Parallel Strategy of Peak Identification Algorithm Based on Region Growing Method

Ick-Hoi Kim, Chen-Chieh Feng

Department of Geography, National University of Singapore, 1 Arts Link, Kent Ridge, Singapore
Telephone: (65) 6601-2999
Fax: (65) 6777-3091
Email: geoihk@nus.edu.sg, chenchieh.feng@nus.edu.sg

1. Introduction

This paper introduces a parallel strategy to process a large digital elevation model (DEM) dataset to identify mountain peaks. It aims to develop a means to help explore the differences of human conceptualizations of convex landscape features in Malaysia, Indonesia, Singapore, and Brunei, the neighboring countries in Southeast Asia where variants of Bahasa are used. ASTER GDEM (http://www.jspacesystems.or.jp/ersdac/GDEM/E/index.html) was downloaded and used for the state purpose. The sizes of DEM for the countries in the study area were from 16GB to 24GB. As pointed out by Mark and Turk (2003), the landscapes and its elements can be conceptualized in different ways due to cultural and linguistic variations. Further complicating the problem is the terminological incompatibilities (Smith and Mark 2001). To understand the kinds of cultural and linguistic variations in identifying mountain peaks in neighboring countries, we need an efficient geo-computation tool that handles large size DEM and considers parameters adaptive to cultural and linguistic differences in human conceptualization of mountain peaks.

According to Pike et al. (2009), Geomorphometry is "the science of quantitative land-surface analysis" and extract land-surface features from DEMs. In general, land-surface features are peak, pass, pit, ridge, channel, and plane (Wood 1996). For this paper, we focused on the spatial cognition of mountain peaks. Researchers proposed diverse peak identification algorithms. Most of the algorithms rely on local analysis (Peucker and Douglas 1975, Wood 1996). Wood (2004) illustrated a region growing algorithm to identify peaks and summits by using a relative drop. Recently, Sinha and Mark (2010) demonstrated an algorithm combining field representation and object-based conception with the region growing algorithm to identify topographic eminence.

2. Preliminary Work

The peak identification algorithm that we used was based on the region growing method by Wood (2004). This algorithm extracts peaks and summits from a DEM using the neighborhood operation (Pike et al. 2009) and is therefore a region growing algorithm. The neighborhood of each region is extended based on a relative drop. In Wood's algorithm, the summit is the cell of local maxima, and the peaks are the cells surrounding the summit with the elevation difference lower than the relative drop. The relative drop can be assigned as an input parameter. The search for a peak starts from a cell with a fixed size of a neighborhood window (e.g., 3 x 3). All neighboring cell values are added to a priority queue. The queue is sorted by the elevation value. The search repeats from the cell with the highest value if the elevation difference is less than the relative drop. The neighboring cells are also added to the priority queue and sorted. Thus, the neighborhood is growing. The search stops when the cell is at the edge or the cell value is higher than the starting cell value. The algorithm provides different number of peaks based on the relative drop parameter. Thus, we have plans to test the algorithm for the DEM in the study area with different relative drop values, such as 50m, 100m, 150m, and 200m.

One issue is that the processing of the DEM in the study area requires large compute time due to its large file size. It is difficult to reduce the compute time by using faster processors because CPU clock speeds have stayed recently (Schiele et al. 2012). To address data intensity issues, the parallelization of GIS models would be a possible solution. Without the parallelization, it takes much time for each run of the peak identification algorithm. Thus, we have parallelized this peak identification algorithm to handle large datasets.

We are working on the implementation of a parallel algorithm for distributed memory parallel processing. This is a challenging problem because of data dependencies across neighboring geographic regions. In this algorithm, we do not know the extent the peak areas can grow. In particular, Wood's algorithm is designed to stop the search of the peak when one of the neighborhood cells is at the edge. So, the peaks or summits cannot be identified when the cells are located at the edge of partitioned data.

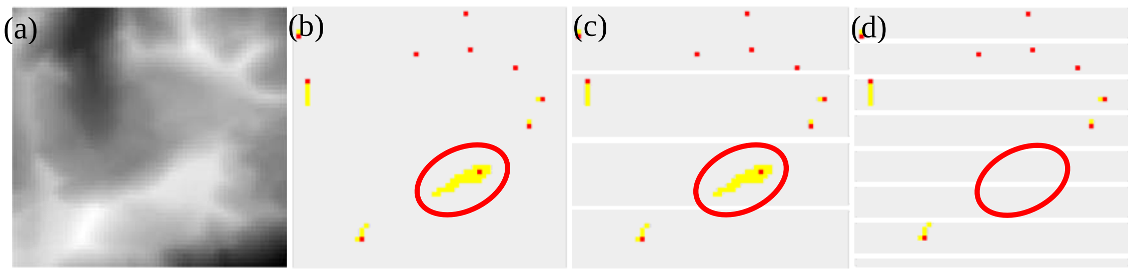

Figure. 1. The comparison of sequential and parallel processing of peak identification using Wood's algorithm. Sub-figure (a) shows the sample DEM data. Sub-figure (b) – (d) show peak identification results by sequential processing, parallel processing with 4 partitions, and parallel processing with 8 partitions. Red points indicate the summit, and the yellow points are the peak area of the summit.

