User Authentication using Fusion of Face and Palmprint*

Ajay Kumar, David Zhang

Department of Computing
The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong.
Tel:(852)-2766-7276, Fax:(852)-2774-0842
Email: ajaykr@ieee.org, csdzhang@comp.polyu.edu.hk

Abstract

This paper presents a new method of personal authentication using face and palmprint images. The facial and palmprint images can be simultaneously acquired by using a pair of digital camera and integrated to achieve higher confidence in personal authentication. The proposed method of fusion uses a feed-forward neural network to integrate individual matching scores and generate a combined decision score. The significance of the proposed method is more than improving performance for bimodal system. Our method uses the claimed identity of users as a feature for fusion. Thus the required weights and bias on individual biometric matching scores are automatically computed to achieve the best possible performance. The experimental results also demonstrate that Sum, Max, and Product rule can be used to achieve significant performance improvement when consolidated matching scores are employed instead of direct matching scores. The fusion strategy used in this paper outperforms even its existing facial and palmprint modules. The performance indices for personal authentication system using two-class separation criterion functions have been analyzed and evaluated. The method proposed in this paper can be extended for any multimodal authentication system to achieve higher performance.

---

* An abbreviated version of this paper was orally presented in the *Intl. Workshop on Multimodal User Authentication*, Santa Barbra, CA (USA), and is referred as [1].

1. Introduction

The technology for trusted *e*-security is critical to many business and administrative process. There has been a newfound urgency after September 11 attacks to develop cutting-edge security technologies. However, the performance of currently available technology is yet to mature for its broad deployment in real environments. The physiological- and behavioral-based human features, *i.e*. biometrics, are unique in an individual. Unlike password or PIN, biometrics cannot be forgotten or lost and requires physical presence of the person to be authenticated. Thus personal authentication systems using biometrics are more reliable, convenient and efficient than the traditional identification methods. The financial risks in personal authentication are high; the double dipping in social welfare schemes are estimated around $40 billion and 40-80% of IT help desk calls are attributed to forgotten passwords [2].

The fingerprint and hand geometry based biometric systems account for more than 60% of current market share [3]. However, the security evaluation against attacks using fake fingerprint or hand has been rarely disclosed. The risk analysis by Matsumoto *et al*. [4], using artificial fingers made of cheap and readily available gelatin, have shown extremely high acceptance on available fingerprint systems. The multimodal biometrics system allows integration of two or more biometric in order to cope up with the stringent performance requirements imposed for high security access. Such systems offer high reliability due to the presence of multiple piece of evidence and are vital for fraudulent technologies as it is more difficult to simultaneously forge multiple biometric characteristics than to forge a single biometric characteristic. One of the recent research problems [5] in the design of multimodal biometrics system concerns with information fusion, *i.e*. how the individual modalities should be combined to minimize errors and

achieve high accuracy. In this work, we investigate a new bimodal biometric system and propose a new method to integrate individual biometric matching scores for multimodal authentication.

1.1 Prior work

Multimodal biometric systems have recently attracted the attention of researchers and some work has already reported in the literature [6]-[14]. Hong and Jain [6] describe a bimodal biometric system that uses fingerprint and face to achieve significant improvement in identification accuracy. Ben-Yacoub *et al*. [7] combined face with voice and achieved best results from support vector machine classifier while Chatzis *et al*. [8] demonstrated this from radial basis function network. Recently, bimodal biometric systems using face and iris [9], palmprint and hand-geometry [10], have also shown promising results. Jain and Ross [11] achieve higher performance in combining face, hand-geometry, and fingerprint by learning user-specific parameters. Osladciw *et al*. [12] have presented a framework for multimodal biometric system that is adaptively tunable to the security needs of user. Verlinde *et al*. [13] achieve decision level fusion by using parametric and non-parametric classifiers. Kittler *et al*. [14] have mathematically shown that the sum rule is most resilient for the estimation of errors in biometric fusion provided the errors in the estimation of posteriori probabilities are small.

1.2 Proposed System

This paper investigates a new bimodal biometric system using face and palmprint [1]. The highest user-acceptance [12] for face is attributed to comfort and ease in the acquisition of facial images. People have lot of concerns about hygiene, especially due to recent spread of SARS (Severs Acute Respiratory Syndrome), while using biometric sensors *e.g*. fingerprint sensors. However the face and palmprint images can be

conveniently acquired from the touchless sensors such as digital camera. The face image of user can be captured while the user is positioning his palmprint to the imaging system. The lighting, background, and the pose for the face image acquisition can be controlled as the user is positioned in the restricted space while acquiring his palmprint image. The automated acquisition of palmprint images from the digital camera is detailed in [10]. Thus a bimodal biometric system that uses face and palmprint will be perceived to be more user-friendly and hygienic.

One of the important features only available in personal authentication, but not in recognition, is the claimed user identity. The claimed user identity is unique for every user and can be used to restrict the decision space, *i.e*. range of matching scores, in user authentication. The claimed user identity can be suitably coded and then used as a feature to classify the genuine and impostor matching scores and is investigated in this paper. The contributions of this paper can be divided in two parts; (i) investigate a new bimodal biometric authentication system by integrating face and palmprint features, and (ii) propose a new decision level fusion strategy, which combines claimed user identity, to achieve improved performance in multimodal authentication system. While formulating the fusion strategy, we also demonstrate that the performance in unimodal authentication system can be improved by integrating claimed user identity with matching score. These consolidated matching scores, when employed for Sum, Max, or Product rules, can achieve much better performance in any multimodal system. In addition, the usefulness of two-class feature separation criterion functions to evaluate the performance of biometric authentication system is also demonstrated.

