Practical Measurement and Neural Encoding of Hyperspectral Skin Reflectance

Xiaohui Li[1,2], Giuseppe Claudio Guarnera[2,3], Arvin Lin[1,2] and Abhijeet Ghosh[1,2]

[1]Imperial College London [2]Lumirithmic Ltd. [3]University of York

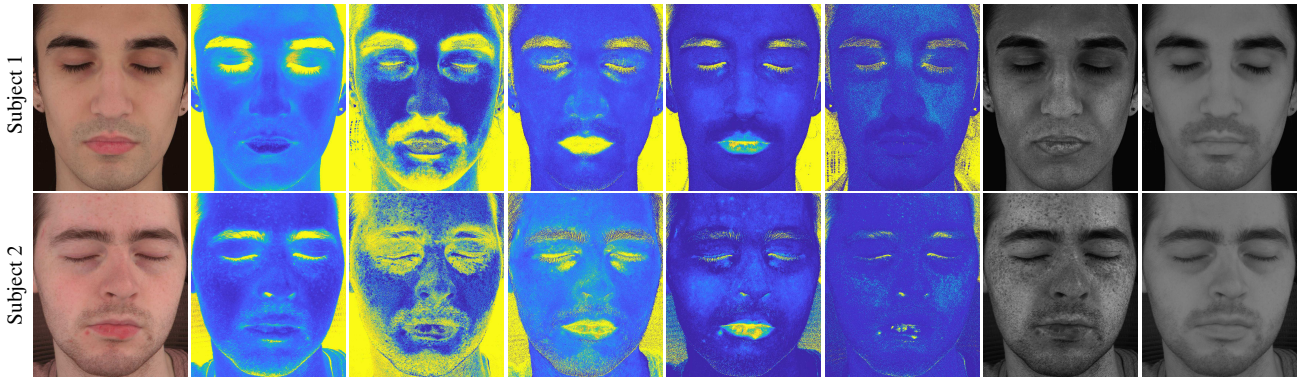


Figure 1. Reconstructions of acquired faces of subject 1 (Mediterranean skin type, top row) and subject 2 (Caucasian skin type, bottom row) under uniform white illumination (a), estimated skin chromophore maps (b–f), and predicted UVA and NIR responses (g, h) using our proposed practical spectral measurement and modeling approach.

Abstract

We propose a practical method to measure spectral skin reflectance as well as a spectral BSSRDF model spanning a wide spectral range from 300nm to 1000nm. We employ a practical capture setup consisting of desktop monitors to illuminate human faces in the visible domain to estimate five parameters of spectral chromophore concentrations including melanin, hemoglobin, and β carotene concentration, melanin blend-type fraction, and epidermal hemoglobin fraction. The estimated parameters make use of a novel three-stage lookup table search for faster parameter fitting, and drive our skin model for accurate reconstruction of facial skin reflectance response in both the visible domain as well as in the UVA and near-infrared range. We also propose a novel neural network architecture that given our measurements, predicts the five chromophore parameters of our model at the encoder stage and full hyperspectral reflectance response as the output of the decoder stage.

1. Introduction

Accurate facial appearance modeling has been a topic of extensive research in computer vision and graphics. Several biophysically-based spectral skin reflectance models have been proposed in recent years [1, 7, 10–12, 17, 18]. Previous works have either proposed complex biophysical models with a large number of parameters to represent various chromophores in the human skin [7, 11, 18], or employed simpler models with fewer parameters that can be practically measured [12, 17], albeit with reduced model accuracy. These works have mainly focused on the visible domain, in which some facial features that are crucial for skin appearance and protection might be concealed [7].

In this work, we employ a diffusion-based hyperspectral BSSRDF skin appearance model spanning a wide spectral range from 300nm to 1000nm, thus encompassing Ultra Violet (UV), Visible (Vis) and Near InfraRed (NIR). While the employed model makes use of the spectral chromophore parameters proposed in [11], we propose a novel extension of their model to predict skin reflectance response in the UVA and NIR domain (Section 3). Therefore, the employed parameters are melanin concentration C_m,

dermal hemoglobin concentration C_{hd}, melanin blend-type fraction β_m, epidermal hemoglobin fraction C_{he}, and β carotene C_{bc} concentration. The skin model additionally includes two fixed parameters: dermal water fraction C_{wd} and epidermal water fraction C_{we}. The inclusion of β carotene in the model enables more accurate skin appearance simulation, particularly for yellowish and olive skin types. Further, we present a practical measurement approach for acquiring spectral skin reflectance based on our model. Our capture setup consists of a set of desktop LCD monitors for illuminating a face with standard visible spectrum RGB illumination in order to acquire high-quality spectral skin chromophore maps.

Previous works estimate spectral skin parameters to achieve high-quality appearance reconstructions by searching in a LookUp Table (LUT) of pre-computed skin tones the closest RGB values of facial albedo images, on a pixel by pixel basis [12, 17]. Such an approach is computationally expensive, particularly for models with higher complexity. To address this issue, we propose a novel three-stage LUT search to efficiently estimate the parameters of our model (see Section 4).

An additional issue of the LUT search approach is related to the quantization of the results it produces, due to the sampling resolutions of parameters used to generate the LUT. To overcome this limitation, we design an Encoder-Decoder network to estimate the five spectral parameters at the encoder stage ($ChromNN$), and a decoder network ($SpectraNN$) for reconstructing the full hyperspectral skin reflectance accurately. Here, hyperspectral reflectance of skin is first compressed using Principal component analysis (PCA), which allows reducing the size of the $SpectraNN$ network, thus speeding up training and testing (Section 6). We validate our model parameters estimated from RGB measurements using comparison of hyperspectral skin appearance reconstructions to reference photographs under various spectral illumination conditions.

In summary, our main contributions are as follows:

- We propose a practical hyperspectral skin reflectance model driven by five chromophore parameters spanning a wide spectral range from 300 to 1000nm, and a practical measurement approach for estimating model parameters employing RGB illumination emitted by desktop LCD monitors.
- We introduce a novel three-stage lookup table search for speeding up conventional LUT-based parameter fitting that takes advantage of our spectral measurements.
- We design a neural network architecture to estimate the chromophore parameters of our skin model and reconstruct full hyperspectral reflectance from the input RGB measurements based on a novel PCA analysis that enables reduction of network complexity.

2. Related Work

Here, we give a brief overview of bio-physical skin appearance modeling, spectral skin appearance measurement, and hyperspectral appearance modeling (including UV and IR) in computer vision, graphics and tissue-optics.

Bio-physical Skin Appearance Modeling: Jensen et al. [16] proposed a practical BSSRDF model to approximate the subsurface scattering for translucent materials with dipole diffusion theory. This work was extended to multi-pole diffusion theory to account for multi-layered translucent materials, such as human skin [9]. Building on these, they introduced a two-layer bio-physical skin model with three chromophore parameters to simulate realistic skin appearance [10]. This model was then extended and simplified by Jimenez et al. [17] to make it suitable for practical measurement and prediction of skin color change during facial animation using a four parameter model with only two free parameters used for fitting using a lookup-table (LUT) search. Recently, Gitlina et al. [12] employed the more complete spectral model of [17], with four free parameters to more accurately reconstruct skin appearance using measurements in a multispectral LED sphere. Besides employing LUT searching, they also proposed a cascaded neural network for model fitting and RGB albedo reconstruction.

