CHAPTER VII

CONCLUSIONS AND TOPICS FOR FUTURE RESEARCH

This chapter summaries the contributions of the dissertation together with general directions for future research.

7.1 Conclusions

In this dissertation, we have introduced new classification algorithms based on the uses of local discriminant basis feature extraction algorithm. Four generalized multiple classifier systems have been explored in detail for the MSTAR SAR ATR and UCI Satimage problems: (1) the system based on the overcomplete framework of *local discriminant frame expansion*, (2) the system based on *generalized code concatenation schemes*, (3) the system based on *ridge prediction optimization method*, and (4) the system based on *local discriminant basis neural network ensemble*. Each of these systems can be classified as either the coverage optimization or the prediction optimization methods, which are generally defined in the MCS framework. It was also shown that these new contributions have significantly improved performance over the existing MCS. Furthermore, the presented systems further help reduce the computational complexity in several cases.

7.2 Contributions of Dissertation

The originality of the dissertation are as follows:

- A new method of coverage construction of multiple classifier systems is developed and applied to the public MSTAR database. This method is conceptually related to signal-domain channel coding. In particular, the method of using the LDB feature extraction algorithm in an MDC framework have been presented. To combat classification error, predefined amount of redundancy are added to the original data during the feature extraction process. Unequal protection is employed by varying the amount of redundancy with the important of data (through the use of local discriminant frame expansion). Classification accuracy is increasingly obtained when more descriptions are constructed and used.

- New methods for coverage construction of multiple classifier systems are developed and applied to the public MSTAR database and the UCI repository data set. The main task of these methods is to construct multiple classifier systems based on two extensions of the ECOC approach: (1) classical code concatenation and (2) generalized code concatenation so that the highest classification performance is attained. An evaluation of the proposed methods on the classification of two public data sets provides additional proofs of the improvement of our multiple classifier systems.

- A new method of prediction optimization of multiple classifier systems is developed and applied to the public MSTAR database. The method works effectively regardless of the amount of variation caused from using different number of most discriminant basis. This estimation method suggests a robust and cost-saving solution over existing weight-combining methods. Furthermore, the experiments also suggested that the classification system based on our multiple description method can generate an effective ensemble consisted of high-accuracy base classifiers that were nearly independent to each other.

- A new method of coverage construction of multiple classifier systems is developed and applied to the public Yale face database. Several frameworks are considered and discussed for their equivalencies with an ensemble of transform networks derived using local discriminant basis algorithm. In addition to these discussions, a proof is provided that the linear combination of individual network weights of an ensemble of transform networks is a more generalized representation for multiple classifier systems than other simple methods, e.g., constant or weighted sample mean of the weights. The experiments suggested that our system performs quite well when only high discriminant data subbands are included and presented to classifiers in the ensemble.

7.3 Topics for Future Research

Based on the development of the systematic framework and the experimental results covers in this dissertation, this section concludes with a list of future research topics which could not be treated here:

- The method in Chapter 3 uses a frame expansion to produce descriptions. It gives descriptions that are deterministically correlated. There are several other multiple description coding methods that focusing on statistical correlation, e.g. correlated transform [44]. Thus, it is possible to apply these methods to multiple description model for multiple classification systems.

- Equal protection is implemented without emphasizing on a region of interest (ROI) or important information in data (or target class information in classification). Intuitively, most of the redundancy should be concentrated in the ROI. As a consequence, the ROI of target pattern should be therefore more heavily highlight for classification than the other parts of the target. Recently, ensemble feature selection strategies [84, 109] are proposed for searching the best collection of feature subsets. Thus, the unequal protection of ROI in MDC for MCS can be performed a priori or by using the ensemble feature selection methods to allocate the optimal number of the MDB for each part of the target. This way, an ensemble is expected to be effectively constructed. Moreover, it is analogous to the optimal rate distortion design (redundancy allocation) in joint source-channel coding.

- The generalized encoding method in Chapter 4 is one example of *extended* coverage construction approaches suitable for multiple classifier systems. Future research should start by generalizing the framework of this approach to other coding schemes, e.g., space-time coding.

- The decoding rule implemented in Chapter 4 is mainly focused on distance. There are other decoding rules that can also be used for MCS. Future work should include the decoding rule based on probability estimation. In this approach, *least square* method is used to find an estimate of an optimal posterior probability vector that minimizes the errors. However, there is a serious practical limitation with this least square decoding rule. The probability estimates can be highly variable when there are a smaller number of columns, relative to the number of classes. This is analogous to the situation in least square regressions where the number of data points is small compared to the number of predictor variables. A common approach to this problem in regression is to use ridge least squares. Thus, the future work is to apply the least square method with adapted ridge parameter presented in Chapter 5 as the decoding rules for ECOC and its extensions.

- There are a number of ridge parameter estimation methods [92], e.g., RIDGM, Golub, Heath, and Wahba, and Bayesian methods. These methods can be explored for the prediction optimization of MCS. Moreover, any improvement in the least squares techniques, e.g., the covariance shaping least square technique in [97] can be of interest to the optimal combining-weights techniques. Especially, if we can further include the ridge parameter estimation method within the covariance shaping least square method.

- Another avenue of future work is to apply our regularization least square methods for other applications such as image superresolution, system identification, data mining, and multiuser detection. The least square method is well-suited to these applications because it provides a better way to estimate the ridge parameter than the conventional cross-validation method.