Face Recognition using Multimodal Biometric Features

N. S. LakshmiPrabha, J. Bhattacharya, S. Majumder
Surface Robotics Lab
Central Mechanical Engineering Research Institute
Mahatma Gandhi Avenue,
Durgapur - 713209, West Bengal, India.
email: n_prabha_mech@cmeri.res.in, bjhilik@cmeri.res.in, sjm@cmeri.res.in

ABSTRACT
This paper presents a new multimodal biometric approach using face and periocular biometric. The available face recognition algorithm performance in presence of multiple variations such as illumination, pose, expression, occlusion and plastic surgery is not satisfactory. Also, periocular biometrics face problems in presence of spectacles, head angle, hair and expression. A method which can extract multiple feature information from a single source and can give a satisfactory performance even with less number of training images is desirable. Thus combining face and periocular data obtained from the same image may increase the performance of the recognition system. A detailed performance analysis of face recognition and periocular biometric using Gabor and LBP features is carried out. This is then compared with the proposed multimodal biometric feature extraction technique. The experimental results obtained using Muct and plastic surgery face database shows that the proposed multimodal biometric performs better than other face recognition and individual biometric methods.

KEY WORDS
Multimodal biometric, Face recognition, Periocular recognition, Gabor wavelets, Local Binary Patterns.

I. INTRODUCTION

Face recognition has been an important area for its growing demand in the field of security and surveillance. This is mainly preferred over other biometrics as it can be carried out using available resource and without subject’s participation. A new challenge to face recognition is plastic surgery [1] which allow evaders to roam without any fear about the face recognition systems. Facial plastic surgery is used to correct facial feature anomalies or cosmetic to improve the appearance and these changes are long-lasting or even permanent. Face recognition is a task performed by human with ease in day to day life. Human beings can recognize faces by considering many features such as view face at different angle, skin color, hair information, facial mark etc and tries to compare with different expression. This task also requires a capability to synthesize faces with different emotions, illumination, pose etc. Making an automated system to do the same, requires the capability to see the things in different way and obtain more information from the available resource. Existing face recognition algorithm encounters difficulty in the presence of multiple variations such as illumination, pose, expression and occlusion. These problems are reduced to some extent by using large number of distinct features.

Periocular area (i.e. a region surrounding the eye) is considered to be one of the most discriminative regions of a face [2] and is used for identifying an individual [3], [4], soft biometrics [5] for example. Face recognition considers the entire eye region as one of the many discriminative features [7]. This periocular biometric can also be fused with other biometric such as Iris [6]. Despite of being a discriminative region the periocular biometrics faces problems in presence of head angle or pose variations, spectacles, hair and facial expression for example. First two issues are more serious than the latter ones. Typical examples of above mentioned cases are shown in figure 1. Shape and texture of the lips also differ from individual to individual and may be used to provide additional support in increasing the performance. However, the information provided by the lips region is useful only when the lips are closed otherwise the portion of the teeth may create problem. When

Fig. 1. Typical cases in Periocular biometric.
there are different expressions or when the subject is talking, the performance of this method will decrease drastically. Thus the distinct features obtained from periocular biometric and lips data cannot serve as an individual biometric. Merging face recognition with other biometric may increase the performance, at the same time requiring another acquisition to obtain that information. A multi-modal system combining two different biometric from one source is more attractive. Fusing face and periocular region feature obtained from the same image may perform better. Thus this paper proposes a new multimodal biometric using face and periocular biometric.

This paper is organized in the following way, section II elaborates on the proposed multimodal biometric including feature extraction, feature dimension reduction and classification. The experimental results obtained using Muct [17] and Plastic surgery face database [1] is discussed in section III.

II. MULTIMODAL BIOMETRIC USING FACE AND PERIOCULAR

A Multimodal biometric combining face and periocular region features may perform well. This biometric method makes use of distinctive features from periocular region and overcomes the problem faced by periocular biometric such as head angle, expression and spectacles. The basic block diagram of the multi-modal biometric system proposed in this paper is shown in figure 2. Two types of feature extraction methods such as Gabor wavelets and Local Binary Pattern (LBP) are explained in section II-B. Feature extraction is done for both face and periocular images followed by feature dimension reduction using PCA. Classification is performed using Euclidean distance. Features from the face images are matched first and when there is a negative result periocular image features are matched. An attempt has been made in combining face and lips region images and the results obtained is discussed in section III.

