A Biometric Verification System Based on the Fusion of Palmprint and Face Features

Slobodan Ribaric, Ivan Fratric and Kristina Kis
Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia
slobodan.ribaric@fer.hr, ivan.fratric@fer.hr, kristina.kis@zg.htnet.hr

Abstract

This paper presents a bimodal biometric verification system for physical access control based on the features of the palmprint and the face. The system tries to improve the verification results of unimodal biometric systems based on palmprint or facial features by integrating them using fusion at the matching-score level. The verification process consists of image acquisition using a scanner and a camera, palmprint recognition based on the principal lines, face recognition with eigenfaces, fusion of the unimodal results at the matching-score level, and finally, a decision based on thresholding. The experimental results show that fusion improves the equal error rate by 0.74% and the minimum total error rate by 1.72%.

1. Introduction

Biometrics [7, 25] is an emerging technology that utilizes distinct behavioural or physiological characteristics in order to determine or verify the identity of an individual. Biometric systems that use a single trait are called unimodal systems, whereas those that integrate two or more traits are referred to as multimodal biometric systems. A multimodal biometric system requires an integration scheme to fuse the information obtained from the individual modalities. Various levels of fusion are possible [12, 20], from the feature-extraction to the decision level. Multimodal biometric systems based on palmprint and hand-geometry features [14, 19], face, fingerprint and hand-geometry features [9, 20] and fingerprint, face and speech [8] have been described.

In this paper we describe a prototype of a bimodal biometric system based on palmprint and facial features.

The palm is the inner surface of the hand between the wrist and the fingers [27]. The palm area contains a large number of features that can be used as biometric features, such as principal lines [3, 14, 19, 21, 24], geometry [14, 19, 21], wrinkle [3], delta point, minutiae, datum point features [27] and texture [5, 29]. In addition to the approaches based on these palmprint features, other approaches have been developed for palmprint-based biometric systems as well, such as eigenpalms [15], fisherpalms [23] and 2D Gabor phase encoding [26].

From the numerous methods developed for the purpose of face recognition [28], the use of eigenfaces [22] is one of the most popular. Some other recent face recognition approaches include fisherfaces [2], support vector machines [11] and elastic graph matching [13].

2. A bimodal biometric system

2.1. System overview

Fig. 1. shows the block-diagram of the proposed bimodal biometric verification system.

In the image-acquisition phase the palm and facial images are acquired using a low-cost scanner and a camera, respectively. The processing of these images, up until fusion, is carried out separately in the palmprint recognition and the face recognition subsystems. In the first phase of the palmprint recognition process the area

