

Efficient Estimation of Spatially Varying Subsurface Scattering Parameters

Sarah Tariq^{1,3}, Andrew Gardner², Ignacio Llamas^{1,3}, Andrew Jones², Paul Debevec² and Greg Turk¹

Georgia Institute of Technology¹, Institute for Creative Technologies University of Southern California²,
NVIDIA Corporation³

Abstract

We present an image-based technique to efficiently acquire spatially varying subsurface reflectance properties of a human face. The estimated properties can be used directly to render faces with spatially varying scattering, or can be used to estimate a robust average across the face. We demonstrate our technique with renderings of peoples' faces under novel, spatially-varying illumination and provide comparisons with current techniques. Our captured data consists of images of the face from a single viewpoint under two small sets of projected images. The first set, a sequence of phase-shifted periodic stripe patterns, provides a per-pixel profile of how light scatters from adjacent locations. The second set of structured light patterns is used to obtain face geometry. We subtract the minimum of each profile to remove the contribution of interreflected light from the rest of the face, and then match the observed reflectance profiles to scattering properties predicted by a scattering model using a lookup table. From these properties we can generate images of the subsurface reflectance of the face under any incident illumination, including local lighting. The rendered images exhibit realistic subsurface transport, including light bleeding across shadow edges. Our method works more than an order of magnitude faster than current techniques for capturing subsurface scattering information, and makes it possible for the first time to capture these properties over an entire face.

1 Introduction

Rendering human faces realistically has been a longstanding problem in computer graphics and a subject of recent increased interest. There are several important areas that require the realistic rendering of human faces, including computer games, animated feature films, and special effects for movies. One of the most crucial factors in creating convinc-

ing images of faces is realistic skin rendering. This is a hard problem because skin reflectance consists of many complex components, including light inter-reflection and subsurface scattering.

Subsurface scattering is the phenomenon of light entering one point on a surface of a material and scattering inside it before exiting at another point. In this process, light is not only scattered, but may also be partially absorbed by the material and these effects typically vary for different wavelengths. Visually, subsurface scattering results in a softening of the appearance of the material, color bleeding within the material, and diffusion of light across shadow boundaries. The human visual system can easily notice the absence of these cues. Light diffusion across illumination boundaries is especially important for cinematic lighting, where faces are often in partial lighting and the resulting shadow edges are soft and blurred.

The dipole diffusion model [6] has been proposed as a fast closed-form approximation for calculating the outgoing radiance due to subsurface scattering, and is capable of producing very realistic images. Rendering with this model, however, requires an estimate of the scattering and absorption parameters of the skin. Current methods for capturing these parameters require setups not suitable for faces [6], or specialized devices [17]. More importantly, these methods are too slow to capture the parameters over an entire face, providing only a sparse set of parameters. Skin, however, is not a homogeneous material, and its scattering properties change across different people and across a face; cheeks have more blood, giving them a reddish appearance, areas with whiskers tend to be darker, and there may be uneven patches in the skin. Spatially varying parameters are hence important for capturing the appearance of skin, and, as shown in [5], are already used in the entertainment industry for rendering realistic faces. Due to the absence of a method for estimating spatially varying scattering properties of skin, these techniques could not use

independent spatially varying parameters (see Section 2 for details). Figures 4 and 7 show the importance of using the spatially varying scattering parameters estimated by our method in producing an accurate image.

Our Approach. We propose a technique that measures the spatially varying scattering properties of a subject, using just a camera and projector, with acquisition times under a minute. This time frame is small enough to permit a human subject to remain sufficiently still. We reduce the number of images needed by making key assumptions about light scattering in skin, which are that the scattering is isotropic, local, and has low spatial frequency.

We estimate a point’s scattering properties using a *Scattering Profile* that encodes the amount of light that scatters to it from nearby locations. We measure one such profile per pixel (see Figure 2), using information derived from an image of the subject taken under shifting patterns of black and white stripes. The spatially varying scattering properties of the face are derived from these *Scattering Profiles* by matching them with a lookup table (LUT) of pre-rendered profiles with known absorption and scattering properties. The per-pixel estimated properties and captured geometry are then used to render the face under new local illumination.