The spectral skin chromophores modeled in this work are mainly related to melanin and hemoglobin concentrations. However, there are some other blood-borne pigments, such as beta-carotene and bilirubin, which also play important roles in skin color [7, 11, 18]. Krishnaswamy et al. [18] introduced a parametric five-layer skin model called BioSpec to approximate the light interaction within human skin. Donner et al. [11] later proposed a layered heterogeneous reflectance model with lateral inter-scattering of light between skin layers using six parameters and an additional inter-layer absorption to model veins and tattoos in skin [11]. Although these models can simulate very realistic skin appearance, high model complexity makes measurements less practical.

Our work follows the line of previous diffusion-based bio-physical skin appearance models. We employ a practical hyperspectral skin model with five free parameters to balance model sophistication and complexity for measurements. We also propose a novel Encoder-Decoder network to predict chromophore maps and reconstruct full spectral skin reflectance given practical RGB measurements. Similar to our work, recent methods have employed neural networks to predict chromophores [12]. However, their network architectures focus on reconstructing RGB facial appearance from the estimated chromophores, under a specific illumination spectrum (e.g., D65). Closer to our approach, Aliaga et al. [1] presented a network that directly predicts spectral skin reflectance in the Vis and NIR range. In comparison, our proposed network architecture

allows predicting hyperspectral reflectance in a wider range (UV+Vis+NIR), while using only a fraction ($\sim 2\%$) of the overall number of network parameters.

Spectral Skin Appearance Measurement: Jensen et al. [16] proposed a simple technique for measuring RGB optical parameters of materials, including two skin samples, by illuminating the surface of a sample with a tightly focused beam of white light and acquiring a photograph with a color camera to observe the radiant exitance across the entire surface. Tsumura et al. [23] used independent component analysis on photographs to estimate melanin and hemoglobin maps, assuming such chromophores have independent effects on skin color and ignoring the influence of scattering and depth. Donner et al. [11] employed multispectral images of skin patches illuminated by a broadband flash and nine different chosen narrow spectral bands to measure spectral skin reflectance and drive parameter estimation for their model. Jimenez et al. [17] used a non-contact SIAscope^{TM} system [8] with cross-polarized flashes to measure the hemoglobin and melanin concentrations of facial skin while Alotaibi and Smith [2] used a single LED illumination spectrum from photographs. Gotardo et al. [13] proposed a multi-view setup with static illumination for dynamic skin albedo estimation. Park et al. [21] used multiplexed spectral illumination for skin reflectance estimation. Closer to our approach, Gitlina et al. [12] designed a measurement approach using two complementary broad and narrow-band spectral illumination conditions using a multispectral LED sphere. They also proposed a practical skin patch measurement approach using an off-the-shelf dermatological imaging device (Miravex Antera 3D camera). Recently, Aliaga et al. [1] employed RGB albedo measurements of facial skin to estimate chromophore parameters for their Monte Carlo simulation based skin model. Our method takes as input photographs under complementary broad and narrow-band illumination conditions produced by LCD screens to estimate parameters of a diffusion-based hyperspectral skin reflectance model.

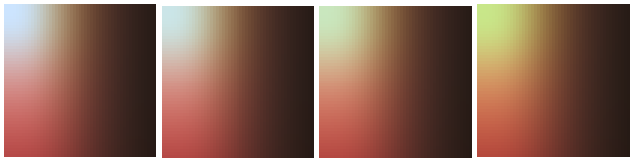
Ultraviolet(UV) and Infrared(IR) Modeling: Some deep skin damage gradually changes facial appearance and skin health due to UV light exposure. The prediction of the pigmentation irregularities such as freckles is beneficial for detection of skin aging [6, 15] and protect skin from photodamage effects and skin cancer risk [25]. In addition to pigmentation, skin hydration is another significant factor to describe skin status [7]. Water displays strong absorption in the IR domain, thus IR photographs can provide information about skin hydration [4].

Chen et al. [7] first introduced a hyperspectral skin appearance model to computer graphics in order to comprehensively forward simulate spectral and spatial distributions of light interacting with human skin from the UV to IR domain involving a large set of skin chromophores. While we borrow from this work, we employ a more practical model with fewer parameters to model skin reflectance over a wide spectral range spanning UVA-Vis-NIR domains.

3. Bio-physical Skin Model and Measurement

3.1. Overview

We propose a diffusion-based hyperspectral skin appearance model driven by five biophysical parameters, namely melanin concentration C_m, melanin type blend β_m (blend between eumelanin and pheomelanin), dermal hemoglobin concentration C_{hd}, epidermal hemoglobin concentration C_{he}, and β carotene fraction C_{bc}. This work adapts and extends the work of [10, 11] into the hyperspectral domain using spectral absorption profiles of the constituent chromophores in the UVA and NIR bands borrowed from [7]. Additionally, we add two parameters C_{wd} and C_{we} to represent water in the epidermis and dermis, respectively. The inclusion of β carotene, an important skin chromophore, extends the skin model color gamut, thus improving the simulation of yellowish and olive human skin tones [3, 11]. Figure 2 shows a small change in the β carotene fraction significantly contributes to the resulting skin color.



(a) $C_{bc} = 0$ (b) $C_{bc} = 0.0013$ (c) $C_{bc} = 0.0042$ (d) $C_{bc} = 0.01$

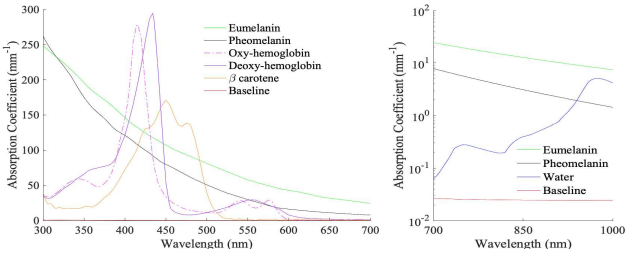
Figure 2. Spectral skin reflectance variations due to C_{bc}. C_m increases along the x-axis, C_{hd} along the y-axis; β_m and C_{he} are kept fixed to 0.5 and 0.25, respectively.

