The Light Stages and Their Applications to Photoreal Digital Actors

Paul Debevec*
University of Southern California Institute for Creative Technologies

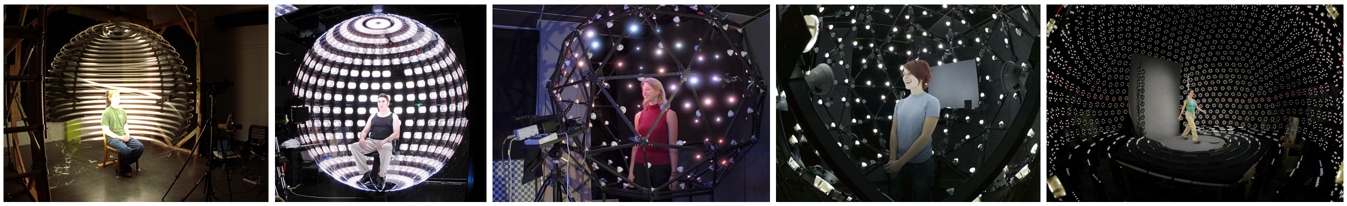

Figure 1: *Left to Right: Light Stage 1's spiraling spotlight records a reflectance field in 60 seconds; Light Stage 2 records actor Alfred Molina for* Spider-Man 2*; Light Stage 3 illuminates an actor with a reproduction of the colorful light of the Grace Cathedral HDRI map; Light Stage 5 uses high-speed photography to record an actor's reflectance with time-multiplexed illumination; Light Stage 6, at 8m in diameter, allows performance relighting for the whole human body.*

Abstract

The Light Stage systems built at UC Berkeley and USC ICT have enabled a variety of facial scanning and reflectance measurement techniques that have been explored in several research papers and used in various commercial applications. This short paper presents the evolutionary history of the Light Stage Systems and some of the techniques and applications they have enabled.

Keywords: 3D scanning, facial animation, computational illumination

1 Introduction

A *Light Stage* is a device which lights a subject with controllable illumination, providing the ability to create or simulate any combination of colors, intensities, and directions of illumination over the range of directions from which light can come. Just as a "sound stage" is a place where actors can be recorded with control over the sound, a Light Stage provides control over the illumination. Numerous forms of light stages have been constructed with different types and quantities of lights and varying mechanical degrees of freedom, and the useful applications of light stages range from image-based relighting to high-resolution facial geometry capture to surface reflectance measurement. This article will describe several light stages built since in the last decade or so and the applications they have enabled.

2 Light Stage 1: Acquiring the Reflectance Field of the Human Face

The first Light Stage was envisioned as a way to apply image-based lighting (IBL) to real-world subjects rather than computer-generated models. In IBL, [Debevec 1998], a panoramic, high dynamic range (HDR) image is projected onto surfaces surrounding the computer-generated object to create a spatially modulated area light which mimics the color, intensity, and directionality of the recorded incident illumination. When the light from this IBL environment is simulated with global illumination, a rendering of the object is produced as if it were illuminated by the recorded real-world lighting. This can add an additional level of realism to the appearance of the object, adding subtle variation to its shading, highlights, and shadows, and is also useful for lighting the object consistently with a scene into which it is being composited, allowing it to appear as if it were really there.

The first concept for a light stage used projectors to display imagery on a concave screen surrounding the subject to illuminate it with the colors, intensities, and directions of the light in a recorded lighting environment. However, the complexity of constructing such a screen and issues with interreflection and contrast of the projected imagery were daunting[1]. Instead, the first Light Stage spiralled a single incandescent spotlight around a person's face so that a digital video camera could record them lit from every direction light could come from, acquiring the 4D *reflectance field* of the actor [Debevec et al. 2000] in less than a minute. Using the linearity of light transport, the color channels of this set of images can be scaled according to intensity of the corresponding color and direction of the light in the scene, and the sum of these scaled images produced simulated images of the subject under novel illumination environments. Although only a 2D process, the results exhibit correct diffuse and specular reflection, subsurface scattering, self-shadowing, and interreflection since they are linear combinations of imagery exhibiting all of these physically accurate effects. This process is called Image-Based Relighting.

