A NEW STEREO MATCHING APPROACH IN IMAGE/OBJECT DUAL SPACES

Yaonan Zhang

Centre for Computer Graphics and Mapping
Thijsseweg 11, Delft University of Technology
2629 JA Delft, The Netherlands

Commission V

ABSTRACT:

The stereo matching (or correspondence) remains one of permanent problems in Photogrammetry and Computer Vision. This paper presents a new approach to solve the problem, which incorporates the image space based matching techniques with the high level knowledge about the objects. The low-level processing (edge detection, feature extraction) and candidate matching are carried out in image space, while the final matching is determined in object space as solving a consistent labelling problem which results from the integration of candidate matching, high level constraints of objects and other constraints of image matching. One of the innovative features in our approach lies in back-projecting (back tracing) the line pairs from candidate matching into the object (scene) space, and combining all the constraints in a unified process. We substitute the concept of "figure continuity" usually used in the image matching with the high level knowledge from the object space.

KEYWORDS: Image processing, Image Matching, Line detection.

1. INTRODUCTION

The research on matching has been taken in computer vision and photogrammetry society for quite a long time. According to the space where the matching takes place, the existing techniques for solving the matching problem roughly fall into two categories: image space based and object space based. In the image space based matching, the primitives of one image are compared with ones on the another image. Many solutions to the matching have been proposed in the image space. The methods vary with different choice of primitives: area-based (intensity-based), feature-based and structure-based (relational matching). Recently, several articles are devoted to the object space based matching. This method emerged originally from the task of reconstructing digital terrain model from a pair of digital images, independently developed by Wrobel[1] and Helava[2], etc. Helava used the concept of "groundel" as a unit in object space similar to the "pixel" in the image space. The image intensities corresponding to each groundel can be analytically computed, if all pertinent geometric and radiometric parameters (including groundel reflectance, etc.) are known. A least square method is adopted to determine a set of unknown quantities or improvements to their approximate values used in the analytical prediction process.

Although the progress is undoubtedly made, most of the algorithms are still task and domain dependent, many parts of the problem still need full exploration of our human intelligence. In solving this difficult problem, we propose a novel approach which unify the techniques in image and object space, combining the geometric knowledge of scenes. The motivation behind this is that a general solution for ill-posed problem such as matching is to use additional constraints or knowledge to restrict possible solution[3]. The existing constraints used in the matching are the uniqueness, smoothness, ordering, figure continuity and camera geometry (epipolar geometry), etc. In this paper, we introduce a new concept of "general geometric constraints of scenes". In the problem of reconstructing digital terrain model from digital image, the smooth constraints of surface can be adequately assumed[1,2,4,5]. But in the most application of computer vision, industrial robot vision and automated close-range photogrammetric system, the scenes are full of lines, shapes and structures, it is impossible to unify all the information in the object space in a straight-forward way. The reason is that the original digital image consists of only raster pixels which can not directly provide much structural information required heavily by later analysis. Without low-level processing in the image space, it is practically impossible to get more structural description of the images. In our approach, we back-project (back trace) the image primitives (from low-level processing) into the object space, and carry the matching in the object space combining the available knowledge from scene with other constraints. We implement the line-based matching in the image space in order to find the candidate line pairs and then project the these line pairs into the object space. The final matching became a consistent labelling or constrained satisfaction problem. We use a relaxation procedure to get the final lines in object space.

The main idea described in this paper has been previously reported by author in 1991[25]. This paper includes the description on line detection and grouping, back projection for horizontal image lines, as well as more experimental results.

2. LINE DETECTION AND GROUPING

Despite the large amount of research, effective extraction of straight lines has remained a difficult problem. Usually a local operator is used to detect the local discontinuities or rapid changes in some image features, followed by aggregating procedure to link the local edges into more global structures. These methods include Hough transforms[16,17,18], edge tracking and contour following[18], curving fitting[26], etc. In our research, the following methods have been implemented.

