Searching the Optimal Threshold for Voxel Coloring in 3D Reconstruction

Young-Youl Yi[1], Hyo-Sung Kim[1], Soo-Young Ye[1], Ki-Gon Nam[1]

[1] Department of Electronic Engineering, Pusan National University

Email: mustapha79@pusan.ac.kr

Abstract

Voxel coloring is one of the well-known methods for reconstructing a 3D shape from 2D images. The conventional methods cause a trade-off problem between precision and stability, when they reconstruct 3D shapes. In this paper, we present a novel approach to solve the trade-off problems. This method searches the real surface voxel on comparing the photo-consistency of an inside voxel on the optic ray with the surface voxel of a center camera. As iterating proposed voxel coloring, the method can search the optimal threshold by itself. The graph cut method is also used for reducing the surface noise.

Keywords: Voxel Coloring, optimal threshold, photo-consistency, optical ray, 3D reconstruction

1 Introduction

The desire of human being craves not the media on 2D plane surface but the media in 3D space, because of the improved computer performance and the wide spread of high speed Internet. The virtual reality that is embodied in 3D space is still in the beginning level, but it is already used in some fields, for example multi media contents, game, movie, and education/training simulation. In near future, virtual reality will be used in all kinds of fields.

The most important thing in virtual reality technique is constructing the 3D model. The image-based 3D shape reconstruction has been studied for a long time. The techniques of reconstructing 3D model can be classified into two large groups. One is active sensing, and the other is passive one. The active sensing analyzes structured light that is reflected on real object. The passive sensing analyzes images that are acquired under general illumination or natural light. The passive sensing has lower precision than active sensing, but it is handy method using only general CCD camera, so it can be widely used in many fields.

In this thesis, we propose a novel method to reconstruct 3D shape from multi-view silhouette-based images. The previous voxel coloring method measures photo-consistency of single surface voxel and compares it with pre-established single threshold, then decides to eliminate the voxel or not. It is a very strenuous work to find the best single threshold. Even if the best threshold has been found, applying it to all surface voxel makes a tradeoff problem between precision and stability. In the proposed method, we compare the photo-consistency of surface voxel with its neighbor inside voxel, and eliminate surface voxel if its photo-consistency is lower than its neighbor. A graph cut method is also used to reduce the irregular noise of surface.

2 Previous work

As mentioned previously, there have been many trials of reconstructing 3D shape from multi-view silhouette-based images. Baker[1] used the silhouette of an object rotating on a turntable to construct a wire-frame model of the object. Martin and Aggarwal[2] used volumetric descriptions to represent the reconstructed shape. Potmesil[3] suggested an octree model using arbitrary views to speed up shaping from silhouette. For each of the views, he constructed the octree representing from conic volume and intersected octrees. Szeliski[4] first created a low resolution octree model quickly and then refined this model iteratively, by intersecting each new silhouette with the already existing model. Generally, the voxel carving method using silhouette images can quickly reconstruct the 3D shape of an object in voxel level, but the method also has some problems Seitz and Dyer[5] proposed the voxel coloring method using photo-consistency without volume carving. It can reduce model errors. The voxel coloring method has a disadvantage that the position of the camera having to satisfy the ordinal visibility constraint. Culbertson and Malzbender[6] proposed a generalized voxel coloring method (GVC) that can be used with a randomly positioned camera.

3 Our approach

The reconstructed 3D shapes using volume intersection have errors in image acquisition and in the concave portion of the object modeling. In figure 1 (a), we assume that *surface_a* is reconstructed surface that is derived by the volume carving method and *surface_b* is the real surface of object at the camera C_{n-1}, C_n, C_{n+1}. A voxel, V_n^{xy}, is defined as the voxel located on the optical axis of center camera C_n to be located by coordinate (*x*, *z*) of *x-z* axis. Given any voxel, we can obtain two voxels on the real surface through the voxel, V_n^{xy}, from two camers, C_{n-1}, C_{n+1}. Using the photo-consistency of the two voxels, we can obtain the dissimilarity of the voxel. In the Figure 1(b), using the voxel, V_n^{i4} and two cameras, C_{n-1}, C_{n+1}, we can obtain two voxels, V_{n+1}^{h6} and V_{n-1}^{j6} on the real surface. From the photo-consistency of V_{n+1}^{h6} and V_{n-1}^{j6}, we can get the dissimilarity of V_n^{i4}. V_{n+1}^{h6} and V_{n-1}^{j6} are close to V_n^{i7} on the real surface and the dissimilarity of V_n^{i4} is more decreasing.

