Multispectral Skin Color Modeling*

Technical Report MS-CIS-01-22

Elli Angelopoulou
Stevens Institute of Technology
elli@cs.stevens-tech.edu

Rana Molana
University of Pennsylvania
molana@grasp.cis.upenn.edu

Kostas Daniilidis
University of Pennsylvania
kostas@grasp.cis.upenn.edu

June 22, 2001

Abstract

The automated detection of humans in computer vision as well as the realistic rendering of people in computer graphics necessitates improved modeling of the human skin color. In this paper we describe the acquisition and modeling of skin reflectance data densely sampled over the entire visible spectrum. The data collected through a spectrograph allows us to explain skin color (and its variations) and to discriminate between human skin and dyes designed to mimic human skin color. We study the approximation of these data using several sets of basis functions. Our study shows that skin reflectance data can best be approximated by a linear combination of Gaussians or their first derivatives. This result has a significant practical impact on optical acquisition devices: the entire visible spectrum of skin reflectance can now be captured with a few filters of optimally chosen central wavelengths and bandwidth.

1. Introduction

An accurate description of the color of human skin is key for both human detection/identification in computer vision and for realistic rendering in computer graphics. Ideally, a good skin color descriptor should be physically accurate, numerically advantageous and flexible in order to allow for skin tone variations.

The majority of existing skin color data is in the form of Red-Green-Blue (RGB) triplets. The prevalence of color descriptors as a combination of these three color primaries can be partly attributed to the influence of the tristimulus human visual perception mechanism. For most cases, describing color in the tristimulus space is sufficient for communicating color information to a human observer. However, if the whole computation of light reflection and transfer is performed in tristimulus space (see section 2), significant color distortions are introduced [6, 7]. Furthermore, the use of tristimulus representations introduces metameric colors, i.e. colors with the same tristimulus values but with distinct spectral power distributions [20, 11]. Thus, it is impossible in vision algorithms to differentiate spectrally distinct colors which give rise to the same RGB values.

A more accurate color descriptor that directly depicts the optical properties of the material is the ratio of the amount of reflected light divided by the amount of incident light over a continuum of wavelengths. Spetrophotometers can measure this color quantity at very high spectral resolutions (resolutions finer than 0.3nm are not uncommon). Thus, an accurate objective color descriptor for the skin should incorporate spectrophotometric measurements. In the last five years, vision researchers like Ohtsuki and Healey [13] and Stoerring et al [17, 18] have looked into exploiting skin spectrograph data in skin color detection and skin modeling. However, their techniques are still limited to using input images in RGB form. Furthermore, their skin spectrograph data come from small datasets (approximately 7 subjects) [2, 5] collected in the late 1920-30s, with inferior quality optical equipment.

Nowadays, spectrographs have become off-the-shelf equipment and the emergence of new technologies allows for more compact, more robust and higher accuracy spectrophotometric equipment. The first step in our research was to collect new skin color data with modern equipment for multiple skin samples (see section 3). As expected, the captured color data revealed details which can not be captured with a typical RGB sensor. Analysis of the sprectrophotometric data allowed us to formulate a biological explanation of skin color and its variations (see section 5). It also enabled us to discriminate between human skin and dyes designed to mimic human skin color.

*. This material is based upon work supported by the National Science Foundation under Grant No. 0085864.

As revealing as raw spectrograph data is, it is often unwieldy and may not reveal 2D/3D information. The compromise between tristimulus data collection and spectrographic information is the use of color filters specifically selected to give maximal response only in the presence of human skin. The choice of filters, in terms of number of filters, their central wavelength and their bandwidth, is crucial and can only be made after thorough analysis and modeling of the skin spectral data.

Similarly, the use of spectrophotometric data in graphics rendering is usually avoided because it is prohibitively expensive both space- and complexity-wise. The preferred representation should have low storage and computation requirements and should be easily incorporated in the rendering pipeline.

Thus, the key to utilizing the high accuracy spectrophotometric data is to build a multispectral (i.e. beyond RGB) model for skin color which is based on spectrograph observations but can be employed in skin recognition and rendering. The challenge is: given dense spectral data, e.g. samples every 0.3nm, to model the empirical data in a numerically advantageous and rigorous manner. Moreover, the modeling of the data should accommodate color processing/analysis in the vision and graphics communities. The derived model should be computationally efficient; should involve few intuitive parameters; and adhere to physical constraints (e.g. not allow negative spectral values, or negative wavelengths).

