Supplemental Document: Polarimetric BSSRDF Acquisition of Dynamic Faces

HYUNHO HA, KAIST, South Korea
INSEUNG HWANG, KAIST, South Korea
NESTOR MONZON, Universidad de Zaragoza - I3A, Spain
JAEMIN CHO, KAIST, South Korea
DONGGUN KIM, KAIST, South Korea
SEUNG-HWAN BAEK, POSTECH, South Korea
ADOLFO MUÑOZ, Universidad de Zaragoza - I3A, Spain
DIEGO GUTIERREZ, Universidad de Zaragoza - I3A, Spain
MIN H. KIM, KAIST, South Korea

ACM Reference Format:
Hyunho Ha, Inseung Hwang, Nestor Monzon, Jaemin Cho, Donggun Kim, Seung-Hwan Baek, Adolfo Muñoz, Diego Gutierrez, and Min H. Kim. 2024. Supplemental Document: Polarimetric BSSRDF Acquisition of Dynamic Faces. *ACM Trans. Graph.* 43, 6, Article 275 (December 2024), 16 pages. https://doi.org/10.1145/3687767

1 Related Work Overview

We summarize the contributions of related methods for face acquisition in Table 1.

2 Polarization and subsurface scattering

2.1 Stokes-Mueller Formalism

A Stokes vector represents the polarization state of a light wave and is denoted as $\mathbf{s} = [s_0, s_1, s_2, s_3]^\intercal \in \mathbb{R}^{4\times1}$. The elements of the Stokes vector include: $s_0 = L$, the intensity of the light; $s_1 = L\psi\cos 2\varsigma \cos 2\chi$, and $s_2 = L\psi \sin 2\varsigma \cos 2\chi$, the power of the 0° and 45° linear polarization components, respectively; and $s_3 = L\psi \sin 2\chi$, the power of the right circular polarization component. ς is the polarization angle, χ is the ellipticity angle, and $\psi = \sqrt{s_1^2 + s_2^2 + s_3^2}/s_0$ is the degree of polarization (DoP), defined as the ratio of the magnitude of the polarized vector elements to the intensity of the light. The effect on the polarization of the interaction between light and any element can be represented by a Mueller matrix $\mathbf{M} \in \mathbb{R}^{4\times4}$, that transforms a Stokes vector $\mathbf{s}_{\text{in}}$ into $\mathbf{s}_{\text{out}}$ as $\mathbf{s}_{\text{out}} = \mathbf{M}\mathbf{s}_{\text{in}}$. For a complete description of polarized light, see the works of Collett [2005] and Wilkie and Weidlich [2012].

Authors' Contact Information: Hyunho Ha, KAIST, South Korea, hhha@vclab.kaist.ac.kr; Inseung Hwang, KAIST, South Korea, ishwang@vclab.kaist.ac.kr; Nestor Monzon, Universidad de Zaragoza - I3A, Spain, nmonzon@unizar.es; Jaemin Cho, KAIST, South Korea, jmcho@vclab.kaist.ac.kr; Donggun Kim, KAIST, South Korea, dgkim@vclab.kaist.ac.kr; Seung-Hwan Baek, POSTECH, South Korea, shwbaek@postech.ac.kr; Adolfo Muñoz, Universidad de Zaragoza - I3A, Spain, adolfo@unizar.es; Diego Gutierrez, Universidad de Zaragoza - I3A, Spain, diegog@unizar.es; Min H. Kim, KAIST, South Korea, minhkim@kaist.ac.kr.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
© 2024 Copyright held by the owner/author(s).
ACM 1557-7368/2024/12-ART275
https://doi.org/10.1145/3687767

Table 1. Comparison with other face capture methods. A green check mark indicates that the component is acquired by the corresponding method. While many dynamic face acquisition methods obtain specular albedo by leveraging polarized light, none of them can obtain a polarimetric BSSRDF parametrization, including the linear polarization components of reflectance. Furthermore, our method is the first to measure all five listed BSSRDF parameters simultaneously with the biophysical parameters of dynamic human faces. LP stands for linear polarization filter. BP stands for bandpass filter for multispectral acquisition. Cyan check marks on the diffuse column means the method does not explicitly model the diffuse appearance. On the dynamic column, the cyan check marks mean the method does not perform tracking. In the specular and subsurface scattering columns, cyan check marks indicate the use of global (or fixed) parameters for humans.

	Method	Camera	Filter	Geometry	Diffuse	Dynamic	Polarization		Face BSSRDF parameters					
							Polarized light	Polarimetric reflectance	Biophysical Params.	Specular albedo	Specular roughness	Single scattering	Subsurface scattering	Refractive index
Photometric stereo	Weyrich et al. [2006]	RGB	–	✔	✔	–	–	–	–	✔	✔	–	✔	–
	Ma et al. [2007]	RGB	LP	✔	✔	–	✔	–	–	–	–	–	–	–
	Ghosh et al. [2008]	RGB	LP	✔	✔	–	✔	–	–	✔	✔	✔	✔	–
	Ghosh et al. [2011]	RGB	LP	✔	✔	–	✔	–	–	✔	–	–	–	–
	Fyffe et al. [2011]	RGB	–	✔	✔	✔	–	–	–	✔	–	–	–	–
	Fyffe and Debevec [2015]	RGB	LP	✔	✔	✔	✔	–	–	✔	–	–	–	–
	Gotardo et al. [2015]	RGB	LP	✔	✔	✔	✔	–	–	–	–	–	–	–
	Fyffe et al. [2016]	RGB	–	✔	✔	–	–	–	–	✔	✔	–	–	–
	LeGendre et al. [2018]	Mono	LP	✔	✔	–	✔	–	–	✔	–	–	–	–
Learning	Li et al. [2020]	RGB	LP	✔	✔	✔	✔	–	–	✔	–	–	–	–
	Bi et al. [2021]	RGB	–	✔	✔	✔	–	–	–	–	–	–	–	–
	Liu et al. [2022]	RGB	LP	✔	✔	✔	✔	–	–	✔	–	–	–	–
	Zhang et al. [2022]	RGB	LP	✔	✔	✔	✔	–	–	✔	–	–	–	–
Biophysical	Preece and Claridge [2004]	Mono	BP	–	–	–	–	–	✔	–	–	–	–	–
	Donner et al. [2008]	Mono	BP	–	–	–	–	–	✔	–	–	–	–	–
	Jimenez et al. [2010]	RGB	–	–	✔	✔	–	–	✔	–	–	–	–	–
	Alotaibi and Smith [2017]	RGB	–	✔	✔	–	–	–	✔	–	–	–	–	–
	Gitlina et al. [2020]	RGB	–	–	✔	–	–	–	✔	–	–	–	–	–
	Aliaga et al. [2022]	RGB	–	–	✔	–	–	–	✔	–	–	–	–	–
	Aliaga et al. [2023]	RGB	–	–	✔	–	–	–	✔	–	–	–	–	–
Stereo matching	Bradley et al. [2010]	RGB	–	✔	✔	✔	–	–	–	–	–	–	–	–
	Beeler et al. [2010]	RGB	–	✔	✔	✔	–	–	–	–	–	–	–	–
	Beeler et al. [2011]	RGB	–	✔	✔	✔	–	–	–	–	–	–	–	–
	Gotardo et al. [2018]	RGB	–	✔	✔	✔	–	–	–	✔	–	–	–	–
	Riviere et al. [2020]	RGB	LP	✔	✔	✔	✔	–	–	✔	✔	–	✔	–
	Azinović et al. [2023]	RGB	LP	✔	✔	–	✔	–	–	✔	–	–	–	–
	Ours	Polar	BP	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔

2.2 Fresnel Equation

The change of Stokes vectors by the transmission and reflection of light can be represented using the Fresnel Mueller matrix $\mathbf{F}^{F\in\{\mathcal{T},\mathcal{R}\}}$ that takes into account the different effects on light polarized along the plane of incidence ($F^{\parallel}$) and light polarized perpendicular to it ($F^{\perp}$). Here $F \in \{\mathcal{T},\mathcal{R}\}$ refers to the Fresnel transmission ($\mathcal{T}$) or reflection ($\mathcal{R}$) coefficients. The Fresnel Mueller matrix is given by

$$\mathbf{F}^{F\in\{\mathcal{T},\mathcal{R}\}} = \begin{bmatrix} \frac{F^{\perp}+F^{\parallel}}{2} & \frac{F^{\perp}-F^{\parallel}}{2} & 0 & 0 \\ \frac{F^{\perp}-F^{\parallel}}{2} & \frac{F^{\perp}+F^{\parallel}}{2} & 0 & 0 \\ 0 & 0 & \sqrt{F^{\perp}F^{\parallel}}\cos\delta & \sqrt{F^{\perp}F^{\parallel}}\sin\delta \\ 0 & 0 & -\sqrt{F^{\perp}F^{\parallel}}\sin\delta & \sqrt{F^{\perp}F^{\parallel}}\cos\delta \end{bmatrix}, \quad (1)$$

where δ is the retardation phase shift. The value of δ is 0 when the incident angle is larger than the Brewster angle, and π otherwise.

The Fresnel coefficients for reflection and transmission, denoted as $\mathcal{R}^{\perp}$, $\mathcal{R}^{\parallel}$, $\mathcal{T}^{\perp}$, and $\mathcal{T}^{\parallel}$, can be calculated as

$$\mathcal{R}^{\perp} = \left(\frac{\eta_1\cos\theta_1 - \eta_2\cos\theta_2}{\eta_1\cos\theta_1 + \eta_2\cos\theta_2}\right)^2, \mathcal{R}^{\parallel} = \left(\frac{\eta_1\cos\theta_2 - \eta_2\cos\theta_1}{\eta_1\cos\theta_2 + \eta_2\cos\theta_1}\right)^2, \quad (2)$$

$$\mathcal{T}^{\perp} = \left(\frac{2\eta_1\cos\theta_1}{\eta_1\cos\theta_1 + \eta_2\cos\theta_2}\right)^2, \mathcal{T}^{\parallel} = \left(\frac{2\eta_1\cos\theta_1}{\eta_1\cos\theta_2 + \eta_2\cos\theta_1}\right)^2. \quad (3)$$

These coefficients describe the polarization state of light after being reflected or transmitted at an interface, and depend on the refractive indices of the media on either side of the interface (η_1 and η_2) as well as the incident (θ_1) and exitant (θ_2) angles. We also define $\mathcal{T}^+ = (\mathcal{T}^\perp + \mathcal{T}^\parallel)/2$ and $\mathcal{T}^- = (\mathcal{T}^\perp - \mathcal{T}^\parallel)/2$ using the Fresnel transmittance coefficients, respectively.

