3-D Object Modeling by a Camera Equipped on a Mobile Robot

Kimitoshi Yamazaki, Masahiro Tomono, Takashi Tsubouchi and Shin'ichi Yuta
Intelligent Robot Laboratry, Tsukuba University
1-1-1 Tennodai, Tsukuba Ibaraki, 305-0001, Japan
Email: {yamazaki, tomono, tsubouchi, yuta}@roboken.esys.tsukuba.ac.jp

Abstract— **This paper describes 3-D shape modeling performed by a mobile robot. We propose a method in which the robot in the real environment acquires the accurate shape and constructs a 3-D model of an object using Computer Vision techniques under the condition that position and shape of the object is unknown. 3-D modeling is performed from image streams which are captured by a camera placed on the robot. The results show the effectiveness of our method.**

I. INTRODUCTION

A mobile robot moving in realistic world needs to observe its environment, decide its motion plan and implement it. One of the functions that is needed for the robot is to acquire the information of the environment. Recently, there are several researches on map building in unknown environment based on information obtained from sensors on the mobile robot [2][5][12]. Most of them are primarily aimed at sensing broad environment information and navigation by a mobile robot. However, they can not necessarily provide enough information to handle a specific object.

We are interested in developing an autonomous mobile robot which manipulates an object in real environment such as object pickng up. It is necessary to hold detail shape information of each object for manipulation, however it is difficult for the robot to retain shape information of all objects which exist in the environment in advance. We take such an approach that detailed model of the target object nor its position is not given in advance, but the robot system reconstructs the 3D object shape model each time the object is found. The model will be reconstructed by taking several images around the object with different viewpoints.

This paper describes how to build 3-D shape model of small object which can manipulate by a manipulator on the mobile robot. A single camera (pan-tilt camera) is placed on the robot, and image streams of the target object are captured by the camera. We propose a method of 3-D shape modeling from image streams, which is based on the various techniques of Computer Vision, such as Factorizaion, Motion Stereo and Space Carving.

Our approach is emphasized as follows:

- The robot is not provided with any position of the target, and hence the robot finds the object autonomously by itself.
- The shape model of the object is acquired while the robot moving.

- No artificial marks is attached on the target object.
- The shape modeling process is preformed using a series of images including the object. Motion model of the robot nor position based on odometry is not incorporated.

This paper is organized as follows: In the next Section, related works are described. Section 3 presents a scheme of our research and problem definition. The outline of 3-D shape modeling is described the Section 4, and detail of dense reconstruction which we propose is described the Section 5. Section 6 provides experimental results with 3-D shape model to demonstrate the effectiveness of our proposed method. Section 7 concludes this paper.

II. RELATED WORK

A large number of issues to reconstruct 3-D shape of the object from several images are contributed in the field of Computer Vision.

Factorization[1][7] and related methods depending on Epipoler Geometry are useful under the condition that correspondence of feature points are known among the images, and sparse 3-D shape and camera pose for each image are obtained. On the other hand, a method to reconstruct dense shape of the object was presented such as Dense Stereo Matching [6] and Voxel Coloring [10]. These dense shape reconstructions are assumed that all camera poses are known for all the source images. In the issues to reconstruct dense shape from image streams without any information of camera pose or motion, Sato et al.[11] constructing 3-D shape model of an outdoor scene. They estimate camera poses from marks which are assigned in one of the images by hand in advance and feature points which are tracked automatically, and integrate reconstructed data based on Multi Baseline Stereo to obtain dense shape of the object in voxel space. Sainz et al.[9] present 3-D shape modeling from uncalibrated multiple view images. It can be emphasized that a projective Factorization is performed to calculate camera poses, and afterward Space Carving is performed to recover dense 3-D shape. On the other hand, our approach in 3-D shape modeling does not assume the use of artificial marks on the object to track the feature point. The modeling is applied to real environment.

III. SCHEME AND PROBLEM DEFINITION

A. Problem definition

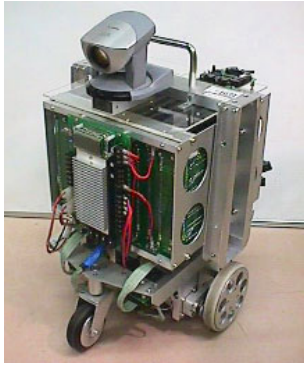Shape modeling is preformed based on image streams which are captured from several viewpoints of the camera which is

Fig. 1. Mobile Robot

equipped on a mobile robot (Fig.1).

