Simulating Spatially Varying Lighting on a Live Performance

Andrew Jones Andrew Gardner Mark Bolas† Ian McDowall‡ Paul Debevec

USC Institute for Creative Technologies
University of Southern California†
Fakespace Labs‡

Abstract

We present an image-based technique for relighting dynamic human performances under spatially varying illumination. Our system generates a time-multiplexed LED basis and a geometric model recovered from high-speed structured light patterns. The geometric model is used to scale the intensity of each pixel differently according to its 3D position within the spatially varying illumination volume. This yields a first-order approximation of the correct appearance under the spatially varying illumination. A global illumination process removes indirect illumination from the original lighting basis and simulates spatially varying indirect illumination. We demonstrate this technique for a human performance under several spatially varying lighting environments.

Keywords: rendering, relighting, environmental illumination, image-based rendering, reflectance models

1 Introduction

Light in the real world is both subtle and complex. As a person moves through space, they interact with changing shadows, dappled lighting, and shafts of light. Cinematographers frequently use such spatially varying light as a visual storytelling tool. In the virtual world, it is desirable to be produce these same lighting effects.

Recently, image-based relighting techniques have been presented that allow realistic illumination to be virtually simulated on photographically acquired subjects. As such renderings are derived from real photographs, the realism of the subject's appearance is retained under the novel illumination. However, image-based relighting techniques to date have been limited to simulating only the effects of distant illumination [18] or relighting only static subjects [9, 17, 1]. For image-based relighting to match the generality of either traditional computer graphics lighting or real-world cinematography, we must further develop techniques that are able to synthesize lighting that varies spatially across the subject.

In this paper, we present a new image-based relighting technique which is able to simulate spatially varying (SV) lighting effects on a live-action performance. We use a fixed set of structured illumination patterns and a directional lighting basis to simulate the effects of spatially varying illumination from arbitrary lighting directions. To do this, we derive a geometric model of the dynamic subject. We then relight each pixel in each frame according to its 3D position in a spatially varying light volume. We also employ a novel global illumination process to remove the indirect illumination from the original lighting basis and to approximate the indirect illumination which would result from the spatially varying incident illumination. Using high-speed video and image projection, we achieve an effective capture rate of 24 frames per second suitable for a live-action performance.

2 Related Work

Image-based relighting techniques synthesize novel renderings of a subject using measurements of radiant light under different incident illumination conditions. If one were to measure the radiant light field [8] for each incident ray of light in a scene, one could simulate arbitrary spatially varying conditions directly from this eight-dimensional reflectance field [4]. However, measuring and storing the entire 8D set of lighting samples is prohibitively expensive and researchers typically focus on capturing lower-dimensional subsets of the reflectance field.

A common reduction is to restrict the reflectance field to a 4D subset from a single image viewpoint considering only the subject's response to directional illumination. [4] captures such datasets for human faces using a rotating directional light source. [18] captures similar 4D datasets for a dynamic subject using a high-speed camera and a sphere of time-multiplexed LED light sources. However, neither technique can simulate SV lighting since no such information is present in the lighting basis.

Other work has acquired 6D reflectance fields that capture the effects of SV illumination from a fixed viewpoint by including SV conditions in the lighting basis. To distinguish between incoming rays, [9] and [17] replace the moving light source with a video projector. To capture the 6D dataset efficiently, [9] makes assumptions about the degree of interaction of the indirect illumination in the scene and uses a relatively coarse lighting basis, whereas [17] adaptively detects overlapping regions of indirect illumination to capture a high-resolution lighting basis. Nonetheless, acquiring even a reduced 6D reflectance field requires capturing a large quantity of data and significant online computation, both of which preclude their direct application to capturing a live-action performance.

Given a subset of a reflectance field, it is possible to interpolate or extrapolate additional information. Previous work has combined reflectance field data with geometry from structured light [4], visual hulls [10], and multi-view stereo [6]. However, the goal of these works was to extrapolate additional viewpoint dimensions or resolution. One recent work has dealt with interpolating the lighting basis. Instead of using proxy geometry, [1] solves for lighting flow between projectors. The projectors provide detailed spatial lighting resolution while the sparse angular dimensions are interpolated. As before, the use of adaptive patterns makes live performance capture difficult.

