GENERATION OF DIGITAL SURFACE MODEL FROM HIGH RESOLUTION SATELLITE IMAGERY

Chunsun Zhang, Associate Professor
Geographic Information Science Center of Excellence (GISCeCE)
South Dakota State University
1021 Medary Ave, Brookings SD 57007, USA
chunsun.zhang@sdstate.edu

Clive Fraser, Professor
Department of Geomatics, The University of Melbourne
Parkville, VIC 3010, Australia
c.fraser@unimelb.edu.au

ABSTRACT

Elevation data is an important component of geospatial database. This paper focuses on digital surface model (DSM) generation from high-resolution satellite imagery (HRSI). The HRSI systems, such as IKONOS and QuickBird have initiated a new era of Earth observation and digital mapping. The half-meter or better resolution imagery from Worldview-1 and the planned GeoEye-1 allows for accurate and reliable extraction and characterization of even more details of the earth surface. In this paper, the DSM is generated using an advanced image matching approach which integrates point and edge matching algorithms. This approach produces reliable, precise, and very dense 3D points for high quality digital surface models which also preserve discontinuities. Following the DSM generation, the accuracy of the DSM has been assessed and reported. To serve both as a reference surface and a basis for comparison, a lidar DSM has been employed in a testfield with differing terrain types and slope.

INTRODUCTION

With the launch of the IKONOS and QuickBird, which produce high-resolution satellite imagery below 1m resolution in the panchromatic mode and 4m resolution in multi-spectral mode, a new era in earth observation, digital mapping and application development has begun. The half-meter or better resolution satellite imagery from Worldview-1 and the planned GeoEye-1 allows for accurate and reliable extraction and characterization of even more details of the earth surface. The possibility of the high-resolution satellite sensors, such as IKONOS and QuickBird to change their viewing angle in one orbit, gives them the capability to obtain stereo or even triple-overlapped images from the same orbital pass. Therefore, imagery collected from high-resolution satellite sensors can alleviate temporal variability concerns as the momentary separation between in-track scene capture allows consistent imaging conditions. These superior characteristics make high-resolution satellite imagery well suited for DSM generation (Toutin, 2004; Zhang, 2005; Poon et al., 2005; Krauss et al., 2005; Sohn et al., 2005; Zhang and Gruen, 2006; Poon et al., 2007) and feature extraction (Hu and Tao, 2003; Di et al., 2003; Zhang et al., 2005). This paper deals with IKONOS Geo stereo imagery for accurate digital surface models generation. After this introduction, we briefly describe the sensor model for image georeferencing. Then, we concentrate on our image matching approach. The approach was developed for automatic DSM generation and provides dense, precise and reliable results. Our approach uses a coarse-to-fine hierarchical solution with a combination of several image matching algorithms. Afterwards, the experiment is conducted and the results is presented using IKONOS Geo stereo imagery in a test site of Hobart, Australia, with large height range and very variable terrain relief and land cover. Finally, the detailed DSM accuracy evaluation is given using a lidar DSM as reference.
SENSOR MODELING AND IMAGE ORIENTATION

The IKONOS satellite operates with a linear array scanner where images are obtained with the pushbroom sensor. Thus, the imagery is composed of consecutive scan lines where each line is independently acquired and has its own time dependent attitude angles and perspective centre position. The imaging geometry of the IKONOS is characterized by nearly parallel projection in along-track direction and perspective projection in cross-track direction. To describe mathematically the object-to-image space transformation, the rational function model has been universally accepted and extensively used (Baltsavias et al., 2001; Jacobsen, 2003; Grodecki and Dial, 2003; Fraser et al., 2002; Fraser and Hanley, 2003; Poli, 2004; Eisenbeiss et al., 2004). The rational function model is the ratio of two polynomials and is derived from the physical sensor model and on-board sensor orientation (Grodecki and Dial, 2003). In IKONOS, the rational polynomials coefficients (RPCs) are supplied with the imagery.

