Journal of Zhejiang University SCIENCE A
ISSN 1009-3095 (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Variable refractive index in environment matte*

ZHAO Ming-tian†[1], XIAO Shuang-jiu[1], YANG Xu-bo[1], MA Li-zhuang[2]
([1]*Digital Art Laboratory, Shanghai Jiao Tong University, Shanghai 200240, China*)
([2]*Department of Computer Science, Shanghai Jiao Tong University, Shanghai 200240, China*)
†E-mail: zamba@sjtu.edu.cn

Received Apr. 5, 2006; revision accepted Apr. 19, 2006

Abstract: Environment matting and compositing is a technique to extract a foreground object, including color, opacity, reflective and refractive properties, from a real-world scene, and synthesize new images by placing it into new environments. The description of the captured object is named environment matte. Recent matting and compositing techniques can produce quite realistic images for objects with complex optical properties. This paper presents an approximate method to transform the matte by simulating variation of the foreground object's refractive index. Our algorithms can deal with achromatous-and-transparent objects and the experimental results are visually acceptable. Our idea and method can be applied to produce some special video effects, which could be very useful in film making, compared with the extreme difficulty of physically changing an object's refractive index.

Key words: Environment matting and compositing, Matte, Refractive index, Image-based rendering, Environment map
doi:10.1631/jzus.2006.A1160 **Document code:** A **CLC number:** TP391

INTRODUCTION

Digital matting and compositing is a popular technique in the film industry, with which we can separate a foreground object, including both its color and opacity, from a background image. The description of the separated foreground object is named digital matte. Smith and Blinn (1996) proposed the blue screen matting technique. They formally described the mathematical solution to the digital matting problem, when either some constraints on the foreground color or a pair of shots against completely different backgrounds are given.

As an extension to the conventional digital matting technique, environment matting can deal with not only color and opacity but also more complicated effects like reflection and refraction exhibited by optically active (Wexler *et al*., 2002) foreground objects. The foreground object with more information than digital matte is called environment matte. Environment matting and compositing technique enables production of realistic images in new environments for objects like glass and polished metal (Zongker *et al*., 1999; Chuang *et al*., 2000). Moreover, through well designed experiments, it is even possible to reproduce dynamic stream effects in front of new backdrops (Chuang *et al*., 2000).

However, most past researches focused on the extracting part of the matting problem, namely, to obtain more precise matte parameters from relatively small number of facile image samples. Although there were also researches on the compositing part, for example, anisotropic filters (Perona and Malik, 1990), few efforts have been made in the direct manipulation of matte data. In this paper, we present an approximate method to transform the matte by simulating the foreground objects' variation of refractive index. We were inspired by the widely used fade effects in films and computer games, which are based on alpha blending. By gradually changing the refractive index,

* Project supported by the National Natural Science Foundation of China (No. 60403044) and Microsoft Research Asia (PROJECT-2004-IMAGE-01)

it is possible to fade out a transparent foreground object by decreasing its refractive index to be the same as air, or fade it in with an opposite process. Fig.1 (see page 1166) illustrates the effects of such a transformation, with some tricks applied for invalid matte regions. Details of this example are explained in the RESULTS section.

In the rest of the paper, an overview of previous work is given first. After that, the composition of environment matte is introduced, followed by the detailed solutions to variable refractive index. Then experimental results are shown and potential future work is discussed.

RELATED WORK

The conventional matting and compositing technique has been used in the film and video industries under patents protection for decades. Vlahos used a single blue screen to get the color and opacity of foreground objects, which sufficiently differ from the blue background. Smith and Blinn (1996) proposed another method to extract the matte with two different backdrops, which has almost no special restrictions on the color of foreground objects.

Although digital matting has become very popular and successful in application, it cannot reproduce optically active effects like reflections or refractions. As a physical simulation of light transport related to foreground objects, Zongker *et al*.(1999) introduced environment matting, which can deal well with these relatively more complicated phenomena. In their method, every pixel of the foreground object is supposed to be mapped to a single axis-aligned area on the backdrop. A series of two-color backdrop patterns with different sized square waves are displayed on a CRT screen to help approximating the mapping area for each foreground pixel, which acts like a box filter to calculate its color by averaging the backdrop area.

