MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

13—30 Model Driven 3-D Shape Reconstruction by Matching Synthesized Images with Actual Images

Hirokazu MATSUI * School of Eng. Nagoya University

Yoshihiko NOMURA † Dept. of Mechanical Eng. Mie University

Seizo FUJII ‡ Dept. of Mechanical Eng. Daido Institute of Technology

Abstract

In this paper we propose a 3-D shape reconstruction method from multi-viewpoint images under the condition that the parameters such as poses of light sources and a moving camera are known. It is based on a principle of a model-synthesized image matching in which the shape of an object model is changed so that the model-synthesized images are matched with the multi-viewpoint input images. A 3-D target object is modeled as a triangular polygon. For matching, it is deformed by moving the vertices with a non-linear least squares method, and, if necessary, it is changed by dividing a triangle into some triangles and by merging some triangles to a triangle.

keywords Shape reconstruction, Image matching, Multi-viewpoint, Model-based vision , Polygon model

1 Introduction

This paper proposes a 3-D shape reconstruction method from multi-viewpoint images. The method has been extended from that for the 2-D object presented by the authors[1]. We use a polygon model composed of triangles, which is a kind of wire-frames. A triangular polygon sustains geometric consistency, even if each vertex moves independently. There are some researches on stabilizing the estimation of a shape for the wire-frame. In a research[2], the global shape is stably deformed by using a finite element method, which is a kind of vibration analysis. On the other hand, model-synthesized image matchings like the proposed method take enormous computational cost in the case that the number of vertices becomes a lot. Because the number of the synthesized images are proportional to the number of parameters for the representation, when making difference images. Considering the computational cost together with the stability, we change the model at first roughly and then gradually in detail. We apply a non-linear least squares method to the estimation to cope with a problem: all the vertices positions, or unknown parameters, interfere with one another. Furthermore, we propose a method of dividing triangles and that of deleting triangles; the former enables detail representations, and the latter suppress geometrical inconsistencies and redundancies caused by the estimating process.

*Address: Furou-chou, chikusa-ku, Nagoya, 464-8603 Japan. E-mail: hirokazu@mitsuya.nuem.nagoya-u.ac.jp

†Address: 1515 Kamihama-chou, Tsu-shi, Mie 514-8507 Japan. E-mail: nomura@mach.mie-u.ac.jp

‡Address: Daido-chou, Minami-ku, Nagoya 457 Japan. E-mail: fujii@daido-it.ac.jp

2 Modeling

We explain the models of a camera, lights and a target object. As for the camera, we regard the imaging model as a perspective projection. Each pose of the camera is determined by six extrinsic parameters, i.e., the position and the orientation. The light sources are assumed to be parallel light sources like the sunshine. Each light source has two parameters, i.e., the orientation and the intensity. A 3-D target object is assumed to be covered by perfect-diffuse surfaces with uniform reflectance characteristics. We model the target object in a polygon composed of triangles, and a shape of the triangular polygon is defined by the positions of all the vertices.

3 Object reconstruction method

3.1 Principle

A 3-D target object is modeled in a 3-D polygon composed of triangles. The model polygon is changed so that each synthesized image is matched with the corresponding input one, as shown in **Fig.1**. The synthesized image is projected from the model polygon, considering (a) lights' directions and brightness and (b) the camera pose corresponding to the input image.

3.2 Evaluation function

We define an evaluation function to measure an inconsistency between a model object and a target object, which is the sum of squared residuals between the intensities of the synthesized images and the input images for all the corresponding pixels.

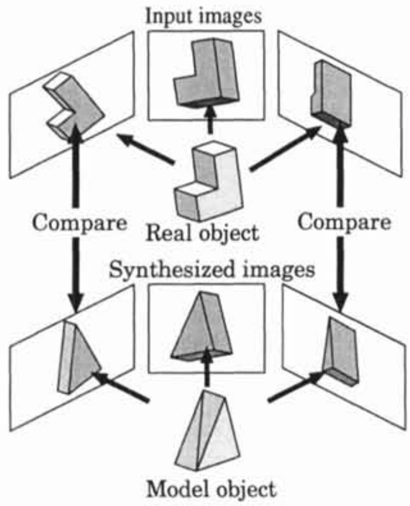


Figure 1: Concept for the proposed 3-D shape reconstruction method

The evaluation function is a function of the shape parameters for the model object, and the shape is optimized by minimizing the evaluation function value.

