Advanced Robotics, Vol. 4, No. 1, pp. 29–42 (1990)

Paper

Three-dimensional data acquisition by trinocular vision

YOSHIFUMI KITAMURA* and MASAHIKO YACHIDA†

Department of Control Engineering, Osaka University, 1-1 Machikaneyama, Toyonaka, Osaka 560, Japan

Received for *JRSJ* 16 October 1986; English version received 10 February 1988

Abstract—For recognition of three-dimensional (3D) shapes and measurement of 3D positions of objects it is important for a vision system to be able to measure the 3D data of dense points in the environment. One approach is to measure the distance on the basis of the triangulation principle from the disparity of two images. However, this binocular vision method has difficulty in finding a correspondence of features between two images. This correspondence problem can be solved geometrically by adding another camera, i.e. by trinocular vision. This paper presents the principles and implementation details of trinocular vision. On the basis of the proposed method, we carried out several experiments, from which we found that many correct correspondences could be established, even for images of a complex scene, by only the geometrical constraint of trinocular vision. However, when there are dense points in the image, multiple candidate points are found and a unique correspondence cannot be established. Two approaches to solve this problem are discussed in this paper.

INTRODUCTION

For recognition of 3D shapes and measurement of 3D positions it is important for a vision system to be able to measure the 3D data of dense points in the scene [1]. One approach is to measure the distance on the basis of the triangulation principle from the disparity of two images taken at two different positions, which is well known as stereo vision. Usually, feature points are first detected in both images and then the correspondence of these features is found between the two images. However, it is not easy to find the correspondence of feature points. To solve this problem, several methods have been studied, such as the use of the dynamic programming method [2] and the coarse-to-fine method [3]. These methods are not totally satisfactory because they require much computation time or are based on heuristic rules.

In this paper, we propose a new approach whereby the correspondence problem can be solved geometrically by adding an additional camera, i.e. by trinocular vision [4–6]. The principles and implementation details of this trinocular vision are presented. On the basis of the proposed method, we performed several experiments, from which we found that many correct correspondences can be established even for images of a complex scene by only a geometrical constraint of the trinocular vision. However, when there are dense points in the image, multiple candidate points are found and a unique correspondence cannot be established. To solve this problem, two approaches are examined.

*Present address: Canon Inc. Information Systems Research Center, 2-9-4 Shimomaruko Ohta-Ku, Tokyo 146, Japan.

†To whom correspondence should be addressed.

2. PRINCIPLE OF TRINOCULAR VISION

Figure 1 shows the principle of trinocular vision. F_1, F_2, and F_3 are the lens centres of cameras 1, 2, and 3, respectively. Let us denote the image of a point P in the scene to the image plane 1 as p_1 and that to the image plane 2 as p_2. Then to find the corresponding point of point p_1 in the image plane 2, we search on an epipolar line l_2 which is the image of the line of sight $F_1 p_1$ to the image plane 2.

Let us call the plane made of two lines—a base line connecting the two lens centres F_1 and F_2 and the line of sight $F_1 p_1$—an epipolar plane. Then the epipolar line can be easily obtained as the intersection between the epipolar plane and image plane 2.

If there is only one point on the epipolar line l_2, it is easy to find the corresponding point of p_1. However, there are usually many points on the epipolar line because all visual points on the epipolar plane project on to the same epipolar line l_2. For example, in Fig. 1, the point R, which lies on the epipolar plane, projects onto the same epipolar line l_2. Then it is not easy to determine which of p_2 or r_2 is the true corresponding point of p_1.

This problem can be solved by introducing another camera 3. In Fig. 1, p_3 and r_3 are the images of points P and R to the image plane 3, respectively. Because p_3 is the image of point P, p_3 must lie on the epipolar line l_3, which is the image of the line $F_1 p_1$ to the image plane 3. However, the other points do not lie on l_3. Furthermore, if p_2 and p_3 are true corresponding points, p_3 must lie on the epipolar line m_3, which is the image of the line $F_2 p_2$.

In summary, if the point p_2 is the true corresponding point of p_1, the image of point P must exist exactly at the intersection of the two epipolar lines l_3 and m_3 on image plane 3. In Fig. 1, since p_1 and p_2 are true corresponding points, an image of point P, p_3, exists at the intersection of l_3 and m_3. Thus, it is concluded that the three points p_1, p_2, and p_3 are the corresponding points. Then the 3D position of point P is determined as the intersecting point of the three lines $F_1 p_1$, $F_2 p_2$, and $F_3 p_3$.

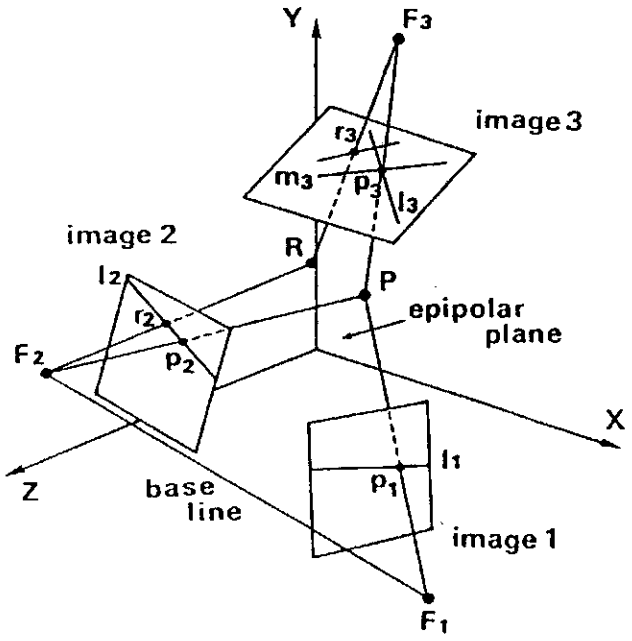


Figure 1. Principle of trinocular vision.

