Imaging artwork in a studio environment for computer graphics rendering

Brittany D. Cox and Roy S. Berns

Munsell Color Science Laboratory, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY, USA 14623

ABSTRACT

When viewing a reproduction of a painting in textbooks and online, important visual information, including surface texture, can be lost. Providing an experience to viewers that can convey some of this lost information without significantly increasing the necessary equipment and training possessed by a typical studio photographer would enrich education, documentation, conservation, and presentation of artwork for the public. A modified photometric stereo technique coupled with the rendering software mental ray packaged with Maya® is presented as a means of capturing surface normal maps and diffuse color information used in the rendering of realistic attributes of paintings. mental ray's ability to realistically render two different paintings with different gloss properties was evaluated by comparing the proposed capture technique and a previously published technique that employs cross-polarization, demonstrating that traditional imaging is a viable technique for generating input data for computer graphics rendering software.

Keywords: Photometric Stereo, Rendering, Fine Art Imaging

1. INTRODUCTION

It is common practice in museum culture for paintings to be lent to other institutions or removed from exhibition for the purpose of addressing conservation needs, leading to a desire of conservators and curators to track changes of the physical appearance and structure of artwork. In addition, there is a growing interest in digital galleries, illustrated by the recent popularity of sites such as Google Cultural Institute, a collaboration between Google Inc. and hundreds of museums and cultural archives, to allow public, digital access to a growing number of fine artworks, historic photographs, and religious artifacts.[1]

Rendering a visually acceptable virtual painting requires information about color (diffuse albedo), microstructure (responsible for gloss), and macrostructure (responsible for depth). Surface normal information, a vector perpendicular to the surface plane, can be used to convey information about the surface macrostructure. This is an acceptable substitute for a depth map when used with a rendering software, such as Maya®. Thus, development of an easily implemented studio imaging system for capturing diffuse albedo and surface normal information of paintings is very valuable.

A frequently used technique for capturing surface information is photometric stereo, first proposed by Woodham in 1980, with the assumption of a Lambertian surface, that is, a surface that does not produce specular highlights.[2] According to this technique the first step in calculating a surface normal map is to determine the lighting direction for each light. The lighting direction, $\mathbf{L} = [l_x, l_y, l_z]$, for a point with a given surface normal $\mathbf{N}$ and diffuse surface albedo ρ is used to define the intensity of each pixel value according to **Eq. 1**.

$$I = \rho \mathbf{L} \cdot \mathbf{N} \tag{1}$$

This equation can be solved for $\rho\mathbf{N}$ using linear regression resulting in **Eq. 2** (written in a form that accounts for multiple light sources), with the constraint of the unit vector, $\sqrt{N_x^2 + N_y^2 + N_z^2} = 1$.

$$\begin{bmatrix} \rho N_x \\ \rho N_y \\ \rho N_z \end{bmatrix} = \begin{bmatrix} l_x^1 & l_y^1 & l_z^1 \\ l_x^2 & l_y^2 & l_z^2 \\ \vdots & \vdots & \vdots \\ l_x^n & l_y^n & l_z^n \end{bmatrix}^{\dagger} \times \begin{bmatrix} I^1 \\ I^2 \\ \vdots \\ I^n \end{bmatrix} I \tag{2}$$

A common practice to determine surface normal is to employ a three-light photometric stereo technique.[2,3,4] Research in this area has focused on increasing the number of lights to improve the accuracy of the resulting normal map. Rushmeier, *et al.* used a photometric stereo technique with five lights, a fixed video camera and a video capture board.[5] Their method eliminated the maximum and minimum values assuming that highlights and shadows are represented by these values, respectively. This method was capable of representing fine scale and self-shadowing surfaces but was inadequate in its ability to represent objects with widely varying BRDF values across their surface.

Sun, *et al.* expanded the traditional thee-light photometric stereo technique to include the use of six lights and a hierarchal sorting technique.[4] This selection method was used to exclude highlights and shadows. Six lights were incorporated into their technique to guarantee that each location on a sample was illuminated by at least three lights once highlights and shadows were excluded. The authors considered six lights to be the minimum number of lights necessary to insure a three light minimum illumination of each area.