At a voxel, V_n^{i7}, in Figure 1(c), we can obtain the lowest dissimilarity, because the voxel V_n^{i7} is correct voxel on real surface. At a voxel V_n^{i9}, in Fig. 1(d), the dissimilarity was calculated by higher value than V_n^{i7}.

The smaller value of the dissimilarity, the closer the voxel is located at the real surface. Therefore, the dissimilarity of any voxel V_n^{xy} has to compared with the dissimilarity of another voxel on the optical axis. And if the dissimilarity of the V_n^{xy} is larger than the dissimilarity of next voxel on the optical axis, the voxel should be eliminated. This process should iteratively be performed until finding the minimum dissimilarity to estimate the voxel. In the Figure 2, the characteristic change of dissimilarity was shown. When depth index is 7, the dissimilarity of a voxel has the lowest value. So we decided the voxel V_n^{i7} as real surface voxel. From all center camera positions, dissimilarity was calculated, and then the optimal threshold value was decided.

This method can be decrease in modeling error comparing with conventional method using the single-fixed threshold because of multi-variable threshold of the all voxels.

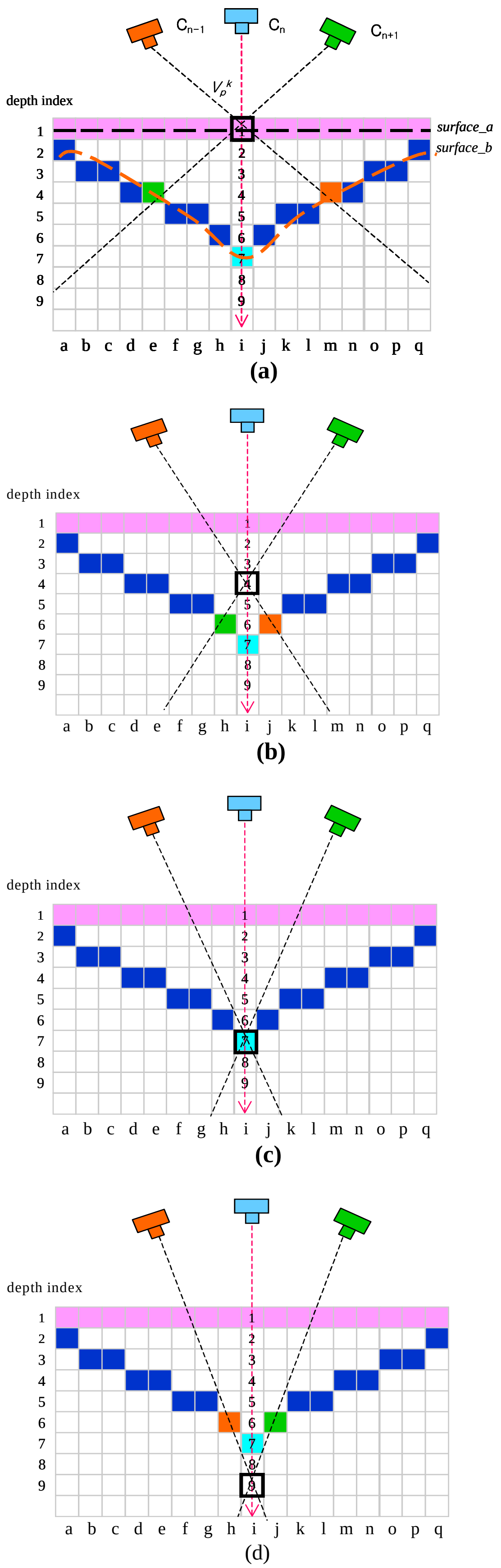

Figure 1: Dissimilarity calculation at the center camera on the optical ray.