In this way, by using three cameras, we can solve the correspondence problem uniquely.

3. PROCEDURE OF THE TRINOCULAR VISION SYSTEM

The procedure to determine the 3D position of each point by the trinocular vision system is described below. First the camera parameters are calculated analytically by viewing a known object placed at a known position [7]. Images are taken at three positions and then the flow chart shown in Fig. 2 is followed:

(1) *Edge detection at each image.* A Sobel operator is applied to each image and non-maximal suppression is applied in order to make the image narrow.
(2) *Generation of epipolar lines.* Epipolar lines (l_1 and l_2) are generated in images 1 and 2. We compute the epipolar lines l_1 and l_2 as the intersecting lines between the epipolar plane and the image planes 1 and 2. This epipolar plane is rotated from the upper side of the image plane to the lower side of the image plane by a certain angle as shown in Fig. 3. The angle is set to a value which corresponds to one picture element. The base line is the axis of this rotation.
(3) *Finding edges on the epipolar line l_1.* We search for edges on the epipolar line in image 1. If there are no edges on the epipolar line, we look for them on the next epipolar line. If there are edges on the epipolar line, go to step 4 and find the correspondence.
(4) *Finding edges on the epipolar line l_2.* We search for edges on the epipolar line in image 2. If there is only one edge point at each image, its 3D position is determined as the intersecting point of the two lines F_1p_1 and F_2p_2. Otherwise, we establish correspondence using the third image.
(5) *Establish correspondence.* From the leftmost point on the epipolar line l_1 of image 1, we try to determine the correspondence geometrically.

Now, let us concentrate on finding a corresponding point for point p_i. If there are n edge points ($q_1, \ldots, q_n$) on the epipolar line l_2 in image 2, this means that there are n candidate points for point p_i. For each point q_j, we draw the epipolar lines $m_1, \ldots, m_n$ which are images of the line $F_2q_1, \ldots, F_2q_n$. Then there is

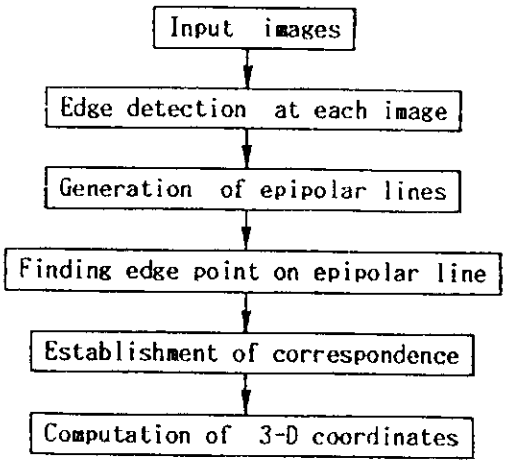


Figure 2. Procedure of trinocular vision.

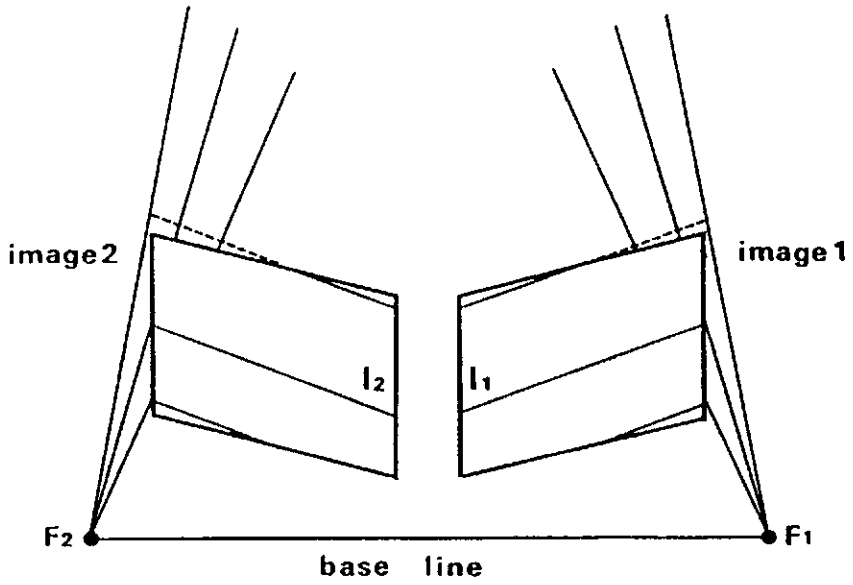


Figure 3. Generation of epipolar lines.

usually an edge point at the intersecting points between line l_3 and lines $m_1, \ldots, m_n$.

This point is the corresponding point of p_i. From the epipolar line number where the corresponding point is found, we can also determine the corresponding point q_j in image 2. Then the 3D position of point P can be obtained as the intersecting point of the three lines p_1F_1, p_2F_2, and p_3F_3.

(6) When we have established all the correspondence of all the edge points on epipolar line l_1, go to step (2), draw the next epipolar line, and then compute all edge points in image 1.

