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

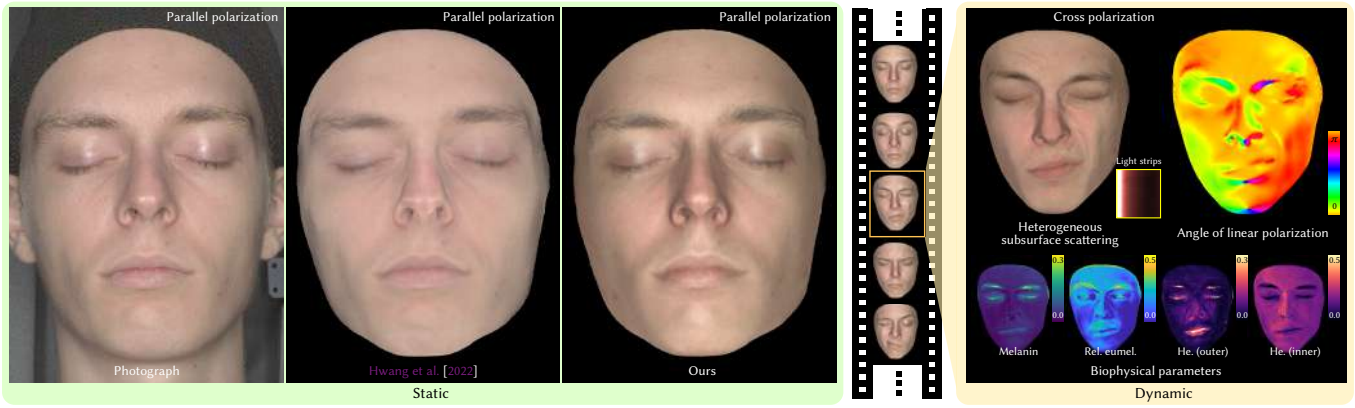


Fig. 1. Current polarimetric acquisition systems are limited to static, opaque objects. In this work, we present a novel appearance acquisition method that allows us to obtain biophysically-based polarimetric BSSRDF and surface geometry of dynamic faces. Our polarimetric appearance parameters include index of refraction, specular roughness, single scattering roughness, specular intensity, and single scattering intensity. We also introduce an end-to-end multispectral optimization with heterogeneous subsurface scattering, revealing biophysically-based skin parameters including inner- and outer-layer hemoglobin, eumelanin and pheomelanin. Our model is compatible with popular human skin models in graphics. Please refer to the supplemental video for dynamic results.

Acquisition and modeling of polarized light reflection and scattering help reveal the shape, structure, and physical characteristics of an object, which is increasingly important in computer graphics. However, current polarimetric acquisition systems are limited to static and opaque objects. Human faces, on the other hand, present a particularly difficult challenge, given their complex structure and reflectance properties, the strong presence of spatially-varying subsurface scattering, and their dynamic nature. We present a new polarimetric acquisition method for dynamic human faces, which focuses on capturing spatially varying appearance and precise geometry, across a wide spectrum of skin tones and facial expressions. It includes both single and heterogeneous subsurface scattering, index of refraction, and specular roughness and intensity, among other parameters, while revealing biophysically-based components such as inner- and outer-layer hemoglobin, eumelanin and pheomelanin. Our method leverages such components' unique multispectral absorption profiles to quantify their concentrations, which in turn inform our model about the complex interactions occurring within the skin layers. To our knowledge, our work is the first to simultaneously acquire polarimetric and spectral reflectance information alongside biophysically-based skin parameters and geometry of dynamic human faces. Moreover, our polarimetric skin model integrates seamlessly into various rendering pipelines.

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.

ACM 1557-7368/2024/12-ART275
https://doi.org/10.1145/3687767

CCS Concepts: • **Computing methodologies → Reflectance modeling**.

Additional Key Words and Phrases: Polarization imaging, multispectral imaging, skin reflectance modeling

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. Polarimetric BSSRDF Acquisition of Dynamic Faces. *ACM Trans. Graph.* 43, 6, Article 275 (December 2024), 11 pages. https://doi.org/10.1145/3687767

1 Introduction

Polarization can provide valuable information about the shape and physical characteristics of an object. As a result, polarization-based systems have become increasingly important in computer graphics. These require capturing and processing four-dimensional Stokes vectors to account for all potential polarization states.

Furthermore, the capture of polarimetric reflectance requires control over incident and outgoing light, which are both Stokes vectors, so the reflectance function is often represented by a four-by-four Mueller matrix that links each component of the incident light with its outgoing counterpart. The complexity of this matrix increases the computational cost of estimating polarization appearance parameters, requiring additional observations with different incoming and outgoing directions and polarization states. Ellipsometry is the most common technique to estimate this matrix, using structured optical measurements to characterize how light interactions affect the polarization state. Current methods to capture spatially varying polarimetric reflectance techniques need to apply strong assumptions to make the problem tractable. As a result, they are still limited to acquiring polarimetric appearance information from static opaque objects.

In this work, we lift these restrictions and capture the polarimetric reflectance of dynamic deformable objects with strong, spatially-varying subsurface scattering. We specifically target the challenging case of human faces, which are both deformable and translucent. Figure 1 compares two different polarization rendering results by a state-of-the-art method [Hwang et al. 2022] and ours. By accounting for heterogeneous subsurface scattering, our model yields more precise results, closer to the photographic reference.

Estimating such subsurface scattering is challenging since light scatters through the multiple translucent layers of skin. We base our subsurface scattering appearance model on biophysical components, in particular melanin (eumelanin and pheomelanin) and hemoglobin (oxy-hemoglobin and deoxy-hemoglobin). Our hardware setup captures six spectral observations, and we rely on the unique spectral absorption profiles of these biophysical components to estimate their individual contribution to the final appearance.

Our algorithm is made up of two stages. First, we capture polarimetric observations of the subject's static face from a wide range of angles, from which we optimize the appearance parameters that serve as initialization for the next stage. In the second stage, for each frame, we follow a similar optimization, starting from the outcome of the first stage. With this approach, we obtain per-frame high-quality dynamic geometry, as well as spatially-varying appearance parameters represented as texture maps. We use polarization imaging for estimating specular reflectance, single scattering and geometrical detail, while we approximate the biophysical parameters from the spectral observations of subsurface scattering. To our knowledge, our technique is the first to capture polarimetric reflectance on dynamic deformable objects. We show results across a wide spectrum of skin tones and facial expressions (both in the main paper and supplemental material). Moreover, our polarimetric skin model integrates seamlessly into many existing rendering pipelines. Our code is available for research purposes[1].

[1] https://github.com/KAIST-VCLAB/polarimetric-bssrdf-dynamic-face.git

2 Related Work

Polarimetric imaging. Polarimetric imaging has widely been used in computer graphics. Passive systems use cameras fitted with polarizers, positioned in front of the lens (e.g., [Atkinson and Hancock 2006; Cao et al. 2023; Cui et al. 2019; Deschaintre et al. 2021; Huynh et al. 2013; Kadambi et al. 2015; Miyazaki et al. 2003; Riviere et al. 2017; Tozza et al. 2017; Zhu and Smith 2019]), or the image sensor [Ba et al. 2020; Dave et al. 2022; Lei et al. 2022; Zhao et al. 2022], while active systems incorporate both polarized light sources and polarized cameras (e.g., [Azinović et al. 2023; Ghosh et al. 2010, 2008; Ma et al. 2007; Riviere et al. 2020]). In general, these setups only measure specific polarization states, such as linear polarization at particular angles or circular polarization.

Other works aim to capture polarimetric appearance across different polarization states [Baek and Heide 2021, 2022; Baek et al. 2018]. Hwang et al. [2022] combined a polarization-array camera with a polarized flashlight. However, these approaches are limited to static scenes and do not take subsurface scattering explicitly into account. In contrast, our work allows us to capture polarimetric information on dynamic faces, including the effects of subsurface scattering.

Face acquisition. Numerous methods have been developed to acquire high-quality geometric shapes and the appearance of static faces (e.g., [Azinović et al. 2023; Debevec et al. 2000; Fyffe 2010; Fyffe et al. 2016, 2011; Ghosh et al. 2010, 2008; Imai 1998; LeGendre et al. 2018; Ma et al. 2007; Shrestha and Hardeberg 2010; Weyrich et al. 2006]). Since they all require multiple structured light patterns and/or input from various viewpoints, they are unsuitable for dynamic captures.

