Three-dimensional reconstruction of environment using factorization method and EPI analysis

Kamijima, Shigeharu, Saito, Hideo

Shigeharu Kamijima, Hideo Saito, "Three-dimensional reconstruction of environment using factorization method and EPI analysis," Proc. SPIE 4572, Intelligent Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision, (5 October 2001); doi: 10.1117/12.444177

Event: Intelligent Systems and Advanced Manufacturing, 2001, Boston, MA, United States

Three-dimensional reconstruction of environment using factorization method and EPI analysis

Shigeharu Kamijima and Hideo Saito†

Department of Information and Computer Science of Keio University
3-14-1 Hiyoshi Kohoku-ku, Yokohama, 223-8522, Japan

ABSTRACT

We propose a method for three-dimensional(3D) reconstruction of environment from image sequence taken with handy video camera. The convention EPI analysis is a popular method for 3D reconstruction from motion camera image sequence, but the camera needs to be moved linearly with constant velocity. We propose a method for enabling the EPI analysis to be applicable to image sequence taken with a handy video camera of unknown motion. The proposed method removes irregular shakings by hands from input image sequence by using the factorization method and tracking feature points. We apply the convention EPI analysis to the corrected image sequence as if which were taken by camera moved horizontally with constant velocity. The method provides 3D points of the feature points in the scene. In addition, we estimate surface of environmental structure from the 3D points obtained with the method. In the surface estimation, the 3D position of the feature points can be adjusted by matching the texture on the input images. Finally, the 3D model represented by triangle meshes is reconstructed. By rendering the corresponding image region to each mesh, the proposed method made it possible to reconstruct 3D scene with more reality than traditional one.

Keywords: 3D Reconstruction of Environment, Factorization Method, EPI Analysis, Motion Image Processing

1. INTRODUCTION

Recently, ideas such as virtual art museum or cyber shopping are matter of concern by diffusion of the Internet. To realize such ideas, construction of realistic cyber space is one of important subjects. Presently, the 3D model of the cyber space is mostly constructed by manual operations, which require skilled human's many works. It is needed to automatically reconstruct the realistic 3D model from images.

There are many researches, which automatically reconstruct 3D structures of observed objects from image sequence.[6,5] Most of these researches deal with a single object, and can't reconstruct large-sized objects such as buildings or rooms. The EPI analysis[7] is one of popular method for automatic reconstruction of 3D structure of observed scene from image sequence. The key idea of EPI analysis is the fact that a line on an EPI provides the 3D position of a feature point corresponding to the line. Since a line segment on EPI is sufficient for providing 3D position, every feature point does not have to be captured in all the images in input image sequence. Therefore, such EPI analysis is suitable for reconstruct large-sized objects of which some partial segments can be captured in images. However, it is hard to apply the EPI analysis to image sequence taken with handy video camera, because it requires a horizontal camera motion with uniform velocity.

In this paper, we propose a new method for EPI-based 3D reconstruction from handy video camera sequence. Since our input is image sequence taken with handy video camera, it contains naturally irregular shakings from hands. It is obviously hard to apply the EPI analysis to the input. To overcome the problem, we remove the irregular shakings from the input sequence and generate a corrected image sequence, which match the constraint of the EPI analysis. We use the factorization method and tracking feature points for this correction. Applying the EPI analysis to the corrected image sequence, we are able to obtain 3D structures of the scene.

Many of these researches dealing with EPI aim to reconstruct structure of observed scene with 3D feature points. In addition to this, our purpose is to reconstruct structure of observed scene in surface with texture. We also propose

E-mail address
Shigeharu Kamijima: sigeharu@ozawsa.ics.keio.ac.jp
Hideo Saito: saito@ozawa.ics.keio.ac.jp
†PRESTO, Japan Science and Technology Corporation(JST)

a method that automatically recovers surface and texture from the 3D feature points obtained by the EPI analysis. The surface is composed of triangle meshes with texture. As a method for triangulation, a Delaunay triangulation is the most well-known method. We applied the method to our 3D points model by the EPI analysis, but the result was not so good. Although the Delaunay triangulation is effective in dense points, our 3D points model is composed of sparse 3D points. Therefore we propose a special triangulation to our 3D points model. The triangulation is determined by EPI's properties, such that points reconstructed by EPI analysis are dense in perpendicular direction to movement, and are sparse in parallel one to movement.

Some results of experiments demonstrate that our proposed method can reconstruct surface model of 3D structure of large-sized scene such as room interior from the image sequence, which contains irregular shakings.

2. PROPOSED METHOD

In this section, we briefly review the factorization method and the EPI analysis, which are the well-known method and very important methods for the proposed method.