Chuang *et al*.(2000) introduced extensions to environment matting towards higher accuracy and real-time capture respectively, to enhance its usability. The former extension aims to enhance the accuracy of area mapping, in which the axis-aligned rectangles are replaced with multiple oriented elliptical areas while the box filters are replaced with anisotropic Gaussian filters corresponding to sizes and orientations of those areas. Chuang *et al*.(2002) thought this extension is a physically better simulation, and they used Gaussian strips instead of square waves to achieve better approximation effects. In the latter extension, they simplified the environment to a single background. Then it is possible to extract the matte with a single shot against the specially designed backdrop, in which colors of every pair of pixels are different.

Wexler *et al*.(2002) used a different model to describe the contribution of the backdrop to foreground pixels. They assumed that for the color of a foreground pixel, every pixel on the backdrop has got a contribution. And they used a probabilistic method to calculate the contribution to extract the matte.

Peers and Dutré (2003) addressed another solution to the environment matting problem. They took the matte as a filter of background images. With captured input and output images, it is possible to calculate response parameters of the filter.

Zhu and Yang (2004) proposed a robust matte extraction method based on frequency. Unlike other methods taking images as isolated samples, they put them together to form a time sequence. By assigning different frequencies to the sequences corresponding to different backdrop pixels, and using Fourier analysis to calculate frequency responses of foreground pixels, contributions of all backdrop pixels to every foreground pixel can be easily calculated. Frequency analysis showed that this method is robust to noise. Moreover, compared to methods of Zongker *et al*.(1999) and Chuang *et al*.(2000) in spatial domain, the frequency-based method can find the exact mappings instead of applying non-linear approximations.

Both Peers and Dutré(2003)'s method and the frequency-based solution (Zhu and Yang, 2004) require huge number of image samples, which makes capturing difficult. Although the former use wavelet method to compress the samples, it is still not ideal.

There were also researches in constructing multi-directional environment mattes by Matusik *et al*.(2002a; 2002b) who used controllable lights and PDP monitors fixed on rotating platforms as well as multiple cameras to extract mattes of optically active objects in multiple directions. And with proper interpolation methods, they could demonstrate rotation of the objects in new environments.

ENVIRONMENT MATTE

The environment matting equation used by Zongker *et al*.(1999) indicates that the final color of a foreground pixel gets contributions from different elements:

$$C = F + (1-\alpha)B + \sum_{i=1}^{m} R_i \mathcal{M}(\boldsymbol{T}_i, \boldsymbol{A}_i).$$

F and α are the intrinsic color and opacity of the foreground pixel, and B is color of the background pixel behind it. If we ignore other contributions, we will have the traditional matting composition equation:

$$C=F+(1-\alpha)B.$$

Environment matting can deal with optically active objects by describing reflection and refraction with area mappings. In the environment matting equation, $\boldsymbol{A}_i$ in $\boldsymbol{T}_i$ are areas in textures that the foreground pixel is mapped to. Function $\mathcal{M}$ is used for compositing pixels inside one area, with R_i being called reflectance coefficient, describing the amount of $\boldsymbol{A}_i$'s contribution. The equation assumes that there is just one mapping area in a texture. So what environment matte covers include F, α, R_i and $\boldsymbol{A}_i$, if $\mathcal{M}$ and number of textures m are pre-defined.

In order to simplify analysis of the results of variable refractive index, we will just refer to the second extension to environment matting introduced by Chuang *et al*.(2000) in the following sections:

$$C=\rho \boldsymbol{T}(\boldsymbol{c}),$$

in which ρ, $\boldsymbol{T}$ and $\boldsymbol{c}$ represent reflectance coefficient, backdrop texture, and center of mapping area on the backdrop. In this simplification the foreground object should be achromatous-and-transparent. Then only two key features of the matte, ρ and $\boldsymbol{c}$, remain to be considered when refractive index changes. In fact, algorithms to transform unsimplified matte are very similar.

