A proposed search approach for image matching using epipolar constraint

Mohamed I. Zahran
Surveying and Photogrammetry, Faculty of Eng. at Shoubra, Zagazig University, Zagazig, Egypt

Using the concept of epipolar constraint in limiting the matching search space to one-dimensional space leads to more efficient matching process especially in the applications including a huge number of points. Resampling of stereo digital imagery into epipolar geometry can be applied after relative orientation so that the rows of pixels in both images lie along epipolar lines. However, original image content might be distorted due to the resampling process. In this paper, the concept of image epipolarity is employed for matching. But here the images are used directly without resampling them to epipolar geometry. In addition to saving epipolar resampling time, this would have the advantage of preserving the image radiometric characteristics unchanged. The proposed approach is based on finding the epipolar line in one image that corresponds to any specified point on the other image. The epipolar line is searched pixel by pixel, within a search range, for the conjugate point. Experimentation using a stereo pair of digital aerial imagery has shown that matching employing one-dimensional and two-dimensional search spaces yield slightly different results. The average difference in the y-direction is nearly half a pixel. However, one-dimensional matching results can be enhanced, by following a 3-row matching strategy.

استخدام مفهوم أو فكرة epipolar constraint في تقييد مجال البحث في عملية توافق الصور ليصبح البحث في بعد واحد يزيد من كفاءة عملية التوافق بالنسبة للتطبيقات التي تشتمل على عدد هائل من النقاط. تطبق هذا الفكرة بعد اتمام التوجيه النسبي بتحويل الصور الرقمية المتداخلة بحيث تصبح الصفوف المتناظرة في الصورتين على استقامة واحدة. في هذا البحث يستخدم هذا المفهوم في عملية التوافق ولكن بدون تحويل الصور الى وضع epipolar geometry مما يوفر زمن عملية التحويل ويحافظ على الخصائص الراديومترية للصور بدون تغير. الأسلوب المقترح يعتمد على ايجاد epipolar line في صورة والمناظر لأي نقطة يتم تحديدها في الصورة الأخرى. .يجرى بعد ذلك البحث على طول هذا الخط في مدى بحثي معين عن النقطة المناظرة للنقطة المحددة. أوضحت الأختبارات على زوج من الصور الجوية الرقمية الأختلاف البسيط (نصف pixel تقريبا) بين نتائج التوافق المتحصل عليها في حالتي البحث في بعد واحد والبحث في بعدين. ومع ذلك فانه يمكن تحسين نتائج الحالة الأولى باجراء التوافق على ثلاث صفوف.

Keywords: Epipolar geometry, Digital photogrammetry, Image matching, Reversed, Matching, Search space

1. Introduction

As generally recognized, digital image matching is a key problem in automating photogrammetric processes. Many researchers have dedicated to develop effective and robust matching techniques. Earlier, area-based matching was used. However, due to its limitations, it was advanced to feature-based matching. Many approaches that use either feature-based matching or a combination of area-based and feature-based matching methods, have been developed [4, 9]. Nevertheless, there are still problems that disallow image matching to be applied as a fully automated procedure. Some problems concerning area-based matching are the need of good approximations, optimal window sizes, effective search strategies, and reliability of matching results.

To effectively search for conjugates in the image to be matched, the search space can be constrained to a one-dimensional space by using of the epipolar geometry, thus leading to decreased computational cost and limited search space [6- 7]. This is critical in the applications including a huge number of points, such as at generating digital elevation models. Epipolar resampling of stereo images is performed after relative orientation so that the rows of pixels in both images lie along epipolar lines. It is a process in which original stereo images are normalized so that the conjugate points have no y-parallaxes [3, 10]. This permits the left and right images to be presented with only x-parallaxes unresolved. However, pixel intensity values of normalized images are obtained by interpolation from the neighboring pixels, which might degrade the content of the normalized images relative to the original image. Another procedure utilizing equations of epipolar lines in digital imagery is introduced in section 2.

