Systems and Computers in Japan, Vol. 36, No. 14, 2005
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J84-D-II, No. 8, August 2001, pp. 1864–1872

Fast Image Synthesis of Virtual Objects in a Real Scene with Natural Shadings

Imari Sato,[1] Morihiro Hayashida,[2] Fumiyo Kai,[3] Yoichi Sato,[4] and Katsushi Ikeuchi[5]

[1]The National Institute of Informatics, Tokyo, 101-8430 Japan

[2]Bioinformatics Center, Institute for Chemical Research, Kyoto University, Kyoto, 611-0011 Japan

[3]Sekisui Chemical Co., Ltd., Urban Infrastructure & Environmental Products Headquarters, Kyoto, 601-8105 Japan

[4]Institute of Industrial Science, The University of Tokyo, Tokyo, 153-8505 Japan

[5]Interfaculty Initiative in Information Studies, The University of Tokyo, Tokyo, 113-0033 Japan

SUMMARY

We propose a new method for superimposing synthetic objects with natural shadings and cast shadows onto a real scene whose illumination condition is dynamically changing in this paper. In general, high computational cost for rendering virtual objects with convincing shading and shadows, such as interreflections or shadows under area light sources, prohibits real-time synthesis of such composite images with superimposed virtual objects. To overcome this limitation, we take advantage of the linearity of the relationship between brightness changes observed on an object surface and change of illumination radiance in a scene and introduce an efficient rendering technique based on a linear combination of pre-rendered reference images of the scene. We have successfully tested the proposed method in a natural illumination condition of an indoor environment to demonstrate how fast the method could superimpose a virtual object onto the scene with highly realistic shadings and shadows. © 2005 Wiley Periodicals, Inc. Syst Comp Jpn, 36(14): 102–111, 2005; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/scj.10155

Key words: computer graphics; computer vision; mixed reality; illumination distribution estimation.

1. Introduction

Techniques for merging virtual objects with a real scene attract a great deal of attention in media industries such as film making, television broadcasting, game making, and advertising. The synthesized world called augmented reality allows us to see a real scene with superimposed virtual objects and to handle phenomena not only in the real world but also in a virtual world, while virtual reality technologies immerse a user in a fully computer-generated scene.

To enhance the quality of synthesized images in the augmented reality, three aspects have to be taken into account [16, 17]: *geometry*, *illumination*, and *time*. More specifically, the virtual object has to be positioned at a desired location in the real scene, and the object must appear at the correct location in the image (*consistency of geometry*). Also, shading of the virtual object has to match that of other objects in the scene, and the virtual object must cast a correct shadow, such as a shadow whose charac-

teristics are consistent with those of shadows in the real scene (*consistency of illumination*). Lastly, motions of virtual objects and real objects have to be correctly coordinated (*consistency of time*).

Consistency of geometry and consistency of time have been intensively investigated in previous studies.* In those studies, 3D position sensors of various modalities are used for measuring the position and orientation of a user's viewpoint in real time, so that a virtual object can be superimposed onto the image that the user is looking at [2, 19, 3, 11].

For the goal to achieve consistency of illumination, several methods have been proposed for rendering virtual objects superimposed onto real scenes using real illumination distributions of the scenes that are either directly measured [10, 14, 5, 9, 23] or indirectly estimated [22]. Those methods have been successfully demonstrated to be capable of synthesizing composite images with convincing qualities.

However, all of the previously proposed techniques suffer from a common difficulty to be used in applications that require real-time processing. High computational cost of rendering virtual objects with convincing shading and shadows, for example, interreflections or shadows under area light sources, prohibits real-time synthesis of high-quality composite images with superimposed virtual objects. In the case where composite images have to be generated in real time, simple rendering algorithms supported by a commonly available graphics hardware need to be used, and thus high qualities in the synthesized images are hardly expected.

Alternative approaches have been proposed for re-rendering a scene as a linear combination of a set of pre-rendered basis images of the scene [7, 15, 8]. These approaches are based on the linearity of scene radiance with respect to illumination intensities. Since this linearity holds for scenes with complex geometry and complex photometric properties such as interreflections between objects and cast shadows, photo-realistic appearance of a scene can be synthesized based on this simple linear combination operation.

Most of the previously proposed methods, however, have been developed for the task of interactive lighting design. Therefore, basis lights under which a set of basis images are rendered are intentionally positioned at the desired locations, so that a scene under desired lighting configurations can be efficiently synthesized. Recently, in the context of synthesizing photo-realistic appearance of human skin, Debevec and colleagues introduced a method for re-rendering a human face based on a linear combination of face images taken under densely sampled incident illumination directions [6]. This method further considered a model of skin reflectance to estimate the appearance of the face seen from novel viewing directions and under arbitrary illumination.

