DEM generation from an IKONOS stereo pair using EpiMatch and Graph-Cut algorithms

Taejung Kim, Yong-Jo Im, Ho Won Kim*, In So Kweon*

Satellite Technology Research Center, *Department of Electrical Engineering and Computer Science

KAIST, 373-1 , Yusung-gu, Taejon, 305-701, ROK

E-mail : {tjkim, yjim}@satrec.kaist.ac.kr, imp97@kaist.ac.kr, iskweon@ee.kaist.ac.kr

ABSTRACT: In this paper, we report the development of two DEM (digital elevation model) generation algorithms over urban areas from an IKONOS stereo pair. One ("EpiMatch") is originally developed for SPOT images and modified for IKONOS images. It uses epipolar geometry for accurate DEM generation. The other is based on graph-cut algorithm in 3D voxel space. This algorithm is believed to work better on height discontinuities than EpiMatch. An IKONOS image pair over Taejon city area was used for tests. Using ground control points obtained from differential GPS, camera model was set up and stereo matching applied. As a result, two DEMs over urban areas were produced. Within a DEM from EpiMatch, small houses appear as small "cloudy" patches and large apartment and industrial buildings are visually identifiable. Within the DEM from graph-cut, we could achieve better height information on building boundaries. The results show that both algorithms can generate DEMs from IKONOS images although more research is required on handling height discontinuities (for "EpiMatch") and on faster computation (for "Graph-cut").

KEY WORDS: DEM, Camera Model, Stereo Matching, Graph Cut, IKONOS

1. Introduction

A DEM over urban areas is a useful tool in many civil engineering applications and techniques for generating DEMs are an active reseach theme for many years. In the past, these DEMs were prepared from digitizing contour lines of topographic maps and by stereo analysis of low or middle resolution satellite images. However, the public release of 1m resolution spaceborne images proposes a precise and accurate generation of urban DEM from high-resolution stereo images.

1m resolution images, on the other hand, possess unique geometric and radiometric characteristics due to shadows, occlusions and height discontinuities. These make the application of traditional stereo matching techniques for mid or low resolution difficult. This paper reports the development of two stereo matching algorithms to generate DEMs over urban areas from IKONOS images.

One algorithm can be regarded as modification of the traditional stereo matching algorithm. This algorithm ("EpiMatch") finds correspondences by local search within a search range carefully guided by the epipolar geometry of the stereo pair. This algorithm was originally developed for SPOT image and modified here for IKONOS images. The other algorithm ("Graph-cut") generates DEMs by global energy minimization scheme. It defines edge capacity in 3D voxel space and calculate the maximum flow within the whole voxel space using graph-cut algorithm. Each algorithm is described next.

2. Algorithm Description

2.1. DEM extraction by EpiMatch

Generally, the epipolar relation can be established in stereo images and this can be very useful in stereo matching process. The EpiMatch algorithm utilizes this relation to guide local search for correspondences. The epipolarity for linear push-broom type camera is different from frame grabber type camera (Kim, 2000). For IKONOS images, we used the epipolarity based on the linear pushbroom direct linear transformation (DLT) model (Gupta and Hartley, 1997). Defining the epipolarity and the corresponding epipolar curves can limit matching around epipolar curves and hence ease the problem of stereo matching to a great deal (see figure 1).

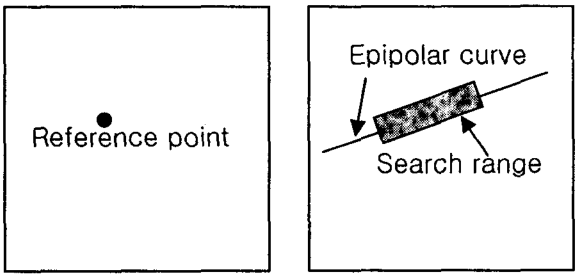


Figure 1. An epipolar curve and search range

The determination of patch size is also an importance factor determining the accuracy. The EpiMatch uses a uniform size for the template (or left) patch but defines the right patch size adaptively taking the scale difference between the left and right. The scale difference may occur due to different tilt angles of the satellite. The left and right patches are further rotated along with epipolar curves so that their ground foot prints coincide.

