Removal of Specular Reflection Component Using Multi-view Images and 3D Object Model

Shu-Kam Chow and Kwok-Leung Chan

Department of Electronic Engineering, City University of Hong Kong, 83 Tat Chee Avenue, Kowloon, Hong Kong {shukchow,itklchan}@cityu.edu.hk

Abstract. Image-based 3D model reconstruction method can use the same multi-view image sequence of the object for the generation of both the geometry model and texture map. Texture is very critical for virtual exhibition of 3D model and should be of high quality comparable to the geometry data. One problem is that the object surface may exhibit specular reflection of illuminated light. The texture extracted directly from the images can be unnatural. We propose a method for the removal of specular reflection component in each image. Each camera view is calibrated and a 3D mesh model of the object is generated. For each triangle patch, the projected colors on all visible views are found. The specular chromaticity is replaced by the corresponding diffuse chromaticity. We test the method on image sequences of synthetic and real objects. The diffuse image sequence can be used to generate the texture map.

Keywords: specularity removal, dichromatic reflection model, 3D model reconstruction, texture mapping.

1 Introduction

To create photorealistic three-dimensional (3D) model of real scenes and objects is an old and challenging computer vision problem. We intend to develop a system for the reconstruction and interactive exhibition of 3D object models on the web. Real object models can be reconstructed automatically using active and passive methods [1]. Object range scanning by laser and structured light are typical examples of the active methods. One of the most significant advantages of laser scanners is their high accuracy in geometry measurements. The passive methods exploit images of the object acquired by digital cameras locating at different viewpoints and reconstruct the 3D model using photogrammetry. They are of low cost and useful when direct access to the object is prohibited. While there are methods for 3D digitization of objects with high geometric resolution, there are still limitations in generating high quality texture for virtual exhibition. Texture is very critical for virtual exhibition of objects such as cultural artefacts and often should be of high resolution comparable to the range data [2]. One problem is that the object surface exhibits specular reflection of illuminated light during the acquisition of surface texture. The texture reconstructed from the multi-view image sequence, such as by texture stitching, can be unnatural. Also, the lighting during texture capturing of one object may not be the same as that of other objects. This is certainly undesirable for a virtual scene with many objects.

To solve the problem, we need to remove the specular reflection and relight the texture mapped 3D model. We propose a method for the removal of specular reflection component in each image. Each camera view is calibrated and a 3D mesh model of the object is generated. With the availability of the 3D model and multi-view images, the correspondence of the image pixels is established. For each triangle patch, the projected colors on all visible views are found. The specular chromaticity is replaced by the corresponding diffuse chromaticity. No assumptions on the light source characteristics are made.

A review of previous work on specularity removal is presented in the following section. Section 3 presents the steps involved in the reconstruction of 3D object from a multi-view image sequence. Section 4 explains our proposed technique for the removal of specularity reflection component in each image. We test our method on image sequences of synthetic and real objects. Finally we conclude the paper and suggest some future work.

2 Related Work

Many methods for the identification and removal of specular reflection component demand polarizer filter, light source arrangement or rely on information obtained in one or more images. Due to page limit, we only present a review on the latter approach in here.

Klinker et al. [3] first introduce the single-image approach for specularity removal. Based on the dichromatic reflection model (most applicable to dielectrics) [4], diffuse and specular pixels form linear clusters of skewed-T in the color histogram. The limitations are that surfaces should be homogeneous (non-textured, locally valid if presegmentation is performed) with negligible image noise, specular lobe is sufficiently narrow, and diffuse (object) and specular (illuminant) colors are different. Tan et al. [5] introduce an inpainting technique for specularities removal. Information for determining the diffuse reflection of specular pixel is obtained based on the dichromatic reflection model and the assumption of a uniform illuminant color in the highlight. The method can preserve the diffuse shading in the highlight region which has been delineated manually. Problems may arise from discontinuities in surface colors. Later, they [6] extend the method by examining higher-order color data in the form of spatial color distributions. This data, which can be obtained outside the highlight region, can provide valuable information in determining diffuse colors within the highlight region. Tan and Ikeuchi [7] propose an iterative method to separate diffuse and specular reflection components. The image is normalized by the illumination chromaticity. This image is then transformed to a specular-free image by shifting the intensity and maximum chromaticity of the pixels nonlinearly while retaining their hue. Diffuse pixels are identified by using intensity logarithmic differentiation on both the normalized image and its specular-free image. The restrictions are that diffuse reflection component always exists in each pixel, and the surface color is chromatic.