Our method simulates indirect SV illumination using synthetic global illumination simulations. Little previous work exists on how to best integrate global illumination simulations with image-based relighting. In this regard, our method has some similarity to [3] which uses a global illumination simulation to add virtual objects and photographs, but [3] do not relate this process to the general case of image-based relighting.

3 Data Acquisition and Processing

Our setup (Fig. 1) consists of a Luxeon V LED-based lighting apparatus as in [18], an 800×600 resolution Vision Research Phantom 7.1 high-speed camera, and a specially modified 1024×768 pixel high-speed DLP projector [11]. We capture images of the actor under a repeating set of 53 lighting conditions. First the subject is lit by a series of 29 basis lighting directions (top of Fig 3), followed by 24 structured light patterns (bottom of Fig 3). The structured light consist of horizontal and vertical binary patterns that uniquely identify each projector pixel. All components in the system are synchronized at 1500fps, allowing the full lighting sequence to repeat 24 times per second. The high-speed camera stores the video locally in internal memory. After the perfomance is complete, this data is downloaded to a computer for processing.

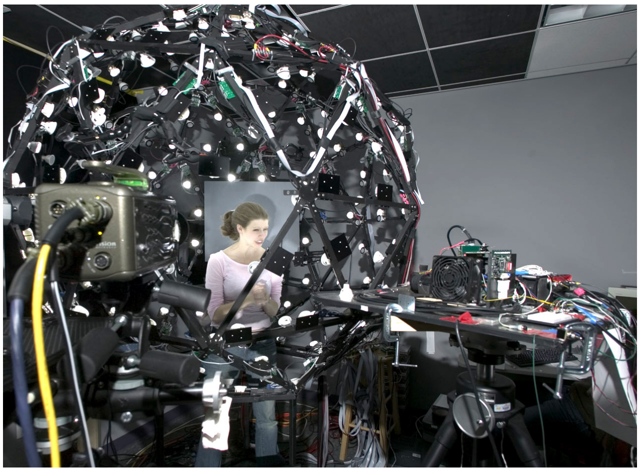
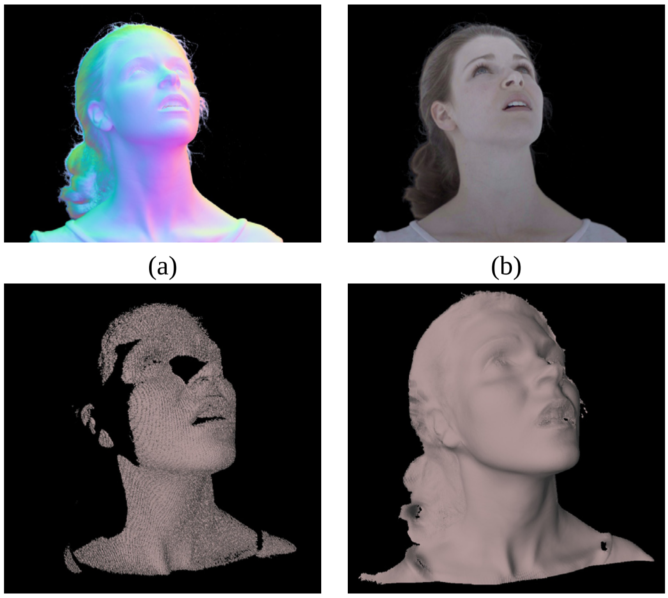


Figure 1: An actor is filmed inside the sphere of LED lights. The high-speed camera is seen on the left and the high-speed DLP projector is on the right.

For each iteration, we recover triangulated geometry from the structured light pattern images for the face regions visible to both the camera and the projector. Subpixel localization of the structured light patterns allows for the recovery of scene geometry shown in Fig. 2 (c). We also perform an initial diffusion pass to fill holes due to occlusion [2].

Using structured light alone does not take full advantage of the 29 basis lighting directions, so we perform photometric stereo on the lighting basis images to generate per-pixel estimates of surface normals (Fig. 2 (a)) and diffuse albedo (Fig. 2 (b)) [18]. The surface normals reveal high-resolution detail not captured by the structured-light system. We could simply integrate the normals to recover geometry, but unfortunately, photometric stereo normals tend to exhibit low-frequency distortion. Instead we use the linear mesh optimization technique outlined by Nehab et al. [13] to combine the high-frequency content of the recovered normals with the low-frequency content of the scanned geometry. The algorithm also conforms any new hole-filling geometry to the known surface normals. The final smooth geometry is based on both the reflectance information and the structured light patterns (Fig. 2 (d)).



