The Little Algorithm that Grew:

Scaling the Morphological Image Compositing Algorithm to meet the Challenges of Processing Large Image Data Sets

C. Bielski[1], J. Grazzini[1] and P. Soille[1]

[1]Institute for Environment and Sustainability
DG Joint Research Centre, European Commission
I-21020 Ispra (Va), Italy
Telephone: (+39) 0332 789 233
Fax: (+39) 0332 786 325
Email: {Conrad.Bielski; Pierre.Soille; Jacopo.Grazzini} @jrc.it

1. Introduction

Image mosaicing (also called image compositing) is a technique used to join two or more images together. Recently, Soille (2006) developed a morphological image compositing algorithm which is being used to automatically generate a pan-European image mosaic based on fine spatial resolution satellite imagery. While this is not the only image mosaicing algorithm available today (Szeliski 2004, Soille 2006, Price 2006) it does generate a mosaic with the following required characteristics: visually pleasing and research grade. Visually pleasing in this case means that the seam lines are difficult to detect visually and research grade means that the radiometry of the original image data is not changed in the resulting mosaic. This talk will describe the evolution of the morphological image compositing algorithm in order to deal with computer memory limitations to make it work on large image data sets and the adaptations to make it available within a grid computing environment. The challenges and solutions of this approach will be described as well as the benefits of this algorithm for use in processing large image data sets.

1.1 Algorithm Description

The original algorithm, presented in Soille (2006), basically works like this:

1. Input is a list of geometrically corrected and re-sampled to the same grid overlapping images. The only constraint is that none of the images are fully included inside another image;

2. Compute the Marker Image which provides for each pixel within the mosaic the number of overlapping images;
3. Compute the Mask Image which is the point-wise minimum between the morphological gradients of all images;
4. Delineate the seam line based on morphological watershed region growing (Vincent and Soille 1991) which is seeded by the Marker Image and constrained by the Mask Image. The initial seeds are the Marker Image pixels where there is no overlap and continues step wise to regions with greater number of overlapping images.

2. Algorithm Evolution

The first implementation of this algorithm was tested on small sets of Landsat 7 ETM+ scenes. It was quickly made apparent however, that applying the algorithm to image sets that generated a mosaic larger than the 4 GB limit of our computer memories was not possible because the entire image mosaic had to be stored in memory for processing. The solution to this problem was dubbed the order independent image compositing algorithm (Bielski and Soille 2005). This algorithm was able to overcome the computer memory limitation because the automatic seam line delineation could be successfully applied on a scene by scene basis which requires a much smaller memory footprint.
The next step in the evolution of the morphological image compositing algorithm began with access to a grid computing environment and the Image 2006 project. The Image 2006 project was proposed in the context of the Global Monitoring for Environment and Security (GMES) Fast Track Land Service (FTLS) where satellite imagery covering 37 European countries will be used to create a mosaic. The fine spatial resolution imagery will be acquired by the SPOT and IRS series of satellites and will provide two temporal coverages at a pan-European extent with over 3000 individual scenes. Processing such a large number of scenes in a short period of time required greater computing power which was made available through blade servers. The blades are accessed through the Sun Grid Engine (SGE 2007), an open source batch-queuing system. In order to take full advantage of this computer power, the Java port of the Distributed Resource Management Application API (DRMAA 2006) was used to adapt the morphological image compositing algorithm.

2.1 Grid Implementation Details

The current implementation of morphological image compositing algorithm is able to take advantage of the processing power of the grid computing environment due to some grid specific pre-processing in order to manage those processes that could be run in parallel with those that could not. Two things cannot happen during image compositing:

regions with greater overlap images must not be processed until all regions with lesser overlapping images have been processed and scenes that share an overlapping image cannot be processed at the same time because the updates may become corrupt (fig. 1). Both issues are resolved by pre-processing the image list and creating batch processes based on overlap region and the image overlap list for any given scene to be processed. Consequently, large batches of jobs can be sent to the grid as synchronised processes without the fear of corrupting any of the updated images. This works best with large numbers of images covering large extents because there is less chance of sharing overlapping images.

When this project began, 838 Landsat 7 ETM+ scenes were processed on three machines separately for almost a week whereas today the same mosaic can be processed on four nodes (or eight processors) of the blade in just over a day. Processing time could also be reduced in critical cases by submitting the jobs to more nodes with the permission of the SGE administrator.

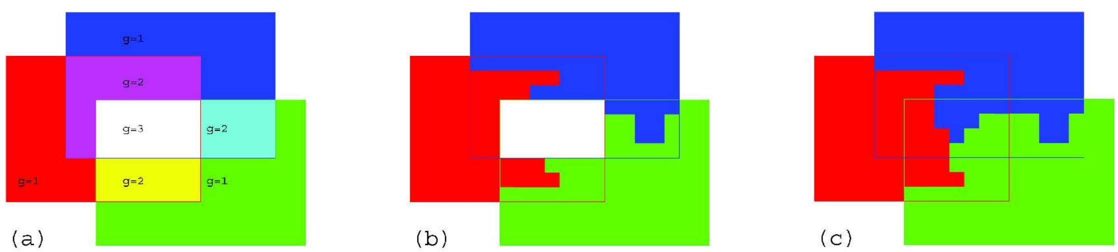

Figure 1. Three overlapping images are shown in (a) (red, blue and green) with the number of overlaps equal to g. In order to avoid corrupted files due to simultaneous updating of the same file, only the red image can be processed in a single synchronised batch process. The blue and green images must be sent as separate jobs within different synchronised batch jobs. Prior to delineating seam lines in the region g=3, all seam lines must be delineated within regions labeled g=2 (b). The pre-processing step orders the jobs and places them into the correct synchronised batches. In (c), the final region is processed.

3. Discussion

These are not the only benefits of the morphological image compositing algorithm. The algorithm also permits easy updating of specific regions of the mosaic, consequently image updates do not require that the entire image data set be re-composed each time new imagery is inserted or removed from the image data set (Bielski and Soille 2007). The algorithm has also been adapted to take advantage of other information during the seam delineation process. For example, if a cloud mask is available the algorithm, if possible, will cut out the clouds from the final mosaic. This same technique could also be used to favour different information layers when generating mosaics in order to make sure that only regions of a certain date, or viewing angle, etc. be used in the final mosaic. Currently, the authors are adapting the algorithm for the mosaicing of Digital Elevation Models (DEM).

4. References

Bielski C and Soille P, 2005, Order independent image compositing. *Lecture Notes in Computer Science*, 3617:1076-1083.

Bielski C and Soille P, 2007, Adaptive mosaicing: Principle and Application to the mosaicing of large image data sets. *Lecture Notes in Computer Science*, 4431(II): 500-507.

DRMAA, 2007, Distributed Resource Management Application API. http://drmaa.org/wiki/

Price, K, 2006, Computer vision bibliography. http://iris.usc.edu/Vision-Notes/bibliography/contents.html

Soille, P, 2006, Morphological image compositing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28:673-683.

SGE, 2007, Sun Grid Engine. http://gridengine.sunsource.net/

Szeliski, R, 2004, Image alignment and stitching: A tutorial. *Microsoft Research*, Technical Report MSR-TR-2004-92.

Vincent L and Soille P, 1991, Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13: 583-598.