2.1. The Factorization Method

The factorization method was developed by Tomasi and Kanade.[1] It is a robust and efficient method for accurately recovering both the shape of an object and the camera's motion from image sequence. It requires at least 4 points tracked in 3 frames. The accuracy and robustness are achieved by applying the singular value decomposition(SVD) to a large number of images and feature points, and by directly computing shape without computing the depth as an intermediate step. Since the factorization method developed by Tomasi and Kanade assumes an orthographic projection model, it cannot deal with notion of the distance from the camera to the object. Presently, for this problem, the factorization method, which assumes a weak perspective model or a paraperspective one, is developed.[2,8]

Tracking P feature points for F frames, the result of tracking, the camera's motion and the shape of an object can be represented by Eq. 1.

$$W = MS + T\,[1 \cdots 1] \tag{1}$$

where W is the $2F \times P$ measurement matrix whose elements indicate the coordinates of tracked feature points at image coordinate system. M is $2F \times 3$ motion matrix whose elements indicate camera's motion at world coordinate system. S is the $3 \times P$ shape matrix whose elements indicate the positions of feature points at world coordinate system. T is the $2F \times 1$ translation vector whose elements indicate the camera's position at world coordinate system. Tomasi and Kanade position the world origin at the center of mass of the object.

To obtain a registered measurement matrix W^*, $T\,[1 \cdots 1]$ is subtracted from W. W^* is factorized into a registered motion matrix $\hat{M}$ and a registered shape matrix $\hat{S}$ by the SVD. Using the constraints of motion matrix M that is called the metric constraints, the decomposition is only determined up to a linear transformation. As a result, the shape of an object and camera's motion are recovered from an image sequence.

2.2. The EPI Analysis

EPI analysis proposed by Bolles et al. is a method, which reconstructs 3D structure of observed scene from image sequence.[7] It assumes perspective projection model, which cannot be handled by the factorization method. The EPI analysis assumes that the camera is moved horizontally with constant velocity, and the captured scene is static. Objects move faster as it gets closer to the camera. A distance from camera to the object is relative to its velocity in the image sequence. EPI can be generated from image sequence by arranging the images in direction of time and cutting them along epipolar lines. Then EPI is generated as the cross section of them, which is illustrated in Fig. 1. A point in the scene corresponds to a track in the EPI, which is a straight line. The slope of the track corresponds to the depth from camera to the point. Therefore 3D position of the point is calculated by Eq. 2, 3, 4.

Occlusions can be detected easily in EPI because a track with gentler slope must occlude tracks with sharper slope in their intersections. In this way, the EPI analysis is suitable for global reconstruction and often used for large-sized scene.

$$X = (u \cdot l + n) \cdot d' \tag{2}$$

$$Y = v \cdot l \cdot d' \cdot \frac{k_x}{k_y} \tag{3}$$

$$Z = l \cdot k_x \cdot F \cdot d' \tag{4}$$

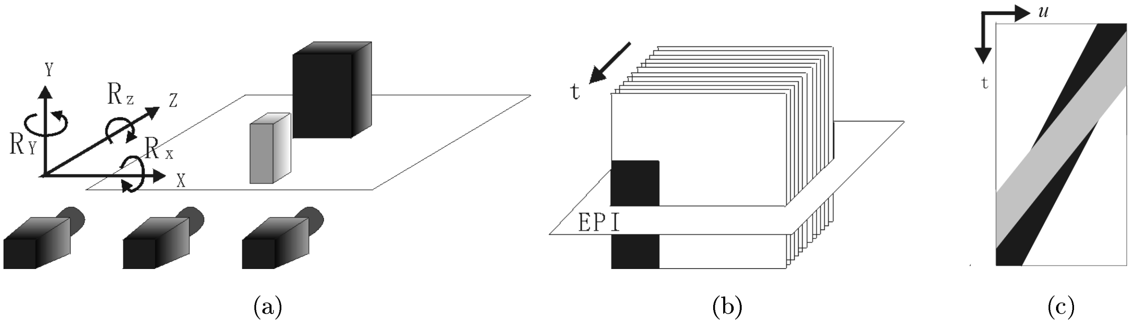


Figure 1. The process to generate EPI. (a)the scene of taking a image sequence, (b)the images arranged in direction of time, (c)the generated EPI

In Eq. 2, $(X,\ Y,\ Z)$ is 3D position of the point at world coordinate system. n is the frame number. d' is pitch, with which camera is moved. $(u,\ v)$ is the 2D point at image coordinate system. F is focal length. l is the slope of the track on the EPI. k_x/k_y is aspect ratio.

3. CORRECTION

Since our input is image sequence taken with handy video camera, it contains naturally irregular shakings from hands. Since the camera isn't moved horizontally with constant velocity, it is hard to apply the EPI analysis to the input image sequence. In this section, we correct the input image sequence so that the corrected image sequence can be applied the EPI analysis. This process is divided into two stages. In the first stage, we correct the image sequence by camera's direction which is obtained by the factorization method. In the second stage, we correct the image sequence by camera's position, which is obtained by tracking feature points. For this correction, feature points in observed scene are tracked twice in our proposed method. We call these feature points as "guiding point", because these points are the guide to correct the input image sequence. In both corrections, we deal with image sequence by dividing it with some blocks depending on the results of tracking.

