Korean Journal of Geomatics
Vol. 1, No. 1 / December 2001
pp. 43~49

Post Processing to Reduce Wrong Matches in Stereo Matching

Hee-Ju Park* and Suk-Bae Lee**

Abstract

Although many kinds of stereo matching method have been developed in the field of computer vision and photogrammetry, wrong matches are not easy to avoid. This paper presents a new method to reduce wrong matches after matching, and experimental results are reported. The main idea is to analyze the histogram of the image attribute differences between each pair of image patches matched. Typical image attributes of image patch are the mean and the standard deviation of gray value for each image patch, but there could be other kinds of image attributes. Another idea is to check relative position among potential matches. This paper proposes to use Gaussian blunder filter to detect the suspicious pair of candidate match in relative position among neighboring candidate matches. If the suspicious candidate matches in image attribute difference or relative position are suppressed, then many wrong matches are removed, but minimizing the suppression of good matches. The proposed method is easy to implement, and also has potential to be applied as post processing after image matching for many kinds of matching methods such as area based matching, feature matching, relaxation matching, dynamic programming, and multi-channel image matching. Results show that the proposed method produces fewer wrong matches than before.

Keywords : *stereo matching, post processing, Gaussian blunder filter, wrong match*

1. Introduction

Stereo matching is one of the major subjects in the field of digital photogrammetry and computer vision. Also is in the field of remote sensing since high resolution imagery is available. The major goals of it are DTM/DEM/DSM generation, semi-automatic inner orientation, automatic inner orientation, relative orientation, absolute orientation, semi-automatic object detection, automaticobject detection, object recognition, etc.

The cross correlation method is a simple statistical method to determine the correspondence between image patches based on correlation of gray value between two image patches when there are enough contrast in both image patches to be compared. It checks only the local correlation between the reference image patch and the target one. However, there are many image patch pairs that have high correlation, but are not conjugate image patches.

Therefore many wrong matches are found in the results of matching by correlation methods. When the search space to find conjugate image patch gets bigger, the ratio of wrong match becomes bigger. If the image patch size (reference and target image patch size) is bigger, the ratio of wrong match is reduced, but the results are not always good. Many techniques have been developed to overcome the weakness of this method. Searching along the epipolar line is a good method to reduce the search space.[1),2)] It reduces search space from 2 dimension to 1. The hierarchical image matching or coarse to fine matching strategy with image pyramid is highly recognized. It contributes to reduce the search space at each pyramid level, and thus it saves time and also reduces wrong matches. The least square matching is highly considered method for precision improvement.[3)] It greatly improves the precision of image matching when the approximate corresponding position is known, and also the surface of the region is homogeneous and has enough contrast. The relaxation, or relational matching infers the relation among points, lines, and objects.[4)-6)] It considers not only local similarity but also global similarity. However, many wrong matches for real images are found in that results. Sometimes there are no corresponding point to some points in real images, or the effect of relief on geometric distortions is too big. But it is not easy to know automatically whether a point has conju-

*Prof. of Department of Cadastral Information, Shingu College, Korea (Email : heeju@shingu-c.ac.kr)
**Prof. of Department of Civil Engineering, JeonJu Technical College, Korea (Email : sblee@jtc.ac.kr)

The manuscript for this paper was submitted for review on December 10, 2001.

gate points or not, whether similarity measure to a matching entity is proper or not, whether a match is wrong or good. Many methods were devised how to suppress wrong matches after main part of matching is over. Relative orientation is used to suppress them. But it is meaningless to epipolar imagery, where geometric constraint is already used. If multi-photo is available to a certain area, another geometric constraint can be used. But it is not for stereo images. Median filter was used to suppress the blunders of DEM after image matching.[7] But median filter requires grid type data. Also some wrong matches may play bad roles during interpolation of DEM generation. How to efficiently suppress wrong matches after main stereo matching is the main focus of this paper. Fundamental idea is to detect blunders in the histogram of image attribute difference, and to detect blunders of parallax difference comparing to neighborhood candidate matches by Gaussian filter. This paper deals with the case there are stereo imagery and its orientation parameters are available, and main part of matching is by correlation method or by least square method. Therefore, searching on epipolar line, and hierarchical image with image matching is assumed to be performed in main part of matching.