(a) (b)

(c) (d)

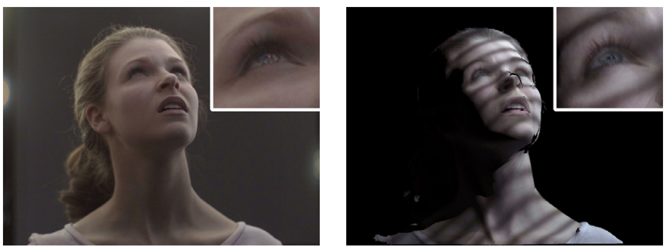
Figure 2: Using photometric stereo we recover (a) color-coded surface normals and (b) surface albedo. These normals are used to optimize and extend the recovered geometry [13], seen (c) before smoothing, (d) after smoothing.

4 Relighting

Using linear combinations of the LED basis images we can relight the subject with distant illumination. However, this cannot produce the lighting effects of an SV lighting environment. One such environment, shown in Figure 4,

Figure 3: Example image sequence repeatedly captured at an interval of 1/24 of a second. (top) 29 basis lighting directions from the LED sources. (bottom) 24 binary structured light patterns from the video projector.

contains strong alternating shadows cast by a set of Venetian blinds. Illuminating the subject with a light probe captured from just one point within this illumination shows accurate subject reflectance (Fig. 4 (a)) but the effect of the SV lighting is lost. Alternately, if the geometric model is texture-mapped with the recovered albedo, we can simulate SV illumination but also lose all complex reflectance properties (Fig. 4 (b)) which is unrealistically diffuse. Our method combines the lighting flexibility of the 3D geometric model with the realism of the image-based reflectance field.



(a) (b)

Figure 4: (a) 4D image-based relighting exhibiting realistic reflectance but not SV light. (b) Global illumination rendering of measured geometry and albedo, exhibiting SV light but inaccurate reflectance.

4.1 Spatially Varying Direct Illumination

We first consider how to apply spatial lighting variation to a single basis image under homogenous LED illumination. The simplest form of SV illumination is to project an image onto the actor from a single LED position, similar to slide projector. If we assume that the visible scene contains no interreflections, generating such lighting becomes a simple per-pixel scaling operation:

$$D_{spatial} = \alpha D_{existing}. \tag{1}$$

$D_{existing}$ represents a photograph of the subject under the existing LED illumination with no interreflections. $D_{spatial}$ represents the subject lit by the equivalent SV light with no interreflections. α represents the relative direct ray intensity at each pixel under SV illumination. We can relight the scene using only Equation 1 by substituting the actual basis image for $D_{existing}$. When the effects of indirect illumination are minimal, the result is a close approximation of SV illumination (Fig. 5 (c)). The geometric model provides the mapping between image coordinates and light rays by indicating where each point in the image lies within the three-dimensional volume of light. We generate α by projectively texture mapping the geometry from the position of the SV light source with the desired SV lighting pattern.

Intuitively, if there are no interreflections, the 4D light transfer matrix contains only diagonal elements and can be measured in a single image. Previously, both [9] and [17] noted that for regions where local interreflections are minimal, large bundles of light rays can be measured simultaneously. A benefit of this approach is that while we have only one physical projector, we can simulate any SV lighting condition from any of the measured 29 LED lighting directions.

Figure 5: Ignoring the effect of indirect illumination, an image can be relit using a per-pixel scale according to Equation 1.

4.2 Correcting for Indirect Illumination

This relighting process above does not correctly model secondary lighting effects resulting from the subject's interreflection. For example, if a SV light source illuminates just a person's shoulder, indirect illumination reflected from the shoulder would also illuminate the underside of the chin. In this case, the one-to-one correspondence between surface locations and light rays no longer holds as there are additional indirect rays from other locations. To account for this, both the original basis image ($L_{existing}$) and relit image ($L_{spatial}$) should be represented as a sum of direct (D) and indirect (I) components:

$$L_{existing} = D_{existing} + I_{existing} \tag{2}$$

$$L_{spatial} = D_{spatial} + I_{spatial}. \tag{3}$$

By combining equations 1 and 3, we express the relit image in terms of the existing direct illumination ($D_{existing}$) and the SV indirect illumination ($I_{spatial}$):

$$L_{spatial} = \alpha D_{existing} + I_{spatial}. \tag{4}$$

Then by using equation 2, we express the relit image in terms of the original lighting basis image:

$$L_{spatial} = \alpha(L_{existing} - I_{existing}) + I_{spatial}. \tag{5}$$

This equation subtracts the existing indirect illumination ($I_{existing}$), then modulates the result by the direct SV rays (α), and finally adds in the desired indirect SV illumination ($I_{spatial}$). The entire process is simple linear combination of basis photographs and indirect illumination (shown in Fig. 6).