Fig. 1 shows the preliminary result of the parallelization effort. The sub-figure (a), (b), (c) – (d) show the input DEM, peak identification result using the original Wood's algorithm (sequential processing), and parallel processing of the data (4 and 8 partitions), respectively. Note the area in the red circles in Figure 1(b) – (d). A peak area is recognized in sequential processing. When the data was decomposed by 4 partitions, the peak area still exists. However, when the data was decomposed by 8 partitions, the peak area could not be recognized because the peak search stops at the edge. Even the summit disappeared because it can be recognized only after the completion of the peak identification. The sub-figures (c) and (d) demonstrated that while parallelizing an existing sequential processing algorithm is possible, the result may be unsatisfactory, as demonstrated in sub-figure (d). Improvement to the existing algorithms is needed.

Current work is to address this issue by exchanging the data between the parallel processors. We made a parallel strategy with 2 steps (fig. 2). If the search for a peak area stops at the edge, the peak area is checked, and the data are sent to the adjacent parallel processor. Step 2 is to continue the search for the peak area in the adjacent parallel processor.

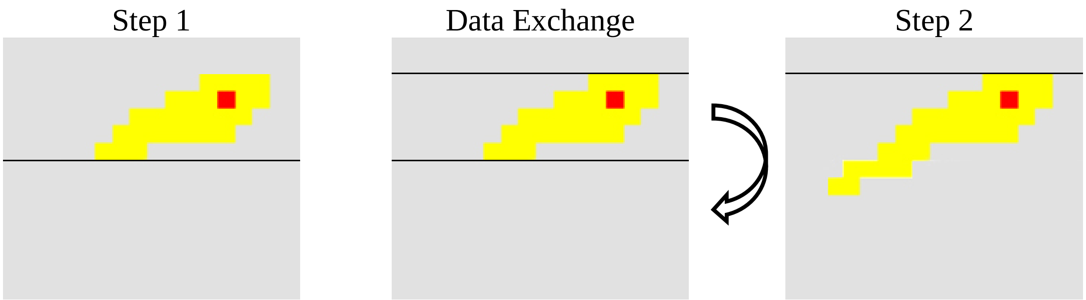

Figure. 2. Modified parallel peak identification algorithm with 2 steps.

3. Conclusion and Future Work

When the development of this parallel algorithm is completed, we will derive mountain peaks in the study area. Then, the locations of the mountain peaks will be compared to locations of place names categorized as mountain in the study area. The future work will provide a better understanding regarding how compatible is the mountain peaks identified by algorithm and by human. The knowledge gained not only helps understand terminological and conceptual incompatibilities for mountain peak names in Southeast Asian countries in which variants of Bahasa are spoken, but also how to fine-tune the peak identification algorithms so that it is more adapted to local cultures and languages.

4. References

Mark, D.M. and Turk, A.G., 2003, Landscape categories in Yindjibarndi: Ontology, environment, and language. In: Kuhn, W., Worboys, M. and Timpf, S. (eds), *Spatial Information Theory: Foundations of Geographic Information Science*, Springer-Verlag, 31-49.

Peucker, T.K. and Douglas, D.H., 1975, Detection of surface-specific points by local parallel processing of discrete terrain elevation data. *Computer Graphics and Image Processing*, 4(4): 375-387.

Pike, R.J., Evans, I.S. and Hengl, T., 2009, Geomorphometry: A brief guide. In: Tomislav, H. and Hannes, I.R. (eds), *Developments in Soil Science*, Elsevier, 3-30.

Schiele, S., Möller, M., Blaar, H., Thürkow, D. & Müller-Hannemann, M., 2012, Parallelization strategies to deal with non-localities in the calculation of regional land-surface parameters. *Computers & Geosciences*, 44:1-9.

Sinha, G. and Mark, D.M., 2010, Cognition-based extraction and modelling of topographic eminences. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 45(2):105-112.

Smith, B. and Mark, D.M., 2001, Geographical categories: an ontological investigation. *International Journal of Geographical Information Science*, 15(7):591-612.

Wood, J., 1996. *The Geomorphological Characterisation of Digital Elevation Models*. PhD dissertation, University of Leicester, UK.

Wood, J., 2004, A new method for the identification of peaks and summits in surface models. In: *Proceedings of the 3rd International conference on GIScience 2004*, Baltimore, MD, USA, 227-230.