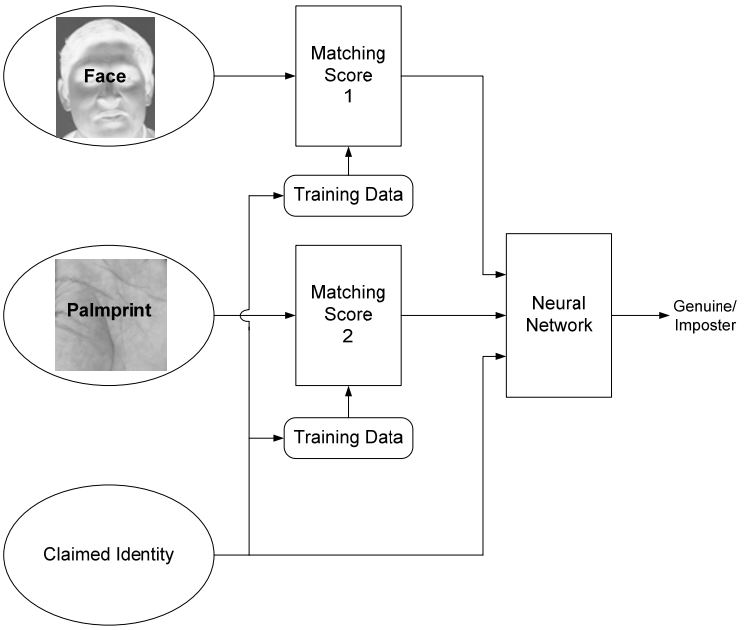


Figure 1: Personal Authentication using Face and Palmprint.

2. Bimodal User Authentication

The proposed biometric authentication system is based on decision level fusion model where the integration is performed on the individual matching scores to generate a composite decision score. These decision scores are used to classify the user into genuine or imposter class. Figure 1 shows one of the possible variant of the proposed system that has been investigated in this work. The acquired grey-level images from the palmprint and face are presented to the system. In addition, each of the users also presents its claimed identity to the system. Each of the acquired images are used to generate respective matching score from the data stored during training. These matching scores are presented to a trained feed-forward neural network classifier. As shown in figure 1, the claimed user identity is also used as a feature to neural network classifier. The weights and bias on individual matching scores are automatically computed from the training data during training phase. Thus a reliable decision score is generated from the output of the trained neural network which is used to assign genuine or imposter class label to the user.

2.1. Face Matching

Several face recognition algorithms have been proposed in the literature [15]. Among these, the appearance based face algorithms are most popular and have been installed in real-environments [16]. The appearance based face authentication approach used in this work employed eigenfaces [17]. Each of the $M \times N$ grey-level face images from every subject are represented by a vector of $1 \times MN$ dimension using row ordering. The normalized set of such training vectors is subjected to principal component analysis (PCA). The PCA generates a set of orthonormal vectors, also known as eigenfaces, which can optimally represent the grey-level information in the training dataset. The projection of users training face image on eigenfaces is used to compute the characteristic features. The matching score for every test face image is generated by computing the similarity score between the feature vectors from the claimed identity (x_c) and computed characteristic feature vector (x_q).

$$\eta_1 = \frac{\sum x_q x_c}{\sqrt{\sum x_q^2 \sum x_c^2}} \tag{1}$$

2.2 Palmprint Matching

Palmprint contains several complex features, *e.g*. minutiae, principal lines, wrinkles and texture, which have been suggested for personal identification. The palmprint matching approach used in this work is same as detailed in [10]. Four directional spatial masks are used to capture line features from each of the palmprint images. The combined directional map is generated from voting of the resultant four images. The standard deviation of pixels, from each of the 24 × 24 pixel overlapping block with 25% overlap, in the combined image is used to form characteristic feature vector. The palmprint

matching scores are generated by computing the similarity measure η_2, similar to η_1 (1), between the feature vectors from acquired image and those stored during training.

3. Decision Fusion

Decision level fusion that can consolidate the decision scores from multiple evidences has shown [3], [18]-[19] to offer radical increase in performance. Therefore we investigated decision level fusion for the proposed authentication system. The decision level fusion for the authentication of user *U* can be formulated as follows: Given $\{\eta_1, \eta_2\}$, assign user *U* to one of the two possible classes $\{\varphi_1, \varphi_2\}$, where φ_1 and φ_2 signifies genuine and imposter classes respectively, *i.e.* assign $U \rightarrow \varphi_x$ if,

$$P(\varphi_x|\eta_1, \eta_2) = \max\{P(\varphi_1|\eta_1, \eta_2), P(\varphi_2|\eta_1, \eta_2)\} \quad x = 1, 2 \tag{2}$$

where $P(\varphi_1|\eta_1, \eta_2)$, $P(\varphi_2|\eta_1, \eta_2)$ are the posteriori probability of genuine and imposter classes, for given $\{\eta_1, \eta_2\}$, respectively. Kittler *et al.* [14] have provided a common theoretical model for the decision level fusion. This model has simplified the combination rules, although with several implicit assumptions, which have been used in this work. Our fusion strategies can be classified into three categories.

3.1 Unimodal Fusion with claimed user identity

The matching scores from each of the two modalities, *i.e.* Face and Palmprint, were further consolidated by integrating with the claimed user identity. In the context of decision level fusion, our strategy is to use the claimed identity of every user as a feature to FFN as shown in figure 2. Each of the matching scores η_1 or η_2 are concatenated with claimed user identity, say u_c, to form a 2-D feature vector. These 2-D feature vectors from each of the two genuine and imposter classes, for every user, are used to train a

Feed-Forward Neural network (FFN). The architecture of employed FFN was same as used and detailed in section 3.3. The trained FFN was used to generate the decision scores for each of the user samples from the test data. Two distinct experiments, each for palmprint and face, were performed to observe the performance of proposed fusion strategy using claimed user identity.

