Chapter 8

Conclusions
Chapter 8

Conclusions

In this thesis, we have investigated the two major research problems such as reconstruction of 3D face surface and face recognition, which are aims to obtain an accurate 3D surface model of face and recognition of face accurately. We have proposed novel approaches and demonstrated that it is possible to reconstruct 3D face surface using stereo images based on stereo vision technique and also computed depth map helps to increase the recognition rate by combining its information with 2D face image.

8.1 Conclusions

We have presented an approach for estimating the disparity map of color stereo face images using multi wavelet based coarse-to-fine technique. The discrete wavelet transform is applied to stereo images to have multiple spatial frequency channels. In order to achieve a better results at featureless regions color face images in spatial domain is transformed to frequency domain by applying discrete wavelet transform. Then adopted a coarse to fine strategy to refine the disparity map up to the finest level. We have shown that our approach works well compared to state of art local stereo correspondence approaches.

We introduced color-segment based stereo correspondence for face images in frequency domain. Segment based methods are attracted attention
due to their good performance in handling boundaries and texture less regions. They are based on the assumption that the scene structure can be approximated by a set of non overlapping planes in the disparity space and that each plane of reference image is coincident with at least one homogeneous color segment in the reference image. Segment based methods perform well in reducing the ambiguity associated with texture less regions and enhancing noise tolerance. The performance of the proposed approach is estimated by varying the illumination and poses of the face images.

We proposed an approach for estimating the disparity map for face images using adaptive weight based method. The segment based method is very difficult to dealing with highly textured regions. The use of adaptive weights in segment based method enforce smoothness over highly textured planes. In the proposed approach, the adaptive weight is assigned to each pixel of segments in order to reduce the foreground fattening effect and also to obtain accurate results in homogeneous region. From the experimental results, we have shown that the proposed approach yields superior results for face images. We also showed that the proposed approach works well for illumination and pose variation face images.

We have presented an approach for reconstructing the 3D surface of face images using stereo vision. The proposed approach consist of different steps i.e. stereo image rectification, stereo correspondence using adaptive weight based approach, depth map estimation using triangulation principle, texture mapping. From experimental results, we have shown that the proposed approach for 3D face surface reconstruction yields the better 3D surface model for face images. The proposed approach is helpful in reconstructing face images. Using the reconstruction method, we are able to create many depth map (3D) of human faces in our 3D database. Then, we used the depth map information to do the 3D face recognition.

Our proposed stereo matching methods reduces the foreground fattening effect and produces smooth disparity at featureless regions with little more computation time.
We have introduced a multi-modal face recognition method based on fusion of intensity (2D) image and stereo depth map (3D) based on curvelet transformation. The extraction of reliable features from 2D images is difficult and is suffers from a variety of possible interpretation errors. The recognition performance of 2D face recognition systems is limited. The use of additional 3D information is expected to improve the accuracy of the recognition system, due to their relative independence from illumination and pose variations. Facial feature extraction is the most important step in face recognition. The curvelet transform is the multiresolution technique has improved directional elements and better ability to represent edges and singularities along curves of face images. The proposed method is verified experimentally using benchmark face database, and results are compared with monocular face recognition technique.

We have developed another technique for multi-modal (3D and 2D) face recognition based on the shearlet transform. Curvelets need to be band-limited and can only have very good spatial localization if allows high redundancy. But, the shearlet transform produces a low redundancy sparse and provides enhanced directional and edge representation. The experimental result shows that the performance of the proposed method is high when compared to wavelet, ridgelet and curvelet based face recognition systems.

The computation cost of our proposed multi-modal face recognition methods is more compared to single-modal face recognition methods with high recognition accuracy.

8.2 Future Directions

The proposed research work has been carried out on static images of stereo face database i.e both camera and objects are kept at a fixed distance while acquiring an images. The outcome of the research work clearly shows better 3D face surface model and recognition rate for face images. The proposed
research work can be extended for video sequences, where instead of using still images, we can work on video frames.

Further improvements could be made to the implementation through the addition of a true n-view reconstruction algorithm. At present the reconstruction rig is considered as 2 independent stereo pairs. This could lead to significant improvements in the correlation system by providing additional data to increase the accuracy of the match confidence estimates.

The image which consist of foreground and background object with same color will introduce much interference for distinguishing the distances. Even though both the foreground and background object having same color are having different disparities, as both objects are classified as one object, they are assigned same disparity. So in future we will try to solve this problem.

We only concentrated on the depth map of the face and ignored left, right, back and top sides of the model. As a result, the 3D reconstruction is done only for the face, not for the complete head. The task of discovery of an automatic system to construct a complete head model remains as a topic for future work.

Recognizing face from video is probably the most difficult problem. One should be able to track the location of face, estimate the pose of face and then recognize face. The proposed work can be expand to a recognizing face from video, face tracking, pose estimation and expression recognition.

We examined the fusion of 2D and 3D face modalities. The fusion of these two modalities has provide a consistent performance improvement. Other modalities could be included to improve performance and robustness.