Abstract— An accurate automatic personal identification is critical in a wide range of application domains such as access control and surveillance systems. Biometrics technologies have been increasingly adopted to provide identification with a high degree of confidence. A number of biometric traits exist and are in use in various applications. As one of the developing biometric techniques, palmprint identification is becoming a popular and convincing solution for identifying persons’ identity since palmprint is proved to be a unique and stable personal physiological characteristic. In this paper, two representations of the unique palmprint are integrated in order to construct an efficient multimodal identification system. For that, the two feature vectors are extracted and used for training two different classifiers. Thus, for each palm image, two feature extraction techniques are used, these are: 2D Block based Discrete Cosine Transform (2D-BDCT) and 2D CoNtourlet Transform (2D-CNT). Subsequently, the two feature vectors have been modeled as Hidden Markov Model (HMM). The two sub-systems outputs are fused at the matching score level. The experimental results showed that the designed system achieves an excellent identification rate and provide more security than unimodal biometric-based system.

Keywords—Biometrics, identification, Palmprint, HMM, 2DBDCT, Contourlet, Data fusion.

I. INTRODUCTION

This template, modified in MS Word 2003 and saved as “Word 97-2003 & 6.0/95 – RTF” for the PC, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons; (1) ease of use when formatting individual papers, (2) automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and (3) conformity of style throughout a conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, such as multi-leveled equations, graphics, and tables are not prescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.
II. PROPOSED MULTIMODAL IDENTIFICATION SYSTEM

Fig. 1 shows the block-diagram of the proposed multi-representations based multimodal identification system. For the both systems, to enroll into the system database, the user has to provide a set of training images. Typically, an observation vector is extracted from each image which describes certain characteristics using the both methods, 2D-BDCT and 2D-CNT technique, and modeling using HMM, then, the models parameters are stored as references models. For identification, the same observation vectors are extracted from the test images and the log-likelihood probability is computed using all of models references in the database. For both sub-systems, and based on the resulting fusion of the obtained normalized matching scores, a decision of accepting or rejecting the user is then made.

III. PALMPRINT PREPROCESSING

In order to localize the palm area, the first step is to preprocess the palm images; we use the preprocessing technique described in [6] to align the palmpints. In this technique, Gaussian smoothing filter is used to smoothen the image before extracting the ROI sub-image and its features. After that, Otsu’s thresholding is used for binarized the hand. A contourfollowing algorithm is used to extract the hand contour. The tangent of the two stable points on the hand contour (they are between the forefinger and the middle finger and between the ring finger and the little finger) are computed and used to align the palmpint. The central part of the image, which is 128 × 128, is then cropped to represent the whole palmpint. Fig. 2 shows the palmpint pre-processing steps.

IV. FEATURE EXTRACTION TECHNIQUES

The feature extraction module processes the acquired biometric data and extracts only the salient information to form a new representation of the data. In our method, the ROI sub-image is typically analyzed using the both 2D-BDCT and 2D-CNT techniques transform. After construction of the observation vectors from the ROI sub-image, an HMM model of each observation vector is constructed.

A. 2D-DCT based feature extraction

The block-based approach partitions the input image, with size \( H \times W \), when \( H = 128 \) and \( W = 128 \), into small non-overlapped blocks; each of them is then mapped into a block of coefficients via the 2D-DCT [7]. Most popular block size is...
commonly set to $M \times M$ with $M = 8$. The number of blocks extracted from each ROI sub-image equals to:

$$\eta = [\eta_1] \times [\eta_2] = \left[\frac{128}{8}\right] \times \left[\frac{128}{8}\right] = 16 \times 16 = 256 \text{ blocks} \quad (1)$$

Then, we form a feature vector from the 2D-DCT coefficients of each image block. The 2D-DCT concentrates the information content in a relatively few transform coefficients top-left zone of block, for this, the coefficients, where the information is concentrated, tend to be grouped together at the start of the reordered array. Thus, a suitable scan order is a zigzag starting from the DC (top-left) coefficient [8]. Starting with the DC coefficient, each coefficient is copied into a onedimensional array. So, each block can be represented by a vector of coefficients:

$$O_{ij} = [F_{ij}(0,0) \ F_{ij}(0,1) \ F_{ij}(1,0) \cdots F_{ij}(u,v)]^T \quad (2)$$