Edge detection and filtering

Sobel operator is employed to detect local edges. In most of practical situations, the image data are noisy and, since edge are high spatial-frequency events, edge detections enhance the noise. In order to get reliable global information for later processing, a optimization method developed by Duncan[24] has been implemented, which is aimed at providing a bridge between local, low-level edge and line detection and higher-level object boundary using a new form of continuous labelling.

Extract straight lines from edge direction

Based on the observation by Burns[19], edge gradient orientation can serve as a very good base to extract line-support region. In our research, we have used Duncan's technique to bridge the edge orientation gap caused by noise and possible irregularity of the object boundary. Following this, there are four steps in extracting straight lines: 1). segment and label pixels into line-support regions based on similarity of gradient orientation. 2). use least square method to accurately allocate the straight line position within each line-support region. 3). verify the straight lines by comparing the difference between the allocated line and the contour which has average intensity grey value, passing through the line-support region. 4). calculating attribute for each line, e.g. contrast, length, orientation, etc.

Perceptual grouping

linear structures are usually broken due to a variety of factors, such as markings, noise in the image, and inadequacies in the low-level processes. Additionally, some of the breaks are due to real structures in the image. Because Duncan filter can only bridge the gap within one or two pixels, additional perceptual grouping in vector form is required. Grouping of straight lines has been the subject of much investigation. The reader is referred to some references[21,22,23]. Because we only use straight lines in our current implementation for matching, only collinearity is considered in grouping image line segments. A more precise way to implement these decision would be to use 3D information if the two line segments are on the same plane, which is in turn based on the matching result.

3. CANDIDATE MATCHING IN IMAGE SPACE

Feature-based matching is very common technique in the image space[6,7,8,9]. The commonly used features have been edges detected in the image. However, edges as primitives may be too local. In our approach, we match straight lines, which consists of connected edges, and hence the inter-scanline connectivity is implicitly in the matching process.

Constraints of matching

Marr[10] and Poggio have suggested use of the following two constraints:

1). Uniqueness. Each point in an image may be assigned at most one disparity value.
2). Continuity. Matter is cohesive, therefore disparity values change smoothly, except at a few depth discontinuities.

In our image/object dual space matching approach, we unify the uniqueness constraint in the relaxation procedure, but substitute the continuity to the general geometric constraints of scenes.

Baker[6] also proposed another "ordering" constraints for un-transparent objects which is valid in the most of cases. We keep this constraints in the image space because this is more easier to implement than in the object space.

Matching attributes

After the low-level processing, the line segments are described by

-- coordinates of the end points
-- orientation
-- strength(average contrast)

Matching criteria:

On this aspect, we use some of criteria developed by Medioni and Navatia[9].

- overlapping: detail of explanation is referred to Medioni. Actually, this is the another version of epipolar geometric constraints. We calculate the corresponding threshold by following formula:

$$t_1 = \alpha W + \delta y \qquad (1)$$

where

α is the estimation of error in the angle parameters of camera geometry, it should be noted that there are totally three angel parameters to describe the orientation of a camera. Here α is a overall estimation of these three angle errors.

W is the width of the photograph.

δy is the estimation of error of y-direction shift between two images.

- disparity limit: disparity is the displacement along the epipolar line of the corresponding points. The object space always has limited depth, this means, the disparity changing on the images also has limited range. The introduction of disparity limit can significantly reduces the search range of matching, therefore reduce the processing time. The efficiency of candidate matching depends on two aspects: guaranteeing that the all the possible matching pairs of lines are in the result of candidate matching, and performing the task as quick as possible. Our experiments have turned out that disparity limit is a very important parameter for the success of final matching.
- orientation: the orientation difference of two corresponding lines is not only due to the camera geometry, but also to the original direction of object line. In our algorithm, we set the maximal allowable orientation difference between the two corresponding lines in order to eliminate wrong line pairs.

In this step, one line on the image can be assigned to more than one corresponding lines on the other image.