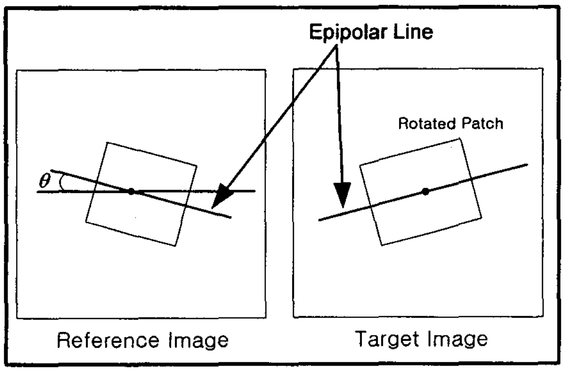


Figure 2. The definition of left and right patch

With the search range and patches defined above, the EpiMatch determines the correspondences by finding a point with the maximum normalized cross correlation value.

This algorithm has been applied to SPOT images (10m resolution) and KOMPSAT-1 EOC images (6.5m resolution) successfully. This algorithm was NOT designed to handle the height discontinuities of 1m resolution images. Height retrieval on building boundaries are not expected. In order to test how well (or poorly) this algorithm can work on such problems, the EpiMatch was applied to IKONOS images.

3.2. DEM extraction by Graph-cut

A new DEM generation scheme was developed. Instead of local search, this scheme uses global energy minimization to find optimal DEM for a given urban scene. Such global approaches in finding stereo correspondences are gaining recent research interests. Among many, the energy minimization scheme by graph-cut (Roy and Cox, 1998) is one popular example. They defined 3D energy spaces using the 2D image space and the direction of disparity appropriately. They then generated optimal disparity maps by estimating capacity of the nodes of 3D energy space and calculating maximum flow within the energy space. This method was reported to work successfully for multiple cameras with a short baseline.

Satellite images in general have very large baseline and the previous method cannot be applied directly. In case of the short baseline, search region is small and the problem of local minima is not severe. In case of satellite images, baseline is large (typically of several hundreds of kms), there are homogeneous intensity patterns (in particular on shadows) and hence the problem of local minima is severe. In addition to these, occlusions due to tall buildings create another problem of information missing.

In this paper, we, instead of using the point-noded 3D energy space, used 3D voxel space and the edge capacity between voxels. Suppose there are two adjacent voxels, u, and v. Using the camera projection equation $p_i^j = P^j(i)$, we can project voxel centers onto the left and right images as shown in figure 3 and obtain the projected pixels $p_u^l, p_u^r, p_v^l, p_v^r$. Also we can define two lines, $l_{u,v}^l$ and $l_{u,v}^r$.

$$l_{u,v}^l = \alpha \cdot \vec{l}_{u,v}^{\,l}$$
$$l_{u,v}^r = \beta \cdot \vec{l}_{u,v}^{\,r} \quad (1)$$

$\vec{l}_{u,v}^{\,i}$ in the above expression is a unit direction vector.

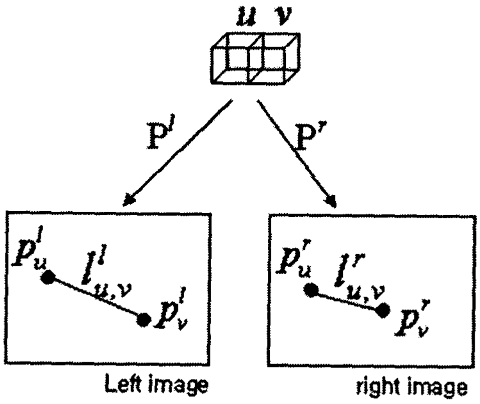


Figure 3. Voxel projection

The distance between adjacent voxel centers is constant but the distance of the corresponding projected distance is not. To take this into account, we normalize the lines as

$$\hat{l}_{u,v}^r = \frac{\alpha}{\beta} \cdot l_{u,v}^r \quad (2)$$

We can define the edge capacity $c(u,v)$ between voxels u, v as

$$c(u,v) = \frac{\sum_{i=0}^{\alpha} \left(I_l(l_{u,v}^l, i) - I_r(\hat{l}_{u,v}^r, i) \right)^2}{\alpha} \quad (3)$$

Since we are using Euclidean voxel space, we can define the edge capacity with six neighboring pixels (see figure 4). To all six, the same weight is assigned.

