A PIXEL VOTING METHOD TO RECOVER 3D OBJECT SHAPE FROM 2D IMAGES

Kenpo TSUCHIYA* Shuji HASHIMOTO* Toshiaki MATSUSHIMA**
*Waseda University **Toho University

3-4-1, Okubo, Shinjuku-ku
Tokyo, 169 JAPAN

2-2-1, Miyama, Funabashi-shi
Chiba, 274 JAPAN

ABSTRACT

The purpose of this study is to extract 3D information from 2D images passively. Though the voting technique is one of the passive methods to obtain the 3D information effectively, it cannot be applied to objects with smooth surfaces. It can measure only the feature points such as the edges which were previously extracted from each image. In this paper, we propose a new method to measure the 3D shape, which we call "a pixel voting method". In this method, the voting is simply done by using the information from the pixels of the images to measure the whole surface of the 3D object without feature extraction.

1. INTRODUCTION

Many methods have been proposed to extract 3D information from 2D images. These are grouped into two categories. One is the active sensing method that uses special purpose apparatus such as laser scanners and spatially coded projectors. Another is the passive sensing method which deal with images with no special lighting mean.

Though the active sensing method is often applied to industrial purposes, because it has better precision, there exist some problems in it. For example, the system tends to have a complex structure and the lighting condition is limited. On the other hand, the passive sensing method can be used in general cases. The binocular stereo technique is a typical passive sensing method. It is a technique which applies triangulation using two images from different viewing points. However, it has a serious problem of feature point matching that is to determine the corresponding points in each image.

The voting technique was proposed to avoid the feature point matching. If we know the camera's positions and its orientations, 3D information can be extracted by using many images from the different camera positions.[1]-[6] Since the usual simple voting technique measures only the feature points, such as edges previously extracted from each image, it is difficult to recover the shape of the curved surface or the surface where the feature points cannot be extracted.

In this paper, we propose a new method to measure the 3D shape, which we call a pixel voting method. In this method, the voting is simply done by using the information from the pixels of the images. It is not necessary to extract the feature points from the images at all, which enables us to measure the whole surface of the 3D object.[7]

2. BASIC IDEA OF THE PIXEL VOTING

The simple voting method can be summarized as follows. If we know the camera's position and its orientation, we can calculate a line for each feature point in an image connected to center of the lens. Each line goes through a real feature point on the object in the object space. We call this line the back projection line. If we have images from different viewing points, the back projection lines corresponding to the same feature point cross at the real feature point in the object space. Fig.1 shows the back projection lines crossing at the real point. Therefore, if we draw the back projection lines for many object images, and find the point where many lines cross together in the object space, the positions of feature points on the object in 3D space can be extracted without difficulties of stereo matching.

In the simple voting method mentioned above, the whole object shape cannot be measured, since only the previously extracted feature points are back projected into the object space. On the contrary, the pixel voting method, proposed here, can recover the whole object surface by using all the

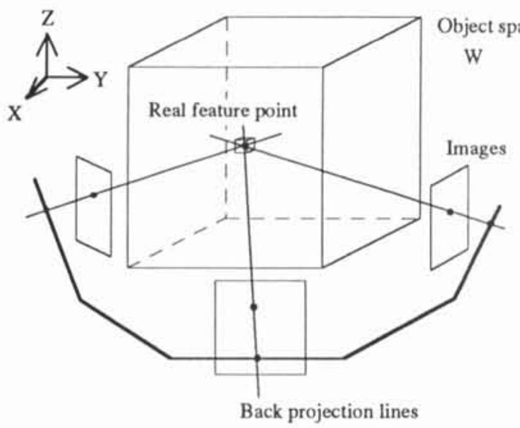
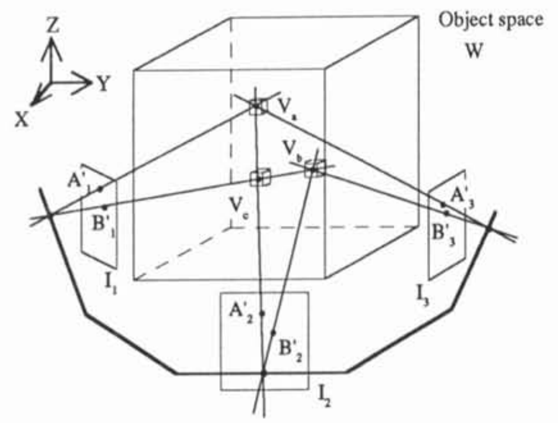


Fig. 1. The simple voting with back projection lines.



I_i : Image plane.
A'_i : Real surface point in an image whose pixel data is 100.
B'_i : Real surface point in an image whose pixel data is 80.
V_a : Voxel including the real surface point. Its assigned data should be 100, 100, 100 (from I_1, I_2, I_3).
V_b : Voxel including the real surface point. Its assigned data should be 80, 80, 80 (from I_1, I_2, I_3).
V_c : Voxel including the false surface point. Its assigned data should be 80, 100 (from I_1, I_2).

