Ensemble of multiple Palmprint representation

Loris Nanni *, Alessandra Lumini

DEIS, IEIIT – CNR, Università di Bologna, Viale Risorgimento 2, 40136 Bologna, Italy

A R T I C L E I N F O

Keywords:
Biometric verification
Palm authentication
Gabor filters
Discrete cosine transform
Multi-matcher

A B S T R A C T

In this work, a new approach for personal authentication using palm image is presented. We design three ensembles of matchers which employ different feature representation schemes of the images: discrete cosine coefficients; invariant local binary patterns; Gabor filters. Each ensemble is obtained by varying the features used to train their matchers. Experimental results confirm that the three methods give complementary information which has been exploited by fusion rules. Finally, we combine our Palm based method system with other biometric characteristics that can be extracted from the hand (middle finger, ring finger, hand geometry), obtaining a further improvement of the performance.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

The several approaches proposed in the literature for Palmprint verification have revealed that the Palmprint contains mainly three types of information: texture information; line information; appearance based information. For this reason the available methods for Palmprint authentication can be divided into three categories on the basis of the extracted features (Kumar & Zhang, 2005):

- texture-based approaches, e.g., Gabor filters (Kong, Zhang, & Li, 2003; Zhang, Kong, You, & Wong, 2003), discrete cosine coefficients (Kumar et al., 2004), Wavelets (Zhang & Zhang, 2004);
- line-based approaches, e.g., Line matching (Zhang & Shu, 1999), line detection (Kumar, Wong, Shen, & Jain, 2003);

Most of the papers proposed in the literature for Palmprint verification are based on a single Palmprint representation, and this fact can be considered as a bottleneck for the performance. An ideal Palmprint verification system should be based on the fusion of several Palmprint representations (Kumar, Zhang, & Kamel, 2006; Kumar & Zhang, 2005; Kumar et al., 2004).

In Kumar and Zhang (2005) from each palm image three feature vectors are extracted: Gabor filters, line detectors, and principal component analysis (PCA). Three different matchers are trained each using a different feature vector and then combined at the score level.

In other papers (e.g., Kumar et al., 2003, 2004, Kumar, Wong, Shen, & Jain, 2006; Li, Qiu, Sun, & Wu, 2004, 2006) the authors try to improve the performance combining different biometric characteristics that can be extracted from a given hand. In Kumar et al. (2006) The Palmprint and the handshape images are used to extract salient features which are combined to improve the verification performance.

This paper presents a new approach for personal authentication using Palm images. The proposed method attempts to improve the performance of each Palmprint-based verification system using an ensemble of matchers (Lumini & Nanni, 2006), finally the three ensembles (each based on a different feature extraction method) are combined together.

Our idea for building the ensemble of classifiers is simple: from a Palm image we extract other five sub-images and from each new image we extract five different feature vectors for each of the three feature extraction methods tested in this paper, the final score is obtained combining the scores from these several Palmprint representations.

Moreover, we combined our Palm matchers with other biometric characteristics that can be extracted from the hand (middle finger, ring finger, hand geometry), without requiring acquisition at different resolution (as fingerprint (Kumar et al., 2006)) or to ask the users to use different biometric sensors ( Kittler et al., 2002). Our experiments on an image database of 100 users achieve promising results and suggest our fusion of matching scores may improve the performance.

The paper is organized as follows: in Section 2 our approach is presented, in Section 3 the experimental results are discussed, finally, in Section 4 some concluding remarks are given.

* Corresponding author.
E-mail address: lnanni@deis.unibo.it (L. Nanni).
2. System proposed

The proposed biometric authentication system, whose block-diagram is shown in Fig. 1, is composed by three steps: localization and image pre-processing, feature extraction, classification and fusion. The hand image is taken using a low-cost scanner, then the area of interest is extracted and processed to reduce the lighting effects. We extract the palm from the hand image using the datum points obtained by a method similar to that used in Li et al. (2004), considering the $x$ coordinate of the hand contour which contains enough information of the location of datum points. The extracted palm image is resized to dimension $100 \times 100$. From this image we extract five overlapped sub-images of dimension $80 \times 80$ (see Fig. 1). From each sub-image $I_i (i=1, \ldots, 5)$ and from the whole palm image $I_0$, we extract three different feature vectors, one for each of the three feature extraction methods tested in this paper: discrete cosine transform ($DCT$), invariant local binary patterns ($LBP$), Gabor filters ($FG$).

Fig. 1. System proposed.

Fig. 2. Hand geometry features, the length of the green lines that link two green balls are the extracted features.
Let $A$ be a feature extraction approach ($A \in \{\text{DCT}, \text{LBP}, \text{FG}\}$), the $i$th feature vector $f_i(A)$ is built by the first $K_A$ features with higher variance among those extracted by the feature extraction method $A$ in the $i$th sub-image.