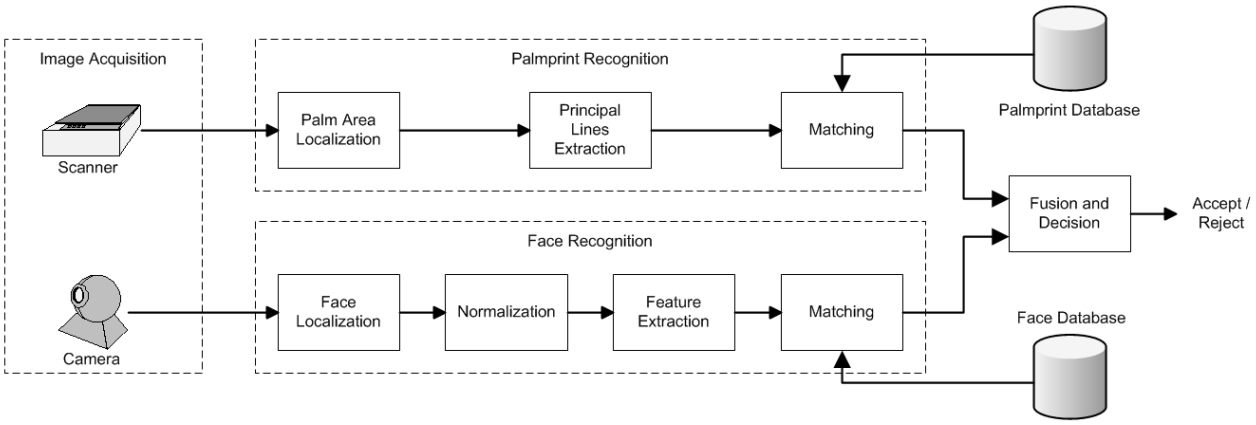


Figure 1: Block-diagram of the proposed multimodal biometric verification system

of the palm is located on the basis of the hand contour and the stable points. In the second phase the principal lines of the palm are extracted using line-detection masks and a line-tracking algorithm. Finally, a live-template based on the principal palm lines is matched to the templates from the palmprint database using an approach similar to the HYPER method [1].

The process of face recognition consists of four phases: face localization based on the Hough method [6]; normalization, including geometry and lighting normalization; feature extraction using eigenfaces; and finally, matching of the live-template to the templates stored in the face database.

Matching scores from both recognition modules are combined into a unique matching score using fusion at the matching-score level. Based on this unique matching score, a decision about whether to accept or reject a user is made.

2.2. Palmprint recognition

In order to localize the palm area, the first step is to preprocess the palm images; this involves Gaussian smoothing and contrast enhancement. Standard global thresholding is used for the segmentation of the hand. After that, a contour-following algorithm is used to extract the hand contour. The two stable points on the hand contour are found [19]: (i) The gap between the little finger and the ring finger, and (ii) The gap between the index finger and the middle finger. Based on the stable points on the contour, the palm area, which is approximated by a hexagonal area, is determined. Figure 2 shows the phases of palm-area localization.

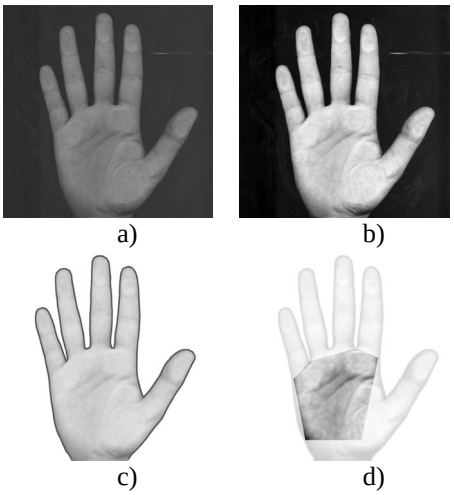


Figure 2: An example of palm-area localization, a) original image, b) image after preprocessing, c) extracted hand contour, d) localized area of the palm

Process of principle lines extraction begins with convolving the grey-scale palmprint area by four line detection masks [29].

After applying the modified line-tracking algorithm, based on [16], a set of lines is obtained. Examples of the palm-line extraction are presented in Figure 3.

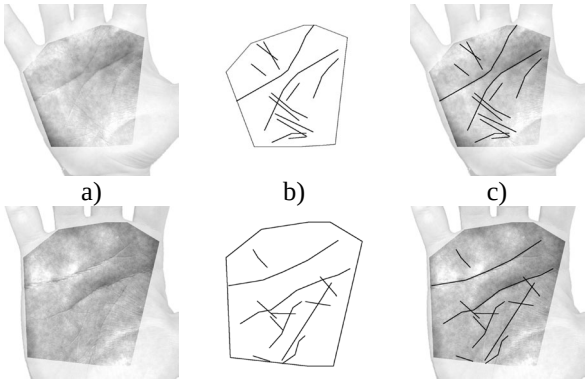


Figure 3: Two examples of palm-line extraction, a) palm area, b) extracted lines, c) overlapped image a) and b)

The extracted lines are described in a hand-coordinate system that is based on the stable points on the hand contour; this makes them invariant to hand translation and rotation. The lines are represented by means of the line-segment sequence, where each segment is described by a four-tuple (x, y, l, α), where x and y are the coordinates of the segment mid-point, l is the length of the segment and α is the segment orientation. The method used for palm matching is based on the adapted HYPER method [1].

