Carved Visual Hull using Correlation of Projected Pixels

Woo-Seok Jang and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro, Buk-gu, Gwangju, 500-712, South Korea
{jws, hoyo}@gist.ac.kr

Abstract - ***In this paper, we propose a 3-D object reconstruction method. First, Initial silhouette of a 3-D object is generated by a shape from silhouette technique, it is carved to represent details of the 3-D object using color consistency. A surface point of the 3-D object obtained by visibility information is projected on the 2-D image during carving process. We select the projection samples using correlation of all pixels in projected region. These samples are considered to form the shape of the 3-D object. We obtain more reasonable threshold value of object carving through this process. Thus, it provides a efficient 3-D object carving process. After mapping proper color values to the surface points of the 3-D object for the more realistic 3-D model, it generates a good result of 3-D object reconstruction.***

I. Introduction

Three-dimensional (3-D) object reconstruction from multiple two-dimensional (2-D) images is one of the most important topics in computer vision. This technique focuses on producing 3-D reconstructed model from calibrated images of an object. To achieve this, we try to acquire 3-D geometry and color information of the object using multiple images and camera parameters. Generally, camera parameters for multiple images are obtained by a camera calibration technique [1]. The relationship between a 3-D point (X) and its projected point(x) on a certain view is represented as [2]

$$x = PX \tag{1}$$

Where P is a camera projection matrix obtained from a camera calibration technique. Each view has a own camera projection matrix. A camera projection matrix is a useful tool for 3-D reconstruction.

Figure 1 shows one point in the 3-D space is projected on the image plane of the 2-D images by camera parameters. Even though we know the camera projection matrix for certain view, we can not exactly back-project the points on the 2-D image since we do not know the 3-D geometry information. The back-projected point of certain 2-D point can be located on the line passing through the 2-D point and its 3-D point and we define this line as a line of sight for a given 2-D point. Then, we can uniquely determine the 3-D point of the certain 2-D point by checking the intersection of two lines of sight for the given certain 2-D point and its corresponding point in the other view. This is called as a triangulation process [3]. Figure 1 also shows the 3-D point can be obtained as the intersection of the two lines of sights.

3-D reconstruction methods have traditionally been based on image matching, using either intensity-based direct methods or feature-based methods. This approach is especially effective for techniques which simplify the correspondence problem. However, finding the corresponding points between

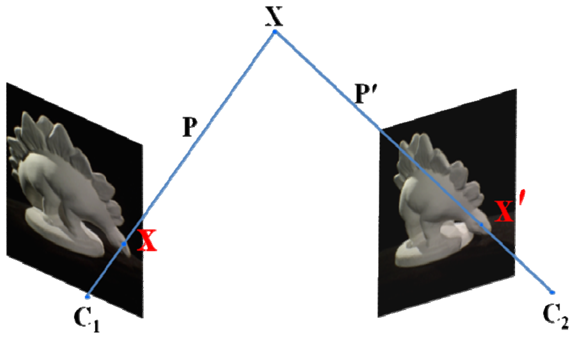


Figure 1. Decision of 3-D point through triangulation

two views is not easy and generates high complexity. Furthermore, it has restrictions that views must often be close together. To avoid the disadvantages listed above, we propose the other approach. It generates occupancy decisions about whether a volumetric primitive contains objects in the scene.

In this paper, we present a 3-D reconstruction algorithm including two common methods. These are shape from silhouette and voxel coloring. From these methods, we obtain a initial silhouette of a 3-D object and it is corrected to represent accurate details of the 3-D object using color consistency. Proposed algorithm uses new sampling method for the appropriate representative value of each occupied voxel.

II. Background

A. Shape from silhouette

The shape from silhouette method reconstructs a 3-D object by creating the visual hull from input images. The visual hull is an approximation of the 3-D object shape. To obtain this approximation, we create a silhouette image of the object in the acquired input images, extract contour lines related to the silhouette, and back-project points on the contour line on the 3-D space using camera calibration data.

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 the multiple silhouette images [4]. Although the visual hull is a good approximation that encloses the entire object, details of the object cannot be reconstructed by the visual hull alone. In spite of this weak drawback, many researchers still use the shape from silhouette method for 3-D object reconstruction because of its simplicity and fast implementation.

