Chapter 1

Introduction

As humans, it is easy (even for a child) to recognize letters, objects, numbers, voices of friends, etc. However, making a computer solve these types of problems is a very difficult task. Pattern recognition is the science with the objective to classify objects into different categories and classes. It is a fundamental component of artificial intelligence and computer vision. Pattern recognition methods are used in various areas such as science, engineering, business, medicine, etc. Interest in pattern recognition is fast growing in order to deal with the prohibitive amount of information we encounter in our daily life. Automation is desperately needed to handle this information explosion. This thesis investigates the application of an efficient optimization method, known as Particle Swarm Optimization, to the field of pattern recognition and image processing. PSOs solve optimization problems by simulating the social behavior of bird flocks.

1.1 Motivation

There are many difficult problems in the field of pattern recognition and image processing. These problems are the focus of much active research in order to find efficient approaches to address them. However, the outcome of the research is still unsatisfactory.

Local search approaches were generally used to solve difficult problems in the field of pattern recognition and image processing. However, the selected set of

problems in this thesis are NP-hard and combinatorial. Hence, evolutionary algorithms are generally more suitable to solve these difficult problems because they are population-based stochastic approaches. Thus, evolutionary algorithms can avoid being trapped in a local optimum and can often find a global optimal solution. A PSO is a population-based stochastic optimization algorithm modeled after the simulation of the social behavior of bird flocks. PSO is easy to implement and has been successfully applied to solve a wide range of optimization problems [Hu 2004]. Thus, due to its simplicity and efficiency in navigating large search spaces for optimal solutions, PSOs are used in this research to develop efficient, robust and flexible algorithms to solve a selective set of difficult problems in the field of pattern recognition and image processing. Out of these problems, data clustering is elaborately tackled in this thesis specifically image data. The motivation for the focus on data clustering is the fact that data clustering is an important process in pattern recognition and machine learning. Actually, clustering is a primary goal of pattern recognition. Furthermore, it is a central process in Artificial Intelligence. In addition, clustering algorithms are used in many applications, such as image segmentation, vector and color image quantization, spectral unmixing, data mining, compression, etc. Therefore, finding an efficient clustering algorithm is very important for researchers in many different disciplines.

1.2 Objectives

The primary objectives of this thesis can be summarized as follows:

- To show that the PSO can be successfully used to solve difficult problems in pattern recognition and image processing.

- To develop an efficient clustering algorithm based on PSO.
- To develop a tool that can aid researchers in the unsupervised image classification field to test their algorithms, compare different clustering algorithms and generate benchmarks.
- To develop an efficient dynamic clustering algorithm that can find the "optimum" number of clusters in a data set with minimum user interference.
- To develop a PSO-based approach to tackle the color image quantization problem.
- To develop an efficient end-members selection method based on PSO for spectral unmixing of multispectral imagery data.

1.3 Methodology

Algorithms proposed in this thesis are first presented and discussed. Experimental results were then generally obtained using various synthetic images with well-known characteristics in order to show the accuracy and efficiency of the proposed algorithms.

In addition, natural images from different areas such as medical images and remotely sensed satellite images were also used to show the wide applicability of the proposed approaches.

The results of *state-of-the-art* algorithms when applied to the same test images were also reported to show the relative performance of the proposed approaches when compared to other well-known approaches.

For the task of unsupervised image classification, attempts were made to find the best values for the PSO parameters.

Due to the stochastic nature of the proposed algorithms, all the presented results are averages and standard deviations over several simulations. However, due to the computational expensive nature of the simulations, results were generally taken over 10 or 20 runs.

1.4 Contributions

The main contributions of this thesis are:

- The development of an efficient clustering algorithm based on the PSO that performs better than *state-of-the-art* clustering algorithms when applied to the problem of unsupervised image classification.
- The development of a simple tool for synthetic image generation and verification. This tool can be used as a preliminary test to compare different unsupervised image classification algorithms. In addition, it can be used to generate a set of benchmark images that can be used by the researchers in the field of unsupervised image classification.
- The development of an efficient dynamic clustering algorithm based on the PSO that is able to simultaneously cluster a data set and find the "optimum" number of clusters in the data set.
- The development of an efficient color image quantization algorithm based on the PSO which is capable of generating high quality quantized images.
- The development of an efficient end-members selection method for spectral unmixing of multispectral satellite imagery data which is based on the PSO. The efficiency of the algorithm is demonstrated by applying it to test imagery from various platforms.

1.5 Thesis Outline

Chapter 2 briefly reviews the subject of optimization. This is followed by a brief discussion of traditional and stochastic optimization methods. Evolutionary Algorithms (EAs) (with more emphasis on Genetic Algorithms (GAs)) are then presented. This is followed by an elaborated discussion of particle swarm optimization and its various modifications. PSO is a model from the swarm intelligence paradigm. Therefore in order to provide a complete coverage of swarm intelligence background, a brief overview of another swarm intelligence model, Ant Colony Systems, is given.

Chapter 3 reviews the problems addressed in this thesis in sufficient detail. First the clustering problem is defined and different clustering concepts and approaches are presented. This is followed by defining image segmentation in addition to presenting various image segmentation methods. A survey of color image quantization and its approaches is then presented. This is followed by a brief introduction to spectral unmixing.

Chapter 4 presents a clustering method that is based on PSO. The algorithm finds the centroids of a user specified number of clusters, where each cluster groups together similar patterns. The application of the proposed clustering algorithm to the problem of unsupervised classification and segmentation of images is investigated. To illustrate its wide applicability, the proposed algorithm is then applied to synthetic, MRI and satellite images.

Chapter 5 presents a new automatic image generation tool tailored specifically for the verification and comparison of different unsupervised image classification

algorithms. The usefulness of the tool is demonstrated in this chapter with reference to the well-known K-means clustering algorithm and the PSO-based clustering algorithm proposed in the chapter 4.

Chapter 6 presents a new dynamic clustering approach based on PSO. This approach is applied to unsupervised image classification. The proposed approach automatically determines the "optimum" number of clusters and simultaneously clusters the data set with minimal user interference. The proposed approach is then applied to synthetic, natural and multispectral images. A genetic algorithm and a random search version of dynamic clustering are presented and compared to the particle swarm version.

Chapter 7 presents PSO-based approaches to tackle the color image quantization and spectral unmixing problems. The proposed approaches are then applied on different image sets to show their applicability and they are compared with other *state-of-the-art* approaches.

Chapter 8 highlights the conclusions of this thesis and discusses directions for future research.

The appendices present a definition of frequently used terms and symbols and a list of publications derived from the work introduced in this thesis.