280 (4 parameters ×70). This set of features has been reduced to 39 features after using Fisher discriminant analysis. Other experiments have been conducted in order to identify the most relevant features among these 39 Fisher discriminants. Better discrimination has been obtained for various values ranging from 20 to 35. We kept only the 20 first Fisher discriminants in order to get a more efficient PNN model.
The second set of experiments has been devoted to test the recognition abilities. Different sizes of training set have been used. The corresponding recognition rates are gathered in Table 1. Best results were obtained when using more than 60% of the dataset for training.

**Table 1.** Recognition rates for different size of training set.

<table>
<thead>
<tr>
<th>Training set</th>
<th>Test set</th>
<th>LDA + PNN (recognition rate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40% (320)</td>
<td>60% (480)</td>
<td>69.58%</td>
</tr>
<tr>
<td>50% (400)</td>
<td>50% (400)</td>
<td>87.75%</td>
</tr>
<tr>
<td>60% (480)</td>
<td>40% (320)</td>
<td>96.25%</td>
</tr>
<tr>
<td>70% (560)</td>
<td>30% (240)</td>
<td>99.58%</td>
</tr>
<tr>
<td>80% (640)</td>
<td>20% (160)</td>
<td>100%</td>
</tr>
</tbody>
</table>

As a third set of experiments and for sake of comparison, three strategies have been investigated: 1. recognition using LDA with Euclidean distance, 2. recognition using only PNN and 3. recognition using LDA combined with PNN. Obtained results using 80% of dataset for training are shown in Table 2. Best results were achieved with the combination LDA and PNN. In our experiments, only one tunable parameter needs to be set: the smoothing parameter used in PNN. Best results were obtained for the values ranging from 0.03 to 0.05.

**Table 2.** Recognition rates for different strategies.

<table>
<thead>
<tr>
<th></th>
<th>LDA with Euclidean Distance</th>
<th>PNN</th>
<th>LDA + PNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set</td>
<td>98.12%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Test set</td>
<td>95.63%</td>
<td>98.12%</td>
<td>100%</td>
</tr>
</tbody>
</table>

6. CONCLUSIONS

In this paper, we described a new approach for automatic signature recognition. Using invariant dynamic parameters with normalization and sampling namely distance to center of gravity (G), pen pressure, azimuth and altitude, the proposed method makes use of Fisher discriminant analysis to reduce the dimensionality of feature space by identifying the representative features that best discriminate between classes. Projections of signatures on the obtained Fisher space are used as inputs to a probabilistic neural network dedicated to effective classification purposes. The proposed approach LDA with PNN is very attractive since it has many interesting characteristics. It uses only one tunable parameter for which the value is easily set experimentally. There is no need to use any threshold for recognition processes. The training of PNN is very fast and converges to global optimum (no local minima issue). The obtained results show that the proposed approach for online signature recognition competes with existing techniques.

REFERENCES