Because RPCs are derived from orientation data originating from the satellite ephemeris and star tracker observations, without reference to ground control points (GCPs), they can give rise to geopositioning biases. These biases can be accounted for by introducing additional parameters (Fraser and Hanley, 2003; Fraser et al., 2006). After the bias compensation process, bias-corrected RPCs can be generated by incorporating bias compensation parameters into the original RPCs, allowing bias-free application of RPC positioning without the need to refer to additional correction terms (Fraser and Hanley, 2003; Grodecki and Dial, 2003; Fraser et al., 2006). It has been shown in previous research that with bias-corrected RPCs, 1 pixel level geopositioning accuracy can be achieved from high-resolution satellite imagery (Dial and Grodecki, 2002a; Fraser and Hanley, 2003; Fraser and Hanley, 2005; Baltsavias et al., 2005; Fraser et al., 2006). As demonstrated in Dial and Grodecki (2002b), sub-pixel accuracy is usually obtained from IKONOS imagery shorter than 50km.

IMAGE MATCHING

To generate a DSM using digital photogrammetry, automatic image matching is essential to collect the data points. Image matching has been an active topic in photogrammetry and computer vision. The goal of image matching is to automatically find the correspondences on overlapping images, thus it is a critical technique in many applications, including digital surface model generation. A wide variety of approaches have been developed, and for DSM generation, some packages are commercially available. However, a fully automatic, precise and reliable image matching method, to adapt to different images and scene contents, does not yet exist. The limitations arise mainly from an insufficient understanding and modeling of the underlying process and lack of appropriate theoretical measures for self-tuning and quality control. The difficulty of image matching comes from, for example, radiometric distortion, geometric distortion, occlusion, repeated pattern and lack of features. The recent research trend in image matching is towards hierarchical solutions with a combination of several algorithms and automatic controls.

We have developed an image matching approach for automatic DSM generation from high-resolution satellite images, which has the ability to provide dense, precise, and reliable results. The approach uses a coarse-to-fine hierarchical strategy with several image matching algorithms, essentially combines the matching results of the feature points, grid points and edges. The general scheme is outlined in Fig. 1.

We employ coarse-to-fine hierarchical strategy where image matching necessarily follows an image pyramid approach. Therefore, the solution of correspondences is found from the top of the image pyramid progressively to the bottom of the pyramid which is the image of the original resolution. During the process, the result from a higher level of the pyramid is used as an approximation and adaptive computation of the matching parameters and search range at the subsequent lower level. Matching continues until the lowest level of the pyramid is reached, where the highest accuracy results are also obtained. In addition, the DSM is firstly generated from the feature points, progressively incorporating more features such as grid points and image edges. Again, as in the image pyramid approach, the resulting DSM from previous feature matching is served as a guide in matching successive features, while the DSM itself will be augmented with the new features, resulting in a denser and denser DSM which allows for better characterization of the terrain. The algorithms for the extraction and matching of feature points, grid points and image edges are described below. More details can be found in Zhang et al. (2007).
Figure 1. Image matching strategy and work flow for DSM generation from high-resolution satellite imagery.

Feature points are the points positioned on features that are randomly distributed over the image. They have been used for generating DSM (Krzystek and Ackermann, 1995; Hsia and Newton, 1999). We applied a new version of the Wallis filter (Baltasvias, 1991) to optimize the images for feature point extraction and subsequent image matching. This filter enhances features in images and therefore enables improved feature point extraction. Furthermore, since the filter is applied in both images using the same parameters, naturally occurring brightness and contrast differences are corrected. Following the image enhancement process, feature points are extracted using the well-known Förstner operator.