VARIABLE REFRACTIVE INDEX

In this section we look into what happen to the matte parameters when the refractive index is changed linearly. Our transformation algorithm is based on one basic assumption that the initial matte extracted in advance mainly represents refractive mappings rather than reflective or other kinds of mappings. We think this assumption is reasonable for common purely transparent materials without too high refractive indices, such as glass and water.

Move the mapping center *c*

Mapping center $\boldsymbol{c}$ is the center of the mapping area directly affected by the variation of refractive index. For clarity, we use 2D illustration to specify the 3D geometry of refraction, in which the original 2D vector $\boldsymbol{c}(x, y)$ on the backdrop becomes a 1D scalar $c(x)$ which can be described by another scalar s representing the mapping offset relative to the pixel's direct background point (In Fig.2, P_3 is P_1's mapping center and P_1' is P_1's direct background point).

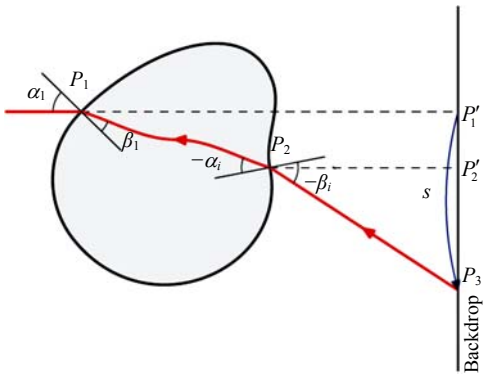

Fig.2 A simplified 2D illustration of the geometric process of refraction in environment matting. We assume angles of incidence and refraction have signs to indicate on which side of the normals they are located. *s* is named the mapping offset

From the geometric viewpoint, s is caused by multiple refractions when light passes between surface points P_1 and P_2 (including the two points themselves). By ignoring specific processes of internal refractions, we have

$$s = x(P_3) - x(P_1') = \boldsymbol{P_1'P_2'} + \boldsymbol{P_2'P_3}. \quad (1)$$

Given α_1, when we change the refractive index, β_1 will change according to refractive law, which causes P_2 to move and α_i and β_i to change synchronously. Because the precise shape of the foreground object is

unknown, we cannot accurately calculate the new value of s. However, if we simplify the object to known shapes like spheres or cubes, we can estimate an approximate $\tilde{s} = s$, which is proved to be acceptable for visual application. In our experiments, we assume all objects have a shape similar to spheres, because we think sphere is the most typical shape satisfying the laws of positive, continuous and self-closing surface normals. Solid objects and hollow objects are considered separately below.

For object simplified to solid sphere, we are able to calculate an approximation of s by taking Eq.(1) a little further. According to Fig.3, we have

$$\overline{P_1P_2} = 2R\cos\beta,$$
$$\overline{P_2P_2'} = d - R\cos(2\beta - \alpha),$$

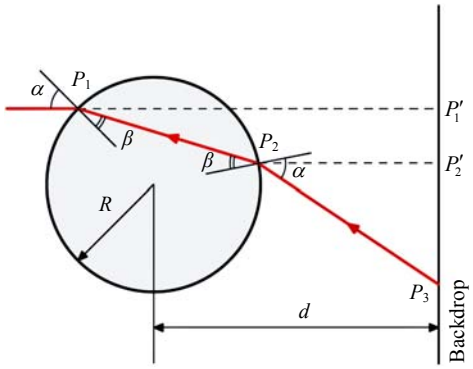

Fig.3 The 2D geometric illustration for solid objects simplified as solid spheres. Due to sphere symmetry we have a pair of αs and a pair of βs

so immediately we get

$$\tilde{s} = 2R\cos\beta\sin(\alpha-\beta)+[d-R\cos(2\beta-\alpha)]\tan2(\alpha-\beta), \quad (2)$$