In this study, we generalize the approach based on the linearity of scene radiance with respect to illumination radiance and present an efficient technique for superimposing synthetic objects with natural shadings and cast shadows onto a real scene whose illumination is dynamically changing: we consider that a scene consists of both real objects and synthetic objects with fixed scene geometry, and the image of a scene viewed from a fixed viewing point is synthesized under dynamically changing illumination.

2. Linearity between Scene Radiance and Illumination Irradiance

Let us first describe the linearity of the relationship between change of brightness observed on an object surface and change of illumination radiance in a scene in this section.

Consider one case where a light source L_1 illuminates an object [Fig. 1(a)] and another case where another light source L_2 placed at a different location illuminates the same object [Fig. 1(b)]. If we let I_1 be the image observed in the first case and I_2 be the image observed in the second case from the same viewpoint, a novel image I, for which both L_1 and L_2 illuminate the object, can then be synthesized as the sum of I_1 and I_2 [Fig. 1(c)]:

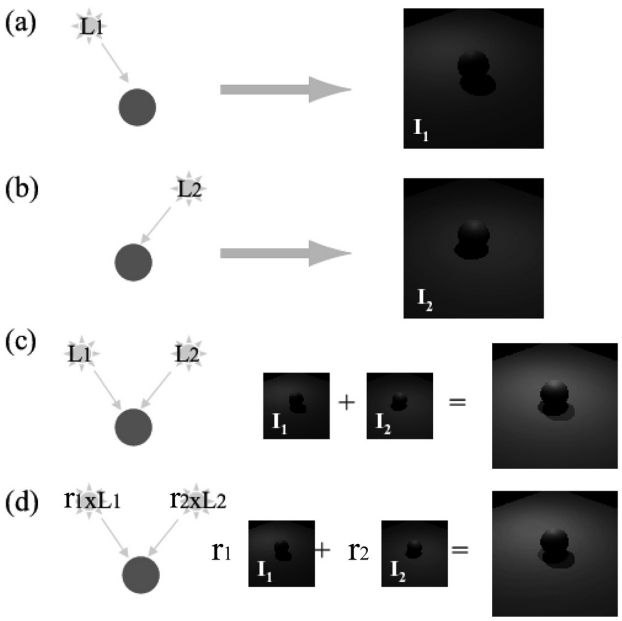

Fig. 1. Linearity between illumination radiance and scene radiance.

*For a good survey of augmented reality technologies, see Ref. 1.

$$I = I_1 + I_2 \tag{1}$$

Similarly, we can obtain an image I', which should be observed when the radiance values of the light sources L_1 and L_2 change by considering a linear combination of I_1 and I_2 as

$$I' = r_1 I_1 + r_2 I_2 \tag{2}$$

where r_1 and r_2 are scaling factors representing the changes of the radiance values of L_1 and L_2, respectively [Fig. 1(d)]. For instance, if the radiance of L_1 decreases by half, then r_1 is set to 0.5. It should be noted that this linearity holds not only for direct illumination from lights but also for indirect illumination such as interreflections between or among objects.

Here distant illumination of the scene is assumed; light sources in the scene are sufficiently distant from the objects, and thus all light sources project parallel rays onto the object surface. Namely, the distances from the objects to the light sources are not considered.

2.1. Approximation of illumination distribution

We generalize the approach based on the linearity of scene radiance with respect to illumination radiance further so that synthetic objects can be synthesized with natural shadings and cast shadows onto a real scene under natural illumination conditions of a real scene.

In order to take illumination from all directions into account, let us consider an infinitesimal patch of the extended light source, of a size $\delta\theta_i$ in polar angle and $\delta\phi_i$ in azimuth as shown in Fig. 2. Seen from the center point A, this patch subtends a solid angle $\delta\omega = \sin\theta_i \delta\theta_i \delta\phi_i$. Let $L_0(\theta_i, \phi_i)$ be the illumination radiance per unit solid angle coming from the direction (θ_i, ϕ_i); then the radiance from the patch is $L_0(\theta_i, \phi_i) \sin\theta_i \delta\theta_i \delta_i \phi_i$ and the total irradiance of the surface point A is [12]

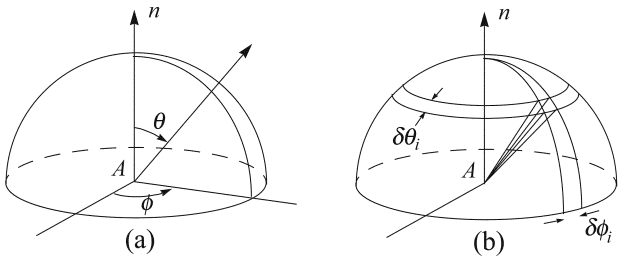

Fig. 2. (a) The direction of incident and emitted light rays. (b) Infinitesimal patch of an extended light source.

