REMOVAL OF SHADOWS AND SHADINGS FROM TEXTURE IMAGES FOR ARCHIVING HISTORICAL RUINS

Ryo FURUKAWA,

Hiroshima City University, Na ryo-f@cs.hiroshima-cu.ac.jp

Yutaka KIUCHI, Nara Institute of Science and Technology yutaka-k@is.aist-nara.ac.jp Rieko KADOBAYASHI Communication Research Laboratory rieko@crl.go.jp

KEY WORDS: Shadow Removal, Shade Removal, Reflection Model

ABSTRACT

We are digitizing a whole terrain and ruined buildings of a medieval harbor city on the island called Gemiler Island, as 'Digital Gemiler Island project.' To make textured 3D models of the ancient buildings, we have measured shapes of some of the buildings on the site with a laser range scanner and taken high-resolution images for textures with digital cameras. Since we've taken the photos in the daytime, the texture images are affected with shadows and shadings by the sunlight. To render the 3D model of the buildings on arbitrary lighting conditions, these effects of shadows and shadings must be removed from the images. Although there are several techniques for removing shadows and shadings from images, most of them need several images with different lighting conditions is very time consuming. In Digital Gemiler Island project, we can use the globally registered 3D model of the buildings, which is measured with 3D laser scanner and GPS, the direction of light (the sun), and camera pose information for each images. In addition, variations of the surface materials of the ruined buildings are very limited. In this paper, effects of shadows and shadings are removed taking advantages of these characteristics. We assume a parametric reflection model and estimate the parameters of the model fitting the model to the images, by which separating effects of shadows/shadings from reflectance. The method needs only single image for each scene.

1 INTRODUCTION

Archiving huge historical buildings or ruins as 3D models is attracting considerable attention of archaeologists or computer engineers. Potential applications vary from preservation to online museums accessible for everyone.

We have started 'Digital Gemiler Island project' which aims to make textured models of a whole terrain and ruined buildings of a medieval harbor city on the island called Gemiler Island.

For the project, shapes of some of the buildings on the site have been measured with a laser range scanner. Also, as a raw data for the textures, we've taken high-resolution images of the buildings with digital cameras. Since we've taken the photos in the daytime, the texture images are affected with shadows and shadings by the sunlight. To render the 3D model of the buildings on arbitrary lighting conditions, these effects of shadows and shadings must be removed from the images, leaving only material colors for the surface.

Separating effects of shadows/shadings from material colors (i.e. reflectance information) has been a major interest in the community of computer vision (Barrow and Tanenbaum, 1978). Because such a problem is ill-posed, some additional constraints should be assumed to solve it. Shinha et al. (Shinha and Adelson, 1993) assumed scenes made of polyhedra without occlusions and cast shadows. Land et al. (Land and McCann, 1971) proposed retinex theory, in which spatial changes of shading are assumed to be much slower than those of reflectance. The restrictions of those methods are too strong and they are inappropriate for removing effects of strong cast shadows.

Using several images with the same camera positions and

different lighting conditions, we can separate reflections and shadings/shadows without assuming strong constraints about the scene. There are some methods with this approach (Szelksi et al., 2000),(Weiss, 2001). Although these approaches are fairly general, taking several photos with different light conditions for each scene is sometimes timeconsuming. If the light source was the sun (like in our project), it would take prohibitively long. A technique for removing shades/shadows with single lighting condition is desired.

Tappen et al. (Tappen et al., 2003) classified edges into those caused by reflection changes and those caused by shadings. The classification is based on clues of color changes and a method of machine learning. Although very sophisticated, their method still relies on uniformity of colors in each region, which may be unreliable for textured surfaces (like in our project).

Finlayson et al. (Finlayson et al., 2002) assumed the light source described by "black box model" which is widely applicable for outside scenes. Using the restriction in the spectral domain, they extracted the edges which are caused by cast shadows. Their method seems promising. However, they did not show how to deal with effects of shadings.