We assume target objects as follows:

1) The object is placed on the floor.
2) Curved surface is allowed and texture on the surface of the object is essential.

The second assumption arises from the fact that texture is necessary to make correspondence of the same feature points of the object in a series of images. The correspondence is also crucial to our aim – dense 3-D shape modeling.

B. Challenges

We challenge the object modeling under the following concerns:

Object Condition
We do not give special condition to help distinguishing an object from others, nor give special marks on the object. Therefore, it is our challenge that the robot must identify one object to be modeled from other objects in the realistic environment by itself. During the process of the shape modeling, all the necessary information must be obtained or extracted from a series of images including the target object.

Error of Camera Pose
Odometry (a method of self position estimation by integrals of wheel rotation using encoders to measure) is the only way to directly obtain the pose of camera of the mobile robot. However, the camera pose estimated from odometry is not accurate enough because of several factors, such as body slip, camera angle error and missed of synchronization between odometry data and the captured image.

For these reasons the robot primarily has to detect the target object by searching in the real environment widely. After detecting the target object, the robot moves around it obtaining a detailed model of it by a robust method overcoming camera pose error.

C. Approach

Our proposed method for object modeling is divided into two steps:

1) Searching for the target object, while the robot moves in the environment.
   (*Object Finding*)
2) Modeling of the found target object
   (*Object Shape Modeling*)

In Object Finding process, the robot identifies its target object automatically while moving in the real environment. The robot moves, searching the target object and calculates the accurate position and rough shape of the target object. In Object Shape Modeling process, the robot acquires the accurate shape model moving around the target object. (A detailed description of Shape Modeling process is presented in Section 5.)

Our approach for object shape modeling is Motion Stereo based on the known camera poses which are actually estimated with respect to odometry. The precision of the reconstructed model depends on the accuracy of the camera positions. However, the camera position with respect to odometry position is considered not to be for the precision of the recovered model. We take other approach that not only the shape of the object but also the camera positions and poses are simultaneously obtained from a series of images.

IV. SHAPE MODELING OF A TARGET OBJECT

Requirements for modeling process are as follows:

1) The process must be executed in real time.
2) The process must be sequential. The motion of the robot can be changed depending on the reconstruction shape of the object by the time.

The robot only knows a 3-D part of the target object, when it finds the object in the environment. It is necessary to acquire a full and an accurate model by the modeling process. One of the purposes of this research is object manipulation (picking up an object), therefore the reconstructed shape must be highly accurate.

In this Chapter, a novel algorithm to reconstruct a 3-D model of the target object will be presented. The proposed algorithm can simultaneously acquire camera poses and the shape of the target object using image streams. We refer to the techniques of Computer Vision which are utilized in our algorithm in Appendix for the readers' convenience.

A. Outline of shape modeling from image streams

In this Section, shape modeling outline from image streams is described.

1) Extraction of feature points and tracking: As the target object is assumed to have a texture, feature points can be extracted from the captured image and tracked where they are in other images. In order to extract and track feature points, we employ KLT-Tracker[3]. Using this tracker, transition of feature points are recorded.

2) Structure from motion method: In this method camera poses are estimated from image streams while modeling the target object, because camera poses which are estimated by a mobile robot include error. The camera poses are estimated by the result of tracked feature points with KLT-Tracker and by making use of the 3-D position of the tracked feature points

obtained during the process. In the meantime, unknown 3-D position of other feature points are calculated from the estimated camera pose. Nonlinear Minimization[8] is used for estimation of camera pose, and Motion Stereo[8] is used for calculation of 3-D feature points. In some cases, Factorization method is used to calculate camera poses and 3-D position of feature points simultaneously.

3) Dense shape reconstruction of target object: Dense shape reconstruction is possible if the camera positions and poses are known when images area captured, taking a method such as stereo matching. We propose a novel method to reconstruct dense shape based on correspondence of feature points of the same object in a series of images together with known camera poses. A detailed discussion is presented in Section 5.

