Automatic Virtualization of Real Object Based on Shape Knowledge in Mixed Reality

Kenji Funahashi[1], Kazunari Komura[2], Yuji Iwahori[1], and Yukie Koyama[1]

[1] Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya 466-8555 Japan
kenji@center.nitech.ac.jp
http://www.center.nitech.ac.jp/~kenji/
[2] MEITETSU COM Co.,Ltd., Japan

Abstract. This paper describes a new method to realize a virtual representation of real object. A special device or a special environment is not required. This method reconstructs the object shape automatically from a few input images using the knowledge of shape for the artificial or the real object. For a polyhedron object, the proposed method can estimate the hidden shape from the front view images. The developed system can virtualize the right regular pyramid or prism from two real images taken by a digital still or video camera. The camera parameters are also estimated through the system, and roughly drawn shape of a real object is represented in virtual space with the texture mapping.

1 Introduction

In recent years, many researches have been proposed for the mixed reality [1][2]. In these researches, the camera motion is limited to get real images and the camera parameters are given manually. The main goal of this paper is the reconstruction (virtualization) of real object shape from two images without a special device or a special environment such as a range finder, a turn table, or many fixed cameras. During this research, an interactive manipulation between a real object and a virtual object is considered. An object is represented as a roughly drawn polyhedron for an interactive manipulation.

This paper shows the method to reconstruct the visible shape of an object from two real input images. The shape reconstruction is performed through the processes of feature points extraction, stereo matching and making 3D convex hull [3].

For the implementation of the above approach, the camera parameters are required. In the conventional researches, users set the camera parameters manually considering the capturing device and its environment [4]. On the other hand, this paper proposes a system in which users can get the camera parameters automatically using the general device and environment.

In general, only a visible shape of an object is recovered, and the hidden vertices and surfaces of the object cannot be recovered. Some researches requires many real images to recover the whole shape of an object [5]. Capturing many images requires large cost and much time.
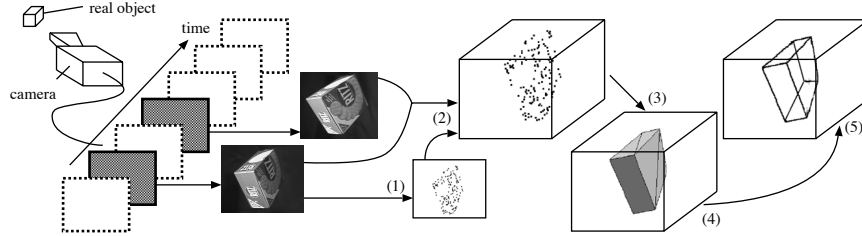
Fig. 1. Process of virtualization of real object

However, human can recognize the object shape and size from its 2D images. Furthermore, the whole shape of an object is estimated including the hidden shape from the visible surface. That is, the knowledge for the shape are used to estimate the hidden shape. In this paper, the right regular pyramid or prism are treated as a target object since these objects often exist in real world as artificial objects. The camera parameter can be obtained automatically using the knowledge and the whole shape of the object can be virtualized with a small cost.

2 Process of Virtualization

The system virtualizes a real object according to the following processes (Fig.1). A digital still camera or a digital video camera is used to take at least two images of a real object with a parallel translation along the horizontal axis of the image. Here, the translation is assumed to be small and parallel.

Two frames, i.e., the first frame and the following frame from video sequence are extracted using the condition that the sum of the difference value at each pixel between these two frames is greater than the pre-defined threshold value. The virtualization consists of the following processes.
(1) Feature point extraction from a 2D image
(2) Stereo matching using two images and feature points
(3) Making 3D convex hull of 3D feature points
(4) Camera parameters estimation based on knowledge
(5) Hidden shape estimation based on knowledge

3 Visible Shape Reconstruction

3.1 Feature Point Extraction

Feature points of a real object are extracted from an image. First, the method extracts the edges in the image by sobel filter and thinning algorithms. To extract the relevant feature points, the number of feature points is decreased by taking one feature point for 8 neighborhood pixels on the edge based on the regular sampling. Then, as shown in Fig.1(1), a set of feature points is obtained as the candidates of the vertices of the object.