2.3 Coordinate Conversions in Polarization

Different from conventional BRDF formulation, polarimetric rendering requires a coordinate conversion matrix $\mathbf{C}(\vartheta)$ for a given angle ϑ:

$$\mathbf{C}(\vartheta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\vartheta & \sin 2\vartheta & 0 \\ 0 & -\sin 2\vartheta & \cos 2\vartheta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

The polarimetric BRDF should be defined with respect to the coordinate systems of the incident Stokes vector and exitant Stokes vector. A common coordinate system often used for polarimetric BRDFs consists of three orthonormal vectors [Hwang et al. 2022]: the z-axis follows the direction of light propagation, the y-axis ($\bar{\mathbf{y}}_{i,o}$) is aligned with the camera up vector and the x-axis ($\bar{\mathbf{x}}_{i,o}$) is perpendicular to both. The plane of incidence of the specular lobe and single-scattering lobe is defined with respect to the halfway vector $\mathbf{h}$ while the diffuse lobe is defined by the surface normal $\mathbf{n}$.

2.4 Polarimetric Reflectance Model

We adopt the specular and single scattering terms of the polarimetric reflectance model from the recent state-of-the-art model by Hwang et al. [2022].

Specular term. The polarized specular reflection $\mathbf{P}_s$ is defined as

$$\mathbf{P}_s = \kappa_s \mathbf{C}_{h\to o}(-\tilde{\varphi}_o)\mathbf{F}^{\mathcal{R}}(\theta_d;\eta)\mathbf{C}_{i\to h}(\tilde{\varphi}_i), \tag{5}$$

where $\theta_d = \cos^{-1}(\mathbf{h}\cdot\boldsymbol{\omega}_i)$ is the zenith angle between incident light $\boldsymbol{\omega}_i$ and the halfway vector $\mathbf{h}$ [Rusinkiewicz 1998], $\mathbf{F}^{\mathcal{R}}$ is the Mueller matrix form of the Fresnel reflection coefficients $\mathcal{R}$ and $\mathbf{C}_{h\to o}(-\tilde{\varphi}_o)$ and $\mathbf{C}_{i\to h}(\tilde{\varphi}_i)$ are the coordinate conversion matrices. The rotation angles are given as $\tilde{\varphi}_{i,o} = \varphi_{i,o} - \pi/2$, where $\varphi_{i,o} = \tan^{-1}((\mathbf{h}\cdot\bar{\mathbf{y}}_{i,o})/(\mathbf{h}\cdot\bar{\mathbf{x}}_{i,o}))$. The term $\kappa_s = \rho_s\frac{\mathcal{D}(\theta_h;\alpha_s)\mathcal{G}(\theta_i,\theta_o;\alpha_s)}{4(\mathbf{n}\cdot\boldsymbol{\omega}_i)(\mathbf{n}\cdot\boldsymbol{\omega}_o)}$ is the specular reflection term, where $\theta_h = \cos^{-1}(\mathbf{h}\cdot\mathbf{n})$ is the zenith angle between the normal $\mathbf{n}$ and $\mathbf{h}$, $\mathcal{D}$ represents the GGX distribution function [Walter et al. 2007], α_s is specular roughness term, $\mathcal{G}$ is Smith's geometric attenuation function of shadowing/masking term [Heitz 2014], and ρ_s is the specular albedo.

Single scattering term. The practical single scattering term, on the other hand, is defined as

$$\mathbf{P}_{ss} \approx \kappa_{ss}\mathbf{C}_{h\to o}(-\tilde{\varphi}_o)\mathbf{F}^{\mathcal{R}}(\theta_d;\eta)\mathbf{C}_{i\to h}(\tilde{\varphi}_i), \tag{6}$$

where $\kappa_{ss} = \rho_{ss}\frac{\mathcal{D}(\theta_h;\alpha_{ss})\mathcal{G}(\theta_i,\theta_o;\alpha_{ss})}{4(\mathbf{n}\cdot\boldsymbol{\omega}_i)(\mathbf{n}\cdot\boldsymbol{\omega}_o)}$ is the single scattering reflection term, and α_{ss} and ρ_{ss} represent roughness and albedo of the single scattering term, respectively.

Subsurface scattering term. Refer to the main paper.

2.5 Human Skin Rendering with Subsurface Scattering

For translucent materials, exitant radiance $L_o(\mathbf{x}_o,\boldsymbol{\omega}_o)$ is computed by convolving the incident light $L_i(\mathbf{x}_i,\boldsymbol{\omega}_i)$ with a bidirectional scattering surface reflectance distribution function (BSSRDF) Ψ [Nicodemus et al. 1977]:

$$L_o(\mathbf{x}_o,\boldsymbol{\omega}_o) = \int_A\int_{2\pi}\Psi(\mathbf{x}_i,\boldsymbol{\omega}_i;\mathbf{x}_o,\boldsymbol{\omega}_o)L_i(\mathbf{x}_i,\boldsymbol{\omega}_i)(\mathbf{n}\cdot\boldsymbol{\omega}_i)d\boldsymbol{\omega}_i dA(\mathbf{x}_i). \tag{7}$$

Donner and Jensen [2005] approximate the BSSRDF of multi-layered translucent homogeneous materials using the multipole diffusion model:

$$\Psi(\mathbf{x}_i, \boldsymbol{\omega}_i; \mathbf{x}_o, \boldsymbol{\omega}_o) = \frac{1}{\pi}\mathcal{T}_i^+(\boldsymbol{\omega}_i; \eta_i) R(||\mathbf{x}_i - \mathbf{x}_o||) \mathcal{T}_o^+(\theta_o; \eta_o), \tag{8}$$

where R is the diffuse reflectance profile and $\mathcal{T}_i^+$ and $\mathcal{T}_o^+$ are the Fresnel transmittance at the incident point $\mathbf{x}_i$ and the exitant point $\mathbf{x}_o$.

Given the absorption coefficients σ_a, reduced scattering coefficients $\sigma_s^{'}$, refractive index η, and the thickness of the outer layer d, the multipole diffusion approximation gives the forward reflectance profile $R_{\text{out}}^{\text{f}}$ and forward transmittance profile $T_{\text{out}}^{\text{f}}$ of the outer layer

$$R_{\text{out}}^{\text{f}}(r) = \sum\nolimits_{k=-n}^{n} \left(P(\sigma_{\text{tr}}, z_{r,k}) - P(\sigma_{\text{tr}}, z_{v,k})\right), \tag{9}$$

$$T_{\text{out}}^{\text{f}}(r) = \sum\nolimits_{k=-n}^{n} \left(P(\sigma_{\text{tr}}, d - z_{r,k}) - P(\sigma_{\text{tr}}, d - z_{v,k})\right), \tag{10}$$

where $z_{r,k}$ and $z_{v,k}$ are the positions of the k-th positive and negative point sources, respectively. $P(\sigma, z) = \frac{\alpha' \cdot z(1+\sigma \cdot d_z)}{4\pi d_z^3} e^{-\sigma \cdot d_z}$ is the influence by the point source. $d_z = \sqrt{r^2 + z^2}$ is the distance between the surface of the object and the point source. $\alpha' = \sigma_s'/\sigma_t'$ is the reduced albedo, $\sigma_{\text{tr}} = \sqrt{3\sigma_a \sigma_t'}$ is the effective transport coefficient, and $\sigma_t' = \sigma_a + \sigma_s'$ is the reduced extinction coefficient.

By solving the boundary conditions about the extrapolated boundaries using a multipole expansion, $(2n + 1)$ multipoles are placed as

$$\begin{aligned} z_{r,k} &= 2k(d + z_b(0) + z_b(d)) + l, \\ z_{v,k} &= 2k(d + z_b(0) + z_b(d)) - l - 2z_b(0), \end{aligned} \tag{11}$$

where $l = 1/\sigma_t'$ is the mean free path, $z_b(0) = 2A(0)D$ and $z_b(d) = 2A(d)D$ are extrapolation distances at depth $z = 0$ and $z = d$, respectively. $A(0) = \frac{1+F(0)_{\text{dr}}}{1-F(0)_{\text{dr}}}$ is the change due to internal reflection at the surface, $D = 1/3\sigma_t'$ is the diffusion constant, and $F(0)_{\text{dr}}$ is average Fresnel reflectance [Egan et al. 1973]:

$$F(0)_{\text{dr}} \approx \begin{cases} -0.4399 + \frac{0.7099}{\eta(0)} - \frac{0.3319}{\eta^2(0)} + \frac{0.0636}{\eta^3(0)}, & \eta(0) < 1 \\ -\frac{1.440}{\eta^2(0)} + \frac{0.710}{\eta(0)} + 0.668 + 0.0636\eta(0), & \eta(0) > 1 \end{cases} \tag{12}$$

where $\eta(0)$ is the relative refractive index over surface $z = 0$.

For the backward reflectance and transmittance profiles of the outer layer, we can simply swap the upper and lower surfaces. Forward reflectance profile at the inner layer can be computed by assuming the thickness of the layer is infinite $d = \infty$ and dipole approximation $n = 0$ using Equation (9).

Convolutional form of rendering equation. Computing the analytic form of bidirectional scattering reflectance is too expensive, so Donner et al. [2008] propose an efficient method that approximates the reflectance and transmittance profiles of multi-layered heterogeneous materials by constraining the variation of parameters to be slow relative to the mean free path, which means that properties are locally homogeneous. The efficiency of this formulation of skin rendering is especially important in our iterative optimization framework.

