Chapter 7

Shearlet Based Multi-modal Face Recognition
Chapter 7

Shearlet Based Multi-modal Face Recognition

This chapter presents a face recognition based on the fusion of the intensity image and stereo depth map at the score level. The depth map is obtained by using triangulation method (Chapter 5). The shearlet transform is a multiresolution method, which is used to extract features from depth and intensity images independently. The principal component analysis (PCA) is used to reduce the dimension of feature set. The extracted Eigen features of intensity and depth map are combined at decision level to improve the face recognition rate. The combination of intensity and depth map is verified experimentally using benchmark face database. The experimental results show that the proposed multi-modal method is better than individual modality.

7.1 Introduction

Human face recognition is a complicated problem in computer vision (Shang1). Face recognition is a challenging task, because it is a real world problem. The human face is a complex natural object that tends not to have easily identified edges and features. Because of this, it is difficult to develop a
mathematical model of the face that can be used as prior knowledge when analyzing a particular image. Recently face recognition is attracting much attention in society of network, multimedia and information access. Areas such as network security, content indexing and retrieval.

Multiresolution analysis tools like wavelets have been found very useful in analyzing the image contents, hence they widely used in image processing, pattern recognition and computer vision. Over the years other multiresolution techniques such as contourlets, ridgelets, curvelets etc., were developed. In chapter 6 we discussed curvelet based multi-modal face recognition techniques. Curvelets provides optimally sparse approximations of anisotropic features. But it has two drawbacks, firstly the curvelet system is not singly generated, i.e., it is not derived from the action of countably many operators applied to a single or finite set of generating functions. Secondly its construction involves rotations and these operators do not preserve the digital lattice which prevents a direct transition from the continuum to the digital setting. Shearlet transform is a recent addition. It has been used in image denoising (Glenn et al., 2009), edge analysis (Sheng et al., 2009) and image separation (Kutyniok et al., 2012) but not much work has been done to solve pattern recognition problems.

Shearlets provides a unified treatment of continuum models as well as digital models, allowing, for instance, a precise resolution of wavefront sets, optimally sparse representations of cartoon-like images, and associated fast decomposition algorithms (Kutyniok et al., 2010). Shearlet systems are generated by one single generator with parabolic scaling, shearing, and translation operators applied to it, in the same way wavelet systems are dyadic scalings and translations of a single function, but including a directionality characteristic owing to the additional shearing operation and the anisotropic scaling. The shearing operation in fact provides a more favorable treatment of directions, thereby ensuring an unified treatment of the continuum and digital realm as opposed to curvelets which are rotation-based in the continuum realm. Mohamed et al.(2009) used the combination
of the shearlet transform to capture geometric information to derive a very efficient representation of facial templates and the PCA-based approach to design a fusion step by a refined model. Mohamed et al. used shearlet network to extract features which takes advantage of the sparse representation properties of shearlets in biometric applications.

In this chapter, we introduced a shearlet coefficients based multi-modal 2D +3D face recognition technique. 3D data or depth map of face images is extracted from stereo face images using adaptive weight based stereo matching method (chapter 4). The multiresolution technique such as shearlet transform is applied independently to depth and intensity image of faces to extract shearlet features. The size of shearlet features is reduced through Principal Component Analysis (PCA). The KNN classifier is used to measure decision scores of 2D and 3D facial images independently. The obtained decision scores of 2D and 3D features are combined at the decision level using OR rule. The proposed method is evaluated by conducting experiments on the stereo face database.

The remaining sections of the chapter are organized as follows. In section 7.2, a brief description of shearlet transforms is presented. The experimental results are presented in section 7.3. Finally, the chapter summary is given.

7.2 Shearlet Transformation

Kutyniok and Labate (2012) proposed a new multi-resolution analysis tool such as shearlet transforms. Shearlet transform is multiscale and multidirectional transform. The shearlet transform is a new tool for analyzing the intrinsic geometrical features of an image using anisotropic and directional window functions. The directionality of shearlet transform is achieved by applying integer powers of a shear matrix, and those operations preserve the structure of the integer lattice. Shearlets parameterize directions by slope rather than angles. The shear matrix preserves the structure of the
integer grid, which is key to enabling an exact digitization of the continuum domain shearlets. The shearlet transform produces a low redundancy sparse and anisotropic feature representation of the image. Emergence of shearlets transform provides enhanced directional and edge representation. So far, it is used in the field of image denoising, image compression, etc. but not much work has been done in shearlet based pattern recognition. The shearlet based 2D face recognition has been presented in (Zeng et al., 2013; Schwartz et al., 2011). Zeng et al. (2013) proposed a new method for face description and recognition using shearlet transform and principle component analysis. The shearlet transform is applied on face images to exploit directional information, along with conventional scaling and translation parameters. Then face feature is obtained by principle component analysis. The experiment is conducted on ORL and FERET face database. Schwartz et al. (2011) introduced a feature descriptor called Histograms of Shearlet Coefficients (HSC). They employed histograms to estimate the distribution of edge orientations and on the accurate multi-scale analysis provided by shearlet transforms.