3.1. Correction Of Camera's Direction

In this process, we correct the input sequence by camera's direction as if the directions of camera's optical axes for all frames are uniform. We use the factorization method for the correction. The factorization method basically recovers a shape of an object and camera's direction from image sequence. However the factorization cannot be applied to the feature points that vanish in some images because of frame-in/out or occlusion. Therefore it is not suitable to reconstruct large-sized objects such as room interior.

We first decide the guiding points in the first frame in the image sequence, and track the guiding points. While the tracking process, some guiding points will be lost because of occlusion or frame out. As described in Sec. 2.1, the factorization method requires at least 4 points tracked. The tracking continues as long as the factorization is applicable. When the tracking finished before the last frame in the image sequence, the new guiding points are selected in the finished frame of the previous tracking, and the new guiding points are tracked. Repeating this process until the last frame in the sequence, the parameters of the camera's direction in all frames are obtained. However the obtained parameters are relative values from the beginning frame in each tracking block. Since the finished frame of the previous tracking block is the same as the beginning frame of the following one, the parameters of camera's direction in all frames can be fixed to the bases of the first frame in the sequence. Using the parameters calculated by Eq. 5, 6, 7, we apply the projective transformation to each input image as if the directions of camera's optical axes in all frames are parallel.

$$R_X = arccos\frac{k_{fz}}{\sqrt{k_{fx}^2 + k_{fy}^2 + k_{fz}^2}} \qquad (5)$$

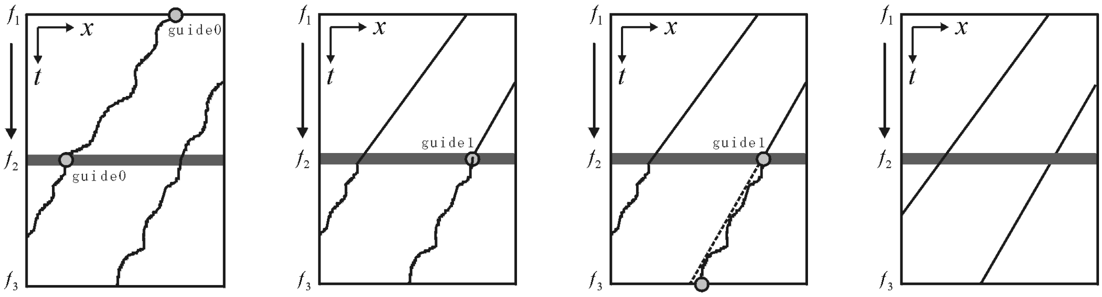
$$R_Y = arccos\frac{\sqrt{k_{fx}^2 + k_{fz}^2}}{\sqrt{k_{fx}^2 + k_{fy}^2 + k_{fz}^2}} \quad (6)$$

$$R_Z = arccos\frac{i_{fx}cos(pan) + i_{fz}sin(pan)}{\sqrt{i_{fx}^2 + i_{fy}^2 + i_{fz}^2}} \quad (7)$$

where (i_f, j_f, k_f) is camera's direction in fth frame.

3.2. Correction Of Camera's Position

In Sec. 3.1, the directions of all camera's optical axes in all frames is parallel. In this stage, we additionally correct the image sequence obtained in the Sec. 3.1, so that the image sequence can be assumed as was taken with the camera horizontal with constant velocity. A process of the correction is illustrated in Fig. 2.



(a) step 1 (b) step 2 (c) step 3 (d) step 4

Figure 2. The process to correct camera's position

In the same way of camera's direction correction, we first decide the guiding points in the first frame in the corrected image sequence and track the guiding points. We obtain a guiding point, which is tracked for a longest time (guide 0 in Fig. 2(a)). Applying the least square method to the tracking results, we can detect a line. When image sequence is just taken horizontally with constant velocity, a point in the scene must correspond to the straight line in the EPI from the image sequence(left straight line in Fig. 2(b)). We correct the images from the first frame to the finished frame of the tracking as the guiding points lies on the line. When tracking has finished before the last frame in the image sequence, we set the new guiding points in the finished frame of the previous tracking, and track the new guiding points. We obtain a guiding point, which is tracked for a longest time(guide 1 in Fig. 2(b)).