Our feature point matching method exploits image pixel grey value similarity and geometrical structure information. The image matching is performed in two steps, where different matching algorithms are employed at each step. After feature point extraction, candidate conjugate points are then located by cross correlation in which the normalized correlation coefficient is used for the similarity measure. This information is then used as prior information in the next step, structural matching. The locally consistent matching is achieved through structural matching with probability relaxation (Zhang and Fraser, 2007).

We use normalized cross-correlation coefficient to quantify the similarity of the candidate matching points. This measure has been shown to be largely independent of differences in brightness and contrast due to normalization with respect to the mean and standard deviation (Zhang and Fraser, 2007; Zhang et al., 2007).

With the computed similarity measures, a matching pool for candidate conjugate points is constructed and a similarity score is attached to each candidate point pair. Although the correlation coefficient is a good indicator of the similarity between points, problems still exist in determining all correct matches. Firstly, there is the difficulty of how to decide on a threshold in correlation coefficients to select the correct matches. The existence of image noise, shadows, occlusions, and repeated patterns exacerbates this problem. Furthermore, matching using a very local comparison of grey value difference does not necessarily always deliver consistent results in a local neighbourhood. In order to overcome these problems, the structural matching algorithm with probability relaxation proposed in Zhang and Baltasvias (2000) has been adopted. The detailed computation of structural matching with probability relaxation is given in Zhang et al. (2007).

Feature points often exist in texture-rich regions, and correspond to points at places with grey value variation. These points are usually suitable for accurate and reliable matching. In case of image regions with poor texture or no texture information, few or even no feature points can be extracted. The image matching with only feature points will leave holes on the DSM in these areas. To solve this problem, grid points can be used and grid point matching...
will be conducted (Hsia and Newton, 1999; Gruen and Zhang, 2005; Baltsavias et al., 2005). Grid points are determined at given positions, uniformly distributed over the whole image. As for feature points, the grid points are matched using cross-correlation and structural matching with epipolar constraint following the coarse-to-fine concept. Since grid points may lie in regions with poor texture, shadows or occlusions, the search for the match of a grid point has a higher possibility to yield ambiguity or no matching candidate. To increase the reliability of the grid point matching, the DSM generated from feature point matching is employed to constrain the matching candidate search. This will further reduce the search space and thus decrease ambiguity while speeding up the matching process.

Image edges are important features. Edges are rich, particularly in man-made environment, and associate with ridge lines and break lines on terrain. Thus, 3D edges essentially characterize surface discontinuity, and are important component of a DSM. In addition, edges play an important role in feature extraction, object recognition, 2D/3D reconstruction of man-made objects, etc.

We employ the edge extraction and matching algorithms developed in Zhang and Baltsavias (2000) in our work. This method was developed for automated 3D reconstruction of man-made objects from airborne and spaceborne images (Zhang, 2003; Baltsavias and Zhang, 2005). The method exploits rich edge attributes and edge geometrical structure information. The rich edge attributes include the geometrical description of the edge and the photometrical information in the regions immediately adjacent to the edge. The epipolar constraint is applied to reduce the search space. The similarity measure for an edge pair is computed by comparing the edge attributes. The similarity measure is then used as prior information in structural matching. The locally consistent matching is achieved through structural matching with probability relaxation. More details of the matching strategy can be found in Zhang and Baltsavias (2000) and Zhang (2003).

Edges are extracted using the Canny operator and then fitted to generate straight line. For each straight edge segment, we compute the position, length, orientation, and photometric robust statistics in the left and right flanking regions. The photometric properties include the median and the scatter matrix.

The epipolar constraint is employed to reduce the search space. The two end points of an edge segment in one image generate two epipolar lines in the other image. With the approximated height information derived from feature point and grid point matching, an epipolar band of limited length is defined. Any edge included in this band (even partially) is a possible candidate. The comparison with each candidate edge is then made only in the common overlap length, i.e. ignoring length differences and shifts between edge segments. For each pair of edges that satisfy the epipolar constraints above, their rich attributes are used to compute a similarity score. Therefore, the similarity score is a weighted combination of various criteria. The detailed computation can be found in Zhang and Baltsavias (2000).