B. Shape modeling algorithm

In this Section, shape modeling algorithm (fig.2 is described. There are three phases in our algorithm, Preliminary Phase, Sequential Phase and Final Phase.

Prerequisite Condition: Taking advantage of the object finding process, the approximate shape of target object and relative positions between the robot and the object are known. In addition, known feature points positions are continuously tracked in subsequent image streams.

Preliminary Phase: Initial value of sequential phase are generated in this phase.

Both, 3-D position of the feature point obtained from the Object Finding process and camera poses obtained by odometry include errors. To re-calculate these parameters, Factorization method is adopted. After that, error of Factorization method is compensated by Nonlinear Minimization, then both accurate camera poses and 3-D positions of several feature points are acquired simultaniously.

In our research, most of feature points are tracked before the robot moves around the target object. Factorization method can be employed in the condition of no occlusion. Moreover, it is a powerful method because the error of feature tracking and digital artifacts are reduced by averaging data. Despite Factorization includes linear approximation error, it can be used for good initial value of Nonlinear Minimization.

Sequential Phase: In this phase, the whole shape of the target object is acquired by image streams around the object. As often as new image is obtained, following processes are implemented.

- Camera pose is estimated from feature points, where 3-D position is known and is tracked between images. Nonlinear minimization is employed.
- 3-D position of new extracted feature points are calculated from estimated camera pose. Motion Stereo is employed.
- Dense reconstruction based on Affine invariance and voting in voxel space are performed.
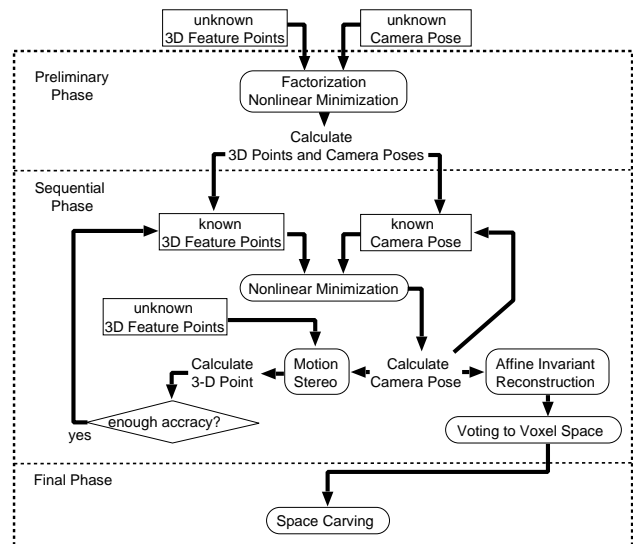

Fig. 2. Shape Modeling Algorithm

When the robot captures images moving around the object, feature points will interchange frequently as the viewpoint of the object changes. In this situation, Motion Stereo is effective because it can calculate the 3-D position of each feature point. Camera pose is estimated from several feature points which have already been calculated, then Nonlinear Minimization is performed.

In the sequential phase, camera pose is estimated from known feature points. Because feature tracker produces tracking error, RANSAC is used to select good feature points for estimating camera poses.

Motion Stereo is performed to calculate 3-D position of unknown feature points by the estimated camera poses. However there are several feature points which could not obtain accurate 3-D position. To discriminate such feature points with inaccurate 3-D position, each obatained 3-D position of the feature points must be reprojected onto the image plane. If the reprojected feature points has good match with the original feature point, it is still accepted as a proper and knwn feature point. In the meantime, dense reconstruction based on Affine invariance and voting its result in voxel space are performed every several images.

In this phase, each process is fast and reconstruction of the target object can be performed sequentially every time a new image is captured. This enables a robot to plan next camera viewpoint to acquire better shape model from the reconstructed shape in realtime.