2. Method to Reduce Blunders

2.1 Epipolar Line Equation

The use of epipolar line in image matching is not a new one. It reduces wrong matches greatly, and saves time. If the orientation parameters for two aerial photographs are known, it is possible to know where epipolar lines are for a given point to be matched. The epipolar lines are straight when there is no distortion during image acquisition, for example, lens distortion and film processing. Say there are two aerial images indices 1, and 2, and the centers of projection and rotation matrixes of them are :

$$O_1 = (X_1, Y_1, Z_1),\ O_2 = (X_2, Y_2, I_2)$$

$$R_1 = \begin{pmatrix} a_{11}\, a_{12}\, a_{13} \\ a_{21}\, a_{22}\, a_{23} \\ a_{31}\, a_{32}\, a_{33} \end{pmatrix},\ R_2 = \begin{pmatrix} b_{11}\, b_{12}\, b_{13} \\ b_{21}\, b_{22}\, b_{23} \\ b_{31}\, b_{32}\, b_{33} \end{pmatrix}$$

f: focal length of camera.

The vector of base :

$$\vec{B} = (B_x, B_y, B_z) = (X_2 - X_1, Y_2 - Y_1, Z_2 - Z_1)$$

Followings are a pair of epipolar line equations for a given point (x_p, y_p) on image 1.[9]

Epipolar line equation on image 1 for (x_p, y_p) on image 1 is :

$$Ax + By + Cf = 0 \tag{1}$$

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} a_{11}\, a_{12}\, a_{13} \\ a_{21}\, a_{22}\, a_{23} \\ a_{31}\, a_{32}\, a_{33} \end{pmatrix}^T \begin{pmatrix} 0 & -B_z & B_y \\ B_z & 0 & -B_x \\ -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} a_{11}\, a_{12}\, a_{13} \\ a_{21}\, a_{22}\, a_{23} \\ a_{31}\, a_{32}\, a_{33} \end{pmatrix} \begin{pmatrix} x_p \\ y_p \\ f \end{pmatrix} \tag{2}$$

Epipolar line equation on image 2 for (x_p, y_p) on image 1 is :

$$A'x' + B'y' + C'f = 0 \tag{3}$$

$$\begin{pmatrix} A' \\ B' \\ C' \end{pmatrix} = \begin{pmatrix} b_{11}\, b_{12}\, b_{13} \\ b_{21}\, b_{22}\, b_{23} \\ b_{31}\, b_{32}\, b_{33} \end{pmatrix}^T \begin{pmatrix} 0 & -B_z & B_y \\ B_z & 0 & -B_x \\ -B_y & B_x & 0 \end{pmatrix} \begin{pmatrix} a_{11}\, a_{12}\, a_{13} \\ a_{21}\, a_{22}\, a_{23} \\ a_{31}\, a_{32}\, a_{33} \end{pmatrix} \begin{pmatrix} x_p \\ y_p \\ f \end{pmatrix} \tag{4}$$

Another method to generate epipolar imagery is found in (Shenk, 1999).[10]

2.2 Candidate Image Points to Match

If all points in image matching are used as candidate points to match, it is time consuming. Furthermore, correlation method requires enough contrast. The Interest Points and edges are good candidates to match. Because they have enough contrast, and are related with big changes in height, and are useful to describe the shape of features.

2.3 Blunders in the Histogram of Image Attribute

2.3.1 The histogram of image attribute and blunder suppression

The cross correlation method is to find the image patch that has maximum cross correlation for gray values between the reference image patch and the target one. Say $s \in \{s\}$ is reference image patch, and $t \in \{t\}$ is target image patch. Then the correlation between the $s \in \{s\}$ and $t \in \{t\}$ for gray values is represented as :

$$Cor(s,t) = \frac{cov(s,t)}{\sigma(s)\sigma(t)} = \frac{\sum_{i=1}^{m}\{G(x_{si}, y_{si}) - \mu(s)\}\{G(x_{ti}, y_{ti}) - \mu(t)\}}{m\sigma(s)\sigma(t)} \tag{5}$$

here, $G(x_{si}, y_{si})$, $G(x_{ti}, y_{ti})$ are gray values of a i-th point in $s \in \{s\}$ and in $t \in \{t\}$.