After the computation of similarity measurement, we construct a matching pool and attach a similarity score to each candidate edge pair. Since matching using a local comparison of edge attributes does not always deliver correct results, the structural matching using probability relaxation, similar to that in point matching, are conducted. The method seeks the probability that an edge in one image matches an edge in the other image, using the geometrical structure information and photometric information of neighboring image edges. Therefore, the correspondences of both individual edges and edge structures are found. As in point matching, the computed edge similarity scores are used as prior information in structural matching. The compatibility function is evaluated using the differences between the relational measurements of two edge pairs in the stereo images. For the details of the definition of relational measurements and evaluation of compatibility function, we refer to Zhang and Baltsavias (2000) and Zhang (2003).

**EXPERIMENT RESULTS**

To test and evaluate the developed algorithms, we applied the matching approach to a set of along track IKONOS Geo stereo images in order to extract a DSM. The test site is in an area around Hobart, Australia. This scene encompasses a total area of 120 km² and consists of a variety of land cover types, including mountainous forest (to a height of 1200 m above sea level), hilly suburban neighbourhoods, parks, urban housing and commercial buildings (Fig. 2). The images were acquired towards the end of the southern hemisphere summer season. Note the cloud cover in the lower left side of the Fig. 2. A more complete description of the scene and the image data can be found in Fraser and Hanley (2005).
First, the vendor-supplied RPCs were refined with the bias compensation model using ground control points. This process corrected the bias in the original RPCs and improved the geopositioning accuracy. The bias-corrected RPC are then used in image matching and for DSM extraction. The process began with feature point matching, and the DSM was progressively augmented with the results from grid matching and edge matching. The matched points and edges were transformed to 3D object space through space intersection using the bias-corrected RPC. Fig. 3 illustrates the generated DSM with a ground sampling distance of 5 meters using the proposed image matching strategy. This DSM is generated from the matched points using bi-cubic interpolation approach. Visual inspection reveals that good results have been achieved. The very dense terrain points enable delineation of the terrain in more detail. By combining results of multi feature matching, particularly the edge features, the fine structures of the terrain including streets, large buildings and other infrastructure are also modeled. Quantitative evaluation of the results using GPS-surveyed ground check points reveals that a height discrepancy of about 1m has been achieved. However, this evaluation represents optimal conditions at check points which are easily identifiable, highly contrasted in the image and located in locally flat areas.

A further quantitative evaluation of the DSM was performed by comparison with a lidar DSM. The lidar DSM is located within the Hobart test site, covers a long strip and contains a diversity of land cover including buildings and suburban housing in central and Southern Hobart (Fig. 4). The elevation of range is about 300m. The lidar data has a 1.25m average ground spacing. The planimetric accuracy for the first-pulse was better than 1m, with standard error of heighting being estimated at 0.25m (AAMHatch, 2004). First, the DSM heights from IKONOS stereo imagery were compared with against the lidar height data. This assessment reveals that the RMS discrepancy was around 4m, indicating that the generated DSM is indeed a good representation of the actual terrain. However, the assessment does not provide an insight into variability of accuracy associated with areas of different land cover.
Figure 3. Generated DSM from IKONOS Geo stereo imagery over Hobart, Australia.
Figure 4. Shaded reference lidar DSM with 2m grid spacing.

In order to take into account the influence of topographic variation and land cover variability on the modeled surface, the lidar strip area was divided into separate sub-areas. These comprised urban and rural areas, with a further subdivision into regions within these two categories. Table 1 gives the DSM accuracy evaluation results. The accuracy of the generated DSM is generally in the range of 2.0 to 6.0m. Height accuracy is better in bare ground areas, while the accuracy degrades in built-up urban areas. The accuracy is worse in forest areas, since image matching is susceptible to difficulties in forest due to the poor contrast of image contents and shadows. The generated DSM is usually higher than lidar reference in forest areas, partially due to the fact that the laser can penetrate into forest canopies. In urban areas, the large discrepancy can be also contributed from lidar when the laser erroneously strikes the vertical profile of an object (e.g. building walls) and is misinterpreted as surface.

