Multiple-Biometric Fusion Methods using Support Vector Machine and Kernel Fisher Discriminant

Nayoung Woo and Hakil Kim
Biometrics Engineering Research Center
Graduate School of Information Technology & Telecommunication
Inha University, Incheon, Republic of Korea
e-mail: {nywoo, hikim}@vision.inha.ac.kr

Abstract: This paper proposes multiple-biometric fusion methods, using Support Vector Machine (SVM) and Kernel Fisher Discriminant (KFD) in order to improve recognition accuracy. The proposed SVM and KFD are non-linear classification methods using Radial Basis Function (RBF) as a kernel function. The RBF kernel is effective with respect to the input data distribution. Experiments have been conducted on NIST BSSR1 (Biometric Scores Set – Release1) data set using SVM, KFD and weighted sum methods, and their performance of multiple biometric fusion has been presented by the HTER (Half Total Error Rate) and the ROC (Receiver Operating Characteristic) curves. The experimental results demonstrate that the non-linear fusion methods provide higher verification performance than linear fusion methods.

Keywords: multiple-biometric, kernel fisher discriminant, fusion, performance evaluation, classification, support vector machine, weighted sum, BSSR1

1. Introduction

Biometrics refers to the automatic identification of an individual using certain physical (i.e. fingerprint, face, iris) or behavioral (i.e. gait, typing habits, signature) traits associated with the individual. As a personal authentication method, biometrics provides higher security and convenience than conventional methods such as using passwords or tokens [1]. Due to these key reasons, security methods are gradually migrating from passwords or keys to biometric methods. However, biometric systems using single biometrics for authentication have various limitations. Single biometric systems are affected by noisy data, performance limitation, circumvention via spoofing, and non-universality.

In order to overcome the technical limitations residing in single biometrics, multiple-biometrics has been studied and various techniques are being developed [2-3]. In general, the performance of a multiple-biometric system is more reliable and provide higher verification rates due to the presence of multiple, independent biometric information for authentication. There are three levels of fusion for combining biometric systems, fusion at the feature extraction level, matching score level and decision level. Among the three levels, fusion at the matching score level is usually preferred because of its readiness and effectiveness when developing fusion methods, relatively less work has been carried out on testing and evaluating the performance of multiple-biometric systems.

Multiple-biometric systems based on score-level are dichotomized according to a combination scheme. First, matching scores of individual systems are combined in order to form a single score and then the system calculates whether the claimed user is genuine or imposter by making a comparison between the fused score and a system threshold. Second, matching scores of individual systems are constructed as a vector, and the system classifies these into two classes, genuine and imposter. The latter is implemented by pattern classification algorithms such as neural network, and discriminant functions [4].

This paper presents the fusion method for multiple-biometric systems using SVM and KFD in order to improve the performance of recognition accuracy. This study presents experimental results based on a large-scale score database. An analysis of the experimental results using the three fusion methods is
presented. The experimental results demonstrate that the non-linear fusion method provides greater recognition accuracy performance over the weighted sum method.

2. Score-level Fusion in Multiple-Biometric System [3]

In score level fusion, each single system provides matching scores indicating the similarity between a pair of biometric templates. These matching scores can then be combined to improve matching performance. The multiple-biometric system based on matching score is generally implemented as presented in figure 1. There are three processes for construction of a multiple-biometric system based on matching score. The first process is Score normalization which maps scores into a domain where they are associated with a common meaning in terms of biometric performance, the second is Score fusion which combines multiple normalized scores to a single score or a vector, and the third is Decision rule by which a pair of corresponding biometric traits are from genuine or imposter.

In the context of verification, there are two distinct approaches to score level fusion. The first approach is to formulate verification as a classification problem, while the second approach is to treat verification as a combination problem. The first method classifies feature vectors constructed by augmenting the matching scores output from the individual systems into one of two classes: genuine or imposter. The latter method combines matching score output from the individual systems into a single score, and then makes a decision by taking the threshold of the multiple-biometric system into account.