Final Phase: In this phase, whole shape of the object which is acquired in the sequiential phase is compensated. First, after voting in voxel space, such voxels whose voted number exceeds a threshold are chosen. Second, Space Carving is performed from the voxels and the final shape model is acquired.
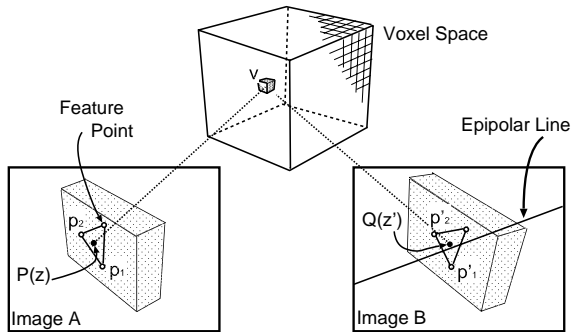
There is an error of Affine approximation in the object

Fig. 3.   correspondance with Affine invariance

Fig. 4.   Space Carving

shape with voxels which are reconstructed in the sequential phase. In the final phase, 3-D object shape is compensated by Space Carving which is based on Photometric nature. In Space Carving, the radiance of a pixel is checked in several different viewpoints from several different images. Final shape model is acquired by the selected voxels which are consistent in all referred images.

V. ACQUIRING DENSE SHAPE

In this Section, a nobel algorithm for dense reconstruction is presented. This method is assumed that camera pose and correspondence of feature points between images are known.

A. Dense reconstruction from two images based on Affine invariance

Fundamentally, dense 3-D shape reconstruction is achieved by a correlation base stereo, where all the correspondence of the pixels in the two images must be established and camera pose of the two images are known. Making correlation is computational power consuming process and takes long time in correlation. In this Section, a smart and faster algorithm for dense 3-D reconstruction is proposed, where sparce correspondence of the feature points which is already obtained in the sequential phase is fully utilized. The crucial point of the proposed approach is to make use of Affine invariance in finding a presumed pixel Q in Image B in Fig.3 when a pixel P in Image A in Fig.3 is assigned in a triangle that is formed by the neighbor three feature points.

The Affine invariance parameter α and β is defined as follows:

$$\mathbf{z} = \alpha(\mathbf{p}_2 - \mathbf{p}_1) + \beta(\mathbf{p}_3 - \mathbf{p}_1) + \mathbf{p}_1 \tag{1}$$

Where $\mathbf{z}$ is a coordinate vector of pixel P and $\mathbf{p}_n$ ($n = 1, 2, 3$) are feature points in Image A in Fig.3. As $\mathbf{z}$, $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$ are known, α and β are invariant against the change of the viewpoint from image A to image B proves that a "presumed" corresponding pixel Q and its coordinate vector $\mathbf{z}'$ in image B in Fig.3 to the pixel P in image A is represented with following equation:

$$\mathbf{z}' = \alpha(\mathbf{p}_2' - \mathbf{p}_1') + \beta(\mathbf{p}_3' - \mathbf{p}_1') + \mathbf{p}_1' \tag{2}$$

Where $\mathbf{p}_n'$ ($n = 1, 2, 3$) in image B in Figure.3 is a corresponding feature point to $\mathbf{p}_n$ ($n = 1, 2, 3$) in image A respectively.
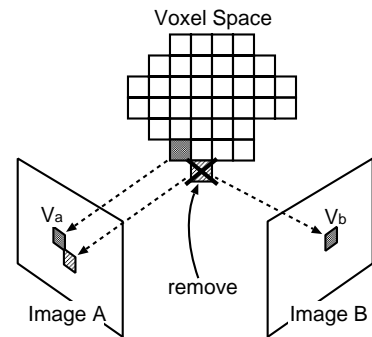
In this approach, we must take notice that the proposed approach employs 2 dimensional Affine transformation, and the presumed point Q contains an error in the coordinate vector $\mathbf{z}'$. Therefore, it is necessary to verify the point $\mathbf{z}'$ with the criteria as follows:

- Distance between presumed pixel $\mathbf{z}'$ of Q to epipolar line in image B in Fig.3 from image A is within a certain threshold.
- A radiance of the pixel Q in image B in Fig.3 is same with the pixel P in image A.

After making the pixel to presumed pixel correspondence in the two images, conventional Motion Stereo method yields dence 3-D object shape reconstruction. Avoiding conventional correlation matching of the pixels in the two images provides computation time merit in the reconstruction process.

B. Integrating into the voxel space

In this step for the shape reconstruction, each 3-D shape which is obtained by a pair of consecutive images in the image streams are voted and integrated into a voxel space. Because of the reconstruction method by Affine inbariance includes 2-D Affine approximation which is described in Section 5.1, reconstruction error will become larger at a scene which has long depth or a target object which has rough surface. There will be phantom particles in shape from the reconstruction by two images. Therefore, voting is effective method to scrape redundant or phantom particles off and to extract a real shape.