3.3 Optimization

3.3.1 The whole algorithm

In this section, we describe an iterative reconstruction algorithm. The model shape is changed from an initial shape by applying the processes in the priority order of (1)(2a)(2b)(3)(4). If the subprocess working at present can not improve the evaluation function value, a subsequent inferior subprocess starts to work. The above processes are iterated until the evaluation function value becomes so small that it can be regarded as noise. These processes are shown in **Fig.2**and are explained in the following subsections.

3.3.2 Deformation by Least Squares method

We use non-linear least squares method to search for analytic optimal solutions. A k-th order iterative solution is given by where A denotes a Jacobian matrix. The entry on the i-th row and j-th column in the Jacobian matrix A is calculated as the intensity change of the i-th pixel in the synthesized-images when varying the j-th parameter value by a small amount.

3.3.3 Suppression of redundant expressions

Redundant expressions for an object shape are classified into two types. One type is that the area of a triangle is very small. The other is that a couple of triangles overlaps at a part. Here, we consider a case that adjacent triangles overlap.

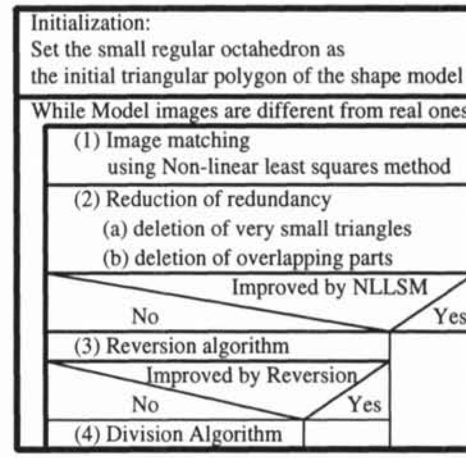


Figure 2: Proposed algorithm

(1)The square of a triangle is very small

The cases are classified into two cases as follows.

(Case 1a)

In a case that a vertex is very near to another in a triangle such as s_1 and s_2 in **Fig.3(a)**, s_1 is moved to the middle of the vertices, and s_2 is deleted as shown in **Fig.3(b)**.

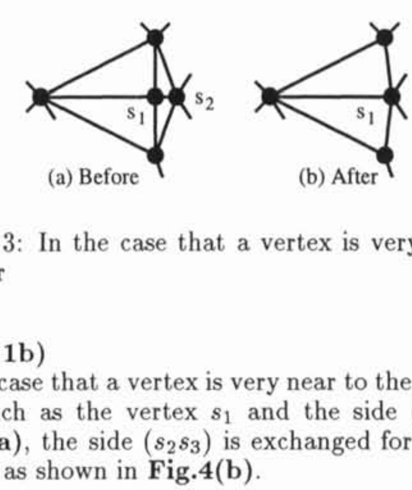


Figure 3: In the case that a vertex is very near to another

(Case 1b)

In a case that a vertex is very near to the opposite side such as the vertex s_1 and the side (s_2s_3) in **Fig.4(a)**, the side (s_2s_3) is exchanged for the side (s_1s_4), as shown in **Fig.4(b)**.

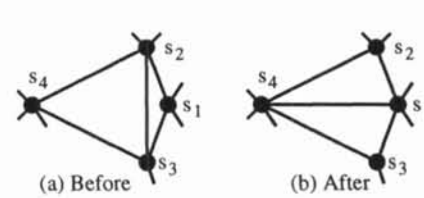


Figure 4: In the case that a vertex is very near to the opposite side

(2)A couple of triangles overlaps at a part

In the case, the overlapping parts should be deleted. They are classified into two cases as follows.

(Case2a)

In a case that a triangle includes the other such as (s_1, s_2, s_3) and (s_4, s_3, s_2), they are changed into the triangles such as (s_1, s_2, s_4),(s_4, s_3, s_1), as shown in **Fig.5**in order to suppress the redundant expressions for the overlapping part such as (s_1, s_2, s_3).

