3D Research (2010) 02: 1-5
10.1007/3DRes.02(2010)1

3DR REVIEW

3-D Object Reconstruction from Multiple 2-D Images

Woo-Seok Jang and Yo-Sung Ho

Received: 22 February 2010 / Revised: 15 March 2010 /Accepted: 10 April 2010

Abstract Three-dimensional (3-D) object reconstruction from multiple two-dimensional (2-D) images is one of the most important topics in computer vision. In this paper, we review two common methods for 3-D object reconstruction: shape from silhouette and voxel coloring. The shape from silhouette method recovers the 3-D shape of an object from silhouette images. This method is based on the visual hull that is formed by intersecting volumes from silhouette cones. Voxel coloring is another reconstruction method that uses the measure of color consistency to build a 3-D model of the object. In this paper, we present a different 3-D object reconstruction algorithm, where the initial 3-D object that was generated by the visual hull is carved to represent details of the 3-D object using color consistency. In this algorithm, visibility checking is proposed to reconstruct the 3-D object well. Experimental results show that the proposed algorithm reconstructs natural 3-D models efficiently.

Keywords 3D reconstruction; shape from silhouette; voxel coloring; visibility checking

1. Introduction

Generally, applications based on computer graphics, such as virtual reality, robotics, games and animations, are based on three-dimensional (3-D) modeling. The 3-D model can be produced by skilled designers using computer graphics software. However, 3-D modeling based on computer graphics needs a lot of time and it is difficult to make detailed 3-D models. In addition, it needs a measurement process one-by-one to model 3-D objects in the real world. Since 3-D modeling by computer graphics has the above shortcomings, different approaches for 3-D object reconstruction have been proposed[1,2]

3-D object reconstruction is a task of recovering 3-D geometry and color information[1]. Various methods for 3-D object reconstruction can be classified into two categories: active method and passive method[2]. The active method generally employs a special device which can obtain more reliable depth information than the passive method. Range scanner, structured light pattern, and depth camera are commonly used in the active method. However, they are very expensive to implement and require special skills for data acquisition. Unlike the active method, the passive method uses mainly digital RGB cameras from different viewpoints. Thus, in the passive method, 3-D information is extracted from multiple 2-D color images.

In this paper, we present two different passive methods for 3-D object reconstruction and a new reconstruction algorithm. Since those methods require calibrated input images, we can use the relationship between a 3-D point and its projected point in each view[3].

The rest of this paper is organized as follows. Section 2 discusses two common methods for 3-D object reconstruction. In Section 3, we present a new 3-D object reconstruction method, which relies on volume intersection and color consistency. Section 4 shows experimental results for two data sets, and we conclude this paper in Section 5.

2. 3-D Object reconstruction methods

2.1 Shape from silhouette

The shape from silhouette method reconstructs the 3-D object by creating the visual hull from input images. The visual hull is an approximation of the 3-D object shape. In order to obtain the visual hull, after we create a silhouette image of the object from the acquired input images, we extract contour lines related to the silhouette and back-project points on the contour line into the 3-D space using camera calibration information. Back-projected silhouette images make cons in the 3-D space. An approximation of the 3-D object shape can be obtained by intersecting such cons created from multiple silhouette images. Although the visual hull is a good approximation that encloses the entire

Woo-Seok Jang (✉) • Yo-Sung Ho (✉)
Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro, Buk-gu, Gwangju, 500-712,
e-mail: jws@gist.ac.kr, hoyo@gist.ac.kr

object, details of the object cannot be reconstructed by the visual hull alone[4]. In spite of this drawback, many researchers still use the shape from silhouette method for 3-D object reconstruction because of its simplicity and fast implementation. Fig. 1 shows the process of volume intersection for the visual hull.