Dynamic face capture methods, on the other hand, often use passive illumination, taking images of objects under uniform lighting conditions. Multi-view camera systems rely on stereo matching for geometry acquisition [Beeler et al. 2010], or tracking in image space [Beeler et al. 2011; Bradley et al. 2010], but do not reconstruct the appearance of skin. Monocular single-shot [Sengupta et al. 2018; Tran et al. 2018, 2019; Tran and Liu 2019], video sequences [Cao et al. 2015; Garrido et al. 2013; Ichim et al. 2015; Shi et al. 2014], or binocular video sequences [Valgaerts et al. 2012] have been used to obtain both geometric information and appearance. These approaches typically assume simplified reflectance models of human skin, for instance, including only diffuse albedo or not taking subsurface scattering into account [Gotardo et al. 2018]. Riviere et al. [2020] developed a passive stereo-capture system to acquire specular reflectance and diffuse albedo. As opposed to our work, the method assumes a pre-determined subsurface scattering profile, while each frame needs to be processed independently for animated sequences. Since single-shot input is ill-conditioned for human skin acquisition, recent research has turned to learning from active multi-view lighting systems [Bi et al. 2021; Li et al. 2020; Liu et al. 2022; Zhang et al. 2022]. These learning-based methods are constrained by the training and test datasets, which do not describe the reflectance of human faces in a physically-based way. In contrast, our multispectral polarimetric subsurface scattering model yields approximate meaningful, spatially-varying, and time-resolved biophysically-based appearance parameters for dynamic faces.

Biophysical appearance acquisition. Existing methods to approximate biophysical parameters of human skin usually rely on simplified models, such as assuming diffuse reflectance, not taking into account subsurface scattering, or not handling dynamic changes in appearance. Tsumura et al. [1999, 2003] created an image-based method to separate the spatial patterns of melanin and hemoglobin in human skin through independent-component analysis of a skin color image. This model was later extended to take into account the more complex properties of skin [Krishnaswamy and Baranoski 2004] based on multispectral images [Chen et al. 2015; Donner et al. 2008; Preece and Claridge 2004], RGB diffuse reflectance images [Aliaga et al. 2022; Alotaibi and Smith 2017], or different lighting conditions [Aliaga et al. 2023; Gitlina et al. 2020; Li et al. 2024]. Some of these methods require precomputed textures for inverse rendering, which may lead to visible discretization artifacts or rely on rendered datasets. Regarding dynamic models, Jimenez et al. [2010] presented a method focused on the acquisition of simplified hemoglobin maps from cross-polarizing filters, requiring multiple captures with the subject repeating the same movements. Later, Iglesias-Guitian et al. [2015] introduced a statistically-based model of human skin that captured the time-varying effects of aging, as the structure of skin and its chromophores change over the years. In contrast, our technique disambiguates the dynamic changes of biophysically-based components, such as oxy-hemoglobin, deoxy-hemoglobin, eumelanin or pheomelanin, as well as their full diffusion profiles, through a multispectral observation of subsurface scattering without requiring impractical repeated motions. Moreover, our system is capable of simulating appearance changes that occur within seconds, instead of decades.

3 Reflectance model of skin

We describe here the main aspects of our reflectance model for skin, including our polarimetric BSSRDF and biophysically-based parameters; please refer to the supplemental document for additional details (Supplemental Section 2).

3.1 Polarimetric BSSRDF Model

A Stokes vector represents the polarization state of a light wave, and is denoted as $\mathbf{s} = [s_0, s_1, s_2, s_3]^\intercal \in \mathbb{R}^{4\times 1}$. Polarized light $\mathbf{s}_i$ reflects off a surface as $\mathbf{s}_o = S\mathbf{P}(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o)\mathbf{s}_i$, 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, and $\mathbf{P}(\boldsymbol{\omega}_i, \boldsymbol{\omega}_o)$ is the polarimetric reflectance model that yields a Mueller matrix for incoming $\boldsymbol{\omega}_i$ and outgoing $\boldsymbol{\omega}_o$ directions [Wilkie and Weidlich 2012].

Subsurface scattering describes how light enters a surface at point $\mathbf{x}_i$ and exits at a different point $\mathbf{x}_o$. Different from existing polarimetric reflectance models [Baek and Heide 2021; Baek et al. 2018; Hwang et al. 2022], we explicitly take into account *heterogeneous* subsurface scattering. This is important for human skin since, although light becomes depolarized during multiple scattering, it gets polarized again when transmitted back out, thus becoming an additional source of reflectance information.

Our polarimetric reflectance model can then be expressed as $\mathbf{P} = \mathbf{P}_s + \mathbf{P}_{ss} + \mathbf{P}_{sss}$, where $\mathbf{P}_s$, $\mathbf{P}_{ss}$, and $\mathbf{P}_{sss}$ represent the specular, single scattering, and subsurface scattering components, respectively. We adopt the specular and single scattering terms from the recent state-of-the-art model by Hwang et al. [2022] (Supplemental Section 2.4), and expand the model with our new subsurface scattering term $\mathbf{P}_{sss}$, described in the following paragraphs.

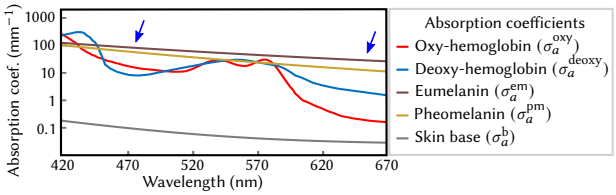


Fig. 2. Spectral absorption coefficients of each biophysical component in the human skin: oxy-hemoglobin, deoxy-hemoglobin, eumelanin, pheomelanin, and skin base.

Subsurface scattering. We model subsurface scattering $\mathbf{P}_{sss}$ from a diffusion profile $\rho_{sss}(||\mathbf{x}_i - \mathbf{x}_o||)$ as

$$\begin{aligned}\mathbf{P}_{sss} &= \sum_{\mathbf{x}_i\in\mathcal{S}} \mathbf{C}_{n\to o}(-\tilde{\phi}_o)\mathbf{F}^{\mathcal{T}}(\mathbf{x}_o,\theta_o;\eta_o)\mathbf{D}(\rho_{sss}(||\mathbf{x}_i-\mathbf{x}_o||))\\ &\quad\cdot\mathbf{F}^{\mathcal{T}}(\mathbf{x}_i,\theta_i;\eta_i)\mathbf{C}_{i\to n}(\tilde{\phi}_i)\\ &= \sum_{\mathbf{x}_i\in\mathcal{S}} \rho_{sss}(||\mathbf{x}_i-\mathbf{x}_o||)\begin{bmatrix} \mathcal{T}_i^+\mathcal{T}_o^+ & -\mathcal{T}_i^-\mathcal{T}_o^+\xi_i & -\mathcal{T}_i^-\mathcal{T}_o^+\zeta_i & 0\\ -\mathcal{T}_i^+\mathcal{T}_o^-\xi_o & \mathcal{T}_i^-\mathcal{T}_o^-\xi_i\xi_o & \mathcal{T}_i^-\mathcal{T}_o^-\zeta_i\xi_o & 0\\ -\mathcal{T}_i^+\mathcal{T}_o^-\zeta_o & \mathcal{T}_i^-\mathcal{T}_o^-\xi_i\zeta_o & \mathcal{T}_i^-\mathcal{T}_o^-\zeta_i\zeta_o & 0\\ 0 & 0 & 0 & 0\end{bmatrix},\end{aligned} \tag{1}$$

where $\mathbf{C}_{n\to o}(-\tilde{\phi}_o)$ and $\mathbf{C}_{i\to n}(\tilde{\phi}_i)$ are the coordinate conversion matrices, $\mathbf{D}$ is the depolarization matrix, $\mathbf{F}^{\mathcal{T}}$ is the Mueller matrix form of the Fresnel transmission coefficients $\mathcal{T}$ that takes into account the different effects on light polarized along the plane of incidence ($\parallel$) and perpendicular to it ($\perp$). The $+/-$ superscript operators refer to $\mathcal{T}^+ = (\mathcal{T}^\perp + \mathcal{T}^\parallel)/2$ and $\mathcal{T}^- = (\mathcal{T}^\perp - \mathcal{T}^\parallel)/2$. Here ζ and ξ are $\sin(2\phi)$ and $\cos(2\phi)$ of the polarimetric azimuth angle ϕ between the light frame and the interaction plane, and $\tilde{\phi}_{\{i,o\}} = \phi_{\{i,o\}} - \pi/2$ is the corresponding rotation angle. The sum $\sum_{\mathbf{x}_i\in\mathcal{S}}$ takes into account the fact that all incoming points $\mathbf{x}_i$ of the surface $\mathcal{S}$ contribute to the outgoing illumination.

The diffusion profile ρ_{sss} is obtained from the absorption and scattering coefficients of two layers, which depend on a set of biophysically-based parameters as described in Section 3.2, by applying the multipole approximation [Donner and Jensen 2005]. We assume that the parameters vary slowly relative to the mean free path of light, and therefore are locally homogeneous. For efficiency purposes, per-layer profiles are approximated to weighted sums of separable Gaussian functions [d'Eon et al. 2007; Donner et al. 2008]. In the rest of the paper, we omit the dependance of ρ_{sss} on the distance $||\mathbf{x}_i - \mathbf{x}_o||$ for the sake of brevity.

