Bounding volumeestimation algorithm for image-based 3D object reconstruction

Tae Young Jang[1], Sung Soo Hwang[1], Hee-Dong Kim[1], and Seong Dae Kim[1]

[1] Department of Electrical Engineering, KAIST / Daejeon 305-701 South Korea
{tyjang1020, dreamerjoe, khdong98, sdkim}@kaist.ac.kr

***** Corresponding Author: SeongDae Kim

Received November 20, 2013; Revised December 27, 2013; Accepted February 12, 2014; Published April 30, 2014

** Short Paper*

***Abstract*: This paper presents a method for estimating the bounding volume for image-based 3D object reconstruction. The bounding volume of an object is a three-dimensional space where the object is expectedto exist, and the size of the bounding volume strongly affects the resolution of the reconstructed geometry. Therefore, the size of a bounding volume should be as small as possible while it encloses an actual object. To this end, the proposed method uses a set of silhouettes of an object and generates a point cloud using a point filter. A bounding volume is then determined as the minimum sphere that encloses the point cloud. The experimental results show that the proposed method generates a bounding volume that encloses an actual object as small as possible.

***Keywords*: Bounding volume, Point consistency, Object reconstruction

1. Introduction

A visual hull is a three-dimensional (3D) entity computed from multi-view silhouettes [1]. Each silhouette of an object defines a generalized cone that encloses the object by back-projecting it to 3D space and a visual hull is generated by the intersection of these cones. A visual hull has been widely used for the image-based reconstruction since it reflects the overall structure of an object and its computation is relatively simple.

Since A. Laurentini proposed the concept of a visual hull, a variety of algorithms have been proposed to compute a visual hull efficiently [2-5]. In most of these algorithms, computing a bounding volume precedes the actual geometry reconstruction. A bounding volume of an object is a 3D space where the actual object is expected to exist [2, 3]. The accuracy of a visual hull increases with increasing number of voxels or segments in a bounding volume since voxels or segments in a bounding volume are used to compute a visual hull. In other words, when the number of voxels or segments is fixed, the more compact a bounding volume is, the more accurate a visual hull is. Hence, to compute an exquisite visual hull, it is important to determine a compact bounding volume.

Although the size of a bounding volume affects the accuracy of a visual hull, few studies have been conducted on the issue. Most studies on visual hull computation handle the bounding volume estimation using simple methods [2-3, 6-9]. Some studies estimated a bounding volume as a 3D space that is visible from all cameras, i.e., a bounding volume is estimated according to the location of the cameras [6-9] (Fig. 1). However, the computed visual hull can be inaccurate if an object is much smaller than an estimated bounding volume since the number of segments or voxels is insufficient to represent a visual hull. Another research [3] estimated a bounding volume using silhouettes. First, it finds rectangles that enclose silhouettes. Thereafter, it estimates abounding volume as an inscribed cylinder of the intersection of back-projected rectangles. Although this method is simple, the estimated bounding volume does not enclose the intersection of back-projected rectangles. Hence, a dead zone exists. If an object exists in a dead zone, the estimated bounding volume might not enclose the actual object(Fig. 2). Moreover, a bounding volume can be unnecessarily large if the silhouettes, which are extracted from the images by segmentation, are noisy.

In this paper, we propose a method that estimates a compact bounding volume which encloses an object. The proposed method generates an initial 3D point cloud using 2D polygons that enclose the silhouettes. Valid 3D points, which represent the volume of an object, are then extracted using a point filter. The point filter verifies the validity of

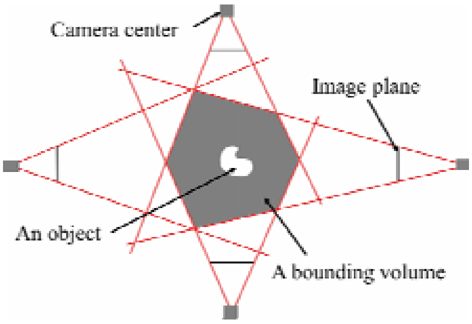


Fig. 1. Bounding volume estimated in [6-9]. If an actual object is much smaller than the estimated bounding volume, the computed visual hull can be inaccurate, since the number of segments or voxels representing a visual hull might beinsufficient.