$F_{ij}$ are 2D-DCT coefficients $u,v$ are chosen as well as the identification rate was maximum. Thus, $u,v \in [0 \cdots 7]$ and the size of $O_{ij}$ is $\tau \times \eta \in [1 \cdots 64]$. Finally, the results $O_{ij}$ of a blocks image are combined in the single template as follows:

$$V_{o} = [O_o \ O_u \ O_v \ O_{\eta}] \quad (3)$$

where the size of resulting observation vector is $[\tau \ \eta]$.

Finally, an HMM model of each observation vector is constructed.

B. 2D-CNT based feature extraction

Discrete contourlet transform is a multi-scale and directional image representation that uses first a wavelet like structure for edge detection, and then a local directional transform for contour segment detection [9]. In general the application of 2D-CNT to an image involves two stages. A Laplacian Pyramid (LP) [10] is first used followed by the application of a Directional Filter Bank (DFB) [11].

To create an observation vector, the ROI sub-image is transformed into a sub-bands form using 2D-CNT technique. Then, the palmprint feature vectors are created by combining same of these bands. Thus, in this work, the ROI subimage is transformed on two 2D-CNT levels. However, an feature vector, $O_{CNT}$, can be extracted based on the different bands in the two levels. Then, $O_{CNT}$ is compressed using Principal Components Analysis (PCA) [12] method and some of principal components are selected for representing the final observation vectors $V_{o}$.

V. MODELING PROCESS

An HMM is a Markov chain with a finite number of states [13]. Although the Markov states are not directly observable, each state has a probability distribution associated with the set of possible observations. Thus, an HMM is characterized by: a state transition probability matrix $(A)$, an initial state probability distribution $(\pi)$ and a set of probability density functions associated with the observations for each state $(B)$. A compact notation $\lambda = (A,B,\pi)$ can represent the complete parameter set of the model. Finally, forward-backward recursive algorithm, Baum-Welch algorithm, and Viterbi algorithm are used to solve evaluating, training, and decoding, respectively.

VI. MATCHING PROCESS

After extracting the observation vectors corresponding to the test palmprint, the probability of the observation sequence given a HMM model is computed via a viterbi recognizer. The model with the highest log-likelihood is selected and this model reveals the identity of the unknown palmprint. Thus, during the identification process, the characteristics of the test image are extraction. Then the log-likelihood score of the observation vectors given each model, $P(V_x|\lambda_j) = \ell(V_x|\lambda_j)$, is computed. Therefore, the score vector is given by:

$$\ell(V_x) = \ell(V_x,\lambda_1) \ \ell(V_x,\lambda_2) \ \ell(V_x,\lambda_3) \cdots \ell(V_x,\lambda_d) \quad (4)$$

Where $d$ represents the size of model database.

An important aspect that has to be addressed in identification process is the normalization of the scores obtained. Normalization typically involves mapping the scores obtained into a common domain. Thus, a $Min-Max$ normalization scheme [14] was employed to transform the log-likelihood scores computed into similarity scores in the same range.

$$\varphi_y = \frac{\varphi - min(\varphi)}{max(\varphi) - min(\varphi)}$$

where $\varphi_y$ denotes the normalized log-likelihood scores. However, these scores are compared, and the highest score is selected. Therefore, the best score is $D_0$ and its equal to:

$$D_0 = max(\varphi_y) \quad (6)$$

Finally, a threshold $T_0$ regulates the system decision. The system infers that pairs of biometric samples generating score, $D_0$, higher than or equal to $T_0$ are belong to the same person. Consequently, pairs of biometric samples generating score lower than $T_0$ are they belong to different persons.