Given the incident flux $\Phi(\mathbf{x}_i, \boldsymbol{\omega}_i)$ at a surface point which can be precomputed by the incident radiance in Equation (7), we can compute the radiant emittance profile, $M(\mathbf{x}_o)$, by convolving the incident flux Φ with the reflectance profile $R_{\mathbf{x}_o}$ at exitant point $\mathbf{x}_o$:

$$M(\mathbf{x}_o) = \iint \Phi(\mathbf{x}_i, \boldsymbol{\omega}_i) R_{\mathbf{x}_o}(||\mathbf{x}_o - \mathbf{x}_i||) dA = \Phi * R_{\mathbf{x}_o}. \tag{13}$$

As opposed to the homogeneous case, the convolution of layer responses in the heterogeneous model depends on the local position on the interface between the layers. For example, at point $\mathbf{x}_o$, the convolution of the heterogeneous profiles of the forward transmission of the outer layer, $T^{\mathrm{f}}_{\mathrm{out}}$, and the reflectance of the inner layer, $R^{\mathrm{f}}_{\mathrm{in}}$ results in

$$(T^{\mathrm{f}}_{\mathrm{out}} * R^{\mathrm{f}}_{\mathrm{in}})_{\mathbf{x}_o}(||\mathbf{x}_o - \mathbf{x}||) = \int T^{\mathrm{f}}_{\mathrm{out},\mathbf{x}_i}(||\mathbf{x}_i - \mathbf{x}||) R^{\mathrm{f}}_{\mathrm{in},\mathbf{x}_o}(||\mathbf{x}_o - \mathbf{x}_i||) d\mathbf{x}_i, \tag{14}$$

which depends on the convolution of the profile of the second layer $R^{\mathrm{f}}_{\mathrm{in},\mathbf{x}_o}$ at $\mathbf{x}_o$ with the transmittance responses of the first layer $T^{\mathrm{f}}_{\mathrm{out},\mathbf{x}_i}$ over all local positions on the interface $\mathbf{x}_i$. Note that $R^{\mathrm{f}}_{\mathrm{in},\mathbf{x}_o}$ and $T^{\mathrm{f}}_{\mathrm{out},\mathbf{x}_i}$ are the profiles at each $\mathbf{x}_o$ and $\mathbf{x}_i$.

Finally, the heterogeneous multi-layered forward reflectance profile R can be computed by accounting for the sum of multiple inter-scattering between the two heterogeneous layers:

$$R = R^{\mathrm{f}}_{\mathrm{out}} + \sum_{i=0}^{n} T^{\mathrm{f}}_{\mathrm{out}} * R^{\mathrm{f}}_{\mathrm{in}} * [R^{\mathrm{b}}_{\mathrm{out}} * R^{\mathrm{f}}_{\mathrm{in}}]^i * T^{\mathrm{b}}_{\mathrm{out}}. \tag{15}$$

To efficiently compute the profiles, d'Eon et al. [2007] use the sum of separable Gaussian functions as an accurate approximation for radially symmetric profiles by minimizing the following equation:

$$\min_{w_j} \int_0^\infty r \left(\{T, R\}^{\{\mathrm{f,b}\}}_{\{\mathrm{in,out}\}}(r) - \sum_{j=1}^{m} w_j G(v_j, r) \right)^2 dr, \tag{16}$$

where v_j and w_j are the variance and weight, respectively, of the Gaussian function $G(v, r) = \frac{1}{2\pi v} e^{-r^2/(2v)}$. After optimization, we can approximate our radially symmetric profiles as the sum of separable Gaussian functions:

$$\{T, R\}^{\{\mathrm{f,b}\}}_{\{\mathrm{in,out}\}}(r) \approx \sum_{j=1}^{m} w_j G(v_j, r). \tag{17}$$

The convolution of separable Gaussian functions can be implemented as two 1D convolutions, which is much more efficient. We also follow Donner et al. [2008] in representing each profile using a fixed set of Gaussians, where the variance of the Gaussian sets is a power of $4^n v_0$, where the initial variance v_0 is 0.01^2 mm, from the mean free path in the outer layer. This results in the following equivalence:

$$\{G(v_0), G(v_0) * G(v_0) \cdots\} = \{G(v_0), G(4v_0), G(4^2 v_0), \cdots\}, \tag{18}$$

which allows us to compute the convolution of the next wider Gaussian function from the results of the previous narrow Gaussian function.

3 Polarimetric Imaging

3.1 Polarimetric Image Formation Detail

We now describe a new coaxial image formation designed for the polarimetric BSSRDF model. In our system, our light sources are equipped with a linear polarizer so that the incident light is linearly polarized, with the Stokes vector being $\mathbf{s}_i = [1, 1, 0, 0]^\intercal$. The Stokes vector $\mathbf{s}_o$ reflected from a surface point can be expressed as

$$\mathbf{s}_o = S\mathbf{P}\mathbf{s}_i = S \begin{bmatrix} \kappa_s \mathcal{R}^+ + \kappa_{ss}\mathcal{R}^+ + \sum_{\mathbf{x}_i \in \mathcal{S}} \rho_{sss}(\mathcal{T}^+\mathcal{T}^+ - \mathcal{T}^-\mathcal{T}^+\xi) \\ \kappa_s \mathcal{R}^+ + \kappa_{ss}\mathcal{R}^+ - \sum_{\mathbf{x}_i \in \mathcal{S}} \rho_{sss}\mathcal{T}^-\mathcal{T}^+\xi \\ - \sum_{\mathbf{x}_i \in \mathcal{S}} \rho_{sss}\mathcal{T}^-\mathcal{T}^+\zeta \\ 0 \end{bmatrix}, \tag{19}$$

where $S = (\mathbf{n} \cdot \boldsymbol{\omega}_i)/\Gamma^2$ is the shading term with attenuation, Γ is the distance between the light source and the surface, κ_s is the specular reflection term of $\rho_s \frac{\mathcal{D}(\theta_h;\alpha_s)\mathcal{G}(\theta_i,\theta_o;\alpha_s)}{4(\mathbf{n}\cdot\boldsymbol{\omega}_i)(\mathbf{n}\cdot\boldsymbol{\omega}_o)}$, κ_{ss} is the single scattering reflection term of $\rho_{ss} \frac{\mathcal{D}(\theta_h;\alpha_{ss})\mathcal{G}(\theta_i,\theta_o;\alpha_{ss})}{4(\mathbf{n}\cdot\boldsymbol{\omega}_i)(\mathbf{n}\cdot\boldsymbol{\omega}_o)}$, $\xi = \cos(2\phi)$, and $\zeta = \sin(2\phi)$.

The multi-layered subsurface scattering light interaction events lead to depolarization. As a result, the difference between Fresnel transmittances for parallel and perpendicular polarized light in both incoming and outgoing directions approaches zero: $\mathcal{T}_o^- \mathcal{T}_i^- \approx 0$. In addition, a near-coaxial setup allows for convenient simplifications in our polarimetric reflectance model [Baek et al. 2018; Hwang et al. 2022]. Geometrically, a coaxial setup results in $\boldsymbol{\omega}_i \approx \boldsymbol{\omega}_o$, $\phi_i \approx \pi - \phi_o$, $\varphi_i \approx 2\pi - \varphi_o$, $\zeta_i \approx -\zeta_o$, and $\xi_i \approx \xi_o$. In addition, the incident angle is, by definition, below the Brewster angle, so $\cos\delta = -1$ for both the specular and the single scattering terms, (where δ is the phase shift delay, $\delta = 0$ when the incident angle is larger than the Brewster angle, $\delta = \pi$ otherwise). Last, while we define Fresnel reflection coefficients as $\mathcal{R}^+ = (\mathcal{R}^\perp + \mathcal{R}^\parallel)/2$, $\mathcal{R}^\times = \sqrt{\mathcal{R}^\perp \mathcal{R}^\parallel}$, and $\mathcal{R}^- = (\mathcal{R}^\perp - \mathcal{R}^\parallel)/2$ ($\mathcal{R}^\perp$ and $\mathcal{R}^\parallel$ being the perpendicular and parallel components, respectively), in a near-coaxial setup the parallel and perpendicular Fresnel reflection coefficients become very similar, thus $\mathcal{R}^\parallel \approx \mathcal{R}^\perp$; this results in $\mathcal{R}^- \approx 0$ and $\mathcal{R}^+ \approx \mathcal{R}^\times$.

Our simplified version of the Mueller matrix thus becomes

$$\mathbf{P} \approx \begin{bmatrix} \bar{\kappa}_{s,ss}\mathcal{R}^+ + \sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}\mathcal{T}^{++} & -\sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}\mathcal{T}^{-+}\xi & \sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}\mathcal{T}^{-+}\zeta & 0 \\ -\sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}\mathcal{T}^{-+}\xi & \bar{\kappa}_{s,ss}\mathcal{R}^+ & 0 & 0 \\ -\sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}\mathcal{T}^{-+}\zeta & 0 & -\bar{\kappa}_{s,ss}\mathcal{R}^+ & 0 \\ 0 & 0 & 0 & -\bar{\kappa}_{s,ss}\mathcal{R}^+ \end{bmatrix}, \tag{20}$$

where $\bar{\kappa}_{s,ss} = \kappa_s + \kappa_{ss}$ is the sum of the specular and single scattering reflection terms, $\mathcal{T}^{++} = \mathcal{T}^+\mathcal{T}^+$ is the multiplication of the positive Fresnel transmission coefficients, and $\mathcal{T}^{-+} = \mathcal{T}^-\mathcal{T}^+$ is the multiplication of the negative/positive coefficients.