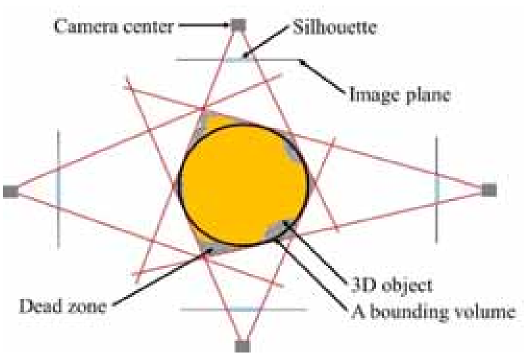


Fig. 2. Bounding volume estimated in [3]. It may not enclose the three-dimensional object.

each 3D point by projecting it into image planes. Subsequently, it generates a bounding volume that encloses valid 3D points. The proposed method estimates a more compact bounding volume than in previous works. Unlike [3], the proposed method is robust to noisy silhouettes since it extracts the valid points from an initial point cloud with the additional help of the point filter.

This paper is organized as follows: Section 2 explains proposed method. Section 3 includes experimental results and Section 4 presents the conclusions.

2. Proposed method

The proposed method consists of three steps, as shown in Fig. 3. In the first step, an initial 3D point cloud is generated using bounding 2D polygons of silhouettes. The validity of each 3D point is then evaluated by the point filter. In the last step, we construct a bounding volume of an object using the valid 3D points.

2.1 Generation of aninitial 3D point cloud

To estimate the valid 3D points that represent the volume of an object, the best method might be to find the intersection volume of back-projected silhouettes, and estimate a bounding volume that encloses the intersection volume. However, this method yields a complex computation when the shape of an object is complicated. Therefore, to avoid these problems, silhouettes are approximated as polygons that enclose the silhouettes.

After the polygons are found, the edges of the polygons are back-projected into the 3D space to generate planes (Fig. 4). Then, three planes out of all planes without ordering are selected. After selecting the three planes, the proposed method calculates a 3D point as follows:

$$Mx=\begin{bmatrix}\pi_1^T\\ \pi_2^T\\ \pi_3^T\end{bmatrix}x=0 \quad (1)$$

where π_1, π_2 and π_3 denote the plane parameter vectors, and x is a three-dimensional point. Note that rank(M) should be three, since π_1, π_2 and π_3 cannot produce a 3D point if these planes are dependent. This process is conducted to all three planes selected out of all planes, and produces a set of 3D points V, which is an initial point cloud (Fig. 5). Note that V includes a set, V', which consists of valid 3D points to represent the volume of an object.