We investigated 3 different modeling techniques (see section 6): linear combinations of Gaussians, wavelets, and principal components. In selecting the best modeling methodology we considered two factors. One has to do with the quality of the approximation and is based on the accuracy, computational efficiency and discernibly power of each model. The other factor is the form of the function itself in terms of how it can best fit the vision and graphics applications at hand. With these criteria in mind, the best skin color model is a five-chromatic model based on linear combination of Gaussians (see section 7).

2. Tristimulus color inaccuracies

In computer graphics determining the color of a surface is a key aspect of the rendering process. Ideally, in order to render an image correctly one has to trace every path of light from its source to the camera, while accurately modeling light reflection and refraction and performing all the computations at each wavelength of the visible spectrum. Once the color of a pixel is determined, one can reproduce the color's *perception* by a combination of three color primaries, typically RGB. In practice, because of time and space concerns, all the data is converted to RGB at the very beginning of the ray tracing process and all the computations are performed in tristimulus space.

The drawback of this simplification is that it introduces color distortions which become more prominent with each subsequent reflection. For example, consider the simplest reflectance model, i.e. the Lambertian model. Let $E(\lambda)$ be the spectrum of the incident light and $\rho(\lambda)$ be the albedo of the surface. The spectrum of the reflected light is:

$$I(\lambda) = E(\lambda)\rho(\lambda)\cos\theta$$

where θ is the angle of incidence. Assume that the surface reflects light only at a relatively narrow band of the visible spectrum, e.g. 400nm-430nm and that the light source does not emit any light in that range,e.g. its spectral range starts from 435nm, then the surface should appear black (see fig. 1a).

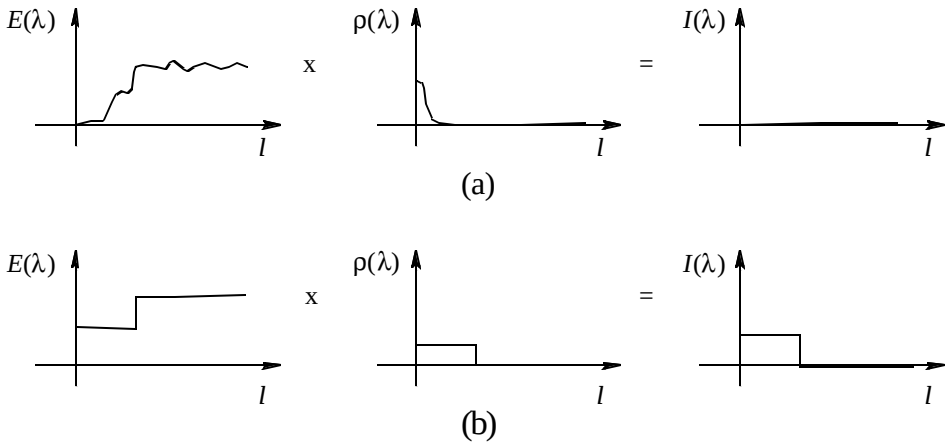


Fig. 1. Pictorial representation of (a) full spectrum color computation vs. (b) color determination after conversion to RGB.

When we reduce a spectral distribution to an RGB triplet, we first compute the XYZ values for the given spectrum and then convert that value to RGB:

$$X = \int_{\lambda} \bar{x}(\lambda)\rho(\lambda)\,d\lambda$$
$$Y = \int_{\lambda} \bar{y}(\lambda)\rho(\lambda)\,d\lambda$$
$$Z = \int_{\lambda} \bar{z}(\lambda)\rho(\lambda)\,d\lambda$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} RGB \\ to \\ XYZ \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

where $\bar{x}$, $\bar{y}$ and $\bar{z}$ are the CIEXYZ tristimulus matching functions [20]. The RGB values are directly related to the area under the curve which results from the convolution of the spectral distribution with $\bar{x}$, $\bar{y}$ and $\bar{z}$ respectively. For the particular example at hand, the area will be non-zero for B, resulting in a surface that has a blue hue, instead of black. For a pictorial representation of the distortion see fig. 1b.

In the field of computer vision, tristimulus color inaccuracies often manifest themselves as loss of information caused by metamerism. When we capture RGB data we end up projecting an infinitely dimensional color space to a three dimensional space. This is a many-to-one projection and results in colors with different spectral distributions giving the same RGB response. This phenomenon is an occurrence of metamerism. Metameric colors can be discriminated only if we use one of the following techniques: a) we capture the reflected light with a spectrograph instead of an RGB sensor; b) we use an illuminant whose spectrum is specifically chosen so that the reflected light will form distinct RGB triplets; or c) we use a different projection to the tristimulus space that will preserve the distinctiveness of the spectral distributions. With the exception of the first methodology, the other two techniques require individual solutions for different sets of metamers.

