Multi-view Facial Capture using Binary Spherical Gradient Illumination

Alexander Lattas
Imperial College London

Mingqian Wang
Imperial College London

Stefanos Zafeiriou
Imperial College London

Abhijeet Ghosh
Imperial College London

(a) spec. norm.: pol. (b) face geometry: pol. (c) face rendering: pol. (d) spec. norm.: bin. (e) face geometry: bin. (f) face rendering: bin.

Figure 1: Comparison of a face acquired using polarized spherical gradient illumination (a, b, c) versus our proposed method using binary spherical gradient illumination (d, e, f). The specular normals acquired with binary gradients (d) exhibit sharper skin mesostructure details, while the reflectance separation enables realistic rendering of facial appearance (f).

CCS CONCEPTS

• **Computing methodologies → 3D imaging**.

KEYWORDS

reflectance separation, face capture

ACM Reference Format:

Alexander Lattas, Mingqian Wang, Stefanos Zafeiriou, and Abhijeet Ghosh. 2019. Multi-view Facial Capture using Binary Spherical Gradient Illumination. In *Proceedings of SIGGRAPH '19 Posters.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3306214.3338611

1 INTRODUCTION

High resolution facial capture has received significant attention in computer graphics due to its application in the creation of photorealistic digital humans for various applications ranging from film and VFX to games and VR. Here, the state of the art method for high quality acquisition of facial geometry and reflectance employs polarized spherical gradient illumination [Ghosh et al. 2011; Ma et al. 2007]. The technique has had a significant impact in facial capture for film VFX, recently receiving a Technical Achievement award from the Academy of Motion Picture Arts and Sciences [Aca 2019]. However, the method imposes a few constraints due to the employment of polarized illumination, and requires the camera viewpoints to be located close to the equator of the LED sphere for appropriate diffuse-specular separation for multiview capture [Ghosh et al. 2011]. The employment of polarization for reflectance separation also reduces the amount of light available for exposures and requires double the number of photographs (in cross and parallel polarization states), increasing the capture time and the number of photographs required for each face scan.

In this work, we adapt our recently proposed diffuse-specular separation technique using binary spherical gradient illumination [Kampouris et al. 2018] for multiview face capture. Instead of relying on polarized illumination, diffuse-specular separation using binary spherical gradients relies on color-space separation of reflectance (assuming dichromatic reflectance of a dielectric material such as skin). Besides requiring acquisition of fewer images for facial capture (with higher light efficiency) than polarized spherical gradients, another advantage of the method is that it can be employed with LEDs that can only switch between a binary on-off state and does not require modulation of intensities to create gray levels. As can be seen in Figure 1, the high resolution facial normal map acquired using binary gradients exhibits sharper skin mesostructure details (obtained from specular reflectance) compared to polarized spherical gradients, while achieving high-quality reflectance separation for realistic rendering of skin appearance. In the following, we discuss some modifications to the processing of data acquired with binary spherical gradient illumination proposed by Kampouris et al. [2018] that we found to be useful for multiview face capture.

2 PROCESSING BINARY GRADIENTS

2.1 Diffuse Normals

The data acquired using binary spherical gradient illumination directly computes a mixed photometric normal (due to mix of diffuse and specular response under unpolarized illumination). Hence, Kampouris et al. first proposed employing rotation of the RGB colors into *suv* space [Mallick et al. 2005] and computing a pure diffuse normal using the magnitude of the chroma components (u, v).

SIGGRAPH '19 Posters, July 28 - August 01, 2019, Los Angeles, CA, USA

ACM ISBN 978-1-4503-6314-3/19/07.
https://doi.org/10.1145/3306214.3338611

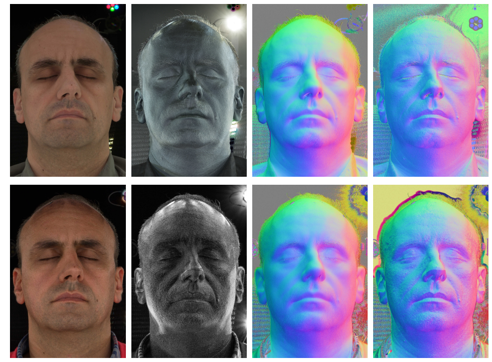
(a) mixed norm. (b) diff. norm.: suv (c) diff. norm.: $R - B$

Figure 2: Comparison of mixed normals (a) obtained directly using binary spherical gradient illumination vs. pure diffuse normals obtained using *suv* projection (b) and our proposed spectral differencing (c).

