Efficient Multispectral Facial Capture with Monochrome Cameras

Chloe LeGendre
USC Institute for Creative Technologies

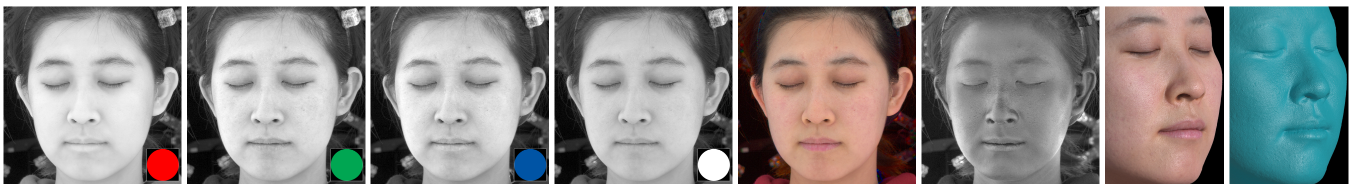
Kalle Bladin
USC Institute for Creative Technologies

Bipin Kishore
USC Institute for Creative Technologies

Xinglei Ren
USC Institute for Creative Technologies

Xueming Yu
Google

Paul Debevec
USC Institute for Creative Technologies



a. red LED b. green LED c. blue LED d. white LED e. diffuse reflection f. specular reflection g. 3D rendering h. 3D geometry

Figure 1: a,b,c,d: Monochrome photographs of a subject lit by red, green, blue, and white LEDs; e: Colorized diffuse reflection image produced by mixing diffuse components of a,b,c, and d; f: Monochrome polarization difference image showing specular reflections; g: Full-color rendering of the subject; h: Geometry rendering of the subject without diffuse albedo. Renderings are produced using only monochrome images.

ABSTRACT

We propose a variant to polarized gradient illumination facial scanning which uses *monochrome* instead of color cameras to achieve more efficient and higher-resolution results. In typical polarized gradient facial scanning, sub-millimeter geometric detail is acquired by photographing the subject in eight or more polarized spherical gradient lighting conditions made with white LEDs, and RGB cameras are used to acquire color texture maps of the subject's appearance. In our approach, we replace the color cameras and white LEDs with monochrome cameras and multispectral, colored LEDs, leveraging that color images can be formed from successive monochrome images recorded under different illumination colors. While a naive extension of the scanning process to this setup would require multiplying the number of images by number of color channels, we show that the surface detail maps can be estimated directly from monochrome imagery, so that only an additional *n* photographs are required, where *n* is the number of added spectral channels. We also introduce a new *multispectral optical flow* approach to align images across spectral channels in the presence of slight subject motion. Lastly, for the case where a capture system's white light sources are polarized and its multispectral colored LEDs are not, we introduce the technique of *multispectral polarization promotion*, where we estimate the cross- and parallel-*polarized* monochrome images for each spectral channel from their corresponding images under a full sphere of even, *unpolarized* illumination. We demonstrate that this technique allows us to efficiently acquire a full color (or even multispectral) facial scan using monochrome cameras, unpolarized multispectral colored LEDs, and polarized white LEDs.

CCS CONCEPTS

• **Computing methodologies → Computational photography**;

KEYWORDS

facial scanning, digital humans, multispectral imaging

ACM Reference format:
Chloe LeGendre, Kalle Bladin, Bipin Kishore, Xinglei Ren, Xueming Yu, and Paul Debevec. 2018. Efficient Multispectral Facial Capture with Monochrome Cameras. In *Proceedings of SIGGRAPH '18 Posters, Vancouver, BC, Canada, August 12-16, 2018,* 2 pages.
https://doi.org/10.1145/3230744.3230778

1 INTRODUCTION

Creating high-quality digital human characters, particularly those based on the likeness of real people, is a long-standing goal in computer graphics, with applications in films, video games, simulations, and virtual reality. Computational imaging and illumination systems have been developed to faithfully capture a subject's facial shape and appearance to produce highly photo-realistic renderings of the subject's digital double. Ma et al. [2007] introduced one such system based on the technique of *shape from shading*, using a series of eight polarized spherical, linear gradient lighting patterns designed to produce surface normal estimates. These patterns were produced using a Light Stage, a spherical lighting rig comprised of broad-spectrum, white LEDs.

SIGGRAPH '18 Posters, August 12-16, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5817-0/18/08.
https://doi.org/10.1145/3230744.3230778

Since colored texture maps are required for rendering a photo-realistic digital double, most facial scanning systems use color cameras and white light sources. In our approach, however, we generate color images by sequentially illuminating a subject with at least three differently colored light sources, capturing images with a *monochrome* camera with a broad spectral response. Photography with monochrome cameras offers a few theoretical advantages. First, most color cameras use RGB filters placed in front of the imaging sensor, which may absorb more than two thirds of the total incident light per pixel. Additionally, these color filters are typically arranged in a Bayer pattern, and the full resolution color image is producing via up-sampling or demosaicing. In contrast, monochrome cameras allow photographers to use comparatively less incident light and produce images of a higher true resolution. In this work, our goal is to gain these benefits of monochrome imaging for high-resolution facial scanning, without sacrificing the color information required for rendering.