From the perspective of having a sensor for depicting a specific set of metamers (e.g. skin vs. mannequin) the use of a different color projection offers the highest versatility and most economical option. Still, in order to design the ideal projection for a specific metamer, we have to know the spectral distribution of the material of interest.

3. Spectrophotometric data

For our analysis we used new skin spectrograph data that we collected at the radiometric facility at the General Robotics and Sensory Perception (GRASP) Laboratory at the University of Pennsylvania. The skin reflectance measurements are for a $0°$ angle of incidence and approximately $4°$ angle of reflectance. Thus, what we are measuring is:

$$BRDF(4,0;0,0;\lambda) = \frac{Reflected(4,0;0,0;\lambda)}{Incident(0,0;\lambda)}$$

where $BRDF(\theta_r, \varphi_r; \theta_i, \varphi_i; \lambda)$ is the five parameter Bidirectional Reflectance Distribution Function. θ_r, φ_r are the spherical coordinates of the angle of reflectance, θ_i, φ_i are the spherical coordinates of the angel of incidence and λ is the wavelength at which the BRDF is measured. For these angles of incidence and reflectance the BRDF is approximately equivalent to the normal albedo [15].

We measured the skin reflectance of the hands of 23 volunteers from 20 to 40 years old (18 males and 5 females). We tried to cover a diverse collection of skin tones: 16 of our subjects were Caucasian (shown in red in the following figures), 3 were Asian (shown in green), 2 were of African origin (shown in magenta) and 2 were Indian (shown in blue). For each subject we took 4 different measurements: from the centers of the back and the palm of their hand, each at 2 different exposure times.

Figure 2 shows the spectrum of the back of the hand for some of our volunteers. All the plots exhibit a gradual increase with respect to wavelength with a dip around 575nm. The skin spectra of various races are interspersed and no clear classification can be done. There is one exception. One can observe that darker shaded skin, independent of race, reflects a smaller proportion of the incident light (which is how a darker surface can be described) and does not exhibit the curvature variation of the other plots. For additional skin data and details on the data collection setup see [1].

.

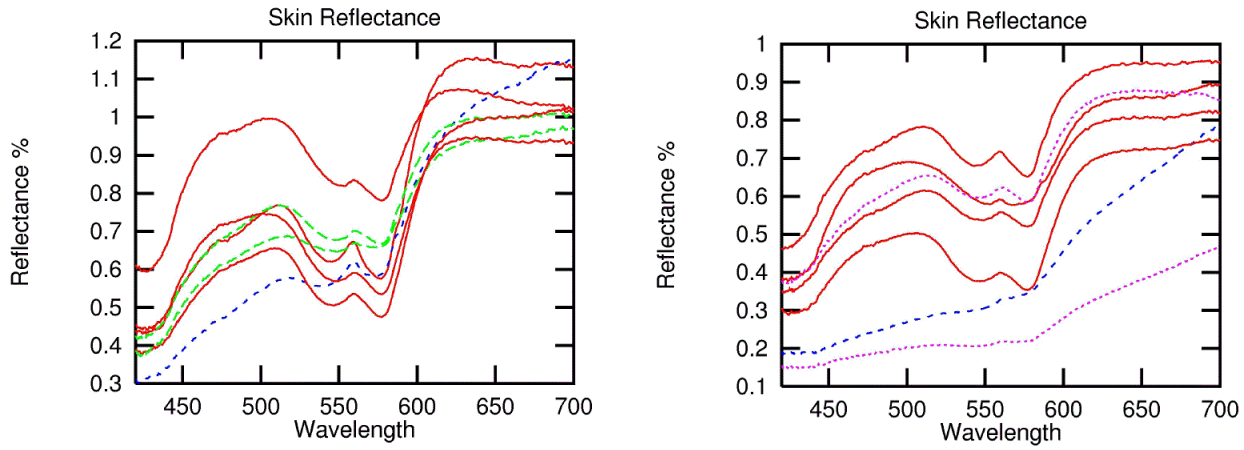


Fig. 2. Reflectance spectra of the center of the back of the hand for different races. Caucasian is shown in red, Asian in green, East Indian in blue and African descent in magenta.

In order to test whether skin spectra variations could be attributed to melanin and local skin structures (hair follicles, hair, pores, etc.) we measured the spectra of the palm of the same volunteers. The measured spectra of the palms still exhibit the same general pattern (fig. 3): a gradual increase with respect to wavelength with a dip around 575nm. As expected, the palm has a more reddish spectrum than the back of the hand. Notice that the spectra are much more closely clustered and exhibit almost identical shape, even for the heavily pigmented individuals.