Figure 3 shows the spectral absorption coefficients of the parameters in our model, spanning from 300nm to 1000nm. The selected model parameters not only determine skin color in the visible domain, but they are also significant absorbers in the UVA range. In the NIR range, while both hemoglobin and β-carotene absorption drop, parameters related to melanin and water still significantly affect skin appearance. Overall, the proposed novel equations of absorption of the epidermis and dermis in the $[300 - 1000]nm$ range are as follows:

$$\begin{aligned}\sigma_a^{epi}(\lambda) =&(1 - C_{we})(C_m(\beta_m\sigma_a^{em}(\lambda) + (1 - \beta_m)\sigma_a^{pm}(\lambda)) \\ &+ C_{he}(\gamma\sigma_a^{oxy}(\lambda) + (1 - \gamma)\sigma_a^{deoxy}(\lambda))) + C_{bc}\sigma_a^{bc}(\lambda) \\ &+ (1 - C_m - C_{he} - C_{bc})\sigma_a^{base}) + C_{we}\sigma_a^{w}(\lambda)\ \mathrm{mm}^{-1}\end{aligned} \tag{1}$$

$$\begin{aligned}\sigma_a^{derm}(\lambda) =&(1 - C_{wd})(C_{hd}(\gamma\sigma_a^{oxy}(\lambda) + (1 - \gamma)\sigma_a^{deoxy}(\lambda))) \\ &+ (1 - C_{hd})\sigma_a^{base}) + C_{wd}\sigma_a^{w}(\lambda)\ \mathrm{mm}^{-1}\end{aligned} \tag{2}$$

where λ is the light wavelength in nanometers, C_m represents melanin fraction, β_m represents melanin type blend between eumelanin and pheomelanin, σ_a^{em} and σ_a^{pm} are the



(a) UVA + Vis (b) Vis + NIR

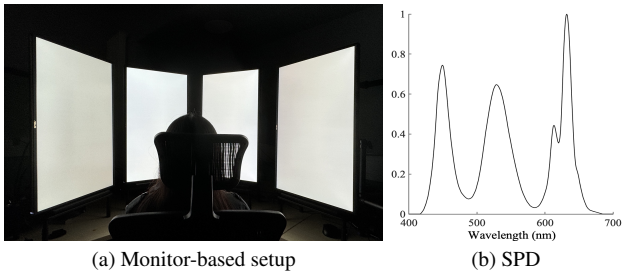
Figure 3. The spectral absorption coefficients curves for the skin chromophores included in our model from 300nm to 1000nm.

absorption coefficients for these two types of melanin. C_{hd} and C_{he} refer to hemoglobin fraction in dermis and epidermis, and their corresponding absorption coefficients are σ_a^{oxy} and σ_a^{deoxy}, with γ being the blood oxygenation ratio between deoxy- and oxy-hemoglobin, fixed as 0.75. C_{bc} and σ_a^{bc} respectively represent the fraction and absorption coefficient of β carotene. C_{we} and C_{wd} are water fractions in the epidermis and dermis fixed as 0.2 and 0.6, respectively; the absorption coefficient of water is σ_a^{w}. σ_a^{base} is the baseline skin tissue absorption. Please refer to the supplemental material for more details of our model and how the LUT data is built.

3.2. Acquisition

We employ a practical setup for our measurements using LCD monitors for illuminating a face similar to the recent work of [19]. Our setup employs four 4K desktop LCD monitors (Asus ProArt PA279CV) and a set of cameras (Canon EOS M6 Mark II) placed in between the monitors for facial measurements. Unlike [19], the cameras in our setup are cross-polarized with respect to the linearly polarized illumination emitted by the monitors (see Figure 4 (a)).

Figure 4 (b) shows the Spectral Power Distribution (SPD) of the screens white illumination (6500K CCT), as measured by a Sekonic SpectroMaster C700 spectrometer. The peak of this blue illumination is at around 450nm. However, according to Preece and Claridge [22], the optimal illumination to measure melanin is narrowband blue with a peak at around 485nm. Therefore, while directly using the narrow band blue illumination in our setup might not be optimal for melanin measurement, it provides cues to estimate β carotene, as the maximal absorption of the latter is closely aligned to the peak of blue illumination. In addition, narrowband green illumination can be used to estimate hemoglobin [22]. As observed in Gitlina et al. [12], directly observing skin response under narrowband illumination can lead to sub-optimal measurements, due to colors outside the gamut of typical off-the-shelf RGB cameras. Hence, we make use of multiple illumination conditions, such as white illumination, a mixture of white and blue illumination, and a mixture of white and green, and employ a computational scheme based on chromatic adaptation transform proposed by [12] to synthesize sharper blue and green narrowband responses from the mixtures (Figure **??**). Therefore, the dimensionality of the input matches that of the estimated parameters:



(a) Monitor-based setup (b) SPD

Figure 4. Monitor-based setup and the SPD of white illumination.

$$(r, g, b, sb, sg) \rightarrow (C_m, \beta_m, C_{hd}, C_{he}, C_{bc})$$

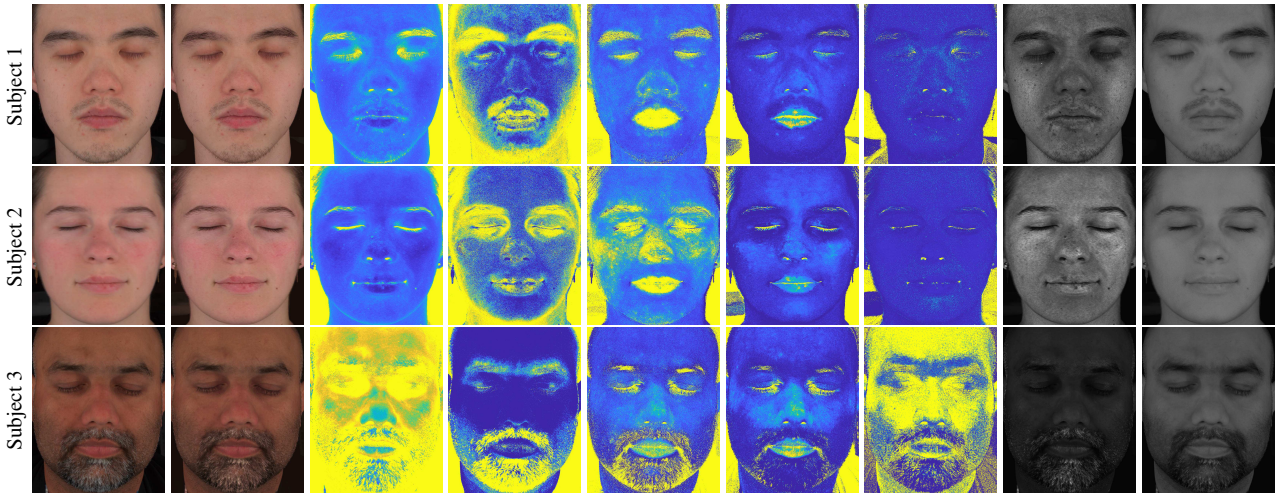
4. Parameter Estimation

The most straightforward approach to parameter estimation is to use pre-computed 5D LookUp Tables (LUTs) of skin colors, one for each of the 3 lighting spectra described in Sec 3.2. For each pixel in the RGB photograph acquired under broadband white light, and in the grayscale synthesized blue and green images, we search the corresponding LUTs, finding the set of skin parameters that minimizes the L2 distance in CIELAB color space (i.e. ΔE_{76} color difference), thus obtaining the estimated skin parameter maps. Radiometric calibration is achieved by scaling the LUTs intensities by constant scalars, derived by taking photographs of the X-Rite color chart under the same illumination conditions. Additionally, we use the optical flow algorithm on the input images to reduce the impact of subjects' movement during the acquisitions.