Unfortunately, it is not trivial to separate the observed basis image into its direct and indirect components. Even if one measures the full 6D lighting transport [9, 17, 1] it is difficult to distinguish changes in indirect illumination from variations in BRDF. Seuitz et al. [16] propose a solution for recovering the indirect component for Lambertian scenes. Nayer et al. [12] use high-frequency projected patterns to separate direct and indirect illumination. This technique could seperate direct and indirect light from the high-speed projector but this is not sufficient. We need to seperate direct and indirect illumination for all possible LED directions.

4.3 Global Illumination Simulation

As we do not independently observe the direct and indirect illumination, we instead simulate it using our best-fit model. An indirect illumination pass can be computed in most standard global illumination rendering engines either by directly caching bounced irradiance or by comparing two renderings of the textured geometry, one with a simulated indirect bounce and one without. Using a standard Monte-Carlo based global illumination algorithm, we simulate both the existing indirect illumination, caused by the LED lighting, and the indirect illumination caused by the synthetic SV light-source.

The virtual LEDs are positioned to match the known geometry of the apparatus. Each light is represented as a point light source with even illumination accross the subject. For more accurate simulation of the existing lights, one could measure the emitted light field of the LED illuminants as in [5]. For the SV light, we project the desired patterns from the position of each LED, then compute a single indirect bounce using the recovered BRDF. The recovered BRDF is used for both the initital bounce and the secondary reflection towards the camera. Each simulation requires approximately 10 minutes. While this rendering time is significant, it can easily be optimized and parallelized across multiple computers.

In these examples, we use the 3D geometry texture-mapped with the recovered diffuse albedo (Fig. 2 (b)) to produce the indirect illumination images shown in Figure 6. By restricting our simulations to the Lambertian case, we show that even a simple reflectance model produces reasonable results for moderately complex scenes such as facial performances. However, Equation 5 is a general relighting equation independent of the BRDF model or underlying acquisition system. Unlike the relighting framework of [1], our technique would extend naturally to more complex BRDF models such as specular indirect illumination or subsurface scattering by simulating such effects in the Monte-Carlo rendering. Using more advanced BRDF fitting such as [7, 20] we could better predict indirect illumination from missing sample directions.

Original image ($L_{existing}$) | Existing indirect light ($I_{existing}$) | Direct spatially varying rays (α) | Spatially varying indirect light ($I_{spatial}$) | Relit image ($L_{spatial}$)

Figure 6: A basis image is modulated to recreate spatially varying direct and indirect illumination using Equation 5

The primary advantage of our method is that it uses the subject's own image-based reflectance (including subsurface scattering and specularity) for the subject's response to the SV direct illumination. Since people's faces are not extremely concave, and since facial albedo is typically less than 50%, this accounts for the majority of the reflected illumination, which is simulated correctly. Since human skin (and most clothing) is relatively diffuse, the Lambertian underlying model we use accounts for the majority of the reflectance of the remaining indirect illumination.

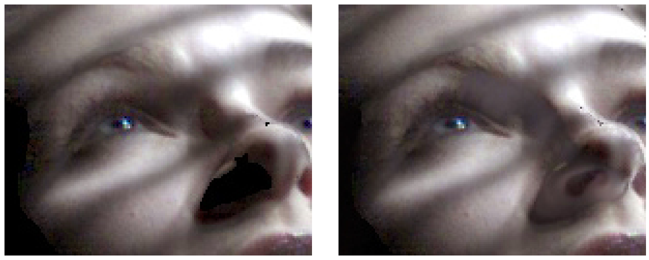
Our technique works best for scenes dominated by direct illumination. For scenes with no indirect illumination, our framework leverages data fully present in the image-based reflectance measurements. For scenes that are almost entirely lit primarily by indirect illumination, our equation more closely resembles a traditional model-based approach.