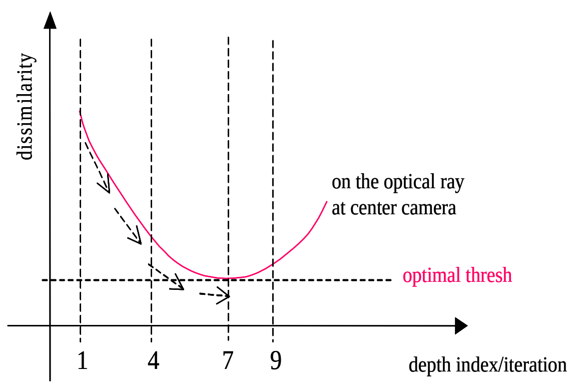

Figure 2: Balance of forces

4 Proposed Voxel coloring Steps

The proposed voxel coloring is basically a form of GVC algorithm[6] and searching optimal threshold method is added. The specific method follows next steps.

4.1 Calculating camera position

Let P^{1T}, P^{2T}, P^{3T} are the row vector of the given camera projection matrix P. $P^{1T}X = 0$ and $P^{2T}X = 0$ mean axis plane. $P^{3T}X = 0$ means principal plane like Figure 3. The camera position C is calculated by Eq. (1).

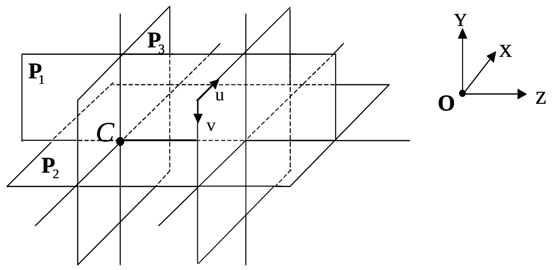

Figure 3: Three planes defined by the row vectors of the camera matrix

$$PC = 0 \tag{1}$$

4.2 Searching surface voxels

We intend to search the surface voxel which are on 3D voxel matrix that is reconstructed by carving method. The 3D voxel matrix has the information that when the voxel is carved, the voxel value is 1, otherwise is 0. To search surface voxel, if the value of itself would be 1 and the one of the 6-connected neighbor voxels would be at least 0, we will allocate a voxel V^k_{pi} to surface voxel, in the Eq. (2).

$$V^k_{sur} = \left\{ V^k_{pi} \middle| V^k_{pi} = 1 \wedge \exists V^k_{pj \in Ni} = 0 \right\} \tag{2}$$

Where, V^k_{pi} is the voxel of 3D voxel matrix at arbitrary position i, V^k_{pj} is the neighbor voxel of V^k_{pi}. Figure 3 is 6-connected neighbor voxels.

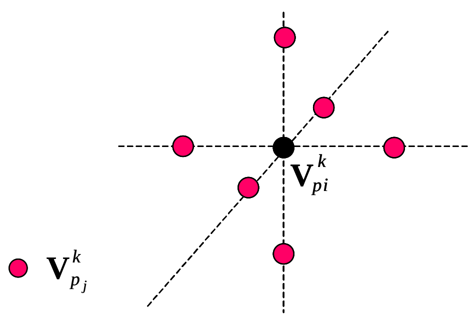

Figure 4: 6-connected neighbor voxels.

4.3 Searching visible surface voxels

The searched surface voxel is projected on an image plane to search visible surface voxel. If two voxels are overlaped, voxel index be saved in the visible index buffer with minimum depth from camera center like Figure 4.

After projecting all surface voxel, there is the only one index of the voxel which is seen from a camera in the visible index buffer. After this process is performed, the information of all voxel is acquired at each camera.

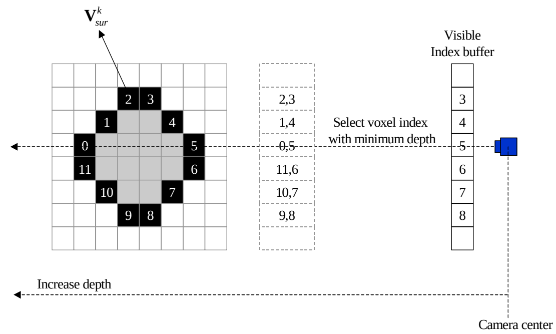

Figure 5: Estimation of a visible surface voxel

4.4 Calculation of the center camera

To decide the center camera $\mathbf{C}^k$, we search visible camera at a voxel surface. If voxel 3 is seen at the 6, 7, and 8 camera, center camera will be 7, like Figure 5.