To cope with 3-D shape error originated from Affine transformation, not only the voxel just on the surface of the reconstructed 3-D shape but also the adjacent voxels are also voted into the voxel space. After finishing the vote from all the reconstructed shapes originated from the image stream around the target object, voxels that has the large voted number exceeding the threshold are saved and other voxels are discarded. The result of reconstruction is presented by a group of voxels which has thickness in its shape.

C. Shape compensation depending on Space Carving

Space Carving[4] is applied to the voxel group obtained by the process described in Section 5.2 to brush the 3-D shape up. The Space Carving itself is originally invented to reconstruct a 3-D shape of an object as voxel group from the multiple color images captured from arbitory viewpoints. Conventionally,

contour of the object is first extracted from several images which were captured from several viewpoints, and produce voxel group based on Silhouette Constraints. Next, the voxels are chosen from the result of reprojection of the voxel groups onto images (Photometric Constraints).

In the frameworks of this research, there are many objects not only the target object because the robot moves in the real environment, and it is difficult to extract only the contour of the target object. For this reason, application of conventional Space Carving to our framework will not provide significant accuracy in the reconstructed 3-D shape. However, the use of reconstructed dense shape based on Affine invariance and integration into the voxel space described in Section 5.2 already provides the approximate shape model as voxel group. This must be suitable to an initial shape for Space Carving, which provides advantages to improve the accuracy of the reconstruction and to save computation time because the shape is already carved roughly provided by the initial shape.

The voxel group which is obtained from dense reconstruction based on Affine inavariant and voting in voxel space are scraped under Photometric Consistency. In the case of Fig.4, a certain voxel V are projected on an image A, and a radiance of projected figure v_a are recorded. In an image B, as well as image A, a radiance of projected figure v_b of V are checked, and if the difference between v_a and v_b is large, the voxel V will be erased from the voxel group. The final object model is produced from the process which is repeated in several images.

No geometrical approximation is included in Space Carving because it is based on Photometric Constraints. For this reason, if the object has texture, the error of shape from Affine invariance could be compensated.

VI. Experiments

The results of shape modeling are shown in Fig.6 and Fig.7. Good result is obtained.

Experiments are performed with a mobile robot "YAM-ABICO" which is equipped with a Canon VC-C4 pan-tilt camera. For inspecting effectiveness of the proposed shape modeling algorithm, the target object is placed on the floor, and the robot is moved there in a circular trajectory with radius 800mm. Images are captured by the image processing module HITACHI "IP7500". In this time the reconstruction process is performed in offline, robot only takes pictures while it is moving. Our experimental image processing is performed on a PC Pentium4 at 2.8GHz CPU.

From image streams which are captured around the target object, the preliminary phase and the sequential phase are performed, and 1000 counts of 3-D feature points and 200 camera poses are reconstructed. In the final phase, 3-D shape model is build from voxels with 2mm in thickness. Finally, the number of voxels of the target object becomes approximately 60000. In the sequential phase, processing time of one image is as follows: camera pose estimation based on RANSAC takes 0.7 sec of computation and dense reconstruction based on Affine invariance takes 0.25 sec of computation. This proves that the process of one image can be executed within 1 sec. In

Former Image
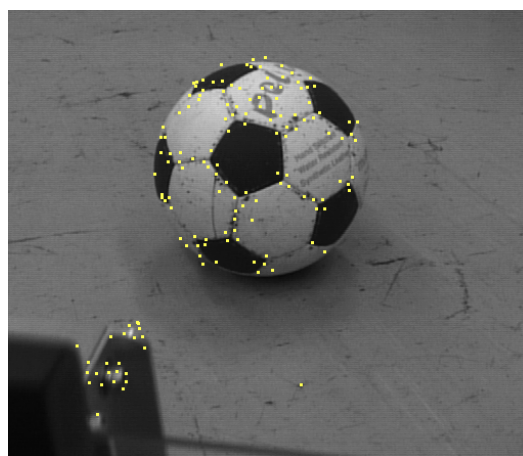
Feature Tracking

Fig. 5. Feature tracking

the final phase, Space Carving is executed in 180 seconds of computation. In terms of accuracy, to the case of real spherical object (Fig.5), the error of the shape model is 10mm at most.

