A SEGMENT-BASED MATCHING ALGORITHM IN TRINOCULAR VISION

SADAAKI YOKOI [1], GERARD MEDIONI [2] and RAMAKANT NEVATIA [2]

(1) NEC Corporation, 484, Tsukagoshi, 3-chome, Saiwai-ku, Kawasaki, Kanagawa, 210 Japan.
(2) University of Southern California, Powell Hall, Los Angeles, California 90089, U.S.A.

ABSTRACT

This paper presents a matching algorithm in three viewed images. Primitives used here are line segments, which consists of connected edges and are derived by Nevatia-Babu method. This algorithm is based on the position and orientation constraints for segments, caused by the epipolar constraint. This algorithm also involves the process to match occluded segments, which are not visible in one image of three. In stereo vision camera calibration is another important factor and we describe our calibration method for experiment. After that, applying this algorithm for real images, which are taken by using a CCD camera attached to PUMA robot, we verify its effectiveness.

INTRODUCTION

Stereo vision has been developed as an very important technique to measure positions and shapes of 3-dimensional object for the intelligent robot in the field of industrial automation. Once corresponding points between different viewed images are found, the depth map, that is, shape description could be computed by triangulation. Therefore, the main problem in stereo vision is to find corresponding points. To reduce search space for finding correspondence, edges in images are used as primitives whose correspondence should be found. Moreover, the basic matching strategies are based on the epipolar constraint, which restrict corresponding points to a point in one image should be on the epipolar line in another image.

Besides this constraint, in conventional binocular vision, hierarchical and global matching algorithms , some of them are based on another constraints and assumptions, are proposed to avoid ambiguous matches (Ohta and Kanade, 1985, Yakimovsky and Cunningham, 1978). However, binocular vision itself has two critical problems. It is very difficult to avoid incorrect matches for those edges which have similar properties along the epipolar line. And also occluded edges, which are not visible in another image, cause false matches.

An extra view is used to solve these problems and trinocular vision algorithm has been proposed to improve matching accuracy (Yachida and Kitamura, 1985). Figure 1 shows the principle of trinocular vision. According to this algorithm, the first step is to find candidates Pu and Pl which are corresponding between the first and second images of three. After that, the third image is examined to check correctness. That is, there should be actual edges in the third image at the position Pr, which could be computed by the epipolar constraint for these candidates. Therefore, it seems very powerful to improve accuracy and to choose one possible match among corresponding candidates. However, if there are still remaining several candidates, another criteria are necessary to choose one of them (Ito and Ishii, 1986, Pietikainen and Harwood, 1986).

In this paper, to avoid ambiguous matches in trinocular vision, we propose a segment-based matching algorithm. Line segments, which consist of connected edges and are derived by using Nevatia-Babu method (Nevatia and Babu, 1980), are used as

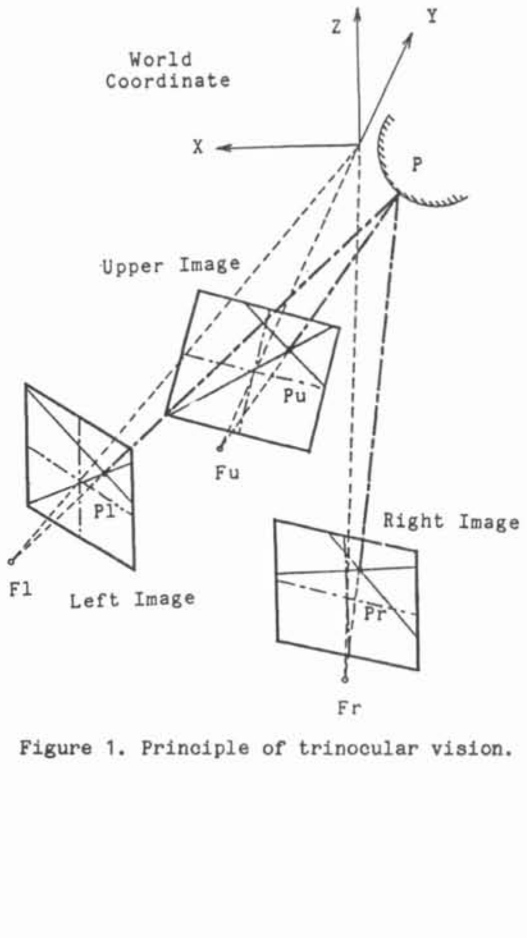


Figure 1. Principle of trinocular vision.

