High Quality Facial Capture using Binary Spherical Gradient Illumination

Alexandros Lattas, Yuan Zhu, Abhijeet Ghosh
{alexandros.lattas17,yuan.zhu18,ghosh}@imperial.ac.uk

Department of Computing,
Imperial College London

1 Acquisition

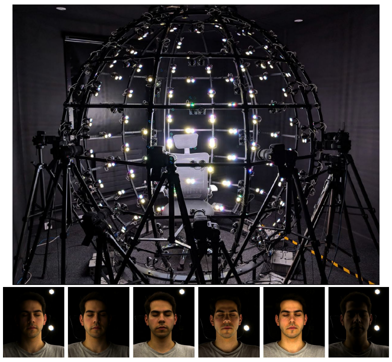
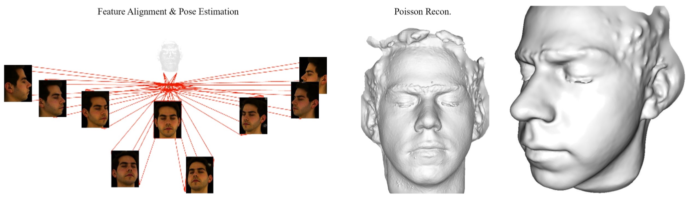


Figure 1: Acquisition setup for multiview facial capture (top) using binary spherical gradient illumination patterns (bottom).

We present a high quality facial capture system that employs unpolarized binary spherical gradient illumination. Our capture setup consists of the Imperial Multispectral LED sphere with 168 LED lighting nodes and a multiview facial capture setup consisting of 9 DSLR cameras (Canon 800D) (Figure 1, top). For acquisition, we employ only the white LEDs on the LED sphere (Philips Color Kinetics iW MR gen3) to illuminate a subject with 6 binary spherical gradient illumination conditions (along principal X, Y and Z directions and their complements) as proposed by Kamprouris et al. [1] (Figure 1, bottom). We rapidly acquire a face under these 6 binary illumination conditions from our 9 cameras operating in burst mode to minimize any subject motion during capture. The data acquired is then employed for both multiview facial geometry reconstruction and for high quality reflectance separation and estimation of separated diffuse and specular photometric normals for rendering.



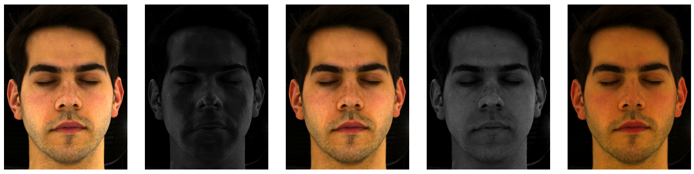
(a) (b)

Figure 2: (a) Acquired facial base geometry using COLMAP. (b) Smoothed base mesh.

Base Geometry Simply adding any two binary gradient and complement pairs results in an observation under uniform spherical illumination condition as seen in Figure 3 (a). While this image is an observation of the mixed (diffuse+specular) albedo, for faces this image is diffuse dominated and can be directly given as input to a multiview stereo reconstruction pipeline for facial geometry reconstruction. We employ the state of the art open source structure-from-motion (SfM) software COLMAP [2] for creating a facial base geometry using the mixed albedo as input (acquired from 9 viewpoints), and apply further smoothing after the initial Poisson reconstruction step to obtain the final the base mesh (see Figure 2).

2 Reflectance Separation

We now present our main contribution of this work which is further improvement in the quality of diffuse-specular separation of the mixed albedo acquired using binary spherical gradient illumination. Our baseline here is the diffuse-specular separation obtain by Kampouris et al. [1] using their proposed analytic solution obtained from HSV color tranformation analysis (see Figure 3 b, c). As can be seen, the quality of separation is quite good at the sides of the face but suffers from poor signal-to-noise (SNR) of the separated specular component near the frontal areas of the face. This results in sub-optimal separation and some specular pollution near the frontal areas of the face in the separated diffuse albedo.



(a) (b) (c) (d) (e)

Figure 3: Mixed albedo (a) obtained from sum of a binary gradient and its complement. (b, c) Specular-diffuse separation of mixed albedo using analytic solution based on HSV color transformation as proposed by Kampouris et al. [1]. (d, e) Improved specular-diffuse separation using our proposed linear system solution under binary spherical gradients.

Instead, we employ a linear system to solve for the specular component based on the observation that under the binary gradient and complement pairs, the brightly lit condition consists of diffuse+specular while the darker condition consists of pure diffuse component. Without loss of generality, assuming a surface appears brighter under gradient G and darker under complement C, we can formulate the linear system solution as follows:

$$\begin{bmatrix} G \\ C \end{bmatrix} = \begin{bmatrix} N_d & 1 \\ (1-N_d) & 0 \end{bmatrix} \begin{bmatrix} D \\ S \end{bmatrix} \tag{1}$$

Here, D is the diffuse component and S the specular component of the signal in a color channel, and N_d is the corresponding component of the diffuse normal (shifted to $[0,1]$) for that principal direction. We solve this linear system for the green channel since it has the best SNR. The diffuse normal N_d is obtained using SUV transformation of the acquired data and employing only the chroma components U, V to compute the photometric normal [1]. We separately solve Equation 1 twice, once each for the X and Y gradient-complement pairs, to obtain two different estimates of D and S. Assuming monochromatic specular reflection, we just compute the final estimate of specular reflctance as the mean of the two solutions for S (Figure 3, d). Finally, we simply subtract the monochromatic specular component S from the mixed albedo to obtain the separated diffuse albedo (Figure 3, e). As can be seen, we obtain a much higher quality of reflectance separation when using the linear system solution.

3 Results

Figure 4: Renderings of two faces acquired using our method.

We present rendering results for faces of a dark-skin subject and a light-skin subject acquired using our facial capture process (see Figure 4). Here, the facial geometry is rendered using the separated diffuse and specular reflectance (albedo) described in Section 3, in conjunction with high quality diffuse and specular normals obtained using binary spherical gradients [1]. As can be seen, we achieve very high quality photorealistic facial renderings with the acquired data.

[1] Christos Kampouris, Stefanos Zafeiriou, and Abhijeet Ghosh. Diffuse-specular separation using binary spherical gradient illumination. In *EGSR (EI&I)*, pages 1–10, 2018.

[2] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.