4. SYSTEM CONFIGURATION

Figure 4 shows the configuration of the trinocular vision system used in this experiment. Input images, taken by three cameras, are digitized by the image input device and then sent to a personal computer where the 3D positions of the scene are

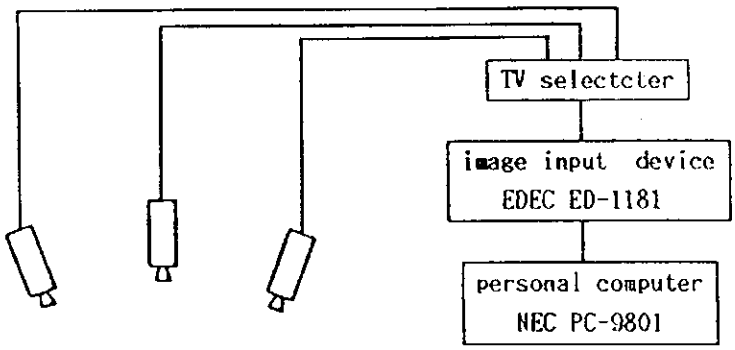


Figure 4. System configuration.

calculated. The programs are coded by Pascal MT + 86 with the operating system CP/M-86.

Figure 5 shows a view of the trinocular vision system and Fig. 6 a picture of the camera stand. After being fixed manually on the pan head, the camera is adjusted precisely.

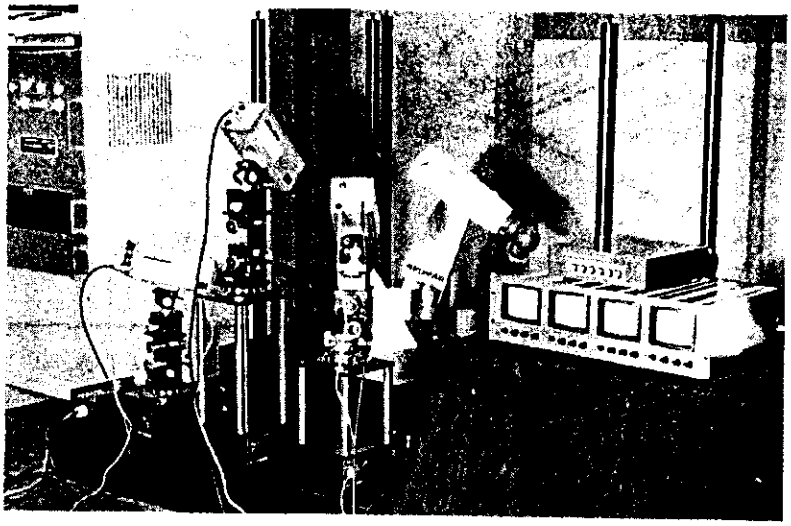


Figure 5. The trinocular vision system.

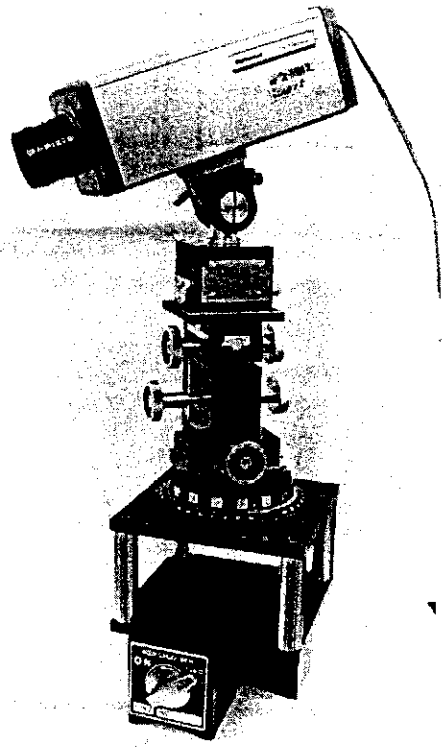


Figure 6. The camera stand.