We then capture the reflected light with a polarization camera that outputs the image $\mathbf{I}$ corresponding to four linear-polarization angles as

$$\begin{aligned} \mathbf{I} = \begin{bmatrix} I_0 \\ I_{90} \\ I_{45} \\ I_{135} \end{bmatrix} &= \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix} \mathbf{s}_o \\ &= \frac{S}{2} \begin{bmatrix} 2\kappa_s\mathcal{R}^+ + 2\kappa_{ss}\mathcal{R}^+ + \sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}(\mathcal{T}^+\mathcal{T}^+ - 2\mathcal{T}^-\mathcal{T}^+\xi) \\ \sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}\mathcal{T}^+\mathcal{T}^+ \\ \kappa_s\mathcal{R}^+ + \kappa_{ss}\mathcal{R}^+ + \sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}(\mathcal{T}^+\mathcal{T}^+ - \mathcal{T}^-\mathcal{T}^+\xi - \mathcal{T}^-\mathcal{T}^+\alpha) \\ \kappa_s\mathcal{R}^+ + \kappa_{ss}\mathcal{R}^+ + \sum\limits_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}(\mathcal{T}^+\mathcal{T}^+ - \mathcal{T}^-\mathcal{T}^+\xi + \mathcal{T}^-\mathcal{T}^+\alpha) \end{bmatrix}. \end{aligned} \tag{21}$$

From the captured images $I_0, I_{90}, I_{45}, I_{135}$, we extract each component of the total reflection. First, I_{90} can be used to extract the unpolarized subsurface-scattering component. We define the unpolarized subsurface scattering observation I_{sss} as

$$I_{sss} = S \sum_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}\mathcal{T}^+\mathcal{T}^+ = 2I_{90}, \tag{22}$$

where $\mathcal{S}$ is the set of the surface points $\mathbf{x}_i$ of the face.

Information about the polarized subsurface scattering term can also be obtained by subtracting I_{135} from I_{45}, which we define as a subsurface scattering polarization observation I_ζ as

$$I_\zeta = S \sum_{\mathbf{x}_i \in \mathcal{S}} \rho_{sss} \mathcal{T}^- \mathcal{T}^+ \zeta = I_{135} - I_{45}. \tag{23}$$

Lastly, subtracting I_0 by I_{90}, we can obtain a combination of specular reflection, single scattering, and oriented subsurface scattering parameters. We define this combination as the specular-dominant polarization observation I_s as

$$I_s = S(\kappa_s \mathcal{R}^+ + \kappa_{ss} \mathcal{R}^+ - \sum_{\mathbf{x}_i \in \mathcal{S}} \rho_{sss} \mathcal{T}^- \mathcal{T}^+ \xi) = I_0 - I_{90}. \tag{24}$$

3.2 Spectral/Geometry Calibration of the System

In order to capture the spectral reflectance information of the human face, we calibrate each Dolby lens transmittance and the polarized camera response function. We use a spectrometer capture device (JETI) with a white Spectralon (99% reflectance) to first capture the transmittance of Dolby lenses by dividing the spectral distributions of the transmitted light by the original light. For the camera response function, we select one of the polarization cameras and capture the Spectralon images, lit by LED lights equipped with a liquid crystal tunable filter (LCTF), which transmits a selected wavelength band. First, we estimate the spectral transmittance of the LCTF filter (similarly to the Dolby lens). Then, we capture the images every 10 nm in range 420 nm – 670 nm. To calibrate across polarization cameras, we capture an image by placing a sphere-shaped Spectralon at the location where the face will be captured. Then, we normalize the captured value to the predicted camera response function with light. For color cameras, we use a color checker to calibrate the color camera by a 3×3 matrix. To calibrate our camera's intrinsic and extrinsic parameters, we use a ChArUco checkerboard. We capture multiple images of varying checkerboard poses and minimize the reprojection errors.

4 Computation Details

4.1 Computing Normals from Heights

The non-unit normal vector $\tilde{\mathbf{n}}$ at each pixel is computed from the displacement map H [Riviere et al. 2020] as

$$\begin{aligned} \tilde{\mathbf{n}} &= (\hat{s}_\mathrm{u} \hat{\mathbf{t}}_\mathrm{u} + \frac{\partial H}{\partial \mathrm{u}} \hat{\mathbf{n}}) \times (\hat{s}_\mathrm{v} \hat{\mathbf{t}}_\mathrm{v} + \frac{\partial H}{\partial \mathrm{v}} \hat{\mathbf{n}}) \\ &= \begin{bmatrix} \hat{\mathbf{t}}_\mathrm{u} & \hat{\mathbf{t}}_\mathrm{v} & \hat{\mathbf{n}} \end{bmatrix} \begin{bmatrix} \hat{s}_\mathrm{u} & 0 & 0 \\ 0 & \hat{s}_\mathrm{v} & 0 \\ 0 & 0 & \hat{s}_\mathrm{u}\hat{s}_\mathrm{v} \end{bmatrix} \begin{bmatrix} -\frac{\delta H}{\delta \mathrm{u}} \\ -\frac{\delta H}{\delta \mathrm{v}} \\ 1 \end{bmatrix}, \end{aligned} \tag{25}$$

where $\hat{\mathbf{t}}_\mathrm{u}$ and $\hat{\mathbf{t}}_\mathrm{v}$ are the finite differences in u and v directions of the tangent vector $\mathbf{t}$ of the initial mesh. $\hat{\mathbf{n}}$ is the unit normal vector of the mesh at the pixel, $\hat{s}_\mathrm{u}$ and $\hat{s}_\mathrm{v}$ are the original lengths of tangent vectors of the initial mesh. We then normalize $\tilde{\mathbf{n}}$ to obtain a unit vector.

5 Optimization Details

5.1 Polarimetric Inverse Rendering Details

For the first polarimetric inverse rendering step, specifically, we minimize the following energy function:

$$\min_{\eta, \alpha_s, \alpha_{ss}, \rho_s, \rho_{ss}, \bar{\rho}_{sss}, H} \lambda_\psi \mathcal{L}_\psi + \lambda_{sss} \mathcal{L}_{sss} + \lambda_s \mathcal{L}_s + \lambda_\phi \mathcal{L}_\phi + \mathcal{L}_\mathrm{reg}, \tag{26}$$

where $\mathcal{L}_\psi$ is the refractive index loss, $\mathcal{L}_{sss}$ is the subsurface scattering loss, $\mathcal{L}_s$ is the specular and single scattering loss, $\mathcal{L}_\phi$ is the azimuthal loss, and $\mathcal{L}_{\text{reg}}$ is the regularizer term, $\lambda_\psi = 0.002$, $\lambda_{sss} = 1$, $\lambda_s = 1$, $\lambda_\phi = 1$ are the weights assigned to each loss, respectively.

Subsurface scattering loss. We formulate a photometric loss of subsurface scattering $\mathcal{L}_{sss}$ by evaluating the rendered subsurface scattering image $\hat{I}^t_{sss}$ with the captured image I^t_{sss} at each frame t of multiview input:

$$\mathcal{L}_{sss} = \sum_t \mathrm{V}^t \left(\hat{I}^t_{sss} - I^t_{sss}\right)^2, \tag{27}$$

where V^t is the visibility texture map at frame t for each view. The visibility map V^t is 1 at the visible pixel region and 0 otherwise.

Solving the subsurface scattering optimization problem directly is computationally expensive and ill-posed. And thus, as mentioned in the main paper, we break the optimization process into two steps to make it more manageable. In the first step, we make the reasonable assumption that the Fresnel transmittance of human skin does not change dramatically across the surface, as the refractive index and roughness of the skin typically change smoothly over the surface. Based on this assumption, we approximate the subsurface scattering reflectance as $I^t_{sss} = S\bar{\rho}_{sss}\mathcal{T}^+\mathcal{T}^+$, where $\bar{\rho}_{sss}$ implicitly encompasses the approximated overall observation of subsurface scattering effects at the exitant point originated from multiple incident locations.

In the following second stage, we use the optimized surface scattering reflectance $\bar{\rho}_{sss}$ and further decompose this value using our novel inverse subsurface scattering optimization method. This two-step approach allows us to address the complex problem of subsurface scattering optimization in a more efficient and comprehensive manner.

Specular and single scattering loss. Current methods for human face skin modeling [Ghosh et al. 2008; Ma et al. 2007; Riviere et al. 2020] and pBRDF optimization [Baek and Heide 2021; Baek et al. 2018; Hwang et al. 2022] often require augmentation or clustering techniques to compensate for the limited number of specular samples per texel when determining specular and single scattering parameters. Thanks to our stereo imaging module, we can obtain a dense set of light-view samples for each texel by merging all video sequence frames into the reference frame as participants rotate their heads. This approach enables a more comprehensive analysis of skin and pBRDF properties, eliminating the need for augmentation or clustering. We formulate the specular and single scattering loss as

$$\mathcal{L}_s = \sum_t \mathrm{V}^t \left(\hat{I}^t_s - I^t_s\right)^2, \tag{28}$$

where $\hat{I}^t_s$ is computed using Equation (24) by using $\bar{\rho}_{sss}\mathcal{T}^-\mathcal{T}^+\xi$ instead of $\sum_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}\mathcal{T}^-\mathcal{T}^+\xi$ and $\xi = \cos(2\phi)$.

Refractive index loss. The refractive index loss is particularly relevant because it globally affects appearance at multiple levels, and our work provides a spatially-varying index of refraction from images. We adopt the refractive-index loss from Hwang et al. [2022] that formulates the degree of polarization (DoP) of the multi-layered subsurface scattering reflections $\psi = |\mathcal{T}^-/\mathcal{T}^+|$ using unpolarized subsurface scattering image I_{sss}, subsurface scattering polarization image I_ζ, and specular polarization image I_s as

$$\psi = \left|\sqrt{(I_\zeta)^2 + (I_\xi)^2}/I_{sss}\right|, \tag{29}$$

where $I_\xi = I_s - \kappa_s S\mathcal{R}^+ - \kappa_{ss} S\mathcal{R}^+ = -S\bar{\rho}_{sss}\mathcal{T}^-\mathcal{T}^+\xi$. With this observed DoP, the refractive index loss term becomes

$$\mathcal{L}_\psi = \sum_t \mathrm{V}^t \left(\hat{\psi}(\eta, \theta^t_o) - \psi^t\right)^2, \tag{30}$$

where $\hat{\psi}$ is the predicted DoP value, which can be formulated using refractive index η and the surface zenith angle θ_o^t [Atkinson and Hancock 2006]. The value of η is only optimized at the static initialization stage, but this loss term is also influenced by the local geometry defined by the displacement map H, which is updated at every frame.