4. BACK-PROJECTION

After the candidate matching is finished, the image line pairs are back-projected into the object space (or scene space). Two cases must be considered, that is, on whether the image lines are horizontal or not, with respect to stereo camera base line. For the non-horizontal lines, the back project principle is described in Fig.1. In the figure 1, let O-XYZ be the object space coordinate system fixed on the ground (scene), S'-x'y'z' and S"-x"y"z" be the left and right camera coordinate system with the origin at the focus point of the camera and z',z" coinciding with the optical axis (z' and z" are not shown in the figure). The (a'b') and (a"b") are the candidate image line pair respectively on the left and right images. W' is the plane formed by left camera focus point and image line (a'b'), similarly with W". For image line pair (a'b') and (a"b"), We define the corresponding object line as L. Very often, the result from line detection is not perfect, so we can not assume the detected lines from image are complete. Also due to the limited resolution of image and the error in the parameters of camera geometry, the two ray lines from the corresponding image points can not be assumed to mathematically intersect in the object space. To avoid such drawback, we use the intersection of two planes W' and W" to get the equation of object line corresponding to image lines (a'b') and (a"b").

In the practice, we calculate the coordinates of point A', A", B', B", put the central point of A' and A" as the starting point of line L, the central point of point B' and B" as the ending point of line L.

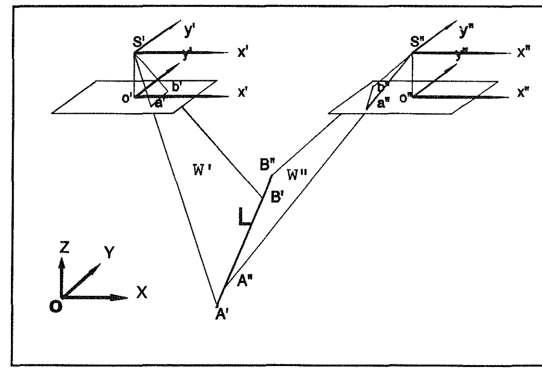

Fig.1 Back project for non-horizontal lines

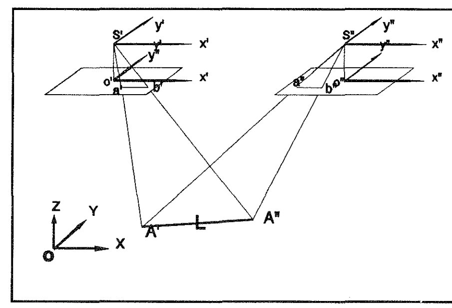

Fig.2 Back projection for horizontal lines

When the image lines are horizontal, the above method would obviously fails. Another method should be used, as illustrated in Fig.2., where A' is chosen as the most closed point to both ray (s'a') and (s"a"), similarly with A".

5. GENERAL GEOMETRIC CONSTRAINTS OF SCENES

The concept of geometric constraints has been used by a number of researchers[11,12]. Their ideas mainly concern the geometric constraints on camera geometry, actually the extension of epipolar line constraints. This kind of conditions have been taken care by our algorithm in the image space. What we propose in this article about GGCS is totally different, which is related to the property of objects in the scene.

An explanation of general geometric constraints of scenes (GGCS)

The idea of utilizing GGCS is to derive as much as possible geometric constraints to globally improve the matching quality. The primitives used in the matching (also in the image analysis) are the intensity, point, straight line, curve, shape, etc. The GGCS concerns the geometric constraints on these primitives themselves as well as the constraints on the relationship between the primitives. So the constraints can be unary, binary, and N-nary. All the these constraints are depicted in object space.

GGCS for lines

Unary constraints :
orientation of line;
length of line;
etc.

Binary constraints :

1). Collinearity.
Collinearity means that two lines have same mathematical equation. It is easy to see that the starting points and ending point for two lines are not necessarily to be the same for the collinearity.

2). Overlapping.
Two lines are said to be overlapping if they are collinear and have common range.

