Efficient Estimation of Spatially Varying Subsurface Scattering Parameters for Relighting

Sarah Tariq, Andrew Gardner, Ignacio Llamas, Andrew Jones, Paul Debevec and Greg Turk

USC ICT Technical Report ICT-TR-01.2006. July 27, 2006

Abstract

We present an image-based technique to rapidly 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 contains structured light and is used to obtain face geometry. We 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 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 application 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 convincing 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 sophisticated setup, [6], or specialized devices, [14]. More importantly, these methods are too slow to capture the parameters over an entire face. Renderings based on this model consequentially can only use pre-measured constant 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. Figures 2 and 1 show the importance of using spatially varying scattering parameters in producing a completely realistic image.

Our Approach. We propose a technique that measures the spatially varying scattering properties

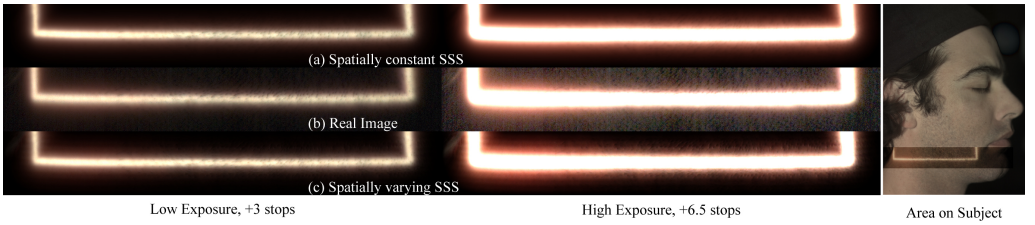


Figure 1: An image of a stripe of light projected onto a male subject's neck(left), and stubble(right) area. Note that the real image(b), exhibits significantly lesser subsurface scattering in the stubble region as as compared to the neck. Our approach(c), using spatially varying scattering parameters 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.

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 4), 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

2.1 Scattering Models

Subsurface transport is described by the Bidirectional Scattering Surface 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 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 [10] 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. [10] 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

σ_t'	reduced extinction coefficient $\sigma_s' + \sigma_a$
α'	reduced albedo σ_s'/σ_t'
σ_{tr}	effective transport extinction coefficient $\sqrt{3\sigma_a\sigma_t'}$
l_u	mean free path $1/\sigma_t'$
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.

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. [6] estimate σ_s' and σ_a of skin by illuminating a small patch on the arm with a thin beam of white light and fitting the resultant image values to Equation 3 by a constrained least squares fit. 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 large acquisition times however make the technique unfeasible for human subjects. [12] 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. [14] 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 derive scattering and absorption parameters for a target face, given its diffuse reflectance, Section 5. In contrast, our technique quickly captures a dense sampling of spatially-varying translucency across the entire surface of the 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 over the entire face. As shown in Figure 2b, using the average parameters from the given face provide a more accurate result compared with using pre-estimated constant parameters,[14].

2.2 Reflectance Fields

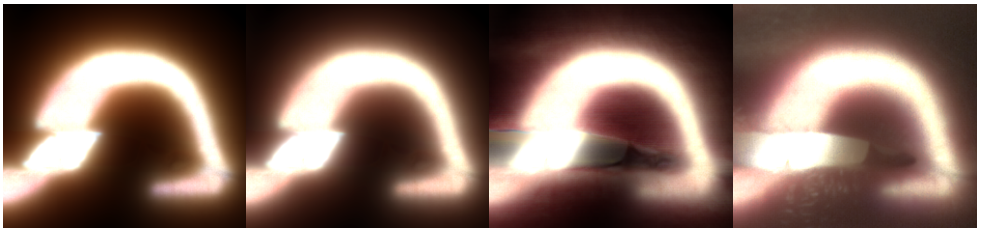
A related concept to the 8D BSSRDF is the 8D reflectance field [1], $R(x_i,\vec{w_i};x_r,\vec{w_r})$, which is a generalization of the BSSRDF to an arbitrary closed surface, A. The reflectance field represents the radiant exitant light field, $R_r(x_r,\vec{w_r})$, from A in response to every possible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

[7] 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. [9] 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 Subsurface Scattering Representation

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



(a) Weyrich06 (b) average translucency (c) spatially varying scattering (d) real image