Azimuthal loss. We implement the azimuthal loss of shape from polarization as proposed by Hwang et al. [2022]:

$$\mathcal{L}_\phi = \sum_{t=1} \mathrm{V}^t \mathrm{W}_\phi^t \left((\hat{I}_\zeta^t - I_\zeta^t)^2 + (\hat{I}_\xi^t - I_\xi^t)^2 \right), \tag{31}$$

where $\hat{I}_\zeta^t = S\bar{\rho}_{sss}\mathcal{T}^-\mathcal{T}^+\hat{\zeta}^t$ and $\hat{I}_\xi^t = -S\bar{\rho}_{sss}\mathcal{T}^-\mathcal{T}^+\hat{\xi}^t$ denote the diffuse polarized images obtained by optimized azimuth angles $\hat{\zeta}^t = \sin(2\hat{\phi}^t)$ and $\hat{\xi}^t = \cos(2\hat{\phi}^t)$ at frame t, respectively. Note that diffuse polarization can be computed as $S\bar{\rho}_{sss}\mathcal{T}^-\mathcal{T}^+ = \sqrt{(I_\zeta^t)^2 + (I_\zeta^t)^2}$. The initial geometry extracted from multi-view stereo effectively resolves the π ambiguity of shape from polarization [Atkinson and Hancock 2006; Kadambi et al. 2015]. We calculate the weight matrix W_ϕ^t by determining the normalized mean value of I_{sss}.

Regularization loss. Our regularization loss term is designed to preserve spatial and temporal consistency and is formulated as

$$\mathcal{L}_{\mathrm{reg}} = \lambda_{H_{\mathrm{treg}}}\mathcal{L}_{H_{\mathrm{treg}}} + \lambda_{H_{\mathrm{sreg}}}\mathcal{L}_{H_{\mathrm{sreg}}} + \lambda_{\alpha_s}\mathcal{L}_{\alpha_s} + \lambda_{\alpha_{ss}}\mathcal{L}_{\alpha_{ss}} + \lambda_\eta\mathcal{L}_\eta, \tag{32}$$

where $\mathcal{L}_{H_{\mathrm{treg}}}$ and $\mathcal{L}_{H_{\mathrm{sreg}}}$ represent temporal and spatial regularization losses for the displacement map, $\mathcal{L}_{\alpha_s}$, $\mathcal{L}_{\alpha_{ss}}$, and $\mathcal{L}_\eta$ are correspond to spatial regularization losses for specular, single scattering roughness, and refractive index, $\lambda_{H_{\mathrm{treg}}} = 1$, $\lambda_{H_{\mathrm{sreg}}} = 1000$, $\lambda_{\alpha_s} = 200$, $\lambda_{\alpha_{ss}} = 200$, $\lambda_\eta = 400$ are the respective weights assigned to each loss.

To preserve the geometry information of our optimized mesh relative to the initial geometry, we apply a temporal regularization loss term to our displacement map:

$$\mathcal{L}_{H_{\mathrm{treg}}} = H^2. \tag{33}$$

For spatial smoothness, we use the Laplacian operator on the displacement map:

$$\mathcal{L}_{H_{\mathrm{sreg}}} = \sum_{\mathbf{x}\in\mathcal{S}} \left(\nabla^2 H(\mathbf{x})\right)^2, \tag{34}$$

where $\mathcal{S}$ represents the valid texture region containing the human face surface.

We assume that local spatial variations in roughness on the human face are minimal, although significant differences can be observed between distinct regions. The local variation of specularity mainly originates from variations in the specular albedo. Additionally, some specific pixels may not have a sufficient number of observations to estimate the parameters. We formulate the spatial smoothness term for the refractive index and the roughness parameter of both specular and single scattering as

$$\begin{aligned}
\mathcal{L}_{\alpha_s} &= \sum_{\mathbf{x}\in\mathcal{S}} \left(\alpha_s(\mathbf{x}) - \bar{\alpha}_s(\mathbf{x})\right)^2, \\
\mathcal{L}_{\alpha_{ss}} &= \sum_{\mathbf{x}\in\mathcal{S}} \left(\alpha_{ss}(\mathbf{x}) - \bar{\alpha}_{ss}(\mathbf{x})\right)^2, \\
\mathcal{L}_\eta &= \sum_{\mathbf{x}\in\mathcal{S}} \left(\eta(\mathbf{x}) - \bar{\eta}(\mathbf{x})\right)^2,
\end{aligned} \tag{35}$$

where $\bar{\eta}(\mathbf{x})$ is the average refractive index values of the neighboring pixels of $\mathbf{x}$, $\bar{\alpha}_s(\mathbf{x})$ and $\bar{\alpha}_{ss}(\mathbf{x})$ are the average specular and single scattering roughness values of the neighboring pixels of $\mathbf{x}$, respectively. We employ a 5×5 window to calculate the average pixel value.

5.2 Dynamic Inverse Rendering Details

By the given roughness parameter from the static reconstruction, we solve the following energy function to estimate the other parameters in the dynamic capture per each frame t:

$$\min_{\rho_s^t,\rho_{ss}^t,\rho_{sss}^t,H^t} \tilde{\lambda}_{sss}\mathcal{L}_{sss} + \tilde{\lambda}_s\mathcal{L}_s + \tilde{\lambda}_\phi\mathcal{L}_\phi + \tilde{\mathcal{L}}_{\text{reg}}, \tag{36}$$

where $\mathcal{L}_{sss}$, $\mathcal{L}_s$, $\mathcal{L}_\phi$ are inherited from the static capture loss (Equation (26)). $\tilde{\lambda}_{sss}=1$, $\tilde{\lambda}_s=1$, $\tilde{\lambda}_\phi=0.2$ are the weights assigned to each loss, respectively.

Here, we defined a dynamic regularization term $\tilde{\mathcal{L}}_{\text{reg}}$ as

$$\tilde{\mathcal{L}}_{\text{reg}} = \tilde{\lambda}_{H^t_{\text{treg}}}\tilde{\mathcal{L}}_{H^t_{\text{treg}}} + \tilde{\lambda}_{H^t_{\text{sreg}}}\tilde{\mathcal{L}}_{H^t_{\text{sreg}}} + \tilde{\lambda}_{\rho_s^t}\tilde{\mathcal{L}}_{\rho_s^t} + \tilde{\lambda}_{\rho_{ss}^t}\tilde{\mathcal{L}}_{\rho_{ss}^t} + \tilde{\lambda}_{\bar{\rho}_{sss}^t}\tilde{\mathcal{L}}_{\bar{\rho}_{sss}^t}, \tag{37}$$

where $\tilde{\mathcal{L}}_{H^t_{\text{treg}}}$ and $\tilde{\mathcal{L}}_{H^t_{\text{sreg}}}$ are the dynamic temporal and spatial regularization loss for displacement map which are similar to Equations (33) and (34), $\tilde{\mathcal{L}}_{\rho_s^t}$, $\tilde{\mathcal{L}}_{\rho_{ss}^t}$, $\tilde{\mathcal{L}}_{\bar{\rho}_{sss}^t}$ are the temporal regularization loss term for specular, single scattering, and subsurface scattering, and $\tilde{\lambda}_{H^t_{\text{treg}}}=0.001$, $\tilde{\lambda}_{H^t_{\text{sreg}}}=500$, $\tilde{\lambda}_{\rho_s^t}=0.01$, $\tilde{\lambda}_{\rho_{ss}^t}=0.01$, and $\tilde{\lambda}_{\bar{\rho}_{sss}^t}=0.01$ are the weights assigned to each loss, respectively.

Our temporal regularization term for specular and single scattering intensity prevents flickering artifacts in the sequence:

$$\tilde{\mathcal{L}}_{\rho_s^t} = \sum_{t=1}\left(\rho_s^t-\rho_s^0\right)^2, \tilde{\mathcal{L}}_{\rho_{ss}^t} = \sum_{t=1}\left(\rho_{ss}^t-\rho_{ss}^0\right)^2. \tag{38}$$

In short-term dynamic sequences, changes in the color of human skin are mainly caused by variations in the hemoglobin ratio, which affects the chromaticity of the skin color. We incorporated it into our temporal subsurface scattering regularization term to minimize the difference between the albedo of the static results and that of the current frame, weighted by the intensity of the albedo:

$$\tilde{\mathcal{L}}_{\rho_{sss}^t} = \sum_{t=1} W^t_{\bar{\rho}_{sss}}\left(\bar{\rho}_{sss}^t-\bar{\rho}_{sss}^0\right)^2, \tag{39}$$

where $W^t=|\dot{\bar{\rho}}_{sss}^0-\dot{\bar{\rho}}_{sss}^t|$ is the weight map which is computed by the difference between the intensity of the average subsurface scattering in the static results $\dot{\bar{\rho}}_{sss}^0$ and the intensity of the average subsurface scattering in the current frame $\dot{\bar{\rho}}_{sss}^t$. Finally, using the estimated average subsurface scattering reflectance $\bar{\rho}_{sss}^t$ at the frame, we optimize the face parameters, which are the same as the static scene reconstruction.

5.3 Optimization of Biophysically-based Parameters Details

In order to optimize biophysically-based parameters using photometric loss from rendering, we propose a coordinate descent method [Wright 2015] using alternating least squares, designed to make this optimization problem manageable. We split our optimization problem into two. The first subproblem is to obtain the spectral weights w_j of the Gaussian functions from the reflectance and transmittance diffusion profiles with initial variables as

$$\min_{w_j}\int_0^\infty\left(\{T,R\}^{\{\text{f,b}\}}_{\{\text{in,out}\}}(r)-\sum_{j=1}^m w_jG(v_j,r)\right)^2 dr. \tag{40}$$

The second subproblem is to optimize the biophysical parameters from the spectral observation $\bar{\rho}_{sss}$ obtained from polarimetric inverse rendering with ρ_{sss}:

$$\min_{C_{\text{hd}},C_{\text{he}},C_{\text{m}},\beta_{\text{m}}}\left(\bar{\rho}_{sss}-\rho_{sss}\right)^2, \tag{41}$$

as rendered with the approximated sum of separable Gaussians.