The obtained line set contains the most prominent palm lines. The number of lines can vary depending on the palmprint texture and wrinkles. Typically, the number of lines extracted from a palm region is between 15 and 20, with 1 to 5 line segments per line.

The matching of the live-template and the template from the database is based on hypotheses generation and its evaluation.

Generating hypotheses

Since the obtained palm lines are invariant to hand translation and rotation, the two lines (one from the live-template and one from the palmprint template-database) can correspond to each other only if they have a similar position and orientation. Every palm line from the live-template is compared to every palm line from the database-template and a decision is made about whether to add this pair to the hypotheses collection.

Let p be the virtual line that connects the midpoints of the first and the last segment of the palm line from the live-template, and let p' be the virtual line that connects the midpoints of the first and the last segments of the palm line from the database-template. The palm lines are compared in the following way:

1. If the absolute angular difference between the lines p and p' is greater than α_{gen_max}, then the palm lines are dissimilar and no further comparison is necessary. Otherwise, Step 2 is performed. In this application the value of α_{gen_max} is set to $\pi/5$, based on the training set of the palmprint database.
2. The average Euclidian distance d_{gen_12} between the line p' and the segment midpoints of the palm line

from the live-template is calculated:

$$d_{gen\_12} = \frac{1}{n}\sum_{i=1}^{n} d(s_{1i}, p'), \qquad (1)$$

where n is the number of line segments and $s_{1i} = (x_i, y_i)$ is the midpoint of the i^{th} segment.

The average Euclidian distance between the line p and the segment midpoints of the palm line from the database-template is calculated in a similar manner:

$$d_{gen\_21} = \frac{1}{m}\sum_{i=1}^{m} d(s_{2i}, p), \qquad (2)$$

where m is the number of line segments and $s_{2i} = (x_i, y_i)$ is the midpoint of the i^{th} segment. The smaller of these distances is taken as a measure of the distance between the two palm lines: $d_{gen} = \min(d_{gen\_12}, d_{gen\_21})$

If d_{gen} is smaller than the threshold d_{gen_max}, then the pair of palm lines is added to the collection of hypotheses; otherwise the lines are considered to be dissimilar. The experimentally selected value for d_{gen_max} is 50 pixels.

Evaluating hypotheses

The hypotheses collection consists of all the line pairs (one line from the live-template and one from the database-template) that satisfy the above conditions. The hypotheses collection can be defined as HC = $\{(L_{i,LT}, L_{j,DB})\}$, $i \le N_{LT}$ and $j \le N_{DB}$, where N_{LT} is the number of lines in the live-template and N_{DB} is the number of lines in the database-template. In general $N_{LT} \ne N_{DB}$.

Evaluating a hypothesis $H = (L_{i,LT}, L_{j,DB})$ consists of comparing every line segment of $L_{i,LT}$ with every line segment of $L_{j,DB}$ and updating the matching measure for each segment. For every line-segment pair (S_k, S_l); $S_k \in L_{i,LT}$, $S_l \in L_{j,DB}$ where $S_k = (x_k, y_k, l_k, \alpha_k)$ and $S_l = (x_l, y_l, l_l, \alpha_l)$, $k = 1, 2 ..., M_{i,LT}$, $l = 1, 2 ..., M_{j,DB}$, where $M_{i,LT}$ is the number of line segments in the line $L_{i,LT}$ and $M_{j,DB}$ is the number of line segments in the line $L_{j,DB}$, the following parameters are calculated:

1. The absolute angular difference $a = |\alpha_k - \alpha_l|$;
2. The Euclidean distance d between the midpoints of the segments;
3. The distance D_{kl} between the midpoint (x_k, y_k) and the virtual line that segment S_l lies on is computed. The distance D_{lk} is computed in a similar way. Parameter D is defined as the minimum of the two distances: $D = \min(D_{kl}, D_{lk})$.

