INVISIBLE LIGHT : USING INFRARED FOR VIDEO CONFERENCE RELIGHTING

Prabath Gunawardane, Tom Malzbender, Ramin Samadani, Alan McReynolds, Dan Gelb, James Davis

prabath@soe.ucsc.edu, tom.malzbender@hp.com, ramin.samadani@hp.com, alan.mcreynolds@hp.com, dan.gelb@hp.com, davis@cs.ucsc.edu

ABSTRACT

Desktop video conferencing often suffers from bad lighting, which may be caused by harsh shadowing, saturated regions, etc. The primary reason for this is the lack of control over lighting in the user's environment. A hardware-based solution to this problem would be to place lights near the video camera, but these would be distracting to the user. We use a set of infrared lights placed around the computer monitor to gather a sequence of frames which is used to infer surface normals of the scene. These are used in combination with a visible spectrum image to create an improved relighting result.

Index Terms— face relighting, image based rendering, video conferencing

1. INTRODUCTION

Well designed, good quality video experiences make a huge difference to the effectiveness of video conferencing. Ideally, this would be in a custom engineered environment, such as HP Halo and Cisco Telepresence. However, those systems easily cost hundreds of thousands of dollars. There is a growing need from both business and consumers for superior video conversations using laptops or workstations. In these more casual settings, where we lack control over the lighting, the video quality currently becomes unpredictable and may degrade the experience. Reliably lit video would improve the effectiveness of casual video conferencing. In this work, we produce good quality video by processing additional information obtained from infrared (IR) lights. This allows us control over lighting without the need to introduce distracting visible fill lighting.

Our approach consists of relighting the input video with the aid of surface normals obtained through photometric stereo. Multiple infrared LEDs are placed around the monitor, and turned on in sequence synced with a high-speed infrared sensitive camera. This set of images, each of which is under a different lighting condition, is then used to compute surface normals of the scene using photometric stereo. A color camera paired with the infrared camera with a beam-splitter is used to simultaneously capture visible spectrum images. These images are then relit using the surface normal information that was obtained from the infrared images, enabling the user to control the lighting environment.

Our main contribution is the novel photometric stereo based relighting method that utilizes infrared illumination, which presents a real-time rendering solution to video conference face relighting.

2. RELATED WORK

Controlled lighting in the visible spectrum : Image based relighting for the purpose of face relighting has been an active topic in research. Most of such papers focus on relighting using controlled lighting in the visible spectrum [1]. Lighting gantries consisting of hundreds [2] or thousands [3] lights have been used for introducing virtually illuminated actors into real environments. While having a large number of illumination conditions makes it easier to interpolate novel lighting, it is not practical for our purpose of desktop face relighting. Geometric models [4] and photometric stereo based approaches [5] [6] have also previously been used to generate synthetically relit models using fewer input images and light sources in the visible spectrum. However such systems with multiple flashing lights would be highly distracting and irritating to the user in the context of video conferencing.

Single image in the visible spectrum : Researchers have also worked on relighting using just a single image or a video stream (in the visible spectrum) as the input. However some of these methods use a prior computed from a collection of faces under different lighting conditions, such as the 3D morphable faces used in [7] and [8]. Structure from motion has been used to construct a 3D geometric model of the face which is then used for rendering purposes [9]. While being convenient from an acquisition hardware perspective, such solutions normally involve solving for a complex lighting model [10] or constructing a geometric model making a real-time implementation harder. A more straightforward approach has been to treat the face as an ellipsoid on a 2D plane and relight it with virtual lights placed in 3D space [11]. However this suffers from not being able to handle the subtle facial features which photometric stereo does a good job of capturing.

Using infrared lighting to improve visible spectrum images : Prior work has focused on relighting in the visble spectrum using infrared illumination. One approach has been to swap out the noisy visibility color channel (in the HSV color space) spectrum on a low lit visible spectrum image based on illumination criteria [12]. This technique is however not well suited for human face relighting, since it leads to a ghastly appearance. Both infrared and ultra-violet illumination was used with constraints based on image gradients in [13]. For desktop video relighting though safety concerns preclude the use of ultraviolet illumination. Controlled lighting in the infrared spectrum for video conferencing has been explored previously in [14]. This is the closest approach to our work, however, their approach focused on generating a uniformly lit output image whereas we allow directional and brightness control over the lighting while using fewer input lighting conditions. This leads to more natural looking illumination, as we will show in section 5, and also is especially important in a group video conference where several participants have to be merged from different environments.