The first light stage project also endeavored to create relightable 3D models of people's faces which could be rendered photorealistically from arbitrary viewpoints. To obtain geometry, a structured light video projector was added to the system, and reflectance estimation algorithms were developed to turn the *reflectance function* of each pixel – the color it appears lit from every direction in the sphere – into maps of the face's diffuse color, specular intensity, and surface normals for use in a traditional computer-graphics rendering process. A hybrid algorithm bridging data-driven and modeled reflectance separated diffuse and specular reflections using col-

*e-mail:debevec@ict.usc.edu

[1]Such a system was eventually built using a fisheye lens video projector and a rough specular dome projection surface [Peers et al. 2006]; a dual to this device [Hawkins et al. 2005] used Helmholtz Reciprocity to record high-resolution reflectance fields by spherically imaging how an object scatters light from a scanning laser beam.

orspace analysis, and rendered the two components separately with image-based relighting, with the specular reflection data driving a view-dependent microfacet distribution model.

3 Light Stage 2: Faster Capture and Facial Animation

The first light stage project intended to animate the facial geometry and reflectance data using motion capture data recorded for the SIGGRAPH 99 animation *The Jester*. However, animating such datasets did not occur until the construction of the Light Stage 2 system at the USC Institute for Creative Technologies in the year 2000. Light Stage 2 made recording the reflectance of a face faster by including a set of thirty strobe-lights on a semicircular arm which rotated around the actor in just eight seconds. Multiview relightable images of an actor were recorded in many facial expressions [Hawkins et al. 2004], and blends between these expressions were driven using motion capture data to create an animated digital character. Since multiple scanned expressions were used, the character realistically exhibited facial wrinkling and dynamic reflectance. Light Stage 2 was used by Sony Pictures Imageworks to record actors Alfred Molina and Tobey Maguire for their digital stunt doubles in the movie *Spider-Man 2 (2004)*, Brandon Routh for the film *Superman Returns (2006)*, and by Weta Digital to record actress Naomi Watts for the movie *King Kong (2005)*. Light Stage 2 was also used to scan cultural artifacts [Hawkins et al. 2001], which were rendered directly from a dense array of lighting directions and viewpoints distributed throughout the sphere.

4 Light Stage 3: Lighting Reproduction

With the advent of color-controllable LED lights, it became practical to surround an actor with light sources which could directly reproduce the colors and intensities of the illumination in an image-based lighting environment. Light Stage 3 used a 2m geodesic sphere of 156 iColor MR red-green-blue LED light sources to reproduce the recorded HDRI lighting conditions on real actors [Debevec et al. 2002], whom could then be composited into real scenes with matching illumination. An infrared matting process was used instead of greenscreen to produce more accurate maps without spilling green light onto the actors. [Wenger et al. 2003] presented solutions for accurately reproducing the effect of arbitrary incident illumination spectra (e.g. incandescent, daylight, fluorescent) with RGB LED light sources. A ”Light Stage 4” conceptual rendering of a studio-sized lighting reproduction system – able to control the environmental illumination on multiple actors and small sets – was presented at SIGGRAPH 2002, but has not yet been built. More recently, a Lighting Reproduction system was used by Lola Visual Effects to help actor Armie Hammer portray both of the Winklevoss twins in David Fincher's *The Social Network (2009)*, and to project properly illuminated faces of full-sized actors, such as Bob Hoskins and Nick Frost, onto shorter on-set actors portraying the dwarves in *Snow White and the Huntsman (2012)*.