Each parameter is upper-bounded by the values a_{max}, d_{max} and D_{max}, respectively. The dissimilarity measure, dm_{kl}, for the segment pair is computed in the following way:

1. If $a > a_{max}$ or $d > d_{max}$ or $D > D_{max}$, then $dm_{kl} = 1$; the segments are entirely dissimilar.
2. Otherwise, set $dm_{kl} = p \cdot a / a_{max} + q \cdot D / D_{max} + r \cdot d / d_{max}$

The parameters p, q and r represent the weightings given to parameters a, D and d, respectively ($p+q+r = 1$, $p \ge 0$, $q \ge 0$, $r \ge 0$). The dissimilarity measure, dm_{kl}, is a number in the range [0, 1], and has a lower value for more similar segments. The experimentally determined parameter values are p=0.5, q=0.4, r=0.1, $a_{max}=\pi/18$, $D_{max}=10$ pixels and $d_{max}= \max(l_k, l_l)$, where l_k and l_l represent the lengths of the segments S_k and S_l, respectively.

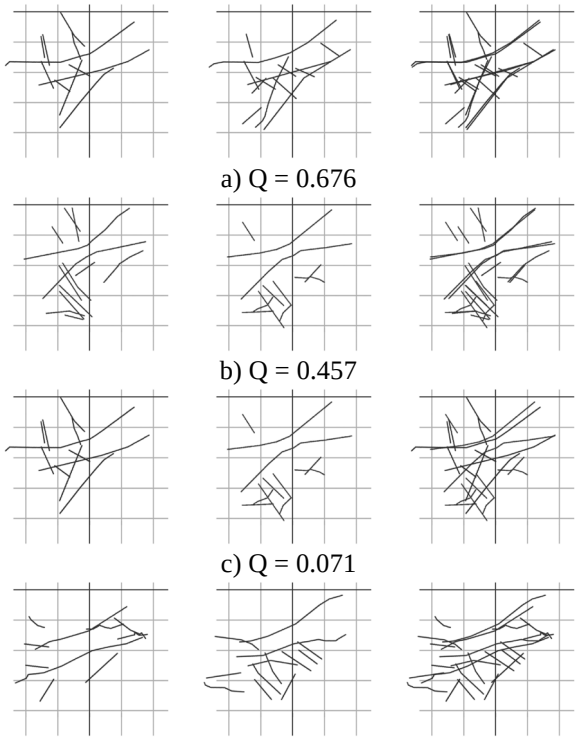
After computing the dissimilarity measure, the measures of matching m_k and m_l for segments S_k and S_l need to be updated. The measures of matching are updated using the following formulas:

$$m_k = m_k + (1-dm_{kl}) \cdot \min(l_k, l_l) \qquad (3)$$

$$m_l = m_l + (1-dm_{kl}) \cdot \min(l_k, l_l) \qquad (4)$$

More than one line segment from $L_{i,LT}$ or $L_{j,DB}$ can contribute to the matching measure of segments from $L_{j,DB}$ or $L_{i,LT}$, respectively. If a line L appears in more than one hypothesis, the matching measures of its line segments are accumulated.

The similarity measure, $Q_{A,B}$, of the palmprint template A and the palmprint template B is expressed in the range [0, 1] and gives an indication of how fully template A is represented within template B. Two similarity measures, $Q_{LT,DB}$ and $Q_{LT,DB}$, are computed. The $Q_{LT,DB}$ is computed as the sum of matching measures of all segments in all lines in the live-template, normalized by the sum of their lengths. Analogously, the similarity measure $Q_{LT,DB}$ is computed (in general, $Q_{LT,DB} \ne Q_{LT,DB}$).