Our approach is very simple. In Digital Gemiler Island project, there are several conditions which we can exploit for the shadow/shade removal. The first of such conditions is that we have the globally registered 3D model of the buildings, which is measured with 3D laser scanner and GPS. The second is that the images for the textures are registered to the 3D model. The third is that the times are recorded when those images are taken. The fourth is that the variations of the surface materials of the ruined buildings are very limited. Because of those conditions, we can acquire 3D model mapped to the objective images and the direction of the light source (the sun). So, the only thing we have to do is assuming a shading model and fit the observed pixel values to the model.

In this paper, effects of shadows and shadings are removed by taking advantages of the conditions. We assume a parametric reflection model for surface of the 3D model, and estimate its parameters. Then we estimate and remove effects of shadows and shading for texture images using the obtained reflection model. Because we can use registered geometorical models, the method only needs a single image for each scene.

2 ASSUMPTIONS

The shadow/shading removal method described in this paper assumes the following conditions, which come from demands of our project.

- The 3D models for the scene are given. In our project, they are measured by using a 3D range scanner.
- The 3D models are globally registered. This means that the latitude and longitude of the site is known, and the pose of the 3D models relative to the Earth is known. In our project, these data are acquired using GPS.
- The objective photo images are registered to the 3D models in advance. In our project, the registration is done by specifying correspondences between 3D models and the pictures. We estimated the projection matrix and improve the solution by non-linear optimization (Faugeras, 1993).
- It is recorded when each the objective photo images is taken.
- There are textures on the surface of the scene. Here, we assume that the distributions of the reflectance are independent of the locations on the surface. This assumption reflects the fact that the walls of the buildings on Gemiler Island are made of stones with similar materials. The textures may have strong edges in the images, but have similar appearances for all the surfaces.

3 ACQUIRING DIRECTION OF LIGHT SOURCE AND SHADOW REGIONS

Since it is recorded when each photo image is taken and the global location of the site is given, the direction of light (the sun) can be calculated. Let us describe the variables used in this calculation. θ_{long} is the longitude of the global location of the site, whose value is the signed angle measured from Greenwich meridian toward the west. θ_{lat} is the latitude with positive direction toward the north from the equator as 0 degrees. d is the number of days passed from the last summer solstice and D is the number of days in a year. t is the time when the objective image is taken expressed as Greenwich time converted to the number of seconds passed from the midnight. T is the length of a day expressed as the number of seconds. θ_{axis} is the angle between the Earth's rotation axis and the normal direction of the plane of the ecliptic.

Now, suppose a local coordinate system called O_l defined at the site as follows: x-axis is directed toward the east, y-axis directed toward the north, and z-axis directed vertically upward. The xy-plane should be horizontal. Let a vector (s_x, s_y, s_z) be the direction toward the sun expressed in the local coordinates O_l . Then, s_x, s_y, s_z can be approximated as the following.

$$\begin{aligned}
\theta_d &= 2\pi d/D_y, \\
\theta_t &= 2\pi t/T - \theta_{long}, \\
A &= \sqrt{\cos^2 \theta_d \cos^2(\theta_{axis}) + \sin^2 \theta_d}, \\
B &= |\cos^2 \theta_d \sin^2 \theta_{axis}| \\
C &= -A \cos \theta_t \\
s_x &= A \sin \theta_t \\
s_y &= YB \cos \theta_{lat} - C \sin \theta_{lat} \\
s_z &= B \sin \theta_{lat} + C \cos \theta_{lat}
\end{aligned}$$
(1)

Since the pose of the 3D model relative to the Earth is known, the model can be expressed by the coordinate O_l . Since the 3D model and the direction of the sun (s_x, s_y, s_z) are expressed in the same coordinates, we can estimate surfaces where shadows are cast. Note that we can not detect shadows which are caused by objects that do not exist in the 3D model.

The surfaces where the shadows are cast are mapped to the processed image based on the registration data, which determines shadow pixel set R_s , where R_s is a set of pixels of the processed image on which one or more shadows are cast.

4 REFLECTION MODEL

To estimate shading effects for the scene on the image, reflection models are used. One of the most frequently used formulations to describe reflection models is the expression of BRDF (bidirectional reflectance distribution function) which is widely used in CG community. This function takes arguments of the direction of the incident ray (w_i) and the direction of the reflected ray (w_r) . It maps the arguments to the ratio of the radiance (the radiant flux of the reflected ray normalized by solid angle and orthogonal area) to the irradiance (light energy of the incident ray coming from the direction w_i).