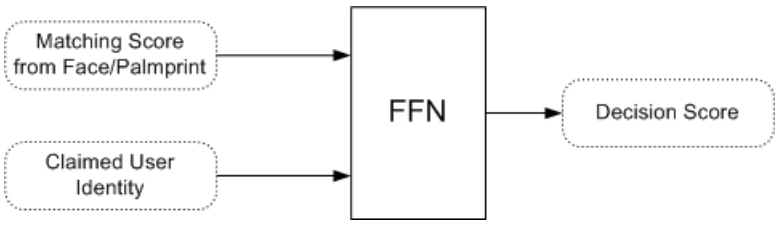


Figure 2: Consolidation of Matching Scores with Claimed User Identity using FFN.

3.2 Sum, Max, and Product Rule

Our second set of experiments in this category uses the decision scores from section 3.1 as consolidated matching scores $\{\eta_1', \eta_2'\}$ from individual biometric modality. Assuming that the representations shown in (2) are statistically independent, the joint probability distribution $p(\eta_1, \eta_2|\varphi_x)$ can be obtained from the straightforward application of estimation theory [14], [20]. Thus the improved estimates of posteriori class probabilities $P'(\varphi_x|\eta_1, \eta_2)$ are obtained from the following combination strategies:

$$\text{Sum Rule,} \quad P'(\varphi_x|\eta_1, \eta_2) = 0.5\{P(\varphi_x|\eta_1') + P(\varphi_x|\eta_2')\} \tag{3}$$

$$\text{Max Rule,} \quad P'(\varphi_x|\eta_1, \eta_2) = \max\{P(\varphi_x|\eta_1'), P(\varphi_x|\eta_2')\} \tag{4}$$

$$\text{Product Rule} \quad P'(\varphi_x|\eta_1, \eta_2) = \{P(\varphi_x|\eta_1') * P(\varphi_x|\eta_2')\} \tag{5}$$

The combination rules (3), (4) and (5) are routinely used in practice. The results achieved from these combination rules provide a plausible justification for underlying assumption behind these rules.

3.3 Decision Rule using FFN

Instead of considering the matching scores $\{\eta_1, \eta_2\}$ as estimates of posteriori probability these are considered as a feature to binary classifier. This scheme is shown in figure 1 and is referred to as *decision rule* in this paper. The matching score from face and palmprint along with the claimed user identity or user code is simultaneously used to authenticate the user. Thus a concatenated feature vector $\{\eta_1, \eta_2, \eta_c\}$ corresponding to genuine and imposter matching scores is used to train the FFN. The decision score output from the trained FFN is used to classify the user into genuine or imposter class.

The execution speed of multi-layer feed-forward neural network is among the fastest of all models currently in use. Therefore this network may be the only practical choice for online personal authentication. The architecture of a Q layer FFN with P_l neurons in the l^{th} (l = 1, …, Q) layer is can be described as follows [21]:

$$\theta_k^l = \sum_{v=1}^{P_{l-1}} b_{vk}^{l\text{-}1,l} w_v^{l\text{-}1}, \quad w_k^l = \psi(\theta_k^l) \tag{6}$$

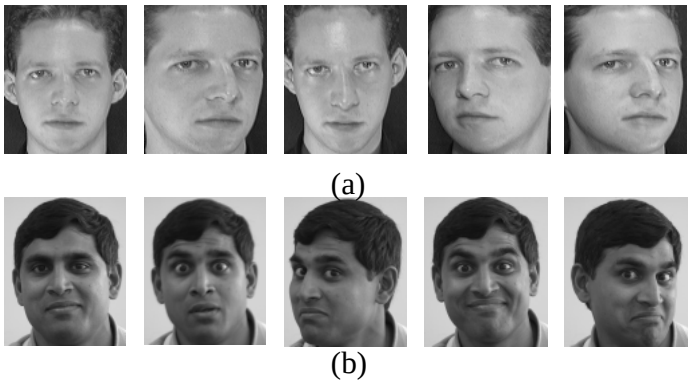
where θ_k^l denotes the sum of weighted inputs for k^{th} $(k = 1, ..., P_l)$ neuron in the l^{th} layer and w_k^l is the output from k^{th} neuron in the l^{th} layer. The weights between the v^{th} neuron from $(l-1)^{\text{th}}$ layer to the k^{th} neuron in the l^{th} layer are denoted by $b_{vk}^{l-1,l}$. The values –1 and 1, corresponding to 'impostor' and 'genuine' responses, were given to the three layers FFN during training as the correct output responses for expected classification from the training data. The hyperbolic tangent sigmoid activation function was selected as its output (-1 to +1) is perfect for learning target (-1 to +1) output values.

$$\psi(\theta_k^l) = tanh(\theta_k^l) \quad \text{for } l = 1, 2. \tag{7}$$

However, a linear activation was selected for the last layer of FFN so that the network output from the test data can take any value (which will be otherwise limited with sigmoid neurons). The back-propagation training algorithm is used for minimizing training function T_e, *i.e.*

$$T_e = \frac{1}{\mathrm{MP_Q}} \sum_{m=1}^{\mathrm{M}} \sum_{k=1}^{\mathrm{P_Q}} (w_{k,m}^{\mathrm{Q}} - y_{k,m})^2 \tag{8}$$

where m is an index for input-output pair and $(w_{k,m}^{\mathrm{Q}} - y_{k,m})^2$ is the squared difference between the actual output value at the k^{th} output layer neuron for pair m and the target output value. The connection weights $b_{vk}^{l-1,l}$ are updated after presentation of every feature vector using a constant learning rate. The training weights were updated by using resilient backpropagation algorithm [22] which achieves faster convergence and conserves memory.



(a)

(b)

Figure 3: Sample face images from the Olivetti research database [20] in (a) and from our database in (b).

4. Experiments

The performance and feasibility of the proposed bimodal system is assessed on a user database of 700 images from 70 users. The ORL face database [23] is composed of 400 images from 40 users with 10 images per user. These images contain uniform black background with variations mainly across pose and expression. We captured another 300

images from 30 users with 10 images per user from the Olympus C-3020 digital camera. All these images were captured against a uniform illumination with white background over a period of two month. Each of these images has spatial and gray-level resolution of 92 × 112 and 256 respectively. The variations in these images are mainly due to facial expression and pose. Thus an integrated face database of 700 images from 70 users was employed that contained reasonable variation in pose, lighting and expression. Figure 3 shows some sample images from this database.