3). Connectivity.
Two lines are connected if lines are mathematically intersected and the intersecting point is within the starting point and ending point for both lines.

4). Coplanarity.
Two lines are coplanarity if they are on the same plane.

5). Parallel.

The general attributes to describe the relationship of two lines can be summarized as following:

distance between a pair of lines (shortest distance)
distance between the endpoints of two line segments

(including shortest distance between the endpoints of two line
segments as well as longest distance)
distance between the midpoints of two line segments
distance from a line to the origin, and
distance from a line to an endpoints of a line segment.
overlapping range
intersecting angle
coordinates of intersecting point
difference of lengths

What type of GGCS should be used is task and scene dependent. The generation of suitable GGCS for specific task, specific scenes or objects will initialize other research issues, that is, the generic modelling, learning mechanism, and related man-machine interface, etc.

6. FINAL MATCHING IN OBJECT SPACE

Because in the image space, it is allowed that one line can be matched with more than one lines on the other image, we need to implement the uniqueness constraints together with GGCS and other constraints in object space.

Such problem is a typical consistent labelling or constrained satisfaction problem which was formulated by Haralick and Shapiro[14,15]. In our current algorithm, we use the relaxation techniques which is frequently used in the computer vision community.

We assign each candidate object line (i) a label LB(i) which ranges from 0 to 1.0 . After the relaxation, the line with label "1" means a true scene line, while a "0" indicate a false line. The initial label value for each line is bound in the image space matching (e.g. from similarity measurement of contrast). The label value is then updated in the iteration by the following formula.

$$LB(i)^{k+1} = (1.0 + \delta_1)\, LB(i)^k + \delta_2 a - \beta_1(b_1 + b_2) - \beta_2 c \qquad (2)$$

where
$LB(i)^{k+1}$ is updated label value
$LB(i)^k$ is old label value
k is the number of iteration
δ_1 is the coefficient for increment from old label
δ_2 is the coefficient for increment from GGCS
β_1 is the coefficient for decrease from uniqueness constraints.
β_2 is the coefficient for decrease from ordering constraints.
a is the measure for the GGCS
b_1, b_2 is the measure for uniqueness constraint
c is the measure for ordering constraint

7. EXPERIMENTAL RESULTS

Experiment on simulated data

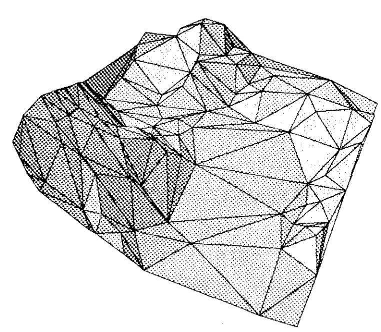

Fig.3 The original TIN data with visualization

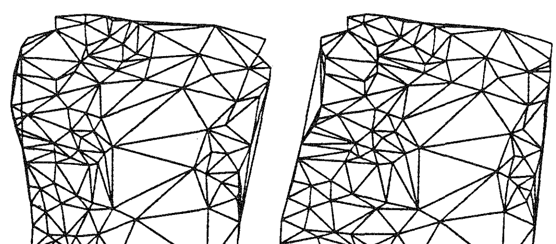

Fig.4 Simulated left and right images (lines)

The data used for simulation is a TIN-based three dimensional lines which is shown in Fig.3 , in which we have used shading to create a depth impression. We first calculate the corresponding stereo image pair by simulating the camera geometry. The simulated image pair is shown in Fig.4.

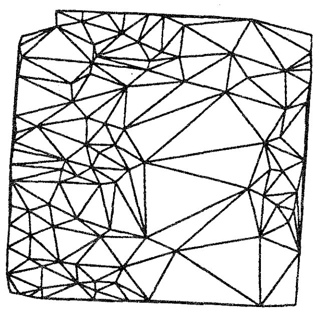

Fig.5 The final matching result (top view of 3-D lines)

In the image space, as described before, we use overlapping, line orientation, disparity limit to find the candidate matching. We have 292 lines from each image, the result of candidate matching contains 3956 candidate line pairs. In this experiment, we use the "connectivity" as our GGCS. The label value for a line is enhanced if it connected to the other line at the endpoints. The final matching result is drown in Fig.5 (here we only give the top view of the three-dimensional data). The final result has 284 lines. The loss of some lines which are in the original TIN data is the result of using ordering constraint.

