

Chapter 8

Conclusion

This chapter briefly highlights the findings and contributions of this thesis and discusses directions for future research.

8.1 Summary

This thesis investigated the application of an efficient optimization method known as Particle Swarm Optimizer to the field of pattern recognition and image processing.

Chapter 4 presented a clustering approach using PSO. The objective of the proposed algorithm is to simultaneously minimize the quantization error and intra-cluster distances, and to maximize the inter-cluster distances. The application of the proposed clustering algorithm to the problem of unsupervised classification and segmentation of images was investigated. The proposed algorithm was compared against *state-of-the-art* clustering algorithms. In general, the PSO algorithms produced better results with reference to inter- and intra-cluster distances, while having quantization errors comparable to the other algorithms. The performance of different versions of PSO was investigated and the results suggest that algorithms that start with high diversity and then gradually go to low diversity perform better than other algorithms. To test its performance on multidimensional feature spaces, the proposed approach was applied to multispectral imagery data.

Chapter 5 presented a tool for synthetic image generation (SIGT). The tool consists of two units: a synthetic image generator and a clustering verification unit.

The first unit allows the user to create a synthetic image based on a user-specified histogram suitable for the required application. The second unit allows the user to measure the efficiency of a clustering algorithm. Different features of SIGT were demonstrated by a set of experiments aided by the K-means clustering algorithm and the PSO-based clustering algorithm proposed in chapter 4. These experiments have demonstrated that the tool can help researchers in the field of unsupervised image classification to generate synthetic images, measure the quality of a clustering algorithm, compare different clustering algorithms and to create benchmarks.

Chapter 6 presented a new dynamic clustering algorithm based on PSO, called DCPSO, with application to unsupervised image classification. DCPSO clusters a data set without requiring the user to specify the number of clusters *a priori*. DCPSO uses a validity index to measure the quality of the resultant clustering. DCPSO has been applied to synthetic images (where the number of clusters was known *a priori*) as well as natural images (including MRI and satellite images), and was compared with other dynamic clustering techniques. In general, DCPSO successfully found the "optimum" number of clusters on the tested images. Genetic algorithm and random search versions of the proposed approach were presented and compared to the particle swarm version with both the genetic and PSO versions outperforming the random search version. The influence of the different DCPSO control parameters was then investigated. The use of different PSO versions was also studied. Finally, to test its performance in multidimensional feature space, the DCPSO was applied to multispectral imagery data.

Chapter 7 addressed two difficult problems in the field of pattern recognition and image processing. The two problems are: color image quantization and spectral unmixing. First, the chapter presented a PSO-based color image quantization

algorithm (PSO-CIQ). The PSO-CIQ algorithm was compared against other well-known color image quantization techniques. In general, the PSO-CIQ performed better than other techniques when applied to a set of commonly used images. The effects of different PSO-CIQ control parameters were studied. The performance of different versions of PSO was then investigated. Chapter 7 then presented a new spectral unmixing approach using PSO (PSO-EMS). The objective of the PSO-EMS algorithm is to determine the appropriate set of end-members for a given multispectral image set. The PSO-EMS algorithm performed well when applied to test image sets from various platforms such as LANDSAT 5 MSS and NOAA's AVHRR. The effects of different PSO-EMS control parameters were then studied. Finally, the performance of different versions of PSO was investigated.

From the results presented in this thesis, it can be concluded that the PSO is an efficient optimization algorithm for difficult pattern recognition and image processing problems. These problems are considered difficult because they are NP-hard and combinatorial problems.

8.2 Future Research

Directions for future research are briefly summarized below.

PSO-based Clustering Algorithm

Although the parametric fitness function used by the PSO-approach contains multiple objectives, no special multi-objective optimization techniques have been used. Future research can investigate the use of a PSO multi-objective approach, which may produce better results. In addition, incorporating spatial information into the PSO-

based clustering algorithm (when used in image segmentation applications) needs to be investigated. One way to incorporate spatial information is to consider the eight neighboring pixels of each pixel as proposed by Liew *et al.* [2000].

SIGT

Future additions to the tool may include a unit to generate a synthetic image from an existing real (synthetic) image by relaxing some constraints. In addition, further studies may use SIGT to do an elaborate analysis and comparison of clustering algorithms.

DCPSO

The application of the DCPSO algorithm (described in Section 6.1) to general data needs to be investigated. Furthermore, the effect of high dimensionality on the performance of the DCPSO should be investigated. Experiments for validating the efficiency of randomly re-initializing M_r (i.e. step 8 in Figure 6.1) need to be conducted. The DCPSO uses the K-means clustering algorithm to refine the cluster centroids. Future research can investigate the use of other more efficient clustering algorithms such as FCM and KHM. In addition, incorporating spatial information into the DCPSO algorithm (when used in image segmentation applications) needs to be investigated.

PSO-CIQ

The PSO-CIQ (described in section 7.1.1) uses the K-means clustering algorithm to refine the color triplets. Future research should investigate the use of other more efficient clustering algorithms such as FCM and KHM. Experiments need to be

conducted to compare the PSO-CIQ with the multi-start K-means (with the best result generated from applying K-means $st_{\max}p_{kmeans}$ times, where each K-means starts from random cluster centroids). Finally, the PSO-CIQ uses the RGB color space. Although the RGB model is the most widely used model, it has some weaknesses. One of these weaknesses is that equal distances in the RGB color space may not correspond to equal distance in color perception. Hence, future research may try to apply the PSO-CIQ to other color spaces (e.g. the L*u*v* color space [Watt 1989]).

PSO-EMS

Experiments need to be conducted to compare the PSO-EMS with the multi-start K-means (with the best result generated from applying K-means $st_{\max}p_{kmeans}$ times, where each K-means starts from random cluster centroids). The performance of the PSO-EMS (described in section 7.2.1) when applied to hyperspectral Satellite imagery is a potential topic for future research.