2003 IEEE International Workshop on Computer Architectures for Machine Perception (CAMP)

Illumination Correction in Digital Images

Somying Thainimit and Chee-Hung Henry Chu

Abstract— **Illumination in an image often needs to be corrected or adjusted for several reasons. Recognition of images, such as those of faces, that are based on image data as patterns depends on having the lighting direction of the query image and that of the database images be aligned. In computational video, acquired images are often texture mapped to blend with computer generated polygonal surfaces. The shading of a polygonal surfaces changes as the light source direction relative to the composite object changes due, for instance, to the object's motion. We experiment with methods that can adjust the texture mapped image according to the illumination change. We test our concept on images of human faces.**

I. INTRODUCTION

Perception of three-dimensional world objects in images is heavily influenced by the shading of the surfaces. The shading of three-dimensional surfaces in an image is a function of the surface orientation, the lighting directions, and possibly the viewing direction. Once an image is acquired, these factors must be uncoupled to adjust the shading of a pixel based on a change in any one of the factors. Consider the recognition of images of structures such as human faces. One approach is to treat the image data as patterns. In this approach, an image to be matched is compared to a library of images. If the lighting direction of these images are not matched, similarity measures may not be easy to define.

Another application that calls for adjusting lighting conditions in an acquired image is in blending an image with computer generated scenes. Digital manipulation of image and video data facilitates special effects operations that often cannot be recorded in the physical world. In many applications, acquired images are texture mapped and blended with computer generated surfaces to form an object or the overall scene. This could be done as an artistic expression or for such practical reasons as balancing the bandwidth with rendering speed. Computer generated surfaces can often be compressed to a much higher degree; on the other hand, the rendering speed can be a concern if the surface is overly complicated. In the latter case, an acquired image may be transmitted instead of the complicated surface since the rendering speed is constant for an image.

The acquisition of the surface data for a realistic rendering might not be possible, necessitating the use of acquired images. One such example is the surface of a human face. An avatar used, e.g., in email and chat sessions can be made with a texture mapped image of a subject's face. The avatar body can be represented as a polygonal Lambertian surface. When the avatar moves in the scene, the lighting direction changes so that the shading on the avatar body changes. The head-proxy with its texture mapped image has a different lighting direction as well, but the image is invariant of the lighting direction change. This can cause the avatar to lack a coherent appearance and be distracting to a viewer.

A solution to this problem is to adapt the texture map according to the avatar's orientation. Mipmapping is the most commonly used pre-calculation technique in texture mapping. It addresses the concern of anti-aliasing by adjusting the texture map according to the size of the texture mapped screen area. A series of images at different resolutions are pre-loaded and the image at the correct resolution is used during rendering. Our solution is therefore to pre-calculate a series of images, each corresponding to a particular orientation relative to the lighting direction. The correct image is then used as the texture map during the rendering process.

The key to the illumination correction approach is acquiring the series of images corresponding to different lighting conditions. A straightforward way to render a series of images of a face under different lighting conditions is to obtain the range and reflected color data using laser-triangulation scanners. These data can be used to build the geometric and material specifications of the surface of a face. Obviously, this direct data acquisition approach is not always possible in many applications.

Grayscale images of an object with a Lambertian sur-

Somying Thainimit is with the Department of Electrical Engineering, Kasetsart University, Bangkok, Thailand.

Chee-Hung Henry Chu is with the Center for Advanced Computer Studies, The University of Louisiana, Lafayette, LA 70504.

0-7803-7970-5/03/$20.00 ©2003 IEEE

face under all possible light sources can be generated by observing that all such images form a convex cone in $\Re^n$, where n is the number of pixels in each image. This convex cone is referred to as the illumination cone [1]. The illumination cone of an object can be constructed using as few as three images of the object with a fixed pose under different lighting directions. Although the human face is not convex, the 3D shape of a face can be recovered from at least three images taken under different, not necessarily known, lighting conditions [2].

An alternative to rendering a human face under arbitrary lighting and viewing directions is to use a basis set called reflectance functions, formed from images of the subject taken from different views and illuminated from different directions. This method requires that a large number of images, as many as 64×32, be recorded and processed. The subject has to maintain his or her pose throughout the acquisition process [3].