Lin and Shum [8] propose a method to perform decomposition of diffuse and specular reflection using two photometric images. Surface reflection is modeled as a linear function of three basis functions. Specular reflection is assumed to have the same spectral composition as the incident illumination. Corresponding pixels in two

images are assumed to be known. Specular highlights do not saturate and do not spatially overlap between the images. Lin *et al.* [9] propose a method based on color analysis and multi-baseline stereo that simultaneously separates the diffuse and specular reflections and estimates the true depth. Specular pixels are identified by a votingbased multiple tri-view color histogram differencing. Disparity of specular pixels is estimated based on neighboring diffuse pixels according to continuity constraint. The corresponding diffuse pixels are found based on the disparity of specular pixel. The method is applicable on grey-scale image. The method is sensitive to color saturation, image noise, and changes of specular color among the viewpoints.

Mallick *et al.* [10] present a unified framework for separating diffuse and specular reflection components in images and videos. Using a known illumination color, each image is transformed into SUV space, and three functions are computed. Specular/diffuse separation is achieved locally by numerically evolving the appropriate multi-scale erosion PDE on one function. Yu *et al.* [11] propose an algorithm to reconstruct the triangulated 3D model and the View Independent Reflectance Map (VIRM) from multiple calibrated images of an object. VIRM represent the joint effect of the linearly related diffuse and specular reflections, and illumination. The reflectance model is designed for objects with the same material and is illuminated by distant lighting with no self-shadowing and inter-reflections. Criminisi *et al.* [12] describe a new approach for analyzing the epipolar-plane-image (EPI) volume generated by the image sequence captured from a linearly translating camera. The EPI volume can be segmented into EPI strips and EPI tubes. Based on the geometric and photometric behavior of specularities in the EPI volume, the static scene can be decomposed into its diffuse and specular components.

3 3D Model Reconstruction

The camera calibration is to obtain the intrinsic and extrinsic parameters defining the internal camera properties and the viewpoint orientation with respect to the object. A box is fabricated which has calibration patterns posted on 5 viewable planes. During image acquisition, the target object is placed onto the top plane of the box (see Figure 1). Both the object and the calibration patterns are captured simultaneously in each image. Therefore, the camera can be placed anywhere and each view can be calibrated independently. Each calibration pattern consists of groups of red and blue lines. Each intersection of a red line and a blue line is a calibration point. We adopt Tsai's method [13] for camera calibration. Depending on the viewpoint, coplanar or non-planar algorithm can be employed.

The volumetric modeling is an important step in the reconstruction of 3D object model from a multi-view image sequence. It represents a world coordinate frame and the volume space in which the object occupies, and makes decision about whether a volumetric primitive (voxel) contains the object. The contour of the object is simply extracted from each of the input images, e.g. by the blue-screen technique with the use of a monochromatic background. We adopt the voting-based shape-fromsilhouette method [14] for volumetric modeling.



Fig. 1. Picture of target object and calibration box

It is more favorable to represent the 3D object model in piecewise-planar surfaces. Majority of rendering techniques such as shading and hidden surface removal are performed on geometrically defined surfaces. Moreover, the polygon is easy to process, and so the rendering and display of object can be speed up. Therefore, in our system, the volumetric model is converted into polygonal surface model using the marching cubes algorithm [15]. The number of triangles generated by the marching cubes algorithm is as much as the number of voxels in the volumetric model. However, such a lot of triangles are not only useless to improve the visual quality of the model, but also burden the overall reconstruction process and even increase the computational time. Therefore, polygon simplification is implemented to simplify the model. When the model is simplified, neighboring polygons sometimes form a sharp transition. This defect is easily observed in the curved surface of the object. To overcome this problem, surface smoothing is performed in order to provide the visually pleasing surface model.

4 Specularity Removal

With the calibrated multi-view images, the correspondence of the image pixels is established. But for real objects, the shape can be complex. Therefore, visibility of the model must be checked. For each triangle patch, the projected colors on all visible views are found. The diffuse chromaticity is determined by searching for the maximum chromaticity. Special care is taken when performing the specularity removal on texture boundary. The specularity removal algorithm is shown below.