Berns, *et al.* employed the use of polarizers with a 4LI, or four-light imaging, setup to remove specular highlights.[6] Cross polarization was used to eliminate first surface reflectance including specular highlights. This technique addressed the problems associated with highlights but did not directly eliminate the presence of shadows. Similar to Rushmeier, a shadow was assumed to occur in only one light direction. That reduced the number of light directions to three since the signal was null. A psychophysical experiment was performed to compare techniques that directly eliminated shadows and those that did not. The results of this experiment showed that the presence of shadows did not negatively affect the appearance of a rendered painting.[7]

The Berns, *et al.* method was implemented in a photography studio by adding two Xenon strobes to the usual set up of a pair of strobes placed on each side of the painting. The expectation was that the opposing-lit images could be used to render images equivalent to conventional imaging without polarizers so that a single set of images could be used for computer graphics and reprographics. Unfortunately, cross polarization resulted in images that appeared blurry.[8] Therefore, reprographics would require additional images with either linear or uncontrolled polarization. This paper describes the results of using a selection technique to remove highlights from images captured using the 4LI system, but without using cross-polarization.

2. IMAGING

Both the 4LI polarization technique and the new technique employed four Broncolor Xenon strobes placed at 45° from the surface of the painting and 90° from each neighboring light along an annulus with the camera, in this a case a Sinar 86H 48 megapixel camera, placed perpendicular to the surface of the painting. Figure 1 is a schematic of the light locations in reference to the painting and camera. The only difference in the physical setup between the Berns method and the new technique is that the Berns, *et al.* method placed a linear polarizer in front of each light and camera while the new method did not.
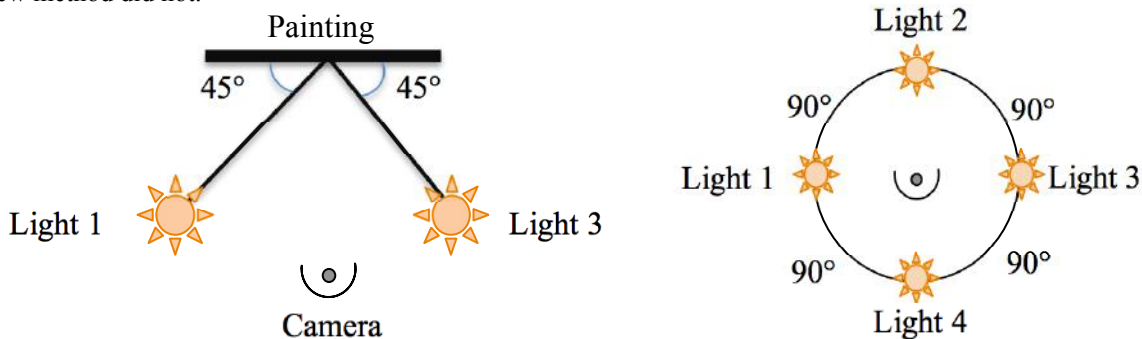

Figure 1. Schematics of the light placement relative to the painting. The figure to the left does not show Lights 2 or 4. Lights are placed 45° from the painting's normal and 90° from each other on a circular path.

The procedure outlined by Berns, *et al.* was performed and repeated for both methods. Due to a limited ceiling height, the bottom strobe was placed at roughly 35° from the painting. Images were taken using each light sequentially of a

uniform white background for flat fielding, a glossy black cue ball to define each light direction, a Xrite ColorChecker Classic for color calibration, and two acrylic paintings, one varnished and one unvarnished. A traditional image of the paintings taken with lighting from both the left and right side is shown in Figure 2.

Figure 2. Traditional image of unvarnished and varnished paintings taken with left and right strobes, simultaneously.

3.   IMAGE ANALYSIS, RESULTS, AND DISCUSSION

Two maps are required for computer graphics rendering: diffuse albedo and surface normal. For the surface normal maps, the computational removal of highlights was achieved by implementing a thresholding method, written mathematically in **Eq. 3**. This method assumed that a highlight is only present in a pixel for one direction at a time. In the thresholding equation, x and y denote a pixel location, imgA is a single image taken from one of the lights directions, avgImg is the calculated average without the maximumum value of all four light directions, and k is a constant, accounting for any skewing that may have occurred if a pixel was in shadow. The constant k was 0.75, selected by visual evaluation. The goal was to avoid any unnecessary removal of pixels, thus preserving as much surface normal information while still eliminating highlights. Pixels above the threshold were assumed to be highlights and removed, resulting in three light directions for calculating the surface normal.

$$if\ imgA(x,y) > avgImg(x,y) + avgImg(x,y) * k,$$
$$then\ imgA(x,y)\ is\ a\ highlight$$
(3)

