Incorporation of Scene Geometry in Least Squares Correlation Matching for DEM Generation from Linear Pushbroom Images

Taejung Kim*, Taehun Yoon*, Heung Kyu Lee**

***Satellite Technology Research Center, **Department of Computer Science**

Korea Advanced Institute of Science and Technology

373-1 Kusung-dong, Yusung, Taejon, 305-701, Republic of Korea

{tjkim|yoonth}@satrec.kaist.ac.kr

Abstract: Stereo matching is one of the most crucial parts in DEM generation. Naïve stereo matching algorithms often create many holes and blunders in a DEM and therefore a carefully designed strategy must be employed to guide stereo matching algorithms to produce "good" 3D information. In this paper, we describe one such a strategy designed by the use of scene geometry, in particular, the epipolarity for generation of a DEM from linear pushbroom images. The epipolarity for perspective images is a well-known property, *i.e.*, in a stereo image pair, a point in the reference image will map to a line in the search image uniquely defined by sensor models of the image pair. This concept has been utilized in stereo matching by applying epipolar resampling prior to matching. However, the epipolar matching for linear pushbroom images is rather complicated. It was found that the epipolarity can only be described by a hyperbola-shaped curve and that epipolar resampling cannot be applied to linear pushbroom images. Instead, we have developed an algorithm of incorporating such epipolarity directly in least squares correlation matching. Experiments showed that this approach could improve the quality of a DEM.

1. Introduction

A DEM (digital elevation model) is one of the key information in analyzing geo-spatial properties of the Earth's surface and this has numerous applications in many fields. Automatic extraction of such DEMs from satellite images is one of the most important tasks in remote sensing, photogrammetry and image understanding communities. Although there have been many researches devoted for this task, it seems full automation has not been achieved yet.

Authors believe one of the reasons for this situation is the lack of understanding of the geometry of satellite images and stereo image pairs, in particular, taken from linear pushbroom-type cameras. Unlike perspective ones, linear pushbroom images do not have universally accepted sensor models [1]. Until recently [2] the epipolarity of linear pushbroom image pairs, the most significant and useful geometric constraint among stereo pairs, has been misled by a (vague) belief that this be the same, (or slightly different but assumed to be the same) as that of perspective image pairs [3,4]. This, however, is not the valid case [2]. There is distinct epipolarity of linear pushbroom image pairs and many of former approaches, such as epipolar resampling, can not be applied to linear pushbroom images [2].

This paper deals with the development of a stereo matching algorithm for the generation of DEMs from linear pushbroom image pairs. The algorithm is modified from the adaptive least squares algorithm proposed by [5] using the epipolarity of linear pushbroom images. Compared with the ones without such epipolarity, this algorithm has achieved greater accuracy, coverage and reliability at the cost of (slight) computation time increase.

We organise this paper as follows: we will first introduce the stereo matching algorithm proposed here. Next, we will describe the experiment set-up and results. We then discuss the performance of the proposed algorithm in comparison with traditional ones and the improvements achieved by incorporating scene geometry in stereo matching algorithms

2. An Epipolarity-based Stereo Matching Algorithm (EpiMatch)

The epipolarity can be stated as this: one point in the left (or right) image is mapped onto a unique curve in the right (or left) image defined by the position and orientation of the left and right camera. According to the findings in [2], epipolar curves of linear pushbroom images can be expressed as

$$y_r = \frac{A_1 x_l + A_2 y_l + A_3}{(A_4 x_l + A_5 y_l + A_6)\sin Q(x_r) + (A_7 x_l + A_8 y_l + A_9)\cos Q(x_r)}$$

where (x_l, y_l) and (x_r, y_r) are the coordinates of left and right image points respectively, $A_1 \sim A_9$ are constants and $Q(x_r)$ is a quadratic polynomial of x_r. From [2], we know that we can approximate the above curve with a line expressed below within a small region of interest.