$\mu(s)$, $\mu(t)$, $\sigma(s)$, $\sigma(t)$ are mean and standard deviation of gray values for the reference image patch $s \in \{s\}$ and the target image patch $t \in \{t\}$.

In the above formula there is compensation for the differences in brightness and in contrast between reference and target image patch. Surely for two overlapping images, one image is brighter than the other, and also one image has higher contrast than the other. Consider the histogram of all $\mu(s) - \mu(t)$ and that of all $\sigma(s) - \sigma(t)$ for all (s, t); $s \in \{s\}$, $t \in \{t\}$ obtained by matching. In normal distribution 99.994%

of all $\mu(s)-\mu(t)$ are within 4 × standard deviation from the median. Also 99.994% of all $\sigma(s)-\sigma(t)$ are within 4 × standard deviation from the median. The pair (s, t) which is one of (6) or (7) case has high possibility of being a wrong matches.

To reduce wrong matches, therefore, author proposes to suppress all the (s, t) pairs which are one of (6) or (7) case to reduce blunders.

$$|\{\mu(s)-\mu(t)\}-\mathrm{med}\{\mu(s)-\mu(t)\}|>4\sigma\{\mu(s)-\mu(t)\} \tag{6}$$

$$|\{\sigma(s)-\sigma(t)\}-\mathrm{med}\{\sigma(s)-\sigma(t)\}|>4\sigma\{\sigma(s)-\sigma(t)\} \tag{7}$$

here, $\sigma\{\mu(s)-\mu(t)\}$ is standard deviation for all cases of $\mu(s)-\mu(t)$ obtained by matching. And $\sigma\{\sigma(s)-\sigma(t)\}$ is standard deviation for all cases of $\sigma(s)-\sigma(t)$ obtained by matching. $\mathrm{med}\{\mu(s)-\mu(t)\}$ is median for all cases of $\mu(s)-\mu(t)$ obtained by matching. $\mathrm{med}\{\sigma(s)-\sigma(t)\}$ is median for all cases of $\sigma(s)-\sigma(t)$ obtained by matching.

However $\sigma\{\mu(s)-\mu(t)\}$ and $\sigma\{\sigma(s)-\sigma(t)\}$ obtained by matching should be smaller than those for non-corresponding pairs, where $s\in\{s\}$ has no correlation with $t\in\{t\}$. If $\sigma\{\mu(s)-\mu(t)\}$ and $\sigma\{\sigma(s)-\sigma(t)\}$ is similar to those for non-corresponding pairs, it is meaningless to suppress (6), (7) case.

2.3.2 A method to get high possibility of corresponding image patch

Regarding standard deviations and medians of (6) and (7), it is highly recommended that they should be similar to the actual corresponding (s, t) combination. If we compute such values using the pairs that have many wrong matches, it is useless. Following step will give a set of (s, t) pairs with few wrong matches.

Step 1 : Generate Interest points both in image 1 and image 2 by Interest Operator such as Foerstner Interest Operator. Then we will get image patch set $\{s\}$ on image 1, $\{t\}$ on image 2 which has enough contrast.

Step 2 : For each $s\in\{s\}$, find the matching position in search space on image 2 by correlation method, and check if the matching position is within 1 pixel from the interest point. If not, suppress it.

Step 3 : Check if vice versa of Step 2 is true : For a t whose center is the matching position to a $s\in\{s\}$, find the matching position in search space on image 1, and check if the matching position is within 1 pixel from $s\in\{s\}$. If not, suppress it.

While processing above 3 steps, image matching with epipolar images (searching on epipolar line) and hierarchical image matching is highly recommended.

