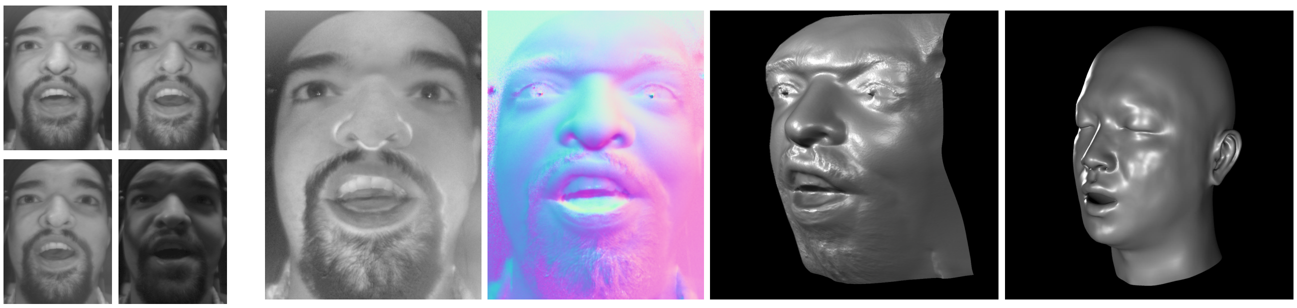
Head-mounted Photometric Stereo for Performance Capture

Andrew Jones Graham Fyffe Xueming Yu Alex Ma Jay Busch Mark Bolas Paul Debevec

University of Southern California Institute for Creative Technologies



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

Figure 1: **(a)** *Lighting conditions (3 directions and ambient only) captured at 120fps* **(b)** *Recovered surface albedo, independent of ambient light* **(c)** *Recovered surface normals* **(d)** *Geometry recovered by integrating surface normals* **(e)** *3D facial rig driven by albedo and normals.*

Head-mounted cameras are an increasingly important tool for capturing an actor's facial performance. Such cameras provide a fixed, unoccluded view of the face. The resulting imagery is useful for observing motion capture dots or as input to existing video analysis techniques. Unfortunately current systems are typically affected by ambient light and generally fail to record subtle 3D shape changes between expressions Artistic interventions is often required to cleanup and map the captured performance onto a virtual character. We have developed a system that augments a head-mounted camera with LED-based photometric stereo. The system allows observation of the face independent of the ambient light and records per-pixel surface normal. Our data can be used to generate dynamic 3D geometry, for facial relighting, or as input of machine learning algorithms to accurately control an animated face.

We use a Point Grey Grasshopper camera mounted 20 cm from the face (Fig. 2). A ring of 12 individually controlled LEDs encircles the lens. The LEDs repeat a sequence of three illumination patterns, followed by an unlit frame to record and subtract ambient light (Fig. 1a); crossed linear polarizers on the lights and lens attenuate specular reflection from the face. The camera and lights run at 120 fps, yielding albedo and normal estimates (Fig. 1b,c) at 30 fps.

Figure 2: *Head-mounted camera and LED light ring*

Traditional photometric stereo [Woodham 1980] uses multiple point lights to recover surface normals by solving linear equations. More recent work (e.g. [Malzbender et al. 2006]) records surface normals in real time. Our system uses three linear gradient intensity ramps across the twelve LEDs, rotated at 0, 120, and 240 degrees. Using the full ring of lights provides more even illumination, reduces shadow artifacts, and emits light from a wider area to increase actor comfort; our video shows results using three individual LEDs for comparison. LED flicker can be greatly reduced by switching lightings patterns at a faster framerate, while using shorter exposure time skip extra frames. To further eliminate distraction from the lights, we built a second lighting rig using invisible infrared LEDs, leveraging the broad spectral sensitivity of the camera. This provided similar results at the expense of some detail in the photometric normals due to increased subsurface scattering.

To correct for subject motion, we compute optical flow between similar illumination patterns to temporally align each set of patterns. Our light sources are close to the face, violating the assumption of distant illumination. To compensate, we compute per-pixel lighting directions relative to a plane approximating the face.

To estimate 3D performance geometry, we integrate the surface normals using Gaussian belief propagation. As expected, the geometry (Fig. 1d) suffers from low-frequency distortions yet reveals expressive performance detail in 3D. The results can be used as input to a machine learning algorithm to drive a facial rig with the performance after an initial training phase. For an initial test, we use an active appearance model [Cootes et al. 2001] to find blendshape weights for a given set of albedo and normal maps. Our initial results are restricted to phonemes (Fig. 1e) and we are working to extend this algorithm to animating the entire face. Initial results show that analysis of normals and albedo provides smoother animation than analysis of albedo alone. We are also working to minimize system weight, which should not be significantly greater than existing rigs as the LEDs weigh just a few grams each.

References

COOTES, T. F., EDWARDS, G. J., AND TAYLOR, C. J. 2001. Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell. 23*, 6, 681–685.

MALZBENDER, T., WILBURN, B., GELB, D., AND AMBRISCO, B. 2006. Surface enhancement using real-time photometric stereo and reflectance transformation. In *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering*, 245–250.

WOODHAM, R. 1980. Photometric method for determining surface orientation from multiple images. *Optical Engineering 19*, 1, 139–144.

SIGGRAPH 2010, Los Angeles, California, July 25 – 29, 2010.
ISBN 978-1-4503-0210-4/10/0007