a) Q = 0.676

b) Q = 0.457

c) Q = 0.071

d) Q = 0.072

Figure 4: Comparison of palmprint templates and similarity measure Q: a), b) – comparison of palmprint templates of the same person; c), d) – comparison of palmprint templates of different people. The first and the second columns represent the individual palmprint templates. The third column represents both templates in the same coordinate system

The final similarity measure, Q, which determines how well the two samples match, is computed in the following way heuristically:

1. If $Q_{LT,DB} > T_H$ and $Q_{DB,LT} > T_H$, then $Q = \max(Q_{LT,DB}, Q_{DB,LT})$;
2. If $Q_{LT,DB} < T_H$ and $Q_{LT,DB} < T_H$, then $Q = \min(Q_{LT,DB}, Q_{DB,LT})$;
3. Otherwise, compute $| T_H - Q_{LT,DB} |$ and $| T_H - Q_{DB,LT} |$ and select Q for which the absolute value is greater,

where T_H is a threshold selected experimentally during the training phase ($T_H = 0.25$). Figure 4 shows the similarity measure Q for several palmprint template pairs.

2.3. Face recognition

Faces in images are localized using an approach [18] that combines the Hough method [6] and skin-colour information [10] for face localization.

Since the K–L transform is used for matching, a normalization procedure is required. Face normalization consists of geometry normalization, background removal and lighting normalization. The images of the faces are normalized to a fixed size of 64x64 pixels. The background is removed by leaving only the image elements inside the elliptical region in the normalized images and setting the rest to 0 (black). In the final normalization step, lighting normalization using histogram fitting [4] is applied. In Figure 5, several images after the normalization phase are shown.

Figure 5: Several faces after the normalization phase

The eigenface technique, used in our system for feature extraction, is a widely used method for pattern recognition [22]. It is based on the K–L transform applied to a set of facial images.

The basis vectors of the K–L transform are calculated by finding the largest *m* eigenvectors of the covariance matrix of the set of images. In the case of facial images, when representing these eigenvectors as images, they will resemble faces, and are called eigenfaces. The subspace these eigenvectors' span is called the face-space. Some of the eigenfaces obtained using our database are presented in Figure 6.

It is clear that the largest eigenvectors (those with the smallest ordinal numbers) look more like faces, while those with the largest ordinal numbers look more like noise. The largest eigenvectors carry the useful information (in the sense of image representation) and only they are used as the basis for the face-space, while the information carried by the smaller eigenvectors is lost in the process of encoding. Based on the preliminary recognition experiments on the training database, we chose $m = 111$ for the face-space dimensionality.

The feature vector from an unknown facial image can be obtained by projecting the image onto a face-space. In this process the image is represented as a linear combination of eigenfaces and the feature vector is made of weightings associated with each eigenface. The face template consists of this 111-component feature vector. The matching score between two face-feature vectors is calculated using the Euclidean distance in the matching phase.

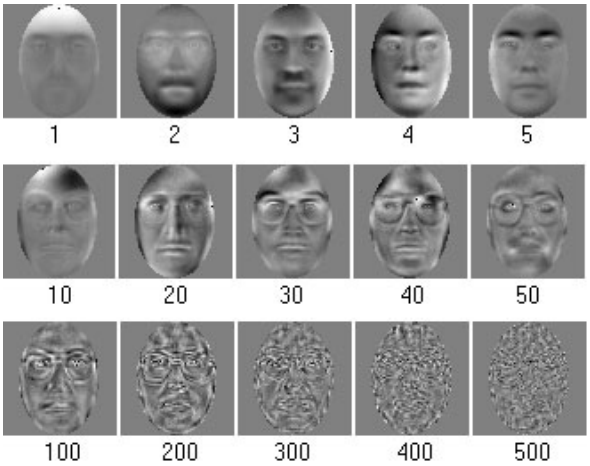


Figure 6: Eigenfaces obtained on our database with appropriate ordinal numbers

2.3. Fusion and decision

In our bimodal biometric system the fusion is performed at the matching-score level.