The hand images from 70 users, with 10 images per user, were acquired using the same digital camera. Since the live feedback of the acquired images was available, the users were guided in placing their hands at the center of imaging board in upright position. Each of these captured hand images were used to automatically segment the 300 × 300 region of interest, *i.e*. palmprint, using the method detailed in [10]. The variations in the palmprint images were mainly due to stretching and misalignment. Figure 4 shows few images in the extraction of palmprint image and its features. The 700 images from face and palmprint were randomly paired[†] to obtain a bimodal image set for each of the 70 users. In all of our experiments, the first four images samples, each from face and palmprint, were used for training and rest six were for testing.

(a) (b) (c) (d)

Figure 4: Feature extraction from the palmprint images; (a) Captured hand image, (b) aligned binary hand image after ellipse fitting, (c) extracted palmprint image, and (d) extracted features from (c).

[†] The mutual independence of biometric modalities [24] allows us to augment two biometric indicators that are collected individually.

Each of the 92 × 112 pixel face image was used to obtain 1 × 70 characteristic feature vector from the 70 eigenfaces. Each of the palmprint images were used to obtain a characteristic features vector of size 1 × 144, from the overlapping blocks of size 24 × 24 pixel with an overlap of 6 pixels (25%). The details of feature extraction method can be found in [10]. The genuine and impostor matching scores from the training set were used to train 18/5/1 neural network as discussed in section 3.3. The learning rate was fixed at 0.01 and the training was aborted when the maximum number of training steps reached to 5000. There is no guarantee that the achieved training error is global and therefore FFN was trained 20 times with the same parameters and the result with the smallest of training errors of all the results are reported [25]. The trained FFN was used to test 420 (70×6) genuine and 28980 (70×69×6) impostor matching scores from the independent test data.

5. Results and Performance

The first set of experiments was performed to evaluate the performance improvement due to the usage of claimed user identity. The receiver operation characteristic (ROC) only from the palmprint test images is shown in Figure 5. The ROC corresponding to only palmprint when claimed user identity is utilized, as detailed in section 3.1, is also shown in Figure 5. Similarly, Figure 6 shows the ROC only from face test images without and with the usage of claimed user identity. The ROC in Figure 5 and 6 illustrate the performance improvement for unimodal authentication when consolidated matching scores are employed. The relative distribution of genuine and imposter matching scores from palmprint and face, for fusion scheme proposed in section 3.1, can be ascertained from the plots in Figure 7 and Figure 8. It is apparent from these figures that the

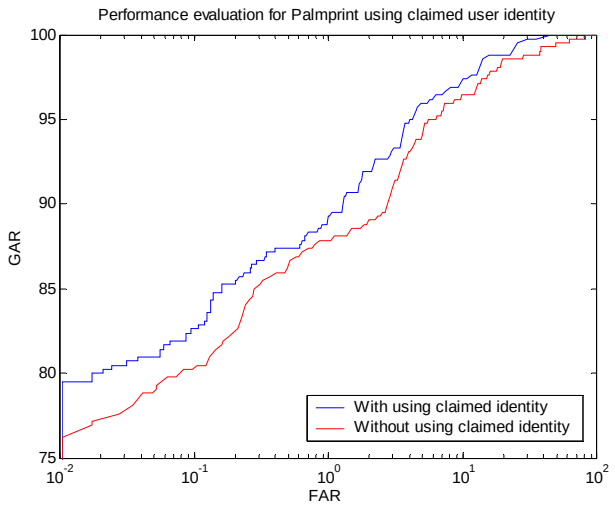


Figure 5: Receiver Operating Characteristics for the user authentication using only Palmprint.

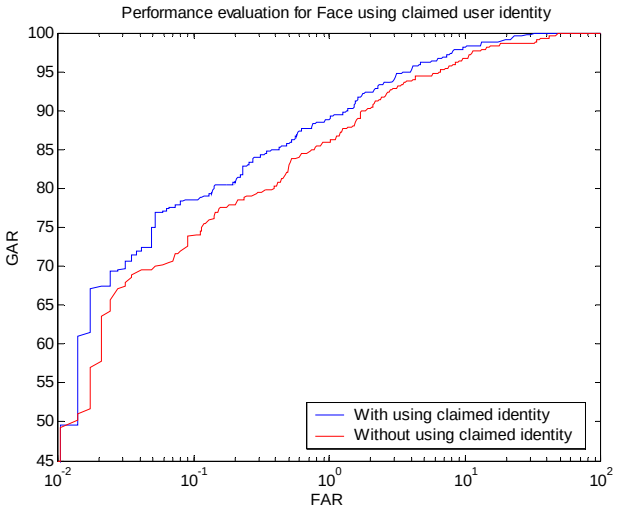


Figure 6: Receiver operating characteristics for the user authentication using only Face.

separation the genuine and imposter matching scores improves with the usage of claimed user identity. Figure 9 shows the combined (using Sum rule) and individual biometric ROC when consolidated matching scores are employed. The variation of false acceptance rate (FAR) and false reject rate (FRR) with decision threshold, corresponding to the combined ROC in Figure 9, is shown in figure 10.