Experiment on real images

A stereo images of a stereo plotter (shown in Fig.6, on the last page of this paper) has been chosen for the test. The images were taken with a normal camera (focus length 28mm and size 36mm X 24mm), followed by the scanning with 100 DPI on positive pictures (about 3 times larger than negative one). The 19 stereo points were visually identified by hands, and used to calculate the orientation parameters, with the result By/Bx = -0.00213, Bz/Bx = -0.08319, phi = -5.194648, omega = -0.126733, kapa = 0.247479 (phi, omega and kapa are in degree).

The program started with sobel operator which produces image gradient magnitude and orientation, later filtered by Duncan's technique. Using the methods described in section 2, the resulted line detection is illustrated in Fig. 7. The short line segments were sorted out in order to reduce the number of candidate matching. The number of image lines involved in the matching is 83 on both images. The final matching has 55 lines, which has turned out to be a correct matching.

8. CONCLUDING REMARKS

This paper has presented a new approach to solve the correspondence problem. The heard of the approach is the back-projection of image space based candidate matching result into the object space and solving the stereo matching as a consistent labelling problem. The experiment has demonstrated that it is a promising idea to tackle this difficult problem. Also, our research has shown the remaining difficulties in robust line detection or extraction, choice of matching attributes and thresholds, etc. The future research will be oriented in dealing with these problems, author wish that more results would come up soon.

9. ACKNOWLEDGEMENT

This research is jointly supported by CCGM (Centre for Computer Graphics and Mapping) and FRANK project, Faculty of Geodesy, TU Delft. FRANK is a registered trademark of FRANK system supported by Geeris Holding Netherlands BV. The computational support from Photogrammetry Laboratory is also appreciated.

10. REFERENCES

1. Wrobel, B., "Facet Stereo Vision (FAST vision) - a new approach to computer stereo vision and to digital photogrammetry", In proceeding of intercommission conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, pp.231-258, 1987.

2. Helava, U.V., "Object-space least-square correlation", Photogrammetric Engineering and Remote Sensing, Vol.54, No.6, Part1, June 1988, pp.711-714.

3. March, Riccardo, "A regularization model for stereo vision with controlled continuity", Pattern recognition letters 10(1989) pp.259-263.

4. Ackermann, F. and Zheng, Y.-J., "Inverse and ill-posed problems in phtogrammetric surface reconstruction", In proceeding of ISPRS Commission III conference, Wuhan, China, 1990.

5. Zheng, Y.-J and Hahn, M., "Surface reconstruction from digital images in the presence of discontinuities, occlusions and deformations", In proceeding of ISPRS Commission II conference, Wuhan, China, 1990.

6. Baker, H. Harlyn., " Surfaces from Mono and Stereo Images ", Photogrammetria, 39 (1984) pp.217-237.

7. Benard, M., "Automatic Stereophotogrammetry: A method Based on Feature Detection and Dynamic Programming ", Photogrammetria, 39 (1984) pp.169-181.

8. Greenfeld, J.S., " Experiments with Edge-Based Stereo Matching ", Photogrammetric Engineering and

Remote Sensing, Vol.55, No.12, Dec. 1989, pp1771-1777.

9. Medioni, Gerard and Nevatia, R., "Segment-based stereo matching", Computer Vision, Graphics, and Image processing 31, 2-18(1985).

10. Marr, D., Vision, Freeman company, San Francisco, 1979.

11. Gruen, A.W. and Baltsavias, E.P., "Geometrically Constrained Multiphoto Matching ", Photogrammetric Engineering and Remote Sensing, Vol.54, No.5, May 1988, pp.633-641.