In order to increase the reliability of matched points, the concept of reversed matching, described in section 3, can be applied. Here the normal matching process is repeated after reversing the template window and search window. Identical matching results yielded by both the normal and reversed matching can be accepted.

2. Matching using epipolar constraint

Given a stereo pair of images with known relative orientation, epipolar lines can be defined by making use of the coplanarity condition see fig. 1. The epipolar plane, for instance the plane O_1PO_2, is any plane containing the two exposure stations and an object point [8, 11]. Epipolar lines are the resulting lines of intersection of the epipolar plane with the left and right image plane. Given the location of point p_1 on the left image, its corresponding point p_2 on the right image is located along the relevant epipolar line. Based only on the location of image point p_1, object point P can be located anywhere along line O_1P, at an arbitrary height. By using the collinearity condition equations, the corresponding location p_2' on the right image for an assumed object point position P', can be calculated.

By comparing small window centered at the location of point p_2' with a corresponding window centered at the location of point p_1, they would not match. This is because the two locations do not correspond to images of object point P. Thus, another point on the right epipolar line would be tried. The search continues along the epipolar line until reaching a match, which occurs at the corresponding image point p_2. Accordingly, the effort of finding conjugate (or matched) points

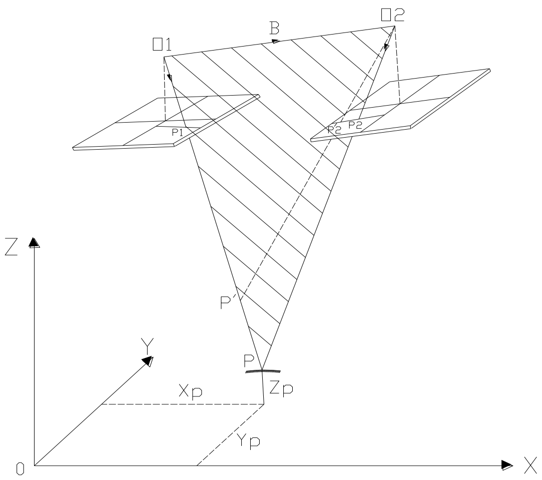

Fig. 1. Epipolar geometry of a stereo pair of images.

is restricted to one-dimensional (1-D) search along the epipolar line, rather than a much more time-consuming two-dimensional (2-D) search throughout windows of considerable sizes.

In this paper, the concept of epipolar geometry is employed for matching. However, the images to be matched are used directly without resampling them to epipolar geometry. Therefore, the actual intensity values of image pixels are kept unchanged. In this approach, for any point $p(x_p,y_p)$ on the left image, its corresponding point $p'(x'_p,y'_p)$ on the right image is located along the epipolar line that can be represented by the following eq. [5]:

$$a_o' x' + a_1' y' + a_2' (-pd) = 0, \quad (1)$$

where

$$\begin{bmatrix} a_o' \\ a_1' \\ a_2' \end{bmatrix} = (\mathbf{R}_2)^T \mathbf{B} \mathbf{R}_1 \begin{bmatrix} x_p \\ y_p \\ -pd \end{bmatrix}, \quad (2)$$

with:

$\mathbf{R}_1$ is the rotation matrix from the left image to the object coordinate system,

B is the 3×3 matrix that is expressed as $\mathbf{B} = \begin{bmatrix} 0 & -bz & by \\ bz & 0 & -bx \\ -by & bx & 0 \end{bmatrix}$ in which bx, by and bz are the components of the vector from the perspective center of the left

image to the perspective center of the right image,
R$_2$ is the rotation matrix from the right image to the object coordinate system,
x_p, y_p are the coordinates of point p in the left image coordinate system, and
pd is the camera's principal distance.