2 METHOD

In the naive approach, facial scanning with monochrome cameras and colored LEDs would require three times the original number of images (eight or more) for RGB output, or, in the multispectral case, *n* times as many for output with *n* spectral channels. Such a large number of images would be impractical to acquire for live subjects, where motion should be minimized. However, color images are not required for each component of a facial scanning pipeline:

Multiview Stereo: Initially, a low-resolution 3D reconstruction of the face is generated using passive multiview stereo, as in Ghosh et al. [2011]. Multiview stereo approaches do not require RGB images; they even operate more efficiently when using only intensity information (one third of the input data).

Specular Normals: Ma et al. [2007] inferred surface normals using polarized gradient illumination patterns, for surfaces that primarily reflected light diffusely (Lambertian) and specularly. They introduced "specular normal maps," showing that polarization difference imaging combined with gradient illumination conditions could yield geometry with resolution comparable to that achieved with laser scanning. Since skin is a dielectric, light reflected specularly from the skin is mostly of the same spectrum as the incident illumination. Accordingly, the specular reflection image of a face produced via polarization difference imaging is largely "colorless," so monochrome imaging can recover the "specular normals" used to emboss a low-resolution mesh with pore-level details.

Diffuse Reflectance: For rendering, artists require a colored texture map of the subject's diffuse reflectance, approximated by color images of the subject lit by a full sphere of cross-polarized white light. If a Light Stage included *polarized colored LEDs* in the same polarization arrangement as the white LEDs, then we could capture *n* cross-polarized images of the subject under a full even sphere of illumination for each of the *n* available spectral channels, generating the subject's multispectral diffuse texture map. For monochrome scanning, these *n* photographs are the only additions.

We also introduce enhancements to a monochrome facial scanning pipeline, described briefly here. For the mathematical details of the following techniques, please see the supplemental materials.

Multispectral Polarization Promotion: Polarizing all the colored LEDs of a lighting rig adds complexity and causes over half of the emitted light to be absorbed. Instead, we can *hallucinate* cross-polarized images for each additional spectral channel from *unpolarized* images. We compute the per-pixel amount of light reflected specularly relative to the quantity of incident light for a given lighting condition using the polarized white LEDs and polarization difference imaging. Since skin is a dielectric, the proportion of incident light reflected specularly should mostly not depend on the spectrum of the incident illumination. Therefore, the specular reflection image is approximately consistent across the other spectral channels, up to a scale factor accounting for the camera's different sensitivity to each spectral channel or differing LED intensities. We can therefore produce a cross-polarized image for each spectral channel by subtracting a scaled per-pixel specular reflection image from a captured photograph of the subject under an unpolarized lighting condition for each spectrum.

Optical Flow: Temporal alignment across photographs is required for all computations using more than one image. For monochrome facial scanning, we need to flow not only across gradient illumination conditions but also across spectral channels, both to account for potential movement between frames and to correct for chromatic aberrations. To flow from an image of a subject illuminated by one spectrum to that of a different spectrum, we extend the technique of "complementary flow" of Wilson et al. [2010] to the multispectral domain, finding pairs of images for different spectral channels that sum to a linear combination of already aligned images.

3 RESULTS

We demonstrate that a full-color, high-resolution facial scan can be achieved using monochrome cameras and multispectral LEDs, with only *n* added images at scan-time, without polarizing any colored LEDs. We captured a scan of a female subject using monochrome Ximea machine vision cameras and unpolarized red, green, and blue LEDs, plus white LEDs polarized in the pattern of Ghosh et al. [2011]. Fig. 1e shows a colorized diffuse reflection image computed using multispectral polarization promotion and metameric reflectance matching to convert multispectral images to RGB [LeGendre et al. 2016]. Fig. 1g shows a full-color rendering of the subject, with geometry in Fig. 1h. In the supplemental materials, we compare the rendering to a photograph of the subject and show colorized images for some of the gradient illumination conditions. We also show insets of the surface normal and texture maps.

REFERENCES

Abhijeet Ghosh, Graham Fyffe, Borom Tunwattanapong, Jay Busch, Xueming Yu, and Paul Debevec. 2011. Multiview face capture using polarized spherical gradient illumination. *ACM Transactions on Graphics (TOG)* 30, 6 (2011), 129.

Chloe LeGendre, Xueming Yu, and Paul Debevec. 2016. Efficient Multispectral Reflectance Function Capture for Image-Based Relighting. In *Color and Imaging Conference*, Vol. 2016. Society for Imaging Science and Technology, 47–58.

Wan-Chun Ma, Tim Hawkins, Pieter Peers, Charles-Felix Chabert, Malte Weiss, and Paul Debevec. 2007. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*. Eurographics Association, 183–194.

Cyrus A Wilson, Abhijeet Ghosh, Pieter Peers, Jen-Yuan Chiang, Jay Busch, and Paul Debevec. 2010. Temporal upsampling of performance geometry using photometric alignment. *ACM Transactions on Graphics (TOG)* 29, 2 (2010), 17.