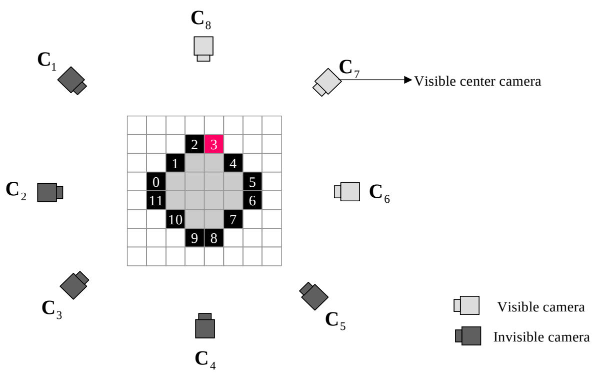

Figure 6: Calculation of center camera.

4.5 Calculation of optical ray

We calculate optical ray from visible center camera. At first, unit vector $\mathbf{n}^k$ is calculated with Eq. (3).

$$\mathbf{n}^k = \frac{\mathbf{C}^k - \mathbf{V}^k_{sur}}{\left\| \mathbf{C}^k \cdot \mathbf{V}^k_{sur} \right\|} \tag{3}$$

Where, $\mathbf{C}^k$ is visible center camera and V^k_{sur} is visible surface voxel.

The Figure 6 shows that $\mathbf{n}^k$ is unit vector on the optical ray at a visible center camera, V^k_{sur} is surface voxel and V^k_{in} is the inside voxel of V^k_{sur}.

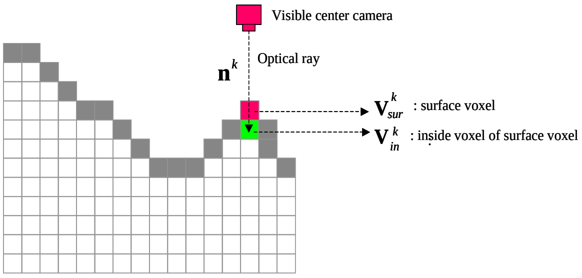

Figure 7: Voxels on the optical ray.

4.6 Calculation dissimilarity on the optical Ray

To decide the threshold value, we calculate dissimilarity on the optical ray. In the conventional voxel coloring method, dissimilarity was calculated from visible surface voxel. But in this paper, in order to solve the single-fixed threshold problem, dissimilarity is calculated from not only surface voxel but also inside voxel on optical ray in the Figure 7.

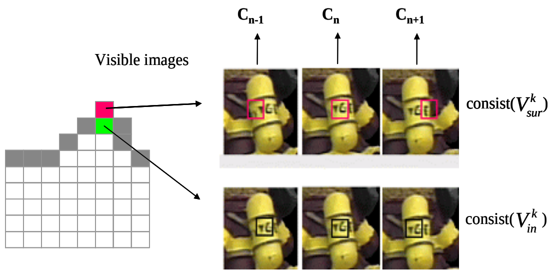

Figure 8: Calculation of dissimilarity for voxels on optical ray.

Dissimilarity is calculated between surface voxel V_{sur}^{k} and inside voxel V_{in}^{k} by using the photo-consistency. The relation of photo-consistency and dissimilarity is as following Eq. (4).

$$consist(\mathbf{V}_n) = \frac{a}{dissimilarity(\mathbf{V}_n)+1} \tag{4}$$

Where, $consist(\mathbf{V}_n)$ is photo-consistency value, a is arbitrary constant, and $dissimilarity(\mathbf{V}_n)$ is dissimilarity value. And $dissimilarity(\mathbf{V}_n)$ can be represented as following equation, Eq. (5).