$$E = \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} L_0(\theta_i, \phi_i) \cos\theta_i \sin\theta_i \mathrm{d}\theta_i \mathrm{d}\phi_i \tag{3}$$

Then the illumination distribution is approximated by discrete sampling of its radiance over the entire surface of the extended light source. This can be considered as representing the illumination distribution of the scene by using a collection of light sources with an equal solid angle. As a result, the double integral in Eq. (3) is approximated as

$$E = \sum_{i=1}^{n} L(\theta_i, \phi_i) \cos\theta_i \tag{4}$$

where n is the number of sampling directions, $L(\theta_i, \phi_i)$ is the illumination radiance coming from the direction (θ_i, ϕ_i), which also includes solid angle ω_i for the sampling direction (θ_i, ϕ_i).

For instance, node directions of a geodesic dome can be used for uniform sampling of the illumination distribution. By using n nodes of a geodesic dome in a northern hemisphere as a sampling direction, the illumination distribution of the scene is approximated as a collection of directional light sources distributed with an equal solid angle $\omega_i = 2\pi/n$.

Some of the incoming light at point A is reflected toward the image plane. As a result, point A becomes a secondary light source with scene radiance. The bidirectional reflectance distribution function (BRDF) $f(\theta_i, \phi_i; \theta_e, \phi_e)$ is defined as the ratio of the radiance of a surface as viewed from the direction (θ_e, ϕ_e) to the irradiance resulting from illumination from the direction (θ_i, ϕ_i). Thus, by summing up the product of the BRDF and the illumination radiance, the scene radiance $I(\theta_e, \phi_e)$ viewed from the direction (θ_e, ϕ_e) is computed as

$$\begin{aligned} I(\theta_e, \phi_e) &= \sum_{i=1}^{n} f(\theta_i, \phi_i; \theta_e, \phi_e) L(\theta_i, \phi_i) \cos\theta_i \\ &= I_1 + I_2, \ldots, I_n \end{aligned} \tag{5}$$

where $I_i = f(\theta_i, \phi_i; \theta_e, \phi_e) L(\theta_i, \phi_i) \cos\theta_i$.

Similar to Eq. (2), we obtain an image I' which should be observed when the radiance values of the light sources L_i change by r_i as a linear combination of I_i as

$$\begin{aligned} I(\theta_e, \phi_e) &= \sum_{i=1}^{n} r_i f(\theta_i, \phi_i; \theta_e, \phi_e) L(\theta_i, \phi_i) \cos\theta_i \\ &= r_1 I_1 + r_2 I_2, \ldots, r_n I_n \end{aligned} \tag{6}$$

3. Overview of Proposed Approach

Taking advantage of this linear relationship between brightness observed on an object surface and radiance values of light sources in a scene, the proposed method synthesizes a new image for novel lighting conditions as described in the following steps.

Step 1: The entire illumination of a scene is approximated as a collection of area sources $L_i(i = 1, 2, \ldots, n)$ which are equally distributed in the scene. In our current implementation of the proposed method, a set of triangular patches of a geodesic dome are used as a collection of area light sources $L_i(i = 1, 2, \ldots, n)$ that approximate the entire illumination distribution (Fig. 3.1).

Step 2: Two images which are referred to as reference images are rendered under each area light source: one with a virtual object superimposed onto the scene O_i, and the other without the object S_i (Fig. 3.2).

Step 3: Scaling factors of the light source radiance values $r_i(i = 1, 2, \ldots, n)$ are measured by using an omnidirectional image of the scene taken by a camera with a fisheye lens (Fig. 3.3).