Figure 2: A curved stripe of light projected over the female subject's lips.(a) and (b) are rendered with constant translucency; parameters from [Weyrich06] 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).

face. Our method of using a lookup table with precalculated profiles is however independent of the model used. It can just as easily be used to fit other models, like the Sum of Gaussians that is used in real time rendering. We choose to use the dipole model since given its wide use, we believe it is important and useful to demonstrate that parameters for this model can be quickly and easily be estimated. Furthermore, the model expresses the data in a compact representation, just two parameters per point, which can be represented as maps over the face that are intuitive and easy to edit and touch up. These maps are a logical extension of existing maps such as specular and diffuse maps.

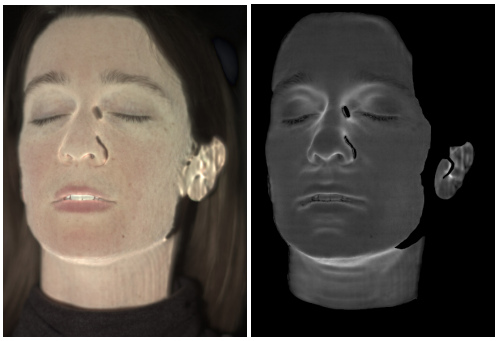
An alternative to fitting the desired model to the data using a lookup table is to estimate the radial scattering profile $R_d(r)$ directly, without making any assumptions about the scattering model. Our decision to forego this approach was based on a couple of drawbacks. The problem of trying to recover 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 considered to be 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 Light Patterns for Parameter Estimation

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] use a constrained nonlinear least squares method for this purpose. A similar approach could be used to estimate spatially varying properties of a surface by capturing images of the surface illuminated at all input points x_i, but would necessitate taking prohibitively many images for a real time capture, [3]. 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 [14]. 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 effect 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

(a) fully lit (b) stripes



(c) flattened (d) indirect

Figure 3: Acquired images, top; fully lit image (a), one of 40 stripe patterns (b), And intermediate images, bottom; fully lit image corrected for light attenuation (c), and estimate of indirect light (d).

thin beam of light [6], we measure the decrease in intensity away from a finite-width *line* of light that is projected onto the face. Further more, 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 subject, Figure 3(c) 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 these low levels of light the

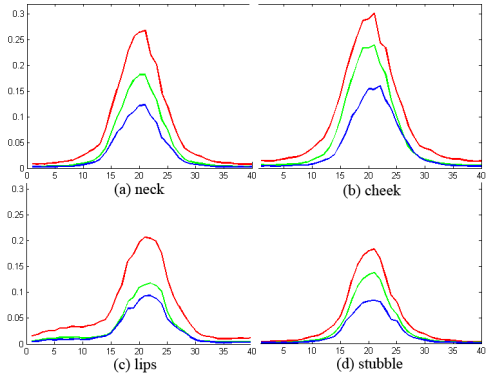


Figure 4: Scattering Profiles of four different points on the face for the three color channels, used for estimating translucency. Note that the lips (c) have a comparatively higher profile in the red channel and the profiles in the stubble region (d) tend to zero more quickly (indicating comparatively lower translucency in all channels).

signal that we are trying to observe, the subsurface scattering, 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 [8], which is based on Hadamard coding, could also be used to generate single pixel width stripes with 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 captured under the projected stripe patterns to create *scattering profiles* for each pixel. A pixel's scattering profile is a 1D array of values representing its reflectance response over time to the shifting patterns. Figure 4 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 [2]. We could also use a least squares fit to the diffusion equation to estimate these para-

meters, 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_a, we use a different parametrization of the dipole model as presented in [5], 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, [5] calculate values of σ_s' and σ_a by inverting Equations 5 and 6.

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

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

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 3 (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).

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

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 it's appearance. The calculation of p depends on the ratio of projector distances to camera distances, which in turn depend on the angle between the projector and the surface point. We assume that the surface is normal to the direction of the projector, which works well for our results. For more accurate results we need to either capture the surface from more angles, or use a 3D table that takes into account the angle of the surface.