$$dissimilarity(\mathbf{V}_r^k) = \sum_{i,j}\left\{\left|\mu_i^{red}-\mu_j^{red}\right|+\left|\mu_i^{green}-\mu_j^{green}\right|+\left|\mu_i^{blue}-\mu_j^{blue}\right|\right\} \tag{5}$$

Where, i, j represents the index of surface and inside voxels . μ_i^{red}, μ_i^{green}, μ_i^{blue} are the average value of RGB of surface voxel. The dissimilarity of V_{sur}^{k} and V_{in}^{k} on the optical ray is calculated. If the dissimilarity of V_{sur}^{k} is larger than V_{in}^{k}, the V_{sur}^{k} is considered as the model error and the voxel is eliminated. The other case, the V_{sur}^{k} is considered as the surface voxel of real object and the voxel is reminded. After this process is performed iteratively, the only real surface voxel is remained at the optimal threshold.

4.7 Decision of voxel elimination using graph cut

We used the graph cut method to finally decide surface voxels. We classified surface voxel into two categories, Opaque and Carving nodes as in Figure 8(a). The result is shown in Figure 8(b). We used the *E(f)* to minimize the energy of surface voxels. The Eq. (6) represents energy function.

$$E(f) = \sum_{\mathbf{V}_n \in \mathbf{V}_{sur}^k} D_{\mathbf{V}_n}(f_{\mathbf{V}_n}) + \sum_{\{\mathbf{V}_n,\mathbf{V}_q\}\in N} V_{\{\mathbf{V}_n,\mathbf{V}_q\}}(f_{\mathbf{V}_n}, f_{\mathbf{V}_q}) \tag{6}$$

In the above energy function, $D_{V_n}(f_{V_n})$ is the expense of data and $V_{\{\mathbf{V}_n,\mathbf{V}_q\}}(f_{\mathbf{V}_n}, f_{\mathbf{V}_q})$ is the expense of smoothing. $D_{V_n}(f_{V_n})$ can be divided into two cases. One is that f_{V_n} was assigned to Opaque, the other is Carving.

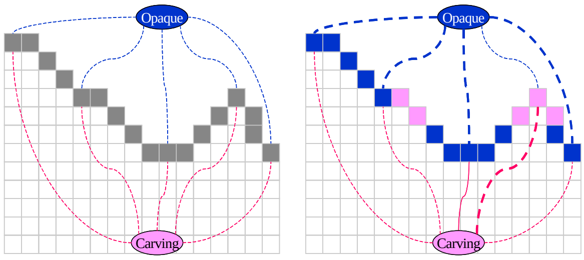

Figure 9: Construction of graph cut for voxel coloring.
(a) Construction of graph.
(b) Voxel labeling resulting graph cuts.

In Eq. (7), the expense of data term has the following form when f_{V_n} is assigned to Opaque.

$$D_{V_n}(f_{V_n}) = \begin{cases} 0 & if\ consist(V_{sur}) > avg\_consist \\ 1 & if\ consist(V_{sur}) \le avg\_consist\ \&\ consist(V_{sur}) \ge consist(V_{in}) \\ 2 & if\ consist(V_{sur}) \le avg\_consist\ \&\ consist(V_{sur}) < consist(V_{in}) \end{cases} \tag{7}$$

Where, the *avg_consist* is photo-consistency value of all surface voxels as in Eq. (8)

$$avg\_consist = \frac{1}{N}\sum_{V_n \in \mathbf{V}_{sur}^k} consist(V_n) \tag{8}$$

In Eq. (7), if the condition is $consist(V_{sur}) > avg\_consist$, the expense of surface voxel is assigned in a low value, 0, not to cut the voxel at the graph because the photo-consistency of V_{sur} is high. If the condition is $consist(V_{sur}) \le avg\_consist$, the expense is decided by considering the photo-consistency of inside voxel on the optical ray. That is, if the condition is $consist(V_{sur}) \ge consist(V_{in})$, the condition is assigned in lower value, 1, more than $consist(V_{sur}) < consist(V_{in})$ because the photo consistency of V_{sur} is larger than V_{in}. If the condition is $consist(V_{sur}) < consist(V_{in})$, the highest expense value, 2, is assigned. If f_{V_n} is assigned to Carving level, the expense value relationships are oppositely from the Opaque.