It is difficult to compare the performance of suggested fusion schemes using ROC predominantly due to its variation over large range. In order to ascertain the quantitative improvement in the performance of suggested fusion schemes, we considered

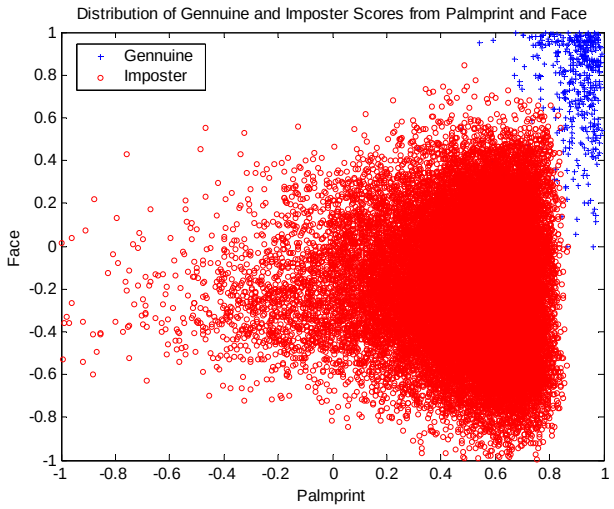


Figure 7: Distribution of genuine and imposter matching scores from Palmprint and Face.

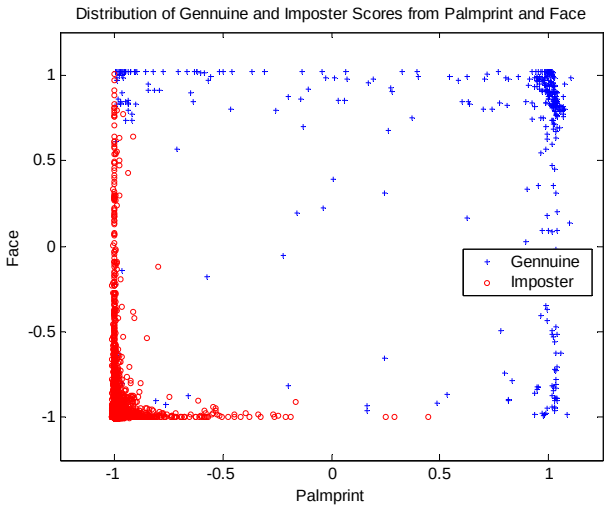


Figure 8: Distribution of consolidated matching scores from Palmprint and Face using claimed user identity.

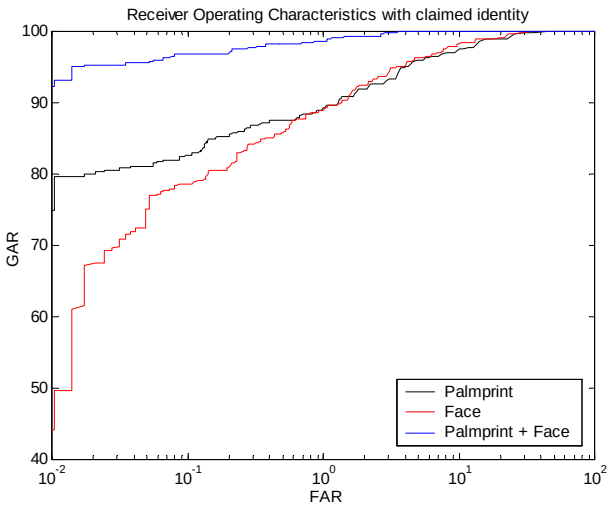


Figure 9: Comparative Receiver Operating Characteristics Using claimed user identity.

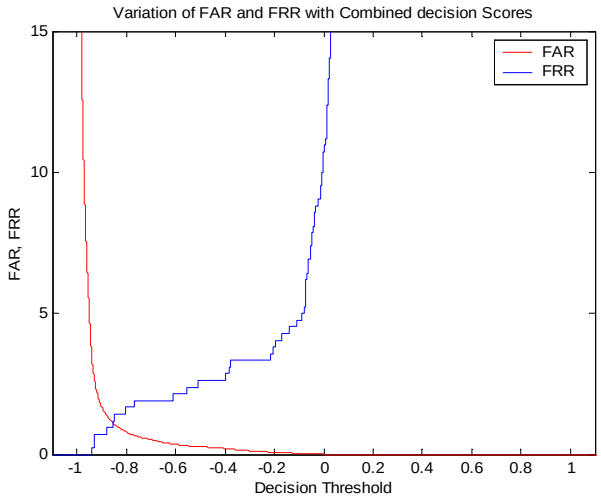


Figure 10: Variation of FAR and FRR scores for the combined decision scores using consolidated matching scores.

the performance index using three criterion functions J_{MS}, J_U, and J_F. These criterion functions have been used to compute the separation of two class texture features for the design of optimal FIR filters for textured defects [26]. The criterion function J_{MS} represents the *ratio* between the average decision scores and was originally suggested by Mahalanobis and Singh [27] for the design of correlation filters.

$$J_{MS} = \frac{\mu_g}{\mu_i} \tag{9}$$

where μ_g and μ_i represent the mean decision scores from the genuine and imposter class respectively. Another criterion function J_U is measure of relative *distance* between average decision scores.

$$J_U = \frac{(\mu_g - \mu_i)^2}{\mu_g \mu_i} \tag{10}$$

The criterion was originally suggested by Unser [28] in the designing of optimal texture transform and user later to design optimal filters for segmentation [29] and defect detection [25]. The main disadvantage of using J_{MS} or J_U as performance index is that

the these functions do not account for the variance of decision scores. Consequently, an authentication system selected on the basis of these performance indices can only achieve large separation of mean genuine and imposter decision scores. However if the variance of these decision scores are large then the feature distribution can considerably overlap. Thus an authentication system should not only produce large separation of genuine and imposter decision scores but also yield their low variance. Another criterion commonly known in pattern recognition literature is Fisher criterion, referred here as performance index J_F, which can also account for the variance of decision scores.

$$J_F = \frac{(\mu_g - \mu_i)^2}{\sigma_g^2 + \sigma_i^2} \tag{11}$$

where σ_g and σ_i are the standard deviation of genuine and impostor decision scores respectively. In this paper, we have evaluated the performance of suggested methods using J_{MS}, J_U, and J_F as performance indices. The scores from these performance indices were computed from the decision scores (scaled to positive axes) from the independent test data. The total minimum error (E_{TM}), *i.e.* minimum sum of FAR and FRR scores, has been used to evaluate the performance in prior work [9]-[10]. Therefore E_{TM} is also quoted to ascertain the performance.