Then for each feature extraction method $A$, a pool of 26 two-class classifiers are trained (the two-classes are genuine/impostor (Lumini & Nanni, 2006a)): the first matcher ($M_{0,0}$) is trained by the Euclidean distance among the features extracted from the whole images, the other 25 matchers $M_{i,j}$ ($i,j = 1,\ldots,5$) are trained considering the Euclidean distance between the features extracted from the sub-images $I_i$ where the feature vectors are built by the first $K_A$ features with higher variance among those extracted by the feature extraction method $A$ in the $j$th sub-image. Finally all the matchers ($26 \times 3$) are combined (Kittler et al., 2002; Roli, Raudys, & Marcialis, 2002) for obtaining a final score.

Our idea is simple to obtain a good ensemble of classifiers it is very important to have a good trade-off between diversity and accuracy of the single classifiers (Rodriguez, Kuncheva, & Alonso, 2006); to this aim we consider slightly perturbated versions of the input image (the 5 sub-images) (Lumini & Nanni, 2006a, 2006b) and we adopt different feature extraction approaches. Please note that the features with higher variance extracted by the sub-image $I_i$ are similar but not equal to the features with higher variance extracted by $I_j$, in this way we obtain classifiers with good performance and enough diversity to be combined and hence our ensemble improves the performance obtained by the matcher trained using only the features extracted from the whole Palm Image.

2.1. Feature extraction: discrete cosine transform

The discrete cosine transform (DCT) is a separable linear transformation; that is, the two-dimensional transform is equivalent to a one-dimensional DCT performed along a single dimension followed by a one-dimensional DCT in the other dimension.

The main interesting property of the DCT (Nanni & Lumini, 2005; Pan, Rust, & Bolouri, 2000) is that the salient information exists in the coefficients with low frequencies, in this way most DCT components are typically very small in magnitude. Take out the small coefficients from the DCT representation we introduce only a small error in the reconstructed images.

We retain the first $K$ coefficients with higher variance.

2.2. Feature extraction: invariant local binary patterns

The invariant local binary patterns (Ojala, Pietikainen, & Maenpaa, 2002) is grey-scale and rotation invariant texture operator. This feature extraction method is robust in terms of grey-scale variations and in Ojala et al. (2002) is shown that it discriminates a large range of rotated textures efficiently. This operator is, by definition, invariant against any monotonic transformation of the grey scale. Moreover, it is possible to achieve the rotation invariance
incorporating a fixed set of rotation invariant patterns (see (Ojala et al., 2002) or the matlab code\textsuperscript{2} for details).

We extract the invariant local binary patterns histogram of 10 bins, 18 bins and 26 bins from each sub-window of dimension $25 \times 25$ of the whole image. To avoid a large number of sub-windows the step of translation is of 10 pixel.

2.3. Feature extraction: Gabor filters

In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave (Daugman, 1988). Image features are extracted by convolving the whole image with a bank of 16 Gabor filters with different wavelengths and orientations (standard deviation $\sigma = 10, 20, 30, 40$ and orientation $= 0, \frac{\pi}{2}, \frac{3}{4}\pi$). To reduce the computational complexity of convolution process images and filters are resized to $8 \times 8$. The Gabor filters are implemented as in Franco, Lumini, Maio, and Nanni (2006).

3. Experimental results

The experiments utilize inkless hand images obtained from digital Camera (Kumar et al., 2003). The database contains seven samples from each user, for 100 users. In our experiments we have included $N$ ($N$ has been varied among 1 and 3 in our experiments) randomly selected samples per user in the training set and the remaining $7-N$ in the test set, then, in order to minimize the possible misleading results caused by the training data, the results have been averaged over ten experiments. Since the training set of each user is composed by $N$ samples, the score for an unknown template against a user is obtained by matching the unknown template against the $N$ templates of the user and fusing the results by the max rule.

In this paper, we use following protocol (similar to that used in Lumini & Nanni (2007), Maio & Nanni (2006)):

**Genuine recognition attempts:** each template that belongs to the test set is matched against the templates of the same user that belong to the training set.

**Impostor recognition attempts:** the first template of each user is matched against the templates of the remaining users.

For the performance evaluation we adopt the Equal Error Rate (EER) (Maio, Maltoni, Jain, & Prabhakar, 2003), and the error under the Receiver Operating Characteristic curve (AUC) (Huang et al., 2005). EER is the error rate where the frequency of fraudulent accesses (FAR) and the frequency of rejections of people who should be correctly verified (FRR) assume the same value. The Receiver Operating Characteristic curve is a two-dimensional measure of classification performance that plots the probability of classifying correctly the genuine examples against the rate of incorrectly classifying impostors examples.