Expense of smoothing term, $V_{\{V_n,V_q\}}(f_{V_n}, f_{V_q})$, is as the following, Eq. (9).

$$V_{\{V_n,V_q\}}(f_{V_n}, f_{V_q}) = \begin{cases} 0 & if\ f_{V_n} = f_{V_q} \\ 1 & if\ f_{V_n} \ne f_{V_q} \end{cases} \quad (9)$$

Where, the V_q means 6-coupled neighbor voxel of V_n. To find the minimum energy, we used the graph cut algorithm that is proposed by Kolmogrov[7].

5 Experiment

In this experiment, the color CCD camera, JAI CV-S3300 was used. The acquisition image is 24bit colors, and its size is 640*480. We used the Visual C++ for compiler and the OpenGL to display 3D image. The Pentium4 computer was used for simulation. We acquired the silhouette images from a real 3D object. The images are 40 image slides with the angle of about 9°. Following images are some of the acquired images, in the Figure 10. We used the images for the input images.

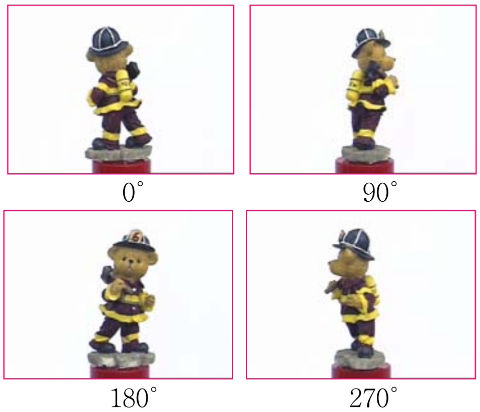

Figure 10: Input images.

Experimental conditions which are used for evaluating the proposed voxel coloring method are shown in the Table 1. In the Table 1, VI means the volume carving method of Szeliski[4] and is used for criterion to evaluate the effect of optimal threshold method.

Table 1: Experimental conditions.

	Algorithm	Threshold	Graph Cuts
(a) VI	Volume Intersection	--	--
(b) GVC_Th50	Generalized Voxel Coloring	50	×
(c) GVC_Th25	Generalized Voxel Coloring	25	×
(d) GVC_GC_Th50	Generalized Voxel Coloring	50	O
(e) GVC_GC_Th25	Generalized Voxel Coloring	25	O
(f) OTVC	Optimal Voxel Coloring	×	O

The experimental conditions (b), (c) used the general voxel coloring method of Culbertson and Malzbender[6], and the threshold value of dissimilarity was set to 50 and 25. Threshold means dissimilarity of the surface voxels. If the threshold value is small, the photo-consistency is high, and the other case the photo-consistency is low. The relation of between photo- consistency and dissimilarity is inverse proportion. We also applied the experimental condition (d), (e) to graph cut method at the same condition (b), (c). The experimental condition (f) is optimal threshold method using voxel coloring.

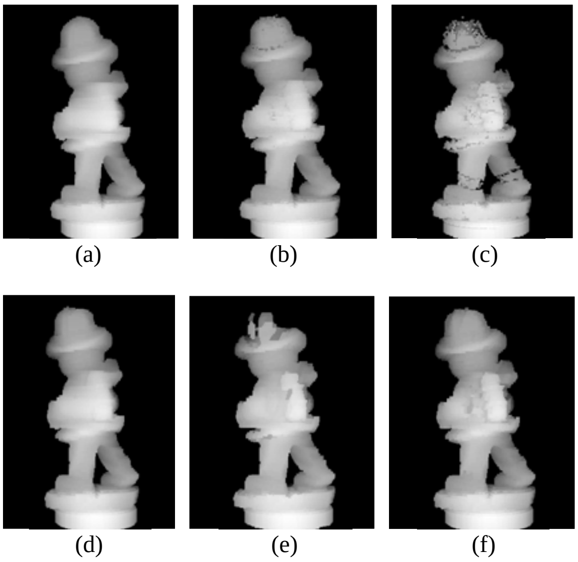

Figure 11: Depth map of the reconstructed of experimental conditions.
(a) VC, (b) GVC_TH50, (c) GVC_TH25
(d) GVC_GC_TH50, (e) GVC_GC_TH25, (f) OTVC.