Our method is based on the quotient image rendering method, which uses a bootstrap set of images to synthesize color images of human faces from a single input image. We describe our method in the following section. Our results are shown in Section III and some concluding remarks are drawn in Section IV.

II. ILLUMINATION CORRECTION

The quotient image rendering technique [4] transforms an input image of a Lambertian object to another image rendered under a different lighting condition, using images of a collection of objects all having the same shape but differing in their surface texture or reflectance. To understand the quotient image rendering technique, consider the interaction between the lighting and the surface geometry that forms an observed intensity value. The reflected light at a particular point on a Lambertian surface Y is given by

$$I_Y = \rho_Y n_Y^T s$$

where ρ_Y is the reflectance, n_Y is the surface normal vector, and s is the vector from the point towards the light source. In the illumination correction approach, we need to synthesize another image I'_y with a different lighting direction s'. The new image is given by

$$I'_Y = \rho_Y n_Y^T s'.$$

The quotient image rendering technique allows us to form I'_Y given only I_Y and neither ρ_Y nor n_Y, under certain constraints and using a database of example images.

Consider an image formed by another object A under the same lighting condition s:

$$I_A = \rho_A n_A^T s.$$

If Y and A have the same geometric structure, i.e., $n_Y = n_A$, then a ratio of the image intensities is the ratio of the reflectances, defined as the quotient image of Y against A:

$$Q_{YA} = I_Y / I_A = \rho_Y / \rho_A.$$

Since the quotient image is a function of the reflectances, it is illumination invariant. Suppose we want to generate an image of Y illuminated by the lighting condition s_ν, denotedI_{Y,s_ν}. This image can be generated in terms of the quotient image Q_{YA} by

$$I_{Y,s_\nu} = \rho_Y n^T s_\nu = (\rho_Y/\rho_A)\rho_A n^T s_\nu = Q_{YA} I_{A,s_\nu}.$$

Hence, if we have an image of A illuminated by the desired lighting condition, then I_{Y,s_ν} can be formed from the quotient image and I_{A,s_ν}. This clearly is not always practical.

But suppose we have K images of A under different lighting conditions, we can generate images of A under s_ν if we can find $x_1, \cdots, x_K$ so that

$$s_\nu = \sum_j x_j s_j.$$

Then,

$$I_{Y,s_\nu} = Q_{YA} I_{A,s_\nu} = Q_{YA} \sum_j x_j I_{A,s_j}.$$

In practice, we start with, instead of a single image I_A, a bootstrap set of images of N subjects, each illuminated by K lighting conditions. The coefficients $x_1, \cdots, x_K$ are found by solving a set of homogeneous equations. The quotient image of an object Y against the bootstrap set can then be found from I_Y, the bootstrap images, and the coefficients $x_1, \cdots, x_K$. An image of Y illuminated by a novel lighting condition can then be synthesized from the quotient image and the set of bootstrap images.

The quotient image rendering method can synthesize a grayscale image under arbitrary lighting condition given an image of the object, and a bootstrap set of images. We extend the quotient image rendering method to handle color images so that we can synthesize a series of color illumination images.

The straightforward approach is to apply the quotient image rendering method to to each color channel independently. By assuming that lighting direction change does not affect the chromatic information of an image taken under it, we apply the quotient image rendering method only to the luminance channel. In our implementations, we convert the RGB data to the hue-saturation-intensity (HSI) and to the YCrCb formats. In either case, a new image for the luminance channel (the I- or the Y-channel) is synthesized.

In the implementation using YCrCb format, because the Y, Cr, and Cb channels are obtained from the R, G, and B channels via a linear transformation, if, at a pixel, the Y value is changed from y to y', the Cr and Cb values should also be scaled by the factor y'/y. The synthesized Y channel is then combined with the properly scaled Cr and Cb channels to form a color image.

III. Results

We apply our color illumination correction method to synthesize a set of images, each corresponding to a different lighting direction. A bootstrap set of three faces varying in illumination directions are taken from the Yale Face Database [5], which consists of images of fifteen individuals. We include seven subjects' faces, each with three images taken under three different light source directions. Examples of images of two subjects used in our bootstrap set are shown in Figure 1. Images of color novel faces are from Purdue University's AR face database [6].