3. Reversed matching

The basic concept of this approach is that the matching function should be one to one. Therefore a successful match from left to right image, if repeated from right to left, should end on the initial pixel from left image. This additional check that is made by reversing the template window and search window would increase the reliability of matched points. This approach is also called double image checking [1]. If the result for each matched point pair is the same for both normal case and reversed case, it is accepted. Matching results are considered similar in both matching cases when having equivalent coordinates of corresponding points as well as comparable estimates of correlation coefficient. Adopting the concept of epipolar geometry described in section 3, eq. 1 can be employed to find the conjugate point on the right image for a point on the left image, which can be regarded as the normal case of matching. In the reversed case, eq. 3 can be utilized to locate the conjugate point $p(x_p,y_p)$ on the left image for the point $p'(x'_p,y'_p)$ determined on the right image by normal case. It is the equation of corresponding epipolar line on the left image, passing with point p.

$$a_o\, x + a_1\, y + a_2\, (-pd) = 0, \tag{3}$$

where

$$\begin{bmatrix} a_o \\ a_1 \\ a_2 \end{bmatrix} = (\mathbf{R}_1)^T\, \mathbf{B}\, \mathbf{R}_2 \begin{bmatrix} x'_p \\ y'_p \\ -pd \end{bmatrix}, \tag{4}$$

with R_1, B, R_2 and pd as defined previously.

4. Proposed search method

This section presents a 1-D search approach that is integrated with correlation matching to match digital images. In this approach, pixel coordinates of any point to be matched in the left image are measured and then transformed into the image coordinate system centered at the principal point. The transformation parameters are computed by utilizing calibrated and measured coordinates of left image fiducial marks in an affine transformation. To find the relevant epipolar line in the other image, eq. 2 is solved for a_o', a_1' and a_2' utilizing known orientation parameters for both images to be matched. The epipolar line is then searched pixel by pixel, within a search range, for the conjugate point. However, this requires that image point coordinates in the epipolar line equation to be expressed in terms of pixel coordinates (column and row coordinates). Using eq.1, the y-coordinate (y') for a point lying on the epipolar line and has an x-coordinate (x') can be found as follows:

$$y' = -[a_o'\, x' + a_2'\, (-pd)] \;/\; a_1'. \tag{5}$$

In order to express eq. 5 in terms of pixel coordinates xc' (column) and yr' (row) that correspond to the image coordinates x' and y', an affine transformation is performed using calibrated and measured coordinates of right image fiducial marks. This would yield six transformation parameters (say a', b', c', d', e', f') that can be used subsequently as follows:

$$\begin{bmatrix} xc' \\ yr' \end{bmatrix} = \begin{bmatrix} a' & b' \\ d' & e' \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} c' \\ f' \end{bmatrix}. \tag{6}$$

The inverse transformation can be expressed as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{1}{a'\, e' - b'\, d'} \begin{bmatrix} e' & -b' \\ -d' & a' \end{bmatrix} \begin{bmatrix} xc' & -c' \\ yr' & -f' \end{bmatrix}. \tag{7}$$

By substituting with x' and y' from eq. 7 into eq. 5, the following equation would be obtained:

$$yr' = f' + [(xc' - c')(a_o' e' - a_1' d') + (a_2' (-pd))(a' e' - b' d')] / (a_o' b' - a_1' a'). \quad (8)$$

This equation determines the row coordinate (yr') for any given point on the right image that lies on the epipolar line and has a column coordinate (xc').

Correlation matching is then employed to find the location of conjugate point in the right image, for any chosen point in the left image. The idea of that technique is to measure the similarity of the reference window in the left image with each of the matching windows within the search window in the right image using the cross-correlation coefficient. Here, the search space is limited to the epipolar line leading to 1-D matching. The reference window is centered at the selected point in the left image whereas the search window is centered at an approximate location of its conjugate in the right image. The size and location of the search window are determined using preliminary knowledge about the photography and the terrain. The cross correlation coefficient is computed for every position of the matching window within the search window. The sub-pixel position yielding maximum correlation coefficient is determined by fitting a one-dimensional polynomial through the correlation values.