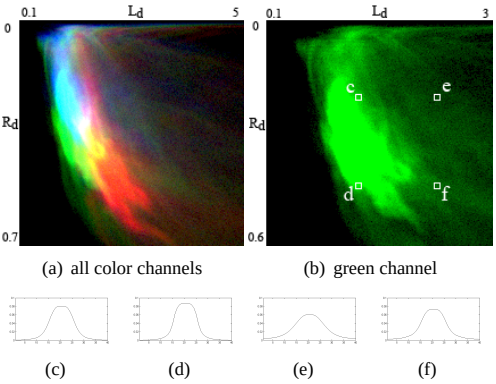


Figure 5: Visual representation of the number of matches found for each entry of the LUT for the three color channels (a). We can clearly see that the red color tends to have a higher mean free path in skin, and also that it has large variations across the face compared to the other two channels. Figures (c)-(f) show the area normalized profiles for four combinations of R_d and L_d shown in (b).

In practice we use three different LUTS, one for each color channel, to maximize the resolution of our estimated parameters, see Figure 5 (a,b). Each LUT has a resolution of 200*x*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 σ_s' and σ_a corresponding to the R_d and l_d values. Four time profiles from different parts of the LUT are shown in Figure 5 (c)-(f).

5.1 Estimating Parameters using 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

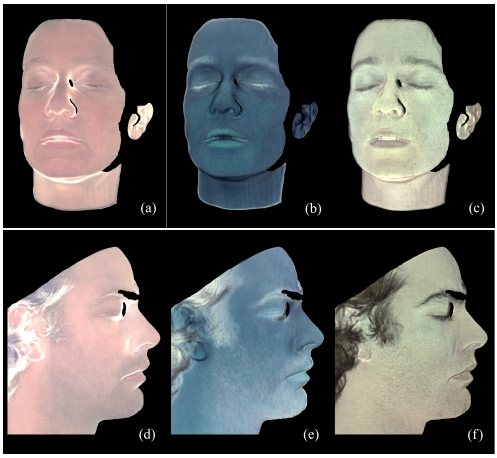


Figure 6: From the top:recovered translucency, absorption σ_a, and scattering σ_s' maps for the three color channels. The white in figures(a,d) represent areas where we cannot estimate good parameters, see section 8

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_a and σ_s' corresponding to the best matching profile are then stored in separate scattering and absorption maps at the pixel location.

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 1. The lips have a higher translucency, and significantly lesser absorption in the green and blue channels. Figure 5 (a) shows the number of times a LUT entry is matched to an input profile, 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 inter-reflected light. The contribution of this light transport has to be removed before we fit the subsurface scattering model to the reflectance values. We make the observation 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 begin projected. 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 indirect light. To compensate for this light we subtract its estimated value from the input profile and then rematch the profile to the LUT. Figure 3 (d) shows the amount of indirect light that we estimated for a particular capture session. Note that the greatest amount of indirect light was calculated to be in concavities such as the eye sockets and under the chin, as is expected.

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 3(a)). Not correcting for this light attenuation would lead us to erroneously conclude that the skin is darker in such areas. Figure 3(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 for all of our input images.

For rendering we need the geometry of the face and the scene. We estimate the geometry by first deriving a sub-pixel accurate projector-camera correspondence image using the method described in [11], 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, [15]. Finally, the geometry is smoothed to reduce high frequency artifacts resulting from the extraction.

		$\sigma_a\ [mm^{-1}]$			$\sigma_s'\ [mm^{-1}]$			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,[14], and [6].

7 Results and Discussion

Setup. Our setup consists of a projector and camera pair aimed at the subject. We use a high resolution camera (2700*x*1700 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 in the projector images. A photo receptor 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 high dynamic range 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 put on diagonally 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 for different subjects and different points on the face. 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 previously estimated parameters, although the color distribution is slightly different. In particular, [6] predict 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.

We validate our parameters with renderings of faces produced under different incident illumination and we compare them to ground truth images captured under cross-polarized illumination to eliminate specular reflections. Our renderings do not reproduce the complete reflectance of a person's face, including the specular surface reflection and indirect light, which are components of a complete face scanning system (e.g. [1], [14]). Using the spatially varying properties that we estimated from section 5 and the geometry information from 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 3×3 color matrix.