2.3.3 Time considerations

If image attributers such as $\mu(s)$, $\mu(t)$, $\sigma(s)$, $\sigma(t)$ are stored in memory or in a file, it doesn't take a long time to check the cases (6), (7). Step 3 takes additional time. However, it is possible to reduce the number of Interest Points to check in Step 3. Then the additional time is not so much.

2.3.4 Use of other image attributes

It is possible to extend this concept to other kinds of image attributes, and to multi-channels of color and multi-spectral image. In that case, it is recommended to check whether the standard deviation of a image attribute difference for corresponding image patch pairs is far less than that for non corresponding ones or not.

2.4 Blunders in Parallax Difference Comparing to Neighborhood Matches

The processing described in 2.3 doesn't consider relative position among corresponding image patches. The parallax difference of a candidate match is linearly proportional to x-disparity difference of it. For potential matching pairs which are obtained by above matching using epipolar images, the weighted mean of relative disparity difference of a point i to its neighboring points can be represented as following:

$$dX_i = \frac{\sum_{j=0}^{m}(x_{si}-x_{ti}-x_{si}+x_{tj})w_{dist}(i,j)}{\sum_{j=0}^{m}w_{dist}(i,j)} \tag{8}$$

$$w_{dist}(i,j) = e^{\frac{-[\min\{d_s(i,j),d_t(i,j)\}]}{2C^2}} \tag{9}$$

here, $\{s\}$ is the reference image patch, $\{t\}$ is target one of each matching pair.

$d_s(i, j)$, $d_t(i, j)$ is the distance between i and j on image 1 and image 2.

C is constant, and it determines the size of this operator.

m is the number of $s_i\in\{s\}$ within a certain range from $s_j\in\{s\}$.

x_{si} , x_{ti} are x coordinates of centers of image patch $s_i\in\{s\}$, and $t_i\in\{t\}$.

x_{sj}, x_{tj} are x coordinates of centers of image patch $s_j\in\{s\}$, and $t_j\in\{t\}$, which are neighboring image patches to $s_i\in\{s\}$, and $t_i\in\{t\}$.

Pixel size in each pyramid level could be used as C. In normal distribution the possibility of following case (10) is less than 0.006%. So to reduce wrong matches, the author proposes to suppress following case of pairs.

$$|dX_i|>4\sqrt{\frac{\sum_i^n dX_i^2}{n}} \tag{10}$$

here, n is total number of potential matching pairs on images.

The form of weight function of (9) is similar to Gauss function, and the background of threshold is normal distribution. So the author calls filter (10) as Gaussian blunder filter. It can be applied to irregular points. Iteration gives better results. Weight to (10) case is assigned 0 during the iteration. Above processing works successfully where the surface is smooth, and wrong matches are not concentrated at certain places. However, there are many break-lines, and occluded regions in real image, and sometimes blunders are concentrated at certain places. Therefore, above formula cannot detect wrong matches in all situations. The problem areas are near the break-line of high building and high tree. Nevertheless, it reduces wrong matches.