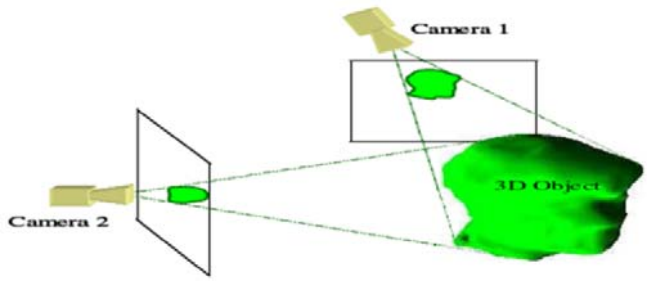

Fig.1 Volume intersection for the visual hull

The shape from silhouette method needs accurate silhouette information of the foreground objects for reliable 3-D object reconstruction. Hwang and Kweon proposed a silhouette extraction method using a statistical model and graph-cut based optimization[5]. Instead of 3-D modeling for the visual hull, a rendering method of the visual hull by arbitrary views is proposed[6]. Recently, several methods have been proposed to reconstruct the visual hull fast and model the 3-D object in real time. Ladikos proposed a real-time 3-D reconstruction system that uses a GPU-based reconstruction method to achieve real-time performance[7]. For higher accuracy of 3-D object reconstruction, some methods use combination of the shape from silhouette method and another method. As we mentioned above, the shape from silhouette method is unable to reconstruct concavities of an object, such as the inside of a bowl. Thus, the shape from structured light method is used in addition to the shape from silhouette method[8]. It can complement the concave parts of the object. Esteban and Scgnutt used the shape from silhouette method as an initial work. They also employed a multi-stereo carving technique using estimation of carving depth, added to the shape from silhouette method [9].

2.2 Voxel coloring

The voxel coloring method depends on color consistency. If the colors of an unoccluded point seen from different camera positions are the same, that point is assumed to be on the surface of the 3-D object. On the other hand, if the colors of the unoccluded point projected on the different cameras are different, that point is assumed to be located in the empty space. Fig. 2 explains how the color consistency can be used to distinguish points on the surface from points not on the surface. On the left in Fig. 2, two cameras see the same colors of the same point. Therefore, this point is on the surface of the 3-D object. However, on the right in Fig. 2, two cameras see different colors of the same point. Thus, this point is not on the surface of the 3-D object.

Seitz and Dyer introduced a voxel coloring method and suggested a constraint on the position of the camera relative to the scene[10]. It is the ordinal visibility constraint to simplify the task of resolving visibility relationship by establishing a fixed depth order enumeration of points in the scene. By extending Seitz and Dyer's voxel coloring method, Culbertson and Malzbender introduced generalized voxel coloring[11]; however, unlike Seitz and Dyer's voxel coloring method, they used images from arbitrary camera locations. Another voxel coloring strategy is based on the color hypothesis test that ensures the consistency of the projected reconstruction with the original images[12]. In this algorithm, a surface voxel list is constantly updated during reconstruction ensuring that only the minimum number of voxels is processed.

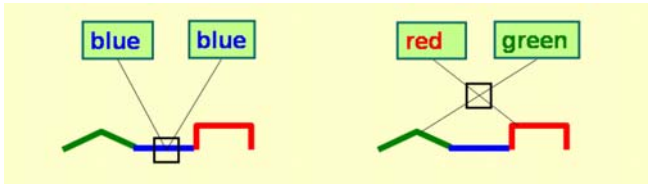

Fig. 2 Color consistency

3. Reconstruction algorithm

3.1 Visual hull method

We apply a voxel structure to the above method to represent the approximation of a 3-D object shape. The 3-D space is partitioned into small voxels. In order to detect the silhouette of the 3-D object, we project each voxel onto the image plane of each camera. If the voxel is projected on the object region of 2-D images, it is regarded as a voxel belonging to the 3-D object. Fig. 3 shows the voxel representation and the projection process from a voxel of the 3-D space onto the 2-D image plane. The 3-D silhouette of the target object is obtained using this process.

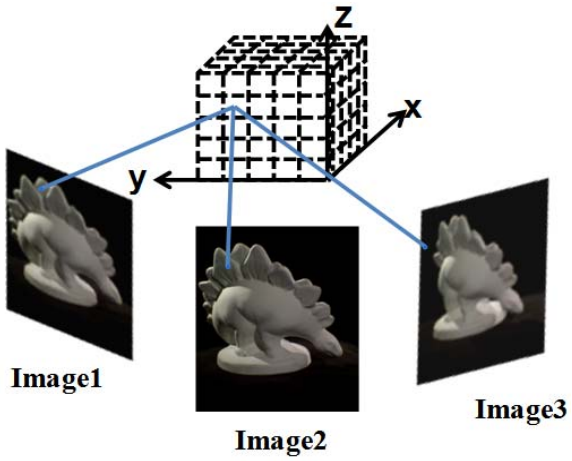

Fig. 3 Voxel structure for the initial 3-D sillouette

In general, this result does not provide an accurate silhouette of the 3-D object since this process cannot recognize the concave part of the 3-D object. Thus, it is difficult to use this result as the silhouette of the real 3-D object. However, this result can be used as the initial silhouette of the 3-D object for a high quality 3-D object reconstruction process.