3. HARDWARE SETUP

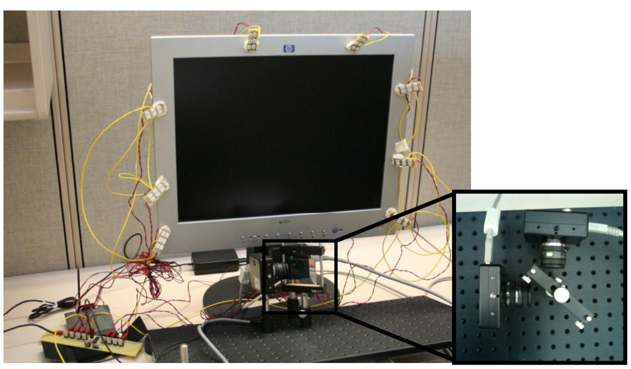

***Figure 1** : Hardware setup for infra-red based relighting. The LEDs are placed around the monitor. The infra-red and color cameras together with the cold mirror is shown in the inset.*

Photometric stereo [15] allows the computation of per-pixel normals and the albedo of a scene given a set of images under different lighting conditions. In our project, we use the same approach but use infrared lighting in place of visible spectrum lighting. We have placed a set of lights around the monitor (Figure 1), which are turned on in sequence and an image captured in the infrared spectrum. Concurrent to capturing these infra-red images, we also capture a visible image. The normals that are obtained from the infrared images are then used to improve the lighting on the visible spectrum image.

Our relighting setup consists of 8 infrared LEDs placed around the computer monitor (although only 4 input lighting conditions were used to generate the results in this paper), synced with a DragonFly Express camera running at 120Hz. The LEDs are lit sequentially with respective frames of the camera, which allows us to keep track of light under which the frame was captured. The DragonFly Express camera has a cold mirror (beam-splitter) mounted in front of it at 45° which reflects visible light (425-675nm) and allows infrared light (800-1200nm) through, ensuring that only the infrared portion of the spectrum reaches the camera. The reflected visible light is captured by a DragonFly 2 color camera, which is running at 15Hz, thereby capturing 1 frame for every 8 frames (a full sequence of infrared LEDs) of the DragonFly Express camera.

4. METHODOLOGY

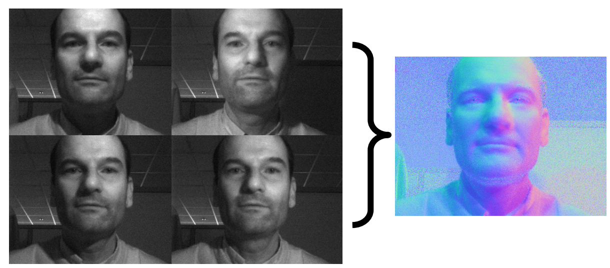

***Figure 2** : Input infrared images and the computed normal map*

Once these infrared images are acquired, we compute surface normals via photometric stereo [15] in real time. The diffuse illumination equation can be written as,

$$I_o = \rho N.L \qquad (1)$$

for a lambertian surface, where I_o is the observed intensity, ρ is the surface albedo, N is the surface normal vector and L is the normalized light direction vector.

After taking the pseudo-inverse of L, we can solve for N and ρ.

$$N' = L^{-1} I_o \qquad (2)$$

and $\rho = \|N'\|$, $N = \dfrac{N'}{\rho}$ (3)

Theoretically this would require only a minimum of 3 different lighting conditions. More input lighting conditions would reduce the noise at the cost of acquisition speed. We have experimented with up to 8 lights and determined 4 lights are a suitable tradeoff to generate acceptable results.

Additionally, even though the color and IR cameras share the same line-of-sight through the beam splitter, they may not be in perfect alignment. We therefore perform an one-time camera registration step to compute a homographic warp from the visible spectrum image frame to the IR image frame.

