

ARBITRARY VIEWPOINTS IMAGE REPRESENTATION BASED ON 3-D GEOMETRIC MODEL AND SPECULAR REFLECTION TRACING

Masaaki Endo, Takahiro Matsushita, Koichi Fukuda, and Akira Kawanaka

Department of Electrical and Electronics Engineering, Faculty of Science and Technology,
Sophia University, Kioi-cho 7-1, Chiyoda-ku, Tokyo

ABSTRACT

In order to develop a system to synthesize arbitrary viewpoint images efficiently, two representative schemes that use the 3-D geometric model to synthesize images have been the focus of attention: model-based rendering and image-based rendering. Image-based rendering is reported to give a better coding performance than the model-based rendering. Although the image-based rendering synthesizes the relatively accurate diffuse reflection component from the camera images at the different viewpoints, it can not synthesize the specular reflection component with sufficient accuracy because the movement of the specular reflection component on the object surface by the viewpoint change is not taken into consideration. Therefore, we propose a method of synthesizing arbitrary viewpoint images by estimating the motion vectors of the specular reflection components for the 3-D model-based synthesized image used as the reference image. Experiments with a carved wooden object show the effectiveness of the proposed method through the coding performance and the subjective evaluation of the synthesized image.

1. INTRODUCTION

There are two representative rendering methods for synthesizing images of a real object in various viewpoints and light conditions: model-based rendering [1] and image-based rendering [2]. The model-based rendering needs the 3-D geometric model and the reflectance characteristics of the object surface. Thus, this method assumes a reflectance model of an object and determines reflectance parameters through detailed reflectance analysis. By using these reflectance parameters, it is possible to integrate synthesized images from the arbitrary viewpoints under the arbitrary illuminations and to make cast shadows efficiently under real light source [3]. However, model-based rendering cannot be applied to complex objects whose reflectance properties cannot be approximated by using a simple reflection model. And it is too difficult to measure these reflection parameters

when changing the direction of the light source and the position of capturing devices. On the other hand, the image-based rendering samples a set of images of the object taken from various viewpoints and under various illuminations. In creating an appropriate environment for computer graphics and virtual reality, arbitrary viewpoint images are synthesized either by selecting camera images from the stored set or by interpolating multi-viewpoint images [2]. Since image-based rendering does not require any detailed analysis of the reflection characteristics of the objects, the method can be applied to any complex objects. Related work has been done by Girod [1][4], who used a 3-D geometric model of a real object to reduce the amount of the data needed to synthesize new images, and compared the model-based coding which leads to texture mapping from camera images with the model-aided coding which performs the disparity compensations among the camera images using a 3-D geometric model. Experimental results applying the model-aided coding show the better coding performance than the model-based coding. However, this method may not accurately represent the specular reflection components of the object accurately, since it doesn't consider the movement of the specular reflection components on the object surface because of the viewpoint change.

In order to overcome these difficulties, we propose a new image synthesis method that uses the motion vectors of the specular reflection components obtained by using the 3-D model-based synthesized image as the reference image for the motion estimation. The block-matching algorithm (BMA), which estimates motion vectors for the fixed-size blocks, is well known as a general approach for estimating motion vectors between two images. Because the regions of the specular reflection components are relatively small and arbitrarily shaped, we introduce a variable block size matching algorithm (VBS-MA). We present the experimental results for a carved wooden object showing coding performance of the proposed method and subjective evaluations of the synthesized image.

2. SYNTHESIZING 3-D MODEL-BASED IMAGES AT ARBITRARY VIEWPOINTS

We consider the synthesis of an arbitrary viewpoint image $f_{\theta}^{(D)}(r)$ using the 3-D geometric model and the camera images $f_{\theta_1}(r)$ and $f_{\theta_2}(r)$, which were taken from the positions close to the desired viewpoint, as shown in Fig.1. Here, θ_1 and θ_2 indicate the direction of the camera images, and the direction of the desired viewpoint is denoted by θ . First we describe the procedure to synthesize the image $f_{\theta}^{(D)}(r)$ using the camera image $f_{\theta_1}(r)$ as follows.

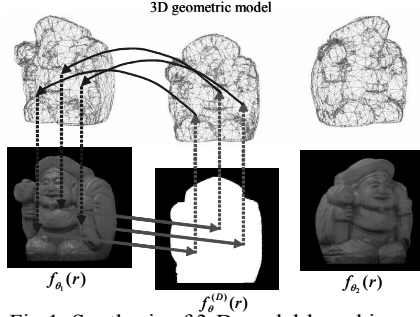


Fig.1 Synthesis of 3-D model-based image.