Instead, we observe that under white illumination, the different color channels of measured skin reflectance have the same amount of specular reflectance mixed with a spectrally varying amount of diffuse reflectance. Hence, we can simply compute a pure diffuse normal using the difference of the brightest and darkest color channel of reflectance, which for human skin are the red and blue channels, respectively. This spectral differencing step completely removes the specular signal, isolating a pure diffuse component. Figure 2 shows that our simple $red - blue$ difference based diffuse normal exhibits identical qualitative softness and blur exhibited by the diffuse normal computed using suv-space projection. The proposed spectral differencing method is much simpler and faster to compute, which is why we adopt it in our face capture pipeline.

2.2 Specular Albedo and Normals

[Kampouris et al. 2018] have shown that a binary gradient and its complement pair can be analyzed for change in color saturation which provides a closed-form solution for estimating the specular albedo. They compute three estimates, one each from the X, Y, and Z gradient-complement pairs and finally compute their median as the final estimate of specular albedo. Instead, we found that we obtain a higher quality estimate of specular albedo by simply averaging the estimate obtained from the local X and Y gradient-complement pairs, ignoring the estimate from the local Z gradient-complement pair. Here, local refers to gradients and complements aligned with respect to the camera viewing direction. However, given the multview capture setup (we employ 9 cameras in our setup), the acquired binary gradients and complements are initially axis aligned with respect to the global coordinates of the LED sphere. Hence, before we can compute diffuse-specular separation of the albedo, we need to computationally rotate the binary gradients and complements into the local coordinate frame of each camera and then employ the local X and Y pairs accordingly to estimate the specular albedo. Note that this rotation of the acquired gradients/complements can be done afterwards as together they form a steerable basis. Finally, the specular normal computation with binary gradients suffers from low-frequency bias, which Kampouris et al. correct by adding a high-pass filter of the specular normals to the diffuse normals. This suffers from slightly noisy results in a multi-view setting. We find that higher quality results are produced by rotating the diffuse normal with a rotation matrix R such that it is aligned with the $+Z$ axis, adding the x and y components of the high-pass to the diffuse normal, and then rotating it back with R^{-1}.



(a) diff. alb. (b) spec. alb. (c) diff. norm. (d) spec. norm.

Figure 3: Comparison of reflectance and normal maps acquired using polarized spherical gradients (top-row), and binary spherical gradients (bottom-row).

3 RESULTS

Figure 3 shows a comparison of the estimated diffuse and specular albedo and normals using polarized spherical gradient illumination versus those estimated using binary spherical gradient illumination. There are qualitative similarities as well as some noticeable differences in the estimated reflectance and normal maps using these two techniques. Interestingly, the specular normal estimated using binary spherical gradient illumination is sharper than that estimated using polarized gradients. Part of the reason here is that the polarized capture requires double the number of photographs and is hence more susceptible to minor subject motion during capture. While the diffuse-specular separation of albedo is of slightly higher quality when employing polarized illumination, the quality of reflectance separation with binary spherical gradient illumination is of sufficiently high quality for enabling realistic rendering of the face scan. Note that the diffuse albedos acquired by the two methods exhibit slightly different color tones due to different types of LEDs used. We correct this by color transformation to sRGB.

ACKNOWLEDGMENTS

The work is supported by EPSRC Early Career Fellowship to Abhijeet Ghosh (EP/N006259/1), and a Google Faculty Fellowship and EPSRC Fellowship DEFORM to Stefanos Zafeiriou (EP/S010203/1).

REFERENCES

2019. SCIENTIFIC & TECHNICAL AWARDS. https://www.oscars.org/sci-tech/ceremonies/2019

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

Christos Kampouris, Stefanos Zafeiriou, and Abhijeet Ghosh. 2018. Diffuse-specular separation using binary spherical gradient illumination. In *Proceedings of the 2018 Eurographics Symposium on Rendering: Experimental Ideas & Implementations (EGSR'18), Karlsruhe, Germany*. 1–4.

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.

Satya P. Mallick, Todd E. Zickler, David J. Kriegman, and Peter N. Belhumeur. 2005. Beyond Lambert: Reconstructing Specular Surfaces Using Color. In *CVPR*.