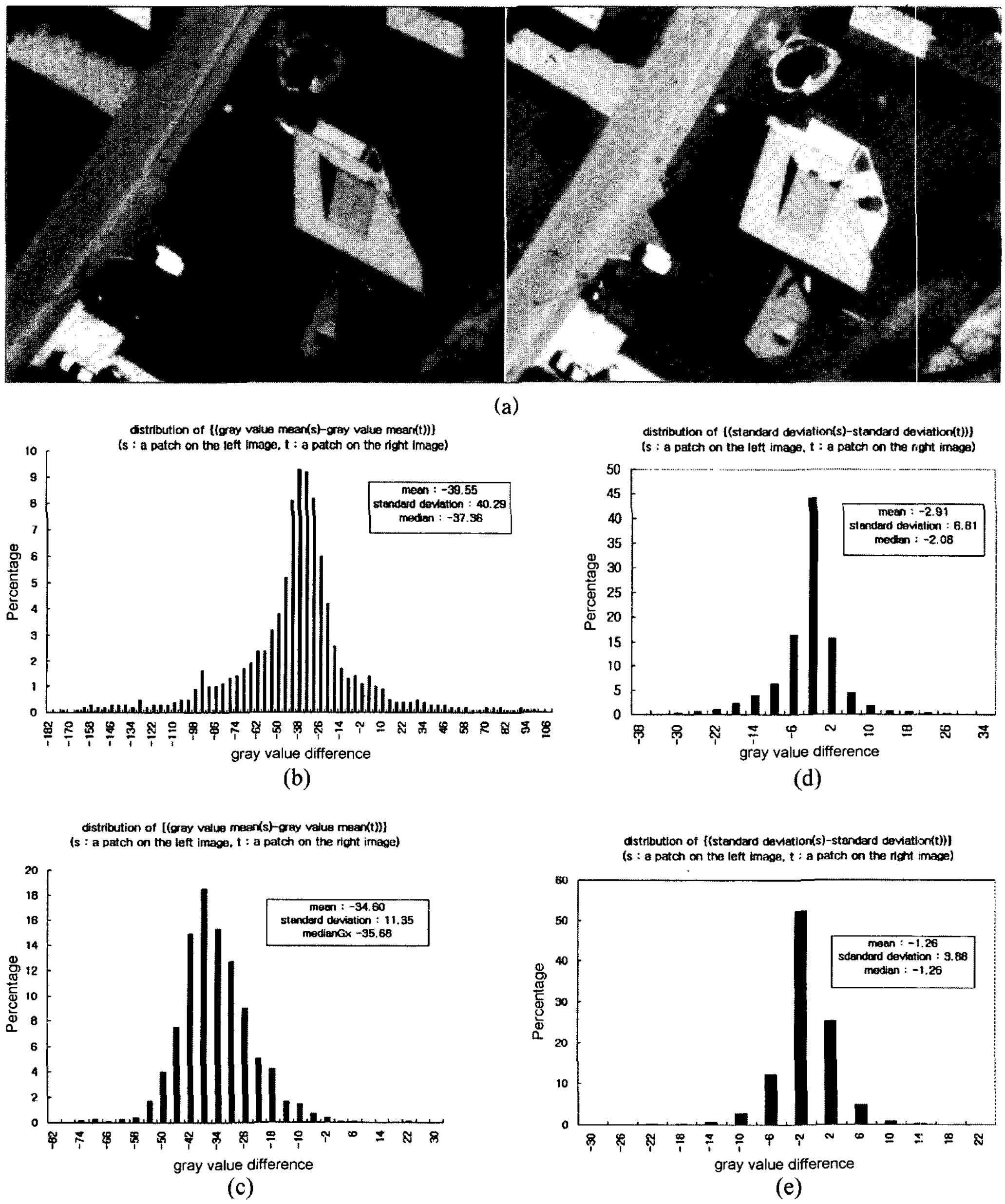


Fig. 1. Example 1 : Image Attribute Difference for not Corresponding Image Patch Pair Sand for Nearly Corresponding Image Patch Pairs.

(a) test image 1 (image size : 640×640), (b) The distribution of mean difference for non-corresponding image patch pairs (image patch size : 5x5), (c) The distribution of mean difference of for nearly corresponding image patch pairs, (d) The distribution of standard deviation difference for non-corresponding image patch pairs, (e) The distribution of standard deviation difference for nearly corresponding image patch pairs.

3. Experimental Results

Tests were conducted for several images, and two sample results are shown here. Fig. 1, Fig. 2 are examples to show the histogram of $\{\mu(s)-\mu(t)\}$, $\{\sigma(s)-\sigma(t)\}$ for non-corresponding pairs, and for nearly corresponding pairs. In these examples, $\sigma\{\mu(s)-\mu(t)\}$ and for nearly corresponding pairs is 1/3 of that for non-corresponding pairs, and $\sigma\{\sigma(s)-\sigma(t)\}$ for nearly corresponding pairs is 1/2 of that for non-corresponding pairs. In these examples, the nearly corresponding pairs are generated by the method described in 2.3.2 step 1, step 2, and step 3. The size of image patch is 5×5.