5 Results

We apply our technique to a four-second sequence of an actor's performance shown in the supplemental video. In the example video, the actor utters several words while continuously rotating her face. By scaling multiple basis images, we can relight the performance with SV light from any direction in the environment. The top two rows of Figure 9 shows the Venetian blinds environment consisting of both SV and directional illumination. The rendering exhibits both environmental illumination and local SV shadows. As the subject moves through the simulated light field, the lighting realistically follows the contours of her face. In the bottom row of Fig. 9 we show a cathedral environment with distant illumination from an HDR environment and local lighting from a focused stained-glass window. To further demonstrate this capability, we show the same pattern projected from several different positions and directions. In all cases, the SV lighting originates from a direction significantly different from the original video projector location.

6 Discussion and Future Work

By comparing a relit scene with and without compensation for indirect illumination (Fig. 7), we can see that indirect illumination has a significant effect in SV lighting environments. Without any correction for indirect illumination, the partially shadowed regions of the nose and eyebrow appear too dark (Fig. 7 (a)) In the worst case, where there is no existing direct illumination, α becomes zero and the relit image appears black. With the addition of simulated indirect illumination, the nose is now lit consistently with the rest of the image (Figure 7 (b)). The small dark spots correspond to unfilled holes from the scanning process.



(a) (b)

Figure 7: (a) Using Equation 1, there is missing and incorrectly scaled indirect illumination (Closeup of Figure 5) (b) Using Equation 5, shadowed regions of the nose are correctly lit (Closeup of Figure 6)

Figure 8 shows a validation test that compares real-world SV light to virtual renderings based on scaled LED illumination. The reference images (Fig. 8 (a) & (d)) are the linear combination of two scaled structured light patterns. As in Figure 7, the addition of indirect illumination approximates the actual indirect appearance. While some specular interreflections are missing, the largest source of visible errors is the low angular sampling of the lighting basis. As several LEDs were combined to recreate the interpolated projector position, we lose some sharp specularities on the lips and accurate shadow boundaries on the cheeks. Our system would easily scale to a more dense lighting system with additional LEDS allowing for better sampling of shadows and specularities. Our technique still preserves the complex reflections already present in the basis such as the specular highlights in the eyes.

Our implementation of the framework has several limitations. Fast movement by the actor could cause misalignment in both the reflectance basis and structured light patterns. This could be overcome by performing optical flow on the

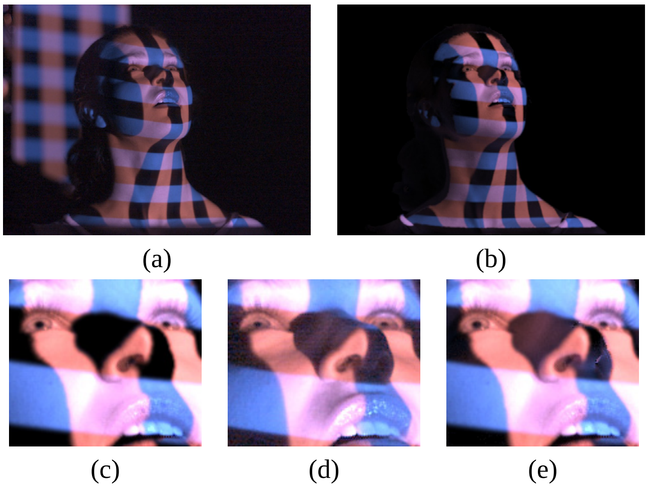


Figure 8: Validation example: (a,d) SV image formed as a linear combination of two real structured light patterns, (c) Virtual SV illumination without indirect illumination simulation, (b,e) Virtual SV illumination with indirect illumination simulation

basis lighting directions [18] and by using structured-light patterns designed for dynamic objects [14, 15, 19]. Geometric fluctuation is occasionally visible in the accompanying video since the hole-filling algorithms of [2, 13] do not enforce temporal consistency. Small temporal changes near occlusion boundaries can cause low-frequency errors in new interpolated geometry. We also have no estimate of surface reflectance or geometry, for regions completely hidden from the camera. Some missing regions could be completed using multiple cameras and projectors to capture a more complete model of subject geometry and reflectance.

