GPU-based Parallelization on Unsupervised Clustering Analysis

Victor J. D. Tsai

Department of Civil Engineering, National Chung Hsing University,
250 Kuokuang Rd., Taichung 40227, Taiwan
E-mail: jdtsai@nchu.edu.tw

Abstract: The functionalities of practical image processing systems are always computational intensive and are limited by the computing performance of the Central Processing Unit (CPU), memory allocation and bandwidth, capacity of storage device, limitation and deficits of the software, and deficiency in processing large volume of remote sensing images. Since there are no theoretical limitation on the number of bands (or features), in principle, any of the conventional algorithms for classifying multispectral data can be directly applied to hyperspectral data. However, these algorithms appear suddenly inefficient when applied to a hyperspectral image with 200 or more closely related spectral bands (Schowenderdt, 1997). The increased number of spectral bands results in a vast increase in the computational load for statistical analysis in order to derive reliable class-specific statistics for maximum-likelihood approach. The access to efficient hardware and software is an important factor in determining the *ease* with which an unsupervised classification can be performed (Lillesand & Kiefer, 2000).

Considering the varieties of image processing functionalities and the parallel computing capability of current multi-core CPU and high-level Graphics Processing Unit (GPU) display card, this research aims on developing both CPU-based and GPU-based parallel processing approaches for fast unsupervised classification of remote sensing images using C++, CUDA (Compute Unified Device Architecture), and OpenMP API. The basic idea of well-known *Iterative Self-Organizing Data Analysis Techniques A* (*ISODATA*) approach was realized and refined in this research for clustering airborne hyperspectral images. The performance of numerical experiments in clustering 240-band airborne Intelligent Spectral Imaging System (ISIS) images into specified number of spectral clusters will be demonstrated for the advantages in computational efficiency of the GPU-based approach in a highly parallel, multi-thread, and multi-core implementation and the refined multi-core CPU-based parallel approach in OpenMP API against traditional CPU-based serial pipeline approach.

Keywords: Clustering, Unsupervised Classification, ISODATA, CUDA, OpenMP