# Representation of Reflection and Transparency by a Photorealistic 3D Capture System

*Hiroto Matsuoka*†*, Tatsuto Takeuchi, and Ikuo Harada*

**Abstract**

We have developed a fast fully automated system that can generate photorealistic three-dimensional computer graphics data from a real object. A major factor in this success came from the findings of our psychophysical experiments, which revealed that human observers do not have an accurate idea of what should actually be reflected from the surface of an object. Utilizing this characteristic of human vision, we developed a rendering algorithm called Alpha Blending with Blurred Background for representing the reflection and transparency of real objects with less computational power. Since this rendering algorithm does not require object-shape data, the cost of object modeling is lower. The system can be applied to many areas, ranging from the rendering of virtual shopping malls for e-commerce to the creation of digital museums, and has many other potential uses.

## 1. Introduction

The synthesis of realistic images is one of the goals of three-dimensional computer graphics (3D CG). One way to generate realistic images is to use image-based techniques, in which images taken from various viewpoints are used as inputs and a photorealistic rendering is produced as an output. Recent advances in these techniques, which include modeling techniques [1]-[3], rendering techniques [4]-[6], and hybrid approaches [7], have made it possible to handle natural lighting conditions such as inter-reflection or transparency. Inter-reflections result from light bouncing between multiple surfaces and illuminating areas that are not directly illuminated by a light source. In the creation of 3D photorealistic computer images, accounting for inter-reflections has proven to be very important. The inter-reflected components can be represented quite accurately by measuring scene radiance and global illumination [8]-[10]. However, the cost of modeling and rendering to represent inter-reflection and transparency is very high because a huge amount of computation is needed.

In this paper, we introduce a new method that represents inter-reflection and transparency with high fidelity and greatly reduces the cost of modeling and rendering. Our method relies on the fact that humans do not recognize all aspects of an image. Many attempts have been made to utilize the characteristics of human vision to reduce the computational load and achieve realism in computer-generated images [11], [12]. Our method is an image-based technique that takes inter-reflections and transparency into account and generates CG data automatically in a short time. In this method, rather than representing true inter-reflections, we calculate pseudo (imperfect) inter-reflections that look realistic enough. This approach was developed by examining the characteristics of human vision for inter-reflections and utilizing knowledge obtained from a series of psychophysical experiments.

## 2. Fully automated 3D CG data generation system

The image-based modeling system we developed is called the photorealistic 3D capture system [13]. It can automatically generate high-fidelity CG data in only half an hour. The system consists of a robot arm with a digital camera, a turntable, and a control PC (personal computer), as shown in **Fig. 1**. The back screen and the turntable are covered with an electroluminescent (EL) sheet [14]. The generation proce-

† NTT Microsystem Integration Laboratories
Atsugi-shi, 243-0198 Japan
E-mail: faq@contents4u.com

dure is shown in **Fig. 2**. First, the system takes images of an object automatically from every viewpoint in the hemispherical area above it. The images are taken from about 100 angles (the interval between angles is about 15°) in half an hour or less. Next, the system extracts parts of the object from each image, obtains information about reflection and transparency, and measures the size of the object. At the time of rendering, the image taken from the nearest viewing angle is chosen from among all the images, resized to fit the scale of the virtual world, and rendered with the reflection and transparency of the object.

The keys to the success of this approach are efficiently and accurately segmenting the object's image from its background and representing reflections and transparency. We achieve accurate segmentation by using an EL sheet instead of a blue sheet as a background screen. The reflection and transparency components are represented by a novel rendering algorithm called Alpha Blending with Blurred Background (αBBB), which is based on the characteristics of human vision.

### 2.1 Segmentation method

To render an object in a 3D virtual space, it is necessary to separate the photographed object into parts. Chromakey with a blue screen is a standard segmentation method that is widely used in the virtual studios of television stations. However, it does not work well when an object has pixels of a similar color to the background. The shadow of an object is recognized as part of the object, as shown in **Fig. 3(c)**. Overcoming this problem requires a studio specially designed to eliminate shadows. In addition, blue lines remain at the edge of the object. Thus, adjustments are necessary and this increases the cost.

