A FAST METHOD FOR EXTRACTION OF 3-D INFORMATION
USING MULTIPLE STRIPES **AND TWO CAMERAS**

Tomio Echigo + and Masahiko Yachida ++

+ **IBM** Japan, Ltd.
Yasu, Shiga 520-23, Japan.

++ Osaka University
Toyonaka, Osaka 560, Japan.

ABSTRACT

This paper describes a method to extract 3-D information from two cameras in the scene on which multiple stripes are projected. Though a single camera cannot identify each stripe in a scene and its image when discontinuities of each stripe occur, with one more camera which our method employs the identification problem can be solved, because multiple stripes and two cameras give two constraints. One is the geometric constraint which gives the necessary condition for identification of each stripe. Another is the local constraint that features between images lie in constant order. After applying the geometric constraint, utilization of the local constraint enables identification of each stripe in a scene and its image. As a result, range data are obtained along multiple stri pes.

We also give a new method for computing camera parameters of 6 degrees of freedom which influence accuracy of 3-D information. They are derived mathematically by seeing the known cube.

1. INTRODUCTION

The importance of 3-D information in robotics has been widely recognized. One approach is to measure the distance on the basis of the triangulation principle from the disparity of two images taken at two different position, which is well known as stereo vision. This method has long been studied; however, it has a few problems; one is detecting features which are easily recognized in both images [1],[2] and the other is finding correspondence of those features between both images[2],[3].

On the other hand, the structured light method which replaces one of the cameras in stereo vision by a spot or a sheet of light projector solves the above problems[A], However, this technique needs much time to extract range data of the entire image, because a sheet of light must be scanned across the scene.

In this paper we project multiple stripes. In this method, however, one image cannot inform of identifying each stripe in a scene and its corresponding image when discontinuities of each stripe occur; for example, when the object is concave or occlusion occurs between objects. In order to identify each stripe in a scene and its image, our method employs one more camera in addition to a light projector and a camera. Then epipolar lines which two cameras produce on both images and multipie sheets of light yield the geometric constraint, which gives the necessary condition for identification of each stripe. In addition, under a certain condition which is described later, the local constraint is useful, which means that some features on epipolar lines lie in constant order between both images. Using the local constraint after the geometric constraint, we can identify each stripe between a scene and its image.

It is important to determine the camera parameters precisely, because they influence the accuracy of the computed value of 3-D position. Gennery[5] determined the most suitable camera parameters which minimize the sum of the errors over the known matching points. However, applying Gennery's method requires that the initial estimate is fairly accurate. Provided an uncertain initial estimate, this method may obtain the incorrect camera parameters as the most suitable ones. On the other hand, our method calculates the camera parameters mathematically from the image of a known cube. Therefore our method does not need the initial estimate and can obtain the camera parameters automatically[61.

2. CALCULATION OF CAMERA PARAMETERS

2.1 Extraction of the line of sight

Fig.1 illustrates a schematic of the object coordinates system(O-XYZ) and the camera coordinates system(F-xyz).

Then, let **P** denote the position vector of the point P and let the coordinate of its image p be (x,y,1) in the camera coordinates, and the line of sight of the point P is expressed as

$$\left.\begin{array}{l}(\mathbf{e_x} - \hat{x}\,\mathbf{e_z})\bullet\mathbf{P} = (\hat{x}_O - \hat{x})\,\mathbf{Po}\bullet\mathbf{e_z} \\ (\mathbf{e_y} - \hat{y}\,\mathbf{e_z})\bullet\mathbf{P} = (\hat{y}_O - \hat{y})\,\mathbf{Po}\bullet\mathbf{e_z}\end{array}\right\} \qquad (2.1)$$

where $x/1=\hat{x}$, $x_O/1=\hat{x}_O$, $y/1=\hat{y}$, $y_O/1=\hat{y}_O$

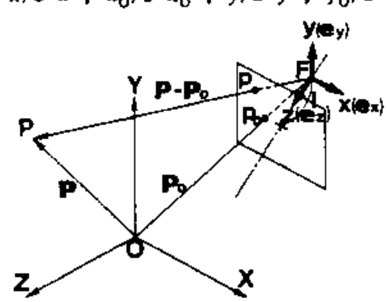


Fig.1 The coordinates systems.