To relight the visible spectrum image, we introduce a synthetic light source, whose position is currently controlled by the user. For every pixel in the visible image that we have acquired, we evaluate a simple diffuse shading model [16] added to the original value as in:

$$I_d = I_v (k_v \cos(\theta) + k_o) \qquad (4)$$

In equation (4) θ is the angle between the computed surface normal and the direction of the synthetic light source, k_v is a scale constant that controls the amount of contribution from the virtual light source, k_o is an offset scale for the original image, I_v is the observed color at the current pixel and I_d is the resultant diffuse shading result. We replace the visibility channel (in HSV color space) of the visible image by this diffuse shading result, effectively brightening areas of the scene that are pointing towards our synthetic and virtual light source. All these computations, consisting of the alignment warp, estimation of surface normals, and relighting are performed on a Nvidia Quadro FX 4800 GPU in real-time. This relighting computation is a good fit to our application because it directly uses the surface normals without a prior, complex surface reconstruction and also allows the amount of relighting and direction of relighting can be controlled interactively.

5. RESULTS

The original images (figure 3) were captured in low light conditions where some or most of the face is in shadow. This is apparent in the contrast adjusted results (b) where even though one side of the face is reasonably well lit, the other side is still relatively dark.

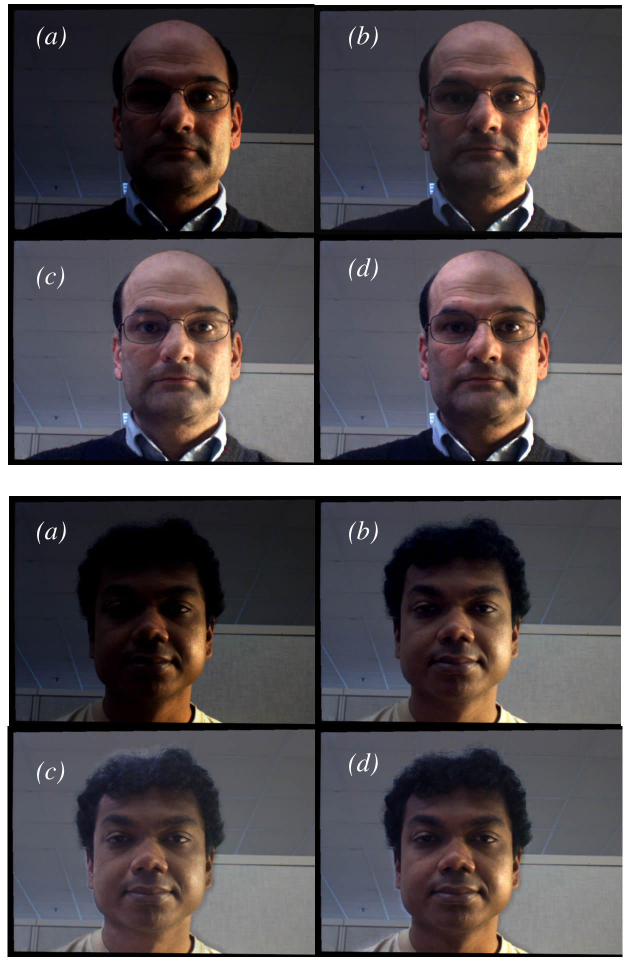

Figure 3 : *Relighting results and comparison of techniques (a) original image (b) contrast adjusted image (c) previous method* [14] *(d) our method*

This is corrected by our algorithm (d) by placing a virtual light on the shadowed side of the face. We also compare our results to the method implemented in [14], where an image set under 8 input infrared lighting conditions was used to generate an evenly-lit facial image. This corrects for the shadowed region but leads to a rather flat looking face relighting result. Since our method uses pixel-based surface normals it generates a relighting result that conforms to facial geometry and thereby appears more realistically lit. Additionally our photometric stereo based method uses just 4 input lighting conditions while avoiding artifacts introduced by ratio imaging.