Figure 1: Process of multiple-biometric system based on matching score [3]

3. Non-linear Fusion Algorithms

3.1 Support Vector Machine [5, 8]

The Support Vector Machine (SVM) is one of the most popular techniques in the field of statistical learning theory. In binary classification, the goal of statistical learning theory is to separate the two classes. Classical learning approaches are designed to minimize the empirical risk, i.e., the classification error of the training set. However, the SVM is based on the principle of Structural Risk Minimization in which better generalization abilities such as performance on unknown test data are achieved through minimization of the upper bound of the generalization error.

Consider the problem of separating a set of training vectors \( x_i \) belonging to different classes \( y_i \in \{-1,1\} \). The linear SVM is defined by Eq. (1) representing a hyper-plane which separates the vectors.

\[
\vec{w} \cdot \vec{x} + b = 0
\]  

(1)

The linear SVM expands into non-linear formulation. In the case where the decision function is not a linear function of the data, the data will be mapped from the input space into a high dimensional
feature space through a non-linear transformation. The non-linear transformation is performed implicitly through a kernel function in order to reduce the dimension of the problem. The kernels must satisfy various constraints in order to achieve validity. For binary classification, various kernel functions such as the RBF kernel \( k(x, y) = \exp\left(-\frac{|x - y|^2}{2\sigma^2}\right) \), polynomial kernel \( k(x, y) = (x \cdot y)^d \), and linear kernel \( k(x, y) = x \cdot y \) are used. An example of non-linear classification using SVM is presented in figure 2.

![Figure 2: Example of non-linear classification using SVM](image)

3.2 Kernel Fisher Discriminant Analysis [6-9]

Fisher’s discriminant is the most well known method for classification. Fisher’s idea is to search for a direction, \( \omega \) that separates the class means effectively (when projected onto the found direction) while achieving small variance around these means. The quantity measuring the difference between the means is named between-class variance and the quantity measuring the variance around these class means is named within-class variance. The goal is to find the direction that maximizes the between-class variance while simultaneously minimizing the within-class variance.

The Fisher linear discriminant is the optimal decision for two Gaussian distributions with equal covariance structure. Regardless the fact that Fisher’s discriminant often yields useful results even when this assumption is violated, the basic limitation is that the discriminating direction is linear. In order to increase the performance, non-linear directions are first found by mapping the data non-linearly into some feature space \( F(\Phi : \mathcal{X} \rightarrow F) \) and computing Fisher’s linear discriminant, thus implicitly yielding a non-linear discriminant in input space.

In order to find the linear discriminant in \( F \), the following criterion function must be maximized:

\[
J(\omega) = \frac{\omega^T S_B^\Phi \omega}{\omega^T S_W^\Phi \omega}
\]

where \( S_B^\Phi = (m_1^\Phi - m_2^\Phi)(m_1^\Phi - m_2^\Phi)^T \) and \( S_W^\Phi = \sum_{i=1,2} \sum_{x \in D_i} (\Phi(x) - m_i^\Phi)(\Phi(x) - m_i^\Phi)^T \) are the between-class matrix and within-class matrix, respectively, and \( m_i^\Phi = \frac{1}{|D_i|} \sum_{j=1}^{n_i} \Phi(x_j) \) are the mean vectors after transformed by a kernel function \( \Phi \).

\( F \) can be extremely high-dimensional that it will be impossible to solve directly. To overcome this limitation, the Kernel Fisher Discriminant makes use of the kernel trick. Instead of mapping the data
explicitly, a formulation of the algorithm which uses only dot-products \((\Phi(x) \cdot \Phi(y))\) from the training patterns is used. This can be achieved using Mercer kernels. These kernels \(k(x,y)\) compute a dot-product of some feature space \(F\), i.e., \(k(x,y) = (\Phi(x) \cdot \Phi(y))\). The kernel functions can use RBF or polynomial kernels.