in which we know from refractive law that

$$\frac{\sin\alpha}{\sin\beta} = n.$$

So if R and d are fixed, we have

$$\tilde{s} = \tilde{s}\,(\alpha, n). \quad (3)$$

Now the problem becomes calculation of Eq.(3) according to Eq.(2), by replacing β with expression $\beta(\alpha, n)$. In fact, Eq.(2) is so complicated in format that we can hardly get a neat expression of Eq.(3). As visual acceptance is preferred to mathematical accuracy, a proper numerical model can do as well as an analytic solution, which then led to our attempt to analyze the numerical distribution of function $\tilde{s}$ on variables α and n, with the results being proved to be quite useful. According to our matte pulling process, we fix d at $3R$ and the original refractive index n_0=1.5 (ordinary refractive index of glass), then we calculate values of function $\tilde{s}$ at different values of α and n. The distribution curves are shown in Fig.4.

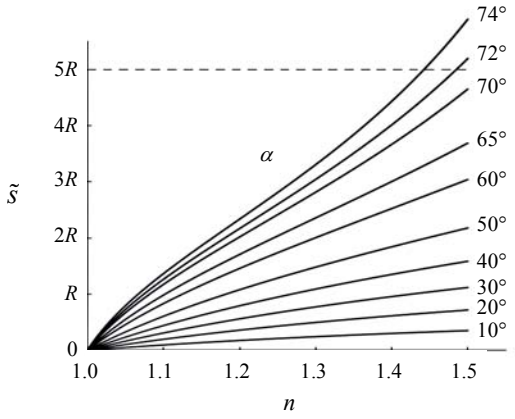

Fig.4 The numerical distribution of the approximate mapping offset $\tilde{s}$ on angle of incidence α and refractive index n for solid sphere. In this graph d is fixed at $3R$

In common cases, when $|s|>>R$, the mapping center will go out of the backdrop region and become invalid. According to our experimental conditions, we can make the assumption from experience that $|s|<5R$. Within this range, $\tilde{s}$ is quite close to a direct proportion function, and so can be approximated by

$$\tilde{s}(\alpha, n) = \frac{n}{n_0}\cdot\tilde{s}(\alpha, n_0), \quad (4)$$

without unacceptable error.

For hollow objects the case is similar except for complication of the angles:

$$\tilde{s} = 2R\cos\gamma\sin(\alpha_1-\gamma)+[d-R\cos(2\gamma-\alpha_1)]\tan2(\alpha_1-\gamma), \quad (5)$$

in which

$$\gamma=\beta_1+\beta_2-\alpha_2,$$

and we know from refractive law that

$$\frac{\sin\alpha_1}{\sin\beta_1}=\frac{\sin\beta_2}{\sin\alpha_2}=n,$$

and from Fig.5 according to sine law that

$$\frac{\sin\beta_1}{r}=\frac{\sin(\pi-\alpha_2)}{R}.$$

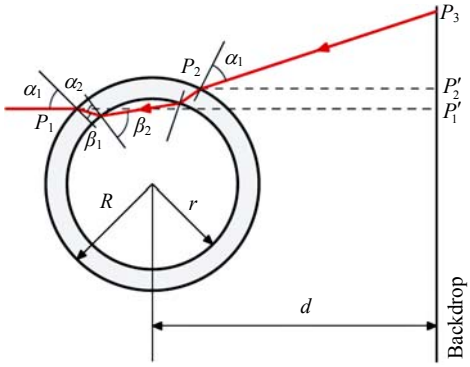

Fig.5 The 2D geometric illustration for hollow objects simplified as hollow spheres. Also, due to sphere symmetry we have all angles of incidence and refraction in pairs