In the same way of camera's direction, the consistency between the previous tracking block and the following one is needed. We track the guiding point in reverse from the frame in which the guiding point is decided. Since the images already have been corrected in the previous tracking, the track must be a straight line. We compute the parameter of the track by applying the least square method to the tracking results(right straight line in Fig. 2(b)). We correct the images, in which the new guiding point was tracked forward, as the new guiding point lies on the line with the parameter of the line obtained by reverse tracking. We are able to obtain the corrected image sequence as if which were taken with camera moved horizontally with constant velocity. We notice that the EPI after applying the correction is obviously better than the EPI before applying the correction for EPI analysis(see Fig. 3).

4. RECONSTRUCTION

In Sec 3, we explain the way of generating corrected image sequence as was taken by camera moved horizontally with constant velocity. We are able to obtain a 3D points model from the corrected image sequence by applying the EPI analysis. Next we estimate a suitable surface and texture from the 3D points model and EPI property.

(a) (b)

Figure 3. EPI (a)before and (b)after applying the correction

4.1. 3D Points Model

We apply the EPI analysis to the corrected image sequence, which was obtained in Sec. 3. Computing the 3D positions in Eq. 8, 9, 10, we obtain 3D points model of feature points in the observed scene.

To detect lines in EPI, we use both the Hough transform and the least square method. We roughly estimate the parameters of the lines by the Hough transform. The results of the Hough transform are discontinuous values up to sampling space. We apply the least square method to edge pixels which are around the line. By using both methods, the results become continuous values.

$$X = \frac{1}{N}\sum(u \cdot l + n) \tag{8}$$

$$Y = v \cdot l \tag{9}$$

$$Z = l \cdot F \tag{10}$$

where all parameters of Eq. 8, 9, 10 is the same as the parameters of Eq. 2, 3, 4. Since d' affect only the scale of reconstructed 3D model, we assume $d' = 1$. We also assume that $k_x = 1$, $k_y = 1$ for simplification of the problem. In Eq. 8, X is averaged for accuracy.

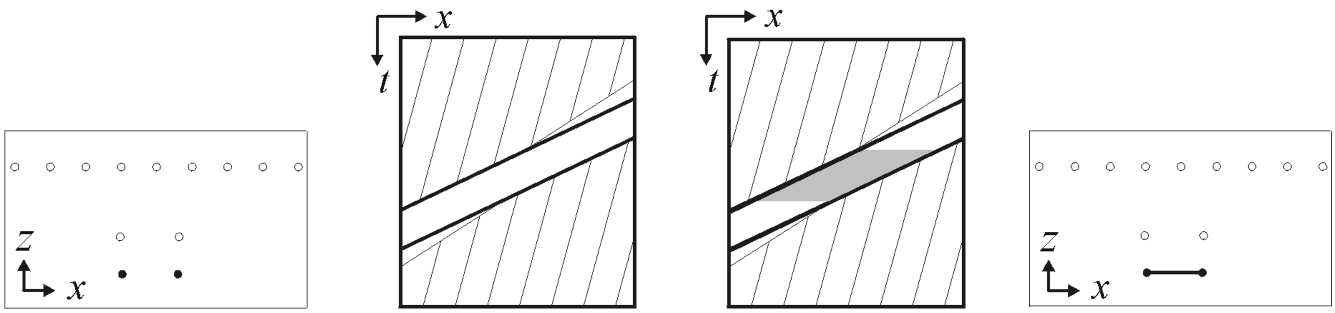
4.2. Surface Model

To generate more realistic model, we estimate the surface of the observed object. In many cases, the surface model is a triangulation of the feature points. The well-known method, which creates a suitable triangulation, is the Delaunay triangulation.[4] The Delaunay triangulation basically generates triangle meshes by connecting the neighboring points. However, connecting neighboring points do not always recover the surface of the object in the case of reconstructed points via EPI analysis, because the density of the 3D feature points is not homogenous. Therefore, the Delaunay triangulation is not suitable to our case.

We develop the triangulation method that takes account into the property of EPI analysis. The triangulation is divided into two stages. In the first stage, we estimate horizontal links of 3D feature points on the certain epipolar plane by referring relations of the tracks on the EPI. In the next stage, 3D feature points are vertically connected with the neighboring points on the epipolar plane. In the course of the process, we adjust 3D positions of the feature points by referring a consistency between the results in the fist stage and texture. The accuracy of the vertical links is improved by this adjustment.