2 Background and Related Work

Scattering Models Subsurface transport is described by the Bidirectional Surface Scattering Reflectance Distribution Function (BSSRDF), S , which relates the outgoing radiance $L_o(x_o, \vec{w}_o)$ at surface point x_o in the direction \vec{w}_o to the incident flux $\phi(x_i, \vec{w}_i)$ at surface point x_i from the direction \vec{w}_i :

$$dL_o(x_o, \vec{w}_o) = S(x_i, \vec{w}_i; x_o, \vec{w}_o) d\phi(x_i, \vec{w}_i) \quad (1)$$

Given S , the outgoing radiance L_o can be computed by integrating the incident radiance over incoming direction and area:

$$L_o(x_o, \vec{w}_o) = \int_{2\pi} \int_A S(x_i, \vec{w}_i; x_o, \vec{w}_o) L_i(x_i, \vec{w}_i) (n \cdot w_i) dw_i dA(x_i) \quad (2)$$

The scattering of light in a material is dictated by two optical parameters per wavelength. These are the scattering σ_s and absorption σ_a coefficients, which indicate the fraction of light that is scattered or absorbed by the medium for each unit length of distance traveled by the light. The 8-dimensional

σ'_i	reduced extinction coefficient $\sigma'_s + \sigma_a$
α'	reduced albedo σ'_s / σ'_i
σ_{tr}	effective transport extinction coefficient $\sqrt{3\sigma_a\sigma'_i}$
l_u	mean free path $1/\sigma'_i$
l_d	diffuse mean free path $1/\sigma_{tr}$
z_r	distance from real source to surface l_u
z_v	distance from virtual source to surface $l_u(1 + 4A/3)$
r	distance $\ x_o - x_i\ $
d_r	distance from x_o to real source $\sqrt{r^2 + z_r^2}$
d_v	distance from x_o to virtual source $\sqrt{r^2 + z_v^2}$

Table 1: Definition of variables.

BSSRDF can be evaluated accurately by solving the radiative transport equation, also known as the volume rendering equation. This is an integral equation and direct methods to solve it, including Monte Carlo simulation [4], finite element methods and photon mapping, are relatively slow.

The diffusion approximation to light propagation in participating media states that for optically dense material the behavior of the material is dominated by multiple scattering [13] and directional dependence is negligible. Hence the BSSRDF can be approximated faithfully by the 4D diffusion approximation, $R_d(x_o, x_i)$, which depends only on the incoming and outgoing points. [13] presents numerical solutions to solve the diffusion equation, but does not present a closed-form solution.

The dipole approximation to the diffusion model is a fast closed-form solution introduced by [6]. It evaluates $R_d(\|x_o - x_i\|)$ for a half infinite slab of homogeneous material as:

$$R_d(r) = \frac{\alpha'}{4\pi} \left\{ z_r \left(\sigma_{tr} + \frac{1}{d_r} \right) \frac{e^{-\sigma_{tr}d_r}}{d_r^2} + z_v \left(\sigma_{tr} + \frac{1}{d_v} \right) \frac{e^{-\sigma_{tr}d_v}}{d_v^2} \right\} \quad (3)$$

Given the scattering and absorption coefficients of the material, $R_d(r)$ gives the fraction of light that exits point x_o given an impulse of incoming light at a point x_i that is distance r away. The total outgoing reflectance at x_o in the direction \vec{w}_o is given by Equation 4. Here $F_{t,i}$ is the Fresnel transmittance at the incident point and $F_{t,o}$ is the Fresnel transmittance at the exitant point.

$$L_o(x_o, \vec{w}_o) = \frac{1}{\pi} F_{t,o}(\eta, \vec{w}_o) \int_A R_d(\|x_o - x_i\|) \int_{2\pi} L_i(x_i, \vec{w}_i) F_{t,i}(\eta, \vec{w}_i) (n \cdot \vec{w}_i) dw_i dA(x_i) \quad (4)$$

Measurement The σ'_s and σ_a of a homogeneous material can be estimated by illuminating a small patch of with a thin beam of white light and fitting the resultant image values to Equation 3, as proposed by [6]. The DISCO method of [3] extends this technique to measure the spatially varying properties of a smooth heterogeneous object. They capture impulse response images of the object to laser light of different wavelengths for all positions x_i . This yields the function $R_d(x_i, x_o)$ for all points x_i and x_o on the surface, which is then used directly to represent the spatially varying subsurface scattering properties. The very large acquisition times, however, make the technique unfeasible for human subjects. In contrast, [10] use a projector and camera setup to project a series of images of point grids on the subject, reducing the acquisition time to a couple of hours. [15] also aim to capture the spatially varying properties of heterogeneous objects, but their focus is on materials with large mesostructures. They factor light transport in such materials into a constant diffuse scattering term and spatially varying mesostructure entrance and exit functions. [17] capture the translucency of skin using a specialized measurement probe, and derive an average value over a sparse sampling of points on several faces. This value can then be used to estimate a single set of scattering and absorption parameters for a target face, given its average diffuse reflectance (see Section 5). Spatially varying diffuse reflectance and a fixed σ'_s have been used by [5] to derive spatially varying σ_a over a face. In contrast to these techniques, our technique *quickly* captures a *dense* sampling of spatially-varying absorption *and* scattering parameters across the entire surface of a target face, using the same camera and projectors usually already present in a typical 3D scanning setup. In addition to being used to render spatially varying sub surface scattering, these parameters can be used to provide a robust average of the scattering parameters for a face. As shown in Figure 7b, using the average of our estimated parameters provides a more accurate result compared with using parameters derived from pre-estimated translucency values (in this case we are using the values from [17]).