To solve these problems, we use two background colors to recognize the area of objects and multi-level silhouettes to separately represent parts of an object and the background in an image. The photorealistic 3D capture system takes two photographs of an object with different colored backgrounds. Since any part of the object can be identified from pixels whose color is the same in the two images, any object with arbitrary colors can be recognized from an image. When the shadow of the object is projected on the background, the color of the shadow area is changed so that it is correctly identified as part of the background. This method does not require a special studio and works in an ordinary room. The background color is changed by using the EL sheet, which changes color when the electric current supplied to it is altered.

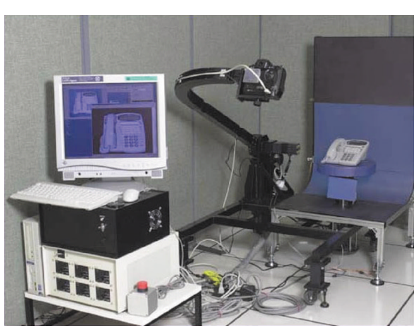


Fig. 1. Photorealistic 3D capture system.

### 2.2 Multi-level silhouette

Glossy objects, such as those made of plastic, porcelain, or metal, cannot be segmented accurately by the above method alone. Traditionally, an object is extracted by using a binary mask to determine whether each pixel in a photograph is part of the object or background.

However, if reflections occur on the surface of an object, it becomes difficult to automatically differentiate the object from the background, and object extraction errors occur, as shown in **Fig. 3(b)**. We have overcome this problem by substituting a 256-

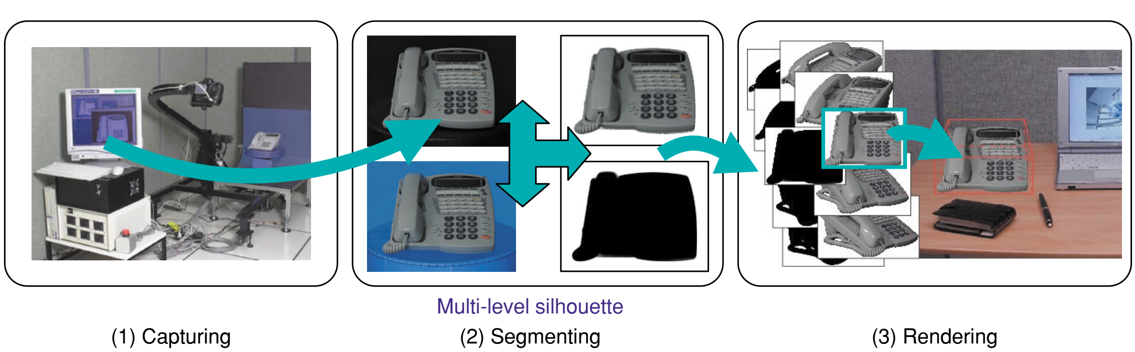
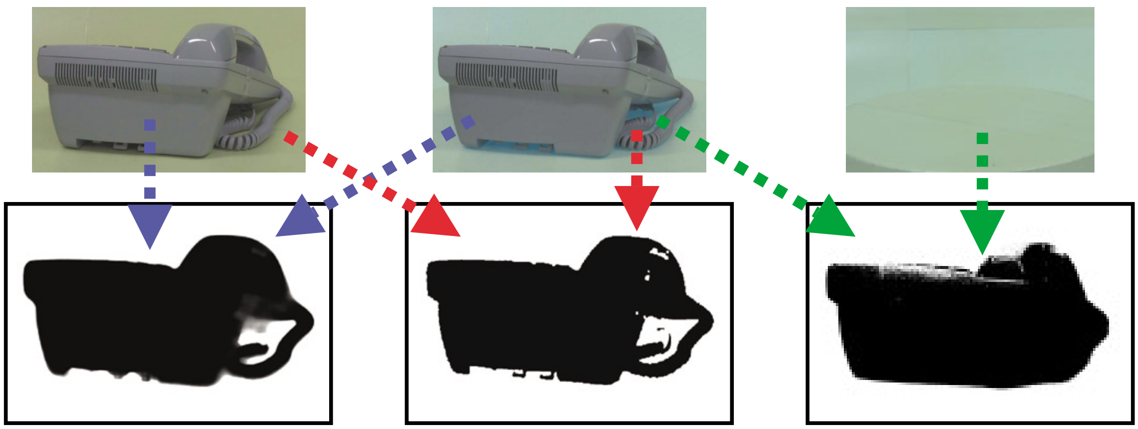
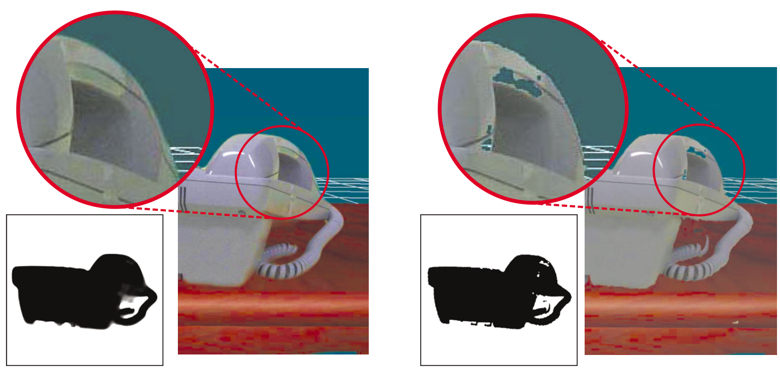