- (1) Determine a point (u, v) on the image plane, as shown in Fig.2.
- (2) Obtain intersection points of the line passing through the point (u, v) on the image plane from viewpoint \mathbf{V} and triangles constructing the polygonal mesh model.
- (3) Select a point \mathbf{p} , which has minimum distance from viewpoint \mathbf{V} .
- (4) Change the direction to the viewpoint θ_1 .
- (5) Project the point \mathbf{p} on the camera image $f_{\theta_1}(r)$, and obtain the intersection point (u_p, v_p) .
- (6) Interpolate the pixel value by using the four nearest pixel points to the point (u_p, v_p) .

Similarly, based on the procedure described above, the synthesized image $f_{\theta}^{(D)}(r)$ is derived using camera image $f_{\theta_2}(r)$. We obtain the arbitrary viewpoint image $f_{\theta}^{(D)}(r)$ as a weighted average of $f_{\theta}^{(D)}(r)$ and $f_{\theta}^{(D)}(r)$ as shown in Eq.(1).

$$f_{\theta}^{(D)}(r) = \frac{|\theta - \theta_2|}{|\theta_1 - \theta_2|} f_{\theta_1}^{(D)}(r) + \frac{|\theta - \theta_1|}{|\theta_1 - \theta_2|} f_{\theta_2}^{(D)}(r) \quad (1)$$

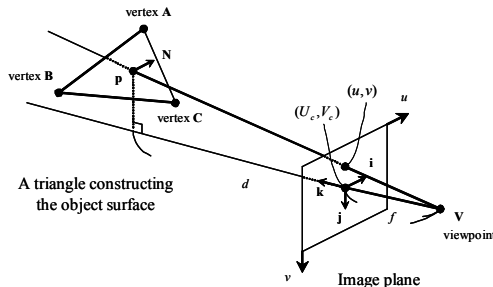


Fig.2 The relation between a triangle constructing the object surface and a pixel on the image plane.

3. SYNTHESIZING IMAGES BY SPECULAR REFLECTION TRACING

3.1 Specular Reflection Tracing by BMA

In the 3-D model-based image synthesis, the light reflection from a point on the 3-D model surface is assumed to be independent of the viewpoint change. The light reflection on the model surface is basically divided into two reflection components: diffuse reflection and specular reflection component. As shown in Fig.3, the intensity of the diffuse reflection component R_d is proportional to the inner product of the direction of the incident light I and the normal vector N of the model surface as shown in Eq.(2).

$$R_d = A_d I \cdot N \quad (2)$$

where A_d is a constant that depends on the intensity of the incident light. On the other hand, the specular reflection component R_s depends on the reflection direction. In the case of an incomplete specular reflection model, the intensity of the specular reflection is represented as follows.

$$R_s = A_s \cos^m(\varphi - \theta) \quad (3)$$

where θ is the incident angle, φ is the reflection angle, $m > 0$ is a constant which determines the divergence of reflection light and A_s is a constant depending on the intensity of the incident light respectively, as shown in Fig.3. Therefore, the diffuse reflection mostly is determined by the light absorption characteristics of the object surface, while the specular reflection changes significantly when the viewpoint changes. In other words, the specular reflection regions move on the target object surface as the viewpoint changes. The 3-D model-based image $f_{\theta}^{(D)}(r)$ is synthesized without reflecting the movement of the specular reflection regions due to the viewpoint changes. In order to represent the specular reflection component correctly, we need to trace the movements of the specular reflection regions.

Therefore, we considered a method to synthesize images using the motion vectors which estimated the movements of specular reflection components using the 3-D model-based image as the reference image for the esti-

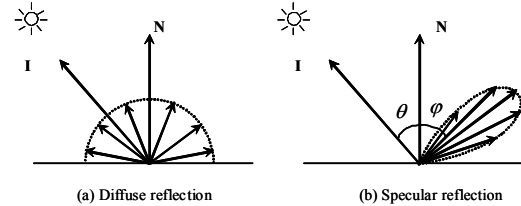


Fig.3 Diffuse reflection and Specular reflection.

mation scheme. In the synthesis method the motion vectors are estimated using the 3-D model-based synthesized image $f_{\theta \leftarrow \theta_1}^{(D)}(r)$ as a reference image. By performing the motion compensation for the estimated motion vectors, the synthesized image from the camera image $f_{\theta_1}(r)$, which denotes $f_{\theta \leftarrow \theta_1}^{(S)}(r)$ is obtained. Similarly another synthesized image $f_{\theta \leftarrow \theta_2}^{(D)}(r)$ is obtained from the camera image $f_{\theta_2}(r)$. We obtain the synthesized image taking the movement of the specular reflection regions into consideration as follows.