3.2 Biophysically-Based Model

Similar to other works [Donner and Jensen 2006; Donner et al. 2008; Jimenez et al. 2010], we adopt a two-layer model, where each layer is characterized by its absorption and reduced scattering coefficients. Figure 2 shows the spectral absorption coefficients of each component, included in our model as explained in the next paragraphs.

Outer-layer absorption. The spectral absorption coefficient of the outer layer σ_a^{out} is mainly due to the presence of melanin in the epidermis, and, to a lesser extent, to the presence of hemoglobin in the upper dermis, and is defined as [Donner et al. 2008; Jimenez et al. 2010]

$$\begin{aligned}\sigma_a^{\text{out}} =& C_{\text{m}}(\beta_{\text{m}}\sigma_a^{\text{em}} + (1-\beta_{\text{m}})\sigma_a^{\text{pm}}) \\ &+ C_{\text{h,out}}(\gamma_{\text{out}}\sigma_a^{\text{oxy}} + (1-\gamma_{\text{out}})\sigma_a^{\text{deoxy}}) + (1 - C_{\text{m}} - C_{\text{h,out}})\sigma_a^{\text{b}},\end{aligned} \tag{2}$$

where C_{m} and $C_{\text{h,out}}$ are the fractions of melanin and hemoglobin in the outer layer, respectively, β_{m} is the fraction of eumelanin in melanin, γ_{out} is the oxy-hemoglobin fraction in hemoglobin. The values of C_{m}, $C_{\text{h,out}}$, and β_{m} are estimated from our captured data while γ_{out} is a constant (see Section 5.3). The different spectral absorption coefficients σ_a of eumelanin (em), pheomelanin (pm), oxy-hemoglobin (oxy), deoxy-hemoglobin (deoxy), and base (b) are given by previous work [Jacques 1998; Prahl 1999].

Inner-layer absorption. The inner layer consists of a rich network of capillaries containing hemoglobin in the dermis. Its absorption coefficient σ_a^{in} is mainly explained by hemoglobin, and can be expressed as [Donner and Jensen 2006]

$$\sigma_a^{\text{in}} = C_{\text{h,in}}(\gamma_{\text{in}}\sigma_a^{\text{oxy}} + (1-\gamma_{\text{in}})\sigma_a^{\text{deoxy}}) + (1 - C_{\text{h,in}})\sigma_a^{\text{b}}, \tag{3}$$

where $C_{\text{h,in}}$ is the fraction of the hemoglobin in the inner layer, and γ_{in} is its oxy-hemoglobin fraction. The value of $C_{\text{h,in}}$ is also estimated from our captured data. We assume that the fractions of oxy-hemoglobin to hemoglobin in the inner and outer layers are the same and have a fixed value $\gamma = \gamma_{\text{in}} = \gamma_{\text{out}} = 0.75$ as other existing models [Donner et al. 2008; Jimenez et al. 2010].

Reduced scattering. The spectral reduced scattering coefficient of the outer layer $\sigma_s^{\text{out}'}$ at wavelength λ nm is defined as [Bashkatov et al. 2005]

$$\sigma_s^{\text{out}'}(\lambda) = 14.74 \times \lambda^{-0.22} + 2.2E^{11} \times \lambda^{-4}. \tag{4}$$

The reduced scattering coefficient of the inner layer is 50% of the outer scattering coefficient, so it does not need to be explicitly estimated.

4 Multispectral Polarimetric Imaging

We summarize here the main aspects of our hardware design and polarimetric image formation model, and refer the reader to the supplemental material (Supplemental Section 3) for more details.

4.1 Capture Hardware

Our capture system is shown in Figure 3. It consists of a multispectral, polarimetric module in the center, and four additional 3D imaging modules surrounding it. The polarimetric imaging module is composed of two polarization machine vision cameras (BFS-U3-51SPC-C), synchronized at 20 fps, each of them fitted with a different multispectral filter from off-the-shelf Dolby 3D glasses. Each polarization camera captures four linearly polarized components (0°, 45°, 90°, 135°) at 2448×2048 resolution. Note that conventional polarimetric cameras have a lower SNR and resolution than color cameras. The cameras are surrounded by 40 linearly-polarized ~1500 lumen LED light sources (CXA-1512 6500K, operating at 350mA with 36V) (eight modules of five LEDs each) in a near-coaxial setup. Each Dolby 3D glass further filters the wavelength range of each of the camera's conventional red, green, and blue filters, effectively halving the range for each channel (the left camera captures the higher half, while the right one captures the lower one), yielding a coverage of the whole spectrum for a total of six samples (see Figure 4). This is particularly useful for the spectral response of human skin, in which the spectral absorption profiles of its components are identifiably different (Figure 2). Finally, each 3D imaging module consists of two machine vision cameras to capture dynamic 3D geometry. All cameras are synchronized (Supplemental Section 3.2).

4.2 Polarimetric Image Formation

Existing polarimetric acquisition methods can capture diffuse and specular information [Baek et al. 2018, 2020] or even single scattering [Hwang et al. 2022], assuming that target objects are opaque. We introduce a novel subsurface scattering term (Equation (1)) to handle translucency, as in human skin.

Since our light sources and cameras are in a near-coaxial setup, we can apply sparse ellipsometry algebraic simplifications to our polarimetric BSSRDF model [Hwang et al. 2022]. Given our linearly polarized captured images from each camera $I_0, I_{90}, I_{45}, I_{135}$, we compute the following observations:

- The unpolarized subsurface scattering observation is defined as $I_{sss} = 2I_{90} = S\sum_{\mathbf{x}_i \in \mathcal{S}} \rho_{sss}\mathcal{T}^{++}$.
- The polarized subsurface scattering observation is defined as $I_\zeta = I_{135} - I_{45} = S\sum_{\mathbf{x}_i \in \mathcal{S}} \rho_{sss}\mathcal{T}^{-+}\zeta$.
- The specular-dominant polarization observation is defined as $I_s = I_0 - I_{90} = S(\bar{\kappa}_{s,ss}\mathcal{R}^{+} - \sum_{\mathbf{x}_i \in \mathcal{S}} \rho_{sss}\mathcal{T}^{-+}\xi)$. Here, $\bar{\kappa}_{s,ss}$ is the summation of the specular reflection term κ_s and the single scattering reflection term κ_{ss}. It contains a combination of specular reflection, single scattering, and multiple subsurface scattering.

Here, $\mathcal{R}^{+} = (\mathcal{R}^{\perp} + \mathcal{R}^{\parallel})/2$ represent Fresnel reflection coefficients, $\mathcal{T}^{++} = \mathcal{T}^{+}\mathcal{T}^{+}$ is the multiplication of the positive Fresnel transmission coefficients, and $\mathcal{T}^{-+} = \mathcal{T}^{-}\mathcal{T}^{+}$ is the multiplication of the negative/positive coefficients. Please refer to the supplemental document for the complete mathematical details of our polarimetric image formation model (Supplemental Section 3.1).

We acquire these observations per frame and use them as input to our optimization algorithm (described in the next section), to obtain our full dynamic data.

5 Reconstruction of Dynamic Skin Appearance and Face Geometry

Our reconstruction algorithm consists of two stages (see Figure 5 for an overview). The first stage is a *static initialization* from multiple views of the same face, rotated thanks to a revolving chair. We first obtain the face's mesh from stereo pairs (Section 5.1); we then simultaneously optimize the displacement map plus the polarimetric appearance parameters (Section 5.2), then optimize the biophysical parameters (Section 5.3). At each iteration, the face is rendered in order to calculate the loss function with respect to the input observations. The second stage is a *dynamic per-frame optimization*, in which, starting with the results of the first stage, we optimize all

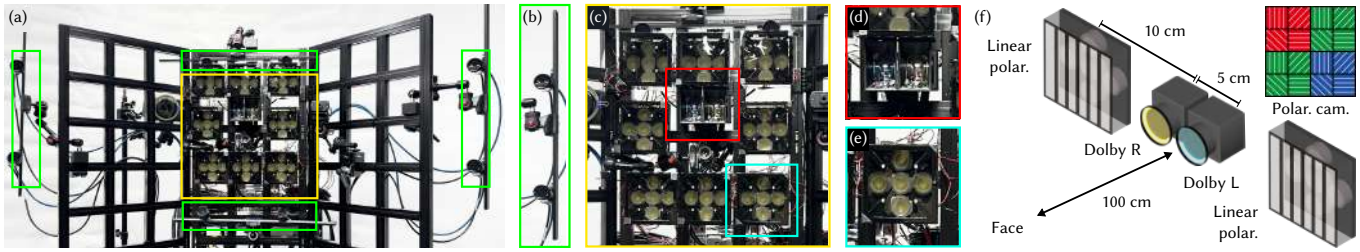


Fig. 3. Our multispectral polarimetric imaging setup. (a) Our system consists of a multispectral polarization module (yellow box) that captures four different linear polarization angles with six multispectral channels and four 3D stereo imaging modules (green boxes). (b) A closeup image of the stereo module composed of two machine vision cameras. Stereo pairs are used to acquire dense depth maps to obtain a complete geometry model and its corresponding texture mapping for each frame. (c)–(e) Closeup images of the multispectral polarization camera module equipped with two polarization cameras (d) covered with two different Dolby filters and 40 LED lights, separated into eight modules of five LEDs, each shown in (e). Our multispectral polarization module captures six different multispectral channels using two different Dolby filters for three RGB channels. Each light is covered with a vertical polarization filter. (f) Schematic diagram of the system configuration. The module also captures four different orientations of linear polarization. The eight light modules are placed at a distance of approximately 10 cm. As the subject is placed 100 cm away, light and camera become almost coaxial at a 5.72° angle. These cameras are synchronized through GPIO cables. We capture images at 20 fps.