B. Voxel coloring

The voxel coloring method depends on color consistency [5]. If the colors of an unoccluded point seen from the different cameras are the same, then that point is assumed to be on the surface point of a 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.

III. Reconstruction Algorithm

A. Initialization

The visual hull of a 3-D object is used to initialize the reconstruction algorithm. We apply voxel structure to the above method to represent the approximation of the 3-D object shape. The 3-D space is partitioned into small voxels which are volume element representing a value on a regular grid in 3-D space. To examine whether each voxel in the 3-D space is a part of the object or not, each voxel is projected on the image plane of each view. The 3-D silhouette of the target object is obtained by using this process. However, this result does not provide the 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 reconstruction result which is added another process.

B. Visibility information

The visibility information gives us 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. Thus, we apply the visibility checking method for visibility information using line equation between a point in the approximation silhouette of the 3-D object and each camera center [6]. Visibility checking method is first applied after initialization. Visibility information is updated using visibility checking method whenever visibility information is changed in a 3-D object carving process.

C. 3-D object carving

To increase the accuracy of the 3-D object silhouette, we process object carving. Up to the previous process, we do not have the accurate silhouette of the object including concave parts of the 3-D object. Thus, we need a carving process for the concave parts and inaccurate parts computed from a finite number of images. Figure 2 shows the process of the 3-D object carving. We have the volume of the object containing true object. We choose a voxel on a visible point of the object using visibility information. We project this voxel onto the visible input images. If color values of each view are different from those of the other views, we process carving. We repeat these processes until convergence and update the visibility information.

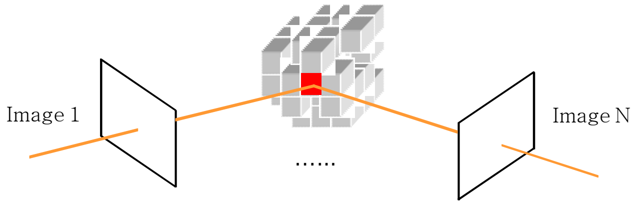


Figure 2. Object carving for real silhouette of the 3-D object

Object carving is processed by color difference using RGB distance. RGB distance is determined by

$$RGB\ distance = \sqrt{(r_k - r_l)^2 + (g_k - g_l)^2 + (b_k - b_l)^2} \qquad (2)$$

Where r_k, r_l are red elements of two views, g_k, g_l are green elements of two views, and b_k, b_l are blue elements of two views. Figure 3 shows the relationship of RGB distance between certain two points in the RGB space.

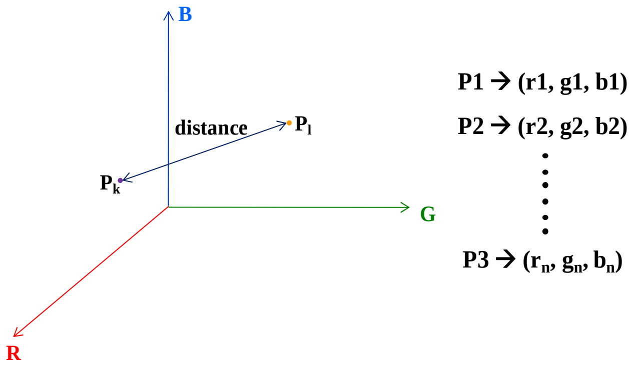


Figure 3. RGB distance

For object carving, we should determine the threshold value of RGB distance. In case of simple voxel structure, even if we use any threshold value of RGB distance, to obtain a accurate 3-D silhouette of the object is difficult. The reason is that the only center value of voxel is used as representative value of whole voxel. In fact, one voxel in the 3-D space is influenced by several pixels of a 2-D image. Figure 4 shows the results of object carving by different threshold values of RGB distance. When threshold value is small, we see inaccurate result. However, actual threshold value is less than 40. These wrong results generate, since the resolution of a 2-D image is higher than that of voxel representation [7]. However, if we set up the higher resolution of voxel representation than that of a 2-D image, it increases the complexity of implementation.