5 Light Stage 5: Performance Relighting

With the advent of bright white LEDs, it became possible to record an actor's reflectance dynamically using time-multiplexed illumination and high-speed video, allowing their performance to be re-lit in postproduction [Wenger et al. 2005]. This process was realized with Light Stage 5, a set of 156 white LED lights added to the geodesic structure of Light Stage 3. Motion compensation algorithms based on diffusely-lit tracking frames allowed the numerous lighting conditions shot during each traditional frame of video to be aligned to the same time instant and then re-illuminated with image-based relighting. An extension [Jones et al. 2006] added a high-speed structured light video projector to record dynamic geometry of the actor's performance (interleaving reflectance and geometry capture) allowing the actor to be computationally illuminated under spatially-varying light such as dappled shadows or projected imagery. [Peers et al. 2007] showed that useful performance lighting could also be achieved from a static facial reflectance field applied through ratio images to a staticly-lit facial performance. Light Stage 5 could also be used for recording traditional one-light-at-a-time reflectance fields, and was used by Sony Pictures Imageworks to digitize the reflectance of actors Will Smith and Charlize Theron for the movie *Hancock (2008)* and a silicone maquette depicting Brad Pitt as an old man for *The Curious Case of Benjamin Button (2008)*. In the latter film, the digital Benjamin effects won Digital Domain a visual effects Oscar in achieving the first photoreal digital main actor in a motion picture.

6 Light Stage 6: Re-lighting the Whole Body

When space became available to build a light stage able to illuminate the whole human body, it was decided to build a large version of Light Stage 5 rather than Light Stage 4. Light Stage 6 was built with 6,666 LumiLEDs Luxeon V LEDs controlled in groups of six. Its 8m diameter was arrived at by scaling up the 50cm diameter working volume of the 2m Light Stage 5 by a factor of four. To avoid suspending the actor in the air, only the upper 2/3rds of the dome was built and the floor was populated with its own set of Lambertian-distribution light sources. The Relighting Human Motion project [Einarsson et al. 2006] recorded actors walking and running on a rotating treadmill inside the stage at 990 frames per second, under a basis of 33 repeating lighting conditions filmed from a vertical array of three high-speed cameras. From this data, a combination of view interpolation [Chen and Williams 1993], light field rendering [Levoy and Hanrahan 1996], and image-based relighting rendered the actors locomoting through new environments under illumination and camera motion matched to arbitrary scenes. Geometry has also been derived from a surrounding set of video cameras using combined cues from silhouettes, computational stereopsis, and photometric stereo [Vlasic et al. 2009] and made to be temporally coherent and complete in [Li et al. 2012]. Most recently, Light Stage 6 has been augmented with a Laser Scanning System from ICON Imaging Studio to scan the geometry and reflectance of actors such as Michael Caine and Dwayne Johnson for *Journey 2: The Mysterious Island (2012)*.

Figure 2: *Animated digital humans rendered from arbitrary viewpoints and illumination from Light Stage 6 data.*

7 High-Resolution Geometry Capture with Polarized Spherical Gradient Illumination

A desire to digitize facial geometry with sub-millimeter precision motivated a rethinking of the light stage acquisition process in the mid-2000's. Although plaster facial cast laser scanning could achieve such detail, the weight of the casting material distorted the face, it could only be used easily to record neutral expressions, and it yielded no reflectance or surface coloration information. Other desirable features of a new acquisition system would be for it to require less data and to work with consumer digital still cameras rather than expensive motion picture cameras. A solution emerged by leveraging the full sphere of LED light sources in Light Stage 5 to create spherical illumination patterns.

Instead of lighting the face one direction at a time, requiring hundreds of photographs to record reflectance functions, key statistics of the reflectance functions would be measured using computational illumination patterns based on the first four spherical harmonics [Ma et al. 2007]. The total energy of a pixel's reflectance function could be measured simply by lighting the face with spherical light from everywhere, corresponding to the 0th order spherical harmonic. For a diffuse surface, this yields the diffuse color. The centroid of the reflectance of each reflectance function could be measured by lighting the face by gradient patterns derived from the first three 1st order spherical harmonics oriented along the three principal coordinate axes. For a diffuse surface, this yields the surface normal, since a Lambertian reflectance lobe is symmetrically oriented along the normal.

