Chapter 14

A Comparative Study of Four Different Satellite Image Classification Techniques for Geospatial Management

Devanjan Bhattacharya
Indian Institute of Science Bangalore, India & University of Pardubice, Czech Republic

Santo Banerjee
Institute for Mathematical Research, University Putra Malaysia, Malaysia & International Science Association, Turkey

ABSTRACT

Satellite imagery interpretation has become the technology of choice for a host of developmental, scientific, and administrative management work. The huge repository of geospatial data and information that are available as satellite imageries datasets from platforms such as Google Earth need to be classified and understood for natural resources management, urban planning, and sustainable development. The classification and analysis procedures involve algorithms like maximum likelihood classifier, isodata, fuzzy-logic classifier, and artificial neural network based classifier. Amongst these classifiers the optimum has to be selected for classifications which involve multiple features and classes. Herein lies the motivation for the present research, which can facilitate the selection of one amongst the many algorithms available to a decision maker/manager. The aforementioned techniques are applied for classification, and the respective accuracies in the classes of forestry, rock, water, built-up area, and dry river bed have been tabulated and verified from ground truth. The comparison is based on time and space complexity of the algorithms considering also the accuracy. It is found that traditional methods like MLC and Isodata offer good time and space consumption performance over the recent more adaptable algorithms as fuzzy and ANN. But the latter group excels in accuracy of assessment. The study suggests points and cases for ranking the techniques as best, 2nd best, and so on, where each technique could be optimally utilised for a given geospatial dataset based on its contents.
A Comparative Study of Four Different Satellite Image Classification Techniques

INTRODUCTION

Classification is the process of partitioning a digital image into multiple classes (sets of pixels). The goal of classification is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze (Ballard & Brown, 1982). More precisely, image classification is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics (Hord, 1982).

The result of image classification is a set of segments that collectively cover the entire image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as colour, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics. Several general-purpose algorithms and techniques have been developed for image classification. Since there is no general solution to the image classification problem, these techniques often have to be combined with domain knowledge in order to effectively solve an image classification problem for a problem domain.

The objective of the present study is to apply and compare the results of the four different image classification techniques namely maximum likelihood classifier (MLC), Isodata, fuzzy-logic classifier (FLC), and artificial neural network (ANN) based classifier, all on the same satellite image of a region acquired from Google Earth for which actual ground truth data has already been collected for later verification of results. The comparative study details out the procedures undertaken, the parameters considered for evaluations, and the standing of each technique, put in a tabular manner. The accuracy assessment has been conducted as per the variations of classified pixels with respect to actual ground truth. The evaluation parameters of time complexity, space complexity, accuracy and upgradability have been considered for performance testing of the techniques wherein the time taken by each algorithm is noted as well as the memory space occupied inside the computing system. Based on the collective results, suggestions are made on the selection of the optimum classification techniques for better land use and land cover management.

BACKGROUND

The algorithms considered for the study are all pixel based techniques that work on raster data sets in the domain of digital image processing and classification. Raster data is a two dimensional array of digital numbers (DN) where each cell of the array corresponds to a pixel of the image. The image could be a gray-scale (b/w) or a coloured image. Accordingly the array values would vary. A gray-scale would have a single layer of digital values and a colour image would consist of hues of RGB. Each DN value represents a unit intensity captured in the image (Lillesand & Kiefer, 1994).

The algorithms try to best utilize the categorization of these DN values. Maximum likelihood, isodata, fuzzy techniques and ANN have been used for numerous satellite image analysis studies but have not been compared amongst themselves on a single platform anywhere. The measures of comparison that effect the performance and utility most are space occupied in memory, time taken for processing, accuracy delivered, and future upgradability. To understand the study better, it is imperative to get a look into the backgrounds related to the study. In this section some common terminologies are discussed to lay the foundation for their applications later.

The simplest method of image classification is called the thresholding method. The key of this method is to select the threshold value (or values when multiple-levels are selected) for each category that exists in a digital image. Several popular methods are used in this field and in industry including the maximum entropy method,