Fig. 2. The pixel voting.

pixel data of the images in a slightly different way of voting.[7]

This method is based on the idea that a point on the object surface gives the similar color information or gray level when it is observed from different viewing points. The back projection lines are calculated for all pixels of the images to find the object surface on which many back projection lines with the same pixel data cross together in the 3D space. Fig. 2 shows a typical example of the assigned data by the pixel voting method. The voxels which include the real surface points, such as V_a or V_b in Fig. 2, should have the similar pixel data from each image, while at the voxels in which the false surface point exists, the assigned data from each image have different values.

Therefore, the whole object surface can be extracted by evaluating the variances of the assigned data at each voxel.

3. THE PIXEL VOTING METHOD

Although the gray level of the object surface may change with the viewing point, we assume that it has little variation in a narrow angle range.

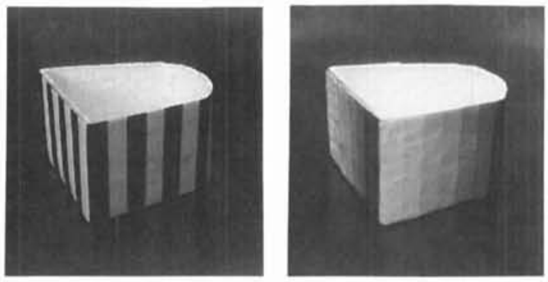
To decrease the variation of the gray level for the difference of viewing angle and the effect of self-occlusion, we will assign the pixel voting in a narrow angle range (we call it the first voting process). Since only a part of the object surface can be extracted in this process, the same process is repeated for another angle to cover the 360 degree view of the object.

If the variation of the gray level on the object surface is little, many false surface points are extracted around the real surface in the first voting process. To eliminate these false points, a second voting process is employed to vote simply (i.e. adding the value 1) to each voxel extracted in the first voting process. Finally we extract the voxels which are highly voted in the second voting.

The proposed algorithm is summarized here as follows.

1. Take N images one after another at every δ degrees of the camera angle.
2. Give the pixel voting (first voting) using images with viewing angles between θ and $\theta+\alpha$.
3. Extract the voxels in which the variance of the pixel data is smaller than the threshold value $t1$.
4. Give the second voting to the voxels extracted in Step 3.
5. Increase θ by β ($\delta < \beta \leqq \alpha$) and go to Step 2 unless $\theta \geqq 360$ degrees.
6. Extract the voxels whose value of the data voted in the second voting is larger than the threshold value $t2$.

4. EXPERIMENTAL RESULTS



(a) (b)

Fig. 3. The experimental object whose surface (a) has an obvious gray level, (b) has small variation of the gray level.

72 images are taken at every 5 degrees around the object. The image size is 512×512 pixels, and each pixel has 8 bit monochrome data. We use the gray level for the pixel data in the basic experiment. The processing parameters (that is α, β and $t1$, $t2$) are chosen empirically in the present study.

First, we used the object which has both the curved surface and the flat surface. Also, as a basic experiment, we made the object surface have an obvious gray level variation as shown in Fig. 3(a). When the object had white and black stripes, each width was about 10 mm, the object surface could be extracted quite accurately for both the curved surface and the flat surface. Fig. 4 shows one horizontal section of the object, (a) is the real one, (b) and (c) are the experimental results. Though we have many false points around the real surface after the first voting, they were restrained by doing the second voting as shown in (b) and (c). Fig. 5 shows the 3D object shape that is sliced along the z-axis to be able to see its shape easily.

However, when the variation of the gray level on the object surface was small as shown in Fig. 3(b), some false points were extracted inside of the real surface, although we

(a) (b) (c)

Fig. 4. One horizontal section for the object in Fig. 3(a). (a) Real object. (b) After the first voting. (c) After the second voting.

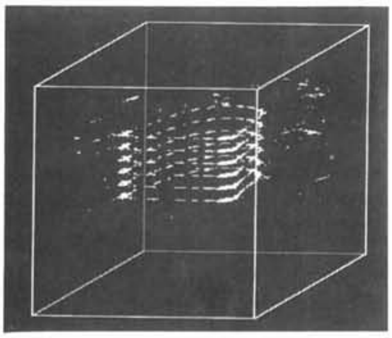
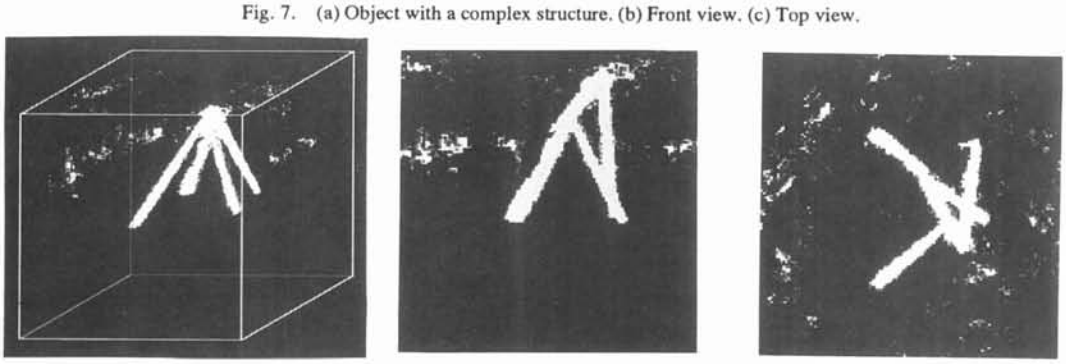


Fig. 5. Result for the object in Fig. 3(a) by doing the pixel voting.