Similarly, we apply numerical analysis on Eq.(5) to find the distribution of $\tilde{s}$ on variables α and n. R is fixed at 1.1r in Fig.6, and for the hollow case a logarithmic function can approximate $\tilde{s}$ well:

$$\tilde{s}(\alpha, n)=\log_{b(\alpha)}n.$$

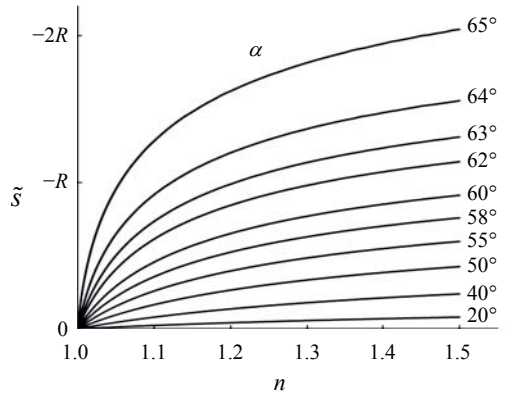

Fig.6 The numerical distribution of the approximate mapping offset $\tilde{s}$ on angle of incidence α and refractive index n for hollow sphere. In this graph d is fixed at $3R$

In fact it is unnecessary to calculate the bases of the logarithmic functions, $b(\alpha)$, because only relative logarithmic values are needed:

$$\tilde{s}(\alpha,n)=\frac{\log_{b(\alpha)}n}{\log_{b(\alpha)}n_0}\cdot\tilde{s}(\alpha,n_0)=(\log_{n_0}n)\tilde{s}(\alpha,n_0). \quad (6)$$

After every new mapping center $\boldsymbol{c}(x, y)$ is calculated according to $\tilde{s}$, we use the neighboring difference method (Chuang *et al*., 2000) to calculate the size of every mapping area, $\boldsymbol{w}(x, y)$, for proper antialiasing. Thus the geometric transformation of the mapping is finished.

Correction of ρ

ρ depends on both material and geometry of the foreground object, as well as path of the light. Theoretically we have $0<\rho<1$ because the light attenuates in its path due to reflection, scattering and absorption. Actually, to accurately measure attenuation of light physically is neither feasible nor necessary for production of visually pleasant images. Going a little further from the simplification made by Chuang *et al*.(2000) for real-time capture, we assume that the attenuation caused by factors other than reflection is a constant ρ_c, while ρ_v caused by reflection can be estimated with knowledge from optics (Born and Wolf, 1999):

$$\rho=\rho_c\cdot\rho_v(n,\theta)=\rho_c\cdot\frac{\sin2\theta\sin2\varphi}{2\sin^2(\theta+\varphi)}\left[1+\frac{1}{\cos^2(\theta-\varphi)}\right], \quad (7)$$

in which θ and φ are the angles of incidence and refraction respectively, and ρ_c can be calculated with the initial value of ρ. In Eq.(7) we just deal with the first reflection by ignoring other reflections that occurred inside the foreground object. It is because we think internally reflected light, if not scattered or absorbed, has other chances to go outside, which may contribute to other pixels. Moreover, in experiments we became aware of the obvious effects of the first reflection.

In fact, in Eq.(7), the value of function ρ decreases while θ increases. It means that the larger is the angle of incidence, the weaker is the influence of

backdrop color. Coincidentally, in Fig.3, function value $\tilde{s}$ in Eq.(2) exceeds $5R$ only when $\alpha>70°$ when ρ is quite small. Another part of the assumption is that $|s|<5R$ because foreground pixels with long mapping distances will not be influenced by the backdrop.

Invalid matte regions

In general cases, there are almost always some invalid matte regions, which are supposed to be mapped outside the backdrop area. And when the refractive index goes below a threshold, the mapping offset s decreases until sizes of these regions become zero. In fact we cannot model this erosion process if geometric parameters of the foreground object are not given, because there exist multiple refractions at unknown inner surfaces. Variation of refractive index may affect not only the angles of refractions, but also positions and times. So size of an invalid region may change monotonously, or may also undulate. The threshold of refractive index to make the whole region valid is also unknown.

Due to such uncertainty, we choose to apply monotonous erosions based on boundary conditions of these regions. If valid mapping offsets near the boundary change rapidly, the erosion is rapid, and vice versa.

