Skin Chromophore Estimation from Mobile Selfie Images using Constrained Independent Component Analysis

Luisa F. Polanía, Raja Bala, Ankur Purwar, Paul Matts, Martin Maltz

Abstract

Human skin is made up of two primary chromophores: melanin, the pigment in the epidermis giving skin its color; and hemoglobin, the pigment in the red blood cells of the vascular network within the dermis. The relative concentrations of these chromophores provide a vital indicator for skin health and appearance. We present a technique to automatically estimate chromophore maps from RGB images of human faces captured with mobile devices such as smartphones. The ultimate goal is to provide a diagnostic aid for individuals to monitor and improve the quality of their facial skin. A previous method approaches the problem as one of blind source separation, and applies Independent Component Analysis (ICA) in camera RGB space to estimate the chromophores. We extend this technique in two important ways. First we observe that models for light transport in skin call for source separation to be performed in log spectral reflectance coordinates rather than in RGB. Thus we transform camera RGB to a spectral reflectance space prior to applying ICA. This process involves the use of a linear camera model and Principal Component Analysis to represent skin spectral reflectance as a low-dimensional manifold. The camera model requires knowledge of the incident illuminant, which we obtain via a novel technique that uses the human lip as a calibration object. Second, we address an inherent limitation with ICA that the ordering of the separated signals is random and ambiguous. We incorporate a domain-specific prior model for human chromophore spectra as a constraint in solving ICA. Results on a dataset of mobile camera images show high quality and unambiguous recovery of chromophores.

Introduction

Human skin is a turbid multi-layered structure. The color and texture of skin is defined primarily by two biological chromophores found in these layers: melanin which is a pigment in the epidermis, and hemoglobin which is found within the vascular structure at the papillary dermis, a sub-layer of skin, in oxygenated and deoxygenated forms and is responsible for red colorations of skin tone. Chromophore concentrations affect skin tone and serve as markers for skin health and disease. Abnormalities in melanin levels result in conditions such as melasma, vitiligo, and sun spots. Skin conditions such as acne, rosacea, and telangiectasia can cause organic changes in vascular structure and elevate the level of hemoglobin present in the dermis, causing a red coloration and uneven skin tone. The analysis and quantification of chromophore levels are thus essential steps for monitoring skin appearance and health, and for recommending treatments.

The most accurate means for measuring chromophores is via direct in-vivo methods using spectrometry and spectrophotometry [15, 13] that involve clinical expertise and special equipment, and are thus of high cost. Alternatively, non-contact methods employ digital cameras for capturing images of skin, and image analysis algorithms to recover spatial melanin and hemoglobin maps. Demirli et al. [7] propose a transform from camera RGB to a proprietary *RBX* color space, where the R and B channels represent respectively hemoglobin and melanin maps. A fundamental assumption in their model is that the RGB images are captured under tightly controlled conditions with a high quality digital camera employing a polarized flash and cross-polarized filters to eliminate specular reflection. Other methods estimate chromophores on the basis of an optical model that describes the reflection, absorption, and scattering of incident light within various layers of skin, as shown in Figure 1. Various researchers have proposed use of the Kubelka-Munk theory [18, 4, 5] or Beer's Law [11] to model the light-skin interaction and estimate melanin and hemoglobin quantities. Direct chromophore estimation from these first-principles models requires precise knowledge of scattering and absorption coefficients and layer thicknesses for melanin and hemoglobin, which is not always readily accessible.
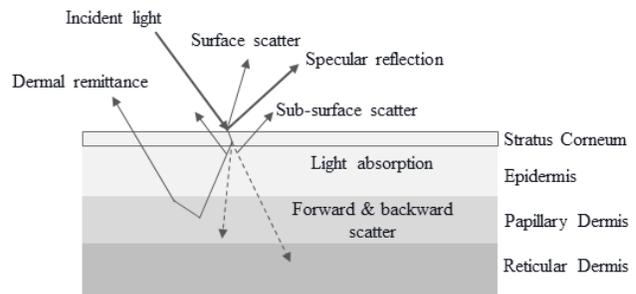