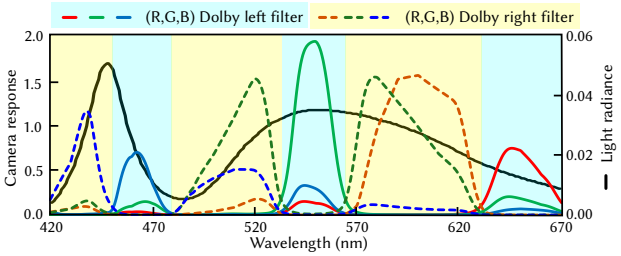


Fig. 4. Multispectral-channel calibration and light spectral power distribution (black line). We disassemble left and right glass lenses from Dolby multispectral 3D anaglyph glasses to use each as a bandpass filter in front of two sRGB cameras. Each color channel is subdivided into two sub-color channels, resulting in six multispectral channels.

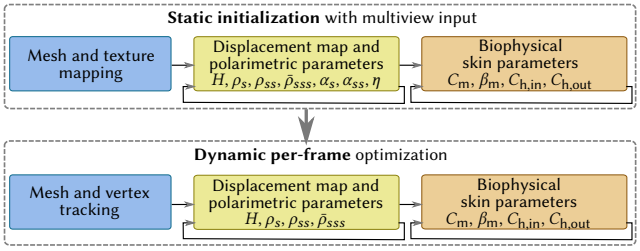


Fig. 5. Overview of our reconstruction algorithm. It consists of two stages: first, the initialization takes multiple views of a static face (top). We then perform a per-frame optimization from a single view (bottom). Both stages follow a similar procedure: first, we obtain a mesh from stereo pairs (left), then we iteratively optimize a displacement map and a set of polarimetric appearance parameters (middle), and lastly, we use inverse rendering to iteratively optimize the biophysical parameters of human skin (right).

the necessary parameters for every frame of the captured video in a similar manner as the first stage.

In more detail, given our hardware setup, from the four linearly polarized images ($I_0, I_{90}, I_{45}, I_{135}$) and the four stereo view pairs, we obtain six-channel multispectral observations I_{sss}, I_ζ and I_s through our image formation model (Section 4.2). We next search the geometric correspondences of spatially varying dynamic appearance changes over time. We aim to couple polarimetric appearance, biophysically-based skin parameters and the surface geometry to track appearance parameters on a face over time. In particular, the index of refraction η (which affects all Fresnel coefficients), albedo (ρ_s and ρ_{ss}) and roughness (α_s and α_{ss}) of the specular (s) and single scattering (ss) components, and the diffusion profile ρ_{sss}, expressed as a function of biophysical parameters (C_m, β_m, $C_\mathrm{h,in}$, $C_\mathrm{h,out}$).

All these parameters, plus a displacement map H encoding geometry details, are spatially varying and are represented as texture maps over the skin's surface. In the following, we explain each stage in more detail and refer the reader to the supplemental material for additional details (Supplemental Section 5).

5.1 Geometry Reconstruction and Tracking

At both stages (static initialization and dynamic per-frame optimization), we first estimate the base mesh geometry using stereo matching [Beeler et al. 2010; Lipson et al. 2021] and Poisson surface reconstruction [Kazhdan et al. 2006]. With this initial geometry, we apply cylindrical texture mapping so that all the spatially-varying appearance parameters and small geometry variations (encoded as displacement map H) are modeled as textures. Note that previous work [Hwang et al. 2022] optimized vertices and their normals directly, requiring an additional Poisson reconstruction process at every iteration, which resulted in a blurrier geometry. By directly optimizing geometric details in the form of a displacement map H, we obtain detailed, more accurate geometrical reconstructions. H is optimized together with the polarimetric appearance parameters (Section 5.2). We assign uv texture coordinates during the initialization stage after the first mesh has been optimized; in the subsequent per-frame optimization we track vertices, but texture coordinates remain unchanged. As a result, textures remain stable along frames, which improves the convergence of our optimization.

During the dynamic per-frame optimization stage, we additionally track corresponding vertices from the initial mesh via optical flow.

We stabilize the tracking by progressively computing the weighted average of the per-frame tracked motion with an anchor-based approach [Beeler et al. 2011], which yields the final per-frame mesh with stable texture coordinates per vertex.

5.2 Optimization of Polarimetric Appearance Parameters

After optimizing the mesh, we optimize the spatially-varying appearance parameters that are related to polarization: index of refraction η, the albedo ρ_s and roughness α_s of the specular component, the albedo ρ_{ss} and roughness α_{ss} of the single scattering component, and a multiple scattering albedo $\bar{\rho}_{sss}$, which is a rough approximation of the diffusion profile $\rho_{sss}(||\mathbf{x}_i - \mathbf{x}_o||)$. This value will be later refined to a full diffusion profile when estimating the face's biologically-based parameters (Section 5.3). We also optimize the displacement map H, which provides the high-frequency details of the geometry. Parameters η, α_s and α_{ss} remain constant in time, and thus they only need to be optimized in the initialization stage.

Inspired by Gotardo et al. [2018], this initialization stage consists on rotating the static face, assuming parameter consistency across frames. As shown in previous work [Nagano et al. 2015], roughness might temporally vary when the skin is stretched or becomes sweaty, but estimating both roughness parameters α_s and α_{ss} from a single view at each frame is an ill-posed problem. The small errors coming from this assumption are compensated by albedos ρ_s and ρ_{ss} and the geometrical variations coming from the displacement map H, all of which are estimated per frame.

In particular, we minimize the following energy function:

$$\min_{\eta,\alpha_s,\alpha_{ss},\rho_s,\rho_{ss},\bar{\rho}_{sss},H} \lambda_\psi \mathcal{L}_\psi + \lambda_{sss}\mathcal{L}_{sss} + \lambda_s \mathcal{L}_s + \lambda_\phi \mathcal{L}_\phi + \mathcal{L}_{\text{reg}}, \quad (5)$$

where $\mathcal{L}_{sss}$ is our subsurface scattering loss, $\mathcal{L}_\psi$ is the refractive index loss, $\mathcal{L}_s$ is the specular and single scattering loss, $\mathcal{L}_\phi$ is the azimuthal loss, $\mathcal{L}_{\text{reg}}$ is the regularization term, and $\lambda_\psi = 0.002$, $\lambda_{sss} = 1$, $\lambda_s = 1$, $\lambda_\phi = 1$ are the corresponding loss weights. We inherit the specular, single scattering, refractive index loss functions from Hwang et al. [2022] and the regularization term (that accounts for spatial and temporal coherency) from Riviere et al. [2020]. Previous work [Hwang et al. 2022] solves these loss terms by alternating the optimization of the refractive index loss ($\mathcal{L}_\psi$) and the specular and single scattering loss ($\mathcal{L}_s$) with azimuthal loss ($\mathcal{L}_\phi$), using a sequential quadratic programming algorithm. In contrast, we use a backward gradient descent-based method that minimizes losses simultaneously. Moreover, while current techniques are limited to polarimetric appearance of static and opaque objects, we capture translucency effects in our initialization stage and track dynamic changes during optimization (Supplemental Section 5).

Subsurface scattering loss. We formulate $\mathcal{L}_{sss}$ by comparing the rendered subsurface scattering image $\hat{I}^t_{sss}$ at time t with the captured image I^t_{sss} as $\mathcal{L}_{sss} = \sum_t \mathrm{V}^t(\hat{I}^t_{sss} - I^t_{sss})^2$, where V^t is the visibility texture map at frame t for each view. Optimizing the full diffusion profile ρ_{sss} along with the rest of the variables is both computationally expensive and ill-conditioned. Therefore, as anticipated earlier, we account for a single multiple scattering albedo $\bar{\rho}_{sss}$ to approximate all the observations of subsurface scattering effects. Since Fresnel transmittance of human skin does not change rapidly along the surface, we approximate the subsurface scattering reflectance as $I^t_{sss} = S\bar{\rho}_{sss}\mathcal{T}^{++}$. Once $\bar{\rho}_{sss}$ and the rest of the polarimetric appearance parameters are optimized, we obtain the full diffusion profile by optimizing the rest of the biophysical parameters, as explained in Section 5.3.