RESULTS

We pulled the simplified environment matte with the method towards real-time capture introduced by Chuang *et al*.(2000), using a 17" LCD monitor to display backdrops and a Sony HC-90E DV to capture images. In our experiments, multiple identical shots were made to reduce noise. As introduced by Chuang *et al*.(2000), anisotropic filter (Perona and Malik, 1990) is applied on mapping center $\boldsymbol{c}$ to obtain neat images. On a Pentium IV 2.8 GHz computer with 1 GB memory, the matte pulling process takes less than 10 s for 512×512 input images and the synthesizing process takes less than one second per frame.

Fig.1 shows a glass cup with its refractive index changed gradually. This cup has sixteen smooth ridges around its external surface. Naturally, we should apply the numerical model for hollow objects. We can see that as the refractive index decreases to 1.0, the appearance of the characters through the cup becomes similar to their original shapes on the background. When n=1.05, the first three letters of the word "Laboratory" at the center part of the cup are already distinguishable. Distortion of the word "Digital" near the top of the cup in the four images can also show the transformation clearly.

Fig.7 is another example of hollow objects. It is a glass bowl like a hemisphere with quite smooth surface. We use the background with black and white grids to illustrate the transformation. It is obvious that the shape distortion becomes weak while the refractive index decreases towards 1.0. Besides, due to the variation of ρ, noticeable color in the fourth image also fades in the former three according to their corresponding refractive indices. This effect is achieved by maintaining separated ρs for different color channels.

Fig.8 is designed to give a careful comparison of the two different numerical models for solid and hollow objects. The first row is synthesized with the solid model, the second row with the hollow model. The differences in distortion are distinguishable, for example, at the lower part of the candlestick when n=1.1. But in visual satisfaction they are both acceptable as an ordinary observer can hardly judge the correctness. Another essential factor is that both the two numerical models ensure the visual continuity to acuum at n=1.0.

Fig.9 is a further example to demonstrate the transformation. Images are also synthesized with a real photo, like in Fig.8. The foreground object is a glass goblet with a knob in its stem. The goblet is half filled with pure water. We apply the solid model because water's refractive index is quite close to that of glass. The visual effect of the cat's whiskers through the goblet with water is quite pleasing. In fact, at the base of the goblet, reflection effects are more than those of refraction, so that our current algorithm must be extended to cover such case in theory. However, in this group of images, colors are not flamboyant, and slight error in brightness is not easily noticeable, so the visual effect is still acceptable.

DISCUSSION AND FUTURE WORK

In this paper, we introduced a novel method to transform the simplified environment matte by simulating the variation of the refractive index. Our

Fig.1 Four images synthesized with a set of environment mattes different in refractive indices. The former three are done with transformed mattes according to our algorithm (from left: n=1.05, 1.1 and 1.3, same for Figs.7, 8 and 9), the last one with the original matte (We assume this glass cup's original refractive index n=1.5)

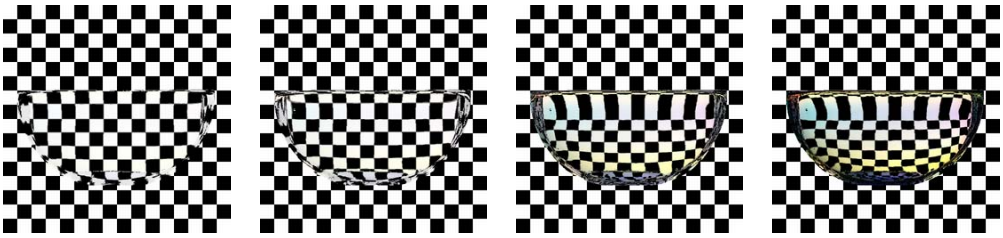

Fig.7 Four images of a glass bowl as the foreground object, only different in refractive indices. The three images on the left are synthesized with transformed mattes. The fourth one is synthesized with the original matte, in which the refractive index n=1.5

Fig.8 A comparison of the two numerical models we applied for solid and hollow objects respectively. The foreground object is a glass candle stick whose upper part is hollow and lower part is solid. This candle stick is formed simply by multiple cylinders. Images in the first row are synthesized supposing the foreground object is solid, while the second row is synthesized using the hollow model