$$f_{\theta}^{(S)}(r) = \frac{|\theta - \theta_2|}{|\theta_1 - \theta_2|} f_{\theta \leftarrow \theta_1}^{(S)}(r) + \frac{|\theta - \theta_1|}{|\theta_1 - \theta_2|} f_{\theta \leftarrow \theta_2}^{(S)}(r) \quad (4)$$

BMA is well known as a general approach for estimating motion vectors between two images [5]. BMA estimates motion vectors for the blocks, which are fixed-size subdivisions of the original image. BMA with a fixed block size can not estimate the motion vector for cases in which the boundaries of the specular reflection and diffuse reflection are included in the block and blocking artifacts will become apparent. Because the regions in which the specular reflection is dominant are relatively small and arbitrarily shaped, we introduce a variable block size matching algorithm (VBS-MA), which estimates the motion vectors based on the adaptively divided block model.

3.2 Estimation of Adaptive Block Model and Motion Vectors

To find an adaptive divided block model, the motion vectors found by BMA in some block sizes are used. We divide a block into smaller blocks repeatedly, by comparing the rate-distortion property for each block size, until a minimum block size is reached. The adaptively divided block model is obtained by minimizing a Lagrangian-type cost function represented by Eq.(5).

$$\min\{d_k + \lambda b_k\} \quad (5)$$

where d_k is the matching error in region k , b_k is the amount of coding bit for the motion vectors, and λ is a Lagrangian multiplier. Adaptive division of each block is based on a conditional equation shown in Eq.(6). When Eq.(6) is satisfied, the block X_{n+1} (block size $2^{n+1} \times 2^{n+1}$) is divided into $X_{n,i}$ (block size $2^n \times 2^n$) shown in Fig.4.

$$\Delta d \geq \lambda \Delta b \quad (6)$$

where Δd is the decrease in the matching error obtained by dividing the block and Δb is an increase in the amount of coding bit. Moreover, we modify estimated motion vectors for the obtained block model based on the rate-

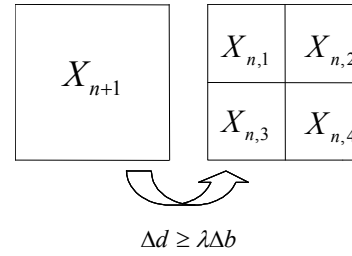


Fig.4 Adaptive block division based on the rate distortion properties.

distortion characteristics, and synthesize the image using these modified motion vectors.

4. EXPERIMENTAL RESULTS

We performed experiments using a carved wooden object in order to evaluate the proposed method of synthesizing the images using the 3-D model and the specular reflection tracing.

In Fig.5 the camera image $f_{\theta=0^\circ}(r)$ of “Daikoku” at the viewpoint $\theta = 0^\circ$ is shown. First the 3-D model-based image $f_{\theta=0^\circ}^{(D)}(r)$ was synthesized from the camera images $f_{\theta_1}(r)$ and $f_{\theta_2}(r)$ at viewpoints θ_1 and θ_2 close to the desired viewpoint θ . The 3-D model-based image $f_{\theta=0^\circ}^{(D)}(r)$ is shown in Fig.7. The diffuse reflection components of $f_{\theta=0^\circ}^{(D)}(r)$ can be synthesized smoothly, but the specular reflection components that appear as the high light regions are represented with displacement and deformation compared to the original image $f_{\theta=0^\circ}(r)$.

So we estimate the motion vectors of the specular reflection components using the 3-D model-based images as the reference image and synthesize the image as a desired viewpoint using the estimated motion vectors. The motion vectors were estimated by using VBS-MA considering the distortion of the synthesized image and the coding rate of the motion vectors. Figs.8 and 9 show the adaptively divided block models obtained by the motion compensation and the motion vectors estimated by VBS-MA, respectively. In Fig.8 and Fig.9, image regions in which the specular reflection components were dominant were divided into smaller blocks and the movements of specular reflection components can be traced as motion vectors. Also, the image $f_{\theta=0^\circ}^{(S)}(r)$ synthesized by using the motion compensation for the 3-D model-based image as the reference image is shown in Fig.10. In Fig.10, the synthesized image $f_{\theta=0^\circ}^{(S)}(r)$ contains few blocking artifacts. The specular reflection components can be represented precisely compared to the 3-D model-based images $f_{\theta=0^\circ}^{(D)}(r)$ shown in Fig.7.

Next, we obtained the relationship between the peak signal to noise ratio (PSNR), which was used for the objective metrics for the synthesized image, and the amount of coding bits required for the motion

compensation errors, the motion vectors and the adaptive block model. Fig.11 shows rate-distortion characteristics which were obtained from PSNR of the synthesized images and the coding rate for three methods: the method that estimates motion vectors of specular reflection components by VBS-MA, the method that estimates the movements by BMA and the model-aided coding which synthesizes a predictive image from a 3-D geometric model and encode the predict error. In Fig.11, there is 0.5 to 1.0 dB gain in PSNR using the BMA over the 3-D model. Moreover, VBS-MA gives better the PSNR performance than BMA by about 0.5 dB.