3.2 Visibility Checking

The visibility checking method has several advantages for 3-D object reconstruction. It can reduce the complexity of the 3-D reconstruction process and increase the accuracy of the 3-D object. For this reason, we make a line between a point in the approximation silhouette of the 3-D object and each camera center. This line equation is carried by

$$X = [X-C] * s + C \quad (1)$$

where the vector X=[x y z]T represents the point in the 3-D object, C denotes the 3-D position of the camera center and s represents a scalar value. This line equation is useful for visibility checking. If the connected lines between one point in the 3-D object and all camera centers include other certain point in the 3-D object, we regard that this point is not visible. Thus, we exclude this point from the visible candidates of the 3-D object. If we apply this process to all points of the initial 3-D object, we can check the visibility reasonably well. The visibility checking method helps object carving and color mapping processes later.

Fig. 4 shows the process of visibility checking. In Fig. 4, A, B, C, and D are points in the 3-D object. To check whether the point A has visibility of the 3-D object or not, we examine the lines which connect the point A. The line AC1 that connects the point A and camera center C1 does not meet any points in the 3-D object. Thus, we regard the point A as a visible point of the 3-D object. Since the number of direct lines that the point B and the point C have is greater or equal to one, we regard the point B and the point C as visible points of the 3-D object. However, the line DC1 meets the points in the 3-D object. The line DC2 and the line DC3 also meet points in the 3-D object; therefore, we know that the point D is not a visible point.

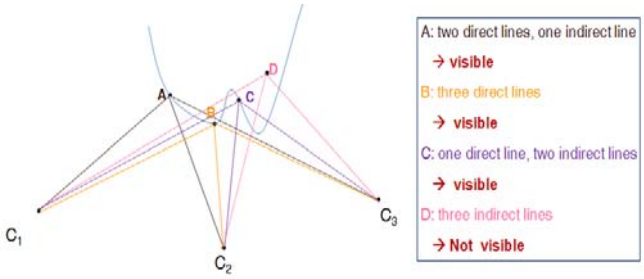

Fig. 4 Visibility checking

3.3 3-D object carving

In order to increase the accuracy of the silhouette of the 3-D object, we apply object carving. Up to the previous process, we do not obtain an accurate silhouette of the object, especially concave parts of the 3-D object. Thus, we need the carving process of the concave parts of the 3-D object. Fig. 5 shows the process of 3-D object carving. We have a volume of the object containing the true object. We choose a voxel on a visible point of the object using the visibility information. We project this voxel onto the visible input images. If the color value of each view is different from those of the other views, we perform carving. We repeat this process to update the visibility information.

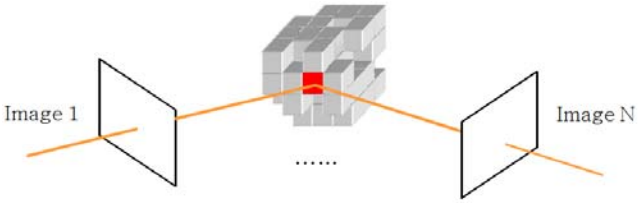

Fig. 5 Object carving for the real silhouette of the 3-D object

Object carving is processed by the color difference using the RGB distance defined by

$$\text{RGB-Distance} = \sqrt{(r_k - r_l)^2 + (g_k - g_l)^2 + (b_k - b_l)^2} \quad (2)$$

where rk and rl are red elements, gk and gl are green elements, and bk and bl are blue elements in two different views. Although a surface point is in the real silhouette of the 3-D object, the RGB distance may be high when one camera is away from the other cameras. In order to solve this problem, we select at most three cameras to calculate the RGB distance. We select the closest camera from the 3-D surface point and its nearby cameras. If a camera is very close to a surface point of the 3-D object, we assume that the color value of the camera reflects the color of the corresponding 3-D point well. Fig. 6 explains the process of 3-D object carving.