Several reflection models are proposed in the form of BRDF. One of the most frequently used models is Oren-Nayar model (Oren and Nayar, 1993). This model expresses diffuse reflections of 'rough' surface. Roughness of surface is modeled by countless V-grooves on the surface, and the roughness is parameterized by a distribution of angles of slopes of the grooves. The model also expresses effects of interreflections caused by rays reflected up to twice.

We used this model on the bases of the following facts. The first reason is that the surfaces of the walls of the objective buildings are mostly stones with rough surface. So, the material itself can be modeled with Oren-Narar model. The second reason is that resolution of the geometrical models acquired by the 3D scanner is larger than the sizes of the bumps of the surface. Since BRDF formulation needs geometrical parameters (normal vectors), we have to sample the intensities of the surface with the same resolution of the geometrical model to fit them to BRDF. So, the bumps of the surface whose sizes are smaller than the geometrical resolution can be considered as "roughness" of the surface, which can be approximated by Oren-Nayar model.

Assuming Oren-Narar model, the value of BRDF f on a surface point is expressed as the following.

$$\begin{aligned} \alpha &= \max(\cos^{-1}(l \cdot n), \cos^{-1}(v \cdot n)) \\ \beta &= \min(\cos^{-1}(l \cdot n), \cos^{-1}(v \cdot n)) \\ \nu_{\perp} &= v - n(n \cdot v) \\ l_{\perp} &= l - n(n \cdot l) \\ \Delta_{\perp} &= \begin{cases} 1 & \text{if } \|v_{\perp}\| \|l_{\perp}\| = 0 \\ \frac{(v_{\perp} \cdot l_{\perp})}{\|v_{\perp}\| \|l_{\perp}\|} & \text{otherwise} \end{cases} \\ C_{1} &= 1 - \frac{\sigma^{2}}{2(\sigma^{2} + 0.33)} \\ C_{2} &= \begin{cases} \frac{0.45\sigma^{2}}{\sigma^{2} + 0.09} \sin \alpha & \text{if } \Delta_{\perp} \ge 0 \\ \frac{0.45\sigma^{2}}{\sigma^{2} + 0.09} \left(\sin \alpha - \left(\frac{2\beta}{\pi}\right)^{3}\right) & \text{otherwise} \end{cases} \\ C_{3} &= \frac{0.125\sigma^{2}}{\sigma^{2} + 0.09} \left(\frac{4\alpha\beta}{\pi^{2}}\right)^{2} \\ f_{1} &= \frac{\rho}{\pi} \left[C_{1} + C_{2}\Delta_{\perp} \tan \beta \right. \\ &\left. + \left\{ (1 - \left| \Delta_{\perp}C_{3} \tan \left(\frac{\alpha + \beta}{2}\right) \right| \right\} \right] \\ f_{2} &= \frac{0.17\rho^{2}}{\pi} \left[\frac{\sigma^{2}}{\sigma^{2} + 0.13} \left\{ 1 - \Delta_{\perp} \left(\frac{2\beta}{\pi}\right) \right\} \right] \\ f(\sigma, \rho) &= f_{1}(\sigma, \rho) + f_{2}(\sigma, \rho) \end{aligned}$$

where l is the direction of the incident light expressed as a direction vector from the surface point toward the light source, v is the direction of the view expressed as a direction vector from the surface point toward the camera, nis the normal vector of the geometrical model at the surface point, σ is the standard deviation of the distribution of the angles of slopes of V-grooves (which are assumed to have normal distribution) representing roughness of the surface, ρ is the reflection rate of the surface of V-grooves. f_1 expresses effects of direct reflection from surfaces of V-grooves, and f_2 expresses effects of interreflection.