4.1. Three-stage LUT Searching Method

As shown in previous work, naïve LUT search can provide high-quality results, which we show can be further improved for yellowish and olive skin tones by including β carotene in the model. However, the increased model complexity leads to larger LUTs and longer matching time. In order to address this issue, we propose a novel three-stage matching method given our proposed spectral measurements to take advantage of the larger parameter space, while reducing the computational cost without affecting the quality of the estimated parameter maps. In the following, we provide an overview of the proposed search strategy.

- Stage 1: we keep the β-carotene value fixed, thus focusing on 4D slices of the LUTs. We then query the broadband white and synthetic blue LUTs to estimate the melanin-related parameters C_m and β_m;
- Stage 2: The estimated C_m and β_m in stage 1 are kept constant while usingwhite and synthetic blue LUTs to



(a) Photo (b) Recon (c) C_m (d) β_m (e) C_{hd} (f) C_{he} (g) C_{bc} (h) UV recon (i) IR recon

Figure 5. Comparison of photos (a) and reconstructions (b) and estimated spectral parameters (c - g) for faces of Subject 1 (Asia skin type), Subject 2 (Caucassian skin type), and Subject 3 (South Asia skin type). The estimated chromophore parameters also predict skin appearance in the UVA and NIR bands (h, i).

search the resulting 3D slices for the optimal C_{bc}. While this search also provides estimates for the hemoglobin-related parameters, these are discarded;

- Stage 3: the values of C_m, β_m and C_{bc} estimated in the previous stages are kept fixed, thus leading to a 2D search space. We then use white and synthetic green to search for the two optimal hemoglobin parameters C_{hd} and C_{he}.

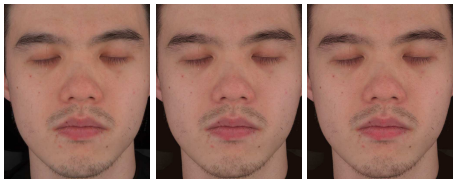
The β-carotene value used in the first stage has been derived as the mode of the distribution of the estimated β-carotene values on a training set containing photographs of subject with different skin types. The estimated values are obtained using the naïve, full search strategy.

Compared to the full search, the above strategy allows a significant speed up of the maps estimation, and allows incorporating knowledge of the most suitable lighting condition to estimate a given skin parameter (see Sec. 3.2). In fact, while blue illumination excite melanin and β carotene (stages 1 and 2), green illumination provides useful information about hemoglobin (stage 3).

4.2. Estimation Results

Figure 5 (a, b) compares the albedo reconstruction using the three-stage matching method with the ground truth photograph, for subjects with Asia, Caucasian, and South Asia skin types, respectively; (c-g) report the estimated chromophore maps while (h, i) displays the UVA and NIR reconstructions. As it can be seen, for all skin types the reconstructions closely match the photographs. The predicted UV response of Subject 3 is lower than that of lighter skin subjects because higher melanin concentration absorbs UV radiation, which is consistent with [7].

Figure 6 compares the reconstruction of an example using the naive full search matching method (b), which inputs the broadband white along with synthetic blue and synthetic green images, with the reconstruction given by the three-stage search (c). As shown in the figure, the results of the two methods are qualitatively similar to the photograph (a).



(a) Photo (b) Full search (c) 3-stage

Figure 6. Comparison of photograph (a), reconstruction using the naive LUT search method, using white, Synth. blue and Synth. green (b), and reconstruction using the 3-stage search (c).

4.3. Skin Appearance Editing

Editing the estimated model parameters allows simulating biophysical changes in skin appearance. By scaling the two melanin-related parameters we can simulate lightening or tanning of skin (C_m), as well as increasing or reducing stubble/facial hair (β_m) present in the input photograph (Figure 7 (a,b)). Similarly, skin appearance can be edited by scaling the beta-carotene parameter C_{bc} (Figure 7 (c)). Editing the two hemoglobin-related parameters C_{hd} and C_{he} allows for changes in facial redness, thus leading to a paler or flushed appearance (Figure 7 (d,e)). Instead of scaling linearly, we apply a simple but effective method to treat two hemoglobin parameters. The scaling equation is $C \times (1 + (s-1) \times i/N)$, where i is the index of the value C in the array of C_{hd} or C_{he} with the size of N while s is the scaling factor. With this scaling, the local variation of

hemoglobin distributions is accentuated in the edits.

(a) Red. melanin (b) Inc. melanin (c) Inc. β carotene (d) Flushed (inc. blood) (e) Pale (red. blood)

Figure 7. Biophysical Skin appearance editing, such as lightening with reduced stubble (a), tanning with increased stubble (b), skin with increased β carotene (c), flushed skin (d), and pale skin (e).

5. Ultraviolet (UV) and Near-infrared (NIR) Validation

As previously noted, the chromophores-related parameters used in our model absorb light significantly both in the visible and invisible range. Hence, once the chromophore maps have been estimated using visible light, we can infer skin appearance in the UV and IR domain by plugging in our forward model the corresponding absorption coefficients.

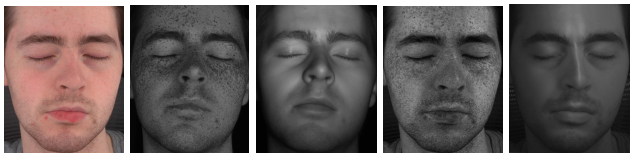
To validate the predicted UVA and NIR appearance, we employ a UV camera (fxo487MXGE) without an IR-cutoff filter to acquire facial images. Subjects are illuminated by five UV LED tubes and an IR LED with emission in the $[300-400]nm$ and $[700-1000]nm$ range, respectively.

Figure 8 (a-c) compares photographs under visible, UVA and NIR lighting, with corresponding albedo reconstructions by our model which assumes uniform illumination (d-e). Consistent with the conclusion that UV photos can display more facial details, Figure 8 (b,d) show some conspicuous spots around the cheek and forehead are much more evident in the UVA domain than in the visible domain.

On the other hand, Figure 8 (c,e) displays the photograph and reconstruction in the NIR domain, where the faces look softer and less influenced by pigmentations, and more visible blood vessels. Note that the UVA and NIR photographs are not cross-polarized and contain specular reflection which is not simulated in our reconstructions. Therefore, in Figure 8 (e) we simulate shading due to small frontal area light and render specular reflections. To account for lateral scattering, we apply an image-space scattering approximation method on the reconstructed albedo. Please see details in Supplemental material.

6. Neural Network

As shown in the previous Section, the use of LUTs to infer model parameters, by searching for the closest match to the color of acquired human skin, can lead to good results. However, even when using the 3-stage approach, parameter estimation is still relatively slow. More importantly, the values contained in the maps are quantized due the discretization of the LUTs. To address these issues, we propose the use of an Encoder-Decoder Neural Network for estimating model parameters and skin reflectance from 300nm to 1000nm, thus covering the UVA + Vis + NIR range.



(a) Vis photo (b) UV photo (c) IR photo (d) UV recon (e) IR recon

Figure 8. Comparisons of Vis, UV and IR photos (a-c) and reconstructions (d, e) using our proposed method.