Fig. 6. One cross section for the object whose surface has little gray level variation as shown in Fig. 3(b).

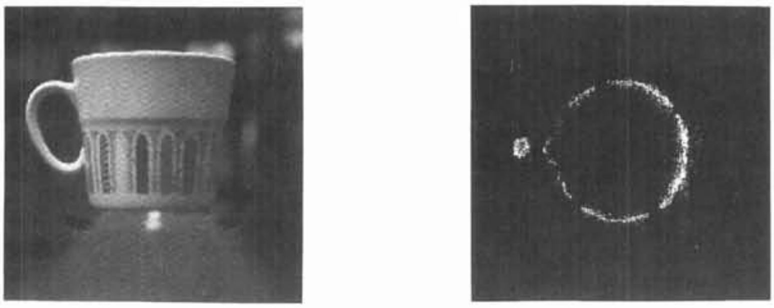
(a) (b) (c)

Fig. 7. (a) Object with a complex structure. (b) Front view. (c) Top view.



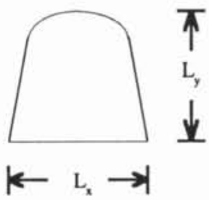
(a) (b) (c)

Fig. 8. (a) Results for the object in Fig. 7 by doing the pixel voting. (b) Front view. (c) Top view.



(a) (b)

Fig. 9. (a) The experimental object; "cup". (b) Result for (a).



	L_x	L_y
Object	81.1 mm	76.0 mm
Result	83.2 mm	77.7 mm
Errors	2.59 %	2.24 %

Fig. 10. Measurement errors for the object in Fig. 4 in the quantitative analysis.

could see the shape of the horizontal section in Fig. 6.

These experimental results prove that the pixel voting method is practical when the object surface has some variation of the gray level.

Next, we used the object which has a complex structure, so that it produce self-occlusion frequently. It was made by joining some branches in three dimensions as shown in Fig. 7. In this figure, (b) is the front view and (c) is the top view. The results are shown in Fig. 8.

From this result, we can see the validity of this method for the object with self-occlusion. Fig. 9 shows the other experimental result for the actual object; "cup".

5. CONCLUSION

The pixel voting method is proposed for extracting 3D information from 2D images. It enables the measurement of all the surface points to recover 3D object shape.

In the former experiment, this method is proved to be practical for the shaded object. The latter result shows the validity for the self-occlusion. In the quantitative analysis, measurement errors are summarized in Fig. 10.

There are some problems, for example, the objects have to be fixed and unreflecting lest the color information or the gray level of the surface should change with the viewing angle, rather long calculation time is required because all pixels of all images are back projected into the 3D object space in this method.

Though only the gray level data on the object surface are used in the present study, measurements of the object shape with smaller gray level variations by introducing the color information to suppress the false surface is planed for future research.

REFERENCES

[1] S.Kawato, "3D Shape Recovery by Octree Voting Technique", Videometrics (SPIE) volume 1820, pp.131-140(1992).

[2] T.Hamano, T.Yasuno, K.Ishii, "Structure from Motion by Voting Algorithm", IEICE Transactions on Information and Systems, Vol.J75-D-II No.2, pp.342-350(1992).

[3] T.Hamano, T.Yasuno, K.Ishii, "Estimation of Structure by Voting Algorithm with Pyramids of Sights", IEICE Transactions on Information and Systems, Vol.J76-D-II No.1, pp.50-58(1993).

[4] T.Kawahara, T.Matsushima, S.Hashimoto, "An Experimental Research on Multi-ocular 3D Vision using Back Projection Method", Proceedings of the 1992 IEICE spring conference D-659, p.7-401(1992).

[5] M.Toya, T.Matsushima, S.Hashimoto, "Multi-ocular 3D Vision using Back Projection Method", IPS Trans. No.2, pp.393-394(1993).

[6] Y.Kurimoto, T.Matsushima, S.Hashimoto, "Multi-ocular 3D Vision using Intersection of Back Projection Lines", Proceedings of the 1994 IEICE spring conference D-545, p.7-281(1994).

[7] K.Tsuchiya, S.Hashimoto, T.Matsushima, "A Pixel Voting Method for Recovery of the Cross Sections of 3D Object", Proceedings of the 1994 IEICE spring conference D-549, p.7-285(1994).