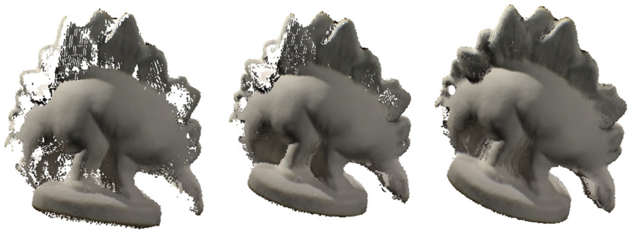


Figure 4. Object carving result by threshold value

Thus, we need projection sampling method to obtain the representative value of voxel. Figure 5 shows two cases. One is that only center value of voxel is projected on a 2-D image, the other is that whole voxel is projected on a 2-D image. From the figure, since we know that one voxel occupies several pixels of a 2-D image, we try to obtain the reasonable representative value of voxel using projection sampling method.

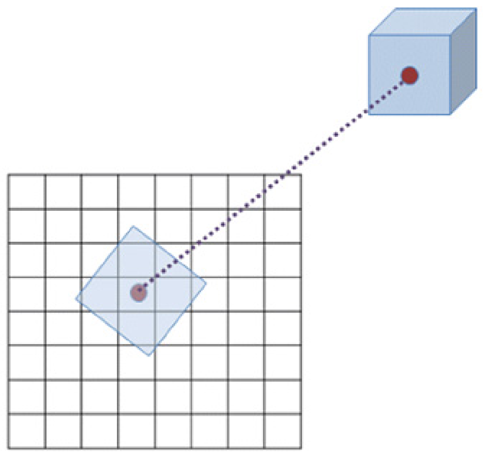


Figure 5: Voxel projection for projection sampling

As I explained before, when one voxel projects on the wider space than the space of one pixel, we should consider not a pixel projected by a central point of voxel but all pixels projected by whole voxel. For this, we use multi-vector cross correlation criterion method [8]. We extract sample pixels worth representative value of corresponding voxel among projected all pixels of whole voxel using multi-vector cross correlation criterion method. Multi-vector cross correlation criterion method is as follows. All projected pixels of a 2-D image generate their vectors. Each vector consists of three components, R, G, B.

$$\mathbf{v}_i = [R_i G_i B_i]^{\mathbf{T}} \qquad i \in \{1,2,.........,n\} \qquad (3)$$

We obtain the vectors like eq. (3) for n pixels. We measure their likeness. If n is equal to two, we use well known normalized cross correlation (NCC) as likeness. Normalized vector is defined by

$$\mathbf{n}_i = \frac{\mathbf{v}_i - \mathbf{m}_i}{\|\mathbf{v}_i - \mathbf{m}_i\|} \qquad i \in \{1,2,.........,n\} \qquad (4)$$

Where vector m_i is comprised by average of R_i, G_i, B_i which are each component of v_i. Furthermore, partial mean sum vector which is average of the other vectors except one vector is defined. This concept is represented by

$$\mathbf{s}_i = \frac{1}{n-1}\sum_{j \neq i} \mathbf{n}_j \qquad (5)$$

We compute the likeness between every vector and the partial mean sum vector like eq. (6).

$$p_i = s_i \cdot n_i \qquad \forall i \qquad (6)$$

We find maximum value of p_i. It is defined by p_{max}. After obtaining p_{max}, only pixels of vector satisfying eq. (7) are extracted as the representative values of corresponding voxel.

$$P_i > T \cdot p_{max}, \qquad T \in [0,1] \qquad (7)$$

Like this, if projection samples are extracted, we set up representative value of voxel using average value of extracted sample pixel values. We can obtain more reasonable threshold value by using this sampling method.

When processing object carving, we find visible views from a 3-D point to decide whether to carve the silhouette of the 3-D object or not. When many views see the 3-D point, even if that point is the real silhouette of the 3-D object, the RGB distance can be high in case that one view is away from the other views. To solve this problem, we select maximum three views to calculate the RGB distance. One of them is the closest camera from the 3-D point. The others are the sides of the closest camera. As the distance between the view and the point of the 3-D object is short, we can assume that the color value of the view reflects the color of the corresponding 3-D point well. If visible views are two, two views are considered and if visible view is just one, object carving process is not allowed.

We repeat the algorithm of 3-D object carving until all RGB distances of any two views are less than threshold value. Whenever the process of object carving is executed, the visibility information from the process of the visibility checking is updated. This process makes the 3-D object carving process efficient. After the whole carving process, visible points of the 3-D object are fixed.