Figure 3 is a false-color map indicating which pixels have been determined to have highlights from a given direction in dark blue (left), green (right), orange (bottom), and light blue (top). The pixels containing no highlights are shown in dark gray. Pixel values for this portion of the procedure come from averaging the R, G, and B values of each pixel in each image.
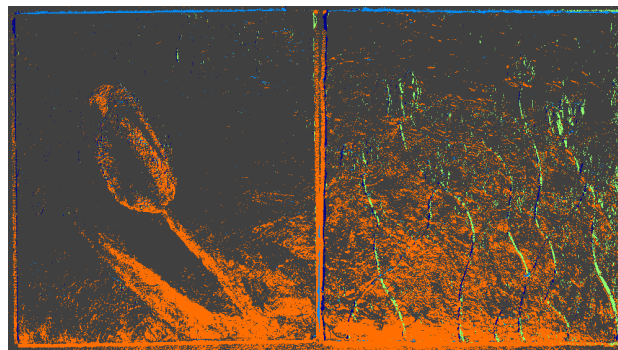

Figure 3. Image identifies which lighting direction was responsible for a highlight. Dark gray pixels indicate that no highlight was present. Orange pixels indicate highlights from the bottom light, green pixels were from the right light, light blue pixels were from the top light, and dark blue pixels were from the left light. The unvarnished painting is on the left and the varnished painting is on the right.

The varnished painting had a glossier surface than the unvarnished paining which naturally produced more highlights. A comparison of the images in Figure 3 where more pixels are categorized as highlights in the image of the varnished painting (right) than of the unvarnished painting (left) where the majority of the pixels are red suggests that the thresholding technique is not just selecting high values and eliminating them, but actually selecting highlights. A large number of highlights come from the bottom light. This is due to the slightly closer placement of this light to the paintings during images. Although not an ideal placement, this will test the robustness of these techniques to lighting based biases.

Once the highlights were computationally removed, a normal map was calculated using information from three or four lights for the appropriate pixel. These calculations, like those done by the 4LI method, were based on the technique proposed by Woodham (equations 1 and 2, above).[2] Figure 4 and 5 contain images of the surface normal map calculated from the threshold method (center) and calculated from cross polarized images taken in accordance with the technique described by Berns, *et al* (left) where N_x, N_y, and N_z were encoded as *R*, *G*, and *B*.

A comparison of these two normal maps indicated that the threshold method for determining the use of three lights when a highlight is present and four lights when a highlight is not present created normal maps with more variation in the Z direction, *i.e.*, out from the painting. This is attributed to the increase in contrast present in these images that made the boundaries of impasto more pronounced. There is more high frequency detail in these maps, observed primarily in the bottom portion of both paintings. Figures 4 and 5 highlight the difference in detail between the two corresponding normal maps in a false color difference map (right). However, the normal maps resulting from the threshold method exhibited more noise while the cross-polarized images generated normal maps that provided smoother transitions in areas with changes in topography and reduced artifacts introduced by pixels in highlighted regions. The lighting bias of the bottom lighting direction had some effect on the images taken without polarization, shown by a greater blue variation at the bottom of both the varnished and unvarnished paintings in Figures 4 and 5 (center).
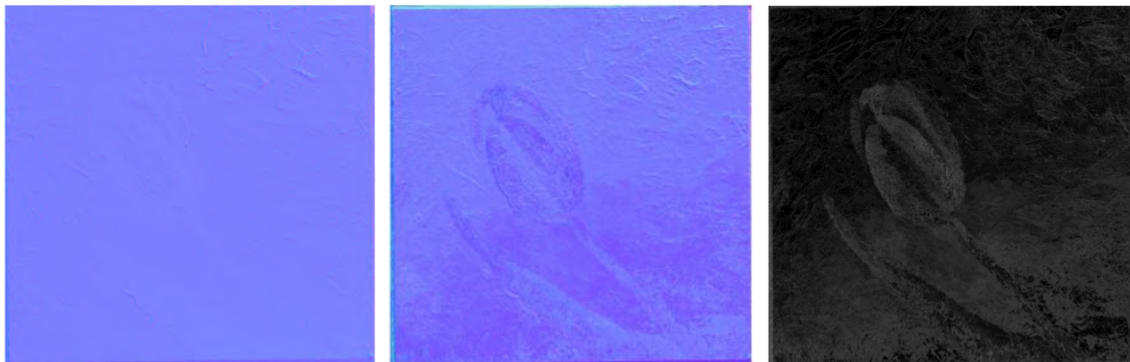