VII. FUSION PROCESS

Combining two or more palmprint representations provides new independent information that efficiency improves identification system performance. Besides complicating spoof attacks, multi-modal biometrics system has many advantages, such as providing greater universality to the system [15]. Moreover, the fusion of two modalities allows better reduction of the impact of system failure than uni-modal biometrics. Information fusion can occur in any of the system levels, in our work, we focus on fusion at matching score level [16]. The fusion is realized using five simple rules. These rules consist of the sum (SUM) and WeightFTed-sum (WHT) of the two similarity measures, their MINimum (MIN) and MAXimum (MAX) of both and finally their MULTIplication (MUL). The final decision of the classifier is then given by choosing the class, which maximizes the fused similarity measures between the sample and the matching base.
VIII. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experimental database

The palmprint images used in the current experiment were selected from the Hong Kong Polytechnic University (PolyU) Palmprint Database [16]. There are totally 1980 palmprint images (grayscale image in BMP format) of 165 different palms in the database. In this database, around 12 samples from each of these palms were collected in two sessions, where around 6 samples were captured in the first session and the second session, respectively. The resolution of all the original palmprint images is 384 × 284 pixels.

B. Simulation Results

In the system-design phase (all experiments), four images are randomly selected of twelve images of each class (person) were used in the enrolment stage to create the system database; the remaining eight images were used for testing. In the following tests, we setup a database with size of 150 classes, which are similar to the number of employees in small to medium sized companies. Thus, the client experiments were performed by comparing eight test images with the corresponding class in the database. A total of 1200 comparisons were made. The impostor experiments were performed by comparing the eight images with each class in the database. A total of 89400 impostor experiments were made.

1) Unimodal test results: At the first stage, we conducted several experiments to selection the best number of 2D-BDCT coefficients in each block. Thus, the 2D-BDCT coefficients reflect the compact energy of different frequencies. Most of the higher frequency coefficients are small and they become negligible. As a result, the features derived from the 2D-BDCT computation is limited to an array of summed spectral energies within a block in frequency domain [17]. In this section, we present the identification accuracy of our system as a function of the number of 2D-BDCT coefficients used. The performance evaluation was repeated for various numbers of 2D-BDCT coefficients, and the results are as a Genuine Acceptance Rate (GAR) are computed. The obtained results shows how the number of 2D-BDCT coefficients used may have an effect on the performance of our system. Thus, the obtained results shows that the identification accuracy becomes very high at certain points, where it actually exceeds 96 % and a slight decrease in identification accuracy as we go to higher numbers of coefficients. Also, note that only 20 coefficients are enough to achieve good accuracy.

At the second stage, we conducted several experiments to investigate the effectiveness of the feature extraction methods. For this, experiment was conducted using the two proposed methods (2D-DCT-HMM and 2D-CNT-HMM). Our goal is to choose the method yield the best performance. Therefore, by varying these methods we can choose the method which minimizes the systems error such as Equal Error Rate (EER).

Thus, in the case of open set identification, the Receiver Operating Characteristic (ROC), which is a plot of False Reject Rate (FRR) against False Accept Rate (FAR), curves for two distinct methods are shown in Fig. 3 (a). This figure compares the identification performance of the system varying feature extraction methods. After analyzing this figure we were able to conclude that the 2D-CNT-HMM method based system achieved the best performance, it can achieve an EER equal to 0.243 % at a threshold To = 0.7077. Poor results are obtained

**TABLE 1: OPEN SET UNIMODAL IDENTIFICATION SYSTEMS PERFORMANCES**

<table>
<thead>
<tr>
<th>DATABASE</th>
<th>2D-DCT-HMM</th>
<th>2D-CNT-HMM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>To FAR FRR</td>
<td>To FAR FRR</td>
</tr>
<tr>
<td>152 Persons</td>
<td>0.9000 4.1409 0.9166</td>
<td>0.6000 0.7925 0.0000</td>
</tr>
<tr>
<td></td>
<td>0.9342 1.5000 1.5000</td>
<td>0.7077 0.2430 0.2430</td>
</tr>
<tr>
<td></td>
<td>1.0000 0.0402 6.3333</td>
<td>1.0000 0.0033 1.667</td>
</tr>
</tbody>
</table>

Fig. 3. Unimodal identification system test results. (a) The ROC curves with respect to the feature extraction methods, (b) The ROC curve for the 2 D-CNT based unimodal identification system and (c) The CMC curves with respect to the feature extraction methods.
when using 2D-DCT-HMM method, in this case the system work with an EER equal to 1.500 % at a $T_0 = 0.9342$. The ROC curve for the best case is displayed in Fig. 3(b).