Let (M^r, M^g, M^b) be the color of the material, (L^r_d, L^g_d, L^b_d) be the intensity of the direct light source (the sun), each

component of which represents light frequency of R,G,B, respectively. (L_e^r, L_e^g, L_e^b) be the intensity of the environment light (such as light from the sky or indirect lights coming from everywhere), and (I^r, I^g, I^b) be the apparent intensity of the scene sensed by the camera (i.e. pixel value). Assuming the reflection of environment lights is proportional to the material colors with ratio of C_e , the pixel value I^c , $(c \in \{r, g, b\})$ can be expressed as

$$I^{c} = M^{c}L_{d}^{c}R(l, v, n, \sigma, \rho)(l \cdot n) + M^{c}C_{e}L_{e}^{r}$$

$$(c \in \{r, g, b\}).$$
(3)

Now, we rewrite the formula using a sampling point p as a variable. p is sampled from the objective image with an appropriate resolution, which is decided considering the resolution of the geometrical models. Then, I^c , v, n, M^r , M^g and M^b are functions of p, expressed as $I^c(p)$, v(p), n(p), $M^r(p)$, $M^g(p)$ and $M^b(p)$, respectively. Since we assume that the materials of the objective buildings are uniform, rough stones, σ and ρ are constant. Then the formula is

$$I^{c}(p) = M^{c}(p)L_{d}^{c}\{K^{\#}(p,\sigma,\rho) + K_{e}^{c}\} \\ B_{d}^{c}(p)\{K^{\#}(p,\sigma,\rho) + K_{e}^{c}\}$$
(4)

where

$$K^{\#}(p,\sigma,\rho) = \begin{cases} 0 & \text{if } p \in R_s \lor (l \cdot n(p)) < 0 \\ R(l,v(p),n(p),\sigma,\rho)(l \cdot n(p)) & \text{.} \end{cases}$$
(5) othewise

Note that R_s means the set of surface with cast shadows. Since l is fixed on the scene and v(p) and n(p) are functions of p, these variables are removed from the arguments of $K^{\#}$.

In the formula 4, specular reflection is neglected. This is justified by the fact that the material we are dealing with is rough stone and there is only small component of specular in its reflections.

By taking log of the formula 4, and defining $\ln \tilde{B}_d^c$ to be the median of the distribution of $\ln B_d^c(p)$, and $\epsilon^c(p) = \ln B_d^c(p) - \ln \tilde{B}_d^c$, then

$$\ln I^{c}(p) = \ln B^{c}_{d}(p) + \ln\{K^{\#}(p,\sigma,\rho) + K^{c}_{e}\} = \ln \tilde{B^{c}_{d}} + \epsilon^{c}(p) + \ln\{K^{\#}(p,\sigma,\rho) + K^{c}_{e}\}.$$
(6)

 $\ln \tilde{B}_d^c + \epsilon^c(p)$ depends on the material of the surface, and $\ln \{K^{\#}(p,\sigma,\rho) + K_e^c\}$ depends on the geometry. We decide parameters σ , ρ , $B_d^c(p)$ and K_e^c so that the samples fit to the model 6.

5 SAMPLING AND FITTING

Acquisition of samples is done as the following. The pixel intensities $(I^r(p), I^g(p), I^b(p))$ can be obtained from pixel

p of the objective image. n(p) and v(p) can be acquired by mapping the geometrical data into the pixels of the objective image, using the camera parameters of registration data for the image. l is acquired from the estimated direction of the sun. n(p), v(p) and l are expressed in the local coordinates O_l ..

Although the samples $\ln I^c(p)$ and $K^{\#}(p, \sigma, \rho)$ can be calculated pixel by pixel of the objective image, we decide to resample them in lower resolution. This is because the resolution of the geometrical model is much lower than that of the images. First, the objective image is divided into square regions (subregions), whose sizes are $N \times N$. Then, $\ln I^c(p)$ and $K^{\#}(p, \sigma, \rho)$ are calculated for each pixels and the median values for each of the regions are taken, representing each subregions.