3.2 Stereo Matching

The stereo matching method is used to obtain the 3D coordinates of the 2D feature points of the object (Fig.1(2)). The focal length of the lenses is f. Let the distance between the center of left and right lenses be d. The point on a real object is projected onto (x_l, y_l) and (x_r, y_r) on each image plane, which are the relative coordinate of each image. Let the pixel size be α, which represents the physical length for each pixel of the projected image (located at a distance of f from the center of the lens). Then, the coordinate of the point i is defined as:

$$X_i = \frac{x_{li}}{x_{li} - x_{ri}} \times d, \ Y_i = \frac{y_{li}}{x_{li} - x_{ri}} \times d,$$
$$Z_i = \frac{1}{x_{li} - x_{ri}} \times \frac{f}{\alpha} \times d. \tag{1}$$

First, the system searches the corresponding points in two images for above feature points with the template matching. After the stereo matching, the excessive feature points are decreased by checking those 3D coordinates.

3.3 3D Convex Hull

The method makes the 3D convex hull with the 3D coordinates of the feature points (Fig.1(3)). However, a set of feature points includes some errors caused by the digitization and the stereo matching, that is, it includes the points which are not the vertices of the object expressed by the polyhedron. So, it is necessary to integrate some patches which form the 3D convex hull for obtaining the vertices. When the normal vectors of the neighboring patches are almost similar to each other, those patches are treated as the same. Here, the feature points are represented by the equations using the parameters as d, f, and α. The solution is described in the following section.

4 Calibration Based on Shape Knowledge

4.1 Camera Calibration

To obtain the exact 3D data from 2D image data, it is necessary to know the geometric relation between 2D image plane and 3D space which including the characteristics of the geometric transformation of a camera. However, human can recognize an object shape from its 2D-image easily. It is considered that human uses the knowledge obtained through the experience. Here, some knowledge for the shape is used for the automatic camera calibration.

4.2 Automatic Camera Parameter Estimation

The 3D coordinates of the feature point of a real object are obtained through the stereo matching. The ratio parameter $\beta (= f/\alpha)$ and the parameter d are

required to get the 3D coordinate values. The parameter d is used to capture the object size, since each X, Y, Z coordinate includes d. While, the parameter β is used to obtain the relation between Z coordinate and X, Y coordinates. However, it is not always necessary to obtain the absolute size of an object, instead, the relative size and shape are enough in many cases. In this sense, it is important to estimate β as the relative shape parameter without having the exact d.

4.2.1 Pseudo Parameter for 3D Convex Hull

For the computation of the 3D convex hull, the parameters d and β need to take some values *a priori*. Here, d is set to be any pre-defined value. Procedure to set the pseudo parameter β is given below. The maximum value max_x and minimum value min_x of X coordinate are searched for an object; a_{i1}. Then the maximum value max_z and minimum value min_z of Z coordinate are searched for an object; a_{i3}. The pseudo parameter β is given by the equation; $max_x - min_x = \beta(max_z - min_z)$, since it is assumed that the object should have some volume. Using the pseudo parameter β, the (pseudo) 3D convex hull is obtained.

4.2.2 Camera Parameter Estimation

The right regular pyramid or prism is considered as a target object in the experiment. This is because many artificial objects are made by such shape parts and it is easy to estimate the symmetric and the square relation of surfaces, edges and vertices.

(1) Decision of Pyramid or Prism

Suppose that the object is a pyramid or a prism. When two or more visible triangle surfaces are observed, the object is recognized as a pyramid. Otherwise it is recognized as a prism.