5.3 Optimization of Biophysically-based Parameters

To estimate a full diffusion profile ρ_{sss} from a multispectral observation of subsurface scattering $\bar{\rho}_{sss}$, we rely on the spectral profiles of the absorption coefficients of oxy-hemoglobin, deoxy-hemoglobin, eumelanin, pheomelanin, and the skin base parameter. Since their absorption coefficients are different with respect to their spectral structure (see Figure 2), we can leverage our multispectral measurements (Figure 4) to disambiguate the concentrations of the different biophysical components of our skin model (C_{m}, β_{m}, $C_{\text{h,in}}$, and $C_{\text{h,out}}$).

We minimize the photometric loss between $\bar{\rho}_{sss}$ and the rendered subsurface scattering using the full diffusion profile. The main challenge is the differentiation of the diffusion profile with respect to the biophysical parameters, since forward optimization is neither efficient nor scalable for high-resolution textures [Donner et al. 2008]. To tackle the lack of end-to-end derivatives, we propose a coordinate descent method [Wright 2015] using alternating least squares. Our optimization is thus split into two subproblems: first, obtaining the weights of the Gaussians that define $\bar{\rho}_{sss}$; second, estimating the biophysical parameters of $\bar{\rho}_{sss}$.

We discretize the spectral absorptions into fifteen multispectral channels. To calculate the photometric loss, we convert these channels into our camera's six channels using our system's spectral calibration functions. For efficiency, early-stage iterations are calculated at a coarser resolution. We use nine Gaussians to approximate each profile and merge this subsurface scattering with the contributions from the specular and single scattering components to render the full appearance model. This efficient method enables fast gradient descent iterations to approximate the skin's biophysical parameters (Supplemental Section 5.3).

6 Results and Validation

We estimate a per-texel polarimetric BSSRDF that consists of a three-by-three Mueller matrix of linear polarization. Note that our reflectance function is nine times larger than the conventional BSSRDF. Our code runs on a machine equipped with an AMD EPYC 7763 CPU of 2.45 GHz and an NVIDIA A100 GPU. In the first stage of our method, polarimetric appearance optimization takes around 180 minutes with 200 frames, while biophysical multispectral optimization takes 50 minutes. In the second stage, polarimetric optimization takes around 180 minutes with 50 frames, and biophysical optimization takes about 20 minutes.

We illustrate the versatility of our reconstructions on eleven subjects with different skin tones, genders, and ethnicities, performing various dynamic facial expressions. Our method requires a near-coaxial light-camera configuration, with the subject's face positioned at the optical center of our polarimetric imaging unit during the capture. In our first initialization stage, participants maintain a neutral facial expression while the lighting and viewpoint angles change by spinning their heads. In the second stage, we instruct

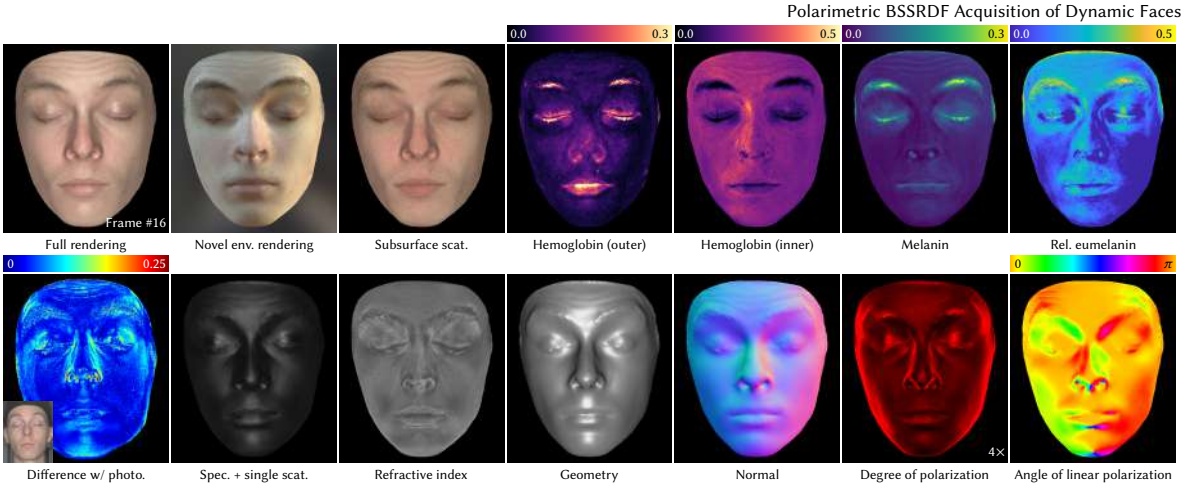


Fig. 6. Our dynamic face reconstruction results. We record data from 11 individuals, each exhibiting diverse skin tones, gender identities, and ethnic backgrounds. Our method successfully captures polarimetric reflectance parameters, biophysical parameters, and geometry with high accuracy. Refer to the supplemental document (Appendix A) and video for more results of dynamic faces.

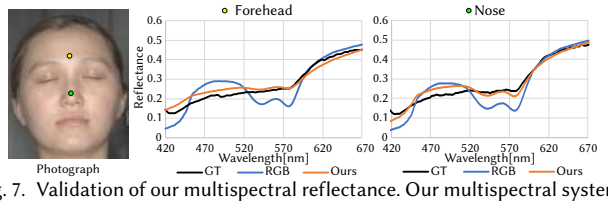
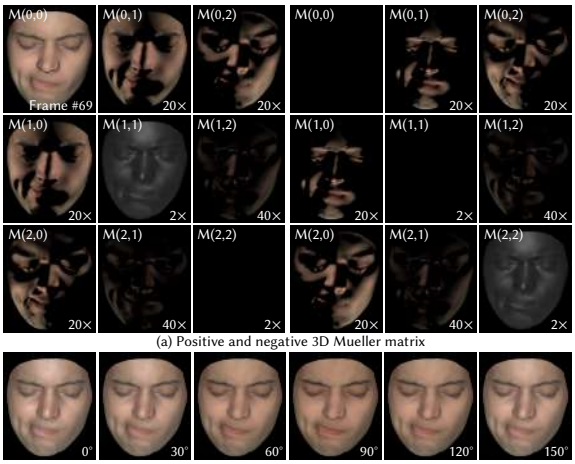


Fig. 7. Validation of our multispectral reflectance. Our multispectral system enables accurate reconstructions of spectral reflectance compared to the ground-truth measurements. Results using trichromatic channels exhibit larger deviations from the ground truth.

the participants to perform various facial expressions. Note that we use a **completely novel viewpoint of an RGB** camera to render and validate the results. Please refer to the supplemental video for the capturing environments of our setup. Figure 6 and the appendix of the supplemental document show, for each subject, the resulting polarimetric BSSRDF, biophysical skin parameters, refractive index, normals, and geometry. These parameters present distinct variations across different subjects, according to their skin tone.

Validation. We first validate our results comparing our estimated spectral reflectance with the ground-truth reflectance measured by a hyperspectral camera (SpecIM). Figure 7 shows how our reconstruction results closely match the ground-truth measurements. In addition, to validate the accuracy of the refractive index that our method estimates, we compare our estimated refractive indices with the reference refractive indices of spherical objects, measured by their Brewster angles [Baek et al. 2020]. As shown in Table 1, our system can measure the refractive indices of objects with high accuracy.

Polarimetric reflectance. Our work is the first to capture polarimetric reflectance functions of human faces in the form of the 3×3 Mueller matrices, as shown in Figure 8(a). It allows us to simulate polarimetric face appearance changes by the linear polarization angle changes on the camera (Figure 8(b)). Also, this enables us to explore other polarization metrics, such as the angle of linear polarization (AoLP) or the degree of polarization (DoP), as shown in the second row of Figure 6. We also show the captured index of refraction, which is crucial for our optimization and is only obtainable thanks to this polarimetric information.



(a) Positive and negative 3D Mueller matrix

(b) Various camera polarization angle

Fig. 8. Our polarimetric reflectance results. (a) Positive and negative values of polarimetric reflectance function as a 3D Mueller matrix. We scale the intensity of each image for visualization purposes. (b) Polarization rendering simulates various polarimetric reflections of linear polarization angles for the camera.

Table 1. Validation on refractive index measurements. We compare our estimations η_{ours} with known refractive indices η_{gt} of ten real-world objects. We achieve a high accuracy with the mean reconstruction error of 0.028.