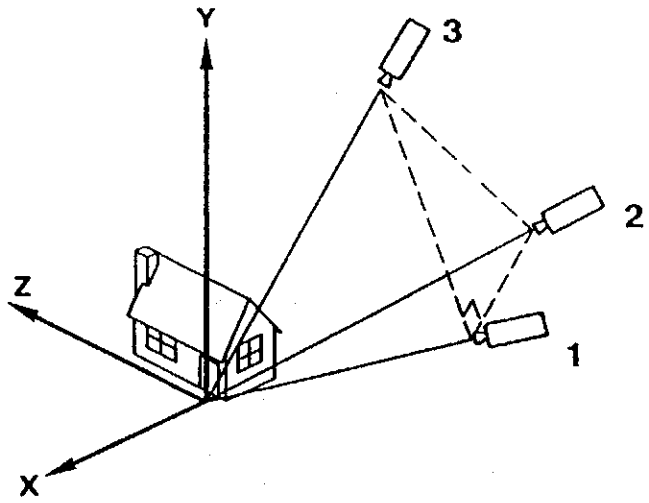
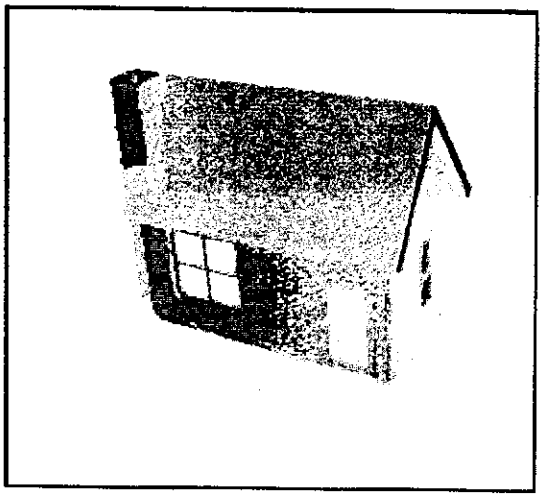
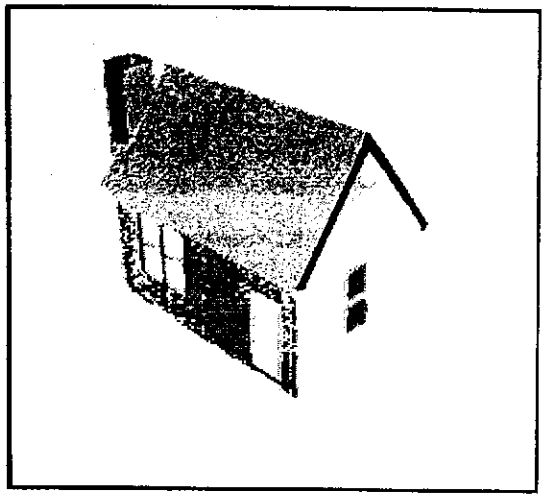


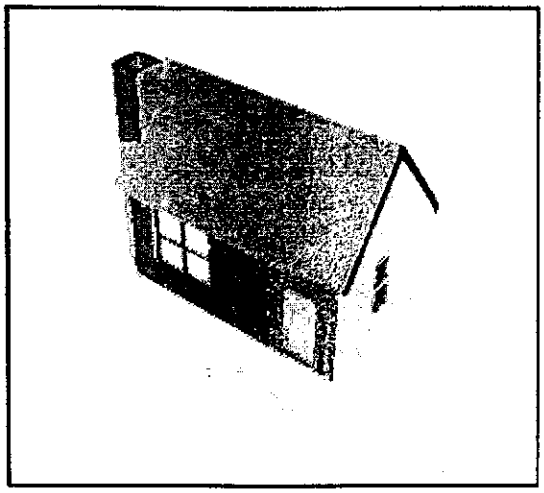
Figure 7. Camera arrangement.



(a) Image 1



(b) Image 2



(c) Image 3

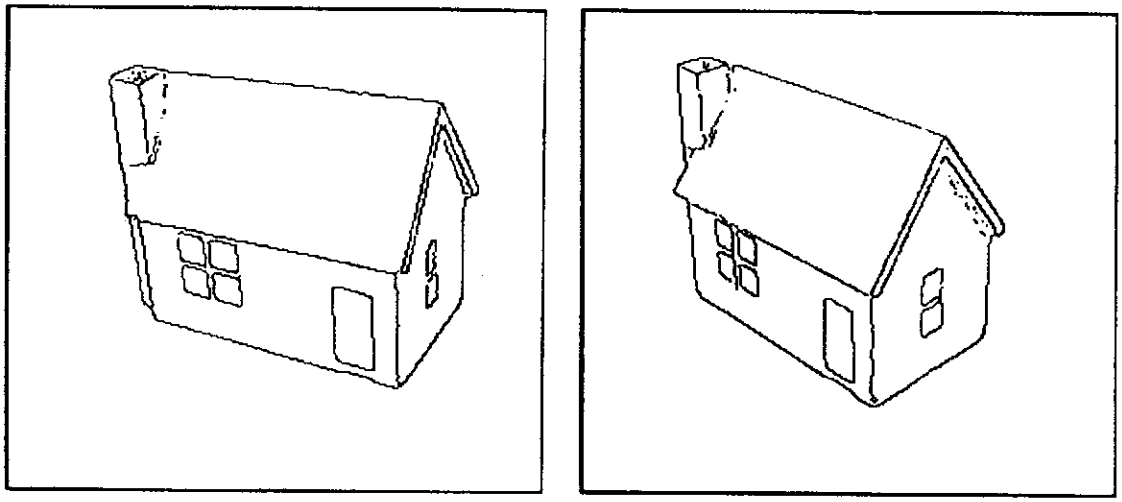
Figure 8. Input images.

5. EXPERIMENTAL RESULTS

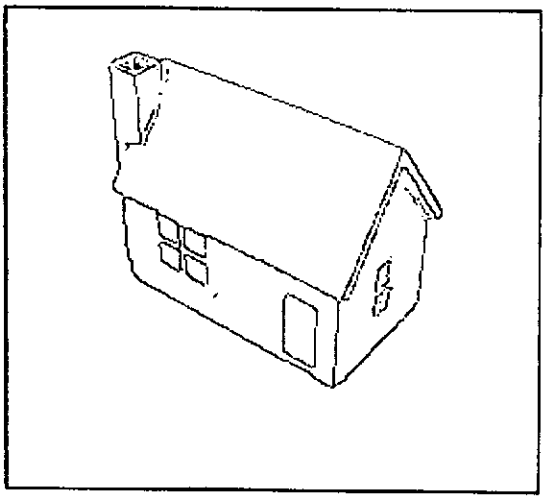
Figure 7 shows the camera arrangement and the world coordinates system. Three cameras were placed at the vertices of a regular triangle of length 10 cm and viewed a model of a house from a distance of 30 cm with a visual field of 15 cm. The measured camera parameters are listed in Table 1.