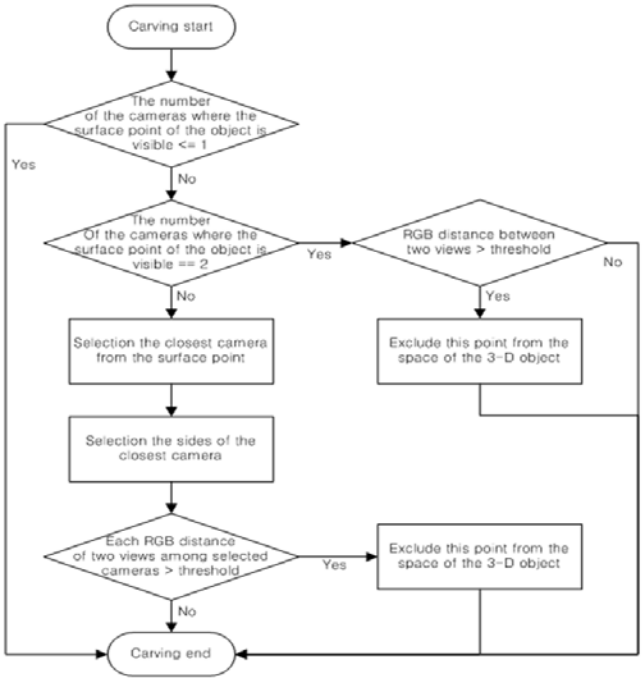

Fig. 6 Flowchart of 3-D object carving

We continue the 3-D object carving operation until all RGB distances of any two views are less than a threshold value. Whenever the process of object carving is performed, we update the visibility information from the process of visibility checking, which makes the 3-D object carving process more efficient. After the object carving process, we fix visible points of the 3-D object.

Fig. 7 Experimental results of 3-D object reconstruction

3.4 Color mapping

The visibility information from visibility checking is also useful for color mapping. In this process, we select cameras that can see the visible point. If several cameras are selected, color differences of the corresponding visible point among those selected cameras may be large. Thus, we select at most three cameras for color mapping. Since the closest camera from the surface of the 3-D object reflects better color, we select the closest camera from the visible point of the 3-D object and its neighboring cameras. After selecting cameras, we perform the blending operation by

$$R = \alpha * R1 + \beta * R2 + \gamma * R3$$
$$G = \alpha * G1 + \beta * G2 + \gamma * G3$$
$$B = \alpha * B1 + \beta * B2 + \gamma * B3 \quad (3)$$

where R is the red value, G is the green value, and B is the blue value of the visible point of the 3-D object. R1, R2, and R3 are red values of each view. G1, G2 , and G3 are green values of each view. B1, B2, and B3 are blue values of each view. If only two cameras are selected, R3, G3, and B3 are equal to zero; if only one camera is selected, the second and third terms of Eq. (3) are zero.

The weighting factors, α, β, and γ, are determined by

$$\alpha = \frac{1/v_1}{1/v_1+1/v_2+1/v_3} \quad \beta = \frac{1/v_2}{1/v_1+1/v_2+1/v_3} \quad \gamma = \frac{1/v_3}{1/v_1+1/v_2+1/v_3} \quad (4)$$

where v1, v2, and v3 are distances between the surface point of the 3-D object and each camera center, respectively. As shown in Eq. (4), the color of the visible point is more influenced by the color of a closer camera.

4. Experimental results

We have tested the performance of the proposed method using two data sets: "DinoSparseRing" and "TempleSparseRing"[13]. We have included 16 views for "DinoSparseRing"; however, for "TempleSparseRing", we have used only 10 views of the same latitude angle. These data sets are sampled on a ring around the 3-D object and the size of each image is 640×480. Fig. 7 shows four input images and experimental results for the two data sets. These results represent the reconstructed model captured from the same viewing point as the input images

In order to evaluate the performance of our method, we have projected 3-D points onto the input images and measured the variance of the projected pixel values in the input images. We have also compared the performance of our method with that of the basic visual hull method. Table 1 lists mean squared error (MSE) values of each RGB value and luminance value. From Table 1, we observe that the proposed method performs better than the basic visual hull method.

Fig. 8 shows the input images, and projected images by the visual hull and our proposed methods. Although those projected images have still some small errors, they look very similar to the original images.