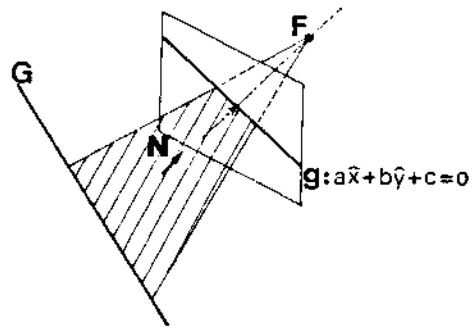


Fig.2 Projection of the straight line.

2.2 Rotation parameters

The unit orientation vectors $\mathbf{e_x},\mathbf{e_y},\mathbf{e_z}$ are expressed by

$$\mathbf{e_x} = \begin{bmatrix} Sx \\ Sy \\ Sz \end{bmatrix} \quad \mathbf{e_y} = \begin{bmatrix} Tx \\ Ty \\ Tz \end{bmatrix} \quad \mathbf{e_z} = \begin{bmatrix} Ux \\ Uy \\ Uz \end{bmatrix} \qquad (2.2)$$

Fig.2 shows that the straight line G in the space is mapped as the line g onto an image plane. The equation of the line g on an image plane is given by $a\hat{x}+b\hat{y}+c=0$. Then the surface normal **N** of a plane including the lens center F and the line G is

$$\mathbf{N} = a\mathbf{e_x} + b\mathbf{e_y} + c\mathbf{e_z} \qquad (2.3)$$

The relationship between the surface normal **N** and the orientation vector **G** of the line G is $\mathbf{N}\bullet\mathbf{G}=0$.

Since there are 3 degrees of freedom for rotation, three known lines of which the orientation vectors are linearly independent yield the rotation parameters. When three straight line G_1,G_2,G_3 are given by (1,0,0),(0,1,0),(0,0,1), let the coefficients of the image g_1,g_2,g_3 of G_1,G_2,G_3 be $a_i,b_i,c_i(i=1,2,3)$ and from $\mathbf{N}\bullet\mathbf{G}=0$ we can obtain the following Eqs.

$$\begin{aligned} a_1Sx + b_1Tx + c_1Ux &= 0 \\ a_2Sy + b_2Ty + c_2Uy &= 0 \\ a_3Sz + b_3Tz + c_3Uz &= 0 \end{aligned} \qquad (2.4)$$

Since the 9 unknown parameters Sxyz,Txyz,Uxyz can be expressed as the 3 rotation parameters $\alpha\ \beta\ \gamma$, the unit orientation vectors can be determined mathematically from Eqs.(2.4).

2.3 Translation parameters

Translation parameters are determined by a known point in the object coordinates. Assuming that the known point P_1 in the scene maps onto the point $p_1(x_1,y_1,1)$ in the image, the following value is obtained from Eq.(2.1);

$$\mathbf{Po}\bullet\mathbf{e_z} = \frac{(\mathbf{e_x}-\hat{x}_1\mathbf{e_z})\bullet\mathbf{P}_1}{\hat{x}_1-\hat{x}_0} = \frac{(\mathbf{e_y}-\hat{y}_1\mathbf{e_z})\bullet\mathbf{P}_1}{\hat{y}_1-\hat{y}_0} \qquad (2.5)$$

Therefore the location of the lens center F (X_f,Y_f,Z_f) is found as

$$\begin{bmatrix} X_f \\ Y_f \\ Z_f \end{bmatrix} = \mathbf{Po}\bullet\mathbf{e_z} \begin{bmatrix} \mathbf{e_x}^t \\ \mathbf{e_y}^t \\ \mathbf{e_z}^t \end{bmatrix}^{-1} \begin{bmatrix} \hat{x}_0 \\ \hat{y}_0 \\ 1 \end{bmatrix} \qquad (2.6)$$

3. EXTRACTION OF 3-D INFORMATION

3.1 Calculation of 3-D location

Assuming that a light source makes light planes in the scene through multiple slit, the location of a point in the scene along the stripe pattern is obtained from the intersection between a plane and a line of sight. With one image of multiple stripes, however, we cannot always identify each stripe with its image. Therefore we use one more camera in addition to a light projector and a camera.