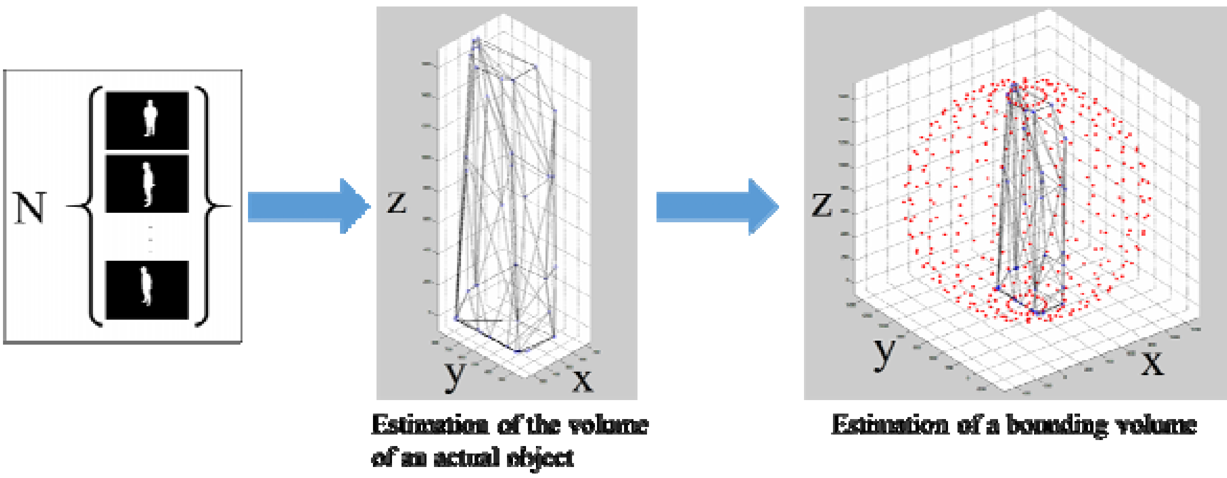


Fig. 3. Overview of the proposed method.

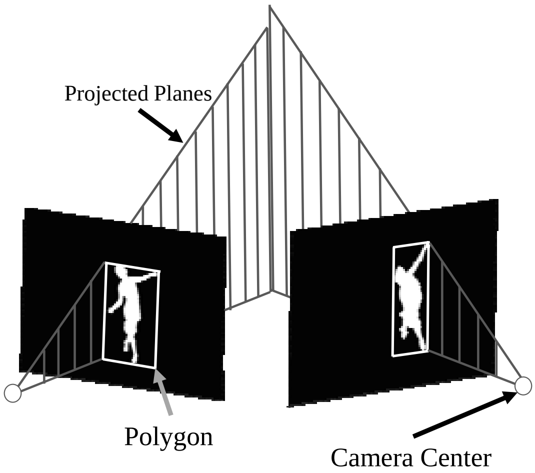


Fig. 4. Planes generated by back-projection edges of polygons.

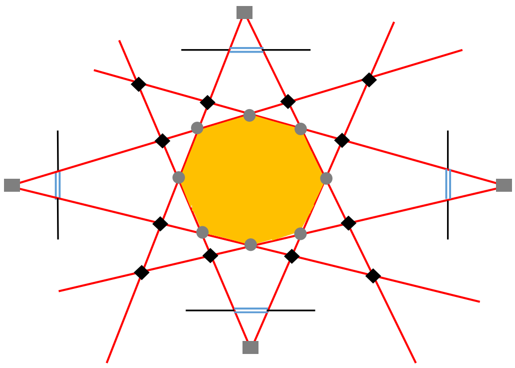


Fig. 5. Initial point cloud generated by three planes out of all planes.

2.2 Evaluation of an initial point cloud using a point filter

Subsequently, to find a set of valid points V', the proposed method applies a point filter to an initial point cloud. The point filter projects each 3D point of V to all image planes. If there is at least a projected element that is located on the outside of polygons, the element of an initial point cloud is an invalid point (Fig. 6). V' can be found after conducting this process on all elements of an initial point cloud. The volume of an object can be estimated from the distribution of the elements of a set V'. Through a point filter, the estimated volume of an object is robust to a noisy silhouette if there is at least one noiseless silhouette.

2.3 Estimation of a bounding volume

After estimating the volume of an object, it estimates a bounding volume. In this process, a variety of methods can be exploited. The simplestmethod is estimating the minimum cube enclosing valid points. The other approach is estimating the minimum enclosing sphere [10-12]. This approach manipulates a bounding volume easily since

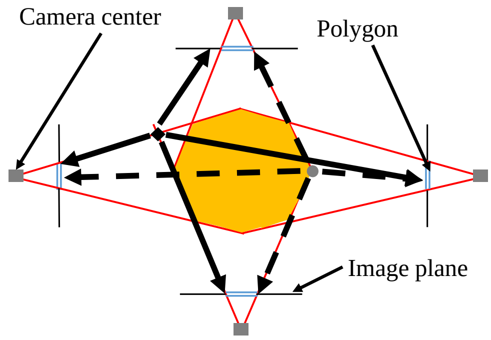


Fig. 6. Evaluation of an initial point cloud using a point filter.

Fig. 7. Standing man (Left), Dancer (Middle), Capoeira (Right) (a) represents the texture, (b) is the silhouette.

there are only four parameters. Hence a bounding volume was estimated as a sphere, and Welzl's method was used to calculate the minimum enclosing sphere [12].This method uses a linear programming algorithm and can estimate the minimum sphere which encloses a set of points in the expected $O(n)$ time.