7 Conclusion

In this paper, we have presented a technique for extrapolating spatially varying lighting information from basis lighting conditions using a small number of structured light patterns to recover scene geometry. Though our technique does not yield exact results, we believe it represents a useful appoximation for live action human performances. The data required is relatively simple and the relighting framework is easily extendable to more complex lighting simulation methods to produce even more accurate renderings. We believe our technique will help advance toward greater unification between computer graphics, lighting, and live-action cinematography.

Acknowledgements

The authors wish to thank Elizabeth Franklin, Diego Nehab, Laurie Swanson, Tom Pereira, Larry Vladic, Vision Research Inc., Bill Swartout, David Wertheimer, Randolph Hall, and Max Nikias for their support and assistance with this work. This work was sponsored by the University of Southern California Office of the Provost and the U.S. Army Research, Development, and Engineering Command (RDECOM). The content of the information does not necessarily reflect the position or the policy of the US Government, and no official endorsement should be inferred.

References

[1] Billy Chen and Hendrik P. A. Lensch. Light source interpolation for sparsely sampled reflectance fields. In *10th International Fall Workshop on Vision, Modeling, and Visualization 2005*, 2005.

[2] James Davis, Stephen Marschner, Matt Garr, and Marc Levoy. Filling holes in complex surfaces using volumetric diffusion. In *Symposium on 3D Data Processing, Visualization, and Transmission*, 2002.

[3] Paul Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proceedings of SIGGRAPH 98*, 1998.

[4] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, pages 145–156, July 2000.

[5] Michael Goesele, Xavier Granier, Wolfgang Heidrich, and Hans-Peter Seidel. Accurate light source acquisition and rendering. *ACM Transactions on Graphics*, 22(3):621–630, July 2003.

[6] Tim Hawkins, Andreas Wenger, Chris Tchou, Andrew Gardner, Fredrik Göransson, and Paul Debevec. Animatable facial reflectance fields. In *Rendering Techniques 2004: 15th Eurographics Workshop on Rendering*, pages 309–320, June 2004.

[7] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatially varying materials. In *Rendering Techniques 2001: 12th Eurographics Workshop on Rendering*, pages 103–114, June 2001.

[8] Marc Levoy and Patrick M. Hanrahan. Light field rendering. In *Proceedings of SIGGRAPH 96*, Computer Graphics Proceedings, Annual Conference Series, pages 31–42, 1996.

[9] Vincent Masselus, Pieter Peers, Philip Dutré, and Yves D. Willems. Relighting with 4D incident light fields. *ACM Transactions on Graphics (SIGGRAPH 2003)*, 22(3):613–620, July 2003.

[10] Wojciech Matusik, Hanspeter Pfister, Addy Ngan, Paul Beardsley, Remo Ziegler, and Leonard McMillan. Image-based 3D photography using opacity hulls. *ACM Transactions on Graphics*, 21(3):427–437, July 2002.

[11] Ian McDowall and Mark Bolas. Display, sensing, and control applications for digital micromirror displays. In *IEEE VR 2005 - Emerging Display Technologies*, pages 35–36, 2005.

[12] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics*, 25(3):935–944, July 2006.