4.2.1. Horizontal Links

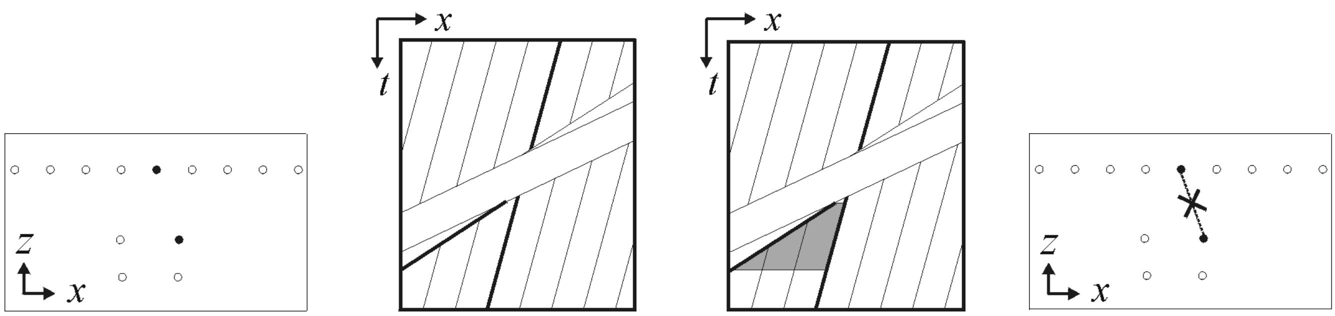
In this stage, we estimate horizontal links of 3D points on certain epipolar plane. We explain a case, in which horizontal link is possible. Now we consider a horizontal link of two black points, which are illustrated in Fig. 4(a). The tracks of the black points in the EPI are illustrated with the thick lines in Fig. 4(b). We consider a region that is caught between the tracks existing jointly in time. The region is illustrated with gray in Fig. 4(c). As described in Sec. 2.2, a track with gentler slope must occlude tracks with sharper slope in their intersections. When another lines doesn't intersect the region, We judge that it is possible to link the points as illustrated in Fig. 4(d).



(a) reffered points (b) tracks of the points (c) region of the points (d) possible link

Figure 4. The possible case.

We next explain a case, in which horizontal link is impossible. As the same as the case that the horizontal link is possible, we consider a horizontal link of two black points(Fig. 5(a)), and find the region(Fig. 5(c)). Since another lines are contained in the region, we judge that it is impossible to link the black points(Fig. 5(d)). By repeating the process for all 3D points, we are able to detect the horizontal links.



(a) reffered points (b) tracks of the points (c) region of the points (d) impossible link

Figure 5. The impossible case.

4.2.2. Adjustment Of 3D Position

Since the observed objects are large-sized such as room interior, the EPI contains many noises. The Hough transform often responds to them and regards a bunch of them as a straight line. There are many 3D points that don't actually exist. To reduce the effects of these noises, we adjust positions of these points by referring the consistency between horizontal links and texture. This adjustment causes accuracy of vertical links to us in the next stage.

We consider three points that are linked horizontally. When the center point is far from two points at ends, we regard the center point as an object of adjustment. We project the two points that are at both ends into each corrected image, in which three points exist, by Eq. 11, 12.

$$u = \frac{F}{Z}(X - nd') \tag{11}$$

$$v = \frac{F}{Z}Y \tag{12}$$

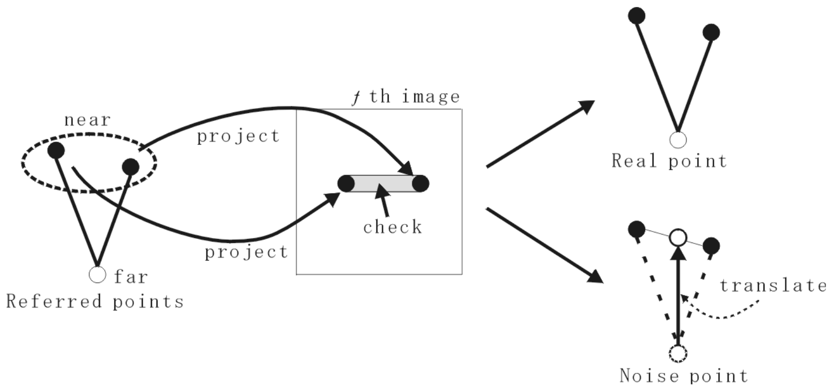


Figure 6. The process of the adjustment.

We register the largest value of horizontal gradient between the projected points in each image. We compare with the parameters of center points by EPI analysis with by applying the least square method to these pixels. If the position of the center point is correct, both the parameters must be near values. If they are not, we judge the 3D points of the noises and translate linearly the center point to a position between the two points.

(a) frame 1

(b) frame 81

(c) frame 161

Figure 7. The input image sequence taken in the scene, in which three boxes are.

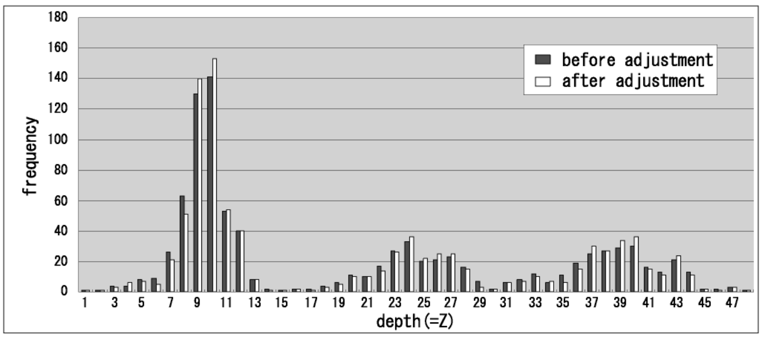


Figure 8. The comparison of the accuracy of 3D points model before and after applying the adjustment to 3D points model from the input image sequence illustrated in Fig. 7.