3. Experimental Results

Several experiments were carried out to evaluate the performance of the proposed method. We compared our method with previous methods [3], and [6] using publicly available datasets and our dataset, i.e., standing man(Fig. 7). A chroma-keying algorithm proposed in [13] was used to acquire silhouettes for our dataset. Silhouettes of publicly available datasets were provided along with the color images. Table 1 summarizes the datasets. In addition, we defined the number of points on a visual hull as the resolution of a visual hull, and the number of voxels or segments in a bounding volume as the resolution of a bounding volume. In the process of generating a point cloud, the rectangle was used as a polygon. To verify the validity of using a rectangle as a polygon, the length of the x-axis, y-axis, and z-axis of the estimated volume of an object using a rectangle and a silhouette was compared (Table 2). From Table 2, we recognized that a rectangle is a proper approximation of a silhouette.

Tables 3 and 4 show the radius of the bounding volume that is estimated by using the methods in [3, 6] and the

Table 1. Data used for the evaluation.

	Resolution	Views
Standing man	1024x768	11
Dancer data	780x582	8
Capoeira	1004x1004	8

Table 2. Validity of using a rectangle as a polygon.

		Rectangle	Silhouette
Standing man	Length of x-axis	0.53	0.45
	Length of y-axis	0.71	0.63
	Length of z-axis	1.97	1.86
Dancer	Length of x-axis	1.32	1.29
	Length of y-axis	0.67	0.58
	Length of z-axis	1.84	1.77
Capoeira	Length of x-axis	0.82	0.78
	Length of y-axis	1.97	1.86
	Length of z-axis	0.80	0.57

Table 3. Radius of the bounding volume estimated using ground-truth silhouettes.

	Standing man	Dancer	Capoeira
Method[6]	2220.16	3.26	2.40
Method[3]	1030.36	1.20	1.12
Proposed Method	922.16	0.95	0.97

Table 4. Radius of the bounding volume estimated using noisy silhouettes.

	Standing man	Dancer	Capoeira
Method[6]	2220.16	3.26	2.40
Method[3]	1298.88	1.72	1.31
Proposed Method	922.56	0.95	0.98

proposed method with ground-truth silhouettes and noisy silhouettes. In the case of noisy silhouettes, the silhouettes provided and acquired were used. Table 3 shows that the proposed method estimates the smaller bounding volume than the other bounding volumes generated using the method in [3] and [6]. In Table 4, we can recognize that the proposed method is robust to noisy silhouettes, since the radius of the bounding volume estimated from noisy silhouettes is similar to that estimated from the ground-truth silhouettes. These results show that the proposed method is robust to noisy silhouettes and estimates a compact bounding volume.

Fig. 8 compares the resolution of a visual hull generatedusing the proposed method, method in [3], and method in [6] to each dataset. As the results show, the proposed method generated a high resolution visual hull for any resolution of a bounding volume. For example, in the dancer dataset, the resolution of a visual hull generated using each method is as follows: the proposed method was 1.30×10^5, the method in [3] was 3.93×10^4 and the