3.2 The geometric constraint

When two cameras are laterally displaced, the geometric constraint can be found on left and right images, as shown in Fig.3 and Fig.4. In Fig.3 a plane containing the stereo pair line produces the intersections across the left image plane and the right one. These intersections are called epipolar lines. All visual points on an epipolar plane must map onto a left epipolar line and a right one. Hence two dimensional matching in two images results in one dimensional one on those epipolar lines.

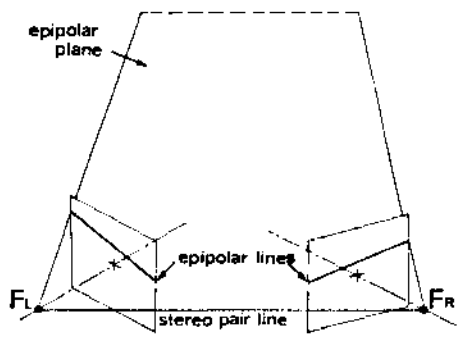


Fig.3 The epipolar geometry.

In Fig.4 an intersection r between an epipolar line and one of stripes in the right image is considered as follows. The line of sight which connects the intersection r with the lens center F_R crosses some light planes Si(i=1,2,..) in the scene. These crossing points Ri (i=1,2,..) in the scene projects its images p_i (i=1,2,..) onto the left image, so that all of the projected points p_i lie on the epipolar line of the left image. Now a stripe s_j in the right image corresponds to a stripe $\overline{s_k}$ in the left one. Then a true matching

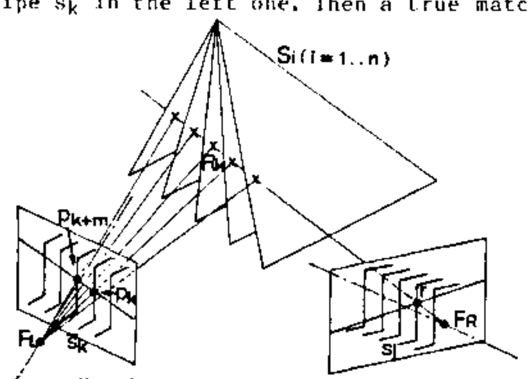


Fig.4 The geometric constraint.

point p_k must correspond to the intersection between the epipolar line and the stripe pattern in the left image as same as the intersection r in the right image. The others ($p_{k-1}, p_{k+1}, \dots, p_{k+x}$) distribute on the epipolar line and few of them like a p_{k+m} in Fig.4 may correspond to the intersection between the epipolar line and a stripe which should not match with the one of stripes investigated on the right image. These projections ($p_k, p_{k+m}, \dots$) corresponding to the intersections in the left image, which contain a true and a false, are treated as matching candidates.

Therefore the geometric constraint cannot match all of them, but rejects false matching candidates remarkably.

3.3 The local constraint

On the ordinary scene, left and right images may exchange the order of some features on two epipolar lines only when some objects exist in front and in the rear in the scene. Fig.5 shows the top view of this situation. Assuming that the point Pi on the surface of the object P will project onto both images, the sub-object Q, which exchanges the order of lines containing Pj and the other features on the epipolar lines in the left image and the right one, should lie in the laterally striped region R_2 in Fig.5. At the same time Q occludes some features in the checked region R_3 on either of images. In Fig.5 the order of the point P_1 and its neighboring features is exchanged between the left image and the right one by Q.

however, the greater Q is, the larger region is occluded by Q. Then a set of the point Pi and its neighboring features whose order is exchanged becomes smaller and at last P_1 may be occluded too. On the other hand, under the condition that the angle made of optical axes of left and light cameras is very small, the laterally striped region R2 becomes narrow.