However, skin reflectance is not Lambertian, and consists of both a specular surface reflection and diffuse subsurface scattering. To measure specular and diffuse reflectance characteristics independently, we extended a polarization difference imaging technique previously explored for facial capture in [Debevec et al. 2000]. With a single linearly polarized light, its polarization-preserving specular reflection can be cancelled or transmitted based on whether a linear polarizer on the camera is perpendicular or parallel, respectively, to the polarizer on the light, and the mostly depolarized subsurface reflection will be partially admitted at either orientation. Thus, a cross-polarized image yields only subsurface reflection by blocking the specular reflection, and subtracting this from a parallel-polarized image yields (for the most part) just the specular reflection. Unfortunately, a single light source reveals specular reflectance information for just small range of surface orientations on a face, so we adapted polarization difference imaging the entire incident sphere of illumination by adding linear polarizer filters to each of the 156 Light Stage 5 light sources. These were placed at appropriate angles so that specular reflections from the face would arrive at the frontal camera vertically polarized. Then, a polarizer in front of the camera which flips between horizontal and vertical orientations tunes in and out the specular reflection of the face from the whole sphere of illumination at the same time.

By shooting each of the four gradient conditions both cross- and parallel-polarized, we can use the gradients to measure the specular and subsurface skin reflectance characteristics such as their albedo and surface normal independently. Interestingly, the surface normals from specular reflection exhibit sharper surface detail than the diffuse normals, owing to the fact that the specular light is a reflection directly from the surface of the skin and the subsurface light has diffused on the order of millimeters within the skin before reflecting back. The specular surface normal detail allows sub-millimeter skin-pore-accurate measurements of local surface shape, which when embossed onto a structured light scan of the face yield 3D geometry nearly approaching the resolution of scanned facial casts. Furthermore, the relatively small number of photographs can be shot with standard digital SLR cameras in just a few seconds, allowing natural facial expressions to be captured, and each scan yields properly aligned texture maps of the skin's diffuse and specular albedo. Combined with global illumination rendering and subsurface scattering simulation [Jensen et al. 2001], realistic facial renderings can be generated from such data.

Figure 3: *A rendering of a photoreal digital face from the Digital Emily project [Alexander et al. 2010].*

This facial scanning process was used as the acquisition process in creating one of the first photoreal digital actors in The Digital Emily Project [Alexander et al. 2010] shown at SIGGRAPH 2008. Thirty-three facial scans of actress Emily O'Brien were used by facial animation company Image Metrics to rig an animatable blendshape model of Emily's face, which was driven using their video-based performance capture process. When composited into a live-action scene illuminated with proper image-based lighting, the results were generally accepted as indistinguishable from video of Emily herself. The polarized gradient illumination facial scanning process was also applied on the movie *Avatar (2009)*, with high-resolution scans of being used by Weta Digital to create realistic digital doubles of the principal cast. A digital Sam Worthington, for example, was seen on screen in an extended shot at the end of the film. Through technology licensees Aguru Images and Light Stage, LLC, the gradient illumination facial scanning process contributed to realistic digital human characters in *GI: Joe (2009)*, *Endhiran (2010)*, *TRON: Legacy (2010)*, *X-Men: First Class (2011)*, and *The Avengers (2012)*.