VII. Conclusion

In this paper, we presented the shape modeling method of the target object for the mobile robot. Experimental results with 3-D shape model are demonstrated and prove the effectiveness of our method.

As a future work, more faster processing is necessary yet, and to realize the shape modeling by the all in one system on the autonomous mobile robot.

References

[1] C. Tomasi , T. Kanade: "Shape and Motion from Image Streams under Orthgraphy: a Factorization Method", IJCV,9(2), pp.137–154, 1992.
[2] J. J. Leonald and H. F. Durrant-Whyte: "Simultaneous Map Building and Localization for an Autonomous Mobile Robot", Proc of IROS'91, pp.1442–1447, 1991.
[3] J. Shi , C. Tomasi: "Good Features to Track", CVPR, 1994.
[4] Kutulakos K. and Seitz S.: "A Theory of Shape by Space Carving", IJCV, 38(3), pp.198–218, 2000.
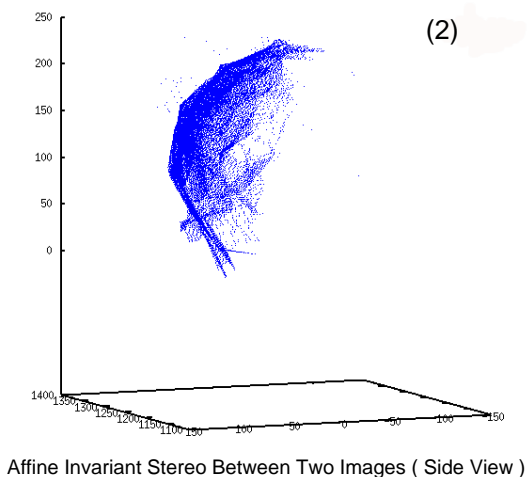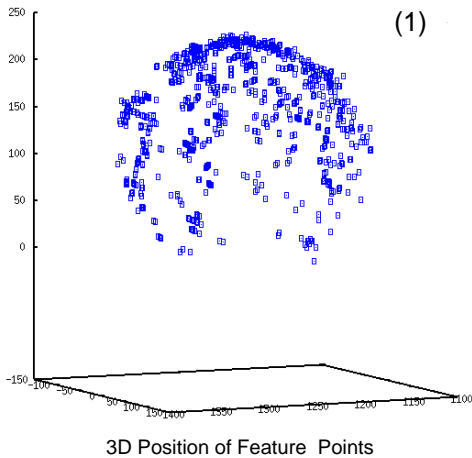[5] L. Iocchi, D. Mastrantuono, D. Nardi: "Visually Realistic Mapping of a Planar Environment with Stereo" ISER 2000.
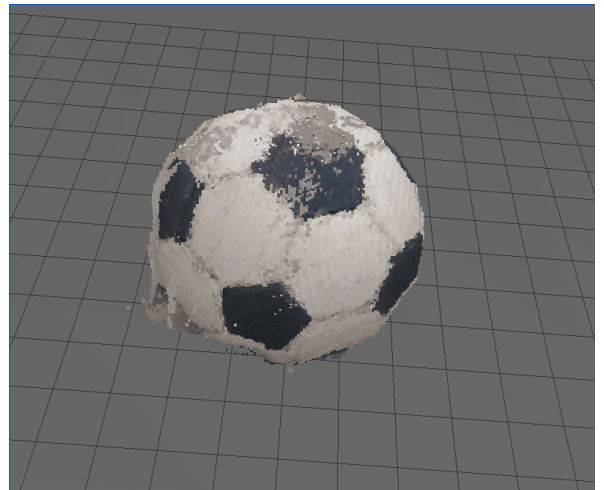
(1)

250
200
150
100
50
0

3D Position of Feature Points

(2)

250
200
150
100
50
0

Affine Invariant Stereo Between Two Images (Side View)