Reflectance Fields A related concept to the 8D BSSRDF is the 8D reflectance field [1], $R(x_i, \vec{w}_i; x_r, \vec{w}_r)$, which is a generalization of the BSSRDF to an arbitrary closed surface, A . The reflectance field represents the radiant exitant light field, $R_r(x_r, \vec{w}_r)$, from A in response to every possi-

ble incident light field, $R_i(x_i, \vec{w}_i)$.

Image-based relighting methods attempt to capture the reflectance field of an object and render from this data. They are capable of producing extremely realistic images, since they inherently encode all the complicated reflectance properties of the scene. [1] capture a subset of the reflectance field of a human face by taking basis images of the subject under directional illumination generated by a movable light source some distance away. An image of the face under any incident field of directional illumination can be approximated by a linear combination of these basis images. However, the restriction of illumination to distant light prevents them from being able to simulate the effects of spatially varying light across the face.

[8] use a camera and projector setup to capture a 6D reflectance field $R(x_i, \vec{w}_i; x_r)$, which can be used to relight the scene with local illumination. The coarse resolution and large acquisition time, however, make this technique inappropriate to apply to capturing faces. [12] use a similar approach but greatly reduce the number of images required, and consequently the acquisition time, by using adaptive refinement and Helmholtz reciprocity. Adaptive refinement, however, requires online computer analysis during capture, which currently cannot be performed in real time.

3 Choice of SSS Parameters

In this work we fit our observed data to the dipole diffusion model in order to estimate the spatially varying subsurface scattering properties of the face. Our method of using a lookup table with pre-calculated profiles is, however, independent of the model used. It can just as easily be used to fit other models, such as the Sum of Gaussians sometimes used in real time rendering. Given the wide use of the dipole method, we use this model in order to show that its parameters can quickly and easily be estimated. Furthermore, the model expresses the data in a compact representation, with just two parameters per point, which can be represented as maps over the face that are intuitive and easy to edit. These maps are a logical extension of existing maps such as specular and diffuse maps.

An alternative to fitting the desired model to the data using a lookup table is to estimate the radial scattering profile $R_d(r)$ directly, without making any assumptions about the scattering model. Our decision to forego this approach was based on a

couple of drawbacks. The problem of trying to recover the function $R_d(r)$ at a point x from images of x under multiple beams of light (our projected patterns) can be stated as recovering the impulse response of a system whose input and output are known, and falls under the area of signal deconvolution. In general this problem is ill-conditioned, and the solutions, especially for impulse responses that are Gaussian-like in shape, are highly sensitive to measurement noise. This is partly because the power spectrum of these functions quickly becomes very small, which leads to issues with numerical stability. By fitting the observed data to a model instead, we reduce the number of unknowns in the system, which makes a robust estimation possible even with noisy data.

4 Illumination Patterns

The optical parameters, σ'_s and σ_a , of a homogeneous translucent material can be accurately determined by illuminating the surface with a thin beam of light and fitting Equation 3 to a one dimensional radial profile of the spread [6]. Extending this approach to capture spatially varying parameters necessitates capturing prohibitively many images for real time capture, [3, 10]. In order to reduce the number of images needed we make three assumptions about the material: that it has slowly varying scattering properties, that the material is optically dense, and that scattering is isotropic [17]. The assumption of an optically dense material implies that its behavior is dominated by multiple scattering, and hence its properties can be faithfully modeled by the diffusion approximation. Furthermore, an optically dense material will have a short mean free path and hence the scattering properties of the material in one part will only affect a small neighboring area.

To estimate per-point optical properties we need to capture images of the face under a small sequence of light patterns. Ideally, each image should yield scattering information for many parts on the face simultaneously. To meet this goal, instead of measuring the radial decrease in intensity away from a thin spot of light [6, 10], we measure the decrease in intensity away from a *finite-width* stripe of light that is projected onto the face. Furthermore, using the assumption that light does not travel far within the skin, we illuminate the skin with multiple stripes of light simultaneously. Hence, our patterns are periodic stripes of light that are swept across the sub-



Figure 1: (a) Face lit by the fully on projector, (b) one of 40 phase-shifted stripe patterns, (c) after correcting for lambertian light attenuation, (d) estimate of indirect light from subtracting the minimum of each pixel’s profile.

ject. Figure 1(b) is an image of one such pattern.