Figure 1: Light-skin interaction model

To address this issue, several researchers have explored statistical techniques such as Independent Component Analysis (ICA) for chromophore recovery [19, 16, 20]. Tsumura et al [19] use the Lambert-Beer light transport model as a foundation and employ ICA to disentangle melanin and hemoglobin channels from the diffuse component of the RGB image. The assumption is that the two chromophore channels are statistically independent non-Gaussian random variables that combine additively to form the diffuse component of skin color. There are two limitations with their approach. First, the Lambert-Beer law posits that the chromophores combine additively in log spectral reflectance space. Tsumura et al. instead perform ICA in log RGB space under a simplifying assumption that the camera's spectral sensitivities can be approximated by Dirac delta functions - an assumption that is far from true in practice. Secondly, the independent components produced by standard ICA do not conform to a unique ordering, i.e. one cannot deterministically identify which of the two components corresponds to hemoglobin versus melanin.

In this work, our goal is to estimate skin chromophore maps from RGB images captured with smartphone cameras "in the wild". Specifically, we extend the work of [19] to overcome the two aforementioned limitations. We perform ICA directly in log spectral reflectance space. This requires a preprocessing step to map camera RGB to spectral reflectance. We employ a linear camera model that relates spectra to RGB. This model requires knowledge of the camera filters and illuminant. The former is obtained from available databases, while the latter is estimated by a novel approach that uses the lip region as a calibration object with known reflectance to recover the lighting vector. Furthermore, we exploit the fact that skin RGB values lie in a low-dimensional manifold in spectral reflectance space [17], and use Principal Component Analysis (PCA) to recover skin spectral reflectance from RGB. Finally, we modify the original ICA algorithm to disambiguate the two chromophore channels by imposing domain-specific constraints on the expected spectral absorption of melanin and hemoglobin.

The paper is organized as follows. We begin by describing a technique to recover spectral reflectances of human skin from camera RGB values, which includes a novel illuminant estimation step. This is followed by a description of the Lambert-Beer light-skin interaction model that relates spectral reflectance to chromophore maps. Subsequently, the ICA algorithm is presented along with our novel use of domain priors for unambiguous chromophore separation. This is followed by experimental results and concluding remarks.

Skin Spectral Reflectance Estimation

We adopt a simple linear model for light captured by a digital RGB camera:

$$\vec{c} = \mathbf{QE}\vec{r} \qquad (1)$$

where $\vec{c} \in \mathbb{R}^3$ is an RGB pixel color recorded by the camera; $\mathbf{Q} \in \mathbb{R}^{3 \times K}$ is the set of R, G, B camera spectral sensitivities sampled over K wavelengths in the visible spectrum between 400-700nm; $\mathbf{E}$ is a $K \times K$ diagonal matrix whose diagonal elements form the incident lighting vector, and $\vec{r} \in \mathbb{R}^K$ is the spectral reflectance of skin. Estimation of illuminant $\mathbf{E}$ is discussed in the next subsection. For $\mathbf{Q}$, we adopt the camera sensitivities reported in [14], and shown in Figure 2.
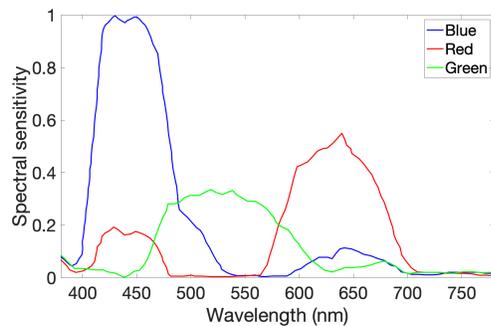

Figure 2: Spectral sensitivities of smartphone camera.

The estimation of high-dimensional spectral reflectance from 3-dimensional RGB is an ill-posed problem. To regularize the task, we restrict the space of skin reflectances to a 3-dimensional manifold. This is achieved by approximating skin reflectance as a linear combination of 3 basis functions obtained by Principal Component Analysis:

$$\vec{r} = \mathbf{B}\vec{w} + \vec{r}_m \qquad (2)$$

where $\mathbf{B}$ is a $K \times 3$ matrix whose column vectors are the bases corresponding to the first 3 principal components of skin reflectances, $\vec{w} \in \mathbb{R}^3$ are the basis weights, and $\vec{r}_m$ is the mean facial skin spectral reflectance vector. We use the 3 primary basis vectors published by Sun and Fairchild [17], shown in Figure 3. The mean skin reflectance vector is also obtained from the same study. Plugging Eqn (2) into (1) we arrive at an expression for $\vec{w}$ in terms of camera RGB pixel values:

$$\vec{w} = (\mathbf{QEB})^{-1}(\vec{c} - \vec{c}_m) \qquad (3)$$

where $\vec{c}_m = \mathbf{QE}\vec{r}_m$ is the mean skin RGB color recorded by the camera. With $\vec{w}$ known, it is trivial to compute spectral reflectance $\vec{r}$ using Eqn (2).
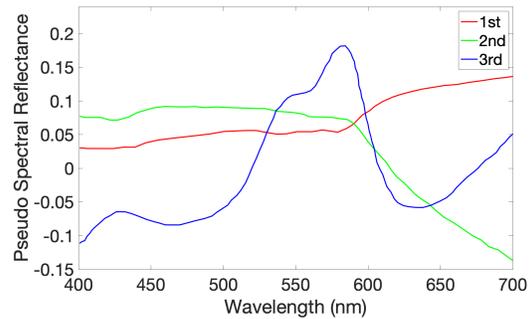

Figure 3: First three principal components of facial skin reflectance adopted from Sun and Fairchild [17].

Illuminant Estimation

A key parameter required for the aforementioned spectral reflectance estimation procedure is $\mathbf{E}$, the spectral power distribution of the incident illuminant. Accurate illuminant estimation under uncontrolled capture conditions is an open research problem [6, 8]. One way to enhance robustness and accuracy of illuminant estimation is to introduce a calibration object into the scene. Bhatti et al [2] propose a scheme wherein subjects hold a calibration chart during image capture. In preliminary experiments we found such an approach to be cumbersome for the subjects. Choi et al [3] propose using parts of the the human eye, namely the sclera and pupil, as calibration objects with reflectances that are relatively stable across individuals.

Our approach is along a similar vein to [3], but adapted to the chromophore separation task. Namely, we observe that the lip region for all humans comprises high concentrations of hemoglobin. Hence we hypothesize that on average, lip reflectance correlates closely with the canonical hemoglobin absorption profile shown in Figure 4. We process camera RGB values from the lip region through the aforementioned spectral reflectance estimation procedure using a candidate illuminant selected from a standard illuminant database [17] to produce a spectral reflectance estimate. The mean square error (MSE) is com-

puted between this estimate and the canonical hemoglobin reflectance vector. This process is repeated for all candidate illuminants in the database, and the one that minimizes MSE is selected as the illuminant for that image.

Light Skin Interaction Model

We adopt the modified Lambert-Beer Law [11] for light transport through skin, which relates diffuse spectral reflectance to the underlying skin chromophores. In the following formulation, spatially varying quantities associated with an image are suffixed with xy, while spectral quantities are denoted as vectors.

$$\vec{r}_{xy} = s_{xy} e^{-\vec{\sigma}_m \vec{l}_m m_{xy} - \vec{\sigma}_h \vec{l}_h h_{xy}} \qquad (4)$$

Here $\vec{r}_{xy}$ is the diffuse spatial reflectance map; s_{xy} is the shading map that represents attenuation of a distant global illumination source; $\vec{\sigma}_m$ and $\vec{\sigma}_h$ are the melanin and hemoglobin spectral density functions respectively, $\vec{l}_m$, $\vec{l}_h$ are the mean spectral path lengths of photons in the epidemis and dermis respectively, and m_{xy} and h_{xy} are the spatial maps corresponding respectively to melanin and hemoglobin concentrations. Taking logs of Eqn (4), we arrive at an additive model:

$$log(\vec{r}_{xy}) = log(s_{xy}) - \vec{\sigma}_m \vec{l}_m m_{xy} - \vec{\sigma}_h \vec{l}_h h_{xy} \qquad (5)$$

The goal is to recover chromophore maps m_{xy} and h_{xy} given the reflectance image $\vec{r}_{xy}$. If all spectral quantities are known, the chromophore and shading maps may be obtained via straightforward linear regression. However in practice one does not have easy access to accurate measurements of chromophore absorption and path lengths for a given subject. Hence we turn to ICA, described next, as a means to simultaneously estimate the chromophore maps and spectral mixing weights.