From the above considerations, it can be said that exchange of the order occurs only when a very small object exists in front of the larger object, and stripe pattern is projected on it.(However, in real scenes actually stripe pattern will seldom appear on it.) Therefore we can use monotony of order for correspondence, that is the order of appearing features on the epipolar lines is not exchanged for the left image and the right one. Similarly if a projector lie in the neighborhood of a camera, the local constraint should be effective between light planes and stripes in the image.

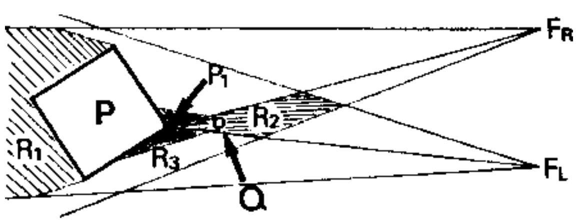


Fig.5 A counter-example of the local constraint.

3.4 Finding correspondences

First of all, lines of sight through intersections between an epipolar line and stripes of a right image are drawn. Then crossing points between light planes and lines of sight are treated as candidates for identification.

Secondly the geometric constraint reduces a lot of improper candidates, using the left image to project all crossing points. Then the local constraint is applied to the rest of candidates, that is when one stripe in the right image has plural candidates as a corresponding light plane, a candidate which is out of line with candidates of neighboring stripes must be canceled. For example, let a stripe image rj and neighbors have the following candidates:

r_{i-1} ------- $S_m, \dots, S_{m+s}$
r_i ------- S_k
r_{k+1} ------- $S_n, \dots, S_{n+t}$

then candidates Sk for ri should be satisfied with

$$S_k = \{ S_x \mid S_m < S_x < S_{n+t} \}$$

This procedure keeps running until each stripe in the right image corresponds to the only one light plane or a number of candidates does not decrease.

4. EXPERIMENTAL RESULTS

4.1 Camera parameters

A cube in Fig.6 is used for determination of camera parameters. Three edges of a cube denote X,Y,Z axes in the object coordinates system.

In order to verify a camera parameter obtained by our method, a calculated parameter is compared

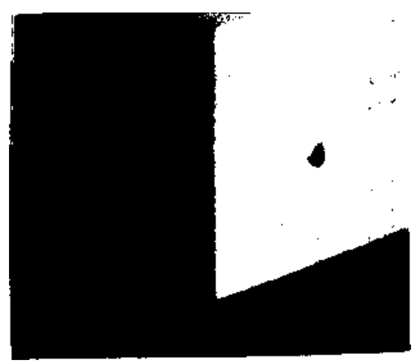


Fig.6 The cube for determination of camera parameters.

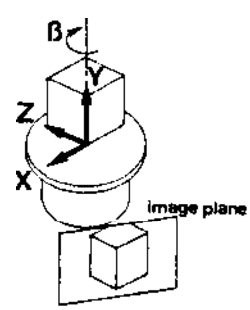


Fig.7 The experiment to examine a camera parameter β derived by our method.

Table 1 Comparison between calculations and measurements.

measurements $\Delta\beta_1$ (degrees)	calculations $\Delta\beta_2$ (degrees)	errors $\Delta\beta_2-\Delta\beta_1$ (degrees)
−10	− 9.85	+0.15
− 5	− 4.85	+0.15
+ 5	+ 4.82	−0.18
+10	+10.18	+0.18
+15	+14.97	−0.03
+20	+19.90	−0.10

with a measured one using the rotary table(Fig.7) as shown in Table 1. Because the rotating axis of the table is made to correspond to Y axis of the cube, rotation angle ft around Y axis becomes independent of the others.

4.2 Extraction of 3-D position

The thinning images of the objects from two cameras are shown in Fig.8. Then a pair of epipolar lines is drawn on the left image and the right one. All crossing points between light planes and lines of sight from the right image are projected onto the left epipolar line (Fig.9). Since one of them matches with an intersection between the left epipolar line and a stripe image in Fig.9, identification between the light plane and its image has accomplished. In case of plural candidates, the local constraint works successfully for reduction of them.