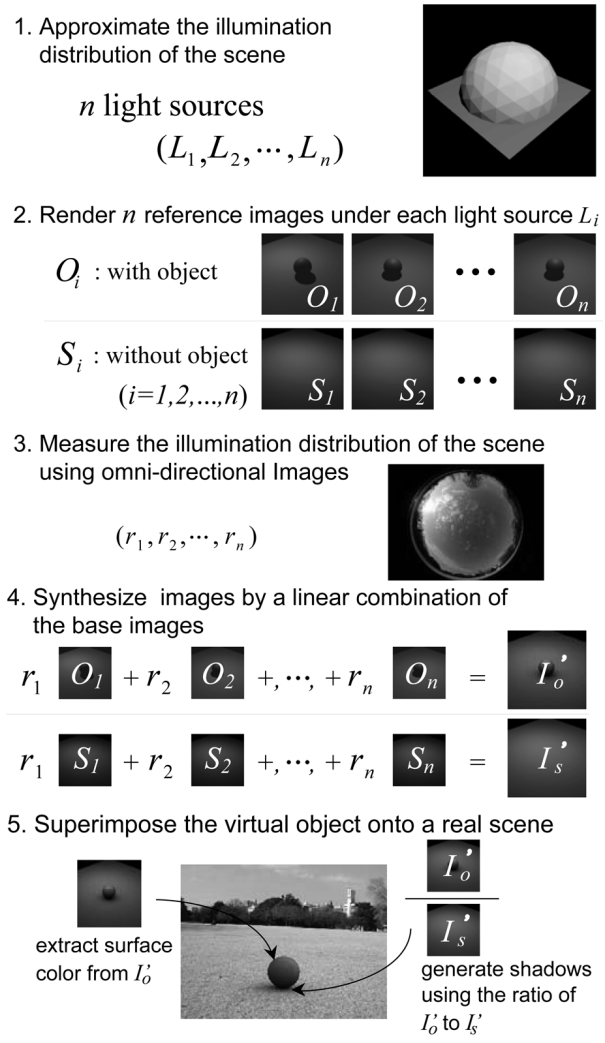

Fig. 3. Basic steps of the proposed method.

Step 4: New images I'_o and I'_s, which should be observed under the current illumination condition, are obtained as a linear combination of O_i's and S_i's with the measured scaling factors r_i's, respectively (Fig. 3.4).

Step 5: Using I'_o and I'_s, the virtual object is superimposed onto the image of the scene along with natural shading and shadows that are consistent with those of real objects (Fig. 3.5). The ray casing algorithm is imposed here; if an image pixel corresponds to the virtual object surface, the color of the corresponding pixels in I'_o is assigned as the value of the pixel. Otherwise, the effects on the real objects caused by the virtual object, that is, shadows and secondary reflection, are added by multiplying the pixel value by the ratio of I'_o to I'_s (Fig. 3.5).

The main advantage of the proposed method is that composite images are synthesized by simple linear operations based on reference images pre-rendered as an off-line process. From this the quality of output images is not affected by the real-time processing requirement on-the-fly.

In the following, Section 4 explains how to prepare the reference images, and Section 5 describes a method for measuring an illumination distribution of a real scene to determine the radiance scaling factor r_i's. Then Section 6 explains the entire process of superimposing a virtual object onto the real scene by using our proposed method.

4. Rendering Reference Images

The image acquisition system used in our experiment is shown in Fig. 4.* A 3CCD color camera (*Camera A*) is

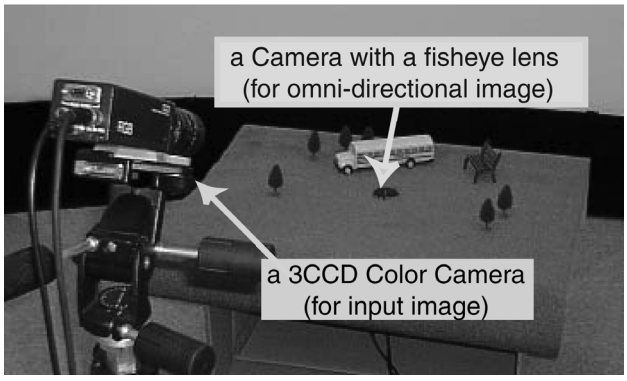

Fig. 4. Image acquisition system. A 3CCD color camera is used for taking an image of the scene, and another color camera with a fisheye lens is used for capturing illumination of the scene.

*A similar image acquisition setup was used in Ref. 13 to superimpose synthetic objects onto a real scene under real-world illumination.

used for taking an image of a real scene onto which a virtual object is superimposed. In addition, another color camera with a fisheye lens (*Camera B*) is used for taking an omnidirectional image of the scene which is then used for determining the illumination distribution of the scene.

First, we estimate the camera parameters of Camera A which will be later used for rendering synthetic images. This is necessary for making virtual objects appear at a desired position in final composite images, that is, *consistency of geometry*. In our method, this camera calibration was done by using Tsai's camera calibration method [21]. This method is known to be able to estimate camera parameters reliably by taking into account various effects causing image distortions.

Then reference images are rendered using the rendering software called *Radiance*, which can accurately simulate reflections of lights based on physical models of lights [24]. Using Radiance, the following two reference images are rendered off-line for each area light source L_i.

1. *Object images* O_i are images of a scene that contains a virtual object [Fig. 5(a)].

2. *Surface images* S_i are images of the same scene without the virtual object [Fig. 5(b)].

For rendering those reference images, we need to know geometric and photometric models of the real scene, that is, shape and reflectance properties of real objects near a virtual object. Recovery of the geometric and photometric model of a real object is not an easy task, and it has been intensively studied in the past. Since geometric and photometric modeling of a real object is not the main focus of our study, those models used for our proposed method are manually specified by a user.