The widths of the stripes and the distances between them are dictated by conflicting requirements. Firstly, to minimize the interference of neighboring stripes, so that points that fall in the middle of two stripes are completely dark, the stripes need to be as far apart as possible. However, to reduce the number of patterns, and hence the time needed to capture all the images, we need to keep the stripes close together. Secondly, the widths of the stripes should ideally be a single projector pixel. Practically, however, at low light levels the subsurface scattering signal that we are trying to observe is too low to be accurately distinguishable from the camera noise. Hence, to increase the signal to noise ratio we keep the stripes several pixels thick. We found that 8 pixel wide stripes that are 32 pixels apart, for a total of 40 patterns, is a good compromise between all these requirements. The approach presented in [11], which is based on Hadamard coding, could also be used to generate single pixel width stripes with a better signal to noise ratio. However, we found the information available in our data was more than sufficient for estimating the scattering parameters of interest.

5 Parameter Estimation

We use the images of a subject under the projected stripe patterns to create *scattering profiles* for each stripe patterns. A pixel’s scattering profile is a 1D array of values representing its reflectance response over time to the shifting patterns. Figure 2 shows the graphs of four such profiles (with their phase shifted so that the peak is centered), that were created from our 40 input stripe patterns.

The per-pixel scattering and absorption parameters are estimated from these profiles by fitting the profiles to the dipole model using a lookup table, similar to the specular lobe parameter estimation

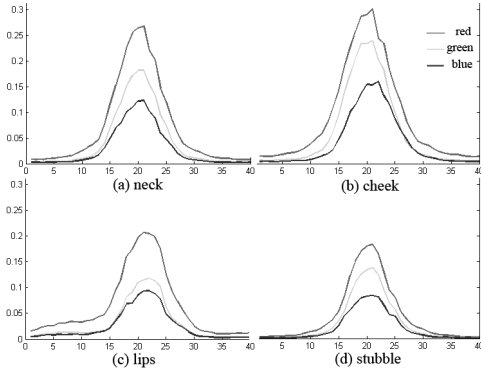


Figure 2: Scattering Profiles of four different points on the face for the three color channels, used for estimating translucency. Note that the lips (c) have a comparatively higher profile in the red channel and the profiles in the stubble region (d) tend to zero more quickly (indicating comparatively lower translucency in all channels).

of [2]. We could also use a least squares fit to the diffusion equation to estimate these parameters, but we found the LUT based approach to be faster. Our LUT contains pre-computed profiles indexed by two optical parameters, and we match a given input profile against the table to estimate its parameters. Instead of using σ'_s and σ'_d , we use a different parametrization of the dipole model as presented in [7], the total diffuse reflectance, R_d , and the translucency, denoted by the diffuse mean free path, l_d . Given values of R_d and l_d , [7] calculate values of σ'_s and σ'_d by inverting Equations 5 and 6.

$$R_d = \frac{\alpha'}{2} \left(1 + e^{-\frac{4}{3}A\sqrt{3(1-\alpha')}} \right) e^{-\sqrt{3(1-\alpha')}} \quad (5)$$

$$l_d = \frac{1}{\sigma'_i \sqrt{3(1-\alpha')}} \quad (6)$$

The advantage of using this re-parametrization is that the total diffuse reflectance value R_d , which is the outgoing radiance of a point when all points around it are illuminated, is available in a fully lit image of the subject after subtracting specular reflection and indirect light (Figure 1(a)). Hence although we build a two-dimensional LUT of R_d vs l_d , for any given pixel we only have to perform a one-dimensional search for the value of l_d , in the row given by its R_d value.

We estimate the total outgoing radiance at pixel x_o resulting from p nearby illuminated pixels using the simplified Equation 7. This equation ignores

Fresnel effects (since these effects are more important at grazing angles where our data is unreliable) and includes the cosine fall-off term inside $L_i(x_i, w_i)$. $R_d^o(r)$ is calculated using Equation 3 and the properties at x_o , and b is the constant amount of indirect light falling on the pixel (section 5.2).

$$L_o(x_o) = b + \sum_{i=0}^p R_d^o(r) L_i(x_i, \vec{w}_i) \quad (7)$$

The time profiles in the LUT are constructed by calculating for each stripe pattern, a pixel x_o 's response to it as given by Equation 7. p are all pixels in the camera image that get direct light from the projector pattern and are also near enough to x_o to influence its appearance. The calculation of p depends on the ratio of projector distances to camera distances, which in turn depends on the angle between the projector and the surface point. In this work we assume that the surface is normal to the direction of the projector. For more accurate results we could either capture the surface from additional angles, or use a 3D table that takes into account the angle of the surface.