$$y_r = E_a x_r + E_b \qquad (1)$$

The original adaptive least squares correlation [5] assumes the relationship between the left and right image pairs as an affine transformation,

$$\begin{aligned} x_r &= s_1 + a_{11} x_0 + a_{12} y_0 \\ y_r &= s_2 + a_{21} x_0 + a_{22} y_0 \end{aligned} \qquad (2)$$

and the brightness difference between the left and right match windows (or patches) be constant. From these assumptions, we can have the following observation equation [5]:

$$\begin{aligned} f - g = {} & g_x ds_1 + g_x x_0 da_{11} + g_x y_0 da_{12} + g_y ds_2 \\ & + g_y x_0 da_{21} + g_y y_0 da_{22} + r_s \end{aligned} \qquad (3)$$

Using equation (1) and (2), we can derive new observation equations between the left and right image pairs:

$$\begin{aligned} 0 = {} & w(E_a ds_1 + E_a x_0 da_{11} + E_a y_0 da_{12} - ds_2 \\ & - x_0 da_{21} - y_0 da_{22}) \end{aligned} \qquad (4)$$

This equation states the final solution shall lie on the epipolar curve (or line). We introduce a weighting factor, w, to determine relative importance between the equations (3) and (4). If we set a large value for w, the equation (4) has more effects on estimation and estimated match pairs tend to lie closer to epipolar curves. If we set a small value for w, the equation (3) has more effects and final solution tends to satisfy the brightness constraint (equation (3)) more than the epipolar constraint (equation (4)). If we set a zero for w, this is identical to the original algorithm proposed by [5].

We now have the following matrix expression for the least squares correlation. We can obtain a solution by repeating the procedure of finding the solution matrix $\mathbf{x}$ and updating the observation matrix $\mathbf{A}$.

$$\mathbf{l} = \mathbf{A}\mathbf{x}$$

$$\mathbf{l} = \begin{Bmatrix} f(x_l, y_l) - g(x_0, y_0) \\ 0 \end{Bmatrix}$$

$$\mathbf{x}^{\mathrm{T}} = (ds_1, da_{11}, da_{12}, ds_2, da_{21}, da_{22}, r_s)$$

$$\mathbf{A} = \begin{Bmatrix} g_x, g_x x_0, g_x y_0, g_y, g_y x_0, g_y y_0, 1 \\ wE_a, wE_a x_0, wE_a y_0, -w, -w x_0, -w y_0, 0 \end{Bmatrix}$$

Compared with the original algorithm, the observation matrix above has doubled its number of rows. However, this increase does not contribute to the computation time significantly since the half elements of the matrix $\mathbf{l}$ are zeros. We can partition the matrices $\mathbf{A}$ and $\mathbf{l}$ to achieve computational efficiency.

3. Intelligent Match Strategy

For accurate DEM generation, match strategy as well as the mathematical formulation of stereo matching problems is very important. Naïve match strategy often results in poor DEMs no matter what stereo matching algorithms are used. This paper, we propose two such strategies: intelligent determination of match window (or patch) shape and match candidates.

3.1. Patch Shape Determination

The proposed algorithm, as well as the original algorithm, defines match windows (or patches) in the left and right images and the elements of the matrices **A** and **l** are calculated on the grid points defined on the patches. The shape and size of patches contribute the overall performance of stereo matching significantly. Before, we defined rectangular patches with an identical size in the left and right image. However, we have found out that there is a better way for this: using epipolar curves and incidence angles of left and right image pairs.

Since we can determine the exact shape of epipolar curves in the left and right images, we can define the left and right patches to lie along with these epipolar curves. By this way, the object surfaces represented by the left and right patches can overlap more. On the other hands, satellite stereo pairs are obtained mostly by tilting left or right or both cameras. Due to these tile angles, the scale of an image (or a patch) changes. Hence, we can define the size of patches according to these scale differences. We tested and compared the performances of the proposed algorithm with the patches whose shape and size are determined in this way and in the traditional way. In the following sections, we show the improvements achieved by this intelligent patch determination.

3.2. Match Candidate Determination

Since the EpiMatch as well as the original least squares correlation deals with non-linear problems and finds solutions through iteration, it requires initial match candidates. The quality of these candidates can effect the overall performance significantly. If one can provide candidates, which are close to true match pairs, stereo matching may find solution at a less number of iteration. On the other hand, if initial candidates are somewhat away from the true pairs, stereo matching will require a larger number of iteration or, in a worse case, fail to find true pairs.

Before, we assumed that once we found one match pair, the points adjacent to this pair could be regarded to have identical affine transformation parameters. We used this assumption to generate initial match candidates.