In addition to the *object image* and the *surface image*, another image which we call the *mask image* is prepared. The mask image is used for determining whether a ray extending from the camera projection center Camera A through each image pixel intersects a virtual object or a real object in the scene.

(a) (b) (c)

Fig. 5. Reference image. (a) *Shadow image*, (b) *surface image*; (c) *mask image*.

5. Measuring Illumination Distribution

Next, we describe how to measure radiance scaling factors r_i which represent an illumination distribution of a scene by using an omnidirectional image taken by a camera with a fisheye lens shown in Fig. 4.*

The fisheye lens used in our system is designed in such a way that an incoming ray to the lens is projected onto a certain location in an omnidirectional image according to a desired projection model. Therefore, for each area light source, projection of three vertices of the area light source patch can be determined based on this projection model of the lens. Then the radiance scaling factor for the area light source is computed by using the average pixel values inside the projection of the area light source in the omnidirectional image.

Note that we also need to determine an additional scaling factor between a unit radiance value used for rendering the reference images and radiance scaling factor r_i that are measured from the omnidirectional image as described above. In our experiment in Section 7, this scaling factor is manually specified so that surface color of a virtual object which is rendered using measured radiance scaling factors becomes a certain value.

In order to automatically determine this scaling factor, it is necessary to perform a photometric calibration between unit radiance value used for rendering reference images and illumination radiance measured from an omnidirectional image.

6. Superimposing Virtual Objects onto a Real Scene

After the radiance scaling factors r_i are determined, the virtual object is superimposed onto the real scene with consistent shading under the captured illumination condition. Note that each of the three color bands (R, G, B) of a composite image is separately synthesized using the corresponding band of the reference images and the scaling factor r_i's. For simplicity of discussion, we do not describe color bands explicitly in this section.

First, a new *object image* and a new *surface image* that should be observed in the measured illumination condition are synthesized as a linear combination of the corresponding reference images O_i's and S_i's, respectively. Let the new object image be I_o' and the new surface image be I_s'; then those images are obtained as

$$I_o' = \sum_{i=1}^{n} r_i O_i \tag{7}$$

*3CCD color camera Victor KYF-57 and fisheye lens (Fit Corporation FI-19 with field of view of 180°).

$$I_s{}' = \sum_{i=1}^{n} r_i S_i \tag{8}$$

Then, for each image pixel in the composite image, we determine whether the pixel corresponds to the virtual object or to a real object in the scene by using the previously prepared mask image. If a pixel corresponds to a point on the virtual object surface, a color to be observed at this point is extracted from the same pixel coordinate of the I'_o and stored in the composite image.

Otherwise, we modify an observed color at the point on the real object surface, so that effects caused by the virtual object onto the real object are represented. For this purpose, the ratio of I'_o to I'_s is first computed. The ratio represents how much of the color at the real object surface should be modified if the virtual object was placed in the scene. Let P be the observed color of the image pixel of Camera A; then the color P' that should be observed if there were a virtual object in the scene is computed by using the ratio

$$P' = P\frac{I_o{}'}{I_s{}'} \tag{9}$$

7. Experimental Results

We have tested the proposed method by using real images of an indoor environment as shown in Fig. 4. In this experiment, SGI Onyx2 with six CPUs was used for capturing a sequence of input images and for synthesizing final composite images. For efficient generation of the composite images, the following two strategies were employed.

1. As described in Section 6, each pixel in the composite image requires a different type of operation, depending on which object surface the pixel corresponds to. Computational cost of synthesizing each image pixel in the composite image is taken into account for distributing computation evenly among all six CPUs.

2. Because effects due to the light sources with low radiance values are negligible in the final composite image, the reference images rendered under light sources whose radiance values are lower than a prefixed threshold value are omitted when a linear combination of the reference images is computed. In this way, we can reduce the number of reference images used for synthesizing final composite images, and achieve required processing time by adjusting the threshold value.

Several examples of synthesized composite images are shown in Fig. 6: (a) represents input images, (b) represents synthesized composite images, and (c) represents omnidirectional images that show changes in the real illumination distribution as time passed. In the composite images shown in this figure, appearance of the virtual object blends well into the scene, and the virtual object casts a