Object	Material	η_{gt}	η_{ours}	Diff.
1	Red billiard	1.485	1.446	0.038
2	Green billiard	1.469	1.516	0.047
3	Blue billiard	1.504	1.503	0.001
4	White billiard	1.463	1.410	0.053
5	POM	1.462	1.447	0.015
6	Fake pearl	2.295	2.263	0.032
7	Yellow silicone	1.303	1.297	0.005
8	Pink silicone	1.177	1.211	0.034
9	White silicone	1.248	1.272	0.024
10	Light green silicone	1.343	1.311	0.032

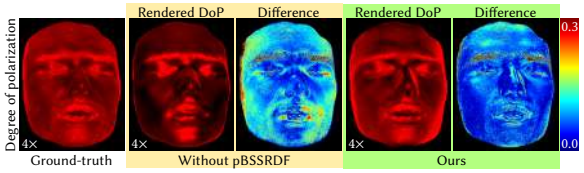


Fig. 9. Validation of our polarimetric reflectance. We compare the degree of polarization of our estimated pBSSRDF with that of the ground-truth measurement by a polarization camera.

Table 2. Ablation study of the impact of each component in the specular-dominant polarization observation. We average the RMSE values over 200 different novel camera poses with 11 different participants.

Refractive index and azimuthal loss	Single scattering parameters	Refractive index parameter	Average RMSE
✔	-	-	$4.44 \cdot 10^{-2}$
✔	✔	-	$4.13 \cdot 10^{-2}$
✔	✔	✔	$\mathbf{3.94 \cdot 10^{-2}}$

Existing methods of face capture [Azinović et al. 2023; Gotardo et al. 2018; Riviere et al. 2020] use a single refractive index value being assumed for the entire face region. In contrast, we optimize spatially-varying refractive-index values as well as polarimetric appearance parameters, which are critical to achieve accurate reconstruction of polarimetric appearance as shown in Figure 9. The skin on the face has varying concentrations of oil and moisture, with higher levels on the forehead and nose compared to the cheeks or lips. In order to validate our pBSSRDF measurement, we compare the degree of polarization of our polarimetric rendering with that of the ground-truth measurement captured by a reference polarization camera. Our polarimetric rendering with the estimated pBSSRDF demonstrates a strong agreement with the ground-truth measurement of the degree of polarization, showing high accuracy. Moreover, as shown in Table 2, leveraging the polarimetric loss term with spatially varying single scattering and refractive index parameters results in the minimum RMSE error on the specular-dominant polarization observation.

Multispectral optimization. Figure 10 shows our multispectral optimization results rendered from our estimated biophysically-based parameters. It can be seen how, as expected, sharper details can be recovered at shorter wavelengths, since longer wavelengths scatter further inside the skin [Donner and Jensen 2006].

Moreover, our method can handle dynamic changes in the appearance of different nature. Figure 11 shows time-varying changes in the distribution of the biophysically-based parameters caused by wrinkles in the forehead, while Figure 12 illustrates changes in the hemoglobin concentration due to applied pressure. At Frame #1, such concentration is lower around the pressed region, but after approximately two seconds, blood re-enters the area (see Frames #10 and #20). As expected, there is no significant change in melanin concentration.

Comparison with prior works. We directly compare our results with the recent, state-of-the-art face acquisition method of Riviere et al. [2020], and the polarimetry method of Hwang et al. [2022] with static scenes as shown in Figure 13. Comparisons with the work by Riviere et al. [2020] are difficult, since there is no publicly available code or dataset. Therefore, we have implemented their method based on our own framework, increasing the number of input images from twelve in the original work to 200, to provide a more fair comparison.

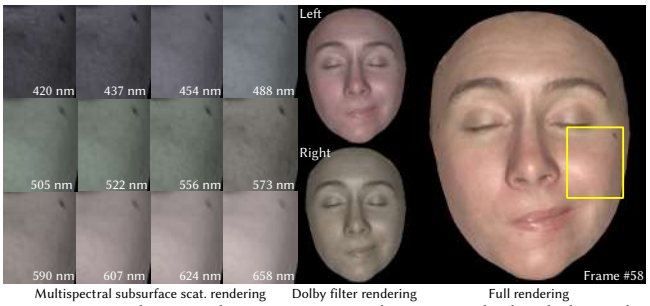


Fig. 10. Our multispectral optimization results. Our method includes multispectral subsurface-scattering rendering across 15 different wavelengths, revealing wavelength-dependent scattering characteristics, like enhanced texture detail at shorter wavelengths. We generate RGB images from these multispectral images using Dolby left/right-filter responses, supplementing the standard sRGB image.

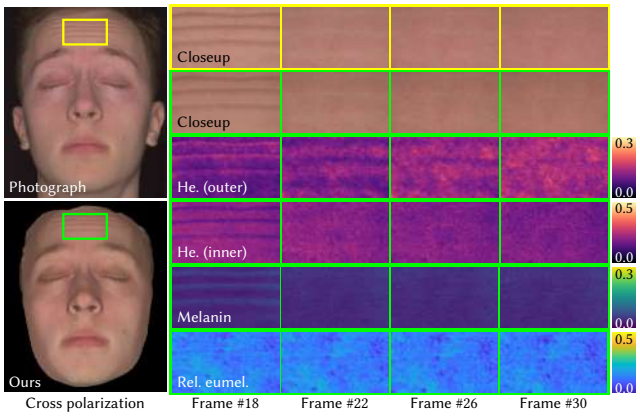


Fig. 11. Dynamic parameter changes. Our captured biophysical parameter map shows both blood flow changes and melanin map changes, caused by the wrinkles, being consistent with the dynamic appearance of the forehead.

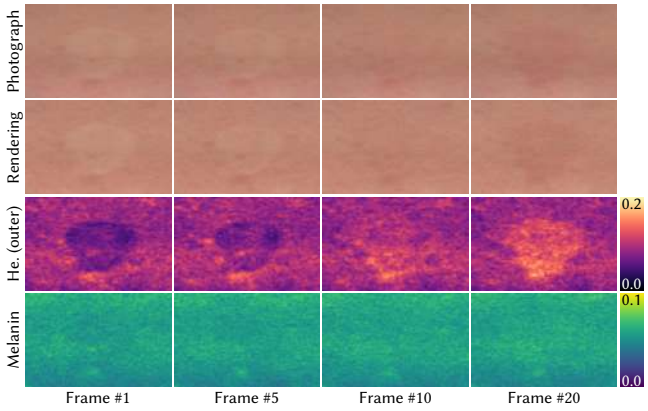


Fig. 12. Blood flow alterations in the forehead region after removing applied pressure. Hemoglobin concentrations in the outer layer return to their previous level as blood re-enters the pressed region.

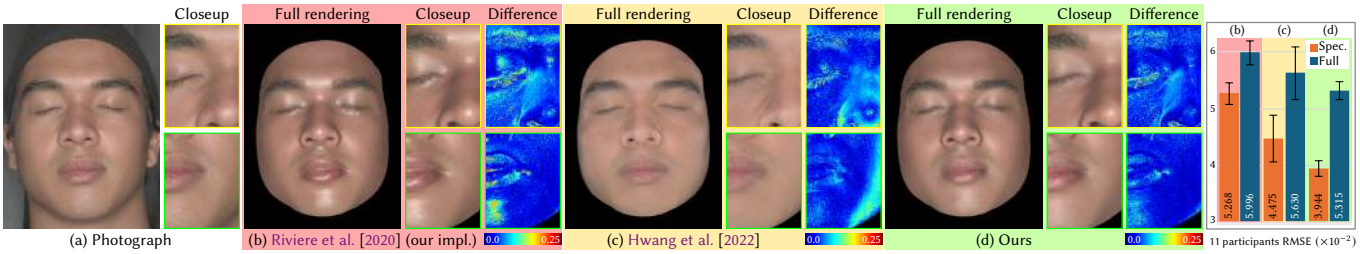


Fig. 13. Comparison of appearance with state-of-the-art face acquisition model [Riviere et al. 2020] using our implementation and static polarimetric acquisition [Hwang et al. 2022]. Riviere et al. [2020] model over-estimated the specular highlights due to the absence of spatially-varying specular roughness, single scattering, and refractive index. Hwang et al. [2022] fails to reconstruct the specular highlights in the forehead and nose (yellow) and baked the specularity in the cheek (green) due to their alternative and cluster-based optimization scheme. Our method successfully reconstructs the highlight details of specular reflection in the nose and forehead region, together with detailed pore-level appearance. We also compute the average RMSE of the 200 different poses on a novel RGB camera on 11 participants, as shown in the right plot. Our method achieves the smallest reconstruction error on both specular and full rendering images.