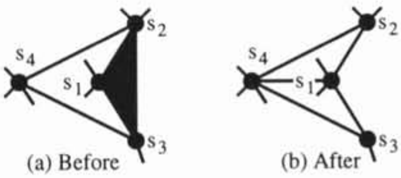


Figure 5: In the case that a triangle includes in the other

(Case 2b)

In a case that neither triangle includes the other triangle such as four triangles (s_1, s_2, s_3), (s_4, s_3, s_2), (s_1, s_3, s_5), (s_6, s_4, s_2) are changed into the six triangles such as (s_1, s_2, s_7),(s_7, s_4, s_3),(s_3, s_5, s_7),(s_7, s_5, s_1), and (s_2, s_6, s_7),(s_7, s_65, s_4), inserting the new vertex s_7, as shown in **Fig.6**in order to suppress the redundant expressions for an overlapping part such as (s_2, s_3, s_7).

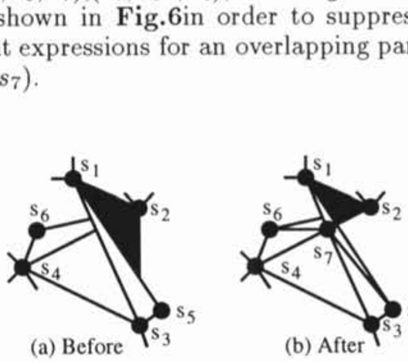


Figure 6: In the case that neither triangle includes the other

3.3.4 Reversion process

There sometimes happen a case that the principle of the shape from shading works alone. In the case, either situation in **Fig.7**could happen. And it is possible that the happened situation accidentally corresponds to a local minimum. A reversion process is applied to improve the ill situation: it checks which situation is true.

Concretely, as shown in **Fig.7**, a vertex s_1 is reversed against the center of gravity of the vertices s_2, s_3, s_4, s_5, each of which is connected to the vertex s1 by a single side.

If the evaluation function value becomes smaller than that of previous, then the position of the vertex is fixed to the reversed state, and this process will be finished. If not, the vertex returns to the previous state, and the remaining vertices are reversed by the same way.

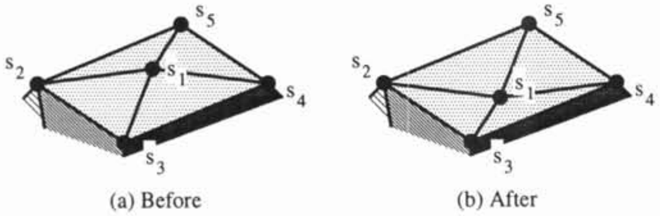


Figure 7: Reversion

3.3.5 Division process

A model polyhedron deforms from coarse to fine in oder to represent the target object in more detail. Concretely, each triangle of the model polyhedron is divided into congruent four triangles, a new vertex being inserted at the middle of each side, as shown in **Fig.8**.

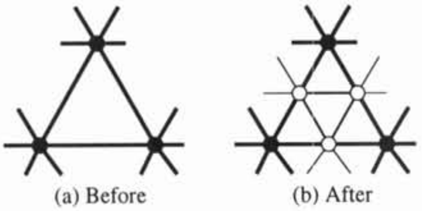


Figure 8: Division

4 Experiment

4.1 Experimental conditions

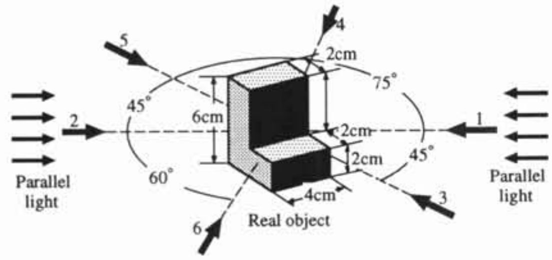
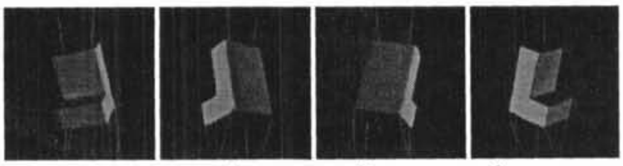


Figure 9: Real object

As shown in **Fig.9**, the target object was set an L-shaped object made of plaster. The light directions and the camera poses were fixed as shown in **Fig.9**. It is assumed that the surface is perfect-diffuse surface, and its reflection coefficient is uniform and known.