Table 1 and Table 2 shows the performance indices obtained from the experiments with and without the usage of claimed user identity respectively. Comparing the entries in the first two rows of these tables it can be noticed that all of the performance indices have shown improvement when the claimed user identity or consolidated matching scores are utilized. This suggests that higher performance can be achieved in unimodal authentication by incorporating claimed user identity. This

Table 1. Performance indices using claimed user identity.

		J_{MS}	J_U	J_F	E_{TM}
Face		4.0398	2.2873	6.1635	8.3161
Palmprint		4.0708	2.3165	4.3515	8.7578
Fusion	Sum Rule	4.0551	2.3017	11.07999	2.0773
	Max Rule	4.6111	2.8280	37.2562	2.2705
	Prod Rule	5.944	4.112	2.3461	1.150

Table 2. Performance indices without using claimed user identity.

		J_{MS}	J_U	J_F	E_{TM}
Face		1.153	0.0203	2.025	9.79
Palmprint		1.2049	0.0348	2.7497	10.4555
Fusion	Sum Rule	1.7513	0.3223	9.7740	3.6473
	Max Rule	1.2156	0.0382	3.7257	5.0414
	Prod Rule	1.7827	0.3436	10.8617	4.3099

Table 3. Error rates using claimed user identity.

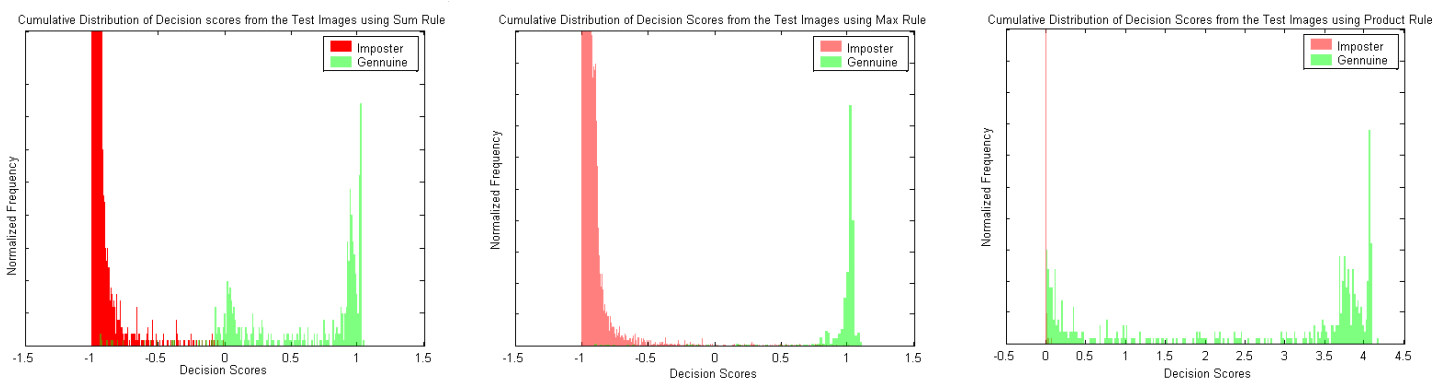
		E_{ER}	$FAR_{E_{TM}}$	$FRR_{E_{TM}}$
Face		4.28	3.0780	5.2381
Palmprint		4.45	4.4720	4.2857
Fusion	Sum Rule	1.13	1.1249	0.9534
	Max Rule	1.4	0.354	1.9165
	Prod Rule	0.715	0.445	0.7143

Table 4. Error rates without using claimed user identity.

		E_{ER}	$FAR_{E_{TM}}$	$FRR_{E_{TM}}$
Face		5.475	4.3202	5.4762
Palmprint		5.237	5.2174	5.2381
Fusion	Sum Rule	2.00	1.32	2.3273
	Max Rule	2.96	1.14	3.9014
	Prod Rule	2.21	2.4051	1.9048

conclusion is significant as the improved performance can be achieved without any extra hardware for feature extraction. Table 3 and Table 4 show the equal error rate (E_{ER}) and FAR/FRR scores at total minimum error (E_{TM}) corresponding to entries in Table 1 and Table 2 respectively. Another conclusion that can be drawn from Table 2 itself is that the fusion (Sum, Max, or Product rule) of two biometric can achieve higher performance than any of these biometric alone as all of the indices have shown improvement.

The performance index from Table 2 and Table 1 for Sum rule suggests that the better performance can be obtained when consolidated matching scores are used in fusion using Sum rule. Similar conclusion can be drawn for the Max rule also. However the performance improvement for J_F using Max rule is significant as it is the highest of all fusion schemes in Table 1. But the improvement in the relative *ratio* or *distance* between two decision scores, *i.e.* J_{MS} or J_U , is not significant. This suggest that the highest scores from J_F is predominantly due to the smaller variance in decision scores from the genuine and imposter class than those from Sum rule. This can also be ascertained from the scaled plots of cumulative decision scores for Max and Sum rule shown in Figure 11. Thus Max rule may be the better choice, as compared to the Sum rule, while combining palmprint and face using claimed user identity.



(a) (b) (c)

Figure 11: Cumulative Distribution of Decision Scores obtained from Sum Rule in (a), Max Rule in (b), and Product Rule in (c).