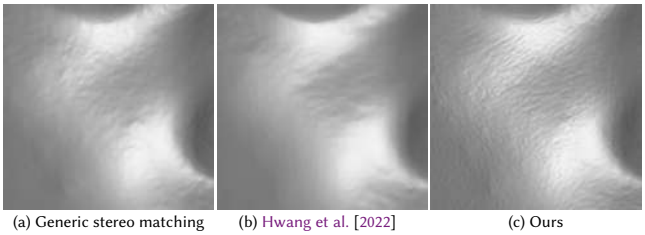


Fig. 14. Comparison of geometry reconstruction. By leveraging multiview polarimetric information in our optimization of the displacement map, our method yields more precise geometric reconstructions. (a) Generic stereo matching (similar to Beeler et al. [2010]'s method without mesoscopic augmentation). (b) Poisson-based inverse rendering [Hwang et al. 2022]. (c) Our method.

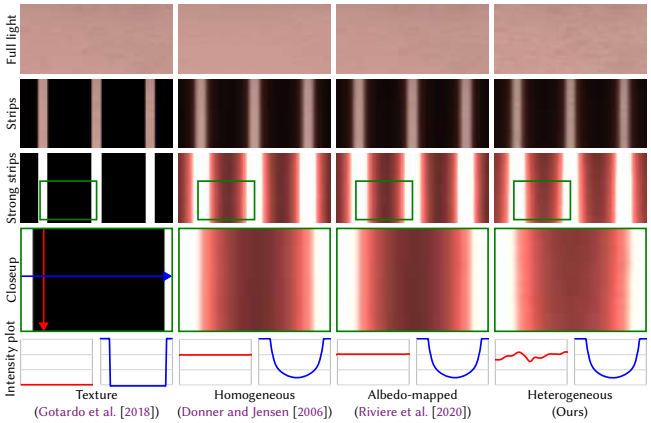


Fig. 15. Full and strip light rendering results of texture-based [Gotardo et al. 2018], homogeneous subsurface scattering-based [Donner and Jensen 2006], homogeneous subsurface scattering with albedo-map based on [Riviere et al. 2020], and our heterogeneous subsurface scattering method using estimated parameter maps. Each plot shows the intensity variation across the vertical cross-section line (red) and the horizontal line (blue).

As shown in the figure, assuming homogeneous specular roughness and refractive index makes Riviere et al. [2020] overestimate specular highlights. On the other hand, Hwang et al. [2022] underestimates them due to their alternative, cluster-based optimization scheme. Our reconstruction of specular highlights and overall reflectance is more accurate, thanks to spatially varying specular roughness, single scattering, refractive index, and joint optimization. We also compute the RMSE of the specularity and the full rendering (right bar plot) images in 200 different poses of the 11 participants. Our method gives the smallest RMSE value on both images.

In terms of geometric accuracy, Figure 14 shows how our approach leads to artifact-free, more detailed reconstructions than previous approaches using generic stereo-matching [Beeler et al. 2010] without mesoscopic augmentation, or Poisson-based inverse rendering optimization [Hwang et al. 2022].

Moreover, Figure 15 compares the impact of our two-layer heterogeneous model as similar to that of Donner et al. [2008], by using structured light patterns. Previous methods based on diffuse albedo [Gotardo et al. 2018; Hwang et al. 2022] cannot accurately simulate subsurface scattering in human skin. Although homogeneous subsurface scattering models [Donner and Jensen 2006] combined with albedo-mapped models [Riviere et al. 2020] can simulate subsurface scattering, they cannot clearly depict the heterogeneity of the spatially varying parameters, as demonstrated in our method.

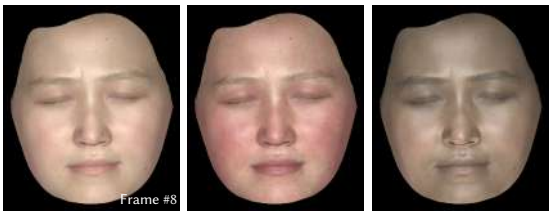


Fig. 16. Editing face biophysical parameters. (a) Rendering with the original parameters. (b) Increased outer layer hemoglobin (2×). (c) Increased melanin (4×).

Face appearance editing. Lastly, our model enables the editing and the exploration of the effect of the different components on the skin's final appearance. Figure 16 shows how changes in hemoglobin and melanin affect such appearance. As expected, when the hemoglobin concentration in the outer layer increases, the subject's skin tone becomes reddish, while an increase in melanin leads to a more tanned appearance.

7 Discussion and Limitations

We have presented a novel polarimetric imaging system that obtains 3D geometry and polarimetric reflectance of dynamic deformable surfaces made of translucent materials, for the particular case of human faces. The system is comprised of multispectral polarization cameras, polarized light sources and stereo imaging modules. Our skin BSSRDF model is two-layered, based on the main biophysically-based components, which we approximate through a multispectral optimization based on their distinctive spectral profiles.

Our system's spatial resolution is half (2K) of 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. Additionally, our off-the-shelf Dolby filters have spectral overlaps at specific wavelengths (570nm–620nm, Figure 4); the three-channel sRGB colors converted from our six-channel multispectral optimization via linear color transformation might thus exhibit a subtle error. Theoretically, this could be fixed by using custom bandpass filters that do not overlap. Last, we do not explicitly include ambient occlusion to make the optimization manageable. As a consequence, for high-frequency geometry variations ambient occlusion shading may be misinterpreted as reflectance, which in turn may lead to small, local melanin variations.

Our optimization obtains a set of parameter maps, both polarimetric and biophysically-based. While we have validated the accuracy of some of the obtained parameters (index of refraction), we cannot claim that each individual parameter (particularly, the biophysically-based ones) is fully accurate. Still, the resulting global polarimetric appearance is a good match w.r.t. the input, and the behavior of each component is plausible.

We have shown results across a wide spectrum of skin tones. However, we have noticed that inner layer components, especially hemoglobin, may be underestimated for subjects with very dark skin (Figure 17). This is because estimating non-invasive in-vivo biophysical parameters relies on the energy returned from the skin. When such energy is low, both existing methods and commercial products [Fawzy et al. 2022; Shi et al. 2022] may fail. This is therefore a common issue affecting very dark skin tones. Our methodology, nevertheless, could be applied to different multi-layered appearance models with different compositions, which is an interesting avenue for future work. Moreover, applying other spectral illumination setups [Aliaga et al. 2023; Gitlina et al. 2020; Preece and Claridge 2004] is an exciting exploration for future research.

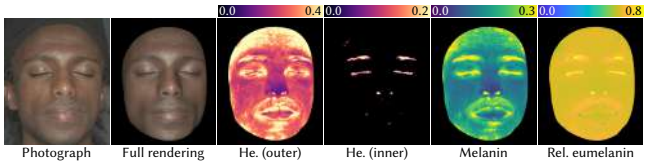


Fig. 17. Limitation example with very dark skin. While our algorithm successfully reconstructs both the geometry and appearance, it underestimates the contribution of the inner layer of hemoglobin because most of the incident light gets absorbed by melanin and eumelanin present in the outer layer.

Acknowledgments

Min H. Kim acknowledges the Korea NRF grant (RS-2024-00357548), the MSIT/IITP of Korea (RS-2022-00155620, RS-2024-00398830, 2022-0-00058, and 2017-0-00072), Microsoft Research Asia, LIG, and Samsung Electronics. Nestor Monzon acknowledges a Gobierno de Aragon predoctoral grant (2023–2027).

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.

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*.

Yunhao Ba, Alex Gilbert, Franklin Wang, Jinfa Yang, Rui Chen, Yiqin Wang, Lei Yan, Boxin Shi, and Achuta Kadambi. 2020. Deep Shape from Polarization. In *Proc. ECCV 2020*. Springer, 554–571.

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 and Felix Heide. 2022. All-photon Polarimetric Time-of-Flight Imaging. In *Proc. IEEE/CVF CVPR*. 17876–17885.

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.

Seung-Hwan Baek, Tizian Zeltner, Hyunjin Ku, Inseung Hwang, Xin Tong, Wenzel Jakob, and Min H. Kim. 2020. Image-Based Acquisition and Modeling of Polarimetric Reflectance. *ACM Transactions on Graphics* 39, 4 (2020), 139.

Alexey N. Bashkatov, E. A. Genina, V. I. Kochubey, and V. V. Tuchin. 2005. Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm. *Journal of Physics D: Applied Physics* 38, 15 (2005), 2543.

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).

Chen Cao, Derek Bradley, Kun Zhou, and Thabo Beeler. 2015. Real-Time High-Fidelity Facial Performance Capture. *ACM Transactions on Graphics* 34, 4 (2015), 9 pages.

Xu Cao, Hiroaki Santo, Fumio Okura, and Yasuyuki Matsushita. 2023. Multi-View Azimuth Stereo via Tangent Space Consistency. In *Proc. IEEE/CVF CVPR*. 825–834.