primitives instead of edges (Medioni and Nevatia, 1985), based on the assumption that the object could be described by its edges. This algorithm is regarded as global matching and is according to the position and orientation constraints for segments, caused by the epipolar constraint. We also describe the matching algorithm for occluded segments, which are not visible in one image of three. Camera calibration is another important factor in stereo vision and we describe our calibration method for experiment.

SEGMENT-BASED MATCHING ALGORITHM

This segment-based matching algorithm consists of three step matching strategies, based on the position and orientation constraints for segments, which are caused by the epipolar constraint.

The first step is to pick up every candidates as shown in fig. 2. One segment picked up in the upper image specifies restricted zone according to the epipolar constraint, in which corresponding segments in the left image should be found. After finding segment in the left image, these paired segments A and B are fragmented, so that the start and end points are coincident with each other. These fragmented segments define the position and orientation constraints for the corresponding segments in the right image. That is, the orientation of the corresponding segments in the right image should be almost parallel to the estimated line, which is computed by the start and end points of the fragmented segments. Moreover, its position also should be closer to this estimated line. If these conditions are satisfied by the segment C, fragmentations are repeated among them and we choose these triple segments as a candidate.

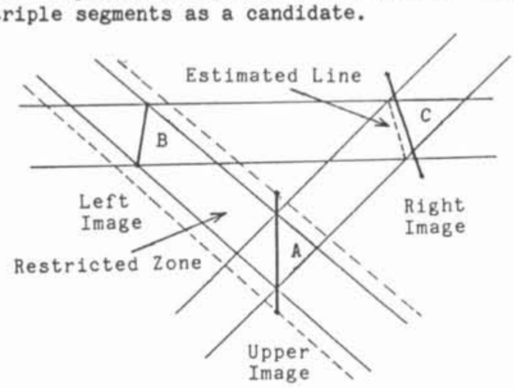


Figure 2. Position and orientation constraints.

Figure 3 shows the second step to select correctly matched segments. In this stage, we apply neighboring relation that, if a certain candidate is assumed to be matched correctly, a neighbored segment should also be matched at each neighboring position. After this process, we can pick up the correctly matched segments A with A', B with B', and C with C' as trinocular matched segments among the bunch of candidates.

The final step is to process occluded segments, which are not visible in one image of three, by applying the binocular vision algorithm as shown in fig. 4. We are only focusing on those segments A'' and B'' which are connected to the trinocular matched segments A', B' and C'. If these segments A'' and B'' satisfy the epipolar constraint between two images with keeping connection, these paired segments are added in the bunch of matched segments and this process is repeated until no more segments are added.

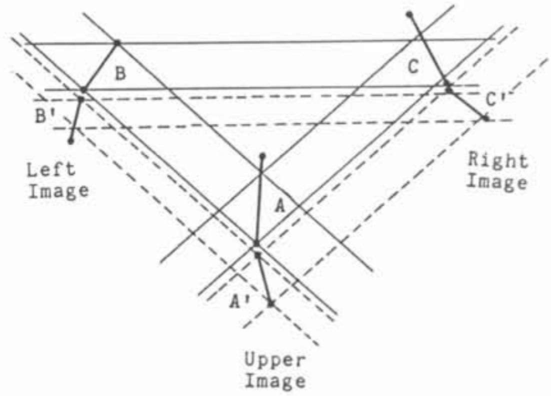


Figure 3. Trinocular matching.