We examine effects of the adjustment. We took an input image sequence with handy video camera in the scene, in which three boxes are(see Fig. 7). We compared the histogram of depth for all the reconstructed points as shown in Fig. 8. Since the object scene consists of three planes as shown in Fig. 7, the histogram should has three peak in the ideal case. We can observe that the peak histogram of after adjustment is sharper and narrower than that of before adjustment. This demonstrates the effect of the adjustment.

4.2.3. Vertical links

To compose a surface of the points, we also need to link the points in the vertical direction, which is perpendicular to epipolar line, for generating triangle meshes. Since the 3D points recovered by the EPI analysis is dense in the vertical direction, the points on the neighboring EPI are connected as shown in Fig. 9(a).

A point, which is end point of a horizontal link on the upper EPI, is detected(black point in Fig. 9(b)). The end point is linked to its nearest point on the lower EPI(dotted line in Fig. 9(b)). According to the number of links of the end point and its nearest point, the vertical link patterns are divided into six patterns shown in Fig. 9(c)-9(h). The pattern 1 is divided into pattern 2 and 5 by comparing the direction of upper vector with the directions of two lower vectors in pattern 1. The pattern 2 is divided into patter 4 or 5 by comparing the length of two dotted line illustrated in Fig. 9(d). The pattern 3 doesn't require any process, and is divided into the pattern 4 and 5. These divisions are continued until all vertical link patterns are represented by pattern 4 or 5, so that a surface is composed of many triangle meshes.

This process is the same as bedding triangle mesh between the upper horizontal link and the lower one. The triangulation is generated up to horizontal link. Therefore the sparseness of the points, which are in parallel direction to movement, doesn't affect the result of the triangulation. However the error of 3D point's positions directly affect the reconstructed 3D surface model.

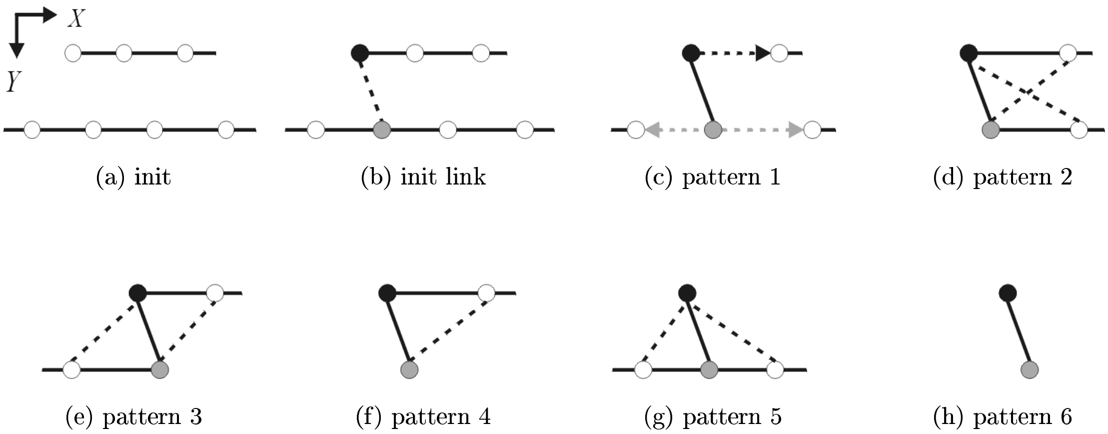


Figure 9. The six patterns of the vertical links

4.3. Texture

We estimate a texture of each triangle mesh by very simple way. We project each 3D mesh into images, in which three vertexes of mesh exist. We assume that the triangle region contains the richer information of texture as the triangle region into which the 3D mesh is projected is the larger area. Comparing the area of the triangle region in each image, we select the texture of the triangle region whose area is the largest. As a result, we are able to reconstruct the realistic 3D model, which is composed of the many 3D meshes with texture.