Figs. 3–5 report the results of a comparison among the following approaches:

- **base(A),** the matchers above denoted as $M_{0,0}$ based on features extracted by the approach $A$ from the input Palm image.
- **base** fusion (by sum rule) among the matchers denoted as $\text{base}(A),$

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig4}
\caption{Performance obtained using in the training set two images for each user (on the left the EER, on the right the AUC).}
\end{figure}
Our ensemble of matchers restricted to those matchers trained by features extracted by the approach \( A \), we combine the matchers \( M_{ij} \) \((i, j = 1, \ldots, 5)\) by sum rule and then we combine \(^3\) this score with \( M_{0,0} \) by max rule.

Our ensemble of matchers (fusion by sum rule).

\( \text{FING3,4} \), a matcher based on the middle finger (Ribaric & Fratric, 2005).

\( \text{FING4} \), a matcher based on the ring finger.

\( \text{GEO} \), a Hand Geometry matcher based on the geometry features shown in Fig. 2 (Franco et al., 2006).

\( \text{FUS} \), fusion (by sum rule) among the last four matchers: \( \text{NEW; FING3, FING4, GEO} \).

In Figs. 3–5 we report the performance obtained using in the training set one, two or three images for each user, respectively. The upper graphs are related to the comparison among the approaches \( \text{BASE}(A) \) and \( \text{NEW}(A) \) by varying the feature extraction \( A \). For each method and for each performance indicator there are three series, obtained varying the value of retained features \( K_A \). In \( \text{S1} \) \( K_{\text{DCT}} = 50, K_{\text{LBP}} = 500, K_{\text{FG}} = 300 \); in \( \text{S2} \) \( K_{\text{DCT}} = 100, K_{\text{LBP}} = 1000, K_{\text{FG}} = 600 \); in \( \text{S3} \) \( K_{\text{DCT}} = 150, K_{\text{LBP}} = 1500, K_{\text{FG}} = 900 \). The lower graphs are a comparison among all the other methods. On the left results in term of EER are reported, on the right in terms of AUC.

The experimental results show that our approach improves the “standard” matchers. The average increase in performance between the matcher based on the extracted Palm Image (BASE) and our ensemble (NEW) is more than 20%. It is interesting to note that the AUC obtained by the standard Gabor-based matcher (the best tested matcher) are: 3.22 (1 image/user in the training set); 1.93 (2 images/user in the training set); 1.5 (3 images/user in the training set). While the AUC obtained by the our ensemble of Gabor-based matcher are: 2.1 (1 image/user in the training set); 1.18 (2 images/user in the training set); 0.95 (3 images/user in the training set).

To validate our idea we have ran other tests. If we combine the six matchers (DCT as features extractor) where the \( i \)th matcher is trained using only the \( K \) features with higher variance among the features extracted from the \( i \)th image, we obtain an EER of 8.7 and an AUC of 3.4 (\( K = 150 \) and 1 image/user in the training set). While if we use Random Subspace (Lumini & Nanni, 2006) for building an ensemble of classifiers (the ensemble is build with 26 classifiers) the best results obtained are (DCT as features extractor and 1 image/user in the training set) an EER of 8.7 and an AUC of 3.2 (while our ensemble of DCT-based matcher obtains an EER of 7.2 and an AUC of 2.34). The best results are obtained when each single matcher is trained using the 50 features with higher variance and other 50 randomly chosen features among the 500 features with higher variance.

Finally, we want to stress that our combination of hand-based matchers (the method named as FUS) obtains a very low AUC, in

3 Before this fusion the scores are normalized to mean 0 and standard deviation 1.

4 The finger image is enhanced using Contrast-Limited Adaptive Histogram Equalization (adapthisteq.m Matlab 7.0 Image Processing Toolbox) and Adjust image intensity values (imadjust.m Matlab 7.0 Image Processing Toolbox). As in Ribaric and Fratric (2005) we use a reduced finger image (it uses 5/6th of the finger). The feature extraction is performed by the discrete cosine transform (DCT), we apply the DCT transform directly to pixels and we retain the coefficients with higher variance.
particular we obtain an AUC of 0.2 when three sample for each user are used in the training set. Please note, that our matchers are based on the simple Euclidean Distance (in our tests the performance using the Euclidean Distance are similar to the performance obtained using the Cosine Distance) and that we do not use supervised feature transform.

4. Conclusions

The objective of this work was to investigate the integration of Palmprint matchers, and to achieve a better performance that may not be achievable with single matcher alone.

This paper has suggested a new method of Palmprint authentication using the combination of Palmprint representations. We want to stress that the results show in this work are obtained using images acquired from simple peg-free setup, since such images are expected to show higher variations as compared to those acquired from fixation pegs (Zhang et al., 2003).

The main contribution of this paper is a new method for creating an ensemble of Palmprint matchers and tested and validate with three different representations of the Palmprint.

Finally, the experimental results show that our combination of hand-based matchers obtains a near zero AUC when three sample for each user are used in the training set. Please note that this low AUC is obtained without to use supervised feature transformations or supervised classifiers.

References