8 Extensions to Gradient Illumination Face Scanning

The basic gradient illumination facial scanning process has been usefully enhanced in several ways. A more expressive model of facial reflectance including a per-region specular model, single scattering, and shallow and deep scattering can be captured to record layered facial reflectance [Ghosh et al. 2008]. Second-order gradients – the next five of the first nine spherical harmonic conditions – allow the estimation of per pixel specular roughness and anisotropy [Ghosh et al. 2009]. A joint photometric alignment process [Wilson et al. 2010] uses complementary gradient conditions, such as forward gradient illumination along $+x$ and a reverse gradient along $-x$, to robustly align reflectance data when the subject exhibits motion; this work technique also facilitates applying the process to video performance capture with a temporally-consistent surface parameterization. Facial Performance Synthesis with Deformation-Driven Polynomial Displacement Maps [Ma et al. 2008] creates animatable facial models with dynamic wrinkle and pore detail from high-resolution sequences of an actor creating various expressions, drivable from just low-resolution motion capture data. And Comprehensive Facial Performance Capture [Fyffe et al. 2011] uses time-multiplexed gradient lighting, heuristic diffuse-specular separation from unpolarized illumination, and multiresolution belief propagation to perform per-frame 3D reconstruction from multiview stereo correspondence and photometric stereo.

9 Multiview Capture and Light Stage X

State-of-the-art high-resolution facial scanning with polarized spherical gradient illumination extends the process to multi-view capture [Ghosh et al. 2011]. The polarization pattern of [Ma et al. 2007] achieves accurate diffuse/specular separation for only a frontal camera position, which required scanning the actor from three different angles to produce an ear-to-ear facial scan. [Ghosh et al. 2011] places a static vertical polarizer on a set of cameras placed around the equator of the light stage, and switches the light stage's illumination from horizontal to vertical polarization to tune the specular reflection in and out, allowing for simultaneous multiview facial capture. This process is implemented by USC ICT's newest Light Stage X facial scanning system, which uses programmable light sources and interchangeable filters to electronically create any color, intensity, direction, and polarization of light across the incident sphere of illumination.

Figure 4: *High-res facial geometry captured in Light Stage X.*

Acknowledgements

We gratefully acknowledge the sponsors which have contributed to this work including the Office of Naval Research, USA RDECOM, the USC Office of the Provost, and Next Media Animation. Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

References

ALEXANDER, O., ROGERS, M., LAMBETH, W., CHIANG, J.-Y., MA, W.-C., WANG, C.-C., AND DEBEVEC, P. 2010. The digital emily project: Achieving a photorealistic digital actor. *Computer Graphics and Applications, IEEE 30*, 4 (july-aug.), 20 –31.

CHEN, S. E., AND WILLIAMS, L. 1993. View interpolation for image synthesis. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, ACM, New York, NY, USA, SIGGRAPH '93, 279–288.

DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. In *Proceedings of ACM SIGGRAPH 2000*, 145–156.

DEBEVEC, P., WENGER, A., TCHOU, C., GARDNER, A., WAESE, J., AND HAWKINS, T. 2002. A lighting reproduction approach to live-action compositing. *ACM Transactions on Graphics 21*, 3 (July), 547–556.

DEBEVEC, P. 1998. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proceedings of ACM SIGGRAPH 98*.

EINARSSON, P., CHABERT, C.-F., JONES, A., MA, W.-C., LAMOND, B., HAWKINS, T., BOLAS, M., SYLWAN, S., AND DEBEVEC, P. 2006. Relighting human locomotion with flowed reflectance fields. In *Rendering Techniques*, 183–194.

FYFFE, G., HAWKINS, T., WATTS, C., MA, W.-C., AND DEBEVEC, P. 2011. Comprehensive Facial Performance Capture. *Computer Graphics Forum 30*, 2, 425–434.

GHOSH, A., HAWKINS, T., PEERS, P., FREDERIKSEN, S., AND DEBEVEC, P. 2008. Practical modeling and acquisition of layered facial reflectance. *ACM Transactions on Graphics 27*, 5 (Dec.), 139:1–139:10.

GHOSH, A., CHEN, T., PEERS, P., WILSON, C. A., AND DEBEVEC, P. 2009. Estimating specular roughness and anisotropy from second order spherical gradient illumination. *Computer Graphics Forum 28*, 4 (June/July), 1161–1170.