In this paper, we developed a new method. According to the findings in [6], it is possible to directly transform a 3D object point onto a 2D image point. We can apply this findings further to find the horizontal coordinates (x, y) of a 3D point if its height (z) and its corresponding 2D image point are known. We can now generate initial match candidates using these findings: Firstly, once a pair is matched, calculate its 3D object coordinates. Secondly, assume that a point in the left (right) image adjacent to this pair has the same height as the pair and calculate the new 3D object coordinates of the point. Thirdly, transform the new 3D object coordinates onto a 2D image point in the right (left) image.

This technique was proved very powerful in that this reduced computation time and improved match accuracy significantly. The following sections discuss this.

4. Experiments

We tested the proposed algorithm (EpiMatch) using sub-scenes of SPOT images. We selected two test sites, one over Kunwee-Whabuk region of the KyungNam province (site 1) and the other over Boruyng-Buyeo region of the ChoongNam province (site 2). Appropriate sub-scenes were extracted from two SPOT stereo pairs. Each pair has different incidence angles of 13.2° and –10.2° for the site 1 and –6° and –25.8° for the site 2. Only the left sub-scenes of the two pairs are shown (figure 1). Both scenes contain hilly regions and rivers. Unfortunately, the stereo pairs for the site 1 have poor contrasts on "rugged" hilly regions. For the site 2, the brightness patterns of the river regions are completely different between the left and right images (in the left the river appeared bright whereas in the right very dark). These factors can degrade the quality of output DEMs. These, however, can provide the limit a stereo matching software can handle and experiments were carried out with these scenes.

For each test scene, ground-truth DEMs were prepared. These DEMs are called as "DTED" (Digital Terrain Elevation Data) and produced by the US army. In order to access the accuracy of these ground-truth DEMs, we compared these with DEMs created from topographic maps at 1:25,000 scale and data collected with differential GPS (Global Positioning System) receivers. We verified that these ground-truth DEMs were within a root mean square error of half a pixel (in meters, 5m). Using these ground-truth DEMs, we compared the DEM automatically produced from our experiments. The comparison was carried out by point-by-point basis, *i.e.*, each and every point in a DEM produced was compared with the ground-truth DEMs.

Table 1 summarises experiment results. We compared the performances of the original adaptive least squares correlation (Method 1), the proposed EpiMatch algorithm (Method 2), the EpiMatch algorithm with intelligent patch determination (Method 3), and the EpiMatch algorithm with intelligent patch determination and match candidate generation (Method 4). As performance measures, we used the average time taken to match one point (speed), the ratio of the number of matched points over the number of maximum match points possible (coverage), and the maximum, minimum, average and root mean squares errors of DEMs produced compared with the ground-truth DEMs (accuracy).

The results of the original adaptive least squares correlation were very striking. These had huge errors in their DEMs, which revealed the limitation of naïve stereo matching algorithms. In fact, the Method 1 was originally developed for perspective scenes, where geometric constraints were not as severe as linear pushbroom ones. Normally, this algorithm is applied to perspective images after epipolar resampling. In our experiments, epipolar resampling was not applied since this process alone could create large errors for linear pushbroom images [2] and to highlight the (poor) results of stereo matching without proper geometric constraints.

The results of the proposed EpiMatch algorithm showed significant improvements in terms of coverage and accuracy for the site 1. The speed had been increased only marginally although the size of matrices was doubled. It was very interesting to notice that the proposed EpiMatch algorithm could not match at all for the site 2. However, we considered this as improvements rather than disappointment. Since the image pairs of the site 2 had very different viewing angles, patches defined in the traditional way were not appropriate. The original stereo matching algorithm (Method 1) could not detect this factor and produced an output with huge errors. The EpiMatch algorithm, instead, did not perform estimation. The reason why the EpiMatch algorithm worked for the site 1 was that the image pairs had viewing angles with similar absolute values

The results of the EpiMatch algorithm with intelligent patch determination also showed improvements. For the site 1, a significant improvement on the accuracy had been achieved by the use of intelligent patches, where the coverage remained the same. For the site 2, the intelligent patch determination enabled successful stereo matching from image pairs with different viewing angles. Of course the accuracy of the output DEM had been improved from that of the Method 1 by a large amount. With these results, we could say that the usual preference of selecting satellite image pairs whose viewing angles are similar in absolute values but opposite in sign be overruled. It was very encouraging to see that the speed had improved for the both sites. This implied that imposing geometric constraints in patch determination as well as mathematical stereo matching formulation could increase the speed, let alone the accuracy.