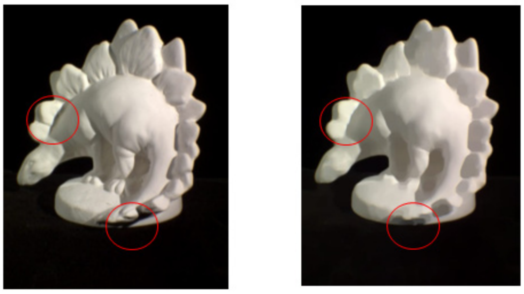
D. Color mapping

We map proper color values to the surface points of the 3-D object as the final step. This process increases realism. This process is performed from the maximum three close views. We determine the color value by blending the selected color values considering the distance between the 3-D point and each view.

IV. Experimental Result

We tested the 3-D reconstruction algorithm with two data sets which are multi-view images. The two data sets are "Dino-sparseRing" and "TempleSparseRing". Both data sets originally include 16 views. However, for "TempleSparse-Ring", we used only 10 views which have same latitude angles for each image. These data sets are sampled on a ring around the 3-D object. The size of each image is 640×480. These images are calibrated.

The data sets used in this paper can separate the objects from the original input images easily since the background is represented with black-tone colors. However, since some shaded regions of the object also have black-tone colors, they can be regarded belonging to background. To solve above the problem, we compensate the shaded regions of the object using preprocessing consisting of image dilation and erosion. Figure 6 shows an example of preprocessing. The circles represent shaded regions of the object in the original image. As you can see in fig. 6(b), the shaded regions are changed with the bright colors in the preprocessed image and can be distinguishable with background. This result is helpful to obtain the more accurate initial silhouette of the 3-D object.



(a) Original image (b) Preprocessed image

Figure 6. Result of preprocessing.

Figure 7 shows four input images and experimental results for the two data sets. The results represent reconstructed model captured from same viewing point of input images.

(a) Input images

(b) Reconstructed model

Figure 7. Result of 3-D object reconstruction

V. Coclusions

We propose a 3-D object reconstruction method using multiple 2-D images. Basically, we used a shape from silhouette method and theoretical foundations of voxel coloring. We apply a visibility checking method for visibility information. It increases the accuracy and efficiency for each stage. In a object carving process, a point in the 3-D space is projected on the 2-D image and we consider difference between the projected pixels. However, if we use only central value of the voxel, to obtain reasonable threshold value of color difference is difficult. The reason is that the resolution of the 2-D image is higher than that of a voxel representation. Thus, we need information of all projected pixel value which affects a 3-D point. To solve this problem, we extract sample by using correlation among projected pixel. From this method, we obtain more reasonable threshold value for object carving. After mapping color value to the 3-D object, we obtain the valid result of 3-D object reconstruction.

Acknowledgments

This research was supported by the MKE (The 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. J. M. Lavest, M. Viala, and M.Dhome, "Do we really need an accurate calibration pattern to achieve a reliable camera calibration?," European Conference on Computer Vision, pp. 158-174, 1998.
2. Y. Lu, J. Z. Zhang, Q. M. Jonathan, and Wu, L. Ze-Nian, "A Survey of Motion-Parallax-Based 3-D Reconstrcution Algorithm," IEEE Transactions on System, Man, and Cybernetics-Part C: Applications and Reviews, Vol. 34, No.4, 2004.
3. R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2003.
4. A. Laurentini, "The Visual Hull Concept for Silhouette-Based Image Understanding," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 2, pp.150-162, 1994.
5. W. Culbertson and T. Malzbender, "Generalized Voxel Coloring," Vision Algorithms: Theory and Practice, vol. 1883, pp. 67-74, 1999.
6. W. Jang, K. Oh, and Y. Ho, "3-D Reconstruction using Surface Detection," 3D System and Application, pp. P2-07(1-4), 2009.
7. A. Broadhurst and R. Cipolla, "A statistical consistency check for the space carving algorithm," 11th British Machine Vision Conference, Vol. I, pp. 282-291, 2000
8. C. Esteban and F.Schmitt, "Multi-Stereo 3-D object reconstruction," International Symposium on 3D Processing, Visualization, and Transmission, pp. 159-166, 2002.