We show the reconstructed results by using depth map of each experimental condition, in Figure 10. Figure10 (a) is the reconstructed shape using VI. It shows model errors because of concave surface.

In experimental conditions (b), (c), if the threshold value is large, the modeling result can be similar to the real object, but concave model error is large. If the threshold is small , concave model error can be small, but the precision of the reconstruction is low. Experimental conditions (d) and (e) are similar to (b) and (c), additionally graph cut method was applied. The result of conditions (d) and (e) shows that the surface noise is eliminated comparing with conditions

(b) and (c). But model error was large. Figure 10(f) was shown the result of using optimal threshold at the condition (f).

We decreased the model error by using the optimal thresholding method, and increased the stability of reconstruction by using the graph cut method. Figure 11 shows the average dissimilarity of reconstructed 3D shape by using the experimental conditions shown Table 1. We know that the smaller the dissimilarity value is the closer the voxel is. And also we found optimal threshold at the minimum dissimilarity.

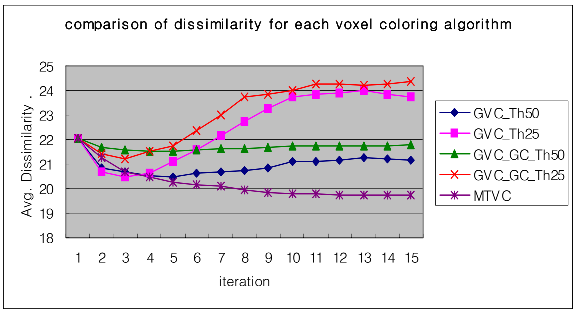

Figure 12: Comparison graph of dissimilarity for experimental conditions.

In this paper, proposed algorithm is better result than convention method.

6 Conclusions

We proposed the improved 'searching optimal threshold' method using the voxel coloring algorithm for the image-based 3D shape reconstruction. The proposed voxel coloring algorithm presented good result comparing with conventional voxel coloring algorithm using the single-fixed threshold value.

The threshold is approached to the optimal value as the dissimilarity of voxel is small. The process is iterated to find out the optimal threshold. And to eliminate the noise of surface voxel, we applied the graph cut method. Graph cut algorithm was used to minimize energy, and irregularities of surface were eliminated by energy of smooth term. Experiments were performed with conventional and proposed method under various conditions. In conventional voxel coloring algorithm, the trade-off problem of accuracy and stability was caused by the single-valued threshold of dissimilarity. We resolved the problem by using optimal threshold and graph cut method. The reconstruction efficiency of proposed algorithm is much better than conventional one.

7 Acknowledgments

This work was supported by "Research Center for Logistics Information Technology (LIT)" hosted by the Ministry of Education & Human Resources Development in Korea.

8 REFERENCES

[1] H. Baker, "Three-dimensional modeling," *Int. Joint Conf. on Artificial Intelligence*, pp. 649-655, 1977.

[2] W. N. Martin and J. K. Aggarwal, "Volumetric description of objects from multiple views," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 5, no. 2, pp. 150-158, 1983.

[3] M. Potmesil, "Generating octree models of 3D objects from their silhouettes in a sequence of image," *Computer Vision, Graphics, and Image Processing*, vol. 40, pp.1-29, 1987.

[4] R. Szeliski, "Rapid Octree Construction from Image Sequences," *Computer Vision, Graphics, and Image Processing*, vol. 58, no. 1, pp. 23-32, Jul. 1993..

[5] S. M. Seitz and C. R. Dyer, "Photorealistic Scene Reconstruction by Voxel Coloring," *Proc. Comprter vision and Pattern Recognition Conf.*, pp. 1067-1073, 1997.

[6] W. B. Culbertson and T. Malzbender, "Generalized voxel coloring," *Proc. of the ICCV*, pp. 100-115, 1999.

[7] V. Kolmogorov and R. Zabih, "What Energy Functions can be Minimized via Graph Cuts?," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2004