In reversed matching, for the point determined on the right image in the normal matching case, the conjugate point on the left image is found hoping that it would be the initial point selected on that image. To find the relevant epipolar line in the left image, eq. 4 is solved for a_o, a_1 and a_2 utilizing known exterior orientation parameters for the image. The epipolar line is then searched pixel by pixel, within a search range, for the conjugate point. Again, the image point coordinates in the epipolar line equation are to be expressed in terms of pixel coordinates (column and row coordinates). Using eq. 3, the y-coordinate (y) for a point lying on the epipolar line and has an x-coordinate (x) can be found as follows:

$$y = - [a_o\, x + a_2\, (-pd)] / a_1. \quad (9)$$

To express eq. 9 in terms of pixel coordinates xc (column) and yr (row) that correspond to the image coordinates x and y, the computed affine transformation parameters (say a, b, c, d, e, f) of the left image can be used as follows:

$$\begin{bmatrix} xc \\ yr \end{bmatrix} = \begin{bmatrix} a & b \\ d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c \\ f \end{bmatrix}. \quad (10)$$

The inverse transformation can be expressed as:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{a\,e - b\,d} \begin{bmatrix} e & -b \\ -d & a \end{bmatrix} \begin{bmatrix} xc & -c \\ yr & -f \end{bmatrix}. \quad (11)$$

By substituting with x and y from eq. 11 into eq. 9, the following equation is resulted:

$$yr = f + [(xc - c)(a_o\, e - a_1\, d) + (a_2\, (-pd))(\, a\, e - b\, d)] / (a_o\, b - a_1\, a). \quad (12)$$

The equation provides the row coordinate (yr) for any given point on the left image that lies on the epipolar line and has a column coordinate (xc).

The entire procedure, comprising normal and reversed matching cases, is implemented on a PC by using MATLAB software.

5. Experimentation and results

The test imagery consists of a stereo-pair of scanned aerial images, covering an urban area, with a 42-µm-pixel size [2]. The images are captured with a camera with a focal length of 153.380 mm and a standard format. The flying height is nearly 800 m above terrain yielding a photography scale of nearly 1:5200. The two images are shown in figs. 2 and 3, respectively. The pixel coordinates of thirty fairly distinct points that appear in the overlap area of the two images are measured in the left image and transformed to related image coordinate system. The first and second columns of table 1 list the measured pixel coordinates.

The conjugate locations of the selected thirty points are found automatically, to sub-pixel accuracy, in the right image using a prototype cross correlation program. The input data to the program are the pixel coordinates of selected points in the left image and of their coarsely- determined conjugates in the right

Fig. 2. The left image of the test image pair.

Fig. 3. The right image of the test image pair.

Table 1
Pixel coordinates of selected points in the left image and their conjugates in the right image, found by correlation matching