Fig.10 shows that some 3-D positions calculated from identified stripes map onto the horizontal plane. Accuracy about the height is less than 2.5% against the visual field.

6. CONCLUSIONS

We have presented a method for extraction of range data in the scene on which multiple stripes are projected. This method combines the structured light, method with stereo image one.

At first, the camera parameters were automatically computed by seeing the known cube. Secondly, in order to identify each stripe in a scene and its image, we utilized two constraints. One is the geometric constraint, which gives the necessary condition for identification of each stripe. This constraint can certainly reduce most of false matching candidates. Another is the local constraint that some features on epipolar lines between both images lie in constant order. The local constraint can reduce almost all of false matching candidates which the geometric constraint does not happen to remove. From the experimental results, we confirmed effectiveness of the above constraints.

This method may be used not only for static but also for dynamic scenes because the depth information can be obtained at one TV frame time.

ACKNOWLEDGEMENT

We would like to thank Prof. S.Tsuji for valuable discussion.

REFERENCES

[1] Moravec,H.P."Obstacle Avoidance and Navigation in Real World by A Seeing Robot Rober.", CMU-RI-TR-3, 1983.

[2] Marr,D., Vision, Freeman, San Francisco, 1982.

[3] Baker,H. "Edge-based Stereo Correlation." Proc. of Image Understanding Workshop, 1980, pp.168.

[4] Shirai,Y. "Recognition of Polyredrons with A Range Finder." Pattern Recognition,4,1972,pp.243.

[5] Gennery,D.B."Stereo-camera Calibration.",Proc. of Image Understanding Workshop, 1979, pp.101.

[6] Kasai,T. et al.,"Measurement System of 3-D MotionUsing A Pair of PositionSensing Detector Cameras.", SICE, vol.19, no.12, 1983, pp.997.

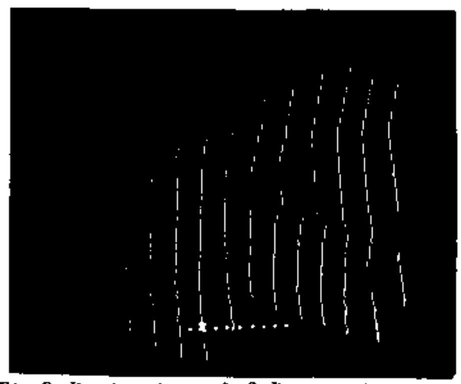
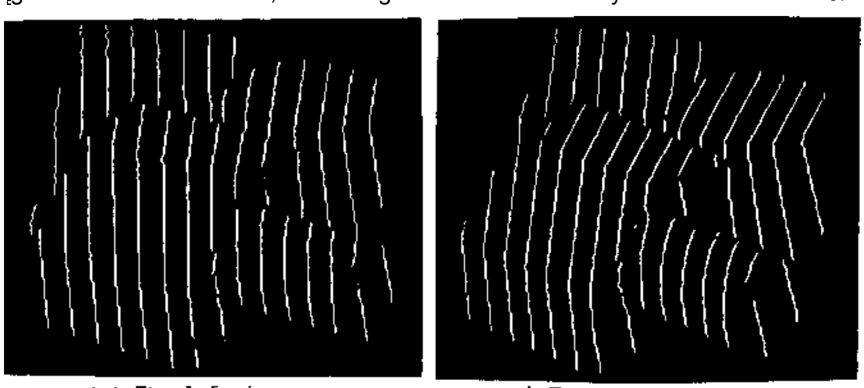


Fig.9 Projection of 3-D crossing points onto the left image.



(a) The left image (b) The right image

Fig.8 The thinning images of stripe pattern.

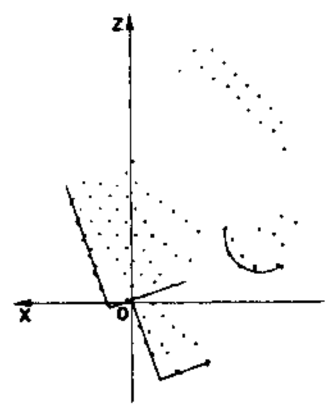


Fig.10 Top view of the scene.