When trying to verify the identity of an unknown sample we receive two sets of scores from the two independent matching modules: (i) Euclidean distances $D(F_x, F_j)$, where F_x is the unknown face-template, and F_j, $j = 1, 2, \ldots, n$ are the face-templates stored in the database under the identity the system is trying to verify; (ii) Similarity measures $Q(P_x, P_j)$ where P_x is the unknown palmprint-template, and P_j, $j = 1, 2, \ldots, n$ are palmprint-templates stored in the database under the identity the system is trying to verify.

In order to generate the unique matching score we need a way to combine individual matching scores from face- and palmprint-matching modules. Since the palmprint-matching scores and the face-matching scores come in different ranges, a normalization has to be performed before they are combined.

The normalization is carried out by means of two transition functions, S^P and S^F. These functions, which map the distances D, and similarity measures Q, into the interval [0,1] were determined experimentally from the training set of the database [19].

The final matching score, expressed as the total-similarity measure (TSM), is calculated as a linear combination of the largest palm- and face-similarity measures:

$$TSM_{xj} = w_P S^P_{xj} + w_F S^F_{xj} \qquad (5)$$

where w_P and w_F are weighting factors associated with the palm and the face, respectively, and fulfil the condition $w_P + w_F = 1$. The weighting factors were set experimentally, based on the preliminary unimodal verification results obtained on the training database, to $w_P = 0.7$ and $w_F = 0.3$.

The final decision about whether to accept or reject a user is made by comparing the TSM with the verification threshold, T. If TSM > T, the user is accepted; otherwise, he/she is rejected.

3. Performance evaluation

To evaluate the performance of the system a database containing palm and face samples was required. The XM2VTS frontal-face-images database [17] was used as the face database. We collected the hand database ourselves using a scanner. The spatial resolution of the hand images is 180 dots per inch (dpi) / 256 grey levels. As the hand and the face databases contain samples belonging to different people, a "chimerical" multimodal database was created using pairs of artificially matched palm and face samples that were made for testing purposes.

The database was divided into two sets: the training set and the testing set. The training set consisted of 440 image pairs of 110 people (4 image pairs per person) and was used as a training database for individual modalities, to get the distributions of the unimodal matching scores used in the decision fusion module and to get the weightings associated with different modalities.

The testing dataset consisted of 1048 image pairs of 131 people (8 image pairs per person) and was used exclusively for the evaluation of the system performance. Out of 8 image pairs for each person, 5 were used in the enrolment stage and 3 were used for testing. The tests involved trying to verify every test pair for every one of the 131 people enrolled in the database. This setup makes for 393 (131 x 3) valid-client experiments and 51,090 (131 x 3 x 130) impostor experiments.

The results of the experiments, expressed in the terms of FRR (false rejection rate) and FAR (false acceptance rate), vary depending on the selected verification threshold T (Figure 7). Our bimodal system can achieve an EER (equal error rate) of 3.08% for T =0.748 and the minimum TER (total error rate) = 5.94% for T = 0.8.

The comparison of both unimodal systems (palm and face modality) and a bimodal system is given in Figure 8. From the results it is clear that the verification based on the palmprint easily outperforms the verification based on the face. It can also be seen that the fusion of palmprint and facial features improves the verification score. The experiments show that EER is reduced by 0.74%, compared with palmprint modality, and the minimum TER is reduced by 1.72%.

4. Conclusion

We have developed a prototype of a biometric verification system based on the fusion of palmprint and facial features. The experimental results show that although palmprint-based unimodal systems significantly outperform face-based unimodal systems, fusion at the matching-score level can still be used to improve the performance of the system.

The other reasons for including the face modality in biometric systems could be in the system usage for physical access where the additional subsystem can log the facial images of the people accessing the secure object. The psychological effects of such multimodal system should also not be disregarded; it is likely that a system using multiple modalities would seem harder to cheat to any potential impostors.