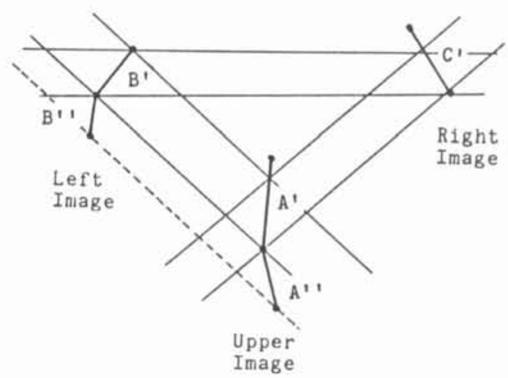


Figure 4. Process for a occluded segment.

CAMERA CALIBRATION

Camera calibration is another important factor in stereo vision to calculate epipolar lines for matching. Calibration errors of camera parameters will degrade matching reliability and cause false matches.

In our experiment, we adopted the camera geometry as shown in fig. 5. Internal camera parameters such as focal length and camera resolution for horizontal/ vertical sampling, are fixed by the initial configurations, so that we could compute these parameters previously by using a simple pattern. Therefore, we only have to estimate four parameters, that is, pan, tilt, swing and distance between focal point and origin in the world coordinate.

To simplify the calculation and reduce calibration errors, we adopted a calibration pattern, which consists of several circles and cross lines, placed on Z=0 plane in the world coordinate. Assuming that these cross lines are placed to coincide with X and Y coordinate in the world coordinate, we first estimate the corresponding camera coordinates. After that we calculate four parameters by using least squares method for the transformation matrix to the center of each circles.

Table 1 shows calibrated camera parameters. Using these parameters, epipolar lines for the experiment can be computed.

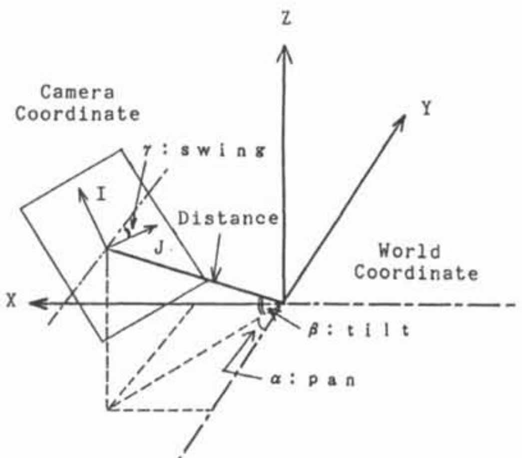


Figure 5. Camera geometry.

Table 1. Calibrated camera parameters.

focal length = 16 mm
camera resolution
horizontal = 0.0162 mm
vertical = 0.0134 mm

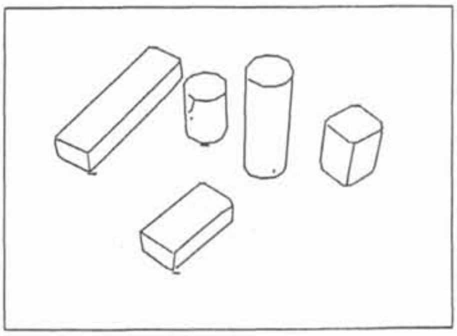
	Upper	Right	Left
Pan (degree)	2.143	23.568	-23.241
Tilt (degree)	53.252	38.616	38.264
Swing(degree)	-1.714	-0.994	0.872
Distance (mm)	476.369	472.894	474.471

EXPERIMENTAL RESULT

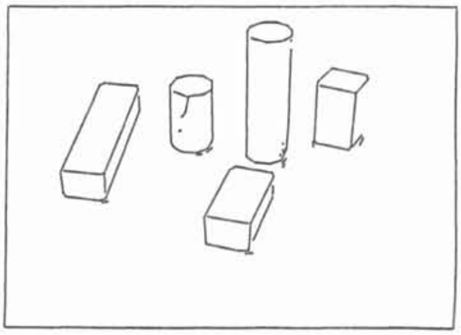
Three viewed images are taken at the same position as camera calibration, by using a CCD camera attached to PUMA robot. These images are segmented by applying Nevatia-Babu method. Figure 6(a), (b) and (c) show segmented images of a block scene, viewed from upper, left and right position respectively.