Fig. 6. Midterm result

Fig. 7. 3-D modeling result

[6] M. Okutomi. and T. Kanade: "A Multiple-Baseline Stereo", PAMI, VOL15,NO.4, pp.353–363, 1993.
[7] Poelman C.J and Kanade T.: "A Paraperspective Factorization Method Shape and Motion Recovery" tech. report CMU–CS–93–219, Computer Science Department, Carnegie Mellon University, 1993
[8] R. Szeliski and S.B. Kang: "Recovering 3d shape and motion from image streams using non-linear least squares", Journal of Visual Computation and Image Representation, vol. 5, pp. 10–28, 1994.
[9] Sainz M., Bagherzadeh N. and Susin A.: "Carving 3D Models from Uncalibrated Views.", CGIM 2002, pp.144–149., 2002.
[10] Seitz M. and Dyer C.: "Photometrealitic Scene Reconstruction by Voxel Coloring", IJCV, 35(2), pp.1063–1073, 1999
[11] T. Sato , M. Kanbara, N. Yokoya, H. Takemura: "Dense 3-D Reconstruction of an Outdoor Scene by Handhelds-baseline Stereo using a Hand-held Video Camera", IJCV,Vol 47, No.1-3, 2002.
[12] S. Thurn , W. Burgard , and D. Fox: "A Probabilistic Approach to Concurrent Mapping and Localization for a Mobile Robot", Machine Learning and Autonomous Robots,31/5, 1-25, 1998.
[13] K.Yamazaki, M.Tomono, T.Tsubouchi and S.Yuta: Sequential Shape Reconstruction of the Target Object for Picking up by a Mobile Manipulator (In Japanese), 21th,RSJ, 2003.
[14] Z. Zhang, R. Deriche, O. Faugeras and Q.-T. Luong.: "A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry", Artificial Intelligence Journal, Vol.78, pp.87–119, 1995.

APPENDIX

Motion Stereo

Motion Stereo is available when correspondence of the same feature point in the two images and camera poses are known. A 3-D position of the feature point by the linear function is calculated as follows:

$$C = \|\mathbf{X} - s_1 \widetilde{\mathbf{m}}_1\|^2 + \|\mathbf{X} - s_2 \mathbf{R}\widetilde{\mathbf{m}}_2 + \mathbf{T}\|^2, \qquad (3)$$

$\widetilde{\mathbf{m}}_1$ and $\widetilde{\mathbf{m}}_2$ are extended vectors which is the coordinates of an correspondent feature point. $\mathbf{X} = (X, Y, Z)$ is the 3-D position of feature points, $\mathbf{R}$ is relative camera poses $\mathbf{T}$ is relative position between the two images.
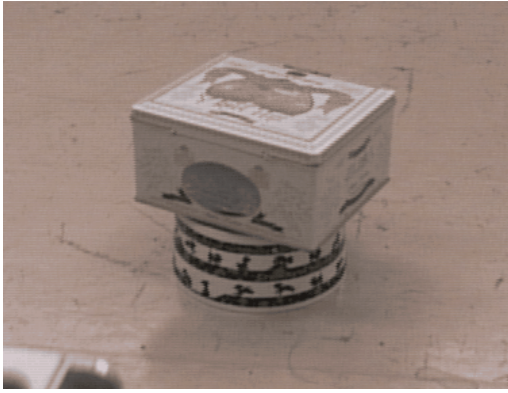
Note that this method can be executed fast and be calculated every point at a time. However, the result of reconstruction are influenced by the camera poses if they have errors.
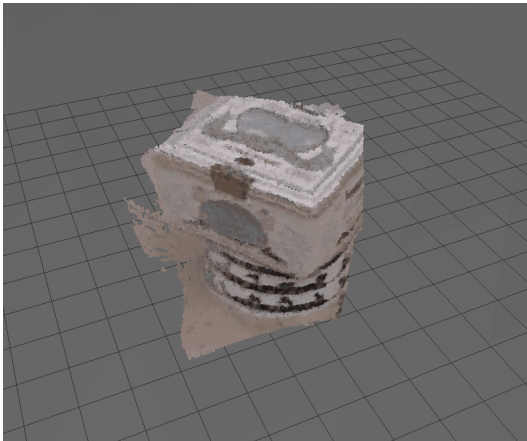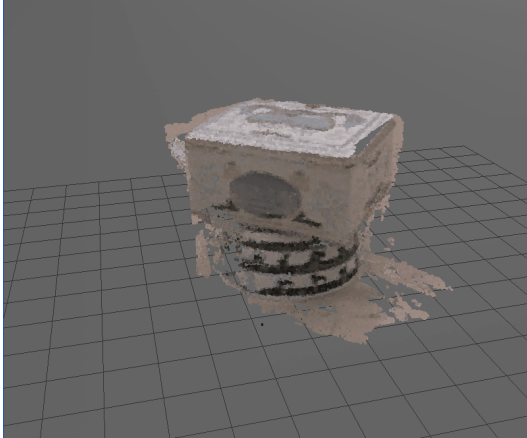
A Factorization Method

A Factorization method can calculate simultaneously camera poses and 3-D position of several feature points from a matrix which is composed several feature points in images.