The performance indices from Product rule does not generate a straightforward conclusion despite the fact that all of the performance indices except J_F, show improvement in Table 1 as compared to those in Table 2. The mean decision scores from the two class, *i.e.* J_{MS} or J_U, have improved significantly for Product rule with claimed user identity and so is the improvement in E_{TM}. However these improvements have been offset by low J_F score which is the lowest score in Table 1. The significant decrease in J_F can be attributed to the high variance in genuine decision scores and this can be noticed from Figure 11 (c). It should be noted that the dynamic range of individual consolidated matching scores lies in the vicinity of -1 to +1 and was translated to about 0-2 range. Thus the resulting dynamic range for decision scores from Product rule is about 0-4 as seen in Figure 11(c). The results in Table 1 suggests that the Product rule is better suited for the fusion scheme using consolidated decision scores, than Max or Sum rule, when the large separation of mean genuine and imposter decision scores is desired. The results from Product rule also suggests that the straightforward conclusion only from J_F can be misleading[‡] as the scores from J_{MS}, J_U, and J_F are distinct. Thus our approach for performance evaluation using all three[§]performance indices is reasonable/justified.

Table 5 shows the performance of fusion scheme discussed in section 3.3 when the decision scores are obtained directly from the trained FFN. Table 6 shows the corresponding E_{ER} and FAR/FRR scores from fusion scheme in Table 5. The performance improvement with the usage of claimed user identity can be observed from Table 5 from each of the performance indices. However a significant improvement (33%)

[‡] The E_{TM} and E_{ER} from Product rule are lowest of all fusion schemes considered in Table 1.

[§] The performance scores from J_{MS} and J_U may not be similar always as this has been illustrated in [25].

Table 5. Performance indices for direct fusion using FFN.

Fusion	J_{MS}	J_U	J_F	E_{TM}
With Claimed Identity	4.9009	3.1049	51.1682	1.7046
Without Claimed Identity	4.8676	3.0731	38.4683	2.1877

Table 6. Error rates for direct fusion using FFN.

Fusion Scheme	E_{ER}	$FAR_{E_{TM}}$	$FRR_{E_{TM}}$
With Claimed Identity	1.19	0.5141	1.1905
Without Claimed Identity	1.72	0.2877	1.9

is shown for J_F which is predominantly due to smaller variance in the decision scores obtained from the usage of claimed user identity. This suggests that the claimed user identity or user-code has been effective in guiding/restricting the genuine and imposter scores, through trained FFN classifier, to their designated trained scores centered at +1 and -1 respectively.

How does the fusion schemes with consolidated matching scores in Table 1 compare with those in Table 5? The fusion scheme detailed in section 3.3 has the best performance when small variance of genuine and imposter decision scores, from their designated training values, is desired as J_F is highest of all schemes considered in this paper. Similarly, Product rule using consolidated matching scores has the best performance when the large separation of mean genuine and imposter decision scores is desired as J_{MS} and J_U are the highest of all the fusion schemes considered in this paper. Both of these conclusions also suggest that the usage of claimed user identity has significant improvement in performance in combining palmprint and face matching scores.

6. Discussion

A summary of prior or related work on bimodal user authentication has been presented in Section 1.1. It is difficult to make a quantitative comparison of these approaches with the one suggested in this paper predominantly due to the varying nature of employed biometric modalities and fusion strategies. Therefore a qualitative comparison is made so as to focus on advantages of our approach. The error rates based on FAR and FRR have been the most common performance criterion used in prior work. In this paper we have suggested and evaluated three two-class feature separation criteria functions as performance indices. Due to the limited size of database used in performance analysis, the quantitative evaluation/comparison of performance based on error rates is cumbersome. Thus an error rate of 1% can have different meaning based on the total number of users/database. However, the performance criteria based on statistical averages from genuine and imposter decision scores, such as the ones suggested in this paper, are expected to remain fairly uniform with the variation/increase in number of the users.

The user-specific thresholds suggested by Jain and Ross [11] are aimed at improving performance, *i.e*. genuine acceptance rate, for a fixed (preset) FAR. However in our work (figure 1), the user-specific information is used to minimize both FAR and FRR simultaneously by training a classifier at decision level fusion. Furthermore, the user-specific thresholds suggested in [11] can also be superimposed on our scheme to further improve the said performance, for the desired FAR, by using user-specific decision thresholds instead of the common decision threshold used in our experiments. Wang *et al*. [9] have combined face with iris and achieved best the performance by using the user-specific radial basis function neural network (RBFN) but without claimed user-identity. The main drawback with RBFN is that they require many times more memory

than a comparable FFN [30] and this requirement becomes enormous when an RBFN is employed for every user. Thus RBFN does not offer an attractive alternative to FFN for online personal authentication. But the usage of user-specific FFN, with claimed user identity, can further enhance the performance when the storage requirements are not very critical. In Summary, our bimodal authentication system compares very well with the prior work. The fusion schemes suggested in this work, namely Sum, Max, and Product rule using consolidated matching scores (section 3.2) and FFN based fusion with claimed user identity (section 3.3), can be applied to similar scenarios in other domains to achieve higher performance.

7. Conclusions

This paper introduces a new bimodal personal authentication method by integrating palmprint with face. The grey-level images of palmprint and face can be simultaneously acquired with a pair of digital camera and used to achieve the performance that may not be possible by single biometric alone. The significance of the proposed method is more than improving performance for a bimodal system. The proposed method utilizes the claimed identity of subjects as a feature for fusion. The employed neural network can automatically compute the weights and bias for the individual biometric matching scores, so as to minimize FAR and FRR and thus achieve the best possible performance. The decision level fusion scheme formulated in this paper (section 3.1) also achieves improved performance by integrating the claimed user identity with biometric matching scores for the unimodal authentication system. Thus our fusion strategy outperforms its existing face and palmprint authentication modules detailed in [11],[17] and [10] respectively. This contribution is of significance as the performance improvement is

achieved without any additional hardware for feature extraction. The experimental results presented in this paper (Tables 1-4) also demonstrate that much better results can be obtained from *consolidated* matching scores using Sum, Max, or Product rule than those obtained from the usage of direct matching scores. This paper has also suggested and evaluated three two-class feature separation criteria as performance index. The analysis of three performance indices in section 5 provides considerable understanding to evaluate quantitative performance of any personal authentication system. However, the disadvantage with these performance indices is that they only account for first two moments and ignore higher order moments. The performance index that can also account for the shape of distribution with higher order moments is perceived to be more reliable and is suggested for the future work.