Independent Component Analysis

ICA assumes that a set of latent source signals mix additively to produce an observed signal. The sources are assumed to be statistically independent and non-Gaussian. ICA attempts to simultaneously estimate the sources and mixing weights by solving an optimization problem that maximizes the statistical independence and non-Gaussianity of the sources [10]. Assume the observed input is an image of P pixels, each pixel being associated with a K-dimensional log reflectance spectrum. We denote this image as a $K \times P$ matrix $\mathbf{R}_{xy}$. Casting the light transport model (5) into the ICA framework yields the following expression for a linear generative model:

$$\mathbf{R}_{xy} = \mathbf{A} \mathbf{V}_{xy} \qquad (6)$$

where $\mathbf{A}$ is a $K \times 3$ mixing matrix, and $\mathbf{V}_{xy}$ is a $3 \times P$ matrix whose rows are the 3 source signals, namely the hemoglobin, melanin and illuminant shading maps. Under the model given by (5), the column of $\mathbf{A}$ corresponding to the shading map is set as a vector of all 1's. The inverse model that describes the chromophores in terms of the observed signal is given by:

$$\mathbf{V}_{xy} = \mathbf{H} \mathbf{R}_{xy} \qquad (7)$$

where in ICA parlance, $\mathbf{H}$ is the unmixing matrix. The ICA optimization solves for the mixing matrix $\mathbf{A}$ from which we compute its pseudoinverse to obtain $\mathbf{H}$, and then readily obtain chromophores and shading $\mathbf{V}_{xy}$ using (7). We adopt the fixed point

ICA optimization algorithm originally proposed by Hyvarinin et al. [10] and applied for skin chromophore separation by Tsumura et al [19]. Specifically, let $\vec{a}_i$ be the ith column vector of mixing matrix $\mathbf{A}$. The ICA algorithm uses a fixed-point iterative scheme to solve:

$$\vec{a}_i = \underset{\vec{a}'_i}{\arg\min}(-|kurtosis(\vec{a}'^{T}_i \vec{x})|) \qquad (8)$$

where $\vec{x}$ is a whitened version of observed data $\vec{r}$ and kurtosis is computed over a set of training samples. (We drop the xy subscript for brevity.) The reader is referred to [19] for details.

Constrained ICA

One practical challenge with the aforementioned implementation is that there is no unique ordering of the resulting independent components; hence there is no deterministic association between the latter and the chromophore signals. To address this issue, we add a constraint to Eqn. (8) as follows:

$$\vec{a}_i = \underset{\vec{a}'_i}{\arg\min}(-|kurtosis(\vec{a}'^{T}_i \vec{x})| + \lambda||\vec{a}'_i - \vec{p}_i||^2_2) \qquad (9)$$

where the second term encourages the ith column, $\vec{a}_i, i = 1, 2$ of the mixing matrix to be similar to domain-specific priors $\vec{p}_i$. These are defined respectively as melanin and hemoglobin spectral absorption curves, as shown in Figure 4. λ is a hyperparameter that controls the tradeoff between independence of the source signals and similarity with the priors. To select this parameter, we adopt the L-curve method [9]: namely, a log-log plot is generated of the norms of the two terms in (9) for the optimal solution obtained with varying λ, and the value that corresponds to the point of maximum curvature is selected. Recall that the third column of $\mathbf{A}$ activates the shading map and is set to $\vec{a}_i \equiv \vec{1}$.
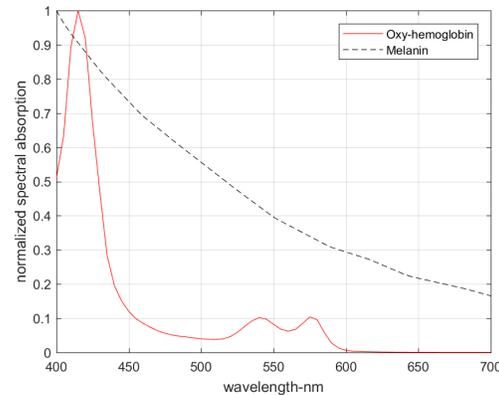

Figure 4: Normalized spectral absorption of oxy-hemoglobin and melanin.