The shearlet systems are generated by parabolic scaling, shearing, and translation operators applied to one single generator. For each $a > 0$ and $s \in \mathbb{R}$ let $A_a$ denote the parabolic scaling matrix and $S_s$ denote the shear matrix of the form

$$A_a = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, \text{ and } S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \quad (7.1)$$

respectively. To provide an equal treatment of the $x$ and $y$ axis, the frequency plane is split into the four cones $C_1$ to $C_4$, defined by
Chapter 7. Shearlet Based Multi-modal Face Recognition

\[ C_l = \begin{cases} 
(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \geq 1, \left| \frac{\xi_1}{\xi_2} \right| \geq 1 : l = 3, \\
(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \geq 1, \left| \frac{\xi_1}{\xi_2} \right| \leq 1 : l = 1, \\
(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \leq -1, \left| \frac{\xi_1}{\xi_2} \right| \geq 1 : l = 4, \\
(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1 \leq -1, \left| \frac{\xi_1}{\xi_2} \right| \leq 1 : l = 2. 
\end{cases} \tag{7.2} \]

For cone \( C_3 \), at scale \( j \in \mathbb{N}_0 \), orientation \( s = -2^j \ldots, 2^j \), and spatial position \( m \in \mathbb{Z}^2 \), the associated shearlets are then defined by their Fourier transforms

\[ \sigma_\eta (\xi) = 2^{-j^3} \psi_1 \left( s + 2^{j^4} \right) \chi C_3 (\xi) e^{2\pi i (A_4 s m^4)}, \tag{7.3} \]

where \( \eta = (j, s, m, l) \) index scale, orientation, position, and cone. The shearlets for \( C_1, C_2, \) and \( C_4 \) are defined likewise by symmetry, as illustrated in Figure 1 and the discrete shearlet system by

\[ \sigma_\eta : \eta \in \mathbb{N}_0 \times \{-2^j, \ldots, 2^j\} \times \mathbb{Z}^2 \times \{1, 2, 3, 4\} \tag{7.4} \]

The definition shows that shearlets live on anisotropic regions of width \( 2^{2j} \) and length \( 2^{-j} \) at various orientations. The discrete shearlet transform algorithm is as follows:

- Classical Fourier transformation.
- Change of variables to pseudo-polar coordinates.
- Weighting by radial density compensation factor.
- Decomposition into rectangular tiles.
- Inverse Fourier transform of each tiles.

7.3 Experimental Results

The performance of the proposed shearlet based multi-modal face recognition method is evaluated by conducting various experiments on stereo face database. The face region is extracted from an image by considering a nose
Figure 7.1: The cones $C_1 - C_4$ and the centered rectangle $R$ in the frequency domain.

tip as reference and the noises in the cropped images are filtered by using median filters. The cropped color images are converted to the gray-level image by averaging R, G and B color components. The K-nearest neighbor (KNN) classifier with Euclidean distance and KNN classifier with cosine distance are used for obtaining the decision scores for each set of features of intensity and depth map.

We considered the left image as the intensity image. The depth image is obtained by employing stereo vision technique explained in the Chapter 4. The shearlet transform is applied on depth and intensity face images to extract the shearlet features. The shearlet transform decomposes the face images into number of horizontal and vertical cones. The shearlet coefficients of horizontal cones and the shearlet coefficients of vertical cones of five levels are concatenated to represent features of face image. The obtained multi orientation information describes face image by a subset of filtered images containing shearlet coefficients which represents the face.
The extracted shearlet features were large in size, which increases the processing time. Increase in the processing time decreases the performance of the overall system. Better performance can be obtained if the processing time is reduced, which can be achieved by reducing the extracted feature vector without losing the important or dominant features. This reduction in features is achieved through Principal Component Analysis (PCA) (Chapter 6).

![Figure 7.2: Recognition rates at various K value for different shearlet levels.](image)

We have conducted several experiments by varying K values from 1 to 6 and shearlet levels from 2 to 5 with number of directions \([9, 9, 5, 5, 3]\) considered. The estimated results of experiments provide us to identify, which K value and shearlet level maximize the recognition accuracy. The plots in Figure 7.2 shows that the maximum recognition rate can be achieved at \(K=2\) and level=5. In our all experiment, the face images are decomposed using shearlet transform at level = 5. The experiments have proved that this level can balance between recognition rate and recognition time, and
larger levels of the transformation will not significantly improve recognition rate. The highest recognition rate 99.38% is obtained at a number of principal components is equal to 40 as shown in Figure 7.3.

