Near-Instant Capture of High-Resolution Facial Geometry and Reflectance

Paul Graham* Graham Fyffe* Borom Tonwattanapong* Abhijeet Ghosh† Paul Debevec*

*USC Institute for Creative Technologies †Imperial College London

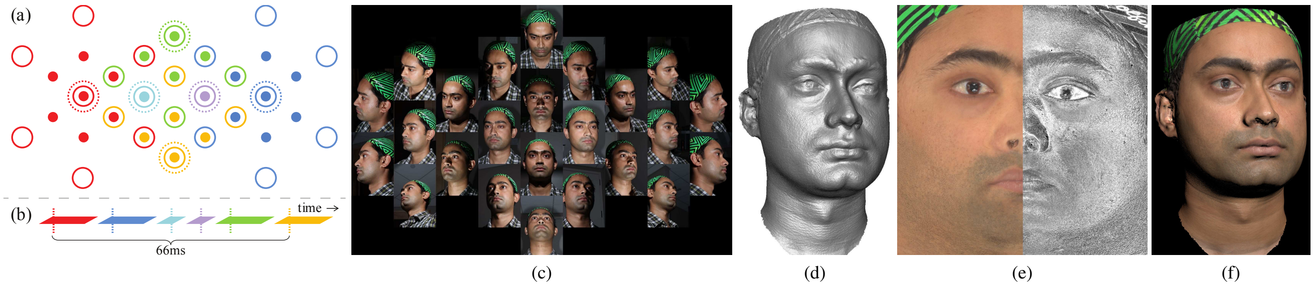

Figure 1: *(a) Schematic showing the location of the flashes (dotted circles), cameras (solid circles), and associated specular highlight half-angles (filled dots; color coded for illustration). (b) The firing sequence for the flashes (dotted lines) and camera exposures (solid strips). (c) A set of images shot using the firing sequence. (d) Final geometry. (e) Diffuse / specular maps. (f) Rendering under novel illumination.*

Overview

Modeling realistic human characters is frequently done using 3D recordings of the shape and appearance of real people, often across a set of different facial expressions to build blendshape facial models. Believable characters that cross the "Uncanny Valley" require high-quality geometry, texture maps, reflectance properties, and surface detail at the level of skin pores and fine wrinkles. Unfortunately, there has not yet been a technique for recording such datasets that is near-instantaneous and low-cost. While some facial capture techniques are instantaneous and inexpensive [Beeler et al. 2010], these do not generally provide lighting-independent texture maps, specular reflectance information, or high-resolution surface normal detail for relighting. In contrast, techniques which use multiple photographs from spherical lighting setups [Ghosh et al. 2011] do capture such reflectance properties, at the expense of longer capture times and complicated custom equipment.

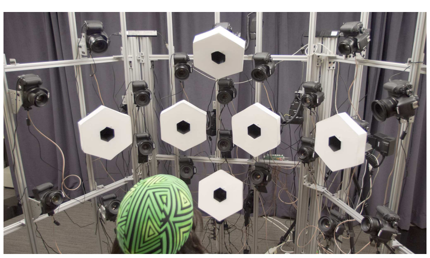

Figure 2: *Our apparatus uses 24 DSLR cameras and 6 ring flashes.*

We propose a near-instant facial capture technique which records high-quality facial geometry and reflectance using commodity hardware. We use a 24-camera DSLR photogrammetry setup similar to common commercial systems and six ring flash units to light the face. Instead of the usual process of firing all the flashes and cameras at once, each flash is fired sequentially with a subset of the cameras, with the exposures packed milliseconds apart for a total capture time of 66ms, which is faster than the blink reflex. This arrangement produces 24 independent specular reflection angles evenly distributed across the face. We use the acquired images to estimate geometry, as well as diffuse color, specular intensity, specular exponent, and surface orientation at each point on the face.

SIGGRAPH 2015 Talks, August 09 – 13, 2015, Los Angeles, CA.
ACM 978-1-4503-3636-9/15/08.
http://dx.doi.org/10.1145/2775280.2792561

Method

Our technique to process the acquired photographs into an accurate 3D model plus maps of diffuse and specular reflectance proceeds as follows: We first leverage passive stereo reconstruction to build a geometric base mesh of the face from the photographs. We then separate the diffuse and specular components of reflectance from the photographs using a novel multiview color-space analysis and further employ photometric analysis to estimate diffuse and specular photometric normals and albedo. We also employ an inverse rendering approach to estimate a per-pixel specular exponent. All reflectance maps are computed in (u,v) texture space for rendering purposes. Finally, we refine the geometric base mesh using the estimated specular photometric normals to produce the final high resolution mesh of the face. Our method produces models suitable for authoring high-quality digital human characters.

Figure 3: *Renderings of additional facial scans.*

References

BEELER, T., BICKEL, B., BEARDSLEY, P., SUMNER, B., AND GROSS, M. 2010. High-quality single-shot capture of facial geometry. *ACM Trans. Graph. 29* (July), 40:1–40:9.

GHOSH, A., FYFFE, G., TUNWATTANAPONG, B., BUSCH, J., YU, X., AND DEBEVEC, P. 2011. Multiview face capture using polarized spherical gradient illumination. *ACM Trans. Graphics (Proc. SIGGRAPH Asia) 30*, 6.