We define the source and sink as in figure 5 and search for the maximum flow over the whole voxel space. In our case, we can set the sink as the reference ground surface of zero height. In order to ease the problem of homogeneous intensity, a shadow mask was provided separately. Height information over the shadow mask was not estimated by Graph-cut but later interpolated as average height of the surroundings.

The next section will show the DEM obtained from this graph-cut algorithm.

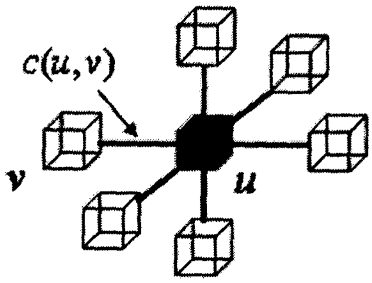


Figure 4. Definition of edge capacitance

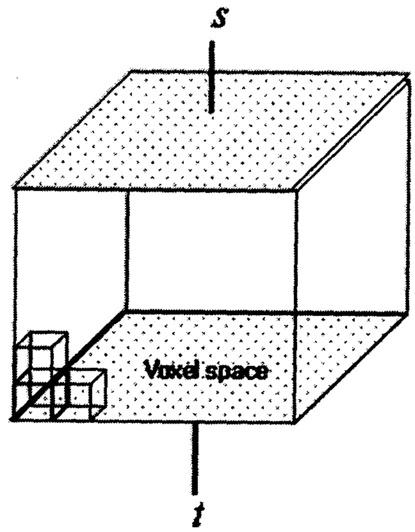


Figure 5. Source, sink and the graph

3. Experiments and Results

IKONOS images over Taejon city were used for the experiments. The left and right pair was acquired on the same day of 19th November, 2001. They were precision-geometrically corrected and then resampled to a quasi-epipolar geometry (provided by the courtesy of e-HD.com). The size of each image was 9700 by 9616 pixels and the images were believed to be subtracted from original full-size images. The images cover approximately 8kms by 8kms on the ground. Figure 6 shows the IKONOS image (left). The image contains dense population of residential houses, apartments and industrial buildings as well as rivers and hills. Automated and reliable DEM extraction from such images is a challenge to any stereo matching algorithms.

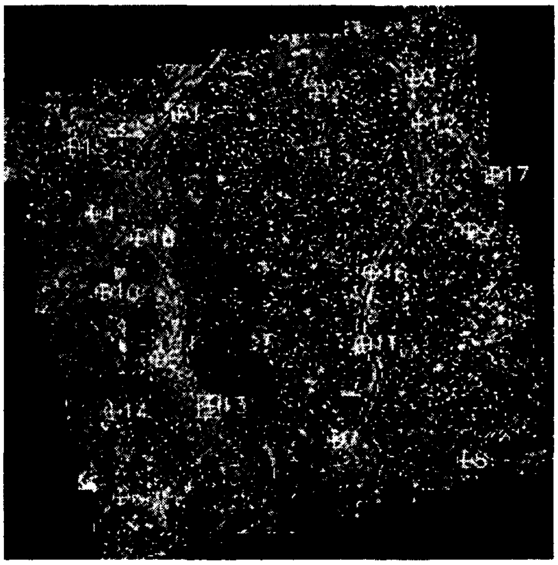


Figure 6. The IKONOS image used for experiments. (Only the left image is shown). Crossed circles indicate ground control points used.