Figures 8 and 9 show the input images and detected edge points upon applying a Sobel operator and non-maxima suppression.

Figure 10 shows an intermediate procedure for finding the correspondence between three images. In order to find the correspondence of point p_1 in image 1(a), the system draws the epipolar line l_2 on image 2(b). There are five candidate points $(q_1, \ldots, q_5)$ for point p_1. To determine the correspondence uniquely, we draw five epipolar lines $m_1, \ldots, m_5$ which are images of the line $F_2q_1, \ldots, F_2q_5$.

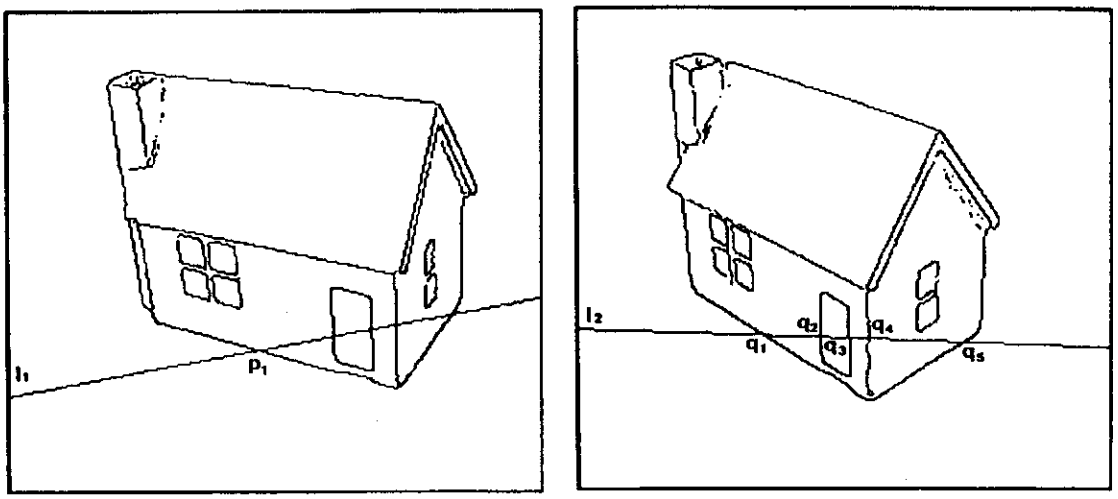


(a) Image 1 (b) Image 2

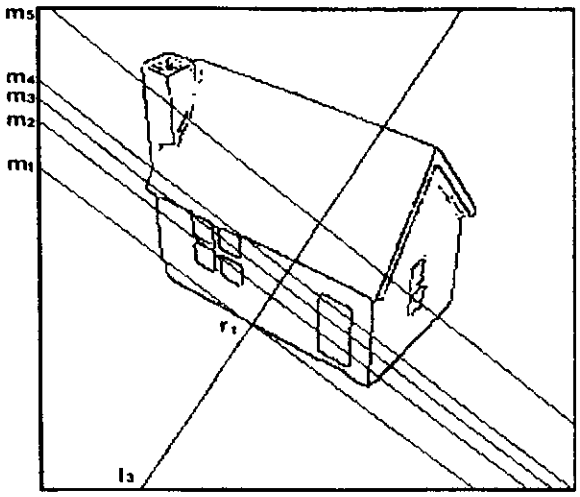


(c) Image 3

Figure 9. Detected edges.



(a) Image 1 (b) Image 2



(c) Image 3

Figure 10. Intermediate procedure to find correspondence.

An epipolar line l_3, generated from the first image, has five intersections with epipolar lines $m_1, \ldots, m_5$, which are generated from the second image. However, since there exists only an edge point at these intersections, it is concluded that r_1 is the corresponding point of p_1. Then the epipolar line number of 'one' indicates that q_1 is also the corresponding point in image 2.

In this way, the system finds the correspondence automatically and calculates the 3D positions of the edge points in the scene. Figure 11 shows the experimental result with the world coordinate (X–Y–Z).

6. DISCUSSION

We could obtain exact 3D data from correct correspondences established by the trinocular vision system as shown in Fig. 11. However, in Fig. 11 there are some

Table 1.
Camera parameters

		Camera 1	Camera 2	Camera 3
Translation	X (mm)	163.59	101.50	136.12
	Y (mm)	178.66	183.77	228.90
	Z (mm)	− 147.12	− 199.69	− 159.56
Rotation	α	0.703	0.879	0.793
		0.101	0.043	0.018
		0.704	0.474	0.609
	β	− 0.406	− 0.265	− 0.385
		0.870	0.871	0.788
		0.280	0.413	0.479
	γ	− 0.584	− 0.395	− 0.472
		− 0.482	− 0.489	− 0.615
		0.653	0.778	0.632

Focus length: 535.5 pixel; image centre: 142.1, 137.0; aspect ratio: 0.983.

points which are not on the edge of the object. These false points are caused by multiple candidates, shown in Fig. 12. When both epipolar lines m_1 and m_2 have intersections with l_1 on the edge points r_1 and r_2 in the third image, the system cannot judge which is the real corresponding point. Therefore, we computed the 3D position for each candidate point as shown in Fig. 11.