We use three different LUTs, one for each color channel, to maximize the resolution of our estimated parameters. Each LUT has a resolution of 200×200 , but covers different areas in the translucency and reflectivity spectra. Each location in the LUT contains an area-normalized time profile, and the scattering and absorption parameters corresponding to the R_d and l_d values. Four time profiles from different parts of the LUT are shown in Figure 3 (d)-(g).

5.1 Parameter Profile Matching

The R_d value of a pixel is estimated by using the pixel's value in the fully lit image, after the incoming light has been normalized to one (Section 7), and angular light fall off has been accounted for (Section 6). We use cross-polarized illumination to ensure that the observed reflectance values do not contain a specular component. To determine a pixel's l_d value we use its R_d value to index into the LUT and perform a one-dimensional search in the corresponding row for the profile that best matches the area-normalized scattering profile of the pixel. Our error metric is the sum of squared differences between the two profiles, and since this metric varies smoothly across the table for a given input profile, our search is implemented using a parabolic minimum finder. The σ'_d and σ'_s correspond-

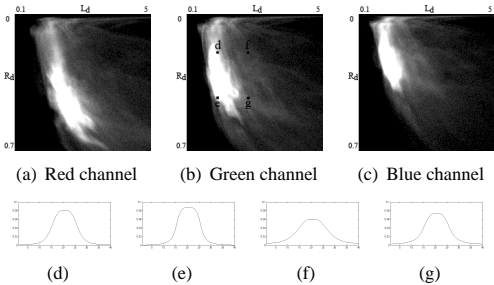


Figure 3: Visual representation of the number of matches found for each entry of the LUT for the three color channels (a,b,c) for the female subject. We see that red has a higher average mean free path, and also that it has large variations across the face compared to the other two channels. Figures (d)-(g) show the area-normalized profiles for four combinations of R_d and L_d shown in (b).

ing to the best matching profile are then stored in separate scattering and absorption maps.

Figure 6 shows some of the different maps that we estimate. We can see that the beard region has higher absorption and lower translucency values, which was also visible in Figure 4. The lips have a higher translucency, and significantly lesser absorption in the green and blue channels. Figures 3 (a-c) shows the number of times a LUT entry is matched to an input profile for the three color channels, intuitively showing the distribution of the parameters R_d and L_d over the face.

5.2 Indirect Light Compensation

In areas of concavities the reflectance profiles that we observe may contain significant amounts of interreflected light. The contribution of this light transport has to be removed before we fit the subsurface scattering model to the reflectance values. We note that most indirect light on a point generally arrives from areas that are relatively far and relatively large compared to the stripe distance: for example, indirect light on a point on the side of the nose arrives from all over the cheek. Thus it is a reasonable assumption that, given the frequency of our stripes, the same amount of indirect light arrives at a particular point on the face regardless of which stripe pattern phase is being projected. The same observation was made earlier by [14] in regard to Gray code stripe patterns. This assumption is violated in areas with local geometric detail (such as skin folds), and we do not expect this technique to produce accurate

results in such regions.

To find the amount of indirect light we first estimate the best matching curve as outlined in section 5.1, and then subtract from the minimum of the input curve the minimum of the matched curve. This gives us the amount of reflectance that the scattering model could not account for, most probably due to the interreflected light. Following [14], we use subtraction to compensate for this indirect light; we subtract its estimated value from the input profile and then rematch the profile to the LUT. Figure 1(d) shows the amount of indirect light that we estimated for a particular capture. Note that the greatest amount of indirect light was calculated to be in concavities such as the eye sockets and under the chin. More recently, [9] uses a similar technique to estimate the direct and indirect light in various scenes.

6 Geometry Acquisition and Use

Since the surface of the face is non-planar, it is important to acquire geometry for both parameter estimation and rendering. For parameter estimation we have to correct for the decrease in light intensity falling on a point as the angle between its normal and the light vector increases (see Figure 1(a)). Failing to correct for this light attenuation would lead us to erroneously conclude that the skin is darker in such areas. Figure 1(b) shows how the fully lit image looks after correcting for the cosine light falloff, by dividing the reflectance value at each pixel by the cosine of the angle between the light and the normal vector. This correction is performed on each of our input images.