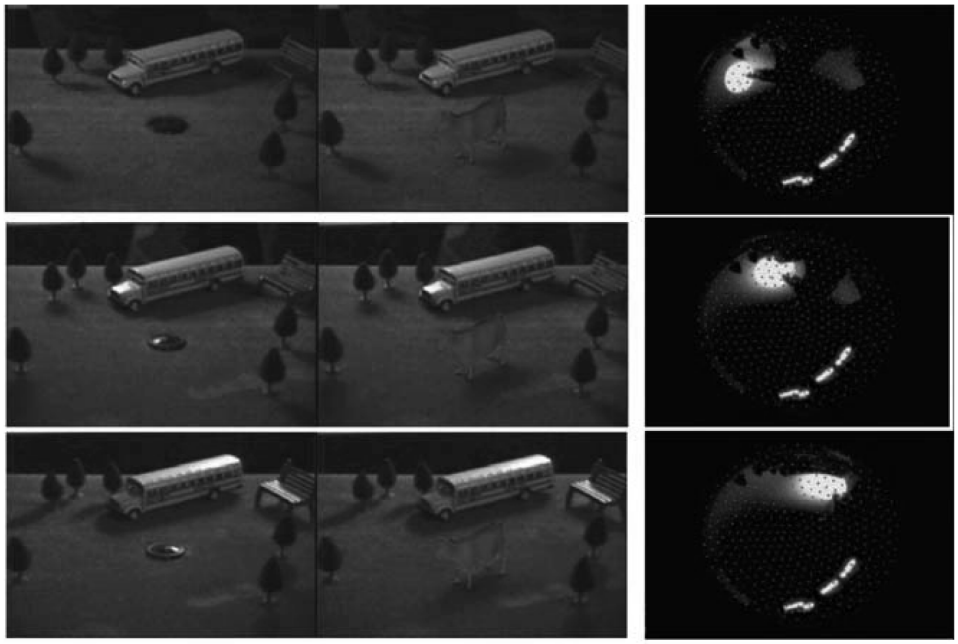

Fig. 6. Synthesized results. (a) Input images; (b) synthetic images; (c) omnidirectional images.

shadow with a soft edge on the grass in the same way as the other objects such as trees in the scene do. Table 1 shows changes in processing time due to the number of reference images.*

Since the composite images are synthesized not by rendering the whole scene over and over but rather based on the linear combination of pre-rendered reference images, the processing time for synthesizing the composite images never affects the achievable image quality. In this example, we approximated the real illumination distribution by using 400 light sources, and light sources whose radiance is greater than a threshold value are selected for the rendering process. In spite of the highly realistic shading and shadows achieved in the composite image, the virtual object was able to be superimposed onto the scene at a frame rate of approximately 3 to 4 frames per second.

In contrast, when light sources were not selected according to the threshold value, it took 1.2 seconds to generate one synthetic image. Here if we use a fewer number of light sources that approximate the real-world illumination distribution in order to speed up the rendering process, the solid angle of each light source ($\delta\omega = 2/n$) increases and this results in blurring the shading and cast shadows of the virtual object, which might differ from those of real objects.

In the experimental example mentioned above, the number of basis images was significantly reduced based on the proposed light source selection scheme and a high frame rate was obtained. However, when the number of basis images to be used for synthesis processing cannot be effectively reduced by this light source selection, a number of basis images need to be used and the synthesis processing speed must be compromised to some extent. We compared this frame rate with that of the conventional rendering technique by assigning a set of light source radiance from the omnidirectional image of the scene. Here all of 400 light sources are used for rendering the scene. We found that it took at least 8 minutes to render natural shading of the virtual object by using the Radiance rendering software on the same SGI-Onyx2 system. Even when light sources were selected based on their radiance value, it took at least 5 minutes to render it.

It was apparent from this that by using the proposed approach, the synthesis processing speed was increased approximately 400 times when light sources were not selected and approximately 1200 times when light sources were selected.† One can say that the proposed technique generated high-quality synthetic images much faster than the conventional technique even when a large number of light sources were used to approximate the illumination distribution of a scene and many basis images were used for the image synthesis process.

Table 1. Processing time due to the number of reference images

No. of reference images	Processing time (s)
6	0.08
25	0.2
90	0.33
190	0.55
400	1.2

In the experiment shown in Fig. 7, the 3D shapes and reflectance properties of the surrounding real objects were modeled by using a photo-modeling tool called Canoma [19]. In Fig. 7, the white hemisphere placed in the center of these images is a synthetic object, and the other blocks are real objects. In the images synthesized by the proposed technique, we see that the brightness of the virtual object changed in a similar manner as the other blocks due to changes in the real-world illumination distribution. In the resulting image on the right, the shadows of the real objects fall naturally on the virtual object. In the synthetic image on the left, interreflections are observed between the virtual object and the cube-shaped block located diagonally in front of the virtual object. A color blending phenomenon in which the color of the cube is reflected on the virtual object is also synthesized correctly.