A. Preprocessing and Normalization

This section explains about the normalization of face image to remove the background and obtaining periocular region from the face images. Preprocessing is performed on the normalized images.

1) Face data: Image normalization of face is done using pixel location of eye center. Head rotation angle is calculated using this eye center pixel location. Image is rotated opposite with this angle which results in horizontal eye alignment. Distance between eyes and the size of the final normalized images are predefined. Horizontal eye aligned image is resized using eye distance. The normalized image is obtained by cropping the resized image to the required size. Histogram equalization is done on the normalized face image in order to spread intensity values of all pixels inside the image. A face image with background and its preprocessed face image is shown in figure 3.

2) Periocular data: There is no database available with periocular region images. Only way to fetch this is using a face image. Periocular biometric is performed in 3 different ways such as overlapping, Non-overlapping, Strip. In this paper all this three periocular regions are experimented and the results are discussed in section III. Extracting this periocular region is a tricky job, as the size of the region increase then there will be intrusion of nose, forehead information. All this three different

Fig. 2. Block diagram of the multi-modal biometric system using face and Periocular biometric.

Fig. 3. A face image and preprocessed face part.

Fig. 4. Process of extracting these three periocular and lip region.
types of periocular regions are obtained using four significant points in the eye region. With these points the regions are warped separately as explained in previous section. The process of extracting these three periocular region is shown in figure 4. There are two levels of features obtained from periocular region such as both upper and lower eyelids, eye folds, and eye corners are level one feature. Skin texture, fine wrinkles, color, or skin pores etc are level two features [7]. In some cases they place mask in the eye region to address eye opening and closing issue. This paper considers both the level one and level two features without placing mask in the eye region assuming eyes are open in all the images. Similarly lips are fetched using two corner points and the figure 5 shows the different conditions of the periocular and lips region along with the face image.

B. Feature Extraction

The feature extraction method which gives more texture information performs better both in face and periocular. This paper considers two feature extraction methods such as Gabor wavelets [8], [9] and Local Binary Patterns (LBP) [10], [11], [12]. The feature extraction methods are explained using face images and the same is applicable for periocular images also.

1) Gabor wavelets: Local features in face images are more robust against distortions such as pose, illuminations etc, and a spatial-frequency analysis are often desirable to extract such features. With good characteristics of space-frequency localization, Gabor wavelet analysis is a suitable choice for face recognition purpose. The Gabor wavelets (kernels, filters) [8] can be defined as follows:

$$\psi_{\mu, \nu}(z) = \frac{|k_{\mu, \nu}|^2}{\sigma^2} e^{-\frac{|z|^2}{2\sigma^2}} [e^{ik_{\mu, \nu}z} - e^{-\frac{z^2}{2}}]$$  \hspace{1cm} (1)

where $\mu$ and $\nu$ define the orientation and scale of the Gabor kernels, the wave vector $k_{\mu, \nu}$, is defined as follows:

$$k_{\mu, \nu} = k_\nu e^{i\phi_\mu}$$ \hspace{1cm} (2)

where $k_\nu = \frac{k_{\text{max}}}{f^\nu}$ and $\phi_\mu = \frac{\pi \mu}{8}$. $k_{\text{max}}$ is the maximum frequency, $f$ is the spacing factor between kernels in the frequency domain and $z = (x, y)$, $\| \cdot \|$ denotes the norm operator. Gabor wavelets at five different scales, $\nu[0, 4]$, and eight orientations, $\mu[0, 7]$ are considered in this work with the following parameters: $\sigma = 2\pi$, $k_{\text{max}} = \pi/2$ and $f = \sqrt{2}$. The size of all this filters is 32 x 32 pixels.

An image can be represented by Gabor wavelet responses by convolving Gabor filters of different scale and orientation. The set of convolution coefficients for kernels at one image pixel is called a jet. The resulting output contains most important face features like eyes, mouth and nose edges, as well as moles, dimples and scars. Magnitude of convolved face image from Muct database is shown in Figure 6. Magnitude information of convolved face image is preferred because it makes data invariant under rotation or translation.