Figure 1 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, similar to the approach proposed in [14]. The presence of the whiskers inside the skin on the right side of the image reduces its translucency, an effect captured by our technique Figure 1 (c). Note that just a constant translucency parameter, Figure 1 (a), cannot account for this. Figure 2 shows that con-

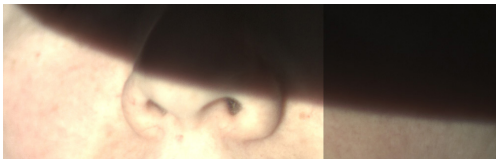
stant parameters also cannot predict correctly the response of skin in the lip region. Figure 7 shows a comparison between a synthetic rendering created using our technique and a real image captured under similar light. Notice that our technique correctly renders the scattering of light across the illumination edge and the skin color in fully lit areas.

8 Limitations and Future Work

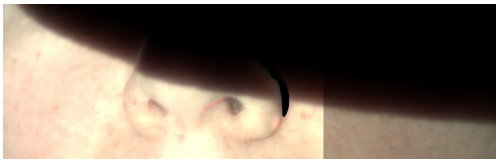
There are several limitations to our method that suggest logical directions for future work. Since we capture images from a single viewpoint both our geometry and estimated parameters are unreliable in areas where the angle between the surface normal and the projector or camera direction is large (for example the white areas in Figure 6 a and b), 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 of the images, but will cause miscalculation of parameters of 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 easy since the 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 enables the capture of spatially varying properties of skin at a high resolution, which was not possible earlier because of the large time requirements of previous approaches. 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 [14]. Our data representation is compact, only two floating-point values per pixel, and can be used to render realistic faces under local illumination in a number of ways. Our data capture approach is a logical addition to existing face-scanning such as [1], [13],[14] and can easily be incorporated in such a system since it only requires a projector and camera.



(a) real image



(b) synthetic rendering

Figure 7: 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 indirect illumination which is not simulated) .

10 Acknowledgements

We thank Sebastian Sylwan, Tom Pereira, and Bill Swartout for their support and assistance with this work. This work was sponsored by the University of Southern California Office of the Provost and USA RDECOM. The content of the information does not necessarily reflect the position or the policy of the sponsors, and no official endorsement should be inferred

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*, pages 145–156. ACM SIGGRAPH, 2000.

[2] Andrew Gardner, Chris Tchou, Tim Hawkins, and Paul Debevec. Linear light source reflectometry. In *Siggraph 2003, Computer Graphics Proceedings*, 2003.

[3] M. Goesele, H. P. A. Lensch, J. Lang, C. Fuchs, and H.-P. Seidel. Disco: Acquisition of translucent objects. In *Siggraph 2004, Computer Graphics Proceedings*, pages 835–844, 2004.

[4] P. Hanrahan and W. Krueger. Reflection from layered surfaces due to subsurface scattering. In *SIGGRAPH 1993, ACM Computer Graphics*, pages 165–174, 1993.

[5] H. W. Jensen and J. Buhler. A rapid hierarchical rendering technique for translucent materials. *SIGGRAPH*, pages 576–581, 2002.

[6] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In *ACM SIGGRAPH 2001*, pages 511–518. ACM Press, 2001.

[7] V. Masselus, P. Peers, P. Dutre, and Y. D. Willems. Relighting with 4d incident light fields. In *Siggraph 2003, Computer Graphics Proceedings*, pages 613–620, 2003.

[8] Yoav Y. Schechner, Shree K. Nayar, and Peter N. Belhumeur. A theory of multiplexed illumination. In *Ninth IEEE International Conference on Computer Vision (ICCV'03)*, page 808, 2003.

[9] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual photography. In *SIGGRAPH 2005, Computer Graphics Proceedings*, 2005.

[10] J. Stam. Multiple scattering as a diffusion process. In *Eurographics Rendering Workshop 1995*, 1995.

[11] Christopher Damien Tchou. Image-based models: Geometry and reflectance acquisition systems. Master's thesis, University of California at Berkeley, 2002.

[12] Xin Tong, Jiaping Wang, Steve Lin, Baining Guo, and Heung-Yeung Shum. Modeling and rendering of quasi-homogeneous materials. In *SIGGRAPH, ACM Computer Graphics*, 2005.

[13] 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, Computer Graphics Proceedings*, 2005.

[14] 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. In *SIGGRAPH*, 2006.

[15] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.