Pt.	Left image coordinates		Right image coordinates using 2-D search space		Right image coordinates using 1-D search space		Right image coordinates using 3-row search space	
	xc	*yr*	*xc*	*yr*	*xc*	*yr*	*xc*	*yr*
1	3659.75	272.39	1410.45	259.29	1410.55	259.29	1410.45	259.29
2	3824.63	125.74	1573.43	116.44	1573.33	115.64	1573.43	116.24
3	2851.41	400.98	612.11	384.98	612.21	385.48	612.11	385.08
4	3089.72	304.54	846.72	290.64	846.82	290.54	846.72	290.64
5	3344.27	583.18	1104.27	563.88	1104.07	564.18	1104.27	563.88
6	2650.49	1192.81	435.80	1165.11	435.50	1165.81	435.70	1165.21
7	2782.18	910.05	558.58	886.25	558.48	886.85	558.48	886.45
8	3609.02	873.30	1377.72	850.00	1377.52	850.30	1377.72	850.00
9	4222.51	1013.15	1988.51	988.15	1989.00	986.97	1988.31	988.05
10	4509.35	1015.19	2270.85	989.79	2270.25	988.99	2270.75	989.69
11	4606.33	1871.63	2389.73	1836.43	2389.23	1835.83	2389.13	1836.33
12	3847.88	1641.95	1632.68	1608.95	1632.78	1608.85	1632.68	1608.95
13	3239.49	1332.65	1021.89	1303.26	1021.89	1303.06	1021.99	1303.16
14	2918.96	1602.14	710.06	1569.14	709.76	1569.84	709.96	1569.44
15	2700.99	2290.23	498.99	2253.03	498.29	2254.13	498.69	2253.43
16	4216.39	1382.67	1989.69	1352.47	1989.79	1351.57	1989.69	1352.27
17	3940.77	2028.83	1730.87	1993.43	1730.67	1992.93	1730.97	1993.33
18	3211.66	2219.71	1006.56	2183.31	1006.46	2182.91	1006.56	2183.31
19	2763.29	2305.51	561.29	2268.21	561.39	2268.41	561.29	2268.21
20	3618.91	2790.91	1419.01	2752.71	1419.21	2752.01	1419.21	2752.51
21	4670.80	2471.22	2464.20	2433.52	2464.00	2432.33	2463.90	2433.12
22	3252.80	3019.50	1055.30	2981.80	1055.70	2980.80	1055.50	2981.30
23	4472.76	2929.00	2271.66	2890.90	2271.66	2890.80	2271.66	2890.90
24	2494.14	3414.99	298.44	3379.19	298.64	3379.19	298.44	3379.19
25	3541.92	3356.23	1347.32	3320.53	1347.72	3319.63	1347.42	3320.13
26	4574.58	3859.46	2385.78	3826.06	2386.18	3825.66	2385.78	3826.06
27	4843.21	3787.08	2649.12	3752.98	2649.22	3752.58	2648.92	3752.28
28	2798.60	3869.16	607.50	3836.06	607.40	3836.06	607.50	3836.06
29	2924.71	4002.98	735.31	3971.08	735.41	3970.98	735.31	3971.08
30	2892.62	4242.27	705.02	4213.47	705.12	4212.97	705.02	4213.47

image. The coarse locations of conjugate points are determined using the magnitude of the photo base. The size of the reference window is specified as 11 pixels by 11 pixels. A matching threshold of 0.7 is selected. Considering those criteria, three matching cases are applied and their results are evaluated. In the first case, a 2-D search space is used. Here, the search window size is taken as 101 pixels by 101 pixels, based on a priori information on the elevation range of the object space. The location of each conjugate point is found, to sub-pixel accuracy, by fitting a two-dimensional polynomial to the nine pixels centered at the position with the highest correlation, and searching for the maximum. The matching results attained in this case are checked against any mismatches in order to have a reliable base with which the other two cases will be compared. Those results are listed in the third and fourth columns of table 1.

In the second case, a 1-D search window of 101 pixel length along the epipolar line is used. Thus, the window has an 11-pixel width that is equal to the size of the reference window. The location of each conjugate point is found by fitting a one-dimensional polynomial to the row pixels centered at the position with the highest correlation. The fifth and sixth columns of table 1 show the related matching results. In the third case, the 1-D search space is expanded one row above and one row below the epipolar line forming a 3-row search space. Here, the search window has a 13-pixel width. Again, the location of each conjugate point is found by fitting a two-dimensional polynomial to the nine pixels centered at the position with the highest correlation. The matching results are exhibited in the seventh and eighth columns of table 1. Table 2 reports the values of average (Ave) and Root Mean Square (RMS) of the computed differences among results of cases 1 and 2 and of cases 1 and 3, respectively.

Comparing the figures of tables 1 and 2, it is notable that the three cases yield slightly different results. The average as well as root mean square values of computed differences are less than a pixel. The figures of case 3 are closer to their corresponding figures of case 1, than of case 2, especially regarding xc coordinates. The justification of having better results in case 3, than in case 2, can be related to the errors in epipolar line computation due to the errors associated with the computation of exterior orientation elements of the stereo pair.