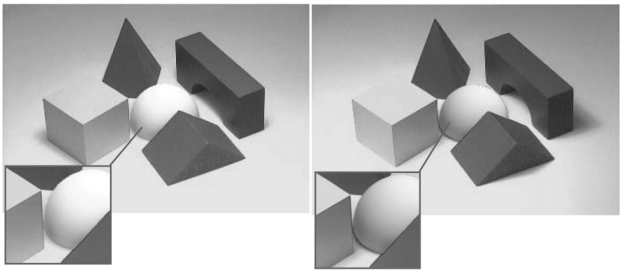

Fig. 7. The 3D shape and reflection properties of the real objects were modeled and used for rendering basis images. The shadows of the real objects fall naturally on the virtual object, and the interreflections are observed between the virtual object and the real cube-shaped block.

*As described before, the number of reference images is equal to the number of area light sources that approximate the entire illumination distribution of the scene.

†However, this comparison does not take into consideration the off-line processing time required when generating the base images. The comparison is simply performed in terms of the synthesis processing time.

8. Discussion

There is one limitation in the approaches based on the linearity of scene radiance with respect to illumination radiance. That is, an image of the scene under novel illumination condition can be synthesized accurately unless this illumination condition lies in the linear subspace spanned by the basis lights under which basis images are rendered. Therefore, it is essential to provide basis lights and their associated weights properly so that images under desirable illumination conditions can be synthesized.

Previous studies investigated how to choose a set of basis lights so that these lights could efficiently represent specific lighting configurations. For instance, Nameroff and colleagues introduced a set of steerable area light sources as basis lights to approximate the illumination effect of daylight [15]. Dobashi and colleagues defined a set of basis lights to represent directionality of spotlights based on spherical harmonics [7]. Later, Teo and colleagues introduced a hybrid method for synthesizing illumination effects from both area sources and directional spotlight sources [20]. In addition, this method adopted a strategy for reducing the number of basis images based on principal components analysis. Recently, the effect of defining a set of basis lights in the frequency domain based on spherical harmonics was demonstrated in Refs. 4 and 18.

9. Conclusions

In this paper, we proposed a new method for superimposing virtual objects with natural shading and cast shadows onto a real scene whose illumination condition was dynamically changing. Based on linearity of the relationship between change of brightness observed on an object surface and change of illumination radiance values in a scene, final composite images are synthesized as a linear combination of pre-rendered reference images of virtual objects placed in the scene in our proposed method. The main advantage of our proposed method is that rendering algorithms that are computationally too expensive for real-time processing can be used for rendering reference images since the rendering process is done off-line. Thus, the image quality is not affected by the requirement for real-time processing, and the proposed method is able to satisfy both high image quality and real-time processing.

REFERENCES

1. Azuma RT. A survey of augmented reality. Presence: Teleoperators and Virtual Environments. 1997;6:355–385.
2. Azuma RT, Bishop G. Improving static and dynamic registration in an optical see-through HMD. Proc SIGGRAPH 94, p 197–204.
3. Bajura M, Fuchs H, Ohbuchi R. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. Proc SIGGRAPH 92, p 203–210.
4. Basri R, Jacobs D. Lambertian reflectance and linear subspaces. Proc IEEE Int Conf Computer Vision '01, p 383–389.
5. Debevec PE. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. Proc SIGGRAPH 98, p 189–198.
6. Debevec P, Hawkins T, Tchou C, Duiker H, Sarokin W, Sagar M. Acquiring the reflectance field of a human face. Proc SIGGRAPH '00, p 145–156.
7. Dobashi Y, Kanade K, Naktani H, Yamashita H. A quick rendering method using basis function for interactive lighting design. Proc Eurographics '95, p 229–240.
8. Dorsey J, Arvo J, Greenberg D. Interactive design of complex time-dependent lighting. J IEEE Computer Graphics and Applications 1995;15:26–36.
9. Drettakis G, Robert L, Bougnoux S. Interactive common illumination for computer augmented reality. Proc 8th Eurographics Workshop on Rendering, p 45–57, 1997.
10. Fournier A, Gunawan A, Romanzin C. Common illumination between real and computer generated scenes. Proc Graphics Interface '93, p 254–262.
11. Fujii H, Kanbara M, Iwasa H, Takemura H, Yokoya N. A registration method using stereo cameras with an inertial sensor for augmented reality. Tech Rep IEICE 2000;PRMU99-192.
12. Horn BKP. Robot vision. MIT Press; 1986.
13. Koudelka M, Belhumeur P, Magda S, Kriegman D. Image-based modeling and rendering of surfaces with arbitrary BRDFs. Proc IEEE Conf Computer Vision and Pattern '01, Vol. 1, p 568–575.
14. Miyama M, Ozawa S, Kondo K. Merging of live action and CG images using estimation of photographic conditions. 2nd Intelligent Information Media Symposium, p 113–118, 1996.
15. Nameroff J, Simoncelli E, Dorsey J. Efficient re-rendering of naturally illuminated environments. Proc 5th Eurographics Workshop on Rendering, 1994.
16. Ohta Y. Topics and expectations in mixed reality techniques. Publication of the First VR Society Mixed Reality Technical Symposium '97.
17. Tamura H. Mixed reality research project. General Conference of the IEICE, ISS-1-4, 1997.