In the future we plan to test whether setting the user-specific weightings to different modalities can be used to improve a system's performance.

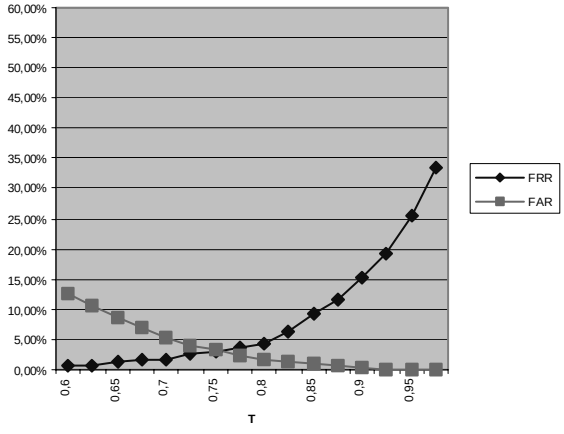


Figure 7: The verification results using the bimodal system depending on threshold

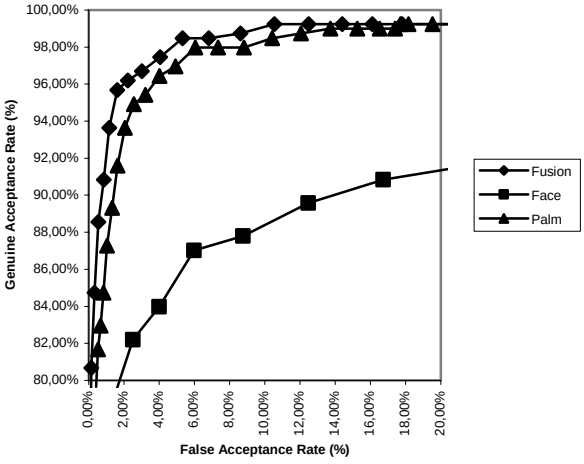


Figure 8: Comparison of unimodal and bimodal system verification results

5. Acknowledgements

This work was supported by the Croatian Ministry of Science, Education and Sport as a part of the projects No. 0036025 and No. 2003-068. “A system for biometric authentication of Internet users based on the fusion of facial and palmprint features.”

6. References

[1] N. Ayache and O. D. Faugeras, “A New Approach for the Recognition and Positioning of Two-Dimensional Objects”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 8, no. 1, 1986, pp. 44-54.

[2] P. Belhumeur, J. Hespanha and D. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, 1997, pp. 711-720.

[3] N. Duta, A. K. Jain and K. V. Mardia. "Matching of Palmprints", *Pattern Recognition Letters*, vol. 23, no. 4, 2001, pp. 477-485.

[4] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, Addison Wesley Publishing Company, Inc., New York, 1993.

[5] C. C. Han, H. L. Cheng, K. C. Fan and C. L. Lin, “Personal Authentication Using Palmprint Features”, *Pattern Recognition*, vol. 36, no. 2, 2003, pp. 371-381.

[6] P. Hough. “Methods and Means for Recognizing Complex Patterns”, *US Patent 3069654*, 1962.

[7] A. K. Jain, R. Bolle and S. Pankanti (Eds.), *Biometrics: Personal Identification in Networked Society*, Kluwer Academic Publishers, USA, 1999.

[8] A. K. Jain, L. Hong and Y. Kulkarni, “A Multimodal Biometric System using Fingerprint, Face and Speech”, *Proceedings of Second International Conference on AVBPA*, 1999, pp. 182-187.

[9] A. K. Jain and A. Ross, "Learning User-Specific Parameters in a Multibiometric System", *Proceedings of the International Conference on Image Processing (ICIP)*, 2002, pp. 57-60.

[10] M. J. Jones and J. M. Rehg, “Statistical Color Models with Application to Skin Deatection”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 1999, pp. 274-280.