It is easily understood that when the density of the edge point increases, the chances of having multiple candidate points also increase. We will discuss the geometrical constraint in Section 6.1, and the quantization error and inaccuracy of the camera parameters in Section 6.2.

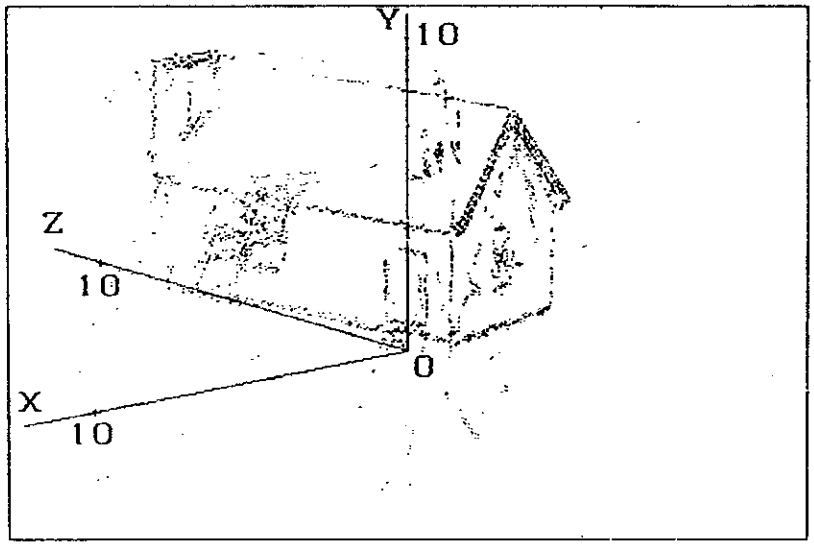


Figure 11. Experimental result using 3 × 3 pixels searching method. Units are in cm.

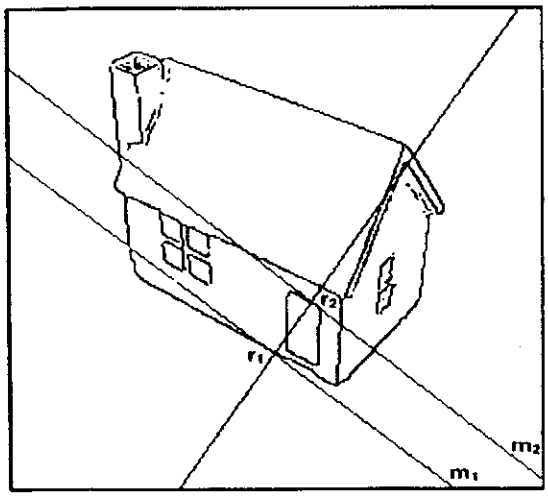


Figure 12. Third image in the case of multiple candidates.

6.1. Estimation of the geometrical constraint

We discuss the probability of being able to establish a unique correspondence by geometrical constraints. When we assume that there are x kilo-points of edge points in the 64 kilo-points (256 × 256) image uniformly, then the density of the edge points becomes $x/64$. Because there are $4x$ edge points on one epipolar line in each image, an edge point on image 1, p_1 has $4x$ candidate points in image 2. When we draw the epipolar lines $m_1, \ldots, m_{4x}$ on image 3 for each point $(q_1, \ldots, q_{4x})$, each line has an intersection with l_3 which is generated from the first image. Figure 13 indicates this process. If the camera parameters are accurate and there is no dead angle, the probability that only one of the $4x$ intersections has an edge point on image 3 is calculated as P.

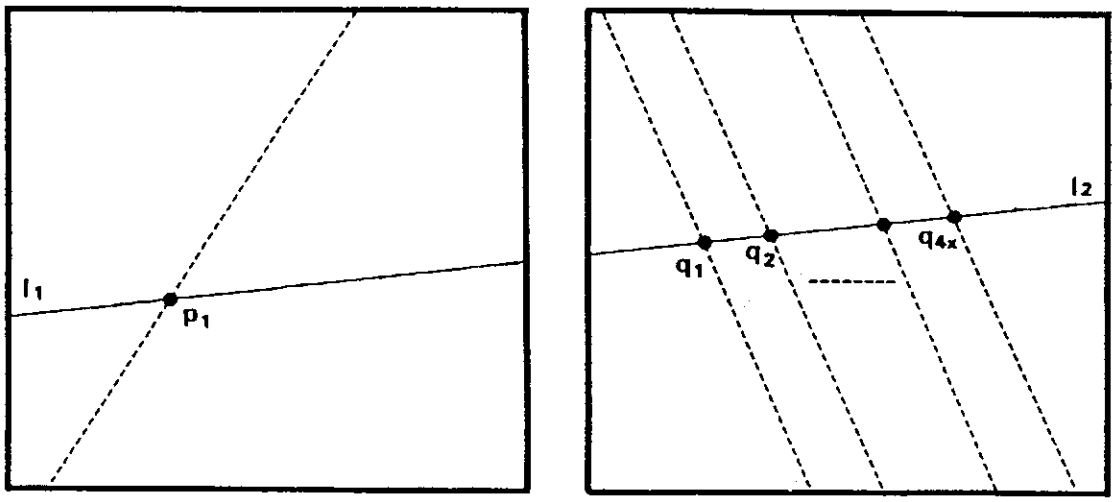
$$P = \frac{x/64(1 - x/64)_{4x}^{4x-1} C_1}{1 - (1 - x/64)^{4x}}.$$

Figure 14 shows the probability of establishing a unique correspondence when the number of feature points increases.