Compared with other existing unimodal systems, the proposed open set identification system has achieved better results expressed in terms of the EER. Finally, Table 1 present the experiments results obtained for the two feature extraction methods in the case of open set identification system.

To further validate our idea we have run other test for the closed set identification case. The result is presented as a Cumulative Match Characteristics (CMC) curve in Fig. 3(c). The best result of Rank-One Recognition (ROR) is given as 99.000 % with lowest Rank of Perfect Recognition (RPR) of 45 in the case of the 2D-CNT-HMM based feature extraction.

1) **Multimodal test results**: The aim of this section is to investigate whether the system performance could be improved by using the integration or fusion of information from each method. Therefore, the results presented by the two methods are fused to make the system efficient using fusion at matching score level.

In the matching score level fusion, the individual matching scores from the two sub-systems are combined to generate a single scalar score, which is then used to make the final decision. During the system design we experimented five different matching score fusion schemes (described in section VII). Therefore, for the open set identification system, an experimental result at the EER point is shown in Fig. 4.(a). The experimental results show that SUM rule based fusion scheme get the best performance with a minimum EER equal to 0.011 % at the threshold $T_0 = 0.9184$, this performance work at a minimum EER equal to 0.1667 % and $T_0 = 0.7230$ when the WHT fusion rule is used. The MIN rule give an EER equal to 0.0833 % at the threshold $T_0 = 0.8306$. The MUL and MAX rules provide an EER equal to 0.0123 % and 0.0833 % with a threshold $T_0 = 0.8614$ and $T_0 = 0.9940$, respectively. Finally, all experiments, in term of EER, are described in Table 2. The results expressed as ROC curve, if the SUM rule is used, is plotted in Fig. 4.(b).

In the case of a closed set identification, a series of experiments were also carried out to select the best method. This has been performed by comparing all fusion rules to determine which fusion rule gives the best identification rate. Table 3 presents the experimental results obtained for all fusion rules. From Table 4, the best result of ROR produces an accuracy of 99.306 % with lowest RPR of 54 in the case of SUM rule. Thus, the CMC curve for SUM rule is shown in Fig. 4.(c).

### Table 2: Open Set Multimodal Identification Systems Performances (Matching Score Fusion)

<table>
<thead>
<tr>
<th>Fusion</th>
<th>SUM</th>
<th>WHT</th>
<th>MAX</th>
<th>MIN</th>
<th>MUL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_0$</td>
<td>EER</td>
<td>$T_0$</td>
<td>EER</td>
<td>$T_0$</td>
<td>EER</td>
</tr>
<tr>
<td>2D-DCT-CNT</td>
<td>0.9184</td>
<td>0.011</td>
<td>0.7230</td>
<td>0.1667</td>
<td>0.9940</td>
</tr>
</tbody>
</table>

### Table 3: Closed Set Multimodal Identification Systems Performances (Matching Score Fusion)

<table>
<thead>
<tr>
<th>Fusion</th>
<th>SUM</th>
<th>WHT</th>
<th>MAX</th>
<th>MIN</th>
<th>MUL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROR</td>
<td>RPR</td>
<td>ROR</td>
<td>RPR</td>
<td>ROR</td>
<td>RPR</td>
</tr>
</tbody>
</table>
IX. CONCLUSION AND FURTHER WORK

In this work, two biometric systems for person identification using palmprint images are proposed. For that, two feature extraction methods are used. The fusion is performed at the matching score level using several fusion rules. We have demonstrated, through the obtained results of the two systems and by establishing a comparison when using each method, that the 2D-CNT-HMM method is reliable for efficient person identification. For further improvement of the system, our future work will focus on the performance evaluation using a large size database, and a combination of palmprint information with other biometrics such as finger-knuckle-print to obtain higher accuracy identification performances.

X. REFERENCES