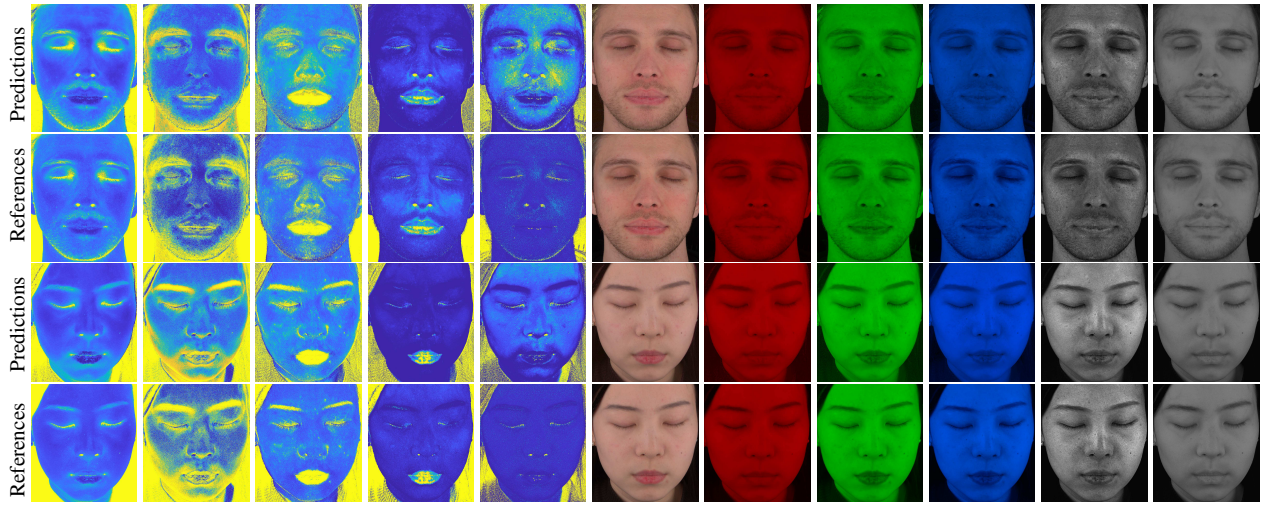
6.1. Network Architecture

Our architecture consists of two networks, as shown in Figure 10. The $ChromNN$ network predicts the five model parameters (C_m, β_m, C_{hd}, C_{he}, and C_{bc}) from the 5D multispectral skin albedo input, that is RGB data from acquired photograph under white, synthetic grayscale blue, and synthetic grayscale green illumination. The estimated parameters are fed to the $SpectraNN$ network which provides a neural approximation of our model to output spectral skin reflectance.

However, directly sampling spectral data densely (e.g. $[300-1000]nm$ in 2nm steps) as output poses issues of larger network and potentially noisy estimates. Therefore, we apply Principal Component Analysis (PCA) to spectral skin reflectance data generated using our model spanning the entire range of parameters, which leads to a training set of 572,220 spectra. We retain the first 10 principal components, shown in the supplemental material, which overall explains $\sim 100\%$ of the total variance. For more details, please refer to the supplemental material.

Training and testing process: When training the $ChromNN$ network, the inputs are values sampled from our 5D LUTs (r, g, b, sb, sg), and the outputs are five model parameters (C_m, β_m, C_{hd}, C_{he}, C_{bc}). These parameters are then fed into the $SpectraNN$ network to predict the reduced 10d PCA space. Utilizing LUT values as a training dataset enables coverage of various skin tones with a lightweight network structure, thereby shortening the training time. To test the $ChromNN$ network, facial images under multiple illuminations are used as inputs to predict five chromophore maps, which are then compared with results from LUT searches. For the $SpectraNN$, inputs can be chromophore values generated either by the $ChromNN$ network or through LUT search approaches, as described in Section 4. The final reconstructions in the predicted 10d PCA basis are compared with photographs under corresponding illuminations.

Implemental details: The $ChromNN$ and $SpectraNN$ are separately optimized using Adam solver with weight decay = 1e-6 and learning rates 3e-4. The activation functions for $ChromNN$ and $SpectraNN$



(a) C_m (b) B_m (c) C_{hd} (d) C_{he} (e) C_{bc} (f) White (g) Red (h) Green (i) Blue (j) UVA (k) NIR

Figure 9. Comparison of the network predictions and references of subject 1 (Mediterranean skin type) and subject 2 (Asian skin type). Figure (a-e) compare the chromophore estimation between the $ChromNN$ outputs and LUT searching methods with the same inputs. Figure (f-i) compare references and visualizations of facial reflectance produced by $SpectraNN$ in the visible domain under white (f), red (g), green (h), and blue (i) illuminations, and in the UVA (j) and NIR (k) domains. Here, the references for the visible domain are photographs, while the references for the UVA and NIR domains are the reconstructions generated by the proposed three-stage LUT searching method.

are Sigmoid and Relu, respectively. All experiments were conducted on Titan X GPUs. The loss function for both networks is given by the Mean Squared Error(MSE).

6.2. Network Results

Figure 9 (a-e) compares the outputs of the $ChromNN$ network and the chromophore maps generated by the proposed three-stage LUT search method of two subjects (Mediterranean and Asia skin types). It is worth noting that the chromophores maps obtained from LUT search are not ground truth, but they can be used as a reference. Generally, the network results have similar quality to the reference maps from LUT search. Chromophore maps from the neural architecture show less noise and less quantization than the ones from LUT search, particularly noticeable in the β_m maps and C_{bc} maps. In order to display clearly, we scaled all chromophore maps, especially C_{bc} maps which should have little difference shown in the range from 0 to 0.01.

After feeding the outputs of $ChromNN$ into the $SpectraNN$, we can obtain the estimated 10-value PCA-based skin reflectance, then expanded into full spectral reflectance. Figure 9 (f-i) compares photographs and corresponding network-generated reconstructions, under white (a), red (b), green (c), and blue (d) visible spectrum illumination produced by the LCD screens. Additionally, in Figure 9 (j,k) we compare network-reconstructed facial appearance in the UVA and NIR domains with the corresponding values provided by LUT search. As the figure shows, the network prediction results are qualitatively very similar

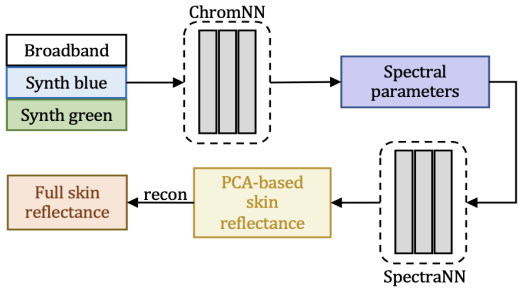


Figure 10. The proposed neural architecture consists of two networks. The $ChromNN$ maps multispectral albedo data to the five spectral model parameters. The model parameters are then fed to the $SpectraNN$ to predict a PCA-based, low-dimensional skin reflectance representation, finally reconstructed using the standard PCA reconstruction technique to obtain the full spectral reflectance in the UVA+VIS+NIR range (300nm to 1000nm).

to LUT search. Note that the comparison is done to LUT search for the UVA and NIR cases in the absence of reference photographs.