To render the subsurface scattering component with optimizing variables, we formulate the total reflectance $\bar{R}$ (or transmittance $\bar{T}$) of each profile as the sum of Gaussians (SoG) as described in Section 2.5:

$$\sum_{j=0} w_{i,\mathbf{x}_o,j} = \bar{R}_{i,\mathbf{x}_o} = 2\pi \int_0^\infty R_{i,\mathbf{x}_o}(r) r dr \tag{42}$$
$$= \sum_{k=-n}^{n} \left(\mathrm{sign}(z_{r,k}) e^{-\sigma_{\mathrm{tr}}|z_{r,k}|} - \mathrm{sign}(z_{v,k}) e^{-\sigma_{\mathrm{tr}}|z_{v,k}|} \right),$$

where $w_{i,\mathbf{x}_o,j}$ is the weight of j-th variance of the i-th layer's SoG at the exitant pixel $\mathbf{x}_o$, r is the distance between the incident surface point and the exitant surface point, sign() is the sign function, and σ_{tr} is the effective transport coefficient. $z_{r,k}$ and $z_{v,k}$ are the positions of the k-th positive and negative point sources in Equation (11), respectively.

Using the total reflectance (or transmittance), we can rephrase the SoG by approximately convolving total reflectance using the normalized SoG as

$$G_{R_{i,\mathbf{x}_o}}(r) = \sum_{j=0} w_{i,\mathbf{x}_o,j} G(v_j, r) \approx \bar{R}_{i,\mathbf{x}_o} \sum_{j=0} \bar{w}_{i,\mathbf{x}_o,j} G(v_j, r) \tag{43}$$
$$= \bar{R}_{i,\mathbf{x}_o} G_{\bar{R}_{i,\mathbf{x}_o}}(r),$$

where $\sum_{j=0} \bar{w}_{i,\mathbf{x}_o,j} = 1$. We can approximate this equation similarly to texture blurring in subsurface scattering rendering methods [d'Eon et al. 2007; Donner and Jensen 2005; Jensen et al. 2001]:

$$R_{i,\mathbf{x}_o} \approx \int \bar{R}_{i,x} G_{\bar{R}_{i,\mathbf{x}_o}}(||\mathbf{x}_o - \mathbf{x}||) d\mathbf{x} = \bar{R}_i * G_{\bar{R}_{i,\mathbf{x}_o}}. \tag{44}$$

Ghosh et al. [2008] measure the translucency of the human skin using a contact probe, and they show that it does not significantly vary spatially. Moreover, state-of-the-art face acquisition methods that consider the blurring due to the subsurface scattering [Riviere et al. 2020] and the human skin rendering techniques [d'Eon et al. 2007; Jimenez et al. 2015] also use a fixed parameter of blurriness.

Following these observations, we assume that the level of blurriness is spatially homogeneous:

$$\bar{R}_{i,\mathbf{x}_o} \sum_{j=0} \bar{w}_{i,\mathbf{x}_o,j} G(v_j) \approx \bar{R}_{i,\mathbf{x}_o} \sum_{j=0} \bar{w}_{i,j} G(v_j) = \bar{R}_{i,\mathbf{x}_o} G_{\bar{R}_i}, \tag{45}$$

where $G_{\bar{R}_i} = \sum_{j=0} \bar{w}_{i,j} G(v_j)$. Then, finally, we can approximate the subsurface scattering of the human skin $\tilde{R}$ as

$$\tilde{R} = \bar{R}^{\mathrm{f}}_{\mathrm{out}} * G_{\bar{R}^{\mathrm{f}}_{\mathrm{out}}} + ((\bar{T}^{\mathrm{f}}_{\mathrm{out}} * G_{\bar{T}^{\mathrm{f}}_{\mathrm{out}}}) \cdot \bar{R}^{\mathrm{f}}_{\mathrm{in}} * G_{\bar{R}^{\mathrm{f}}_{\mathrm{in}}}) \cdot \bar{T}^{\mathrm{b}}_{\mathrm{out}} * G_{\bar{T}^{\mathrm{b}}_{\mathrm{out}}} + \cdots . \tag{46}$$

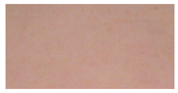
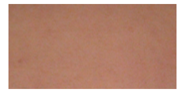
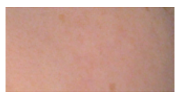
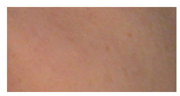
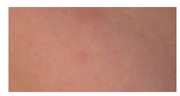
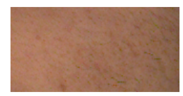
We use gradient descent optimization to acquire the face skin parameters. To estimate the SoG of each profile, we first downsample the images. As the distance of the neighboring pixel becomes larger, we can ignore the spatial blurring. Then, we can render the per-pixel intensity just considering the multi-layer interaction between the total reflectance as

$$\bar{R}_{\mathbf{x}_o} = \bar{R}^{\mathrm{f}}_{\mathrm{out},\mathbf{x}_o} + \sum_{n=0} \bar{T}^{\mathrm{f}}_{\mathrm{out},\mathbf{x}_o} \bar{R}^{\mathrm{f}}_{\mathrm{in},\mathbf{x}_o} [\bar{R}^{\mathrm{b}}_{\mathrm{out},\mathbf{x}_o} \bar{R}^{\mathrm{f}}_{\mathrm{in},\mathbf{x}_o}]^n \bar{T}^{\mathrm{b}}_{\mathrm{out},\mathbf{x}_o}. \tag{47}$$

Then, we fit SoG to each diffusion profile using the median intensity pixel. At the original resolution, we minimize the difference between the rendered images using Equation (46) and the subsurface scattering albedo images from polarimetric inverse rendering.

We use a fixed-size discrete kernel for each SoG. At each pixel $\mathbf{x}_o$, we first compute the distance between the neighboring pixel, which is within the kernel size, and $\mathbf{x}_o$. Using this distance, we can compute the discrete SoG

Table 2. Estimated biophysical parameters (mean and standard deviation) of subjects on the forehead and cheek with different levels of skin tone.

Forehead					
Skin		I	II	III	IV
Photograph					
Skin param.	He. (inner)	0.03691 (0.00563)	0.02072 (0.00792)	0.03622 (0.00903)	0.02654 (0.00837)
	He. (outer)	0.01701 (0.00256)	0.08323 (0.02069)	0.07337 (0.00441)	0.10450 (0.01002)
	Melanin	0.03243 (0.00183)	0.05036 (0.00400)	0.06155 (0.00228)	0.08909 (0.00330)
	Rel. eumel.	0.09607 (0.00393)	0.04759 (0.01404)	0.05304 (0.00658)	0.05291 (0.00629)
Reflec.	Refrac. idx	1.44248 (0.01258)	1.41736 (0.01182)	1.39902 (0.00759)	1.41017 (0.00734)
	Spec. rough.	0.58061 (0.02791)	0.57004 (0.04008)	0.54546 (0.02790)	0.54927 (0.02887)
	SS. rough.	0.96929 (0.00980)	0.97965 (0.00731)	0.95140 (0.01186)	0.98458 (0.00542)
Cheek					
Skin		I	II	III	IV
Photograph					
Skin param.	He. (inner)	0.02672 (0.00645)	0.01722 (0.00643)	0.05753 (0.00970)	0.02135 (0.01118)
	He. (outer)	0.04337 (0.00637)	0.05639 (0.00652)	0.09753 (0.01057)	0.09513 (0.01790)
	Melanin	0.02891 (0.00286)	0.04335 (0.00434)	0.03993 (0.00435)	0.07982 (0.00686)
	Rel. eumel.	0.07224 (0.00821)	0.03684 (0.00971)	0.02355 (0.01014)	0.05559 (0.01228)
Reflec.	Refrac. idx	1.41323 (0.01404)	1.42730 (0.01083)	1.41453 (0.01454)	1.42888 (0.01116)
	Spec. rough.	0.51574 (0.02848)	0.64520 (0.03452)	0.53625 (0.02827)	0.58147 (0.02452)
	SS. rough.	0.96598 (0.00951)	0.95688 (0.01927)	0.97491 (0.00971)	0.97636 (0.00948)

kernel, which represents the subsurface scattering reflectance. To ensure energy conservation, we normalize the kernel.

6 Implementation Details

Optimization. To compute all the losses in each iterative optimization, we use a PyTorch RMSprop optimizer. We use a 2K×2K resolution texture to optimize whole parameters. For the static initialization stage, we use 200 frames that represent different views. We implement a patch-based gradient descent optimization with a 256×256 size of the patch. For the dynamic sequence, we estimate appearance parameters per frame, and we similarly use patch-based gradient descent optimization. Our code runs on a machine equipped with an AMD EPYC 7763 CPU of 2.45 GHz and a single NVIDIA A100 GPU. For the static initialization, the polarimetric inverse rendering takes 180 minutes (150 iterations) on 200 frames (views), and the biophysical multispectral optimization takes 50 minutes (1,000 iterations at coarse resolution and 250 full-resolution iterations). Dynamic inverse rendering takes 180 minutes for 50 frames (150 iterations) in addition to the additional biophysical multispectral optimization of 20 minutes per frame (100 iterations at coarse resolution and 100 full-resolution iterations for subsurface scattering).

7 Additional Results

In this section, we provide additional results of face acquisition.

Biophysical Parameters. We analyze the estimated biophysical parameters of the forehead and cheek areas of subjects with different levels of skin tone. Table 2 shows the estimated parameters as well as captured photographs. We show that, as expected, darker skin exhibits higher concentration levels of estimated melanin. Moreover, the estimated refractive indices of the skin fall into the range from 1.35 to 1.55, which shows a good agreement with previous biophysical studies [Anderson and Parrish 1981; Van Gemert et al. 1989].