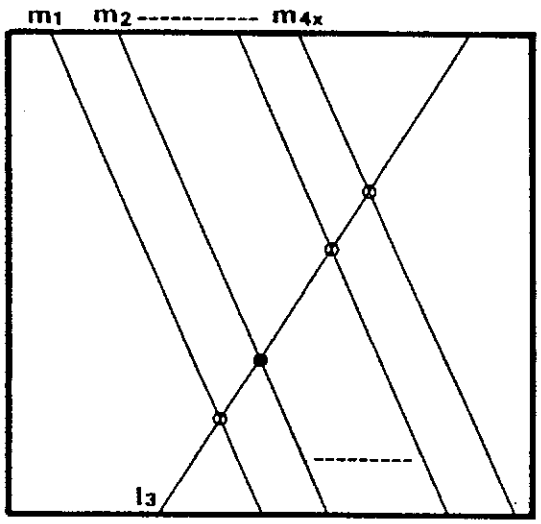
6.2. Estimation of the error

The trinocular vision system determines the correspondence when an edge point exists at the intersecting point of two epipolar lines on image 3. In practice, however, there is not always a true corresponding edge point at the estimated picture element, because of the quantization error and the inaccuracy of the camera parameters. Therefore, we set a 3 × 3 pixels searching area whose centre is the intersection of two epipolar lines, and if an edge point is found in this area, we designate this point as the corresponding point.

However, this method also has a disadvantage. Expansion of the search area contributes to not only finding many true correspondences but also picking up some false correspondences. Figure 11 gives an experimental result using the 3 × 3 pixels searching method and Fig. 15 shows the result using the 1 pixel searching method.



(a) image 1 (b) image 2



(c) image 3

Figure 13. Probability of having multiple candidates.

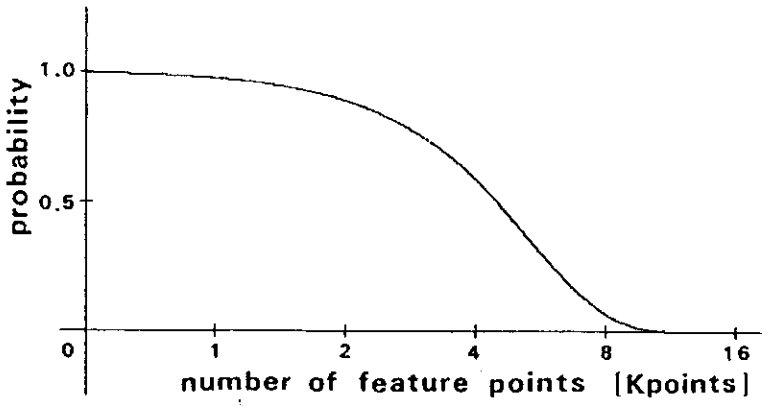


Figure 14. Probability of being able to establish a unique correspondence.

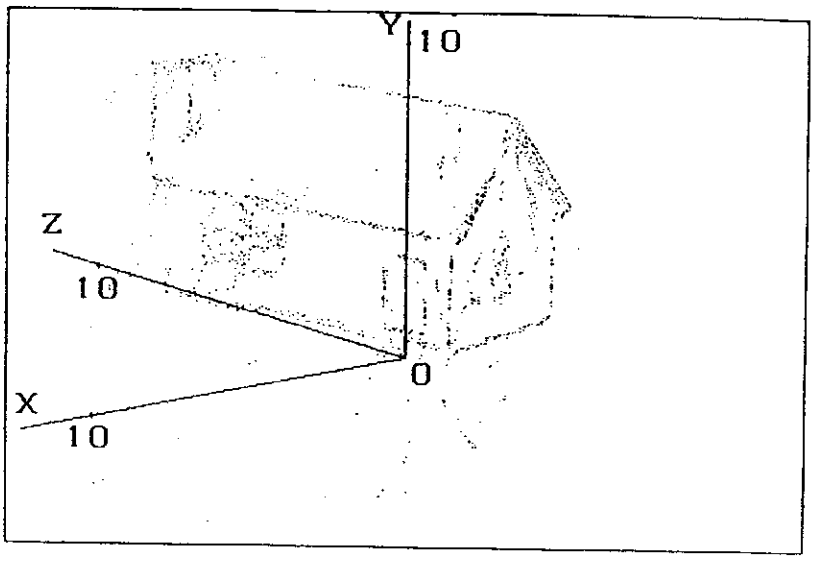


Figure 15. Experimental result using 1 pixel searching method. Units in cm.

Table 2.
Comparison between the two searching methods

	Searching method			
	1 pixel		3 × 3pixels	
Number of feature points in 1st image	1557			
Feature points without correspondence	347	(22.3%)	48	(3.1%)
Feature points with a unique correspondence	522	(33.5%)	166	(10.7%)
Feature points with two candidates	398	(25.6%)	258	(16.6%)
Feature points with more than two candidates	290	(18.6%)	1085	(69.6%)

Table 2 compares the experimental results of these two searching methods. We could not obtain all the necessary correspondences by the 1 pixel searching method. But we could establish both many true correspondences and some false correspondences by the 3 × 3 pixels searching method.