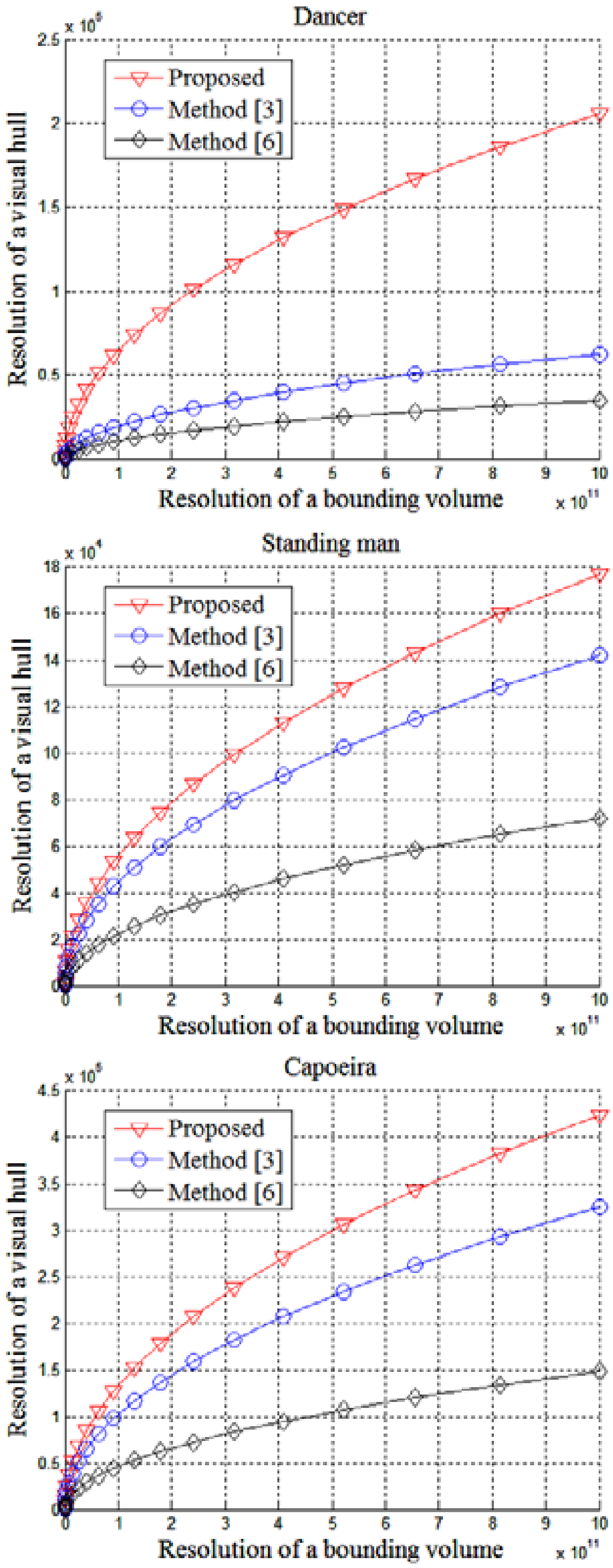


Fig. 8. Experimental results for the three datasets. Thevisual hull computation using the proposed method generates a visual hull with ahigh resolution. In terms of resolution of a bounding volume, visual hull computation using the proposed method represents the geometry component of a visual hull despite having a low resolution.

method in [6] was 2.20×10^4 in the case when the resolution of a bounding volume is 5.00×10^8. In the case of standing man dataset, the resolution of a visual hull was 1.11×10^5, 8.90×10^4 and 4.51×10^4, respectively, and, 2.66×10^5, 2.04×10^5, and 9.30×10^4 in the case of capoeira data, respectively, at a bounding volume resolution of 5.00×10^8. In particular, in the result of the dancer data, the proposed method achieved approximately three to four times higher visual-hull resolution than the previous methods. From these results, the proposed method helped to generate a high resolution visual hull.

4. Conclusion

In this paper, we propose a method that estimates a bounding volume that encloses an object using a point filter. To this end, the silhouettes of an object were used. Through a point filter, the volume of an object was estimated even when the silhouettes are noisy. In addition, a sphere was used as a bounding volume since it is represented by four parameters. Through the proposed method, a visual hull can have high resolution, despite the low resolution of a bounding volume. The experimental results show that the proposed method can be used as a useful tool for image-based object reconstruction.

Acknowledgement

This work was supported by Mid-career Researcher Program through the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education, Science and Technology (MEST)(2011-0016298)

The captured performance data were provided courtesy of the research group three-dimensional Video and Vision-based Graphics of the Max-Planck-Center for Visual Computing and Communication (MPI Informatik / Stanford).

The dancer data used in this project was obtained from http://grimage.inrialpes.fr.