Fig. 3(a) and Fig. 4(a) are prior to the processing pro-

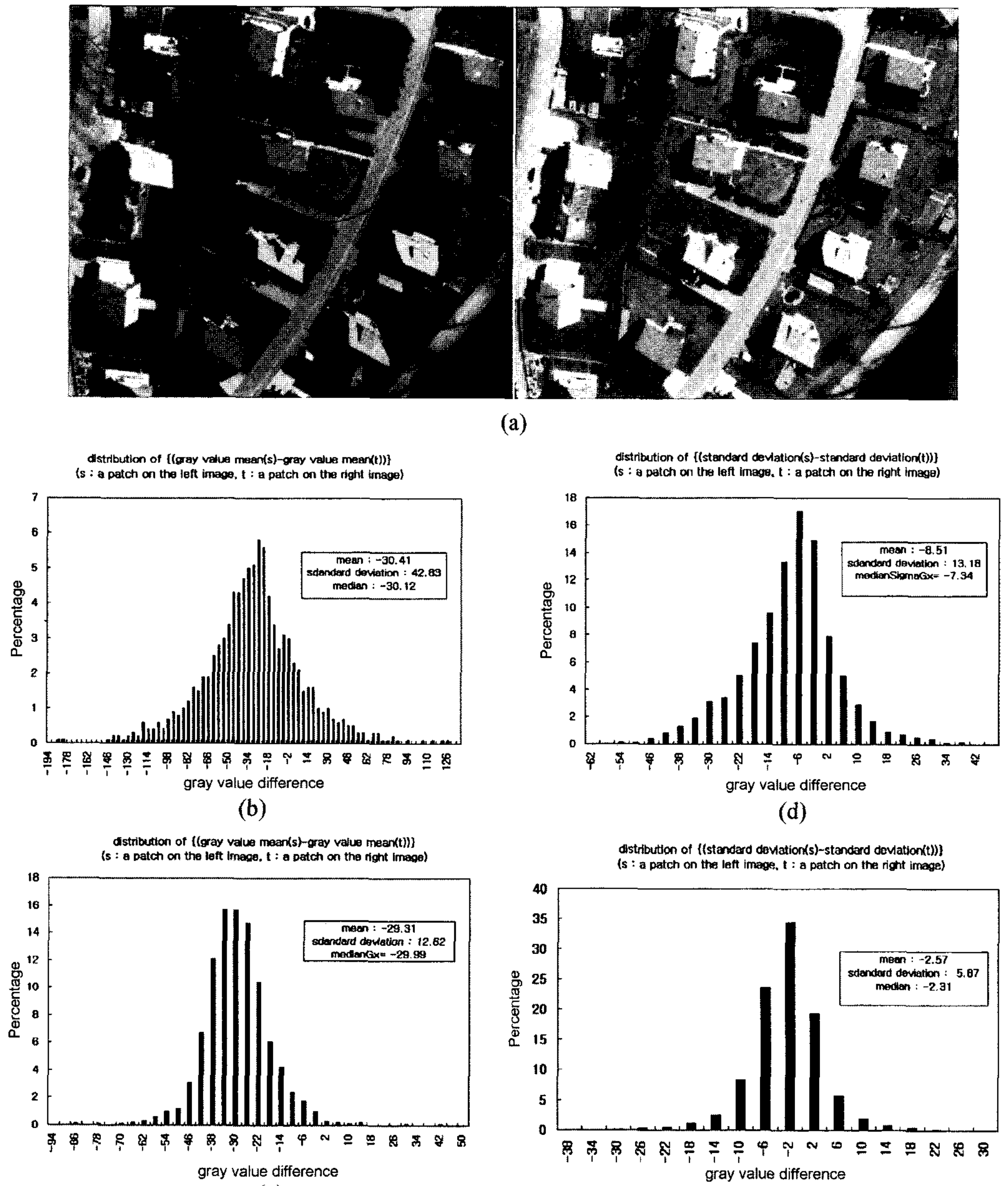


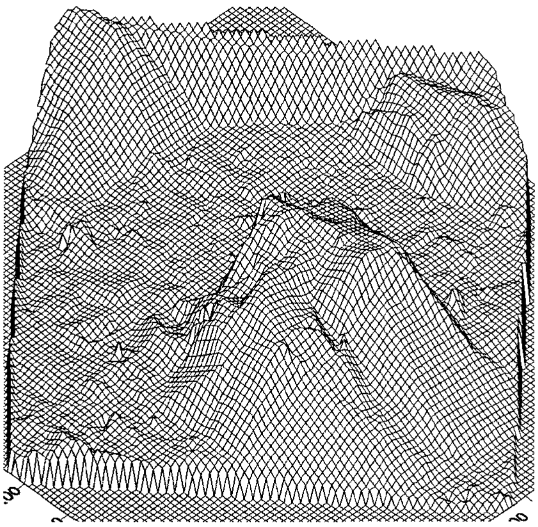
Fig. 2. Example 2 : Image Attribute Difference for not Corresponding Image Patch Pairs and for Nearly Corresponding Image Patch Pairs.

(a) test image 2 (image size : 1800×1800), (b) The histogram of mean difference for non-corresponding image patch pairs (image patch size : 5×5), (c) The histogram of mean difference for nearly corresponding image patch pairs, (d) The histogram of standard deviation difference for non-corresponding image patch pairs, (e) The distribution of standard deviation difference for nearly corresponding image patch pairs.

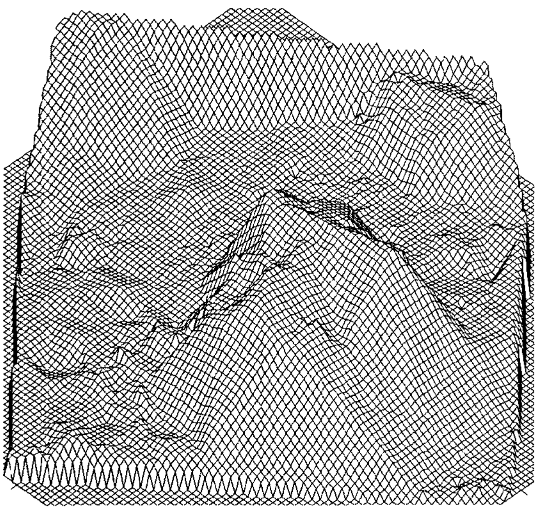
posed in this paper, and Fig. 3(b) and Fig. 4(b) are after the processing. Wrong matches are reduced. For these examples, the size of image patch for correlation is 5×5, and back-matching[10] is applied. In addition, 5 levels of an image pyramid for the image Fig. 1(a) (size 640×640), and 7 levels of an image pyramid for the image Fig. 2(a) (size 1800×1800) are used. Reduction ratio between each level of image pyramid is 1/2. Epipolar line equations of (1), (2), (3), (4) are used to generate epipolar images. Interest Points generated by Foerstner Operator[11] are used to match. $w > \text{mean}\{w\}$, $q \geq 0$, where $w = 2\text{detN}/\text{trN}$ and $q = 4\text{detN}/\text{tr}^2\text{N}$, and mask size of operator is 5×5. If the distribution of points is enough, and the points are a little far from the high building, it works well.

Near a high building, many points have two corresponding points, and relief displacements are big. Therefore, the image matching problem by correlation method happens. It is more difficult on a high tree. These cases are not easy to solve.

The comparison between matching with the grid points, and matching with interest points was conducted. The matching with interest points generally gives better results, but not always. It was also tested for the case of image patch size 7×7. The results are better than in the case of image patch size 5×5.