For each image after convolution there are 40 images containing extracted features. All these 40 images are converted into a feature vector. This increases the time consumption and memory requirements. This can be avoided by taking limited number of pixels from the feature images with regular spacing grids. When a test image is given and that is also Gabor feature extracted followed by vector concatenation.

2) Local Binary Pattern: Local Binary Patterns provides a powerful means of texture description [10]. LBP features are gray scale and rotation invariant texture operator. These features are more widely used for expression recognition [13], [14]. LBP features are also applied for face recognition task [11], [12]. LBP feature extraction is faster than any other feature
Consider a 3x3 pixels with center pixel \((x_c, y_c)\) intensity value be \(g_c\) and local texture as \(T = t(g_0, g_1, g_2, g_3, g_4, g_5, g_6, g_7)\) where \(g_i (i = 0, 1, 2, 3, 4, 5, 6, 7)\) corresponds to the grey values of the 8 surrounding pixels. These surrounding pixels are thresholded with the center value \(g_c\) as \(t(s(g_0 - g_c), ..., s(g_7 - g_c))\) and the function \(s(x)\) is defined as,

\[
s(x) = \begin{cases} 
1 & , x \geq 0 \\
0 & , x \leq 0 
\end{cases}
\] (3)

Then the LBP pattern at the given pixel is defined as an ordered set of the binary comparisons and the resulting value can be obtained using equation (4). An example of LBP operator is shown in figure 7.

\[
LBP(x_c, y_c) = \sum_{i=0}^{7} s(g_i - g_c)2^i
\] (4)

LBP feature extraction is performed on a face image and the resulting features are shown in figure 8 along with the feature histogram. To increase the feature strength and to get more details, the face images are divided into number of blocks. Figure 9 shows the face divided into 5 blocks (totally 25 blocks) and its feature histogram. When a test image is given as input the LBP histogram features are extracted which then used for classification purpose.

C. Feature dimension reduction using PCA

In many cases the size of the features are huge which causes problem such as computational complexity, memory requirement etc. At the same time discarding features will also effects the performance and in turn accuracy of the face recognition method will also be affected. Thus PCA is used here for dimension reduction. The local features extracted using Gabor wavelets and LBP are prone to noise. By performing PCA on the Gabor feature vector these issues can be addressed. This increases the recognition rate as the problem of PCA under illumination variation is eliminated by the Gabor or LBP features.

The deviation of each feature from the mean using the equations (5, 6) is calculated.

\[
\psi = \frac{1}{M} \sum_{n=1}^{M} I_n
\] (5)

\[
\phi_n = I_n - \psi
\] (6)

By taking the eigenvectors of the covariance matrix using equation (7), the variation among the training set can be obtained. This results in the eigenface space [15], [16]. All the training set are projected into the eigenface space using
equation (8).

\[ C = \frac{1}{M} \sum_{n=1}^{M} \phi_n \phi_n^T = AA^T \]  

(7)

\[ \omega_k = u_k \phi = u_k (I - \psi) \]  

(8)

Weight Matrix \( \Omega = [\omega_1, \omega_2, \ldots, \omega_M]^T \) is the representation of a training image in the eigenface space.

When a new test image is to be classified, the feature extracted from Gabor or LBP methods are also mean subtracted using equation (6) and projected onto the eigenface space using equation (8). Weight matrix of the test image is \( \Omega_T = [\omega_1^T, \omega_2^T, \ldots, \omega_M^T] \).

D. Classification

Euclidean distance is used as the classifier to identify the image in the training set to which the test image belong. Classification is performed by comparing the feature vectors (weight matrix) of the images in the training set with the feature vector (weight matrix) of the test image using Euclidian distance, \( \varepsilon_i \).

\[ \varepsilon_i^2 = ||\Omega_T - \Omega_i||^2 \]  

(9)

where \( \Omega_i \) is a vector describing the ith face image in the training set. Test image is classified as belonging to image i when the minimum of \( \varepsilon_i \) is below some chosen threshold value \( \Theta \). Threshold value is defined as half of the distance between the two most distant images in a class. Distance Threshold, \( \Theta = \frac{1}{2} max(||\Omega_i - \Omega_j||) \), where i and j are images from same class.

III. RESULTS AND DISCUSSION

The performance of the proposed multimodal biometric is tested with Muct [17] and plastic surgery face database [1]. From Muct database 500 images of 50 different people face with 10 images each is used. Number of images used for training is 5 per person, so totally 250 images and remaining 250 images are used as test set. Only test set images are used to check the performance of all the above mentioned methods.