18. Ramamoorthi R, Hanrahan P. A signal-procession framework for inverse rendering. Proc SIGGRAPH '01, p 117–128.
19. State A, Hirota G, Chen DT, Garrett WF, Livingston MA. Superior augmented-reality registration by integrating landmark tracking and magnetic tracking. Proc SIGGRAPH 96, p 429–438.
20. Teo PC, Simoncelli EP, Heeger DJ. Efficient linear re-rendering for interactive lighting design. Report No. STAN-CS-TN-97-60, Department of Computer Science, Stanford University, 1997.
21. Tsai R. A versatile camera calibration technique for high accuracy machine vision metrology using off-the-shelf TV cameras and lenses. IEEE J Robotics and Automation 1987;3:323–344.
22. Sato I, Sato Y, Ikeuchi K. Illumination from shadows. IEEE Trans Pattern Anal Mach Intell 2003;25:290–300.
23. Sato I, Sato Y, Ikeuchi K. Acquiring a radiance distribution to superimpose virtual objects onto a real scene. IEEE Trans Visualization and Computer Graphics 1999;5:1–12.
24. Radiance is Unix freeware for lighting design and rendering developed by the U.S. Department of Energy and the Swiss federal government. http://hobbes.lbl.gov/radiance/

AUTHORS (from left to right)



Imari Sato received her B.S. degree in policy management from Keio University in 1994 and M.S. and Ph.D. degrees in interdisciplinary information studies from the University of Tokyo in 2002 and 2005. She is an assistant professor at the National Institute of Informatics. Previously, she was a visiting scholar at the Robotics Institute at Carnegie Mellon University, Pittsburgh, Pennsylvania. Her research interests include physics-based vision and image-based modeling. She is a research fellow of the Japan Science and Technology Agency.

Morihiro Hayashida received his B.S. and M.S. degrees from the Division of Computer Science in the Department of Physical Science at the University of Tokyo in 2000 and 2002 and Ph.D. degree from the Graduate School of Informatics, Kyoto University, in 2005. He is an assistant professor at the Bioinformatics Center, Institute for Chemical Research, Kyoto University. He is engaged in research on the analysis of gene expression networks.

Fumiyo Kai graduated from the Department of Science, Engineering and Information Science of Saga University in 1992. From 1992 to 1994, she worked on developing simulations of applications for factory production lines and also production systems at the Kyoto Technology Center–Factory Automation Technology Promotion Center of Sekisui Chemical Co., Ltd. From 1994 to 1996, she worked on the development of technology information managing systems and technology information collection systems intended for managers at the Kyoto Technology Center and Technology Planning Headquarters of Sekisui Chemical Co.; and from 1996 to 1998, she was engaged in the development of design systems at the Kyoto Technology Center–CAE Research Laboratories of the same company. Since 1998, she has been engaged in research on computer graphics as a member of the Kyoto Research Institute CAE Group of Sekisui Chemical Co., Ltd.

Yoichi Sato (member) received his B.S. degree in mechanical engineering from the University of Tokyo in 1990 and M.S. and Ph.D. degrees in robotics from the School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, in 1993 and 1997. In 1997, he joined the Institute of Industrial Science at the University of Tokyo, where he is currently an associate professor. His primary research interests are in the fields of computer vision (physics-based vision, image-based modeling), human–computer interaction (perceptual user interface), and augmented reality. His awards include the Yamashita Commemorative Prize in 2000, the International Conference on Shape Modeling and Applications '97 Most Valuable Paper Prize, and the Annual Journal of the Virtual Reality Society of Japan Paper Prize in 2000. He is a member of the Virtual Reality Society of Japan, the Association for Computing Machinery, and IEEE.

Katsushi Ikeuchi (member) received his B.E. degree in mechanical engineering from Kyoto University in 1973 and Ph.D. degree in information engineering from the University of Tokyo in 1978. He is a professor at the University of Tokyo's Institute of Industrial Science. After working at the Artificial Intelligence Laboratory at Massachusetts Institute of Technology, the Electrotechnical Laboratory of the Ministry of International Trade and Industries, and the School of Computer Science at Carnegie Mellon University, he joined the University of Tokyo in 1996. He has received various research awards, including the David Marr Prize in computational vision in 1990 and IEEE R&A K-S Fu Memorial Best Transaction Paper award in 1998. In addition, his 1992 paper, "Numerical Shape from Shading and Occluding Boundaries," was selected as one of the most influential papers to have appeared in the *Artificial Intelligence Journal* within the past 10 years. He is a fellow of IEEE.