Fig. 2. Generation procedure.



(a) Multi-level silhouette obtained by comparing two background colors. (b) Binary mask obtained by comparing two background colors. (c) Binary mask obtained using chromakey.

Fig. 3. Segmentation algorithms.



(a) Multi-level silhouette (b) Binary mask

Fig. 4. Synthetic images obtained using a multi-level silhouette and a binary mask.

tone multi-level silhouette for the conventional binary mask as shown in **Fig. 3(a)**. With this method, parts that cannot be distinguished as either background or object are represented by an ambiguous value. Naturally, in the final CG image those parts will also be ambiguous, but humans have an innate tendency to infer the shapes of objects, so the image appears entirely natural to the human eye. Just by regarding the ambiguous parts as reflections, we can obtain images depicting glossy objects that are extremely realistic. **Figure 4** shows examples of synthetic images obtained using a multi-level silhouette and a binary mask. With this approach, we can automatically generate 3D CG data representing plastic and ceramic objects, which used to be very difficult. The multi-level silhouette method can also represent translucency, making it possible to automatically model translucent objects, such as polypropylene containers and glass, which could never be well rendered before.

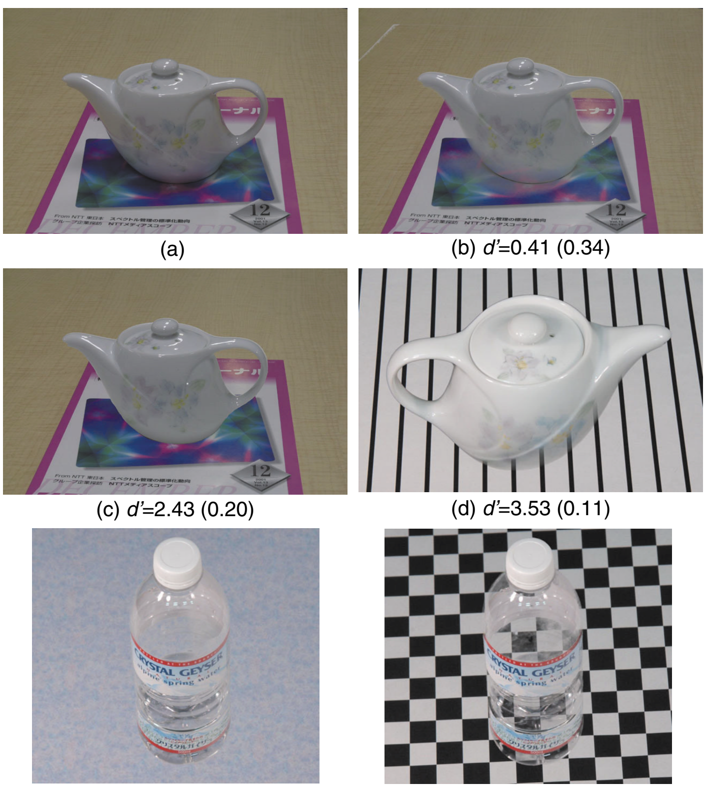
## 3. Psychophysical experiments

Though the multi-level silhouette can represent reflection and transparency on the surface of an object, those representations are not real and may be different from the actual reflection and transparency. The question is whether the reflection and transparency on the object's surface rendered by our method appear natural to human observers.

Using the multi-level silhouette method, we made many images by superimposing different objects having different surface materials against various backgrounds (**Fig. 5**) and conducted psychophysical experiments to evaluate their appearance. In the experiments, 20 participants ranked their confidence regarding an image being real or synthetic. The choices were "definitely real", "maybe real", "unsure", "maybe synthetic", and "definitely synthetic". Some of the images are shown in Fig. 5. Figure 5(a) is a real image of a teapot placed on a magazine.