Fig.9 Four images of a glass goblet with water in it as the foreground object. It has a knob in the stem. The refractive index is transformed using the solid model, because water's refractive index is quite close to that of glass. The three images on the left are synthesized with transformed mattes. The fourth one is synthesized with the original matte with refractive index n=1.5

work focuses on the environment matte simplified for real-time capture (Chuang *et al*., 2000). And in order to achieve the transformation, we also simplified the foreground objects' shapes, and several assumptions on illuminative behaviors. We have demonstrated that with some approximations, it is possible to make the foreground object's refractive index increase or decrease, which is achieved with further assumptions of the environment, compared to the current geometric and optic environment matting and compositing model.

This research leads to useful applications in film and video making, as well as interesting future work. First, as we introduced before, the refractive index fade effects are much more impressive than old effects realized with techniques like alpha blending. Second, it is possible to realize indirect environment matting and compositing for objects difficult or even impossible to acquire. For example, it is almost impossible to obtain a bottle made of diamond, but a glass bottle is available everywhere. We can extract the latter's environment matte and process it with our methods to get an approximate matte for the former. To put this idea into application, more complicated phenomena such as double refraction should be taken into consideration. Third, it is the beginning of a new series of techniques to directly manipulate the environment matte. It is imaginable to apply advance transformations for further simplifications involving approximate 3D models. For example, we can make deformable environment mattes.

For matte extraction and presentation, there is also work to do to enhance transformation. As we mentioned before, our method just covers the simplified environment matting for real-time capture. If the algorithms are generalized for higher accuracy model (Chuang *et al*., 2000), they will be much more useful. Besides, some intrinsic features of environment matting technique make the transformation difficult, such as taking reflection and refraction as a whole. Dealing with the invalid matte regions is also an intractable problem. All these are left to be improved for better transformation effects.

References

Born, M., Wolf, E., 1999. Principles of Optics, 7th (expended) Edition. Cambridge University Press.

Chuang, Y.Y., 2004. New Models and Methods for Matting and Compositing. Ph.D Thesis, Department of Computer Science and Engineering, University of Washington.

Chuang, Y.Y., Zongker, D.E., Hindorff, J., Curless, B., Salesin, D.H., Szeliski, R., 2000. Environment Matting Extensions: Towards Higher Accuracy and Real-time Capture. Proceedings of ACM SIGGRAPH'00, p.121-130.

Ding, Y.F., 2005. A Research on Environment Matting. MS Thesis, Department of Computer Science and Engineering, Shanghai Jiao Tong University.

Matusik, W., Pfister, H., Ngan, A., Beardsley, P., Ziegler, R., McMillan, L., 2002a. Image-based 3D Photography Using Opacity Hulls. Proceedings of ACM SIGGRAPH'02, p.427-437.

Matusik, W., Pfister, H., Ziegler, R., Ngan, A., McMillan, L., 2002b. Acquisition and Rendering of Transparent and Refractive Objects. Proceedings of the 13th Eurographics Workshop on Rendering, p.267-278.

Peers, P., Dutré, P., 2003. Wavelet Environment Matting. Proceedings of the 14th Eurographics Workshop on Rendering, p.157-166.

Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, **12**(7):629-639. [doi:10.1109/34.56205]

Smith, A.R., Blinn, J.F., 1996. Blue Screen Matting. Proceedings of ACM SIGGRAPH'96, p.259-268.

Wexler, Y., Fitzgibbon, A.W., Zisserman, A., 2002. Image-based Environment Matting. Proceedings of the 13th Eurographics Workshop on Rendering, p.279-290.

Zhu, J., Yang, Y.H., 2004. Frequency-based Environment Matting. Proceedings of the Computer Graphics and Applications, 12th Pacific Conference on (PG'04), p.402-410.

Zongker, D.E., Werner, D.M., Curless, B., Salesin, D.H., 1999. Environment Matting and Compositing. Proceedings of ACM SIGGRAPH'99, p.205-214.