The database employed in this work is of reasonable size as our objective was to demonstrate the feasibility of proposed fusion schemes. However, the more reliable estimate on the performance can be obtained if significantly larger database is available and we are working to enroll more users. The proposed method of fusion can be extended to any multimodal system to achieve higher performance. Face, voice, iris, hand-geometry and palmprint, employed in prior work [6]-[11], [14], can be acquired from touchless sensors. However, the user-acceptance for voice and iris is limited due to the efforts required in their acquisition. Therefore a multimodal system that uses palmprint, face and hand geometry can offer high user-acceptance and is suggested for future work.

8. References

1. A. Kumar and D. Zhang, “Integrating palmprint with face for user authentication,” *Proc. Multi Modal User Authentication Workshop*, pp. 107-112, Santa Barbara, CA, USA, Dec. 11-12, 2003.
2. Gartner, Inc., http://www.gartner.com

3. *Biometric Systems: Worldwide Deployments, Market Drivers, and Major Players*, Allied Business Intelligence, http://www.abiresearch.com
4. T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, “Impact of artificial “gummy” fingers on fingerprint systems,” pp.275-289, *Proc. SPIE* 4677, San Jose, USA, 23-25 Jan. 2002.
5. E. P. Rood and A. K. Jain (*Eds.*), *Biometric Research Agenda*: *Report of the NSF Workshop*, http://biometrics.cse.msu.edu/BiometricResearchAgenda.pdf, Jul. 2003.
6. L. Hong and A. Jain, “Integrating faces and fingerprints for personal identification,” *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 20, pp. 1295-1307, Dec. 1998.
7. Ben-Yacoub, Y. Abdeljaoued, and E. Mayoraz, "Fusion of face and speech data for person identity verification," *IEEE Trans. Neural Networks*, vol. 10, pp. 1065-1074, 1999.
8. V. Chatzis, A. G. Borş, and I. Pitas, “Multimodal decision-level fusion for person authentication,” *IEEE Trans. Sys. Man Cybern.*, Part A, vol. 29, pp. 674-680, Nov. 1999.
9. Y. Wang, T. Tan, and A. K. Jain, “Combining face and iris for identity verification,” *Proc. AVBPA*, Guildford (U.K.), Jun. 2003.
10. A. Kumar, D. C. M. Wong, H. Shen, and A. K. Jain, “"Personal verification using palmprint and hand geometry biometric,” *Proc. AVBPA*, pp. 668-675, Guildford, UK, June 2003.
11. A. K. Jain and A. Ross, “Learning User-specific Parameters in a Multibiometric System”, *Proc. ICIP 2002*, Rochester, New York, Sep. 2002.
12. L. Osadciw, P. Varshney, and K. Veeramachaneni, “Improving personal identification accuracy using multisensor fusion for building access control applications,” *Proc. ISIF* 2002, pp. 1176-1183, 2002.
13. P. Verlinde, G. Chollet, and M. Acheroy, “Multi-modal identity verification using expert fusion,” *Information Fusion*, vol. 1, pp. 17-33, 2000.
14. J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, “On combining classifiers," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 20, pp. 226-239, Mar. 1998.
15. M.-H. Yang, D. J. Kriegman, and N. Ahuja, “Detecting faces in images : A Survey,” *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 24, pp. 34-58, Jan. 2002.
16. Visage Technology Inc., http://www.viisage.com
17. M. A. Turk and A. P. Pentland, “Eigenfaces for Recognition,” *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-76, 1991.
18. S. Prabhakar and A. K. Jain, "Decision level fusion in fingerprint verification," *Pattern Recognition*., vol. 35, pp. 861-874, 2002.

19. R. Brunelli and D. Falavigna, “Personal identification using multiple cues,” *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 17, pp. 955-966, Oct. 1995.
20. J. Kittler, J. Mastas, K. Jonsson, and M. V. R. Sánchez, “Combining evidences in personal identity system,” *Pattern Recognition Letters*, vol. 18, pp. 845-852, 1997.
21. A. Kumar, “Neural network based detection of local textile defects,” *Pattern Recognition*, vol. 36, pp. 1645-1659, 2003.
22. M. Riedmiller and H. Braun, “A direct adaptive method for faster backpropagation learning: The PROP algorithm,” *Proc. Intl. Conf. Neural Networks*, vol. 1, pp. 586-591, Apr. 1993
23. The Olivetti Research Database of Faces; http://www.cam-orl.co.uk/facedatabase.html
24. A. Ross, A. K. Jain, J.-Z. Qian, “Information fusion in biometrics,” *Pattern Recognition Lett*., vol. 24, pp. 2115-2125, Sep. 2003.
25. A. K. Jain and K. Karu, “Learning texture discrimination masks,” *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 18, pp. 195-205, 1996.
26. A. Kumar and G. Pang, “Defect detection in textured materials using Optimized filters,” *IEEE Trans. Systems*, *Man*, *and Cybernetics* : *Part B, Cybernetics*, vol. 32, pp. 553-570, Oct. 2002.
27. A. Mahalanobis and H. Singh, “Application of correlation filters for texture recognition,” *Appl. Opt*., vol. 33, pp. 2173-2179, Apr. 1994.
28. M. Unser, Local linear transforms for texture measurements,” *Signal Process*., vol. 11, pp. 61-79, 1986.
29. T. Randen and J. H. Husøy, “Texture segmentation using filters with optimized energy separation,” *IEEE Trans. Image Processing*, vol. 8, pp. 571-582, Apr. 1999.
30. S. Chen, C. F. N. Cowan, and P. M. Grant, “Orthogonal least squares learning algorithm for radial basis function networks,” *IEEE Trans. Neural Networks*, vol. 2, pp. 302-309, Mar. 1991.