All faces in the bootstrap set and the novel faces are roughly aligned using four manually selected facial features, viz. left and right eye corners, and left and right mouth corners. distance between two eye corners and distance between the mid point The aligned faces are cropped to the size of 128 × 128 pixels. After aligning and cropping the faces, the mid point between two eye corners of all faces are approximately at the same pixel location.

A color image with a specified lighting direction is synthesized using the method described in the previous section. In Figure 2, we show an example of a set of color illumination-corrected images synthesized in the HSI color space. The color illumination-adjusted images show the same face with varying illumination directions.

Fig. 1. Two faces taken under three distinct lighting conditions in the bootstrap set.

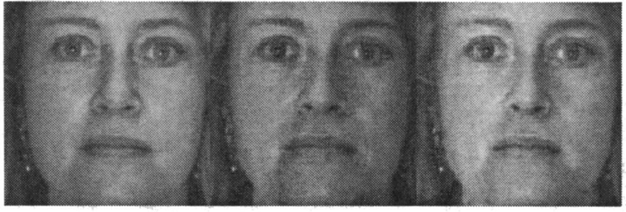
Fig. 2. A novel face image (*top row*) and its synthesized illumination-adjusted images (*middle and bottom rows*).

In Figure 3, we show an example of a novel face image and the images synthesized in the HSI format and in the YCrCb format. In both illumination-adjusted images, the lighting direction was successfully changed from the camera right to the camera left. The quality of the images are different, however.

Figures 4, 5, and 6 show another example of our color face rendering system, illustrating the effect of scaling the chromatic channels. The novel face used in this experiment is captured with a high intensity,

0-7803-7970-5/03/$20.00 ©2003 IEEE

Fig. 3. A novel face image (*left*) and the synthesized illumination-adjusted images obtained in the HSI (*middle*) and in the YCrCb (*right*) color spaces without scaling chromatic information.



non-uniformly distributed light source. The obtained gray scale synthetic faces generated using different color spaces, shown in Figure 5 look similar. However, its corresponding color faces (left and middle of Figure 6) are different. The illumination-adjusted images generated in the YCrCb color space has severe color distortion while that generated in the HSI space remains realistic. Results of new color faces generated using scaled chromatic information (right panel in Figure 6) show improvement.

Fig. 4. A novel face image (*left*) and its quotient images of the luminance channels in the HSI (*middle*) and in the YCrCb (*right*) color spaces.

Fig. 5. Synthesized faces using gray scale quotient image rendering on the luminance channels in the HSI (*left*) and in the YCrCb (*right*) color spaces.

Fig. 6. Illumination-adjusted images in color: rendered in the HSI space (*left*), in the YCrCb (*middle*) space without scaling the chromatic information, and in the YCrCb (*right*) space with scaled chromatic information.

IV. CONCLUDING REMARKS

We extended the quotient image rendering method to handle color images. The new method can synthesize an image of an object under an arbitrary lighting condition given an image of the object and a bootstrap set of images containing surfaces with the same geometrical structure with the object's surface. This latter condition is largely met in the case of human faces; i.e., human face surfaces have similar geometrical structure unless viewed at close distances.

The method allows us to synthesize a series of images, with illumination adjusted, corresponding to images taken with different lighting directions. These images can be used as texture maps to represent an individual using an avatar. As the avatar moves about in the virtual world, the texture mapped face can have the correct shading corresponding to the lighting condition.

REFERENCES

[1] P. N. Belhumeur and D. Kriegman, "What is the set of images of an object under all possible lighting conditions?," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*,June 1996, San Francisco, Calif., pp. 270–277.

[2] A. Georghiades, D. Kriegman, and P. N. Belhumeur, "Illumination cones for recognition under variable lighting: Faces," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* June 1998, Santa Barbara, Calif., pp. 52–58.

[3] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *Proceedings of the 27th ACM Siggraph Annual Conference on Computer Graphics and Interactive Techniques*, July 2000, New Orleans, La., pp. 145–156.

[4] A. Shashua and T. Riklin-Raviv, "The quotient image: Class-based re-rendering and recognition with varying illuminations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 129–139, February 2001.

[5] Yale Face Database, `http://cvc.yale.edu/ projects/yalefaces/yalefaces.html`.

[6] A. Martinez and R. Benavente, "The AR face database," `http://rvl1.ecn.purdue.edu/ alexi/alexifaceDB.html`, 1998.