CHAPTER 7
CONCLUSION AND FUTURE WORK

7.1 INTRODUCTION

To combat terrorism and to protect people and assets, an automated video surveillance has become a necessity and is also a focus of research of image processing community. The existing challenges and issues in video surveillance include, shadow removal for motion segmentation, object classification in the presence of occlusions, consistent labeling in a homogeneous environment, gender classification and human action classification in a crowded environment. All these issues are critical in nature and have limited solutions and hence, they need to be addressed for the successful deployment of the intelligent video surveillance system. In this thesis, an attempt has been made to address these issues by proposing robust algorithms which have been validated with the benchmark datasets.

7.2 WORK DONE

As a first process, an approach for eliminating shadows from a static background using Gabor filter based Gaussian Mixture shadow modeling in optimal color space has been proposed. The experimental results on benchmark datasets and for local environments have shown that the proposed shadow model efficiently models different shadows in various environmental conditions. In order to evaluate the validity of the proposed methodology, a metric called shadow detection rate is obtained and compared with that of existing algorithms, namely conventional GMM and Basic Gaussian shadow modeling, for both indoor and outdoor having different types of shadows. The proposed Gabor filter based Gaussian shadow modeling
attains detection rate of 94.18% on benchmark datasets pertaining to different shadows of both indoor and outdoor scenes.

As the second step in motion analysis, a reliable classification algorithm has been proposed for human and vehicle, which works well in the challenging real-world conditions, including the presence of shadows and groups of people. Based on the experimental results, the overall classification accuracy is 97.7% on PETS 2001 bench mark datasets. The algorithm proves its robustness and correctness in classifying people under occlusion, like when a shadow is merged with a group of people. Subsequently, a novel approach is presented to deal tracking and identifying persons in a homogenous environment using the skin color model in the optimal color space. The Object Tracking Error (OTE) for the proposed method is 4.6% under occlusion.

This thesis has extended the concept of Relevance Vector Machine (RVM) to classify the gender of the people captured in a application of video surveillance environment. The appearance features which are used to recognize the gender of an individual are extracted using elliptical view of seven segmented regions of a gait and are fed as the input into the RVM learning system. By the application of the RVM on the benchmark datasets, it is proved that the gender classification task can clearly handle a large number of subjects successfully and shows accuracy of 90.75% as compared to SVM which produces 80.63%.

A video surveillance system to classify human action as normal or abnormal has also been proposed in this work. The distinct contribution of the proposed work is in classifying the action of an individual in a crowded scene even in the case of partial occlusion. The proposed method is able to detect abnormal actions of an individual such as running, bending down, waving
hand while others walk, and also persons fighting with each other. On an average the performance of the linear Support Vector Machine (SVM) classifier was 84.67% on benchmark datasets CAVIAR, CMU, IBM and Weizmann, where as the performance of the proposed RVM classifier is 93.53%.

### 7.3 Future Work

In video surveillance, many issues are still open and deserve for further research, especially in the areas like security and crowd control. The shadow is modeled for few frames and eliminated thereafter affects the real-time analysis, hence a trade-off between speed and robustness for eliminating the shadows will facilitate the real-time analysis. Introducing multiple camera view of a scene in different angles would improve the object classification performance and further, the classification can be extended to sub-classes like person carrying objects, a person driving bike or cycle and for vehicle classification particularly as car, van, truck to help in high level analysis.

The parametric models and their applications in tracking and identification are important research directions in video surveillance and 3-D modeling deserves more attention in future work. If the distance is very far or the subjects of interest are turned away from the camera, the face features are possibly at too low resolution to recognize. Instead of facial feature, human gait can be a new biometric feature for person specific identification that can be employed for identification of human. Future work on gait recognition will focus on handling the following two problems. First, most experiments are carried out under constrained circumstances like there is no occlusion while objects are moving and the background is simple. Second, the assumption of image scene being filled with less number of people than complex crowded environment.
As a whole, recognition of human and their actions is just in its infancy, and there exists a trade-off between computational cost and accuracy for real time application. Hence, future machine learning algorithms must be able to independently sense the surrounding environment, with the use of multiple cameras with robust identification and data fusion algorithm.