12. Quan, L. and Mohr, R., "Matching perspective images using geometric constraints and perceptual grouping", In proceeding of IEEE second international conference on computer vision, Florida.1988.

13. Kriegman, D.J., Binford, T.O., Sumanaweera, T., "Generic models for robot navigation", In proceedings of DARPA Image Understanding Workshop, 1988.

14. Haralick, R.M. and Shapiro, L.G., "The consistent labeling problem: Part I", IEEE Transaction on PAMI, VOL.PAMI-1, No.2, April 1979.

15. Haralick, R.M. and Shapiro, L.G., "The consistent labeling problem: Part II", IEEE Transaction on PAMI, VOL.PAMI-2, No.3, May 1980.

16. Duda, R.O., and Hart, P.E., Pattern Classification and Scene Analysis, New York, Wiley, 1973.

17. Kang, C.-W., Park, R.-H., and Lee, K.-H., "Extraction of straight line segments using rotation transformation: generalized Hough transformation", Pattern Recognition, Vol.24, No.7, pp.633-641, 1991.

18. Nevatia, R. and Babu, K., "Linear feature extraction and description", Computer Graphics and Image Processing, Vol.13, pp.257-269, 1980.

19. Burns, J.B., Hanson, A.R. and Riseman, E.M., "Extracting straight lines", IEEE Transactions on PAMI, Vol. PAMI-8, No.4, July, 1986.

20. Svalbe, I.S., "Natural representations for straight lines and the Hough transform on discrete arrays", IEEE Transactions on PAMI, Vol.11, No.9, Sep. 1989.

21. Mohan, R., and Nevatia, "Using perceptural organization to extract 3-D structures", IEEE Transactions on PAMI, Vol. 11, pp.1121-1139, 1989.

22. Weiss, R. and Boldt, M., "Geometric grouping applied to straight lines", CVPR, pp.489-495, 1986.

23. Huertas, A., Cole, W., and Nevatia, "Detecting runways in complex airport scenes", CVGIP 51, 107-145 (1990).

24. Duncan, J.S., Staib, L.H. and Birkhoelzer, T., "Edge reinforcement using parametrized relaxation labeling", Proceed. on Computer Vision and Pattern Recognition, San Diego, CA, June, 1989.

25. Zhang, Y., "Matching in image/object dual spaces", SPIE Vol.1526 Industrial Vision Metrology, 1991.

26. Paton, K.A., "Conic sections in automatic chromosome analysis", Machine Intell., Vol.5, 1970.

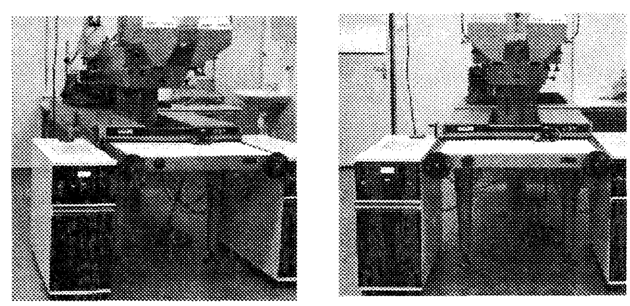

Fig. 6

stereo images on stereo plotter

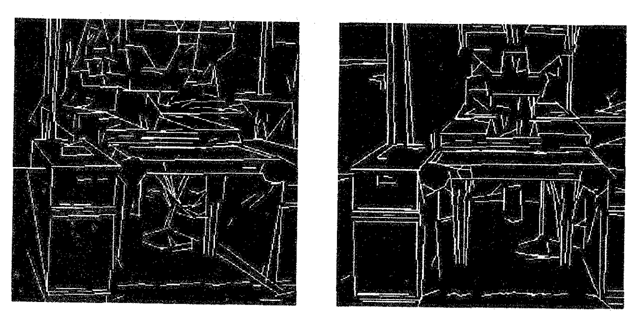

Fig. 7

line detection result