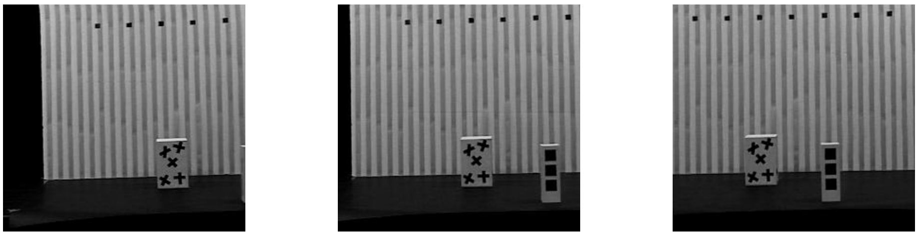
Figure 10. The 3D surface model with texture from the image sequence illustrated in Fig. 7

5. EXPERIMENTS AND DISCUSSION

We tested our proposed method on two real image sequences. The image sequences are taken with handy video camera, which is moved as horizontally with constant velocity as we can with the handy camera. Every images are 256×240 pixels and 8 bit gray scale. The image sequence contains frame in/out.

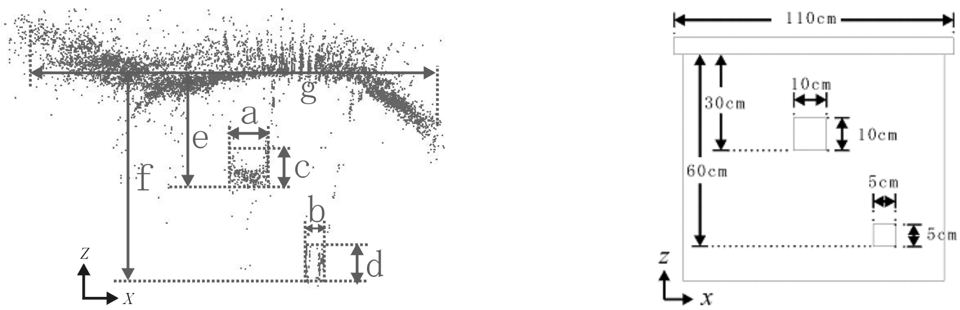
5.1. Two Boxes Model

We tested accuracy of computing depth by our proposed method in this experiment. We take the scene in which there are two boxes in front of a background with a striped pattern (Fig. 11). The image sequence is 150 frames. We have measured the position and shape of the background and two boxes.



(a) frame 1 (b) frame 61 (c) frame 121

Figure 11. The input image sequence taken in the scene, in which two boxes are.



(a) (b)

Figure 12. The measurement of each label. (a)The experimental result, (b)An actual measurement.

Table 1. Comparison between an actual measurement and an experimental result.

	a	b	c	d	e	f	g
actual measurement(cm)	10.0	5.0	10.0	5.0	30.0	60.0	110.0
actual ratio	100.0	50.0	100.0	50.0	300.0	600.0	1100.0
model ratio	100.0	50.49	96.16	98.60	283.69	560.34	1105.88
error(%)	—	0.98	3.84	97.2	5.44	6.61	0.53

When we set a length of label 'a' as basis of unit length, the length of each label in both the models is illustrated in Tab. 1. In addition, we computed the error of each experimental result. The result is going well except label 'c'. The large error of label 'c' issues from the properties both the EPI analysis and the factorization method. In the EPI analysis, information of the plane that is parallel to camera's optical axis is poor. The orthographic factorization method could not account for the scaling and position effects. Since the right box with label 'c' is at near position from camera, the recovered camera's direction is incorrect by these effects. We think that using the factorization method, which assumes the scaled orthographic model or the paraperspective model, will solve this problem. We set overlap of each tracking block 1 frame, and the number of guiding points finished tracking is the minimum number to apply the factorization method or to correct camera's position. This setting causes the speedy and easy process to me, and the decline of the accuracy and robustness of the correction to me. It is an important subject to estimate the suitable number of the overlap and guiding points finished tracking.

The error of the line detection in EPI also causes the error of computing the depth. In the case that camera is moved slowly, even if the amount of change in slope of the line is a little, the one of change in depth is large. Therefore the error strongly affects the accuracy of the 3D points model. For the problem, we use both the Hough transform and the least square method as described in Sec. 4.1. However the results obtained by this process are not enough values. The more robust method for the line detection is required.

5.2. The Room Interior Model

In this experiment, we actually apply the proposed method to the image sequence of room interior (Fig. 13). We take the image sequence of a corner in the room. The image sequence is suitable to check accuracy of the 3D points model reconstructed by our method because there is enough difference as to the depth. The image sequence is 400 frames.

The 3D surface model with texture, which reconstructed by our proposed method, is illustrated in Fig. 15. We are able to recognize some objects in scene, such as a blackboard, a washstand and a door. However the model accuracy is not sufficient. The main reason of the insufficiency is the lack of the points, which are in parallel direction to movement, and usually must reconstruct. The triangulation in the EPI analysis is difficult because of the sparseness of the points, which are in parallel direction to movement. Our triangulation covers the difficulty with the denseness of the points, which are in perpendicular direction to movement. Therefore, the lack of points is a serious problem. The solutions of this problem are two ways. One is the way that detects the lacked points, and interpolates these points. The other is the way that improves our triangulation method. The former must be relatively effective and easy because the way only detects a corresponding edge in the reconstructed 3D points model with an edge in the scene. Since it is difficult to improve our triangulation method while maintaining its advantages, the latter is not better solution. Our triangulation method doesn't consider the length of link and links only two points in an EPI and its neighbor EPI. Therefore our triangulation method is able to divide the process into the six patterns and to obtain the result with a few holes. It is difficult to obtain the surface with a few holes under the situation, in which the Deaulnay triangulation method is also not so effective.