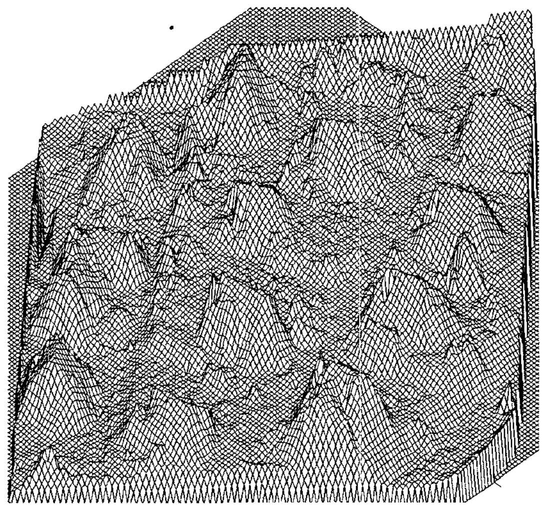


(a)

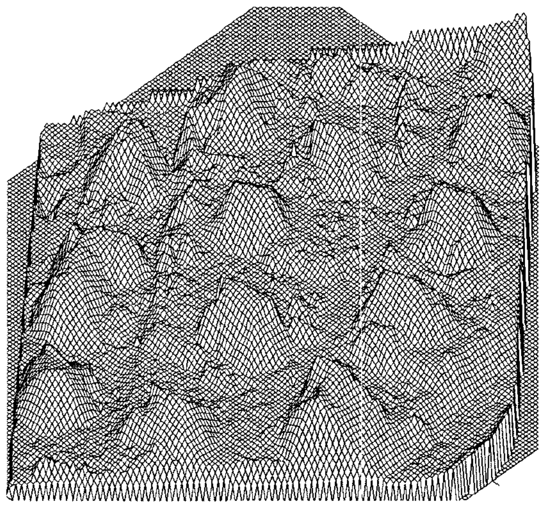


(b)

Fig. 3. The Comparison of DEM Results by Image Matching for the Image of Fig. 1(a)
(a) DEM result before proposed processing, (b) DEM result after proposed processing.



(a)



(b)

Fig. 4 The Comparison of DEM Results by Image Matching for the Image of Fig. 2(a).
(a) DEM result before proposed processing, (b) DEM result after proposed processing.

4. Conclusion

In this paper following two methods are proposed to reduce after the main part of image matching is over. Although these cannot remove every wrong matches in every situation, they are proved to work efficiently for real images.

1) Suppression of suspicious matches in the image attribute difference by histogram analysis of other's.

2) Suppression of suspicious matches in the parallax difference by Gaussian blunder filter.

These can be applied to most of point matching such as correlation method, least square method, interest point matching, relaxation method for stereo image and multi-photo images.

References

1. Schenk, T. (1999). "Digital Photogrammetry", Vol. I, Terra-Science, pp. 301-313.
2. Zhang, Z. (1994). "A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry", INRIA Rapport de recherche no. 2273.
3. Gruen, A. (1985). "Adaptive least squares correlation : a powerful image matching technique", South African Journal of Photogrammetry, Remote Sensing and Cartography Vol. 14, No. 3, pp. 175-187.
5. Haralick, R.M., and Shapiro, L.G. (1993). "Computer and Robot Vision", Addison Wesley, Vol. 2, pp. 342-349, pp. 379-426.
6. Zhang, Z., Zhang, J., Liao, M., and Zhang, M. (2000). "Automatic Registration of Multi-Source Imagery Based on Global Image Matching", Photogrammetric Engineering & Remote Sensing, Vol. 66. No. 5, pp. 625-629.
7. Vosselman, G. (1992). "Relational Matching", Springer-Verlag Berlin Heidelberg.
8. Baltsavias, E.P. (1991). "Multiphoto Geometrically Constrained Matching", PhD. Dissertation, Institute of Geodesy and Photogrammetry, ETH-Zurich.
9. Park, H., and Zimmermann, P. (2000). "Color Image Matching for DTM Generation and House Extraction", ISPRS Vol. XXXIII, Book III B, Amsterdam, pp. 137-1442.
10. Hannah, M.J. (1989). "A System for Digital Stereo Image Match", Photogrammetric Engineering & Remote Sensing, Vol. 55, No. 12, pp. 1765-1770.
11. Foerstner, W., and Guelch, E. (1987). "A fast operator for detection and precise location of distinct points, corners and centers of circular features", In Pro. ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data, 281-305, Interlaken.