Figure 7(a) and (b) show the matched segments in upper image, which are calculated by this segment-based matching algorithm using calibrated camera parameters. The difference between fig. 7(a) and (b) is depending on the different process for occluded segments, because this final process is applied independently between upper and left image [fig. 7(a)] and between upper and right image [fig. 7(b)]. Therefore, the final matched segments, which can be used to calculate depth map, seem to be superimposition of both.

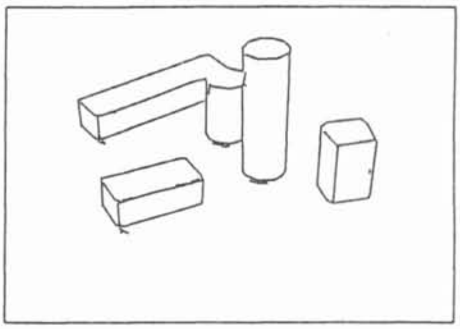
From this result, we assure that examples shown above, which have simple shapes and only a few candidates in the first calculation, could be matched well. Rims



(a)



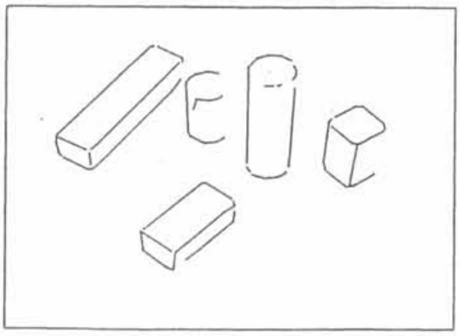
(b)



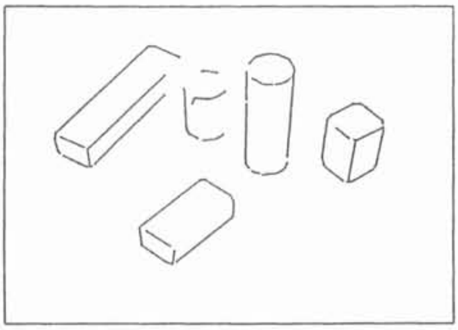
(c)

Figure 6. Segmented images of a block scene.
(a) Upper image. (b) Left image.
(c) Right image.

around the cylindrical objects do not cause any effects, because the camera positions are far from the objects. And also, as is expected earlier, occluded segments are correctly matched except for those segments whose connected segments are not matched.



(a)



(b)

Figure 7. Matched results in fig. 6.
(a) Between upper and left image.
(b) Between upper and right image.

CONCLUSION

In this paper, we proposed a segment-based matching algorithm in trinocular vision, which is based on the position and orientation constraints caused by epipolar constraint. This algorithm also involves the matching process for occluded segments. By applying this for real images, we had verified its effectiveness.

However, we have to pay more attention to process those segments which are parallel to epipolar lines, because there will be a lot of candidates. And also it is necessary to estimate the recognition accuracy, which would be degraded by segmentation, about the 3-dimensional shape of the object.

ACKNOWLEDGMENTS

The author is deeply indebted to Dr. Toru Mikami and colleagues for their efforts to research this work in University of Southern California. He also wish to thank Dr. Andres Heurtas for his helpful discussions and support.

REFERENCES

Ito, M., and Ishii, A., 1986, "Range and shape measurement using three-view stereo analysis", Conference on Computer Vision and Pattern Recognition, pp.9-14.

Medioni, G., and Nevatia, R., 1985, "Segment-based stereo matching", Computer Vision, Graphics and Image Processing, vol. 31, pp.2-18.

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

Ohta, Y., and Kanade, T., 1985, "Stereo by intra- and inter-scanline search using dynamic programming", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 7, no. 2, pp.139-154.

Pietikainen, M., and Harwood, D., 1986, "Depth from three camera stereo", Conference on Computer Vision and Pattern Recognition, pp.2-8.

Yachida, M., and Kitamura, Y., 1985, "Getting 3-D information by trinocular vision (in Japanese)", Proc. 3rd Conference of Robotics Society of Japan, pp.293-294.

Yakimovsky, Y., and Cunningham, R., 1978, "A system for extracting three-dimensional measurements from a stereo pair of TV cameras", Computer Graphics and Image Processing, vol. 7, pp.195-210.