5. CONCLUSION

We have proposed a new method of synthesizing arbitrary viewpoint images using a 3-D geometric model and specular reflection tracing. The experiments showed that it was effective to use the motion compensation for the 3-D model-based synthesized image as the reference image in order to estimate the movements of specular reflection components. Also, we have improved the synthesized image by using VBS-MA, because the specular reflection components of objects appear in small regions and have high intensity compared with other reflection components. Experimental results showed that the coding performance of the proposed method can be improved by determining the adaptive block model for the motion vector estimation considering rate-distortion characteristics of the synthesized images.

6. REFERENCES

- [1] M. Magnor, and B. Girod, "Model-based Coding of Multi-Viewpoint Imagery", Proc. SPIE VCIP-2000, Perth, Australia, vol. 1, pp. 14-22, June 2000.
- [2] W. Matusik, C. Buehler, R. Rasker, S. J. Gortler, and L. McMillan, "Image-Based Visual Hulls", SIGGRAPH '00, July, pp.369-374, 2000.
- [3] K. Nishino, Y. Sato, and K. Ikeuchi "Eigen-Texture Method: Appearance Compression and Synthesis based on a 3-D Model", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No. 11, pp.1257-1265, Nov., 2001.
- [4] M. A. Magnor, P. Eisert, and B. Girod, "Model-Aided Coding of Multi-Viewpoint Image Data", ICIP2000, Vancouver, Canada, pp. 919-922, September 2000
- [5] Y. Suzuki, M. Shirokane, K. Fukuda, and A. Kawanaka, "Efficient Representation of Multiple Images Using 3-D Model", 2002 IEICE General Conference, D-11-147, 2002.



Fig.5 Camera image "Daikoku". Fig.6 Camera image $f_{\theta_1}(r)$ used for synthesizing.

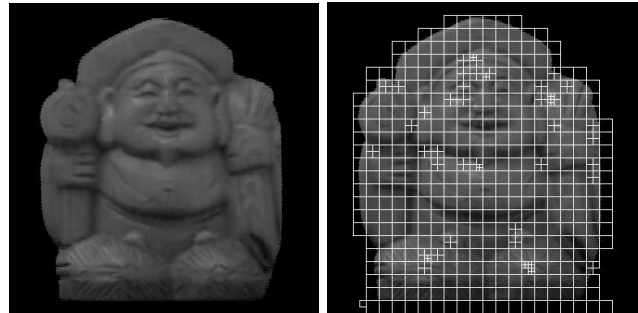


Fig.7 $f_{\theta}^{(D)}(r)$

Fig.8 Adaptive block model.

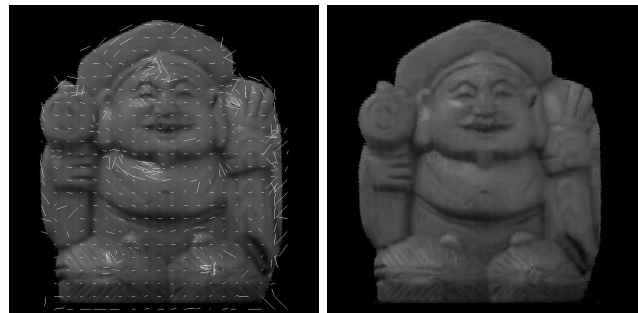


Fig.9 Estimated motion vectors. Fig.10 3-D model + VBS-MA

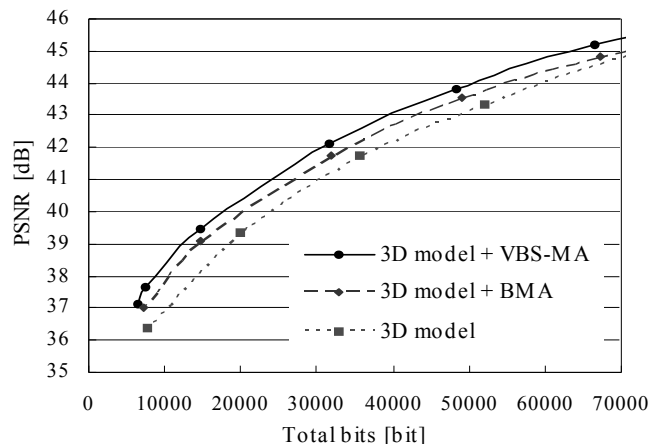


Fig.11 Rate-distortion characteristics of the proposed coding methods.