We estimate the face geometry by first deriving a sub-pixel accurate projector-camera correspondence image using the method described in [14], which reliably distinguishes between illumination edges in the presence of subsurface scattering. We then triangulate the correspondence image using the calibration information of the camera and projector, [18]. Finally, the geometry is smoothed to reduce high-frequency extraction artifacts.

7 Results and Discussion

Setup. Our setup consists of a projector and camera pair aimed at the subject. We use a high resolution camera (2352×1726 Basler A404) to capture detailed images of the skin in a short amount of time. The projector and camera are synchronized by assigning alternating colors to a small square

		$\sigma_a [mm^{-1}]$			$\sigma'_s [mm^{-1}]$			Translucency		
		Red	Green	Blue	Red	Green	Blue	Red	Green	Blue
female	lips	0.0657	0.2505	0.2835	1.102	0.9084	0.7982	2.0844	1.0716	1.0426
	forehead	0.0836	0.1774	0.2554	1.3079	1.5042	1.1741	1.6928	1.0571	0.9555
	cheek	0.0525	0.1653	0.2216	1.2814	1.5000	1.1248	2.1817	1.1004	1.0569
	mean	0.075	0.156	0.2254	1.3881	1.5017	1.0969	1.9282	1.2048	1.0988
	std. dev.	0.0532	0.0652	0.0703	0.3484	0.3737	0.2496	0.3378	0.1986	0.1524
male	lips	0.3227	0.6631	0.6833	0.7819	0.6710	0.5209	0.9670	0.6139	0.6365
	forehead	0.4761	0.3163	0.442	1.1415	1.2267	0.8851	0.6578	0.8264	0.7538
	stubble	0.4056	0.6024	0.6707	0.9237	0.8621	0.6648	0.7863	0.6147	0.6100
	mean	0.2047	0.3595	0.4691	1.1602	1.1586	0.8373	1.1926	0.8164	0.7606
	std. dev.	0.0989	0.1148	0.1179	0.1823	0.1985	0.1354	0.3127	0.15611	0.1325
Weyrich 06								1.8155	1.0213	0.6453
Jensen 01	skin1	0.032	0.17	0.48	0.74	0.88	1.01	3.6733	1.3665	0.6827
	skin2	0.013	0.070	0.145	1.09	1.59	1.79	4.8215	1.6937	1.0899

Table 2: Estimated parameters for different areas on two different subjects, female and male. The last row is provided for comparison against constant parameters estimated in previous work, [17, 6].

in the projector images. A photoreceptor detects these changes and triggers the camera. The setup is capable of displaying and capturing five patterns per second, with two images of different exposures captured for each projected pattern. The differently exposed images are compiled into HDR images after subtracting the camera and projector black level. We calibrate the incident illumination by taking an image of a reflectance standard, and we divide all captured images by its average value. Finally, we cross-polarize the camera and the projector to eliminate the specularly reflected light, which is necessary for correctly fitting our scattering model to the observed values. The polarizers are placed diagonally on the LCD projector and camera to avoid color shifts. We project a total of 88 images: 40 stripe patterns, and 48 structured light images.

Table 2 shows a comparison of the estimated parameters across the face for different subjects. There is a significant difference in both parameters across the male and female subject; the male subject’s skin exhibits higher absorption and smaller scattering. Our parameters are on the same scale as those previously estimated, although the color distribution is slightly different. In particular, [6] predicts a greater σ'_s for the blue channel than the red channel. This might be because of different color response of our systems, or because [6] obtain measurements on the subject’s arm instead of the face.

We validate our parameters with renderings of faces produced under different incident illumination and we compare them to ground truth images (cap-

tured under cross-polarized illumination to eliminate specular reflections). Our renderings do not reproduce the complete reflectance of a person’s face, as we ignore the specular surface reflection and interreflected light, which would be simulated in the context of a complete face scanning system (e.g. [1, 17]). Using the spatially varying properties estimated in Section 5 and the geometry estimated in Section 6, we render the subsurface response of the face using the method outlined in [6]. We perform a color correction step on our renderings using a color matrix to compensate for crosstalk between the color channels of the projector and the camera.

Figure 4 shows a comparison of our approach with an image rendered with constant translucency (an average of the estimated translucency values over the entire face), and spatially varying reflectivity. The presence of whiskers inside the skin on the right side of the image reduces its translucency, an effect captured by our technique (Figure 4(c)). Note that just a constant translucency parameter (Figure 4(a)) cannot account for this. Figure 7 shows that constant parameters also cannot correctly predict the response of skin in the lip region. However, using constant parameters that are the average of our estimated spatially varying parameters for the subject, 7(b), provides a more accurate image compared with parameters derived from previously published translucency values 7(a). Figure 5 shows a comparison between a synthetic rendering created using our technique and a real image captured under similar light. Our technique correctly

renders the scattering of light across the illumination edge and the skin color in fully lit areas.