Figure 4. Normal maps of the unvarnished painting. Left image used cross-polarized data to generate a normal map using four lights. The center image used thresholded data from three or four lights when appropriate. The right image is a difference map.
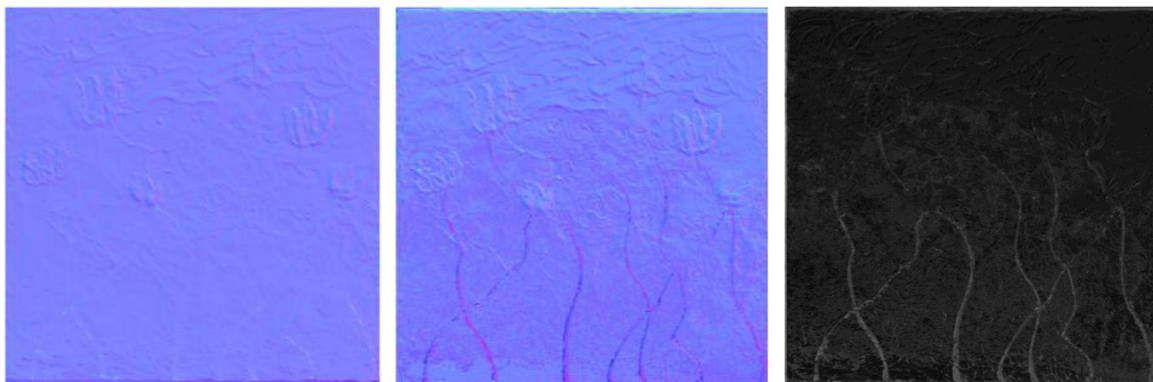

Figure 5. Normal maps of the varnished painting. Left image used cross-polarized data to generate a normal map using four lights. The center image used thresholded data from either three or four lights when appropriate. The right image is a difference map**.**

The second map required for computer-graphics rendering was the diffuse albedo (color) map. The diffuse color of the paintings was calculated by taking an average of each pixel excluding the maximum. Excluding the maximum removes the influence of highlights from the diffuse color. This differed from the 4LI method, in that Berns, *et al.* used all four pixel values, one from each light direction, to calculate the average. The diffuse map is intended to represent the color of each painting without the influence of first surface scattering. To display the color of each painting more accurately the Xrite ColorChecker Classic was used to derive a transform from RGB camera data to XYZ tristimulus values. The tristimulus values were then used to encode into sRGB for display. The transform from RGB camera signals to XYZ was derived by populating a three by three transformation matrix with offset that minimized the mean CIEDE2000 color difference between known CIELAB values of the Xrite ColorChecker Classic and the CIELAB values calculated from the camera.

Using a simple average of all four cross polarized images resulted in images with a decrease in contrast. By examining the images in Figures 6 and 7, a decrease in the amount of black present in cross polarized images compared to the traditional images averaged without the maximum pixel values between the set of four images is easily observable. This decrease in black resulted in the appearance of a less crisp image by comparison. Although the diffuse map calculated from the traditional images appeared sharper, the mathematical removal of highlights via the threshold method did not account for the effect of gloss due to first surface scattering across the entire painting.

The Saunderson equation was used to introduce a gloss compensation factor of 0.02 into the camera RGB to XYZ workflow to calculate internal reflection and simulate spherical illumination with specular reflectance excluded.[9] This gloss factor was chosen based on the glossy nature of the painting as well as the surface roughness. If this painting behaved more like a mirror, a factor equivalent to theoretical values based on the surface's refractive index would have been chosen. Gloss compensation resulted in a diffuse map that was more physically correct but did not suffer from the same lack of contrast that the cross-polarized diffuse map exhibited. The gloss factor compensation was not used for the unvarnished painting. The new diffuse map with gloss compensation is shown on the right in Figure 7. The black levels are now more similar to those depicted in the cross-polarized diffuse image but still have a sharper appearance.

Figure 6. Diffuse image of the unvarnished painting from cross-polarized images (left) and traditional images (right).

Figure 7. Diffuse color image of varnished painting from cross-polarized image (left), traditional images (middle), and with gloss compensation (right).