Figure 5(b) is a synthetic image of the teapot and the magazine created using a multi-level silhouette: the background pattern (the texture of the magazine) passes through (i.e., is not inter-reflected) the surface of the teapot. Figure 5(c), in which inter-reflections on the teapot were erased, was made for comparison. Figure 5(d) is a synthetic image in which the teapot is superimposed by our method against a background of high-contrast lines. Figures 5(e) and (f) are also images made by our method: they show a translucent bottle superimposed against textured patterns.

Values of the sensitivity of discriminating between real and synthetic images, *d'*, calculated for the 20 participants' responses are also shown in Fig. 5 for each image. The *d'* is higher for (c) and (d), which means that the participants easily recognized that they were synthetic. The high *d'* for (c) suggests that inter-reflection is an important factor in making a synthetic image appear realistic. The lower *d'* for (b), (e), and (f) means that those images were often judged to be "real". Thus, in those cases, our method works well, though it should be noted that they were not always judged as real. The high *d'* for (d) suggests that our method has some limitations. For this image, the background pattern was perceived to penetrate the object, which looks very unusual.



(a) (b) *d'*=0.41 (0.34)

(c) *d'*=2.43 (0.20) (d) *d'*=3.53 (0.11)

(e) *d'*=0.74 (0.32) (f) *d'*=0.59 (0.28)

The values in the parentheses denote standard errors.

Fig. 5. Examples of a real image and five synthetic images.

## 4. Alpha Blending with Blurred Background (αBBB)

Through other observations, we found that when the background pattern contained sharp edges, as in Fig. 5(d), it was often perceived to penetrate and pass through the object. Conversely, when there were no sharp edges in the inter-reflected components, the image appeared to be more realistic. This led us to develop the (αBBB rendering algorithm, which removes high-spatial frequency components in the penetrating (or reflected) background image by applying a low-pass filter. **Figure 6** shows how this algorithm works. The surroundings of the object are projected clear-