(2) Extraction of Square Angle

The base surface p_1 of a pyramid or a prism is obtained, then three points which consist of two vectors with the square angle will be obtained.

prism: Find the base of prism which is not a quadrangle, since the laterals are quadrangles. In the case when the object consists of only quadrangles, it is recognized as a cuboid and any quadrangle is taken as its base.
A visible lateral is chosen as p_2. Here, p_1 is the neighbor of the p_2, and the p_1 and p_2 make a square angle. In this case, the vertex of both p_1 and p_2 is labelled as $\mathbf{v_3}$, and the vertices of either of p_1, p_2 are labelled as $\mathbf{v_1}, \mathbf{v_2}$, respectively. That is, each of $\mathbf{v_1}, \mathbf{v_2}$ makes the edge with $\mathbf{v_3}$ on the p_1, p_2, respectively.

pyramid: Find the base of pyramid which is not a triangle, since the laterals are triangles. When only two triangles are observed, the object is recognized as a pyramid. When three or more triangles are observed, the pyramid is recognized to be viewed from a side or a top.
In the case when the base is not triangle, the center point of the base is labelled as $\mathbf{v_3}$. Also any vertex of the base is labelled as $\mathbf{v_1}$, and the common vertex of the laterals is labelled as $\mathbf{v_2}$.

When only two triangles are observed, any triangle is recognized as the base of the pyramid. The center point of the base is labelled as $\mathbf{v_3}$, and the vertex of another surface is labelled as $\mathbf{v_2}$.

When three among more triangles are observed, the common vertex of the laterals is labelled as $\mathbf{v_2}$ and any other vertex is labelled as the vertex $\mathbf{v_1}$.

When $\mathbf{v_2}$ is on the edges which separates from background in 2D image, the center of the vertices of only one visible lateral is labelled as $\mathbf{v_3}$. Otherwise, the center point of all vertices except $\mathbf{v_2}$ is labelled as $\mathbf{v_3}$.

(3) Estimation of Camera Parameter

From the constraint that two vectors $\mathbf{a} = \mathbf{v_1} - \mathbf{v_3}$ and $\mathbf{b} = \mathbf{v_2} - \mathbf{v_3}$ make a square angle, the dot product $\mathbf{a} \cdot \mathbf{b}$ becomes 0. Then, the unknown parameter β can be estimated, that is, the correct Z coordinate for X, Y coordinate can be determined.

5 Hidden Shape Reconstruction

5.1 Whole Shape

When the virtualized real object is rotated in the scene, the system should show the hidden side of the object. Interaction of the real object from the hidden side is also desired. In this section, the method to estimate the hidden shape from the visible shape is described.

5.2 Hidden Shape Estimation

First, the hidden vertices of an object are estimated. Then, the hidden surface of an object is reconstructed by making 3D convex hull using all visible and hidden vertices of the object.

prism: Suppose the right prism whose lateral is rectangle. In this case, the edges on two bases of the prism are parallel each other and those edges have the same length. This vector is labelled as $\mathbf{c}$ and the vertex of visible base p_1 different from other visible laterals, is labelled as $\mathbf{v_i}$. The hidden vertex $\mathbf{v_i'}$ is obtained as $\mathbf{v_i} + \mathbf{c}$.

pyramid: When the base polygon of non-triangle is observed, all vertices of a pyramid appear in an image. When two triangles are observed, the object is recognized as a pyramid whose base is a visible triangle. When no visible bases can be observed and three or more triangles are observed, it is recognized that the pyramid is seen from a top or a side. For the view image from a top, all vertices are observed. In the above cases, estimating the hidden vertices is not necessary. If the pyramid is observed from a side, the center of the vertices contained by only one visible lateral is labelled as $\mathbf{v_m}$, while the vertices contained by only two visible lateral are labelled as $\mathbf{v_i}$. Then, the hidden vertices $\mathbf{v_i'}$ are obtained as $\mathbf{2v_m} - \mathbf{v_i}$ using the rotational symmetrical constraint.
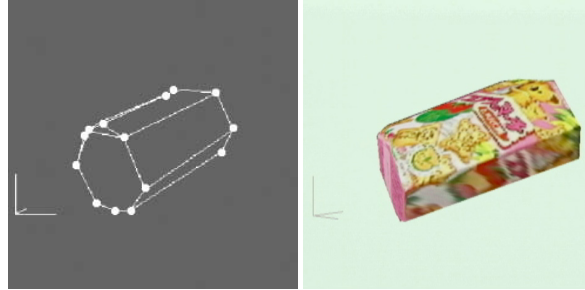
(a) left image

(b) right image

Fig. 2. Input Images (1)

(a) wire frame   (b) texture mapping

Fig. 3. Result Images (1)

6   Experimental Results

An experimental system mentioned above has been implemented using C Language. A real object is virtualized. Using a digital video camera, two input images were obtained automatically for the virtualization.