In order to solve Fisher’s problem in a kernel feature space \(F\), a formulation which makes use of the training samples only in terms of dot products is required. It can be proven that the solution \(\omega \in F\) of (2) can be expanded as

\[
\omega = \sum_{i=1}^{l} \alpha_i \Phi(x_i), \quad \alpha_i \in \mathbb{R}
\]

(3)

And the coefficient \(\alpha_i\) can be found by solving a quadratic programming problem. The projection of a new pattern \(x\) onto \(\omega\) is given by

\[
(\omega \cdot \Phi(x_i)) = \sum_{i=1}^{l} \alpha_i k(x_i, x)
\]

(4)

4. Experimental Results

• Database (Biometric Scores Set – Release 1) [10]

BSSR1 is a set of raw output similarity scores from two face recognition systems and one fingerprint system, operating on frontal faces, and left and right index live-scan fingerprints, respectively. This includes true multimodal score data, i.e., similarity scores from comparisons of faces and fingerprints of the same individuals. The data is intended to permit interested parties to investigate a range of outstanding statistical problems related to biometrics. BSSR1 contains three partitions. Set 1 is comprised of two face scores and one fingerprint score from a population of 517 individuals. Set 2 is comprised of two fingerprint scores from 6000 individuals. Set 3 contains scores from two face systems consisting of images from 3000 individuals. The matching scores of every set are obtained by full cross-comparison.

• Normalization(z-score)

The z-score normalization method is used for score transformation into a common domain. This is the most commonly used score normalization technique using the mean and standard deviation of the distribution of given data. This scheme can be expected to perform well if prior knowledge of the average score and the score variations of the system are available. If the score data is Gaussian distributed, z-score normalization retains the input distribution at the output. This is due to the fact that the mean and the standard deviation are optimal location and scale parameters only for a Gaussian distribution. The normalized score \(S'\) is given by

\[
S' = \frac{S - \mu}{\sigma}
\]

(8)

where \(\mu\) and \(\sigma\) is the mean and standard deviation of the data set \(S\), respectively.

• Evaluation Tool (Multiple-Biometric Testing Tool)

Multiple-Biometric Testing Tool (MBTT) has been implemented by the authors for analysis of experimental results and performance evaluation of score level fusion algorithms. MBTT is developed using Matlab ver.7 by utilizing various built-in mathematical and graphical functions. This tool can be used for performance evaluation of score fusion algorithms with a graphic user interface. Figure 3 shows an example of running MBTT and the results.
The experiments are performed with three fusion methods, linear weighted sum, SVM Table and KFD methods. From BSSR1 Set1 data, four combinations of multi-modal data sets are constructed, Face(C)-Fingerprint(LI), Face(C)-Fingerprint(RI), Face(G)-Fingerprint(LI), and Face(G)-Fingerprint(RI), where C and G represent face recognition systems and LI and RI stand for left index and right index, respectively. In the classification experiments, each data set is divided into a training set and a test set with the portions of 10% and 90%, respectively. The experimental results provide ROC curves and HTER(Half Total Error Rate=(FAR+FRR)/2). Figure 4 shows the ROC curves of four combinations of multi-modal data sets from BSSR1-Set1, and Figure 5 compares FRR and FAR when HTER is minimal. The results demonstrate that the SVM and KFD methods provide lower HTER than the weighted sum method when classifier is applied as optimal. However, the error rates of non-linear methods are more unstable than those of weighted sum method. And, the optimal parameters for non-linear classifier can not be found easily.
Figure 5: Experimental results of the BSSR1 database with fusion algorithms

5. Conclusions

In this paper, various multiple-biometric schemes have been implemented using the SVM and KFD for the purpose of improving the performance of recognition. They have been tested over NIST BSSR1-Set1 data, and the experimental results demonstrated. The experimental results demonstrate that the SVM and KFD methods provide low error rates when classifier is applied as optimal. However, the performance of non-linear methods is more unstable than that of weighted sum method. This research demonstrates that non-linear methods can be a good classification algorithm for multiple-biometric system even though few considerable problems exist such as how to find an optimal parameter, how to make a stable performance.

Acknowledgements

This work was supported in part by the Biometrics Engineering Research Center, KOSEF.

References