8 Limitations and Future Work

There are several limitations to our method that suggest logical directions for future work. Since we capture images from a single viewpoint, both our geometry and estimated parameters are unreliable in areas where the angle between the surface normal and the projector or camera is large (e.g. the white areas in Figures 6a and 6b), or where they are not in focus. In practice the acquisition should be repeated for multiple viewpoints, e.g. left, front, and right, and the results merged so that the most fronto-parallel data is used from each position. Alternatively, we could create a 3D LUT, where the third dimension represents the angle between the normal at a point and the projector direction. Since these angles are already known for each point, the 3D table would not impose any extra cost for searching. Although our LUT-based approach to estimation is extremely fast, it assumes locally flat geometry. This assumption is valid for most image regions but causes miscalculation of parameters for points near geometric discontinuities. For such places, a more accurate but more expensive approach would be to use the geometric information and a least squares solution. A good compromise, however, is to manually correct such errors with information from the neighboring area, which is reasonable since the parameter maps that we estimate are images that can easily be edited and touched up.

9 Conclusion

We have presented an image-based acquisition system that quickly and efficiently estimates the scattering properties of an optically dense translucent non-homogeneous material. Our system allows us to capture the spatially varying properties of skin at a high resolution, which was not possible with earlier approaches because of their large time requirements. Our method can also be used to estimate improved average scattering parameters for the skin more conveniently and efficiently than earlier point-sampling based approaches. The technique is minimally invasive, using neither potentially dangerous lasers [6] or specialized equipment [17]. Our data representation is compact and editable, with just two floating-point values per pixel, and can be used to render realistic faces under local illumina-

tion. Our data capture approach is a logical addition to existing face-scanning approaches e.g. [1, 16, 17] as it requires only a projector and camera.

References

- [1] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *SIGGRAPH*, 2000.
- [2] Andrew Gardner, Chris Tchou, Tim Hawkins, and Paul Debevec. Linear light source reflectometry. *ACM Transactions on Graphics*, 22(3):749–758, July 2003.
- [3] M. Goesele, H. P. A. Lensch, J. Lang, C. Fuchs, and H.-P. Seidel. Disco: acquisition of translucent objects. *ACM Transactions on Graphics*, 23(3):835–844, August 2004.
- [4] Pat Hanrahan and Wolfgang Krueger. Reflection from layered surfaces due to subsurface scattering. In *Proceedings of SIGGRAPH 93*, August 1993.
- [5] Christophe Hery. Implementing a skin bsrdf. *RenderMan: Theory and Practice*, Siggraph Course Notes, 2003.
- [6] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In *Proceedings of ACM SIGGRAPH 2001*, August 2001.
- [7] Henrik Wann Jensen and Juan Buhler. A rapid hierarchical rendering technique for translucent materials. *ACM Transactions on Graphics*, 21(3):576–581, July 2002.
- [8] Vincent Masselus, Pieter Peers, Philip Dutré, and Yves D. Willems. Relighting with 4d incident light fields. *ACM Transactions on Graphics*, 22(3):613–620, July 2003.
- [9] S.K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast Separation of Direct and Global Components of a Scene using High Frequency Illumination. In *SIGGRAPH*, 2006.
- [10] Pieter Peers, Karl vom Berge, Wojciech Matusik, Ravi Ramamoorthi, Jason Lawrence, Szymon Rusinkiewicz, and Philip Dutre. A compact factored representation of heterogeneous subsurface scattering. In *SIGGRAPH*, 2006.
- [11] Yoav Y. Schechner, Shree K. Nayar, and Peter N. Belhumeur. A theory of multiplexed illumination. In *9th IEEE International Conference on Computer Vision (ICCV'03)*.
- [12] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual photography. In *SIGGRAPH*, 2005.
- [13] J. Stam. Multiple scattering as a diffusion process. In *Eurographics Rendering Workshop 1995*, 1995.
- [14] Christopher D. Tchou. Image-based models: Geometry and reflectance acquisition systems. Master's thesis, University of California at Berkeley, 2002, <http://gl.ict.usc.edu/research/3DScanning/Tchou-Thesis-2002.pdf>.
- [15] Xin Tong, Jiaping Wang, Steve Lin, Baining Guo, and Heung-Yeung Shum. Modeling and rendering of quasi-homogeneous materials. In *SIGGRAPH*, 2005.
- [16] Andreas Wenger, Andrew Gardner, Chris Tchou, Jonas Unger, Tim Hawkins, and Paul Debevec. Performance relighting and reflectance transformation with time-multiplexed illumination. In *SIGGRAPH*, 2005.
- [17] T. Weyrich, W. Matusik, H. Pfister, J. Lee, A. Ngan, H. W. Jensen, and M. Gross. Analysis of human faces using a measurement-based skin reflectance model. *ACM Transactions on Graphics*, 25(3):1013–1024, July 2006.
- [18] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.