When there are F images and there are P feature points which are corresponded in all these images and feature point

Former Image

3D Modeling Result

Fig. 8. Other example

p in frame f is described (u_{fp}, v_{fp}), a matrix $\mathbf{W}$ are defined from the several feature points.

$$\mathbf{W} = \begin{bmatrix} u_{11} & \cdots & u_{1P} \\ & \ddots & \\ u_{F1} & \cdots & u_{FP} \\ v_{11} & \cdots & v_{1P} \\ & \ddots & \\ v_{F1} & \cdots & v_{FP} \end{bmatrix} \qquad (4)$$

This $2F \times P$ matrix is named as measurement matrix. A matrix $\mathbf{W}'$ is composed from which the $\mathbf{W}$ is reduced an

average value of its each row. The camera pose represented three orthonormal basis of world coordinates is defined by $\mathbf{i}_f = (i_{fx}, i_{fy}, i_{fz})$, $\mathbf{j}_f = (j_{fx}, j_{fy}, j_{fz})$ and $\mathbf{k}_f = (k_{fx}, k_{fy}, k_{fz})$, besides the coordinate of points in the 3-D space are define at $\mathbf{s}_p = (s_{px}, s_{py}, s_{pz})$, a formation comes as follows: (1) $\mathbf{i}_f$ and $\mathbf{j}_f$ are orthgonal, (2) $\mathbf{i}_f$ and $\mathbf{j}_f$ are unit vectors. It is possible that the relation W is decomposed by M and S,

$$\mathbf{W}' = \mathbf{MS} \qquad (5)$$

The matrix $\mathbf{M}$ ($2F \times 3$) presents camera poses and the matrix $\mathbf{S}$ ($3 \times P$) presents the 3-D position of feature point s.

This method is useful because the reconstruction of camera poses and 3-D positions of feature points are performed simultaneously from only image streams. However there are some drawbacks if we apply it to in our research as follows:

- The complicated algorithm are needed when an occlusion occurs, because this method fundamentally processes many feature points which go on to be tracked from many images all together. In our research, because the images are captured from around the object, feature points are frequently interchanged.
- The result of reconstruction has error which is caused from linear approximation.

Nonlinear Minimization

The result of Motion Stereo when camera poses include the error and camera poses and shape of the object calculated by the Factorization are not enough accurate. In this research, Nonlinear Minimization are performed from evaluation function as follows:

$$C = \sum_{j=0}^{2} \sum_{i=0}^{P} \left(\frac{\mathbf{r}_{xj}^T \mathbf{m}_{ij}}{\mathbf{r}_{zj}^T \mathbf{m}_{ij}} - \frac{X_i + t_{xj}}{Z_i + t_{zj}} \right)^2$$
$$+ \sum_{j=0}^{2} \sum_{i=0}^{P} \left(\frac{\mathbf{r}_{yj}^T \mathbf{m}_{ij}}{\mathbf{r}_{zj}^T \mathbf{m}_{ij}} - \frac{Y_i + t_{yj}}{Z_i + t_{zj}} \right)^2 \qquad (6)$$

Where $\mathbf{r}$ is column vector of rotation matrix $\mathbf{R}$, t_x, t_y and t_z are elements of translation vector. A 3-D position of feature points are represented at X, Y and Z.

In our experimental system, the robot moves on the floor with the condition that the camera on the robot has pan-tilt angle. There are four variant camera parameters, which are tilt angle θ of the camera, pan angle ϕ of the camera and movement elements of the robot (d_x, d_y).

The variables of the optimization are the number of feature points (the coordinates of (x, y, z)) in addition four camera parameters. It is the fact that processing time is varied from the number of feature points.