For each camera view Find triangle patches that are facing backward Find triangle patches that are occluded Create a list of visible surface for that camera view End For For each image For each visible triangle patch Get the projected colors in the current camera view Get the corresponding projected colors in other camera views Determine the diffuse colors End For Replace the specular colors by the corresponding diffuse colors End For

4.1 Visibility of Triangle Patch

Here, we want to find out those triangle patches that are invisible (e.g. facing backward or occluded by other triangle patches) in a particular camera viewpoint. This is very important to eliminate all the impossible projections and make our multi-view specular-to diffuse color transformation reliable.

A back face triangle is invisible in a particular camera view and hence does not provide true projected colors. Basically, our algorithm compares the orientation of the triangles with the camera viewpoint and determines those triangles that cannot be seen. The visibility test is best carried out in the view space. We examine the dot product of the triangle's outward normal vector \hat{N}_t and the vector from the camera's centre of projection \hat{N} as shown in Figure 2. The triangle is visible only if the value of the dot product is larger than the pre-defined threshold (i.e. $\hat{N}_t \bullet \hat{N} > threshold$).



Fig. 2. Identification of back face triangle (\hat{N}_t is the unit normal vector of triangle, \hat{N} is the 'line of sight' unit vector)

For concave or self-occluded object, some triangles may be occluded by other surfaces in a particular camera view. Occluded triangles should be removed from the list of visible surface V for that viewpoint. We identify the occluded triangles by counting the number of non-occluded pixels projected on the camera view C. We assign the triangle ID as a label to each projected pixel of a target triangle T_1 . If another triangle T_2 projects to a region that overlaps with the triangle T_1 , and T_2 is closer to the camera than T_1 , we update the label with the T_2 's triangle ID. If all the labels of T_1 are being overwritten, T_1 will be removed from the list of visible surface V for camera view C.

4.2 Specular-to-Diffuse Transformation

To remove the specular reflection component, we first estimate the illumination color which is used for normalizing and renormalizing the whole set of multi-view images. Various methods exist for illumination color estimation. Here, we adopt the single-image based illumination color estimation method proposed by Tan et al. [16].



Fig. 3. Projection of specular surface on multi-view images

Assume the image to be corrected is I_1 for camera view C_1 as shown in Figure 3. For a visible triangle patch T, the projected color is P_1 . If the object surface is specular and for this orientation of the triangle patch with respect to the light source direction and camera viewpoint, the projected color appears brighter than the actual color of the object. Assume T is also visible in camera view C_2 . The corresponding projected color is P_2 on image I_2 . As C_2 is away from the ideal specular reflection orientation, the projected color is close to the actual color of the object. After checking all the visible views, we obtain a list of projected colors for T which can help us to obtain the diffuse color. Lastly, we replace the specular color by the diffuse color.

We adopt the dichromatic reflection model. Illumination at any point comes from a single light source plus ambient light which is in equal amounts in all directions. Therefore, light reflected from a dielectric object P is a linear combination of diffuse P_d and specular P_s reflections:

$$P = m_d P_d + m_s P_s \tag{1}$$

where m_d and m_s (varying from point to point) are the magnitudes of diffuse and specular reflections respectively. It is assumed that the spectral power distribution of the specular reflection is similar to the spectral power distribution of the incident light. Therefore, P_s can be considered as the illumination color. When the image is normalized by the estimated illumination color, the reflection model is changed to:

$$P' = m_d P'_d + m_s \tag{2}$$

where
$$P' = \frac{P}{P_s}$$
 is the normalized observed color, and $P'_d = \frac{P_d}{P_s}$ is the normalized

diffuse color. Assume we obtain M projected colors in M visible views of the triangle patch T. They all share the same $m_d P'_d$ with varying m_s . As the camera is made rotating around the object, it is very likely that the M projected colors contain the diffuse color. Purely diffuse pixel has the maximum chromaticity. Therefore, to solve for $m_d P'_d$, we search for the maximum chromaticity in the list of visible projected colors.