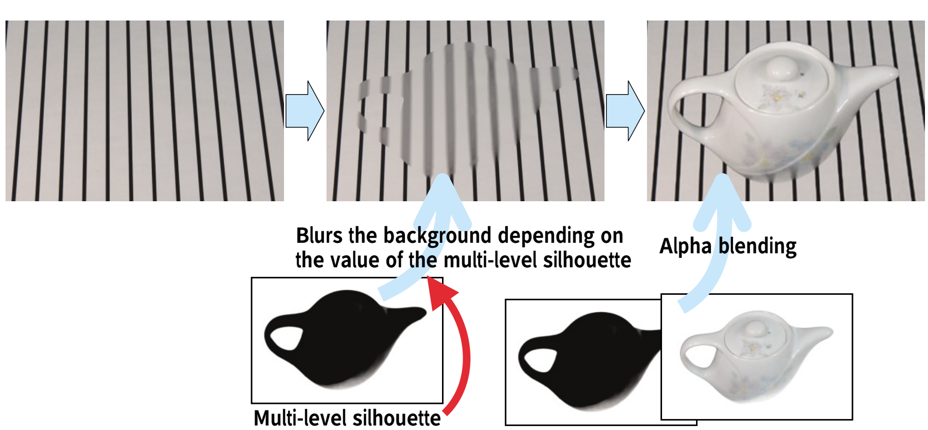
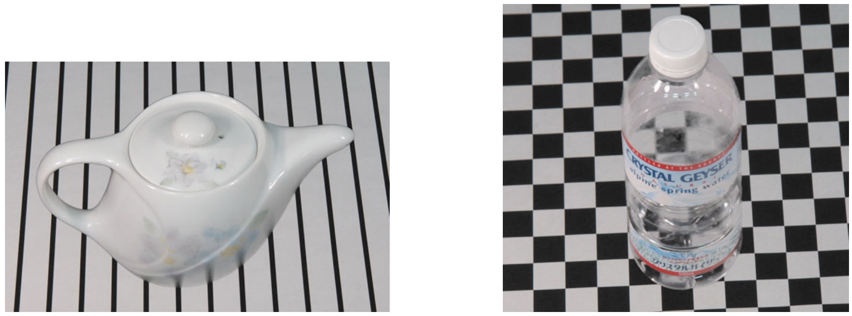



Fig. 6. Algorithm for αBBB.



(a) $d'$=0.46 (0.23)    (b) $d'$=0.50 (0.27)

Fig. 7. Examples of synthetic images made using alpha blending with blurred background.

ly on its surface if the reflection and transparency are strong, while they are projected less clearly if they are weaker. Thus, the gain of the low-pass filter is varied depending on the strength of inter-reflectance or transparency of the object.

Synthetic images made using the method described above were evaluated by the same 20 participants as before. Examples of the synthetic images are shown in **Fig. 7**. In Fig. 7(a), patterns that penetrated from the background in Fig. 5(d) were low-passed filtered. Similar filtering was applied to the image of the translucent bottle in Fig. 7(b). In both cases, *d'* was lower than that for Fig. 5(d), which suggests that the participants judged these synthetic images to appear more natural than those without low-pass filtering.

## 5. Results

**Figure 8** shows a virtual world made from 3D CG data generated using the method described above. All of the objects, excluding the table and floor, were generated automatically by our system. The synthetic images with pseudo inter-reflection and transparency look very natural. The images were generated in real time. It took at most 66 ms to render each of these two virtual world images using a personal computer (CPU: 2-GHz Pentium4, memory: 512 Mbytes, and graphics card: NVIDIA GeForce2).

Fig. 8. Examples of synthetic images.

## 6. Conclusion

We have developed a photorealistic three-dimensional (3D) capture system that can automatically generate photorealistic 3D computer graphics (CG) data based on a real object very quickly. Two important methods made this development possible: an accurate segmentation method using an electroluminescent (EL) sheet and Alpha Blending with Blurred Background (αBBB), which is a rendering algorithm that utilizes the characteristics of human vision. These methods enable synthetic CG images to be made quickly at a lower cost. They can also be applied to various objects having different surface materials, such as plastic, porcelain, translucent glass, or acrylic resin. The system can be used to render virtual shopping malls for e-commerce and to create digital museums. It also has many other potential uses. Our rendering algorithm requires little computational power, so it can achieve realtime rendering on a personal computer. One application of our

Fig. 9. Demonstration of realtime rendering at an exhibition.

method, in which 3D CG data is superimposed on a video image in real time is shown in **Fig. 9**. This application has been introduced at several exhibitions.

## References


[1] A. Laurentini, "The Visual Hull Concept for Silhouette-Based Image Understanding," IEEE Trans. Pattern Anal. Machine Intell., Vol. 16, No. 2, pp. 150-162, 1994.

[2] W. Niem and R. Buschmann, "Automatic Modelling of 3D Natural Objects from Multiple Views," European Workshop on Combined Real and Synthetic Image Processing for Broadcast and Video Production, 1994.