Table 1. Performance comparison

	Visual hull	Proposed method	Error variation (%)
Average MSE (R)	734.79	644.01	-14.10
Average MSE (G)	706.91	627.66	-12.63
Average MSE (B)	686.19	604.44	-13.52
Average MSE (Luminance)	521.32	461.70	-12.91

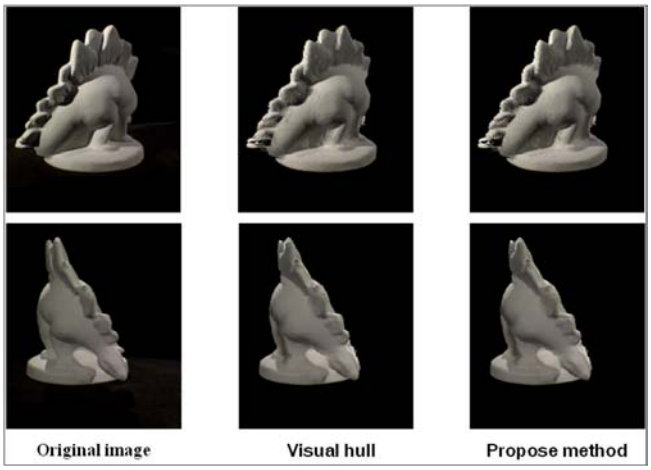

Fig.8 Projected images for performance evaluation

5. Conclusion

In this paper, we described two common methods for 3-D object reconstruction and proposed a new algorithm based on these two methods. Although the shape from silhouette method is simple and the voxel coloring method is efficient, both methods are not good enough for the accurate 3-D object reconstruction. By including the visibility checking process in the two common methods, we improved the accuracy of the 3-D object reconstruction. Experimental results demonstrated reasonably good performances with 3-D object data sets.

Acknowledgement This research was supported by the MKE(Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2010-(C1090-1011-0003)).

References

1. M. Grum and A. Bor (2008) Multiple Image Disparity Correction for Scene Representation, *IEEE International Conference on Image Processing*: 209-212
2. D.G. Lowe (2004) Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, 60: 91-110
3. Y. Lu, J. Z. Zhang, Q.M. Jonathan Wu, and L. Ze-Nian (2004) A Survey of Motion-Parallax-Based 3-D Reconstruction Algorithm, *Proc. of IEEE Transactions on System, Man, and Cybernetics-Part C: Applications and Reviews*. 34: 4
4. A. Laurentini (1994) The Visual Hull Concept for Silhouette-Based Image Understanding, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 15: 150-162
5. Y. Hwang, J. Kim, and I. Kweon (2005) Silhouette Extraction for Visual Hull Reconstruction, *IAPR Conference on Machine Vision Applications*: 39-42
6. W. Matusik, C. Buehler, R. Raskar, S. Gortler, and L. McMilan (2000) Image-Based Visual Hulls, *SIGGRAPH*: 369-374
7. A. Ladikos, S. Benhimane, and N. Navab (2008) Efficient Visual Hull Computation for Real-Time 3-D Reconstruction using CUDA, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*: 1-8
8. S. Tosovic, R. Sablatnig, and M. Kampel (2002) On Combinig Shape from Silhouette and Shape from Structured Light, *The 7th Computer Vision Winter Workshop*: 108-118
9. C. Esteban and F. Schmitt (2002) Multi-Stereo 3D Object Reconstruction, *International Symposium on 3D Data Processing Visualization and Transmission*: 159-166
10. S. Seitz and C. Dyer (1997) Photorealistic Scene Reconstruction by Voxel Coloring, *IEEE Computer Society Conference on Couputer Vision and Pattern Recognition*: 1067-1073
11. W. Culbertson and T. Malzbender (1999) Generalized Voxel Coloring, *Vision Algorithms: Theory and Practice*. 1883: 67-74
12. E. Steinbach, B. Girod, P. Eisert, and A. Betz (2000) 3-D Object Reconstruction Using Spatially Extended Voxels and Multi-Hypothesis Voxel Coloring, *International Conference on Pattern Recognition*: 774-777
13. Middlebury, Dino Spars Ring data set and Temple Spars Ring data set available at: *http://vision.middlebury.edu/mview/data/*