In practice, the recovered columns of the mixing matrix $\mathbf{A}$ may not confirm closely to ideal melanin and hemoglobin absorption spectra due to the fact that there are many simplifying assumptions made with the camera and light-skin models, including knowledge of the illuminant, which may not hold strictly for smartphone images captured in the wild. As a practical alternative, we employ a data-driven approach to define the priors. In

a first training phase, we apply unconstrained ICA on a small batch of training images, and use visual inspection to select a few high-quality examples wherein the estimated independent components effectively highlight melanin-induced features (e.g. sun spots) or hemoglobin-induced features (e.g. acne) in the facial images. We then define the priors $\vec{p}_i$ as the average of the column vectors $\vec{a}_i$ corresponding respectively to high-quality melanin and hemoglobin maps. In the final (inference) phase, we solve the constrained ICA problem in (9) on test images employing the priors from the training phase.

Experiments

Images of 180 female Asian subjects were captured using iPhone5 and iPhone6 devices. In all cases, the rear camera on the smartphone was used. Subjects were chosen to span a variety of ages and skin conditions. There were 32 repeat shots per subject, captured under a variety of indoor and outdoor conditions.
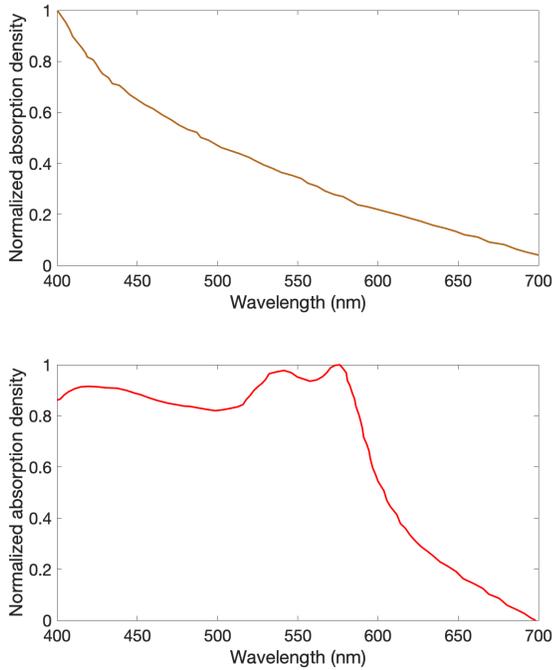

Figure 5: Average prior vectors for melanin (top) and hemoglobin (bottom) obtained from training images yielding high quality source separation.

Figure 5 shows the two prior vectors of the mixing matrix obtained from unconstrained ICA applied on a preliminary set of 30 training samples. We note that these priors approximately exhibit the spectral absorption characteristics of melanin and hemoglobin. This validates the hypothesis that the Lambert-Beer law holds to a first order, and the chromophores do combine additively and independently in log spectral density space. These priors are then incorporated into the constrained ICA optimization (9) with $\lambda = 10$ on subsequent test images to achieve unambiguous ordering of chromophores.

Figure 6 shows examples of RGB images and the estimated source signals corresponding to the first and second columns of **A**.

We note that appropriate melanin- and hemoglobin-induced features are highlighted in the first and second source signals respectively, and that the first and second source consistently correspond to melanin and hemoglobin, thus eliminating the ordering ambiguity inherent in standard ICA. The images in the third row exhibit white spots, which indicate a failure of our model to recover chromophore concentrations in these regions. Potential causes include the presence of specular highlights, incorrect estimation of ambient illumination or camera filters.
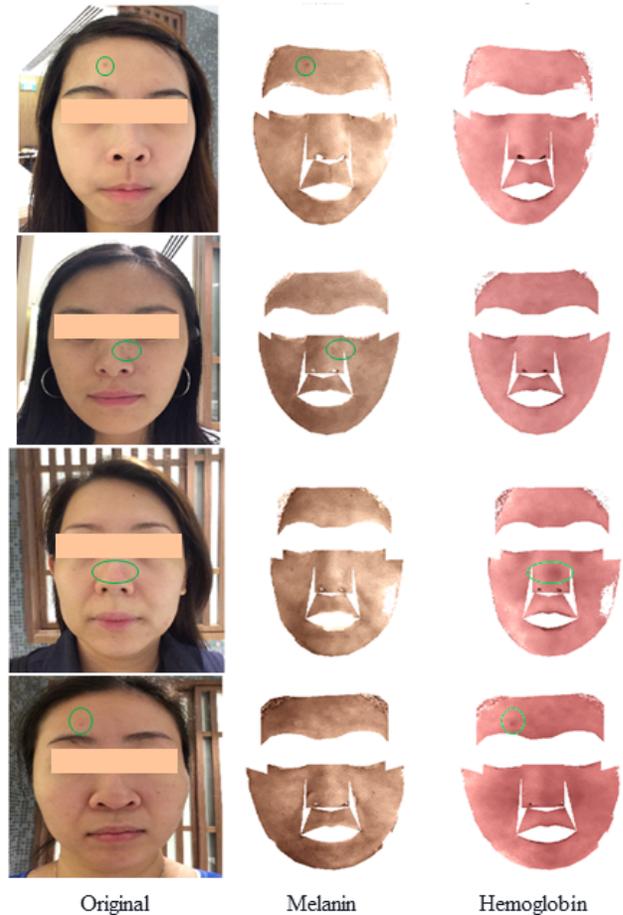