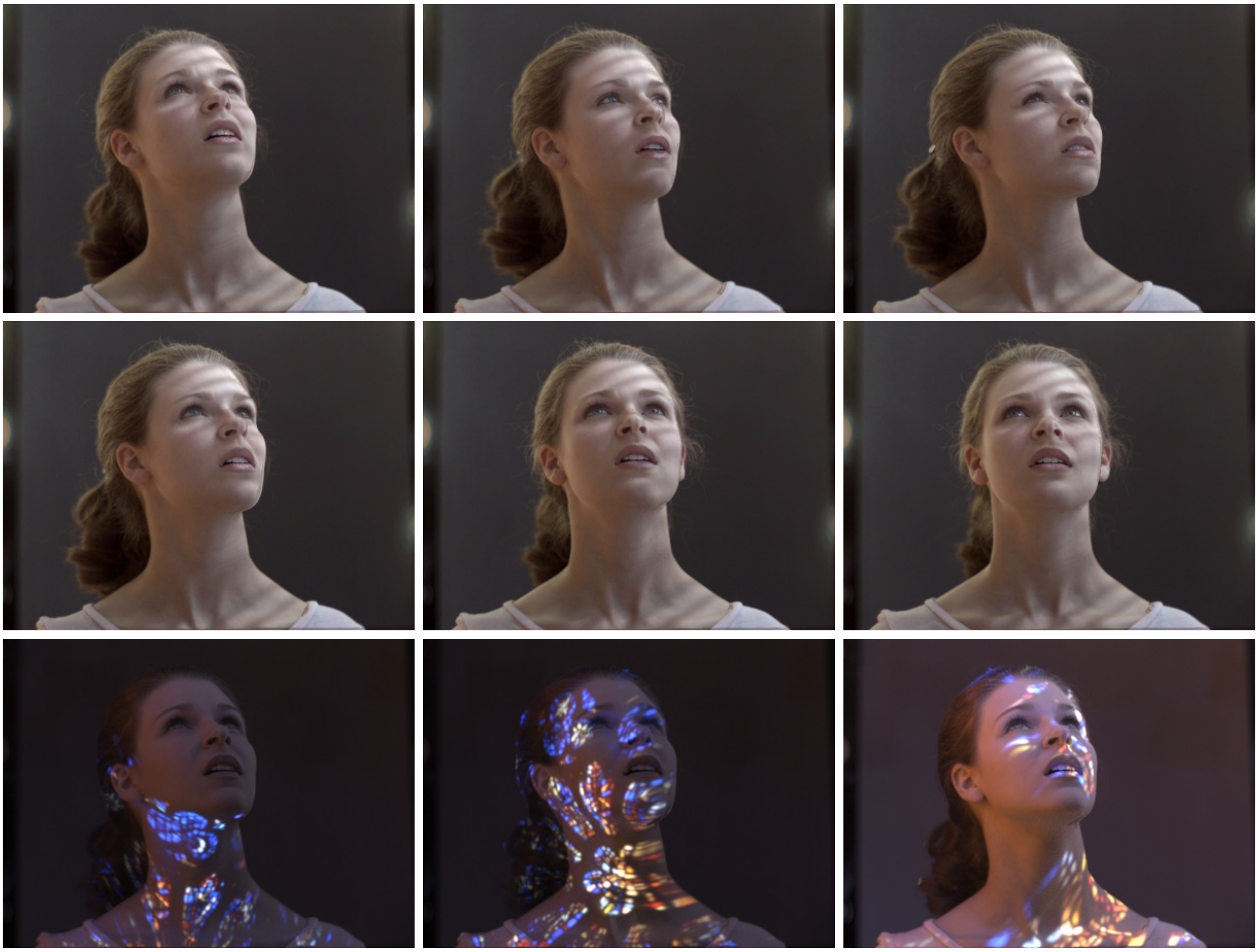


Figure 9: (top,middle) Several frames showing the subject motion relit by Venetian blinds. The final composite combines local and distant illumination. (bottom) The performance lit by a stained-glass window. The three renderings show the same pattern virtually projected from several different LED positions distant from the projector.

[13] Diego Nehab, Szymon Rusinkiewicz, James Davis, and Ravi Ramamoorthi. Efficiently combining positions and normals for precise 3d geometry. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 2005.

[14] Szymon Rusinkiewicz, Olaf Hall-Holt, and Marc Levoy. Real-time 3D model acquisition. *ACM Transactions on Graphics*, 21(3):438–446, July 2002.

[15] Asla Sa, Paulo Cezar P Carvalho, and Luiz Velho. (b,s)-BCSL : Structured light color boundary coding for 3D photography. In *Proceedings of 7th International Fall Workshop on VISION, MODELING, AND VISUALIZATION*, 2002.

[16] Steven M. Seitz, Yasuyuki Matsushita, and Kiriakos N. Kutulakos. A theory of inverse light transport. In *International Conference on Computer Vision (ICCV), 2005*, 2005.

[17] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen Marschner, Mark Horowitz, Marc Levoy, and Hendrik Lensch. Dual photography. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 2005.

[18] Andreas Wenger, Andrew Gardner, Jonas Unger, Chris Tchou, Tim Hawkins, and Paul Debevec. Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 2005.

[19] Li Zhang, Noah Snavely, Brian Curless, and Steven Seitz. Spacetime faces: High-resolution capture for modeling and animation. *ACM Transactions on Graphics (SIGGRAPH 2004)*, 2004.

[20] Todd Zickler, Sebastian Enrique, Ravi Ramamoorthi, and Peter Belhumeur. Reflectance sharing: Image-based rendering from a sparse set of images. In *Rendering Techniques 2005: 16th Eurographics Workshop on Rendering*, pages 253–264, June 2005.