References

[1] Laurentini, Aldo. "The visual hull concept for silhouette-based image understanding." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.16, no.2, pp.150-162, February, 1994. Article (CrossRef Link)

[2] Kutulakos et al., "A theory of shape by space carving." *International Journal of Computer Vision*, vol.38, no.3, pp.199-218, 2000. Article (CrossRef Link)

[3] Lee, Jin-Sung, et al. "Efficient three-dimensional object representation and reconstruction using depth and texture maps." *Optical Engineering*, vol.47, no.1, pp.017204-1-017204-8, January, 2008. Article (CrossRef Link)

[4] Kim, Sujung, et al. "Fast computation of a visual hull." *in Proc. ofComputer Vision–ACCV 2010*. Springer Berlin Heidelberg, pp.1-10, 2011. Article (CrossRef Link)

[5] Matusik, Wojciech, et al. "Image-based visual hulls." *in Proc. of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., pp.369-374, 2000. Article (CrossRef Link)

[6] Matsuyama, Takashi, et al. "Real-time three-dimensional shape reconstruction, dynamic three-dimensional mesh deformation, and high fidelity visualization for three-dimensional video." *Computer Vision and Image Understanding*, vol.96, no.3, pp.393-434, 2004. Article (CrossRef Link)

[7] Grau, Oliver et al. "A combined studio production system for 3-D capturing of live action and immersive actor feedback." *IEEE Transactions on Circuits and Systems for Video Technology*, vol.14, no.3, pp.370-380, 2004. Article (CrossRef Link)

[8] Starck, Jonathan et al. "Surface capture for performance-based animation." *IEEEComputer Graphics and Applications*, vol.27, no.3, pp.21-31, 2007. Article (CrossRef Link)

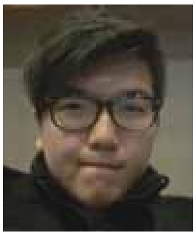
[9] Theobalt, Christian, et al. "High-quality reconstruction from multiview video streams." *IEEE Signal Processing Magazine*, vol.24, no.6 pp.45-57, 2007. Article (CrossRef Link)

[10] Megiddo, Nimrod. "Linear-time algorithms for linear programming in R3 and related problems." *IEEEFoundations of Computer Science, 1982. SFCS'08. 23rd Annual Symposium on*, 1982. Article (CrossRef Link)

[11] Skyum, Sven. "A simple algorithm for computing the smallest enclosing circle."*Information Processing Letters*, vol.37, no.3, pp.121-125, 1991 Article (CrossRef Link)

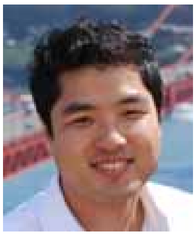
[12] Welzl, Emo. "Smallest enclosing disks (balls and ellipsoids)"*Springer Berlin Heidelberg*, 1991. Article (CrossRef Link)

[13] Hwang, S.S. et al."High-resolution 3D object Reconstruction using Multiple Cameras."*Journal of The Institute of Electronics Engineers of Korea*, vol. 50, no.10 pp.2602-2613, Oct. 2013. Article (CrossRef Link)



Tae Young Jang received his B.S. degree in electrical engineering from Sejong University, Korea, in 2013. Currently he is a graduate student of electrical engineering at Korea advanced institute of science and technology (KAIST) in Korea. He is interested in 3D object reconstruction.

Sung Soo Hwang received his B.S. degree in electrical engineering from Handong University, Korea, in 2008 and the M.S. degree in electrical engineering from KAIST, Korea, in 2010. Currently he is a Ph.D. candidate in electrical engineering from KAIST, Korea. His current research interests include computer vision, and 3D object compression.



Hee-Dong Kim received his B.S. degree in electrical engineering from Busan University, Korea, in 2005 and the M.S. degree in electrical engineering from KAIST, Korea, in 2007. Currently he is a Ph.D. candidate in electrical engineering from KAIST, Korea. His research interests include image processing, and computer vision.

Seong Dae Kim received his B.S. degree in electrical engineering from Seoul National University, Korea, in 1977, M.S. degree in electrical engineering from KAIST, Korea, in 1979 and Dr. Ing. degree in electrical engineering from ENSEEIHT, France, in 1983. Since 1984, he has been a Professor at the department of electrical engineering, KAIST. He was the President emeritus of the Institute of Electronics and Information Engineers (IEIE).His research interests include image processing, and computer vision.