7. Quantitative Analysis and Comparison

7.1. Quantitative Analysis

We report a quantitative analysis of our methods and the conventional LUT search method towards efficiency and accuracy. The results we show below are obtained by testing all the datasets including different skin types and genders.
- Efficiency analysis: For $2K$ images, the conventional LUT search method requires 2.88h while the proposed

three-stage LUT search uses 13.91min and the PCA network only needs 2.27s on average.

- Accuracy analysis: Table 1 compares MSE loss for reconstructing spectral reflectance over the Vis. domain. This table validates our conclusion that the quality of the three-stage method is close to conventional LUT search, while the network addresses the discrete value problems of LUTs and achieves much higher accuracy. Our method is not too biased toward certain skin tones.

Table 1. Accuracy analysis

	conventional	three-stage	our network
Caucassian	1.8345e-04	1.8150e-04	3.11e-08
Mediterranean	1.8916e-04	2.0859e-04	3.02e-08
Asian	2.1900e-04	2.2337e-04	3.33e-08
South Asian	2.4326e-04	2.4697e-04	3.96e-08

7.2. Correspondence Comparison

To validate our model, we compare the generated skin reflectance with the accurately measured data in Leeds skin dataset which only has Vis domain [24] (see Figuree 11). Please see the full hyperspectral skin reflectance of our model in Figure 17. The maximum and average RMSE between our model and all the data in the Leeds dataset is 0.036 and 0.008, respectively, which is lower than that of Aliaga et al. [1] (mostly lower than 0.05).

Figure 12 shows UV photograph and reconstructions using multiple inputs (white, synthetic blue, and synthetic green) and a single input (white) in the UV domain. Although reconstructions in visible domain with two different inputs are both close to the photograph, it has an obvious difference in the UVA domain. Figure (b) with multiple inputs is more similar to photograph (a) than Figure (c), which benefits from narrowband illumination measurements.

Compared with Aliaga et al. [1], we apply a diffusion-based skin model with multiple inputs while they employ a Monte Carlo simulation based model with single input. As illustrated above, our model has higher accuracy and better performance in a wider domain and has been validated by practical measurement. Moreover, compared to the network of [1], our PCA-based network has only $\sim 2\%$ parameters. Please see more details in the supplemental material.

We also compare our method with Gitlina et al. [12], on their data, acquired using their Light Stage setup. In Figure 13 we report the photograph of the subject under white illumination (b), its reconstruction using their 4D model (a), and the reconstruction using our 5D model (c), which additionally includes the parameter C_{bc}, as well as fixed C_{we} and C_{wd}. The comparison of the absolute error maps displayed in the right corner shows that our model provides more faithful reconstructions of the input photograph, as well as not being limited to data acquired with our setup.

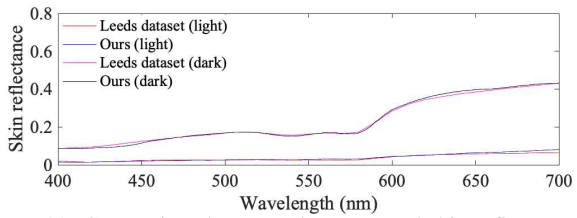
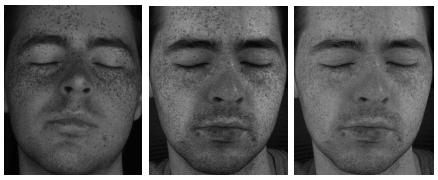


Figure 11. Comparison between the generated skin reflectance of our model and Leeds skin dataset of dark and light skin.



(a) photo (b) recon (wgb) (c) recon (w)

Figure 12. Comparisons of UV photo (a) and reconstruction (b) using broadband white and narrowband blue and green as inputs, and reconstruction using broadband white as input.

(a) 4D Recon (b) Photo (c) 5D Recon

Figure 13. Comparison of a facial photograph captured by Light Stage under white illumination (b), reconstruction using 4D model of [12] (RMSE=0.0311) (a), and reconstruction using our 5D model (c) (RMSE=0.0164), as well as the absolute error maps.

8. Conclusion

We proposed a practical diffusion-based hyperspectral skin reflectance model driven by five chromophore parameters which faithfully represent spectral reflectance from 300 to 1000nm range. To estimate such parameters for faces, we presented a practical measurement setup that employs desktop LCD monitors for spectral illumination. For efficient parameter estimation, we introduced a new three-stage LUT-search method. Moreover, we propose a novel compact neural network to estimate the parameters and reconstruct full hyperspectral reflectance using PCA-bases. Our practical measurements and model allow accurate prediction of skin appearance under a very wide range of spectral illumination conditions, as well as realistic biophysical editing of facial appearance for various applications.

Acknowledgements

This work was partly supported by the BBSRC grant BB/X01312X/1.

References

[1] Carlos Aliaga, Mengqi Xia, Hao Xie, Adrian Jarabo, Gustav Braun, and Christophe Hery. A Hyperspectral Space of Skin Tones for Inverse Rendering of Biophysical Skin Properties. *Computer Graphics Forum*, 2023.

[2] Sarah Alotaibi and William A. P. Smith. A biophysical 3d morphable model of face appearance. In *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, pages 824–832, 2017.

[3] R. Rox Anderson and John A. Parrish. The optics of human skin. *Journal of Investigative Dermatology*, 77(1):13–19, 1981.

[4] Michael Attas, Trevor Posthumus, Bernie Schattka, Michael Sowa, Henry Mantsch, and Shuliang Zhang. Long-wavelength near-infrared spectroscopic imaging for in-vivo skin hydration measurements. *Vibrational Spectroscopy*, 28 (1):37–43, 2002. A Collection of Papers Presented at SHEDDING NEW LIGHT ON DISEASE: Optical Diagnostics for the New Millennium, Winnipeg, Canada, June 25-30, 2000.

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

[6] Laurence Boissieux, Gergo Kiss, Nadia Thalmann, and Prem Kalra. Simulation of skin aging and wrinkles with cosmetics insight. *Computer Animation and Simulation*, 2000.

[7] Tenn F. Chen, Gladimir V. G. Baranoski, Bradley W. Kimmel, and Erik Miranda. Hyperspectral modeling of skin appearance. *ACM Trans. Graph.*, 34(3), 2015.

[8] Symon Cotton, Ela Claridge, and Per Hall. A skin imaging method based on a colour formation model and its application to the diagnosis of pigmented skin lesions. *Proceedings of Medical Image Understanding and Analysis*, 99:49–52, 1999.

[9] Craig Donner and Henrik Wann Jensen. Light diffusion in multi-layered translucent materials. *ACM Trans. Graph.*, 24 (3):1032–1039, 2005.

[10] Craig Donner and Henrik Wann Jensen. A spectral bssrdf for shading human skin. In *Proceedings of the 17th Eurographics Conference on Rendering Techniques*, page 409–417, Goslar, DEU, 2006. Eurographics Association.

[11] Craig Donner, Tim Weyrich, Eugene d'Eon, Ravi Ramamoorthi, and Szymon Rusinkiewicz. A layered, heterogeneous reflectance model for acquiring and rendering human skin. *ACM Trans. Graph.*, 27(5), 2008.