We set the target at the center of a circle with radius of 42cm, and took four pieces of input images as shown in **Fig.10**. Each input image had the size

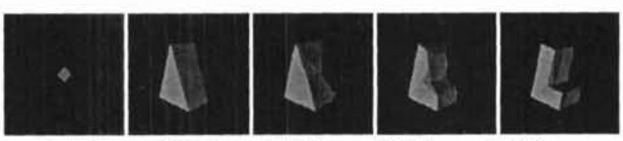


3 4 5 6

Figure 10: Input images

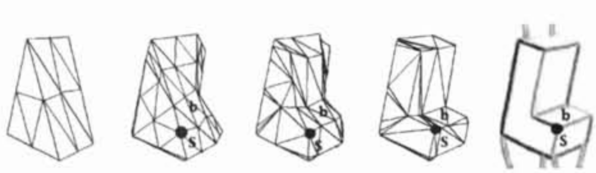
of $128 \times 128[pixels^2]$. The positions of the camera were on the circle to the center, and the orientations were shown by the vectors 3,4,5,6 in **Fig.9**. And the initial polygon model was set as a regular octahedron with six vertices on the spherical surface with radius of 1cm.

4.2 Experimental result



Initial 301st 460th 467th 500th

Figure 11: Process of estimation



301st 460th 467th 500th edge(6)

Figure 12: Deformation of the wire frame

Fig.11and **Fig.12**show the experimental result at the typical estimation stages. In **Fig.12**, an edge image of the input image 6 is added.

Through the estimations from 1st to 301st, the polygon became approximately as large as the target object since the principle of the silhouette contours works to match the outlines of the model object with those of the target object. At the 301st and 460th, the model was divided. Because no vertices could improve, though the synthesized images were still much different from the input. At 467th, the model object was expressing a part of the step in the L-shaped object, since the model had an ability to express the object in more detail through the division process at the 460th stage. Finally, at 500th the model object was almost matched with the target object. The deformation processes of the polygon model show that the proposed algorithm integrates the reconstruction principles based on the silhouette contours, the stereo, and the shape from shading, and is more effective than the method based on the individual principle.

5 Conclusion

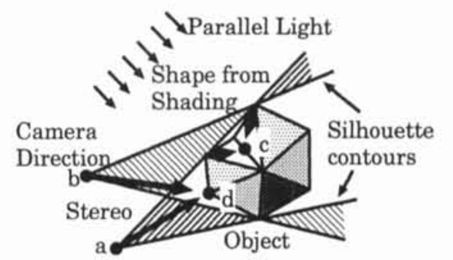


Figure 13: Fusion of stereo, silhouette contours, and shape from shading

The proposed algorithm is based on the matching of images themselves, and therefore, it fuses the three kinds of principles by the simple and natural way: the stereo, the silhouette contours, and the shape from shading. It results in the following advantages:

1. Neither programming for each of the three principles nor considering their priority is needed.

2. The solution is simpler than existing ones in a sense that it just synthesizes the images from models in the forward direction, and never try to solve inverse problems.

3. The 3-D positions of ridges and occluding-contours are measured with high accuracy owing to the stereo and the silhouette-contours.

4. The 3-D measurement process is robust in a sense that it needs no feature extraction owing to shape-from-shading.

5. The 3-D measurement is measured pixel by pixel owing to the shape-from-shading.

As for the disadvantages:

6. Many parameters are needed, e.g., poses of viewing and lighting for image synthesis.

7. Image synthesis needs extensive calculations.

References

[1] H.Matsui, Y.Nomura and S.Fujii: "Model Driven 2-D Shape Reconstruction by Matching Synthesized Images with Actual Images", Asian Conference on Computer Vision, II:787-791 December 1995.

[2] S.Stan, A.Pentland, "Closed-Form Solutions for Physically-Based Shape Modeling and Recognition", Proc. Computer Vision and Pattern Recognition, pp.238-243, June 1991.