6. LIMITATIONS AND FUTURE WORK

In the future we would also like to look at two main issues. First, our system currently allows the user to interactively control virtual lighting direction and intensity with a mouse. However, in an actual videoconferencing deployment, these parameters should be set by default giving the most pleasing lighting conditions available. We plan to use measure the entropy of the resultant image [17] to automatically determine these settings. Secondly, our current system does not make any special provision to detect harsh self shadowing or completely saturated regions. We wish to treat these as special cases and use information from similar facial regions to transfer the reflectance information.

7. CONCLUSION

We have prototyped our approach and demonstrated that relighting based on surface normals derived from infrared LEDs results in improved results for real-time video conferencing. Our technique is simple, fast and non-intrusive to the user and addresses the problem of imperfect lighting in casual video conference settings.

8. REFERENCES

[1] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face" in *SIGGRAPH 2000*, 2000.

[2] T. Hawkins, J. Cohen, and P. Debevec, "A photometric approach to digitizing cultural artifacts" in *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*. Glyfada, Greece: ACM, 2001.

[3] P. Einarsson, C.-F. Chabert, A. Jones, B. Lamond, W.-C. Ma, S. Sylwan, T. Hawkins, and P. Debevec, "Relighting human locomotion with flowed reflectance fields" in *SIGGRAPH 2006 Sketches*. Boston, Massachusetts, 2006.

[4] M. Fuchs, V. Blanz, H. Lensch, and H.-P. Seidel, "Reflectance from Images: A Model-Based Approach for Human Faces" *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, pp. 296-305, 2005.

[5] T. Malzbender, D. G. B. Wilburn, and B. Ambrisco, "Surface Enhancement using Real-time Photometric Stereo and Reflectance Transformation" presented at Eurographics Symposium on Rendering 2006, Nicosia, Cyprus, 2006.

[6] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "Illumination-based image synthesis: creating novel images of human faces under differing pose and lighting" presented at Multi-View Modeling and Analysis of Visual Scenes, 1999. (MVIEW '99) Proceedings. IEEE Workshop on, 1999.

[7] L. Zhang, S. Wang, and D. Samaras, "Face Synthesis and Recognition from a Single Image under Arbitrary Unknown Lighting Using a Spherical Harmonic Basis Morphable Model" in *CVPR '05*: IEEE Computer Society, 2005.

[8] Y. Wang, Z. Liu, G. Hua, Z. Wen, Z. Zhang, and S. D., "Face Re-Lighting from a Single Image under Harsh Lighting Conditions" in *CVPR '07*: IEEE Computer Society, 2007, pp. 1-8.

[9] M. Dimitrijevic, S. Ilic, and P. Fua, "Accurate face models from uncalibrated and ill-lit video sequences" presented at CVPR 2004, 2004.

[10] Z. Wen, Z. Liu, and T. S. Huang, "Face relighting with radiance environment maps" in *CVPR 2003*, vol. 2: IEEE Computer Society, 2003, pp. II-158-65 vol.2.

[11] A. Basso, H. P. Graf, D. Gibbon, E. Cosatto, and S. Liu, "Virtual Light: Digitally-generated Lighting for Video Conferencing Applications" presented at ICIP 2001, 2001.

[12] C. Fredembach and S. Süsstrunk, "Coloring the near infra-red" in *IS&T/SID 16th Color Imaging Conference*. Portland, USA: , 2008, pp. 176-182.

[13] D. Krishnan and R. Fergus, "Dark flash photography" presented at SIGGRAPH 2009, 2009.

[14] O. Wang, J. Davis, E. Chuang, I. Rickard, K. de Mesa, and C. Dave, "Video Relighting Using Infrared Illumination" *Computer Graphics Forum (Proceedings Eurographics 2008)*, vol. 27, 2008.

[15] R. J. Woodham, "Photometric method for determining surface orientation from multiple images" in *Shape from shading*: MIT Press, 1989, pp. 513-531.

[16] J. D. Foley and A. V. Dam, *Fundamentals of interactive computer graphics*: Addison-Wesley Longman Publishing Co., Inc., 1982.

[17] T. Malzbender and E. Ordentlich, "Maximum entropy lighting for physical objects" *Proceedings of SPIE, the International Society for Optical Engineering*, vol. 6060, 2006.