[12] Y. Gitlina, G. C. Guarnera, D. S. Dhillon, J. Hansen, A. Lattas, D. Pai, and A. Ghosh. Practical Measurement and Reconstruction of Spectral Skin Reflectance. *Computer Graphics Forum*, 39(4):75–89, 2020.

[13] Paulo Gotardo, Jérémy Riviere, Derek Bradley, Abhijeet Ghosh, and Thabo Beeler. Practical dynamic facial appearance modeling and acquisition. 2018.

[14] Chet S. Haase and Gary W. Meyer. Modeling pigmented materials for realistic image synthesis. *ACM Trans. Graph.*, 11(4):305–335, 1992.

[15] Jose A. Iglesias-Guitian, Carlos Aliaga, Adrian Jarabo, and Diego Gutierrez. A biophysically-based model of the optical properties of skin aging. *Computer Graphics Forum (EUROGRAPHICS 2015)*, 34(2), 2015.

[16] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*, pages 511–518, 2001.

[17] Jorge Jimenez, Timothy Scully, Nuno Barbosa, Craig Donner, Xenxo Alvarez, Teresa Vieira, Paul Matts, Verónica Orvalho, Diego Gutierrez, and Tim Weyrich. A practical appearance model for dynamic facial color. *ACM Transactions on Graphics*, 29(6):1–10, 2010.

[18] Aravind Krishnaswamy and Gladimir V.G. Baranoski. A biophysically-based spectral model of light interaction with human skin. *Computer Graphics Forum*, 23(3):331–340, 2004.

[19] Alexandros Lattas, Yiming Lin, Jayanth Kannan, Ekin Ozturk, Luca Filipi, Giuseppe Claudio Guarnera, Gaurav Chawla, and Abhijeet Ghosh. Practical and scalable desktop-based high-quality facial capture. pages 522–537. Springer, 2022.

[20] Hubert Nguyen. *Gpu Gems 3*. Addison-Wesley Professional, first edition, 2007.

[21] Jong-Il Park, Moon-Hyun Lee, Michael D. Grossberg, and Shree K. Nayar. Multispectral imaging using multiplexed illumination. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8, 2007.

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

[23] Norimichi Tsumura, Nobutoshi Ojima, Kayoko Sato, Mitsuhiro Shiraishi, Hideto Shimizu, Hirohide Nabeshima, Syuuichi Akazaki, Kimihiko Hori, and Yoichi Miyake. Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. In *ACM SIGGRAPH 2003 Papers*, page 770–779, New York, NY, USA, 2003. Association for Computing Machinery.

[24] Kaida Xiao, Julian Yates, Faraedon Zardawi, Suchitra Sueeprasan, N. Liao, Liz Gill, Chelsea Li, and Sophie Wuerger. Characterising the variations in ethnic skin colours: A new calibrated data base for human skin. *Skin Research and Technology*, 23, 2016.

[25] David Zargaran, Florence Zoller, Alexander Zargaran, Tim Weyrich, and Afshin Mosahebi. Facial skin aging: key concepts and overview of processes. *International Journal of Cosmetic Science*, 2022.

Practical Measurement and Neural Encoding of Hyperspectral Skin Reflectance

Supplementary Material

In this document we provide additional information on several aspect of our work. Part A describes the detailed information of our model and the process of generating Lookup table (LUT) data in Section 3 while Part B shows additional experiments for narrowband measurement. Part C introduces the image-space scattering approximation method used in Section 5. Part D provides additional information about the proposed neural network (described in Section 6 of the main paper), and includes a comparison with [1]. Additionally, we also provide a video including comparison among photograph in the visible domain, and reconstructions in the UVA and NIR domains, as well as more realistic skin appearance editing results.

A. Bio-physical Skin Model

In this section, we provide details on our model and how the LUT data is generated.

- Diffusion BSSRDF: Similar to the work [10, 12], we used a two-layer skin model, which consists in the epidermal and dermal layers, based on diffusion theory. For the epidermis, we employ a multipole diffusion model to simulate subsurface scattering, with thickness fixed to 0.32mm. As for the dermis, simulated as a semi-infinite medium, we employ a dipole diffusion model. The absorption coefficients for epidermis and dermis are shown in Equation 1 and Equation 2. The absorption coefficients of two types of melanin (eumelanin and pheomelanin) are

$$\sigma_a^{em}(\lambda) = 6.6 \times 10^{10} \times \lambda^{-3.33}\ \mathrm{mm}^{-1}, \tag{3}$$

$$\sigma_a^{pm}(\lambda) = 2.9 \times 10^{14} \times \lambda^{-4.75}\ \mathrm{mm}^{-1} \tag{4}$$

The absorption of the baseline skin tissue is

$$\sigma_a^{base}(\lambda) = 0.0244 + 8.53 \times e^{-(\lambda - 154)./66.2}\ \mathrm{mm}^{-1} \tag{5}$$

where λ refers to the light wavelength in nanometers.

The reduced scattering coefficient of the epidermis is

$$\sigma_s'^{epi}(\lambda) = 14.74 \times \lambda^{-0.22} + 2.2 \times 10^{11} \times \lambda^{-4}\ \mathrm{mm}^{-1} \tag{6}$$

The scattering coefficient of the dermis is approximated as half of epidermal scattering [10] [5].

The index of refraction (IOR) of both layers is 1.4. We convolved the spectral profiles (reflectance and transmittance) with the Kubelka-Munk theory [14] to obtain the total diffuse reflectance, then generated a LUT with SPDs of different illuminations.

- Chromophore sampling: For each point in the LUT, we sampled five chromophore parameters in different ranges and methods as shown in Table 2, parameter range and sampling density vary on a per-chromophore basis. C_m, C_{hd} and C_{bc} are sampled cubicly in their respective ranges, according to the following equation:

$$c_i = C_{max} \times (\frac{i}{N})^3 \tag{7}$$

where c_i is the value of the i^{th} sample for chromophore C, $i \in \{0, 1, \ldots, N\}$, $N+1$ is the number of samples and C_{max} is the maximum value C can assume in our model. C_m , C_{hd} and C_{bc} have 51, 51, 5 samples, respectively. β_m and C_{he} are uniformly sampled, respectively with 11 and 4 samples.

Parameters	Sampling	Spectral Range
C_m	0 - 0.5 (cubicly)	UV-Vis-IR
β_m	0 - 1 (uniformly)	UV-Vis-IR
C_{hd}	0 - 0.5 (cubicly)	UV-Vis
C_{he}	0 - 0.3 (uniformly)	UV-Vis
C_{bc}	0 - 0.01 (cubicly)	UV-Vis
C_{we}, C_{wd}	0.2, 0.6	Vis-IR

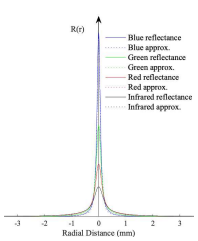
Table 2. Parameters of skin spectral absorption and scattering.