The result of our triangulation depends on the accuracy of the 3D points model. Therefore it is effective in the result of the triangulation to improve the accuracy of the 3D points model. For example we adjust the 3D points model by using the consistency between each mesh and its texture, such as we presently adjust the position of the reconstructed points by using the consistency between the horizontal link and its texture in EPI.[3]

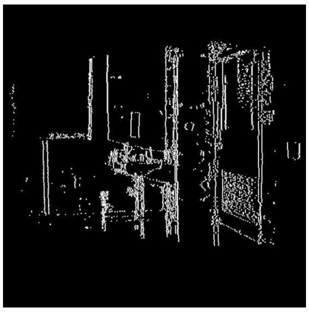
(a) frame 1

(b) frame 101

(c) frame 201

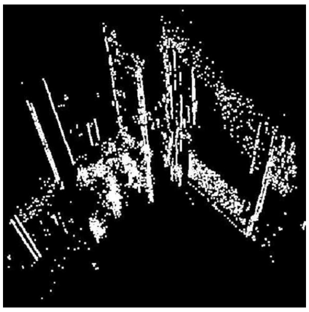
(d) frame 301

Figure 13. The input image sequence taken actually in room interior.



(a) front

(b) over



(c) diagonally over

Figure 14. The result of 3D points models, which were reconstructed by our proposed method.

Figure 15. The 3D surface model with texture from the image sequence illustrated in Fig. 13

6. CONCLUSION AND FUTURE WORK

The EPI analysis is the well-known method, which reconstruct 3D structure model of the scene automatically from image sequence. Since the camera needs to be moved linearly with a constant velocity for applying the convention EPI, it is hard to reconstruct 3D structure of the scene easily from image sequence taken with handy video camera. In this paper, we proposed the improved method, which is able to reconstruct 3D structure of the large-sized scene such as room interior automatically and easily. Before applying the convention EPI analysis to image sequence with irregular shakings from hands, we have generated the corrected image sequence as if which were taken by camera moved in uniform velocity and horizontally. The correction was done by applying the factorization method and tracking feature points.

Many of these researches dealing with EPI aim to reconstruct structure of observed scene with 3D feature points. 3D points obtained by the EPI analysis are dense in perpendicular direction to movement, and are sparse in parallel direction to movement. It is hard to apply the Delaunay triangulation to the 3D points because of the properties. We proposed the special triangulation method to the 3D points model obtained by the EPI analysis. Our proposed method made it possible to reconstruct 3D scene with more reality than convention EPI analysis.

Our future work will be to improve our proposed triangulation method. By taking account into consistency between a triangle mesh and its texture, we will adjust the 3D positions of points obtained by the EPI analysis and refine the initial triangulation. We try to apply our proposed method to the larger or outdoor environment, such as buildings.

REFERENCES

1. C. Tomasi, T. Kanade, "Shape and Motion from Image Streams Under Orthography: A Factorization Method," I JCV, vol.9, no.2, pp. 137-154, Nov.1992.
2. C.J. Poelman and T. Kanade, "A Paraperspective factorization method for shape and motion recovery," IEEE Trans. PAMI, Vol.19, no.3, pp.206-218, 1997.
3. D.D. Morris and T.Kanade, "Image-Consistent Surface Triangulation," CVPR 2000, Vol. 1, pp.332-338, June. 2000.
4. Jonathan Richard Shewchuk, "Delaunay Refinement Mesh Generation," Technical Report in CMU, CMU-CS-97-137, 1997.
5. M. Pollefeys, R. Koch, M. Vergauwen, L. Van Gool, "Automated reconstruction of 3D scenes from sequences of images," in *ISPRS Journal of Photogrammetry & Remote Sensing*, pp. 251-267, 2000.
6. Olivier Faugeras, Luc Robert, Stephane Laveau, Gabriella Csurka, Cyril Zeller, Cyrille Gauclin, and Imad Zoghlami, "3-D Reconstruction of Urban Scenes from Image Sequences," in *Conputer Vision And Image Understanding*, pp. 292-309, 1998.
7. R. Bolles, H. Baker, and D. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion," IJCV, Vol.1, pp.7-55, 1987.
8. T. Morita and T. Kanade, "A sequential factorization method for recovering shape and motion from image streams," IEEE Trans. PAMI, Vol.19, no.8, pp.858-867, 1997.