Table 1. Accuracy evaluation of the generated DSM over Hobart, Australia.

<table>
<thead>
<tr>
<th>Land cover</th>
<th>RMSE(m)</th>
<th>Mean(m)</th>
<th>Abs Max(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBD</td>
<td>4.0</td>
<td>1.6</td>
<td>38.6</td>
</tr>
<tr>
<td>Residential</td>
<td>2.6</td>
<td>0.8</td>
<td>22.8</td>
</tr>
<tr>
<td>University</td>
<td>2.8</td>
<td>0.1</td>
<td>27.0</td>
</tr>
<tr>
<td>Building</td>
<td>2.9</td>
<td>1.1</td>
<td>21.3</td>
</tr>
<tr>
<td>Sporting fields</td>
<td>2.7</td>
<td>0.2</td>
<td>14.6</td>
</tr>
<tr>
<td>Park</td>
<td>3.1</td>
<td>1.2</td>
<td>29.8</td>
</tr>
<tr>
<td>Gardens</td>
<td>2.9</td>
<td>0.9</td>
<td>34.9</td>
</tr>
<tr>
<td>Bare ground</td>
<td>2.1</td>
<td>0.5</td>
<td>8.7</td>
</tr>
<tr>
<td>Sporting fields</td>
<td>2.6</td>
<td>0.2</td>
<td>12.9</td>
</tr>
<tr>
<td>Forest</td>
<td>6.3</td>
<td>1.8</td>
<td>43.6</td>
</tr>
</tbody>
</table>

DISCUSSION AND CONCLUSION

This work has developed algorithms and programs necessary to automatically generate a DSM from high-resolution satellite imagery, consisting of a combination of feature points, grid points, and edges. Key components presented are methods to explore IKONOS Geo stereo imagery for producing dense and detailed DSMs over large areas. First, the vendor-supplied sensor model coefficients must be refined using a bias compensation model to achieve sub-pixel geopositioning accuracy. Then the DSM is automatically generated by a sophisticated image matching approach. The matching approach involves an integration of feature point, grid point and edge matching algorithms, makes use of the explicit knowledge of the image geometry, and works in a coarse-to-fine hierarchical strategy. The
coarse-to-fine strategy allows for the matching process following an image pyramid approach, while progressively reconstructing the surface model from feature points, grid points to edges. This strategy reduces search space, provides more reliable results, and speed up the process. For the matching of each feature, a two-step scheme is employed in which the candidates are first found using normalized correlation coefficient (for points) or by comparing the attributes (for edges), while the final matches are located by a structural matching algorithm with probability relaxation. This scheme avoids a hard threshold in deciding matches which usually causes commission and omission errors, while providing consistent results in a local neighborhood. The integration of multi features for DSM generation is another advantage of the proposed approach. The grid point matching allows for bridging gaps or holes in regions with poor or no texture. The integration of edges in the DSM is particularly useful and preserves the discontinuity of the terrain that allows for better characterization of terrain.

We have presented the result of processing of IKONOS Geo stereo images over a test site in Hobart, Australia with accurate ground control points, nearly 1300m height range and variable land cover. The result was compared with reference data from airborne laser scanning. The general height accuracy is around 4m over topographically diverse areas. The quality and accuracy of the generated DSM improves in the open rural areas. The largest errors are usually found in forest areas or urban centers. There are also contributing errors from the reference data.

ACKNOWLEDGEMENTS

The work described in this paper is partially supported by the U.S. Geological Survey. The lidar data is provided by AAMHatch Pty Ltd through CRC for Spatial Information, Australia.

REFERENCES