The images available in the plastic surgery face database [1] are not properly aligned and there are lots of repetitions of the same face with different id. number (24644, 26302 and 28704 id. belongs to same person face images under rhinoplasty nose surgery for example). This will affect the performance of the face recognition method and the accuracy will also decrease. Totally 500 images from 250 different subject are used for experimentation. Before surgery images are used for training and after surgery images are used for testing the algorithm.

The face images are preprocessed and normalized as in section II-A1 to a size of 150 x 130 pixels. Number of scales and number of orientation is 5 and 8 respectively for Gabor wavelet feature extractor. Number of blocks used for Local Binary Pattern feature extractor is 9 (totally 81 blocks). The results obtained from face images using feature extraction and feature dimension reduction methods are given in table I.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Muct (%)</th>
<th>Plastic surgery (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gabor</td>
<td>69.8</td>
<td>65.33</td>
</tr>
<tr>
<td>LBP-PCA</td>
<td>80.8</td>
<td>71.33</td>
</tr>
<tr>
<td>LBP</td>
<td>70</td>
<td>60.00</td>
</tr>
<tr>
<td>LBP-PCA</td>
<td>80.4</td>
<td>66.00</td>
</tr>
</tbody>
</table>

Feature dimension reduction method performs better than feature extraction method. The size of periocular region such as strip, overlapping and Non-overlapping is 50 x 130, 50 x 65 pixels respectively. The size of extracted lips region is 50 x 100 pixels. The results obtained from periocular biometric and lips region using Muct and plastic surgery face database is given in table II. The results obtained using the proposed multimodal biometric using face and periocular data is shown in table III.

The results from periocular biometric are superior than face biometric in case of Muct database (normal face database) however its performances is not satisfactory for plastic surgery face database. This is possibly because the Plastic surgery face database has one image per person for training (before surgery) and periocular biometric furnishes good results when there more number of images for training. Extracting features from one periocular data will not provide sufficient information for classification and this may be because of head angle, expression, spectacles and hair etc. It’s clear from the table II that LBP-PCA features gives better results than Gabor and Periocular region provides more discriminative features compared to lips region.

In case of multi-modal biometric, Gabor-PCA and LBP-PCA performs equally and the performance of face with periocular

<table>
<thead>
<tr>
<th>Methods</th>
<th>Muct (%)</th>
<th>Plastic surgery (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye Strip</td>
<td>82.40</td>
<td>49.60</td>
</tr>
<tr>
<td>Gabor-PCA</td>
<td>78.00</td>
<td>45.20</td>
</tr>
<tr>
<td>Overlapping eyes</td>
<td>LBP-PCA</td>
<td>82.30</td>
</tr>
<tr>
<td>Overlapping eyes</td>
<td>Gabor-PCA</td>
<td>79.20</td>
</tr>
<tr>
<td>Non-overlapping eyes</td>
<td>LBP-PCA</td>
<td>82.40</td>
</tr>
<tr>
<td>Non-overlapping eyes</td>
<td>Gabor-PCA</td>
<td>77.20</td>
</tr>
<tr>
<td>Lips</td>
<td>52.00</td>
<td>27.20</td>
</tr>
<tr>
<td>Gabor-PCA</td>
<td>51.20</td>
<td>29.00</td>
</tr>
</tbody>
</table>
is better than face with lips data. The results obtained for Muct database (normal face database) and Plastic surgery face database using proposed method is so far the best when compared to other feature extraction, feature dimension reduction methods as noticed from table I, II and III.

### IV. CONCLUSION

The proposed multimodal biometric using face and periocular biometric performs better than individual biometric methods. The recognition rate obtained using Gabor and LBP features are very similar. The accuracy of the periocular biometric is superior when applied to the lips region data. The result from Muct database clearly shows that the discriminative nature of periocular region increases the performance of the recognition system than the face biometric and this is valid only when the number of images is reasonably large for training. In case of plastic surgery, the proposed method performs better than other methods. However it can be improved by considering more number of features.

### ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Mrs. S. Datta for the necessary support and assistance during various stages of work. I would also like to thank Prof. Gautam Biswas, Director of the Institute for his support and encouragement.

### REFERENCES