7. CONCLUSIONS

In this paper we have described the principle and implementation details of the trinocular vision system. As discussed in Section 6, when the density of the edge point increases, the chances of the points having multiple candidate points also increase.

To determine the efficiency of the geometrical constraint of trinocular vision, we did not use existing methods which have been studied for stereo vision. For example, when there are multiple candidate points by geometrical constraints, we can check the similarity of local properties and edge continuity, and can determine the true corresponding point.

Another approach to solving the multiple candidates problem is to compute the 3D position of each candidate point, and to determine later which is the true one when the 3D positions of all the other points have been obtained. Because not all the points have

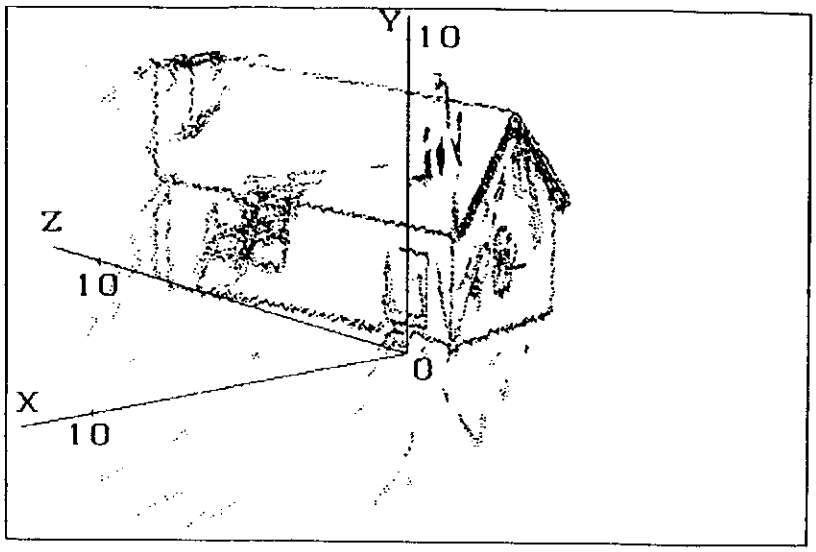


Figure 16. Result derived from unique correspondences. Units in cm.

multiple candidates, the reliable unique correspondences can eliminate false candidates in the neighbourhood. As shown in Table 2, the probability of having a unique correspondence using the l pixel searching method is higher than that using the 3 × 3 pixels searching method. So we can select the true correspondences from the candidates by the 3 × 3 pixels searching method using reliable 3D information obtained from unique correspondences as shown in Fig. 16.

The experimental results of these two approaches will be described in other papers.

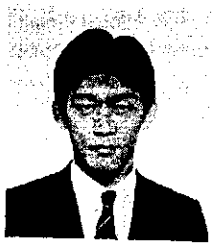
Acknowledgement

We would like to thank Professor S. Tsuji of Osaka University for useful discussions.

REFERENCES

1. M. Yachida, "3-D position and shape measurement by vision and its applications," *Systems Control*, vol. 29, no. 10, pp. 631–638, 1985 (in Japanese).
2. Y. Ohta and T. Kanade, "Stereo by two-level dynamic programming," *Proc. 9th IJCAI*, p. 1120, 1985.
3. W. E. L. Grimson, *From Images to Surfaces*. Cambridge, MA: MIT Press, 1981, pp. 53–58.
4. M. Yachida, "3-D object recognition by multiple views," *Nikkei Mechanical*, pp. 82–91, 1984 (in Japanese).
5. M. Yachida and Y. Kitamura, "3-D data acquisition by multiple views," *Preprints of 3rd ISRR*, 1985. Also published in *Robotics Research*, Cambridge, MA: MIT Press, 1986, pp. 11–18.
6. Y. Kitamura and M. Yachida, "Getting 3-D information by trinocular vision," *Proc. Natl Conf. Robotics Society of Japan*, pp. 293–294, 1985 (in Japanese).
7. T. Kasai, T. Asahi, T. Yoshimori and S. Tsuji, "Measurement system of 3-D motion using a pair of position sensing detector cameras," *SISC* vol. 19, no. 12, pp. 997–1003, 1983 (in Japanese).

ABOUT THE AUTHORS



Yoshifumi Kitamura (S'87-M'88) was born in Hyogo, Japan on 30 July 1962. He received B.Sc. and M.Sc. degrees in control Engineering from Osaka University in 1985 and 1987, respectively. In 1987, he joined Canon Inc., where he is a research scientist in the Machine Intelligence Division, Information Systems Research Center. His current research interests include computer vision, knowledge information systems, and their applications to industrial machine vision. He is a member of the Robotics Society of Japan.

Masahiko Yachida (M'83) received B.E. and M.Sc. degrees in electrical engineering and a Ph.D. degree in engineering science from Osaka University, Osaka, Japan in 1969, 1971 and 1976, respectively. He worked on the Intelligent industrial Project at the Electrotechnical Laboratory of the Japanese government from 1969 to 1970; he was a Research Associate at the Coordinated Science Laboratory, University of Illinois, Urbana, U.S.A. from 1973 to 1974; and was a Research Fellow at the University of Hamburg from 1981 to 1982. Since 1971 he was an Assistant Professor and he is currently an Associate Professor of Control Engineering at the Faculty of Engineering Science, Osaka University. His research interests are in the fields of artificial intelligence, computer vision, picture processing, and mobile robots.