The results of the EpiMatch algorithm with intelligent patch determination and match candidate generation further supported the findings of the Method 3. By generation of match candidates intelligently, the speed, coverage and accuracy were all improved. These results underpinned the motivation of new match candidate generation scheme that this should create more accurate match candidates, reduce the number of iteration required to find solutions, and hence increase the speed of stereo matching.

Due to the poor contrasts in "rugged" areas, the accuracy for the site 1 was larger than for the site 2. Both scenes still had some blunders due to the poor contrasts, opposite brightness patterns in river regions and some errors by the stereo matching algorithm. In figure 2, we showed the DEMs generated by the Method 4. Note that since the DEMs were plotted based on the ground reference system, these were not aligned with image coordinates. The DEM for the site 2 had holes on river regions. This was because stereo matching was not performed on these regions and we applied some intelligent automated interpolation schemes to remove blunders [7]. (The accuracy figures stated in table 1 were before removing blunders).

5. Conclusions

In this paper, we described an approach of incorporating geometric properties of linear pushbroom images in stereo matching and DEM generation. We first introduced a new mathematical formulation of the stereo matching problem by the use of the epipolarity of linear pushbroom images. We further developed matching strategies to incorporate geometric properties in intelligent ways.

Through the analysis of experiment results, we found out that naïve stereo matching without proper geometric constraints could not be used, at least, for linear pushbroom images. We also confirmed that our approach of incorporating scene geometry was valid. More importantly, we learned that geometric constraints could improve the speed of stereo matching as well as the accuracy.

In this paper, we have considered the geometric properties derived only from sensor models. Many other geometric constrains can be derived from object surface properties such as surface slopes, types and so on. Further research on this issue is required.

Acknowledgements

The work described in this paper is supported by the Ministry of Science and Technology of the government of the Republic of Korea under the project, "Development Study of EOC Image Data Processing and Application Technology".

References

[1] McGlone, C., 1996, "Sensor Modeling in Image Registration", Digital Photogrammetry-An Addendum to the Manual of Photogrammetry, edited by Greve, C., published by ASPRS, chapter 5, pp. 115-123

[2] Kim, T, 1999, "A Study on Epipolarity of Linear Pushbroom Images", *Photogrammetric Engineering and Remote Sensing*, (in press)

[3] Al-Rousan, N., Cheng, P., Petrie, G., Toutin, Th., Valadan Zoej, M.J., 1997, "Automated DEM Extraction and Orthoimage Generation from SPOT Level 1B Imagery", *Photogrammetric Engineering and Remote Sensing*, 63(8):965-974

[4] Otto, G.P., 1988, "Rectification of SPOT Data for Stereo Image Matching", *Archives* of the 16th Congress of International Society for Photogrammetry and Remote Sensing, pp. 635-645, Kyoto, Japan, 1-10 July

[5] Gruen, A.W., 1985, "Adaptive Least Squares Correlation: A Powerful Image Matching Techniques", *South African Journal of*

Photogrammetry, Remote Sensing and Cartography, 14(3):175-187

[6] Kim, T., Shin, D., and Lee, Y-R, 1999, "Development of a Robust Algorithm for Transformation of a 3D Object Point onto a 2D Image Point for Linear Pushbroom Imagery", Taejung Kim, Dongseok Shin and Young-Ran Lee, *Photogrammetric Engineering and Remote Sensing*, Submitted

[7] Kim, S.A.B., Kim, T., Park, W., Lee, H-K., 1999, "An optimal interpolation scheme for producing a DEM from the automated stereo matching of full-scale SPOT images", *Proceedings* of the ISPRS conference "Sensors and Mapping from Space 1999", Hanover, Germany