EpiMatch and Graph-cut algorithms were applied to the dataset. Figure 7 shows the DEMs. Both DEMs contained a good topography of the scene but the DEM fro Graph-cut preserves better urban surface patterns. Note that the orientation of the two DEMs appeared different due to the different algorithm setting but they do cover the same ground surface. Figure 8 compares the results more detail. The DEM from the EpiMatch also possesses building heights but the building boundaries appeared as blurred clouds. The DEM from Graph-cut, on the other hands, recovers building heights better than EpiMatch. However, a closer look of the DEM tells that still the height around building boundaries was somewhat spread.

Also it seems that Graph-cut produces some blunders where brightness patterns were homogeneous. These blunders appeared as white blobs in the DEM. For EpiMatch, there are dark artifacts in the DEM. These seemed due to match failure.

In summary, both EpiMatch and Graph-cut produced DEMs from IKNOS image and both contains a good topography. Graph-cut seemed to generate better DEMs over dense urban areas, partly due to the voxel-based global enforcement and partly because the shadow masking. Improvements are required on both algorithms to reduce mismatches and to handle height discontinuities better. One other aspect is computation time and memory requirements. While EpiMatch took less than two CPU hours and required the memory of a few hundreds Mbytes, Graph-cut required a memory of a few Gbytes.It was not feasible to run Graph-cut using the full scene and hence the image in figure 6 was split into smaller segments and Graph-cut applied to each segment.

4. Conclusions

This paper reports the development of two DEM (digital elevation model) generation algorithms over urban areas from an IKONOS stereo pair. The results show that both algorithms can generate DEMs from IKONOS images although more research is required on handling height discontinuities (for "EpiMatch") and on faster computation (for "Graph-cut").

References

Gupta, R., and Hartley, R., 1997. Linear pushbroom cameras, *IEEE Trans. PAMI*, 19(9), pp.963-975.

Kim, T., 2000. A study on the epipolarity of linear pushbroom images, *PE&RS*, 66(8), pp.961-966.

S. Roy and I. J. Cox, "A Maximum-Flow Formulation of the N-camera Stereo Correspondence Problem", *Int. Conference on Computer Vision, Bombay, India, January 1998*