Two images with 320×240 pixels are used for the reconstruction (Fig.2). It is assumed that no other edges are observed except the object. Fig.3 shows the results with the wire frame and texture mapping images, respectively. The result is almost good but the surface integration is incomplete from the point that it contains 14 visible vertices and 7 visible surfaces.

The camera parameter β is estimated based on the shape knowledge. The correct shape is obtained without capturing the absolute size of the object. Then, the hidden shape is estimated based on the shape knowledge.

Fig.4 shows the real images used for the virtualization of a real object. This object has no texture. Fig.5 shows the result of virtualization for this object. The real objects are virtualized automatically when the system can obtain the correct vertices and surface information. It takes around 7 seconds to virtualize an object with the environment of Intel Pentium III 877MHz CPU and 256MB main memory.

7   Conclusion

In this paper, a virtualization technique of a real object at small cost was proposed. In addition to the visible object shape, the experience and knowledge are used to recognize the whole object shape. In the experiment, the right regular pyramid or prism is treated and the camera parameter was estimated automatically. The hidden vertices and surface which are not viewed in the original input images are also estimated using the knowledge for the shape. It is shown that the proposed system can virtualize a real object.
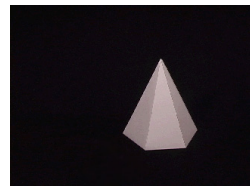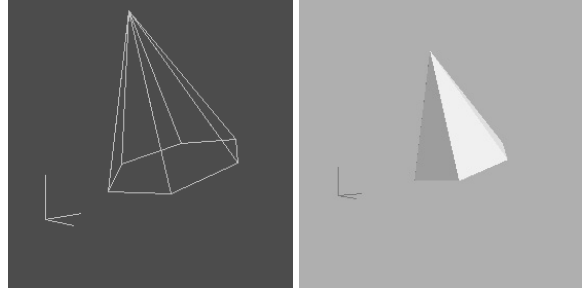
(a) left image
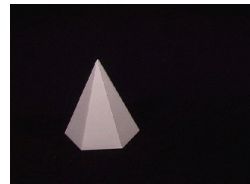
(b) right image

Fig. 4. Input Images (2)

(a) wire frame   (b) surface

Fig. 5. Result Images (2)

Good results were obtained in terms of camera parameter and hidden shape, but there are some points that need improvement. For example, the reconstructed shape contains more vertices after making 3D convex hull. A surface integration is also still incomplete. As a future subject, the above problems remain to be solved. A set of the obtained feature points includes some errors of digitization and stereo matching. The errors should be decreased and corrected more accurately. It also remains as our task that not only a prism or a pyramid, but also other shape objects including the concave polyhedron objects should be virtualized.

References

1. S. Noda, Y. Ban, K. Sato, and K. Chihara. "An optical see-through mixed reality display with a realtime rangefinder and an active pattern light source", Trans. of the Virtual Reality Society of Japan, Vol.4, No.4, pp.665–670, 1999.
2. I. Sato, M. Hayashida, F. Kai, Y. Sato, and K. Ikeuchi. "Fast image synthesis of virtual objects in a real scene with natural shading", The Trans. of the Institute of Electronics, Information and Communication Engineers, Vol.J84-D-II, No.8, pp.1864–1872, August 2001.
3. K. Funahashi, K. Komura, and Y. Iwahori. "Virtualization of real object based on shape knowledge in mixed reality", Proc. of the Virtual Reality Society of Japan the 8th Annual Conference, pp.529–532, September 2003.
4. M. Kanbara, T. Okuma, H. Takemura, and N. Yokoya. "Real-time compsition of stereo images for video see-through augmented reality", The Trans. of the Institute of Electronics, Information and Communication Engineers, Vol.J82-D-II, No.10, pp.1775–1783, October 1999.
5. Y. Nishiyama, H. Saito, and S. Ozawa. "Estimation of 3d model and object color using multiple color images", The Trans. of the Institute of Electronics, Information and Communication Engineers, Vol.J82-D-II, No.6, pp.990–1000, June 1999.