Figure 6: Selected images and their estimated chromophore maps. Brown and red hues are assigned respectively to melanin and hemoglobin grayscale maps for visualization. Green circles denote features of interest identified by clinical experts. The first two rows show features induced by unevenness in melanin pigmentation, while the last two rows show features arising from uneven hemoglobin concentration. (Best viewed zooming in on color display.)

Conclusions

We present a method to estimate facial melanin and hemoglobin maps from smartphone RGB images. The RGB signals are mapped to spectral signals, and source separation is performed in log spectral space in a manner that is consistent with the Lambert-Beer light transport model. The ill-posed mapping from camera RGB to spectra is constrained by operating in a 3D

skin subspace defined by PCA. Camera filters are retrieved from available databases, and the ambient illuminant is estimated by a novel technique that treats the human lip as a single-chromophore reference object. Finally the inherent ambiguity in the order of the source signals produced by ICA is resolved by incorporating domain-specific priors into the optimization. Results show that the chromophore maps estimated by our technique are uniquely ordered, and effectively highlight relevant skin features identified by clinicians.

Future work includes detailed comparisons between ICA performed in RGB vs. spectral space. Improved techniques to remove specular highlights from face images, such as the work of [12] would be beneficial. Robust methods for estimating camera filters and scene illumination across widely varying capture scenarios are needed. While much of the current work is derived on narrow subject demographics, it is imperative that in order to make meaningful advances, model development and experimentation must consider diverse ethnicity, notably dark-skinned subject populations. Finally, a fruitful computational tool that can be brought to bear to facial skin analysis is machine learning, specifically deep learning. The recent work by Alotaibi and Smith [1] is a promising step in this direction.

References

[1] S. Alotaibi and W. Smith. BioFaceNet: Deep biophysical face image interpretation. In *Proceedings of the 30th British Machine Vision Conference*, 2019.

[2] N. Bhatti, H. Baker, H. Chao, S. Clearwater, M. Harville, J. Jain, N. Lyons, J. Marguier, J. Schettino, and S. Süsstrunk. Mobile cosmetics advisor: An imaging based mobile service. In *Multimedia on Mobile Devices*, volume 7542, pages 39 – 49, 2010.

[3] H. Choi, K. Choi, and H. Suk. The human sclera and pupil as the calibration targets. *Electronic Imaging*, 2017(17):200–203, 2017.

[4] S. Cotton and E. Claridge. Developing a predictive model of human skin coloring. In *Medical Imaging 1996: Physics of Medical Imaging*, volume 2708, pages 814 – 825, 1996.

[5] S. Cotton, E. Claridge, and P. Hall. A skin imaging method based on a colour formation model and its application to the diagnosis of pigmented skin lesions. In *Proceedings of Medical Image Understanding and Analysis*, volume 99, pages 49–52, 1999.

[6] P. Das, A. Baslamisli, Y. Liu, S. Karaoglu, and T. Gevers. Color constancy by GANs: An experimental survey. *arXiv preprint arXiv:1812.03085*, 2018.

[7] R. Demirli, P. Otto, R. Viswanathan, S. Patwardhan, and J. Larkey. Rbx® technology overview. *Canfield Systems White Paper*, 2007.