Table 2
Average and root mean square of differences among pixel coordinates resulted by applied matching cases

Statistic	Case1 versus case 2	Case1 versus case 3
Ave (*xc*)	0.22	0.09
Ave (*yr*)	0.51	0.14
RMS (*xc*)	0.28	0.16
RMS (*yr*)	0.62	0.23

It is worth mentioning that the number of operations, in each of which the reference window is moved within the search window and the coefficient of cross correlation is estimated successively, is considerably larger in the 2-D case than in the 1-D case. Assuming that the number of operations in the 1-D case is equal to n, the corresponding number of operations in the 2-D case would be n^2. Since the number n in the setup of this research is 101-11+1 = 91, the number of operations in the 2-D case is 91^2 = 8281. Having a number m of image points to be matched, the total number of operations in the 2-D case goes up to $m.n^2$.

At trying larger search window sizes while applying both 2-D and 1-D matching cases on some other points that are not fairly distinct in the left image, some mismatches have been resulted in the 2-D case. Those mismatches have occurred due to the existence of other points that resemble selected points, within the search window. On the contrary, no mismatches are detected in the 1-D case. In fact, the chances of having repetitive patterns along the same epipolar line are almost improbable. Thus, enlarging search spaces have almost no influence on 1-D matching results.

It has been found also that the 2-D reversed matching of the selected points has similar performance to the 2-D normal matching; slower matching process as well as higher possibility of having mismatches than the 1-D case.

6. Conclusions

In this research a one-dimensional (1-D) search approach is proposed and integrated with correlation matching to match stereo digital imagery. The approach is based on finding the epipolar line in one image that corresponds to a point on the other image. The matching process is carried out subsequently along the epipolar line. Here the images are used for matching directly without resampling them to epipolar geometry, thus keeping the actual radiometric characteristics of the imagery. In accordance with the acquired results, the following conclusions can be formulated:

- Matching results yielded at utilizing 2-D and 1-D search spaces differ within a fraction of a pixel. Size of difference is actually related to the precision of epipolar line computation.
- 1-D matching results can be enhanced, particularly in the y direction by following the 3-row matching strategy.
- Correlation matching utilizing 1-D search space considerably reduces the chances of getting mismatches especially with the existence of repetitive patterns or linear features.
- In contrast to its unfavorable influence regarding 2-D matching, enlarging the size of search space has almost no influence on the performance of 1-D matching.
- Reversed matching with 1-D search space is much simpler than with 2-D search space.

References

[1] D.P. Skarlatos, International Archives of Photogrammetry and Remote Sensing, Vol. 33 (B3), p. 845 (2000).

[2] DVP Geomatic Systems, DVP Software., DVP Geomatic Systems Inc. (1997).

[3] E.M. Mikail, J.S. Bethel, and J.C. McGlone, Introduction to Modern Photogrammetry, John Wiley and Sons, Inc., New York (2001).

[4] F. Samadzadegan et al. Photogrammetric Record, Vol. 17 (100), p. 651 (2002).

[5] H. Park, and P. Zinmmermann, International Archives of Photogrammetry and Remote Sensing, Vol. 33 (B3), p. 697 (2000).

[6] M. Morgan et al., International Archives of Photogrammetry and Remote Sensing, Vol. 35 (B3) (2004).

[7] P.M. Dare, International Archives of Photogrammetry and Remote Sensing, Vol. 35 (B3) (2004).

[8] P.R. Wolf and B.A. Dewitt, Elements of Photogrammetry, McGraw Hill, Inc., New York (2000).

[9] T. Schenk, Digital Photogrammetry, Laurelville, OH: TerraScience (1999).

[10] T. Schenk, Technical Notes Photogrammetry, No. 4, Department of Geodetic Science and Surveying, The Ohio State University, Columbus, OH. (1990).

[11] W. Cho, T. Schenk, and M. Madani, International Archives of Photogrammetry and Remote Sensing, Vol. 29 (B3), p. 404 (1992).

Received March 28, 2005
Accepted May 23, 2006