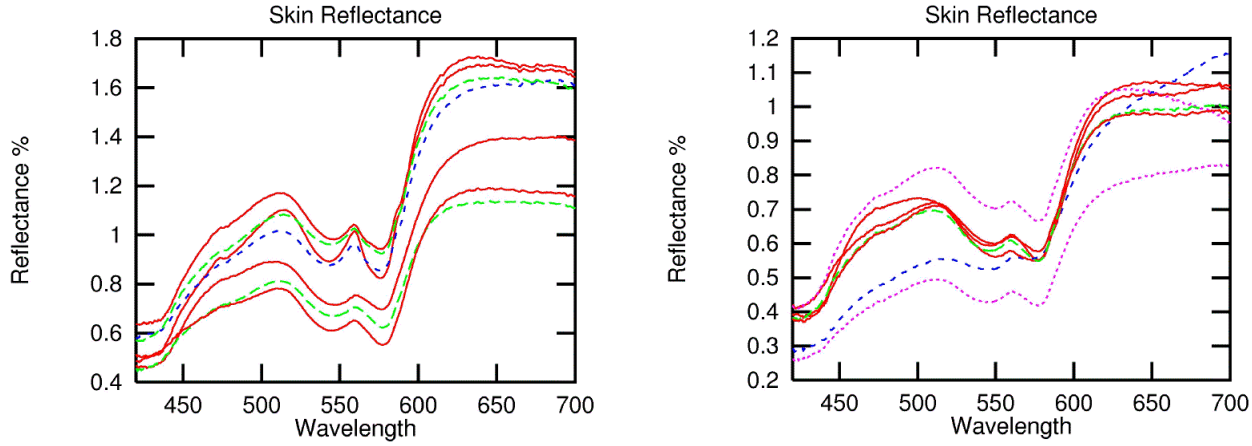


Fig. 3. Reflectance spectra of the center of the palm for different races. Caucasian is shown in red, Asian in green, East Indian in blue and African descent in magenta.

The measured spectra demonstrate that there is a very specific pattern in the spectral distribution of the color of skin. This prompted us to investigate whether materials made to mimick human skin exhibit a similar pattern. We took 5 measurements of the spectrum of light reflected by a mannequin under the same experimental conditions. Figure 4 shows the spectra of the mannequin (in dashed lines) and the spectra of the back of the hand of four of our volunteers of different race (in solid lines). These plots demonstrate that dyes mimicking human skin can have very distinct spectral distributions.

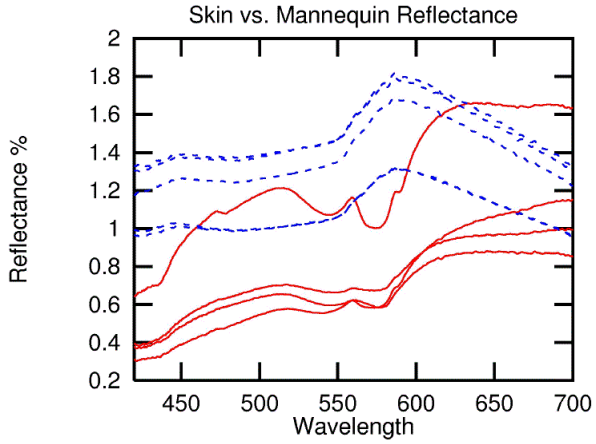


Fig. 4. Reflectance spectra of the back of human hands of different races (shown in red) compared to that of a store mannequin (shown in blue).

4. Analysis of the spectrograph data

The albedo graphs suggest that at approximately 560nm there is a persistent "**W**" pattern (two dips with a bump in the middle), which seems to be unique to the human skin. To verify this hypothesis we used the following algorithm for detecting the location of this "**W**" pattern.

"**W**" pattern detection algorithm:

1. Find all local minima and maxima of the curve.
2. Find the two smallest local minima after 430nm.
   They should be the two dips of the "**W**" pattern.
3. Let λ_1 and λ_3 be the wavelengths of the leftmost and rightmost dips respectively.
4. If λ_3-λ_1<50nm (the maximum "**W**" width)
   Find the max. between λ_1 and λ_3 (middle bump)
   Let λ_2=wavelength of the middle bump.
   else
   No "**W**" pattern is present

Our algorithm detected the "**W**" pattern in 44 out of the 51 samples. The 7 samples where no "**W**" pattern was detected were all 5 of the mannequin samples and 2 from the back of the hands of the most heavily pigmented volunteers. None of these 7 samples exhibited the desired pattern. Thus, our algorithm was 100% successful in locating the "**W**" pattern when present. More importantly, the algorithm allowed us to compute the location of the local minima and maxima that define the "**W**" pattern (see Table 1).