[8] A. Gijsenij, T. Gevers, and J. Van De Weijer. Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, 2011.

[9] P. C. Hansen. The l-curve and its use in the numerical treatment of inverse problems. In *Computational Inverse Problems in Electrocardiology*, pages 119–142. 1999.

[10] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, 1999.

[11] T. Igarashi, K. Nishino, S. Nayar, et al. The appearance of human skin: A survey. *Foundations and Trends in Computer Graphics and Vision*, 3(1):1–95, 2007.

[12] H. Kim, H. Jin, S. Hadap, and I. Kweon. Specular reflection separation using dark channel prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1460–1467, 2013.

[13] N. Kollias and A. Baqer. Spectroscopic characteristics of human melanin in vivo. *Journal of investigative dermatology*, 85(1):38–42, 1985.

[14] C. Matasaru. Mobile phone camera possibilities for spectral imaging. Master's thesis, University of Eastern Finland, School of Computing Department, 2014.

[15] P. Matts, P. Dykes, and R. Marks. The distribution of melanin in skin determined in vivo. *British Journal of Dermatology*, 156(4):620–628, 2007.

[16] N. Ojima, N. Tsumura, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, and Y. Miyake. Measurements of skin chromophores by independent component analysis and the application to cosmetics. In *PICS*, pages 571–574, 2003.

[17] Q. Sun and M. Fairchild. Statistical characterization of face spectral reflectances and its application to human portraiture spectral estimation. *Journal of Imaging Science and Technology*, 46(6):498–506, 2002.

[18] S. Tominaga and B. Wandell. Standard surface-reflectance model and illuminant estimation. *JOSA A*, 6(4):576–584, 1989.

[19] N. Tsumura, H. Haneishi, and Y. Miyake. Independent-component analysis of skin color image. *JOSA A*, 16(9):2169–2176, 1999.

[20] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, and Y. Miyake. Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. In *ACM SIGGRAPH*, pages 770–779. 2003.

Author Biography

Luisa Polanía is a Lead AI Scientist at Target Corporation. Her research interests are signal and image processing, computer vision, and machine learning. She received a PhD in Electrical and Computer Engineering from the University of Delaware.

Raja Bala is a Principal Scientist at the Palo Alto Research Center. His research interests are at the intersection of digital imaging, computer vision, and machine learning. He is a Fellow of IS&T and Senior Member of IEEE. He received a PhD in Electrical Engineering from Purdue University.

Paul Matts is Victor Mill Research Fellow at the Procter Gamble Company. He is also Visiting Professor to the University of the Arts, London and a Fellow of the Royal Society of Chemistry and the Royal Society of the Arts. He received a PhD in Biochemistry from the University of Cardiff, Wales, UK.

Ankur Purwar is a Principal Scientist at the Procter Gamble RD Singapore. His research interests are skin appearance understanding leveraging computer vision and machine learning techniques. He received a PhD in Applied Mathematics from Indian Institute of Technology Kanpur, India.

Martin Maltz has been with Xerox Corporation since 1968. While his primary contributions are in color management, he has worked in many areas including display technologies, Xerographic modeling, security printing, forms analysis, image enhancement, license plate recognition, gait analysis, and skin health analysis. He received a PhD in Electrical Engineering from Massachu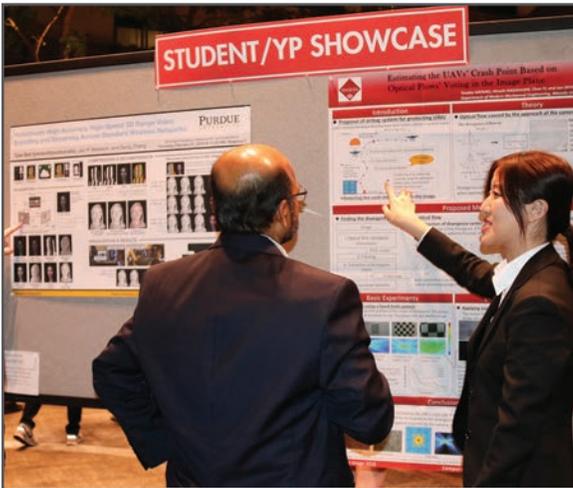setts Institute of Technology.