The regions where $K^{\#}(p, \sigma, \rho) = 0$ is those where the light of the sun does not reach. Failing in fitting the samples at these regions often results in sharp differences at the edge of cast shadows. So, the samples where $K^{\#}(p, \sigma, \rho) = 0$ are very important even if the number is relatively small. Therefore fitting of the samples are calculated separately from the other samples. The calculation is

$$\ln \hat{I}_{e}^{c} = \text{med}\{\ln I^{c}(p) \mid K^{\#}(p,\sigma,\rho) = 0\}$$
(7)

where \hat{I}_e^c is the estimate of $B_d^c K_e^c$ for $c \in \{r, g, b\}$. med means taking a median of a set.

The samples $I^c(p)$ normally include many outliers, and n(p) acquired from the geometrical models also include errors. So, we use LMedS estimation to estimate σ , ρ and \tilde{B}^c_d , which is more robust than the least square methods. LMedS estimation is done by minimizing

$$F(\sigma, \rho, \tilde{B}_{d}^{r}, \tilde{B}_{d}^{g}, \tilde{B}_{d}^{b}) = \operatorname{med}\{(\ln I^{c}(p) - \ln \hat{I}^{c}(p))^{2} \mid c \in \{r, g, b\}, \forall p\} (8)$$

where

$$\ln \hat{I}^{c}(p) = \ln \tilde{B}^{c}_{d} + \ln\{K^{\#}(p,\sigma,\rho) + \hat{I}^{c}_{e}/\tilde{B}^{c}_{d}\}.$$
 (9)

The minimization requires nonlinear optimization. Fixing σ and ρ the value of $\hat{I}^r(p)$ for $c \in r, g, b$ depends only on \tilde{B}^r_d . This means that optimizing the formula with B^r_d , can be done independently of B^g_d and B^b_d , requiring simple 1-dimensional optimization. For each of $\hat{I}^g(p)$ and $\hat{I}^b(p)$, the situation is the same. Expressing the optimal solutions of $\tilde{B}^r_d, \tilde{B}^g_d, \tilde{B}^b_d$ for fixed σ and ρ as $B^{r*}_d(\sigma, \rho)$, $B^{g*}_d(\sigma, \rho)$ and $B^{b*}_d(\sigma, \rho)$, the overall optimization is reduced to minimizing

$$F^*(\sigma,\rho) = F(\sigma,\rho, B_d^{r*}(\sigma,\rho), B_d^{g*}(\sigma,\rho), B_d^{b*}(\sigma,\rho)).$$
(10)

We minimized the 2 variable function $F^*(\sigma, \rho)$ using simplex descending method (Press et al., 1994). Let us define $\hat{\sigma}$ and $\hat{\rho}$ as the solutions of the optimization with σ and ρ , and $B^c_d = B^{c*}_d(\hat{\sigma}, \hat{\rho}), \ (c \in \{r, g, b\})$.

6 REMOVING SHADINGS AND SHADOWS

Using the estimated parameters of the model $(\hat{\sigma}, \hat{\rho}, B_d^c)$, and \hat{I}_e^c , $(c \in \{r, g, b\})$, we remove the effects of shadings and shadows from the objective image. Although we want the reflection rates for each color M^c , $M^c(p)$ and L_d^c are inseparable. So, $B_d^c(p) = M^c(p)L_d^c$ are calculated, instead. Form the form 6, $B_d^c(p)$ is calculated by

$$B_d^c(p) = I^c(p) / \{ K^{\#}(p,\sigma,\rho) + \hat{I}_e^c / \tilde{B}_d^c \}.$$
(11)

for each pixel. In this calculation, $I^c(p)$ and $K^{\#}(p,\sigma,\rho)$ are also sampled for each pixel.

7 EXPERIMENTAL RESULTS

We applied the shade/shadow removal method described in this paper on some images. Figure 1 shows the first example. Figure 1 (a) is the objective image, (b) is the image of detected shadow regions, (c) is the shadings with optimized parameters ($\sigma = 46^{\circ}$ and $\rho = 0.55$), and (d) is the result of shadow/shade removal. Examining (b), we could see that the shadow region on the right side was detected correctly, although there were some errors in boundaries of the estimated shadows at the top of the shadow regions. In the image shown in (d) the intensity of the shadow regions were compensated correctly, except that there were some regions where the pixel values are saturated.