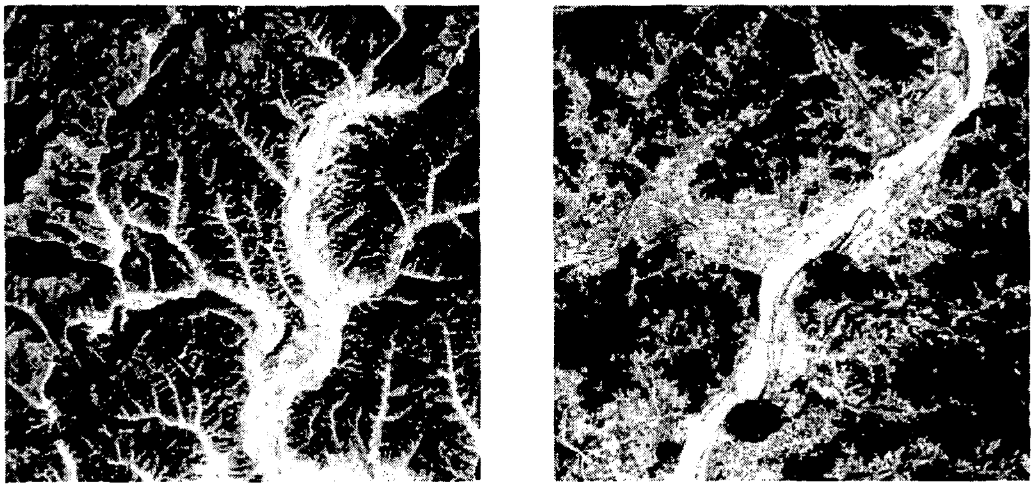

Figure 1. Test scenes of Kunwee-Whabuk (left) and Boryung-Buyeo (right)

Table 1. The comparisons of the performances of the original adaptive least squares correlation (method 1), the proposed EpiMatch algorithm (method 2), and the EpiMatch algorithm with intelligent patch determination (method 3) and the EpiMatch algorithm with intelligent patch determination and match candidate generation (method 4).

	Site 1				Site 2			
	Method 1	Method 2	Method 3	Method 4	Method 1	Method 2	Method 3	Method 4
Speed	220 msec	224 msec	145 msec	49 msec	163 msec	202 msec	140 msec	40 msec
Coverage	26 %	66 %	65 %	76 %	63 %	1 %	82 %	83 %
Accuracy (m)								
Max. Error	-2591.590	-521.239	-578.149	-403.657	-1253.020	-48.420	-595.447	-206.405
Min. Error	2460.410	721.022	346.638	362.128	1476.819	412.224	673.273	170.225
Aver. Error	320.922	43.738	33.812	38.377	99.368	89.363	77.871	17.369
RMS. Error	510.062	86.892	58.294	52.064	191.298	133.347	120.430	29.532

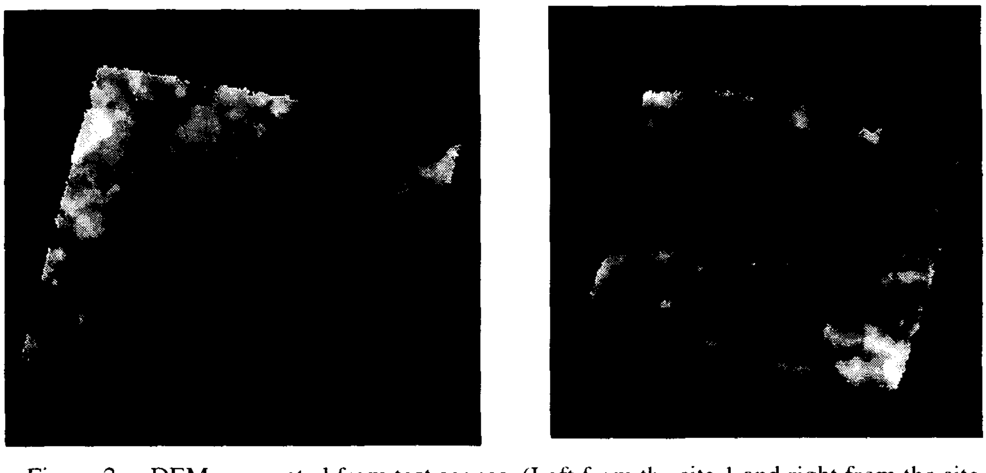

Figure 2. DEMs generated from test scenes. (Left from the site 1 and right from the site 2)