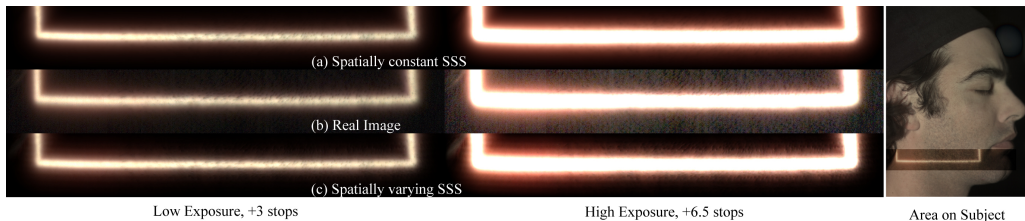


Figure 4: An image of a stripe of light projected onto a male subject’s neck (left), and stubble (right) area. Note that the real image (b), exhibits significantly less subsurface scattering in the stubble region as compared to the neck. Our approach (c), using spatially varying scattering parameters captures this effect, whereas the rendering using constant average translucency (a) overestimates the scattering in the stubble region and underestimates the scattering in the neck.

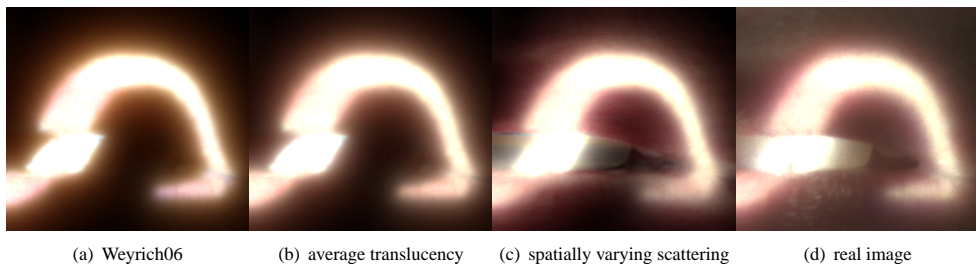


Figure 7: A curved stripe of light projected over the female subject’s lips. (a) and (b) are rendered with constant translucency; parameters from [17] and an average of our estimated parameters, respectively. Note that the scattering in (b) has a yellower tone than the real image. (c) is rendered using spatially varying parameters, and comes closest to capturing the color and influence of the subsurface transport in the real image, (d).

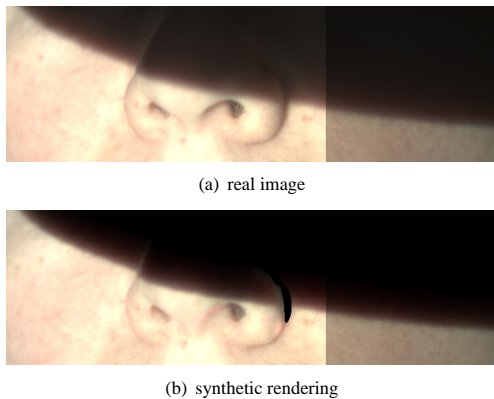


Figure 5: Comparison of a rendering using our estimated parameters with a real image under spatially-varying incident illumination (high exposure on left, low exposure on right). The amount of scattering along the shadow edge is generally consistent between the two, except for the inter-reflected light which is not simulated.

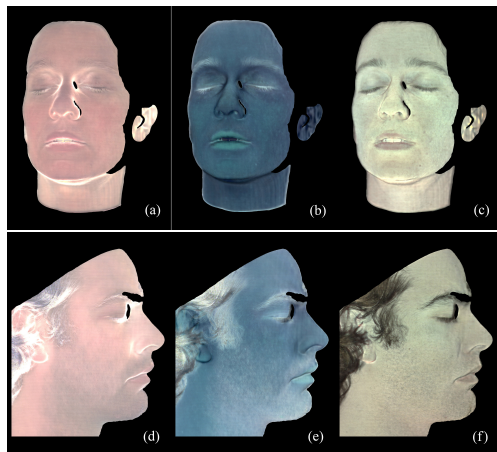


Figure 6: From the top: recovered transmissivity, absorption σ_a , and scattering σ_s' maps for the three color channels. The white parts in (a,d) are areas where we cannot estimate good parameters, see Section 8.