As the mesh model is simplified, the projection of a triangle patch on the image is a region rather than a single pixel. The projected region may contain different colors due to texture or shading of the object. Therefore, we segment the region into different sub-regions. Each sub-region will only contain one color. We avoid using the original colors since specular highlight may occur within the region. Therefore, we group the colors into sub-regions by using the specular-free color. The chromaticity κ is defined as:

$$\kappa = \frac{P'}{\sum_{R,G,B} P'} = \frac{m_d P'_d + m_s}{m_d \sum_{R,G,B} P'_d + 3m_s}$$
(3)

This equation can be reformulated to produce the definition of m_s :

$$m_{s} = \frac{m_{d} \left\{ \left[\kappa \sum_{R,G,B} P'_{d} \right] - P'_{d} \right\}}{1 - 3\kappa}$$
(4)

By substituting Equation (4) into Equation (2):

$$P' = \frac{\kappa m_d \left[3P'_d - \sum_{R,G,B} P'_d \right]}{3\kappa - 1}$$
(5)

In normalized image, we can set $\sum_{R,G,B} P'_d = 1$ without loss of generality. When the pixel is converted to purely diffuse (when specular reflection component is eliminated, $P' = m_d P'_d$), $\kappa = P'_d$ and $m_d = \sum_{R,G,B} P'$. Therefore, Equation (5) can be reformulated as:

$$m_{d} = \frac{P'[3\kappa - 1]}{\kappa[3P'_{d} - 1]}$$
(6)

Purely diffuse pixel has the maximum chromaticity. Referring to Equation (3), P' and P'_d are at their maximum $\overline{P'}$ and $\overline{P'_d}$ respectively. The color of the pixel is specular-free, that is, $m_d = \sum_{R,G,B} P' = \sum_{R,G,B} P'_{sf}$. Therefore,

$$\sum_{R,G,B} P'_{sf} = \frac{\overline{P'}[3\overline{\kappa} - 1]}{\overline{\kappa}[3\overline{P'_d} - 1]}$$
(7)

where $\overline{\kappa} = \frac{\overline{P'}}{\sum_{R,G,B} P'}$. While $\overline{P'}$ and $\overline{\kappa}$ can be obtained from the normalized image,

 $\overline{P'_d}$ is unknown. By setting $\overline{P'_d}$ to a sufficiently large value (e.g. 0.5) for all pixels, we can obtain the sum of color intensities of the corresponding specular-free image using Equation (7). For non-diffuse pixels:

$$m_{s} = \frac{\sum_{R,G,B} P' - \sum_{R,G,B} P'_{sf}}{3}$$
(8)

$$P'_{sf} = P' - m_s \tag{9}$$

Although the specular-free color may not equal to the original color, it is very consistent on different views. If the difference of the specular-free colors of two pixels is less than a pre-defined threshold, we group the two pixels into one sub-region. For each sub-region, we search for the existence of similar sub-regions in all visible views. Using similar idea, we check the specular-free color of sub-regions in other visible views. If the color difference of two sub-regions is less than the sub-region difference threshold, we save the sub-region color into a color list. In order to preserve the texture or pattern on the object surface, we applied an adaptive threshold in this checking. We set a low sub-region difference threshold value initially. Then we increase the threshold if the checking fails to find a sub-region. In case the checking fails after a certain number of iterations, we ignore this view and proceed to the next view.

In real situation, it is common to encounter saturated highlights which violate the assumption of the dichromatic reflection model. While the transformation method mentioned so far can only handle non-saturated specular reflections, we need a special procedure to cater for saturated specular reflections. Before the image is normalised, we detect the existence of grey scale colors by checking the maximum chromaticity of the pixels. The pixel is identified as grey scale if its maximum chromaticity is less than the grey scale detector threshold. We cannot use the specular-free color to find the corresponding pixels since the specular-free color of grey scale pixels are almost black. Instead, we simply get the colors directly from other corresponding and visible views and search for the diffuse color.

5 Results

First, we test our method on a synthetic object. The first row of Figure 4 shows different views of the synthetic teapot model which is illuminated by some light sources. The corresponding diffuse images are obtained as ground truth (second row of Figure 4). The third row of Figure 4 shows the diffuse images obtained by our method. The estimated illuminant color is (R=0.27, G=0.34, B=0.39). The sub-region color difference threshold is initially set as 5 and the increment in each iteration is 5. For better judgment of the specularity removal capability, the difference between the ground truth and the diffuse images obtained by our method are shown in the last row of Figure 4. It can be seen that most of the highlights are removed.