B. Narrowband measurements

As observed in previous work [12], narrowband illumination allows to reduce the baked-in subsurface scattering in the estimated chromophore maps and hence in the albedo, thus leading to sharper details than the corresponding estimates from broadband-only measurements. The increased sharpness is beneficial when rendering skin appearance, as the blurring introduced by subsurface scattering simulation allows a closer match to the ground truth appearance. On the other hand, estimates obtained under broadband illumination will appeared over-blurred when rendered. However, as discussed in the main paper, to address camera color gamut issues we avoid direct observation of skin response under narrowband illumation and use a computational scheme to recover the response from the mixture narrowband + broadband. Figure 14 compares the reconstruction of cheek area using broadband white and synthetic blue and green illuminations (a) with the ground truth photograph (b). In the same figure, we report the reconstruction using broadband white and direct blue and green illuminations. As the figure shows, the reconstruction with synthetic data shows sharper skin details than that reconstructed from measurements direct illumination, particularly visible around skin blemishes (e.g. moles) and pores.

(a) Recon(synth.) (b) Photo (c) Recon(direct)

Figure 14. Comparison of photograph and reconstructions with different measurements. Figure (a) employs broadband white, synthetic blue, and synthetic green illuminations while Figure (c) employs broadband white, direct blue, and direct green illuminations.



(a)

Variance(mm)	Blue	Green	Red	NIR
0.0013	0.0738	0.0106	0.0013	0
0.0116	0.4972	0.2360	0.0715	0.0237
0.0338	0	0	0.0684	0.0746
0.1315	0.4289	0.4158	0.1675	0.1398
0.519	0	0.1777	0.2944	0.3285
1.012	0	0.1598	0	0.0057
2.012	0	0	0.3969	0.4277

(b)

Figure 15. Approximating profiles (a) with sums of seven Gaussians (b).

(a) White (b) Red (c) Green (d) Blue (e) NIR

Figure 16. Comparison of image-space scattering approximation and photos (or channels) under white (a), red(b), green (c), blue (d), IR (e) illumination.

C. Image-space Scattering Approximation

In our work, both the proposed LUT search or the neural network aim at producing sharp reconstructions, thus reducing the effect of lateral scattering of light. As observed in the previous section, this would affect the realism of skin appearance, if the reconstructed albedoes were directly compared with ground truth photographs. To address this, we apply an image-space scattering approximation method on the reconstructed albedo which produces images that better match the validation photographs.

We first select the closest skin reflectance curves in our lookup table to those of the light and dark skin patches in a standard colorcheker chart. For each patch, we select the reflectance values at 450nm, 530nm, 630nm and 835nm, corresponding to the peaks of blue, green, red, and NIR illumination in our model, respectively. Similar to the work of [20], we apply a linear combination of seven Gaussians to fit the curves of each patch. Note that these Gaussians share the same variances, and each profile weights them differently shown in Figure 15. We then blur the reconstructed albedos with each Gaussian and obtain the final result via a linear combination of multiple blur kernels. We do not apply any blur to the UVA reconstruction since this has been predicted from measurements under longer wavelengths (i.e. visible light), thus already encoding some blur.

Figure 16 compares the results and photos under white (a), red (b), green (c), blue (d), and IR (e) illuminations. As for the broadband white image, we process RGB channels separately, since the white illumination of the monitor setup is simply a sum of three separate RGB peaks. Finally, for the IR LED panel, we simulate shading due to small frontal area light and also render specular reflection. We brighten these images to clearly show the details. As the figure shows, adding the image-space blur makes the results qualitatively closer to the photographs.

D. Neural Networks

In this part, we offer detailed information of the PCA-based network, ablation study of PCA, and comparisons with the recent work of Aliaga et al. [1]

We apply Principal Component Analysis (PCA) on the full spectral skin reflectance data spanning the UVA, Vis, and NIR domains to reduce the complexity of the proposed neural network (see Section 6). The retained first 10 principal components, which overall explain $\sim 100\%$ of the total variance, are shown in Figure 18. Using such components, any given skin reflectance spectrum can be reconstructed using a 10d vector containing one weight for each of the 10 principal components, along with the mean spectral reflectance of the dataset.

- Ablation study: Figure 17 compares a ground truth skin reflectance spectra with its PCA reconstruction using this 10d PCA basis. As can be seen, the reconstruction almost perfectly matches the spectral reflectance, sampled at 2nm resolution.

Compared with the recent work of Aliaga et al. [1], which directly estimates reflectance without PCA, our method offers three advantages:

- Our model and PCA-based neural network covers a wider spectral domain spanning from UVA, visible and NIR domains, while Aliaga et al. [1] only predicts reflectance in

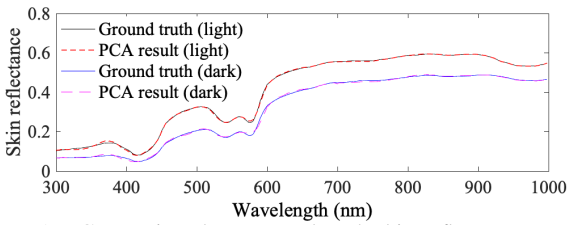


Figure 17. Comparison be a ground truth skin reflectance spectra with its PCA reconstruction using the first 10 PCA components.

the visible and NIR domain.

- The predicted skin reflectance using our model is more accurate. Figure 11 compares the generated spectral skin reflectance with Leeds skin dataset [24] of two examples representing light and dark skin. As can be seen, the generated skin reflectance of our model is close to that measured accurately in the Leeds dataset. As mentioned in Sec 7.2, the maximum and average RMSE of our model is 0.036 and 0.008, respectively while the RMSE between Aliaga et al. is generally lower than 0.05 [1]. The accuracy of the model will decide the performance of the network since the training dataset comes from the LUTs generated by our model.
- Although our method can cover a wider spectral range and at higher resolution, the size of our network is smaller, with significantly fewer parameters (9354 parameters for our network, thus about 98% fewer parameters than the 424758 for the network in [1]).

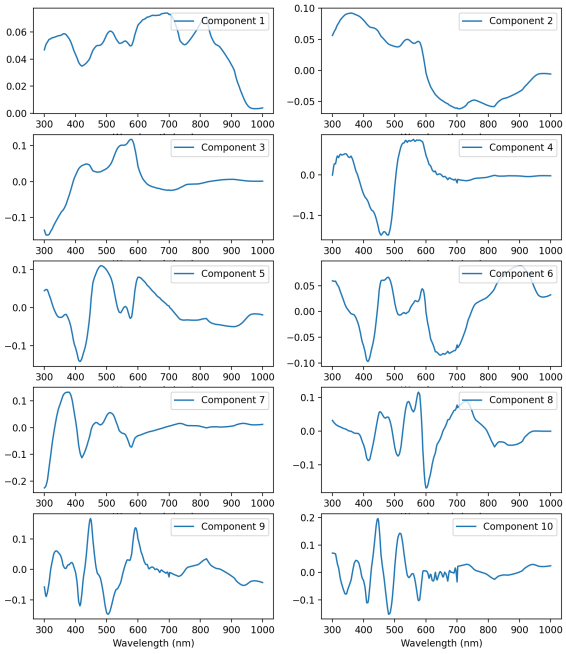


Figure 18. The first ten PCA components derived from our analysis.