![Recognition rate at varying the number of principal components.](image)

**Figure 7.3:** Recognition rate at varying the number of principal components.

**Single Modal Face Recognition**

In the first experiment, the curvelet transform is applied to intensity and depth images. In the second experiment, shearlet features of intensity image and depth image to describe an individual. The rank one recognition rate of first experiment is shown in the first two columns and of second experiment is shown in the last two columns of Table 3. The shearlet features outperform compared to curvelet features for intensity images. Similarly the shearlet features of depth map yields highest recognition rate than curvelet features for depth map. Hence, we can conclude that shearlet features characterize the individual faces thoroughly for face recognition than curvelet features for both intensity and depth images. From the above two set of experiments, we can conclude that, shearlet features of depth images perform better to characterize the faces for face recognition.
Table 7.1: Comparison of Rank 1 recognition rate for stereo face database.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Intensity Curve-face</th>
<th>Depth Curve-face</th>
<th>Intensity shearlet</th>
<th>Depth shearlet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition Rate (%)</td>
<td>88.68</td>
<td>90.13</td>
<td>89.16</td>
<td>90.36</td>
</tr>
</tbody>
</table>

In order to verify the effectiveness of shearlet features for single modal face recognition, we have compared obtained recognition results of shearlet features with PCA features with curvelet with PCA. For intensity face images, the curvelet with PCA features yields 88.68% recognition accuracy whereas shearlet with PCA features gives 89.16% accuracy. Similarly, for depth images, curvelet with PCA based face recognition accuracy is 90.13%, whereas shearlet with PCA based recognition rate is 90.36%. The shearlet with PCA features outperform compared to curvelet with PCA features for intensity images. Similarly the shearlet with PCA features of depth map yields highest recognition rate than curvelet with PCA features for depth map. Hence, we can conclude that shearlet with PCA features characterize the individual faces thoroughly for face recognition than curvelet with PCA features for both intensity and depth images.

Multi-modal Face Recognition

The rank 1 recognition rate of the proposed method increases from 94.6 %, training with one image per subject to 99.83% with four images per subject. The rank 1 recognition rate of shearlet with PCA features of depth, shearlet with PCA features of intensity method increases from 86.6%, 83.3% with the training number 1 to 94.08%, 86.6% with training number 4 respectively. The cumulative recognition rate versus rank is shown in Figure 7.4 for number of training images from 1 to 4. It shows that the proposed method gives the better performance when the number of training images of each subject varies from 1 to 4 and rank varies from 1 to 9.
Chapter 7. Shearlet Based Multi-modal Face Recognition

The performance of proposed methodology has been evaluated using Receiver Operating Characteristics (ROC) and is shown in Figure 7.5.

From the results of Table 7.2, we can conclude that the shearlet based method significantly improves the recognition performance in comparison to wavelet based method and curvelet based method. The curvelet transform can able to capture multidirectional features, but wavelet transform focus mainly on horizontal, vertical, and diagonal features, which are not
dominant in most face images. Shearlet based method had higher recognition rate in comparison to both the wavelet and curvelet due to the representing data with anisotropic information at multiple scales and edges in face region are precisely detected and located. So that features obtained from the Shearlet will have more powerful information compared to the features from the wavelet sub-bands and curvelet. Shearlet based features yielded accuracy rate 99.83%, which significantly improved accuracy ranges for curvelet based features and wavelet based features. The recognition rate for wavelet based method is 92.52%, while curvelet based method is 99.46%. This was also expected since the shearlet and curvelet transform are

\[
\text{Table 7.2: Comparison of recognition rate of multi resolution analysis methods.}
\]

<table>
<thead>
<tr>
<th>Methods (Depth + Intensity)</th>
<th>Recognition Rate(%)</th>
<th>Computation Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet+PCA based</td>
<td>92.52</td>
<td>1.27</td>
</tr>
<tr>
<td>Curvelet+PCA based</td>
<td>99.46</td>
<td>1.42</td>
</tr>
<tr>
<td>Shearlet+PCA based</td>
<td>99.83</td>
<td>1.31</td>
</tr>
</tbody>
</table>

Figure 7.5: The ROC curves.
able to capture multidirectional features in wedges, as opposed to limited directional sensitivity in wavelet transform.

7.4 Chapter Summary

This chapter introduced a novel approach for multi modal face recognition based on exploiting the features using shearlet transform. The size of shearlet features are huge, so that PCA is applied to reduce the size of features. Shearlet based method had higher recognition rate in comparison to both the wavelet and curvelet due to the data with anisotropic information at multiple scales and edges in face region are precisely detected and located. The depth map of a stereo pair is computed using wavelet based stereo matching technique. The KNN classifier is employed independently for extracted shearlet features of intensity and depth map to compute the decision score. These computed decision scores of intensity and depth map are fused at decision level using OR rule. The depth information is more robust than intensity information under pose variations. Thus, the combination of intensity and depth can improve the recognition rate.