Figure 2 shows the results of the second example. The original image is shown in figure 2(a), (b) is the shadings with optimized parameters ($\sigma = 46^{\circ}$ and $\rho = 0.55$), and (c) is the result. The result is magnified in figure 3(a). Examining the magnified result, we could see some errors of boundaries of detected shadow regions. We think that they are caused by errors in global registration of 3D model. Also, the removal of shadow was not good enough, as the brightness of the shadow regions in the result is apparantly darker. We think this is because, for this example, the shadow region was very small and there were not enough samples to estimate correct intensities of shadow regions.

To examine the effects of using Oren-Nayar model, we also tested optimization keeping $\sigma = 0^{\circ}$, which is Lambertian model. Figure 4(a) shows the shading, and (b) shows the result. The magnified result is shown in figure 3(b). Comparing figure 3(a) and (b), we could see that the result of shading removal using Oren-Nayar model was better than that using Lambertian model. Note that, in 3 (b), the intensities at the side of the window are too large.

8 CONCLUSION

In this paper, a method to construct shadow/shading-free texture images for 3D digitized models is described. To separate the effects of shadings using only one image, we fit the pixel values of the image to Oren-Nayar's reflection model. The direction of irradiance is obtained by calculating the direction of the sun from the recorded time when The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 5/W12 and Spatial Information Sciences, Vol. XXIV, Pa

the image is taken. The direction of radiance and the normal vectors of the surface are obtained from globally registered geometrical model and the camera pose. Using these data, unknown parameters of Oren-Nayar's model are estimated. Then the reflectance at each pixel is acquired. Experiments show promising results for this technique, although there are some errors in estimated shadow boundaries and shadow intensities.

For the future work, improving accuracies for shadow detection is important. To do so, it may be a effective method to correct shadow bondaries by searching shadow edges around estimated shadow boundaries using some edge detectors.

REFERENCES

Barrow, H. G. and Tanenbaum, J. M., 1978. Recovering intrinsic scene characteristics from images. In: A.Hanson and E.Riseman (eds), Computer Vision Systems, Academic Press.

Faugeras, O. D., 1993. Three-Dimensional Computer Vision: A Geometrical Viewpoint. MIT Press.

Finlayson, G. D., Hordley, S. D. and Drew, M. S., 2002. Removing shadows from images. In: Computer Vision-ECCV2002, pp. pp. 823–836.

Land, E. H. and McCann, J. J., 1971. Lightness and retinex theory. Journal of the Optical Society of America 61, pp. 1–11.

Oren, M. and Nayar, S. K., 1993. Generalization of the lambertian model. In: Proceedings Image Understanding Workshop 1993, Morgan Kaufmann Publishers, pp. 1037–1048.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannnery, B. P., 1994. Numerical Recipe in C: the art of scientific computing, 2nd edition. Cambridge Univ. Press.

Shinha, P. and Adelson, E., 1993. Recovering reflectance and illumination in a world of painted polyhedra. In: Proc. of Forth Inter. Conf. on Computer Vision.

Szelksi, R., Avidan, S. and Anandan, P., 2000. Layer extraction from multiple images containing reflections and transparency. In: Proceedings IEEE CVPR.

Tappen, M. F., Freeman, W. T. and Adelson, E. H., 2003. Recovering intrinsic images from a single image. In: Advances in Neural Information Processing Systems 15 (NIPS).

Weiss, Y., 2001. Deriving intrinsic images from image sequences. In: Proc. of ICCV 2001.



(a): Original image



(b): Region of shadow



(c): Image of reflection model $(K^{\#}(p, 62^{\circ}, 0.53))$



(d): Result of shadow/shade removal Figure 1: Example 1.



(a): Original image.



Figure 3: Magnified results of example 2. (a): the result of Oren-Nayar model. (b): the result of Lambertian model.



(b): Image of reflection model $(K^{\#}(p, 46^{\circ}, 0.55))$



(c): Result of shadow/shade removal Figure 2: Example 2



(a): Image of reflection model (Lambertian)



(b): Result of shadow/shade removal by Lambertian model

Figure 4: Example 2 processed with Lambertian model.