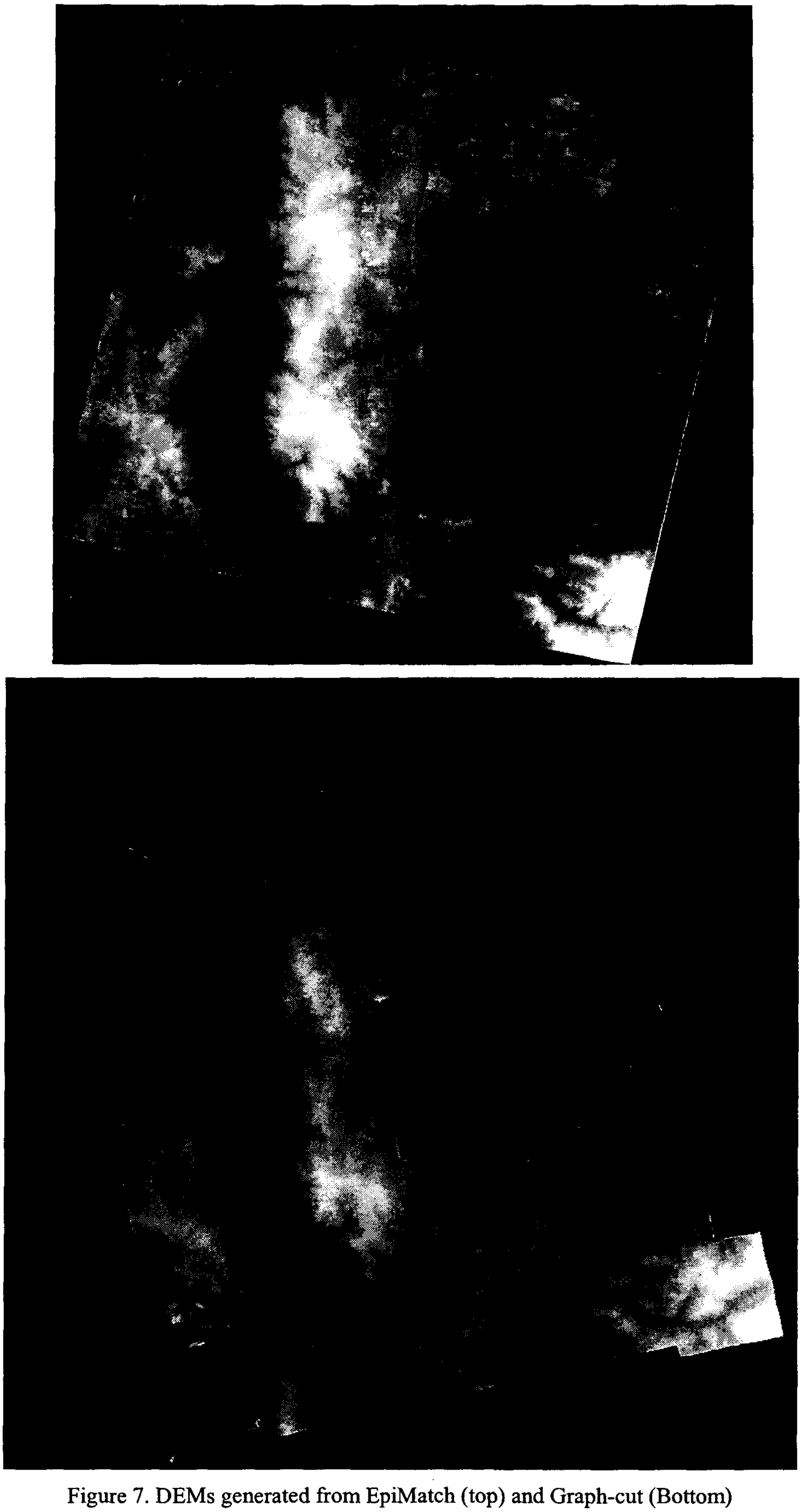
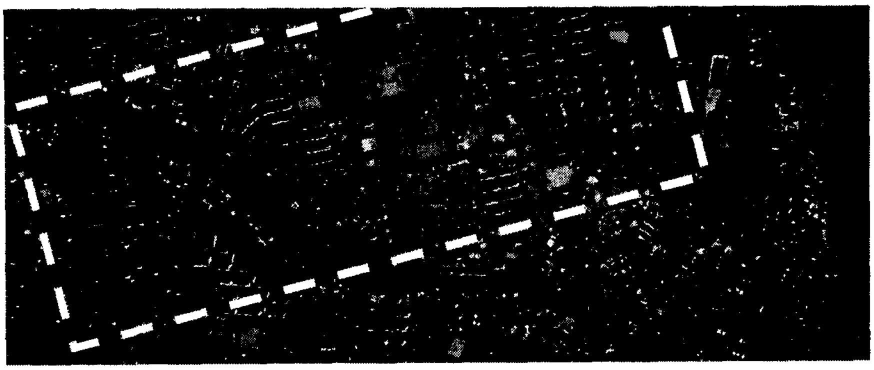
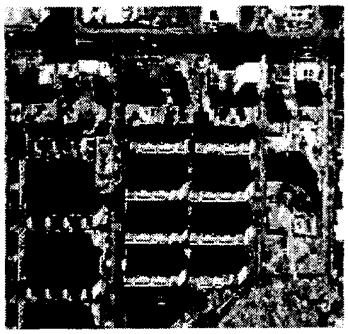


Figure 7. DEMs generated from EpiMatch (top) and Graph-cut (Bottom)





(a) IKONOS image over Dunsan area

(b) Image over apartments

(c) A DEM from EpiMatch over Dunsan area

(d) DEM from EpiMatch

(e) A DEM from Graph-cut over Dunsan area

(f) DEM from Graph-cut

Figure 8. DEMs from EpiMAtch and Graph-cut at larger scale