[3] W. Niem, "Robust and Fast Modelling of 3D Natural Objects from Multiple Views," SPIE, Vol. 2182, pp. 388-397, 1994.

[4] L. McMillan and G. Bishop, "Plenoptic Modeling," Proc. of SIGGRAPH'95, pp. 39-46, 1995.

[5] M. Levoy and P. Hanrahan, "Light Field Rendering," Proc. of SIGGRAPH '96, pp. 31-42, 1996.

[6] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The Lumigraph," Proc. of SIGGRAPH '96, pp. 43-54, 1996.

[7] P. Debevec, G. Borshukov, and Y. Yu, "Efficient View-Dependent Image-Based Rendering with Projective Texture-Mapping," The 9th Eurographics Workshop on Rendering, pp. 105-116, 1998.

[8] P. Debevec, "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography," Proc. of SIGGRAPH '98, pp. 189-198, 1998.

[9] D. E. Zongker, D. M. Werner, B. Curless, and D. H. Salesin, "Environment Matting and Compositing," Proc. of SIGGRAPH '99, pp. 205-214, 1999.

[10] Y. Chuang, D. E. Zongker, J. Hindorff, B. Curless, D. H. Salesin, and R. Szeliski, "Environment Matting Extension: Towards Higher Accuracy and Real-Time Capture," Proc. of SIGGRAPH 2000, pp. 121-130, 2000.

[11] A. McNamara and A. Chalmers, "Seeing is Believing: Reality Perception in Modeling, Rendering, and Animation," SIGGRAPH 2001 Courses Notes 21, pp. 1-68, 2001.

[12] H. Matsuoka, T. Takeuchi, H. Kitazawa, and A. Onozawa, "Representation of Pseudo Inter-reflection and Transparency by Considering Characteristics of Human Vision," Proc. of Eurographics 2002, pp. 503-510, 2002.

[13] http://www.contents4u.com/

[14] http://www.seiko-p.co.jp/E/device/lumi.html


**Hiroto Matsuoka**

Research Engineer, Ubiquitous Interface Laboratory, NTT Microsystem Integration Laboratories.

He received the B.S. degree in electrical and electronic engineering from Ritsumeikan University, Kyoto. From 1989 to 1992, he worked at NEC Kansai. He received the M.S. degree from the Graduate School of Information Science of the Nara Institute of Science and Technology, Nara in 1995. In 1995, he joined NTT and researched circuit designs for microwave equipment. He is now engaged in R&D of three-dimensional computational graphics, especially modeling and rendering algorithms.

**Tatsuto Takeuchi**

Senior Research Scientist, Human and Information Science Laboratory, NTT Communication Science Laboratories.

He received the B.A. and M.A. degrees from Kyoto University, Kyoto and the Ph.D. degree in experimental psychology from the University of Tokyo, Tokyo. He performed postdoctoral research at the University of California at Berkeley. In 1995, he joined NTT Basic Research Laboratories in Atsugi (now NTT Communication Science Laboratories). He is currently working on basic and applied research concerning visual motion and color perception. He is a member of the Association for Research in Vision and Ophthalmology, the Japanese Psychological Association, and the Vision Society of Japan.

**Ikuo Harada**

Senior Research Engineer, Supervisor, Ubiquitous Interface Laboratory, NTT Microsystem Integration Laboratories.

He received the B.S., M.S., and Ph.D. degrees in electronic engineering from Osaka University, Suita, Osaka in 1981, 1983, and 1998, respectively. In 1983, he joined the Atsugi Electrical Communication Laboratories, Nippon Telegraph and Telephone Public Corporation (now NTT), Kanagawa, Japan, where he engaged in research on VLSI layout algorithms. In 2000, he moved to ATR Media Integration & Communications Research Laboratories, Kyoto, Japan, and worked on Kansei information processing. In 2002, he returned to NTT. His current research interests include human-computer interactions and CG content creation. He is a member of IEEE, the Institute of Electronics, Information and Communication Engineers, and the Information Processing Society of Japan.