GHOSH, A., FYFFE, G., TUNWATTANAPONG, B., BUSCH, J., YU, X., AND DEBEVEC, P. 2011. Multiview face capture using polarized spherical gradient illumination. *ACM Trans. Graph. 30*, 6 (Dec.), 129:1–129:10.

HAWKINS, T., COHEN, J., AND DEBEVEC, P. 2001. A photometric approach to digitizing cultural artifacts. In *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, ACM, New York, NY, USA, VAST '01, 333–342.

HAWKINS, T., WENGER, A., TCHOU, C., GARDNER, A., GÖRANSSON, F., AND DEBEVEC, P. 2004. Animatable facial reflectance fields. In *Rendering Techniques 2004.*, 309–320.

HAWKINS, T., EINARSSON, P., AND DEBEVEC, P. 2005. A dual light stage. In *Rendering Techniques 2005: 16th Eurographics Workshop on Rendering*, 91–98.

JENSEN, H. W., MARSCHNER, S. R., LEVOY, M., AND HANRAHAN, P. 2001. A practical model for subsurface light transport. In *Proceedings of ACM SIGGRAPH 2001*, 511–518.

JONES, A., GARDNER, A., BOLAS, M., MCDOWALL, I., AND DEBEVEC, P. 2006. Simulating spatially varying relighting on a live performance. In *Proc. of European Conference on Visual Media Production (CVMP)*, 127–133.

LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, ACM, New York, NY, USA, SIGGRAPH '96, 31–42.

LI, H., LUO, L., VLASIC, D., PEERS, P., POPOVIĆ, J., PAULY, M., AND RUSINKIEWICZ, S. 2012. Temporally coherent completion of dynamic shapes. *ACM Trans. Graph. 31*, 1 (Feb.), 2:1–2:11.

MA, W.-C., HAWKINS, T., PEERS, P., CHABERT, C.-F., WEISS, M., AND DEBEVEC, P. 2007. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Rendering Techniques 2007: 18th Eurographics Workshop on Rendering*, 183–194.

MA, W.-C., JONES, A., CHIANG, J.-Y., HAWKINS, T., FREDERIKSEN, S., PEERS, P., VUKOVIC, M., OUHYOUNG, M., AND DEBEVEC, P. 2008. Facial performance synthesis using deformation-driven polynomial displacement maps. *ACM TOG (Proc. SIGGRAPH Asia)*.

PEERS, P., HAWKINS, T., AND DEBEVEC, P. 2006. A reflective light stage. Tech. Rep. ICT Technical Report ICT-TR-04.2006, ICT-USC.

PEERS, P., TAMURA, N., MATUSIK, W., AND DEBEVEC, P. 2007. Post-production facial performance relighting using reflectance transfer. *ACM Transactions on Graphics 26*, 3 (July), 52:1–52:10.

VLASIC, D., PEERS, P., BARAN, I., DEBEVEC, P., POPOVIĆ, J., RUSINKIEWICZ, S., AND MATUSIK, W. 2009. Dynamic shape capture using multi-view photometric stereo. *ACM Trans. Graph. 28*, 5 (Dec.), 174:1–174:11.

WENGER, A., HAWKINS, T., AND DEBEVEC, P. 2003. Optimizing color matching in a lighting reproduction system for complex subject and illuminant spectra. In *Eurographics Symposium on Rendering: 14th Eurographics Workshop on Rendering*, 249–259.

WENGER, A., GARDNER, A., TCHOU, C., UNGER, J., HAWKINS, T., AND DEBEVEC, P. 2005. Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM TOG 24*, 3, 756–764.

WILSON, C. A., GHOSH, A., PEERS, P., CHIANG, J.-Y., BUSCH, J., AND DEBEVEC, P. 2010. Temporal upsampling of performance geometry using photometric alignment. *ACM Transactions on Graphics 29*, 2 (Mar.), 17:1–17:11.