[11] K. Jonsson, J. Kittler, Y. P. Li and J. Matas, “Support Vector Machines for Face Authentication”, *Proceedings of British Machine Vision Conference BMVC99, British Machine Vision Association*, 1999, pp. 543-552.

[12] J. Kittler and F. M. Alkoot, “Sum Versus Vote Fusion in Multiple Classifier Systems”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 1, 2003, pp. 110-115.

[13] C. Kotropoulos, A. Tefas and I. Pitas, "Frontal Face Authentication Using Morphological Elastic Graph Matching", *IEEE Transactions on Image Processing*, vol. 9, no. 4, 2000, pp. 555-560.

[14] A. Kumar, D. C. M. Wong, H. C. Shen and A. K. Jain, "Personal Verification Using Palmprint and Hand Geometry Biometric", *Proceedings of the 4th International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA)*, 2003, pp. 668-678.

[15] G. Lu, D. Zhang and K. Wang, "Palmprint Recognition Using Eigenpalms Features", *Pattern Recognition Letters*, vol. 24, no. 9-10, 2003, pp. 1463-1467.

[16] D. Maio and D. Maltoni, “Minutiae Extraction and Filtering from Grey-Scale Images”, in [9], 1999, pp. 155-192.

[17] K. Messer, J. Matas, J. Kittler, J. Luettin and G. Maitre, “XM2VTSDB: The Extended M2VTS Database”, *Second International Conference on Audio and Video-based Biometric Person Authentication (AVBPA'99)*, Washington D.C., 1999, pp. 72-77.

[18] N. Pavešić, I. Fratrić and S. Ribarić, "Degradation of the XM2VTS Database Face Images", *Proceedings of the 2nd COST 275 Workshop, Biometrics on the Internet: Fundamentals, Advances and Applications*, University of Vigo, Vigo, Spain, March 25-26, 2004, pp. 15-19.

[19] S. Ribarić, D. Ribarić and N. Pavešić, "Multimodal Biometric User-identification System for Network-based Applications", *IEE Proceedings Vision, Image & Signal Processing*, Vol.,150, No. 6, 2003, pp.409-416.

[20] A. Ross and A. Jain, "Information Fusion in Biometrics", *Pattern Recognition Letters*, vol. 24, no.13, 2003, pp. 2115-2125.

[21] W. Shu and D. Zhang, “Automated Personal Identification by Palmprint”, *Optical Engineering*, vol. 37, no. 8, 1998, pp. 2359-2362.

[22] M. Turk and A. Pentland, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, vol. 3, no. 1, 1991, pp. 71-86.

[23] X. Wu, D. Zhang and K. Wang, “Fisherpalms Based Palmprint Recognition”, *Pattern Recognition Letters*, vol. 24, no. 15, 2003, pp. 2829-2838.

[24] X. Wu, D. Zhang, K. Wang and B. Huang, "Palmprint Classification using Principal Lines", *Pattern Recognition*, vol. 37, No. 10, 2004, pp. 1987-1998.

[25] D. Zhang, *Automated Biometrics: Technologies & Systems*, Kluwer Academic Publishers, USA, 2000.

[26] D. Zhang, W. K. Kong, J. You and M. Wong, "Online Palmprint Identification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, 2003, pp. 1041-1050.

[27] D. Zhang and W. Shu, "Two Novel Characteristics in Palmprint Verification: Datum Point Invariance and Line Feature Matching", *Pattern Recognition*, vol. 32, no. 4, 1999, pp. 691-702.

[28] W. Y. Zhao, R. Chellappa, A. Rosenfeld and P. J. Phillips, “Face Recognition: A Literature Survey”. *UMD CfAR Technical Report CAR-TR-948*, 2000.

[29] J. You, W. Li and D. Zhang, "Hierarchical Palmprint Identification via Multiple Feature Extraction", *Pattern Recognition*, vol. 35, no. 4, 2002. pp. 847-859.