8 Additional Discussion

Impact of multispectral polarimetric imaging. Our system utilizes multiple polarimetric cameras equipped with off-the-shelf multispectral filters. Polarimetric inverse rendering enables us to separate the components of the polarimetric reflectance function. In addition, this yields a refractive index per texel. Accurate values for this refractive index are crucial for the estimation of subsurface scattering, as it disambiguates its contribution to appearance. Then, thanks to the multispectral input, we obtain concentration maps for individual biophysical components. We observe that the combination of both polarimetric and multispectral input is effective in estimating the overall range of subsurface scattering with high accuracy.

Spatial resolution. We utilize a polarization camera with a spatial resolution of 2448×2048. However, due to four linear polarization filters and four color filters (RGBG), its effective resolution is reduced to 612×512. Although we leverage the recent proposed demosaicing algorithm [Morimatsu et al. 2020] to enhance the spatial resolution of the images in 2K, our system's overall spatial resolution is half of that provided by conventional machine vision cameras (4K). We anticipate that the spatial resolution of BSSRDFs can be significantly improved when higher-resolution polarimetric cameras become available in the future.

Near-coaxial setup for polarimetric imaging. Our coaxial imaging configuration has the potential to substantially alleviate the optimization challenges associated with polarimetric inverse rendering as evidenced by Baek et al. [2018] and Hwang et al. [2022]. However, this setup requires only one directional light to be activated when the corresponding directional camera captures the subject. In essence, this constraint prevents multiple polarimetric cameras at different orientations from capturing the subject simultaneously, allowing for a more efficient capture process.

Potential applications of photoplethysmography. Recent progress in photoplethysmography [Vilesov et al. 2022] allows for precise heart rate measurements by integrating a traditional RGB camera with radar signals. This represents a promising future research direction alongside our biophysical component measurements. However, polarization cameras, which include additional polarization filters, tend to have lower light efficiency than standard RGB cameras and often suffer from a low signal-to-noise ratio. This makes it challenging to distinguish temporal changes in skin appearance caused by heart rate fluctuations over time. As advancements continue in polarization camera technology, the prospect of using these devices for photoplethysmography presents an intriguing future research opportunity.

Table 3. Symbols and notations used in the paper.

	Symbol	Description
Vectors/Angles	$\mathbf{n}, \mathbf{h}$	Normal/halfway vector
	ϕ_i, ϕ_o	Azimuth angle between the incident/exitant light along the plane of incidence of the normal vector
	φ_i, φ_o	Azimuth angle between the incident/exitant light along the plane of the incidence of the halfway vector
	$\theta_i, \theta_o, \theta_h$	Zenith angle between the normal and the incident/exitant/halfway vector
	θ_d	Zenith angle between the incident light and the halfway vector
	$\mathbf{x}_i, \mathbf{x}_o$	Incident/exitant point
	$\boldsymbol{\omega}_i, \boldsymbol{\omega}_o$	Incident/exitant light direction
	$\zeta\{i, o\}$	$\sin(2\phi_{\{i,o\}})$
	$\xi_{\{i,o\}}$	$\cos(2\phi_{\{i,o\}})$
	δ	Retardation (delay) phase shift (0 when the incident angle is larger than the Brewster angle, π otherwise)
Polarimetry	$\mathbf{s}_i, \mathbf{s}_o$	Stokes vector of the incident/exitant light to an object surface
	$\mathbf{s}_{\text{in}}, \mathbf{s}_{\text{out}}$	Stokes vector before/after the transformation event
	$I_0, I_{90}, I_{45}, I_{135}$	0/90/45/135 degree linear polarized image
	I_s, I_{sss}	Specular/Subsurface scattering observation image
	I_ζ	Subsurface scattering polarization observation image
	η	Index of refraction
	$\mathbf{M}$	Mueller matrix
	$\mathbf{C}$	Coordinate conversion matrix
	$\mathbf{D}$	Depolarized matrix
	$\mathbf{F}^{F\in\{\mathcal{T},\mathcal{R}\}}$	Fresnel Mueller matrix for transmission ($\mathcal{T}$) / reflection ($\mathcal{R}$)
	$\mathcal{T}^{\parallel}, \mathcal{R}^{\parallel}$	Fresnel transmission/reflection coefficient along the plane of incidence
	$\mathcal{T}^{\perp}, \mathcal{R}^{\perp}$	Fresnel transmission/reflection coefficient perpendicular to the plane of incidence
	$\mathcal{T}^{+}, \mathcal{R}^{+}$	$(\mathcal{T}^{\perp} + \mathcal{T}^{\parallel})/2$, $(\mathcal{R}^{\perp} + \mathcal{R}^{\parallel})/2$
	$\mathcal{T}^{-}, \mathcal{R}^{-}$	$(\mathcal{T}^{\perp} - \mathcal{T}^{\parallel})/2$, $(\mathcal{R}^{\perp} - \mathcal{R}^{\parallel})/2$
	$\mathbf{P}_d, \mathbf{P}_s, \mathbf{P}_{ss}$	Mueller matrix for diffuse/specular/single scattering reflection
	$\mathbf{P}_{sss}$	Mueller matrix for subsurface scattering
BSSRDF	$\mathcal{D}, \mathcal{G}$	Normal GGX distribution/Smith's geometric attenuation function
	κ_s, κ_{ss}	Specular/single scattering term
	$\bar{\kappa}_{s,ss}$	$\kappa_s + \kappa_{ss}$
	α_s, α_{ss}	Roughness parameter of specular/single scattering
	$\rho_s, \rho_{ss}, \rho_d$	Albedo of specular/single scattering/diffuse
	ρ_{sss}	Subsurface scattering reflectance function
	$\bar{\rho}_{sss}$	Averaged subsurface scattering reflectance value
Subsurface scattering parameter	$\sigma_a^{\text{oxy}}, \sigma_a^{\text{deoxy}}$	Spectral absorption coefficient of oxy/deoxy hemoglobin
	$\sigma_a^{\text{em}}, \sigma_a^{\text{pm}}$	Spectral absorption coefficient of eumelanin/pheomelanin
	σ_a^{b}	Spectral absorption coefficient of base human skin
	$\sigma_a^{\text{out}}, \sigma_a^{\text{in}}$	Spectral absorption coefficient of outer/inner layer
	$\sigma_s^{\text{out}\prime}, \sigma_s^{\text{in}\prime}$	Reduced scattering coefficient of outer/inner layer
	α'	Reduced albedo
	σ_t', l	Reduced extinction coefficient and mean free path
	σ_{tr}	Effective transport coefficient
	$z_{r,k}, z_{v,k}$	Position of the positive/negative monopole
	D	Diffusion constant
	$F(0)_{\text{dr}}$	Average Fresnel reflectance at the surface depth 0
	$A(0)$	$(1 + F(0)_{dr})/(1 - F(0)_{dr})$
	$C_{\text{h,out}}, C_{\text{h,in}}$	Fraction of hemoglobin in outer/inner layer
	C_{m}	Fraction of melanin in the inner layer
	β_{m}	Fraction of eumelanin in the inner layer melanin
	$\gamma_{\text{out}}, \gamma_{\text{in}}$	Oxy-hemoglobin fraction in outer/inner hemoglobin
Subsurface Scattering	Φ	Incident flux
	M	Radiant emittance profile
	L_i, L_o	Incident/exitant radiance
	Ψ	Bidirectional scattering-surface reflectance-distribution function
	$R_{\text{out}}^{\text{f}}, T_{\text{out}}^{\text{f}}$	Forward reflectance/transmittance profile of the outer layer
	$R_{\text{out}}^{\text{b}}, T_{\text{out}}^{\text{b}}$	Backward reflectance/transmittance profile of the outer layer
	R_{in}^{f}	Forward reflectance profile of the inner layer
	$\bar{R}_i, \bar{T}_i$	Total reflectance/transmittance profile at layer i
	G	Gaussian function
	v_j	Variance of the sum of the Gaussian at index j
	G_{R_i}, G_{T_i}	Sum of Gaussian of reflectance/transmittance profile at layer i
	$G_{\bar{R}_i}, G_{\bar{T}_i}$	Normalized sum of Gaussian of reflectance/transmittance profile at layer i

References

Carlos Aliaga, Christophe Hery, and Mengqi Xia. 2022. Estimation of spectral biophysical skin properties from captured RGB albedo. *arXiv preprint arXiv:2201.10695* (2022).

Carlos Aliaga, Menqi Xia, Xao Xie, Adrian Jarabo, Gustav Braun, and Christophe Hery. 2023. A Hyperspectral Space of Skin Tones for Inverse Rendering of Biophysical Skin Properties. In *Computer Graphics Forum*, Vol. 42. Wiley Online Library.

Sarah Alotaibi and William AP Smith. 2017. A Biophysical 3D Morphable Model of Face Appearance. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 824–832.

R. Rox Anderson and John A. Parrish. 1981. The Optics of Human Skin. *Journal of investigative dermatology* 77, 1 (1981), 13–19. Publisher: Elsevier.

Gary A. Atkinson and Edwin R. Hancock. 2006. Recovery of Surface Orientation from Diffuse Polarization. *IEEE transactions on image processing* 15, 6 (2006), 1653–1664. Publisher: IEEE.

Dejan Azinović, Olivier Maury, Christophe Hery, Matthias Nießner, and Justus Thies. 2023. High-Res Facial Appearance Capture from Polarized Smartphone Images. In *Proc. the IEEE/CVF CVPR*.

Seung-Hwan Baek and Felix Heide. 2021. Polarimetric Spatio-Temporal Light Transport Probing. *ACM Transactions on Graphics* 40, 6 (2021), 18 pages.

Seung-Hwan Baek, Daniel S. Jeon, Xin Tong, and Min H. Kim. 2018. Simultaneous acquisition of polarimetric SVBRDF and normals. *ACM Trans. Graph.* 37, 6 (2018), 268–1.