Then, we test our method on a real object. The left column of Figure 5 shows different views of the plastic toy object. The middle column of Figure 5 shows the diffuse images obtained by our method. Each image is interactively segmented into 2 parts and processed separately. The right column of Figure 5 shows the mask of one part. The illuminant color estimated in the region of the mask is (R=0.3, G=0.46, B=0.46). The illuminant color estimated in the remaining region is (R=0.32, G=0.27, B=0.39). The sub-region color difference threshold is initially set as 2 and the increment in each iteration is 2. The grey scale detector threshold is 0.35. Figure 6 shows the zoom in views of the results of specularity removal. It can be seen that most of the highlights are removed and replaced by the correct colors.

Original image	
Ground truth diffuse image	
Diffuse image obtained by our method	客たれ
Difference between ground truth and diffuse image ob- tained by our method	

Fig. 4. Results of specularity removal on synthetic object

Original image	Diffuse image obtained by our method	Region mask

Fig. 5. Results of specularity removal on real object



Fig. 6. Zoom in views of the results of specularity removal

6 Conclusion

We propose a method for the removal of specular reflection component on multi-view images. Each camera view is calibrated and a 3D mesh model of the object is generated. For each triangle patch, the projected colors on all visible views are found. The specular chromaticity is replaced by the corresponding diffuse chromaticity based on the fact that a purely diffuse pixel has maximum chromaticity. We test the method on image sequences of synthetic and real objects. In the future, we intend to speed up the method by parallelizing the algorithm. We also want to investigate how good our method performs on other kinds of object.

Acknowledgement

The work described in this paper was fully supported by a grant from the Research Grant Council of Hong Kong Special Administrative Region, China [Project No. CityU 110807].

References

- Pavlidis, G., Koutsoudis, A., Arnaoutoglou, F., Tsioukas, V., Chamzas, C.: Methods for 3D digitization of cultural heritage. Journal of Cultural Heritage 8, 93–98 (2007)
- Blais, F., Beraldin, J.-A.: Recent developments in 3D multi-modal laser imaging applied to cultural heritage. Machine Vision and Applications 17, 395–409 (2006)
- Klinker, G.J., Shafer, S.A., Kanade, T.: The measurement of highlights in color images. International Journal of Computer Vision 2(1), 7–32 (1990)
- Shafer, S.: Using color to separate reflection components. Color Research and Applications 10, 210–218 (1985)
- Tan, P., Lin, S., Quan, L., Shum, H.-Y.: Highlight removal by illuminant-constrained inpainting. In: Proceedings of IEEE International Conference on Computer Vision, pp. 164– 169 (2003)
- Tan, P., Lin, S., Quan, L.: Separation of highlight reflections on textured surfaces. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 1855–1860 (2006)
- Tan, R.T., Ikeuchi, K.: Separating reflection components of textured surfaces using a single image. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(2), 178– 193 (2005)
- Lin, S., Shum, H.-Y.: Separation of diffuse and specular reflection in color images. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 341–346 (2001)
- Lin, S., Li, Y., Kang, S.B., Tong, X., Shum, H.-Y.: Diffuse-specular separation and depth recovery from image sequences. In: Proceedings of European Conference on Computer Vision, pp. 210–224 (2002)
- Mallick, S.P., Zickler, T., Belhumeur, P.N., Kriegman, D.J.: Specularity removal in images and videos: a pde approach. In: Proceedings of European Conference on Computer Vision, Part I, pp. 550–563 (2006)
- 11. Yu, T., Xu, N., Ahuja, N.: Shape and view independent reflectance map from multiple views. International Journal of Computer Vision 73(2), 123–138 (2007)
- Criminisi, A., Kang, S.B., Swaminathan, R., Szeliski, R., Anandan, P.: Extracting layers and analyzing their specular properties using epipolar-plane-image analysis. Computer Vision and Image Understanding 97, 51–85 (2005)
- Tsai, R.Y.: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. Journal of Robotics and Automation RA-3(4), 323–344 (1987)
- Matsumoto, Y., Fujimura, K., Kitamura, T.: Shape-from-silhouette/stereo and its application to 3-D digitizer. In: Proceedings of Discrete Geometry for Computing Imagery, pp. 177–188 (1999)
- 15. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3D surface construction algorithm. Computer Graphics 21(4), 163–169 (1987)
- Tan, R.T., Nishino, K., Ikeuchi, K.: Color constancy through inverse-intensity chromaticity space. Journal of Optical Society of America A 21(3), 321–334 (2004)