The final component to rendering, gloss, is addressed by setting the appropriate properties in Maya®. The diffuse images and their corresponding normal maps were imported into Maya® and the paintings were assigned properties of a matte plastic, a preset material of the software. Matte plastic was chosen because it was the closest preset material with characteristics similar to dry acrylic paint. The index of refraction of the unvarnished painting was set to 1.40 and the index of refraction of the varnished painting was set to 1.52, the approximate index of refraction values of dry acrylic paint and an average varnish, respectively.[10] The reflective glossiness property for the varnished painting was also changed to 0.600, compared to 0.500 for the unvarnished painting. The Fresnel Reflection ray-tracing option was used along with the mental ray rendering software plugin to generate realistic images that mimicked light interaction with the surface of each painting simulating a D65 spotlight 45° to the right of the center of the painting and a camera, or observer 45° to the left of the painting.

Qualitatively, there is little discernable difference between images rendered using the various discussed techniques, as shown in Figures 8 and 9. Imaging paintings without cross polarization (middle) produced a rendered image of the paintings that compared very well in image sharpness and surface texture detail to a rendering generated from cross-polarized images (left). Combining the surface normal map from the cross-polarized technique with the diffuse color map from the sharper thesholding method (right) also produced little discernable difference. Additionally, any effects of the angle of the bottom strobe on the normal maps from non-polarized images are not evident in the renderings of either of the paintings used in this study.

Figure 8. Rendered images of varnished painting from cross-polarized normal and diffuse maps (left), from traditional/thresholded normal and diffuse color maps (middle), and combining the cross-polarized normal map with the traditional/thresholded diffuse color map (right).

Figure 9. Rendered image of the unvarnished painting from cross-polarized normal and diffuse maps (left), from traditional/thresholded normal and diffuse color maps (middle), and combining the cross-polarized normal map with the traditional/thresholded diffuse color map (right).

4. CONCLUSIONS

A new imaging approach has been developed and implemented for capturing input data for computer graphics rendering of paintings. The approach improved previous research by eliminating the need for cross-polarization, reducing studio setup time, and improving image sharpness. Different thresholding and averaging methods for producing surface normal and diffuse albedo maps might produce images that are equivalent to so-called "beauty shots" where the photographer combines diffuse (e.g., umbrellas and diffusers) and direct lighting to maximize the most important properties of the painting when viewing a static image. The most important result is the possibility of having a single studio setup when imaging materials with appreciable micro- and macro-structure, much like a copy stand. Accentuating or diminishing structure becomes a dynamic tool, reducing the need for reshooting for different purposes.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Google Cultural Institute., "Google Cultural Institute," 2013, <https://www.google.com/culturalinstitute/about/> (1 August 2014).
[2]   Woodham, R. J., "Photometric Method for detemining surface orientation from multiple images," Opt. Eng. **19**(1), 139–144 (1980).
[3]   Okatani, T., Deguchi, K., "On Uniqueness of Solutions of the Three-Light-Source Photometric Stereo: Conditions on Illumination Configuration and Surface Reflectance," Comput. Vis. Image Underst. **81**(2), 211–226 (2001).
[4]   Sun, J., Smith, M., Smith, L., Midha, S., Bamber, J., "Object surface recovery using a multi-light photometric stereo technique for non-Lambertian surfaces subject to shadows and specularities," Image Vis. Comput. **25**(7), 1050–1057 (2007).
[5]   Rushmeier, H., Taubin, G., Guéziec, A., "Applying shape from lighting variation to bump map capture," Eurographics Render. Worshop, 35–44 (1997).
[6]   Berns, R. S., Chen, T., Wyble, D. R., Chen, L., "Update:* Practical Total Appearance Imaging of Paintings," IS&T Arch. Conf. Copenhagen(April), 1–6 (2012).
[7]   Chen, L., "Systems Evaluation for Computer Graphics Rendering of the Total Appearance of Paintings by," Master's Thesis, Rochester Institute of Technology (2013).
[8]   Berns, R. S., Chen, T., Chen, L., "Evaluating the Quality of Surfce Normal Maps of Paintings Using a Scanning Linear Source or Four-Strobe Photometric Stereo," Rochester, 1–16 (2012).
[9]   Abed, F. M., Berns, R. S., Masaoka, K., "Geometry-Independent Target-Based Camera Colorimetric Characterization," J. Imaging Sci. Technol. **57**(5), 1–15 (2013).
[10]   Berns, R. S., Rene de la Rie, E., "The Effect of the Refractive Index of a Varnish on the Appearance of Oil Paintings," Stud. Conserv. **48**(4), 251–262 (2003).