Thabo Beeler, Bernd Bickel, Paul Beardsley, Bob Sumner, and Markus Gross. 2010. High-Quality Single-Shot Capture of Facial Geometry. *ACM Transactions on Graphics* 29, 4 (2010), 9 pages.

Thabo Beeler, Fabian Hahn, Derek Bradley, Bernd Bickel, Paul Beardsley, Craig Gotsman, Robert W. Sumner, and Markus Gross. 2011. High-Quality Passive Facial Performance Capture Using Anchor Frames. *ACM Transactions on Graphics* 30, 4 (2011), 10 pages.

Sai Bi, Stephen Lombardi, Shunsuke Saito, Tomas Simon, Shih-En Wei, Kevyn Mcphail, Ravi Ramamoorthi, Yaser Sheikh, and Jason Saragih. 2021. Deep Relightable Appearance Models for Animatable Faces. *ACM Transactions on Graphics* 40, 4 (2021).

Derek Bradley, Wolfgang Heidrich, Tiberiu Popa, and Alla Sheffer. 2010. High Resolution Passive Facial Performance Capture. *ACM Transactions on Graphics* 29, 4 (2010).

Edward Collett. 2005. Field guide to polarization. SPIE Bellingham, WA.

Eugene d'Eon, David Luebke, and Eric Enderton. 2007. Efficient Rendering of Human Skin. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*. Citeseer, 147–157.

Craig Donner and Henrik Wann Jensen. 2005. Light Diffusion in Multi-Layered Translucent Materials. *ACM Transactions on Graphics* 24, 3 (2005), 1032–1039.

Craig Donner, Tim Weyrich, Eugene d'Eon, Ravi Ramamoorthi, and Szymon Rusinkiewicz. 2008. A Layered, Heterogeneous Reflectance Model for Acquiring and Rendering Human Skin. *ACM Transactions on Graphics* 27, 5 (2008), 12 pages.

W. G. Egan, To Hilgeman, and J. Reichman. 1973. Determination of Absorption and Scattering Coefficients for Nonhomogeneous Media. 2: Experiment. *Applied Optics* 12, 8 (1973), 1816–1823. Publisher: Optica Publishing Group.

Graham Fyffe and Paul Debevec. 2015. Single-Shot Reflectance Measurement from Polarized Color Gradient Illumination. In *Proc. IEEE ICCP*. 1–10.

Graham Fyffe, Paul Graham, Borom Tunwattanapong, Abhijeet Ghosh, and Paul Debevec. 2016. Near-Instant Capture of High-Resolution Facial Geometry and Reflectance. In *Computer Graphics Forum*, Vol. 35. Wiley Online Library, 353–363.

Graham Fyffe, Tim Hawkins, Chris Watts, Wan-Chun Ma, and Paul Debevec. 2011. Comprehensive Facial Performance Capture. In *Computer Graphics Forum*, Vol. 30. Wiley Online Library, 425–434.

Abhijeet Ghosh, Graham Fyffe, Borom Tunwattanapong, Jay Busch, Xueming Yu, and Paul Debevec. 2011. Multiview Face Capture Using Polarized Spherical Gradient Illumination. *ACM Transactions on Graphics* 30, 6 (2011), 10 pages.

Abhijeet Ghosh, Tim Hawkins, Pieter Peers, Sune Frederiksen, and Paul Debevec. 2008. Practical Modeling and Acquisition of Layered Facial Reflectance. *ACM Transactions on Graphics* 27, 5 (2008), 10 pages.

Yuliya Gitlina, Giuseppe Claudio Guarnera, Daljit Singh Dhillon, Jan Hansen, Alexander Lattas, Dinesh Pai, and Abhijeet Ghosh. 2020. Practical measurement and reconstruction of spectral skin reflectance. In *Computer graphics forum*, Vol. 39. Wiley Online Library, 75–89. Issue: 4.

Paulo Gotardo, Jérémy Riviere, Derek Bradley, Abhijeet Ghosh, and Thabo Beeler. 2018. Practical Dynamic Facial Appearance Modeling and Acquisition. *ACM Transactions on Graphics* 37, 6 (2018), 13 pages.

Paulo FU Gotardo, Tomas Simon, Yaser Sheikh, and Iain Matthews. 2015. Photogeometric scene flow for high-detail dynamic 3d reconstruction. In *Proceedings of the IEEE international conference on computer vision*. 846–854.

Eric Heitz. 2014. Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs. *Journal of Computer Graphics Techniques* 3, 2 (2014), 32–91.

Inseung Hwang, Daniel S. Jeon, Adolfo Muñoz, Diego Gutierrez, Xin Tong, and Min H. Kim. 2022. Sparse Ellipsometry: Portable Acquisition of Polarimetric SVBRDF and Shape with Unstructured Flash Photography. *ACM Transactions on Graphics* 41, 4 (2022), 14 pages.

Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. 2001. A Practical Model for Subsurface Light Transport. In *Annual Conference Series (Proceedings of SIGGRAPH)*. 511–518.

Jorge Jimenez, Timothy Scully, Nuno Barbosa, Craig Donner, Xenxo Alvarez, Teresa Vieira, Paul Matts, Verónica Orvalho, Diego Gutierrez, and Tim Weyrich. 2010. A Practical Appearance Model for Dynamic Facial Color. *ACM Transactions on Graphics* 29, 6 (2010), 10 pages.

Jorge Jimenez, Károly Zsolnai, Adrian Jarabo, Christian Freude, Thomas Auzinger, Xian-Chun Wu, Javier von der Pahlen, Michael Wimmer, and Diego Gutierrez. 2015. Separable Subsurface Scattering. In *Computer Graphics Forum*, Vol. 34. Wiley Online Library, 188–197. Issue: 6.

Achuta Kadambi, Vage Taamazyan, Boxin Shi, and Ramesh Raskar. 2015. High-Quality Depth Sensing with Polarization Cues. In *Proc. IEEE/CVF CVPR*. 3370–3378.

Chloe LeGendre, Kalle Bladin, Bipin Kishore, Xinglei Ren, Xueming Yu, and Paul Debevec. 2018. Efficient Multispectral Facial Capture with Monochrome Cameras. In *ACM SIGGRAPH 2018 Posters*. Association for Computing Machinery, 2 pages.

Ruilong Li, Karl Bladin, Yajie Zhao, Chinmay Chinara, Owen Ingraham, Pengda Xiang, Xinglei Ren, Pratusha Prasad, Bipin Kishore, and Jun Xing. 2020. Learning Formation of Physically-Based Face Attributes. In *Proc. IEEE/CVF CVPR*. 3410–3419.

Shichen Liu, Yunxuan Cai, Haiwei Chen, Yichao Zhou, and Yajie Zhao. 2022. Rapid Face Asset Acquisition with Recurrent Feature Alignment. *ACM Transactions on Graphics* 41, 6 (2022), 1–17.

Wan-Chun Ma, Tim Hawkins, Pieter Peers, Charles-Felix Chabert, Malte Weiss, and Paul E. Debevec. 2007. Rapid Acquisition of Specular and Diffuse Normal Maps from Polarized Spherical Gradient Illumination. *Rendering Techniques* 2007, 9 (2007).

Miki Morimatsu, Yusuke Monno, Masayuki Tanaka, and Masatoshi Okutomi. 2020. Monochrome and color polarization demosaicking using edge-aware residual interpolation. In *2020 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2571–2575. https://ieeexplore.ieee.org/abstract/document/9191085/

Fred Edwin Nicodemus, Joseph C. Richmond, Jack J. Hsia, Irving W. Ginsberg, and Thomas Limperis. 1977. Geometrical Considerations and Nomenclature for Reflectance. *Final Report National Bureau of Standards* (1977).

Stephen J. Preece and Ela Claridge. 2004. Spectral filter optimization for the recovery of parameters which describe human skin. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 7 (2004), 913–922.

Jérémy Riviere, Paulo Gotardo, Derek Bradley, Abhijeet Ghosh, and Thabo Beeler. 2020. Single-Shot High-Quality Facial Geometry and Skin Appearance Capture. *ACM Transactions on Graphics* 39, 4 (2020), 12 pages.

Szymon M. Rusinkiewicz. 1998. A New Change of Variables for Efficient BRDF Representation. *Rendering techniques* 98 (1998), 11–22.

M. J. C. Van Gemert, Steven L. Jacques, HJCM Sterenborg, and W. M. Star. 1989. Skin Optics. *IEEE Transactions on biomedical engineering* 36, 12 (1989), 1146–1154. Publisher: IEEE.

Alexander Vilesov, Pradyumna Chari, Adnan Armouti, Anirudh Bindiganavale Harish, Kimaya Kulkarni, Ananya Deoghare, Laleh Jalilian, and Achuta Kadambi. 2022. Blending Camera and 77 GHz Radar Sensing for Equitable, Robust Plethysmography. *ACM Transactions on Graphics* 41, 4 (2022), 1–14.

Bruce Walter, Stephen R. Marschner, Hongsong Li, and Kenneth E. Torrance. 2007. Microfacet Models for Refraction through Rough Surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*. 195–206.

Tim Weyrich, Wojciech Matusik, Hanspeter Pfister, Bernd Bickel, Craig Donner, Chien Tu, Janet McAndless, Jinho Lee, Addy Ngan, Henrik Wann Jensen, and Markus Gross. 2006. Analysis of Human Faces Using a Measurement-Based Skin Reflectance Model. *ACM Transactions on Graphics* 25, 3 (2006), 1013–1024.

Alexander Wilkie and Andrea Weidlich. 2012. Polarised Light in Computer Graphics. In *SIGGRAPH Asia 2012 Courses*. 87 pages.

Stephen J. Wright. 2015. Coordinate descent algorithms. *Mathematical programming* 151, 1 (2015), 3–34. Publisher: Springer.

Longwen Zhang, Chuxiao Zeng, Qixuan Zhang, Hongyang Lin, Ruixiang Cao, Wei Yang, Lan Xu, and Jingyi Yu. 2022. Video-driven Neural Physically-based Facial Asset for Production. *ACM Transactions on Graphics* 41, 6 (2022), 1–16.