Tenn F Chen, Gladimir VG Baranoski, Bradley W Kimmel, and Erik Miranda. 2015. Hyperspectral modeling of skin appearance. *ACM Transactions on Graphics (TOG)* 34, 3 (2015), 1–14.

Zhaopeng Cui, Viktor Larsson, and Marc Pollefeys. 2019. Polarimetric Relative Pose Estimation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2671–2680.

Akshat Dave, Yongyi Zhao, and Ashok Veeraraghavan. 2022. Pandora: Polarization-aided neural decomposition of radiance. In *European Conference on Computer Vision*. Springer, 538–556.

Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. 2000. Acquiring the Reflectance Field of a Human Face. In *Annual Conference Series (Proceedings of SIGGRAPH)*. 145–156.

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.

Valentin Deschaintre, Yiming Lin, and Abhijeet Ghosh. 2021. Deep Polarization Imaging for 3D Shape and SVBRDF Acquisition. In *Proc. IEEE/CVF CVPR*. 15567–15576.

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 and Henrik Wann Jensen. 2006. A Spectral BSSRDF for Shading Human Skin. *Rendering techniques* 2006 (2006), 409–418. Publisher: Citeseer.

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.

Ashraf Fawzy, Tianshi David Wu, Kunbo Wang, Matthew L. Robinson, Jad Farha, Amanda Bradke, Sherita H. Golden, Yanxun Xu, and Brian T. Garibaldi. 2022. Racial and ethnic discrepancy in pulse oximetry and delayed identification of treatment eligibility among patients with COVID-19. *JAMA internal medicine* 182, 7 (2022), 730–738. Publisher: American Medical Association.

Graham Fyffe. 2010. Single-shot photometric stereo by spectral multiplexing. In *ACM SIGGRAPH ASIA 2010 Sketches*. 1–2.

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.

Pablo Garrido, Levi Valgaert, Chenglei Wu, and Christian Theobalt. 2013. Reconstructing Detailed Dynamic Face Geometry from Monocular Video. *ACM Transactions on Graphics* 32, 6 (2013), 10 pages.

Abhijeet Ghosh, Tongbo Chen, Pieter Peers, Cyrus A. Wilson, and Paul Debevec. 2010. Circularly polarized spherical illumination reflectometry. In *ACM SIGGRAPH Asia 2010 papers*. 1–12.

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.

Cong Phuoc Huynh, Antonio Robles-Kelly, and Edwin R. Hancock. 2013. Shape and Refractive Index from Single-View Spectro-Polarimetric Images. *International journal of computer vision* 101, 1 (2013), 64–94. Publisher: Springer.

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.

Alexandru Eugen Ichim, Sofien Bouaziz, and Mark Pauly. 2015. Dynamic 3D Avatar Creation from Hand-Held Video Input. *ACM Transactions on Graphics* 34, 4 (2015), 14 pages.

Jose A. Iglesias-Guitian, Carlos Aliaga, Adrian Jarabo, and Diego Gutierrez. 2015. A Biophysically-Based Model of the Optical Properties of Skin Aging. *Computer Graphics Forum (EUROGRAPHICS 2015)* 34, 2 (2015).

Francisco H. Imai. 1998. Multi-spectral image acquisition and spectral reconstruction using a trichromatic digital camera system associated with absorption filters. *Munsell Color Science Laboratory, Rochester Institute of Technology* 2 (1998).

Steven L. Jacques. 1998. Skin Optics Summary. *Oregon Medical Laser Center News* (1998).

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.

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

Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. 2006. Poisson Surface Reconstruction. In *Proc. Eurographics symposium on Geometry processing*, Vol. 7. 0.

Aravind Krishnaswamy and Gladimir VG Baranoski. 2004. A biophysically-based spectral model of light interaction with human skin. In *Computer graphics forum*, Vol. 23. Wiley Online Library, 331–340.

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.

Chenyang Lei, Chenyang Qi, Jiaxin Xie, Na Fan, Vladlen Koltun, and Qifeng Chen. 2022. Shape from Polarization for Complex Scenes in the Wild. In *Proc. IEEE/CVF CVPR*. 12632–12641.

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.

Xiaohui Li, Giuseppe Claudio Guarnera, Arvin Lin, and Abhijeet Ghosh. 2024. Practical Measurement and Neural Encoding of Hyperspectral Skin Reflectance. In *Proc. Int. Conf. on 3D Vision (3DV)*. 1301–1309.

Lahav Lipson, Zachary Teed, and Jia Deng. 2021. Raft-Stereo: Multilevel Recurrent Field Transforms for Stereo Matching. In *Proc. Int. Conf. on 3D Vision (3DV)*. 218–227.

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).

Daisuke Miyazaki, Robby T. Tan, Kenji Hara, and Katsushi Ikeuchi. 2003. Polarization-based Inverse Rendering from a Single View. In *Proc. IEEE ICCV*, Vol. 3. 982–982.

Koki Nagano, Graham Fyffe, Oleg Alexander, Jernej Barbič, Hao Li, Abhijeet Ghosh, and Paul Debevec. 2015. Skin Microstructure Deformation with Displacement Map Convolution. *ACM Transactions on Graphics* 34, 4, Article 109 (2015), 10 pages.

Scott Prahl. 1999. Optical Absorption of Hemoglobin. http://omlc.ogi.edu/spectra/hemoglobin. Publisher: Oregon Medical Laser Center, Accessed: Sept. 08, 2024.

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.

Jérémy Riviere, Ilya Reshetouski, Luka Filipi, and Abhijeet Ghosh. 2017. Polarization imaging reflectometry in the wild. *ACM Transactions on Graphics* 36, 6 (2017), 1–14.

Soumyadip Sengupta, Angjoo Kanazawa, Carlos D. Castillo, and David W. Jacobs. 2018. SfSNet: Learning Shape, Reflectance and Illuminance of Faces in the Wild. In *Proc. IEEE/CVF CVPR*. 6296–6305.

Chunhu Shi, Mark Goodall, Jo Dumville, James Hill, Gill Norman, Oliver Hamer, Andrew Clegg, Caroline Leigh Watkins, George Georgiou, Alexander Hodkinson, Catherine Elizabeth Lightbody, Paul Dark, and Nicky Cullum. 2022. The accuracy of pulse oximetry in measuring oxygen saturation by levels of skin pigmentation: a systematic review and meta-analysis. *BMC Medicine* 20, 1 (Aug. 2022), 267.

Fuhao Shi, Hsiang-Tao Wu, Xin Tong, and Jinxiang Chai. 2014. Automatic Acquisition of High-Fidelity Facial Performances Using Monocular Videos. *ACM Transactions on Graphics* 33, 6 (2014), 13 pages.

Raju Shrestha and Jon Yngve Hardeberg. 2010. Multispectral image capture using two RGB cameras. In *2010 18th European Signal Processing Conference*. IEEE, 1801–1805.

Silvia Tozza, William AP Smith, Dizhong Zhu, Ravi Ramamoorthi, and Edwin R. Hancock. 2017. Linear Differential Constraints for Photo-Polarimetric Height Estimation. In *Proc. IEEE/CVF CVPR*. 2279–2287.

Anh Tuan Tran, Tal Hassner, Iacopo Masi, Eran Paz, Yuval Nirkin, and Gerard Medioni. 2018. Extreme 3D Face Reconstruction: Seeing Through Occlusions. In *Proc. IEEE/CVF CVPR*. 3935–3944.

Luan Tran, Feng Liu, and Xiaoming Liu. 2019. Towards High-Fidelity Nonlinear 3D Face Morphable Model. In *Proc. IEEE/CVF CVPR*. 1126–1135.

Luan Tran and Xiaoming Liu. 2019. On Learning 3D Face Morphable Model from In-the-wild Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 1 (2019), 157–171.

Norimichi Tsumura, Hideaki Haneishi, and Yoichi Miyake. 1999. Independent-component analysis of skin color image. *JOSA A* 16, 9 (1999), 2169–2176.

Norimichi Tsumura, Nobutoshi Ojima, Kayoko Sato, Mitsuhiro Shiraishi, Hideto Shimizu, Hirohide Nabeshima, Syuuichi Akazaki, Kimihiko Hori, and Yoichi Miyake. 2003. Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. In *ACM SIGGRAPH 2003 Papers*. 770–779.

Levi Valgaerts, Chenglei Wu, Andrés Bruhn, Hans-Peter Seidel, and Christian Theobalt. 2012. Lightweight Binocular Facial Performance Capture under Uncontrolled Lighting. *ACM Transactions on Graphics* 31, 6 (2012), 11 pages.

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.

Jinyu Zhao, Yusuke Monno, and Masatoshi Okutomi. 2022. Polarimetric Multi-View Inverse Rendering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022). Publisher: IEEE.

Dizhong Zhu and William AP Smith. 2019. Depth From a Polarisation + RGB Stereo Pair. In *Proc. IEEE/CVF CVPR*. 7586–7595.