Table 1: Wavelengths of local min. and max.

Feature	Location Median λ	Location Mean λ	Standard Deviation
Left Min.	546.42	546.56	2.54
Local Max	559.48	559.72	0.95
Right Min.	576.26	575.45	2.34

5. Human Skin Reflectance

All the analysis thus far indicates that existence of the "**W**" pattern implies the presence of live human skin. Such a strong statement should also be supported by the physics of skin reflectance. In this section we will briefly describe skin chromophores and how they contribute to skin reflectance. A more extensive presentations of skin biology and its optics is beyond the scope of this paper. The interested reader can check [4, 10, 19, 1].

The skin is composed of two layers. The outer layer is the epidermis and the inner layer is the dermis. The main chromophore in the epidermis is melanin. The exact spectrum of melanin in humans is unknown. In general, the absorption spectrum of melanin in the visible range is monotonically decreasing with wavelength. The various degrees of skin pigmentation are directly related to the amount of melanin in the epidermis. Highly pigmented individuals have more than 7 times the amount of melanin-producing cells than lightly pigmented subjects [10].

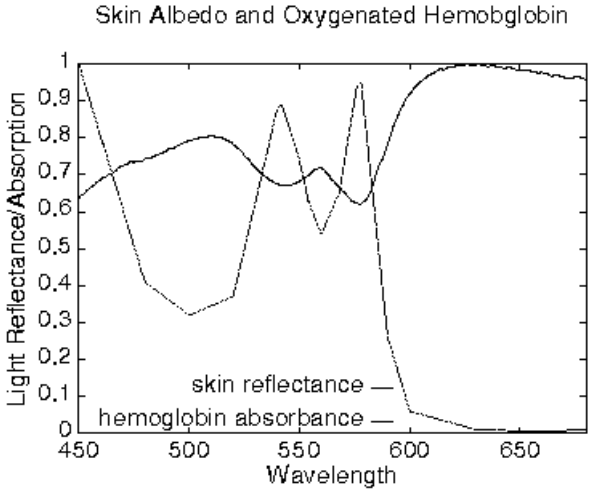


Fig. 5. A sample skin reflectance plot versus the hemoglobin absorption plot.

The dermis is heavily permeated with blood vessels which contain hemoglobin, Hb, a protein which binds very easily to Oxygen. Oxygenated hemoglobin has a unique absorption spectrum with characteristic absorption bands at 420nm, 542nm, 560nm and 576nm, right where the "**W**" pattern is observed (see Table 1). Figure 5 shows the skin reflectance spectrum plotted against the HbO_2 absorption spectrum (HbO_2 data is taken from Zijlstra et al [21]).

To summarize, light incident on skin must first encounter the epidermis. The melanin in the epidermis acts as a filter whose transmission rate increases with wavelength. The light that is not absorbed by melanin penetrates the epidermis and reaches the dermis. In the dermis, the hemoglobin in the blood vessels causes the formation of a "**W**" pattern in skin albedo. Heavily pigmented skin has increased amounts of melanin which absorbs most of the light in the epidermis, allowing a much smaller percentage of the incident light to reach the vasculature of the dermis. Thus, the hemoglobin absorption bands, although still present, are not detectable.

6. Modeling

A natural extension to these observations is to build a model for skin color that encapsulates the "**W**" pattern. Our modeling technique should be applicable to any arbitrary spectral distribution $d(\lambda)$. The resulting approximation could either be tuned for use as an explicit model for rendering, or used to derive a feature vector for recognition purposes.

6.1. Gaussian modeling

We propose approximation of the spectral distribution as a linear combination of N Gaussian functions of unknown means μ_k and standard deviations σ_k with amplitudes γ_k described by:

$$d(\lambda) = \sum_{k=1}^{N} \gamma_k g(\lambda, \mu_k, \sigma_k)$$

An example of fitting Gaussians to skin spectral profile is shown for $N = 4$ in fig. 6(a). We also investigate the performance of the first derivatives of Gaussians as basis functions. Fig. 6(b) shows an example of fitting $N = 4$ such functions to a skin spectral profile after the mean value was subtracted. In both cases, in order to compute the best fit we minimize the L_2 norm, yielding a non-linear optimization with respect to the N triples $(\gamma_k, \mu_k, \sigma_k)$, $k = 1 \ldots N$.

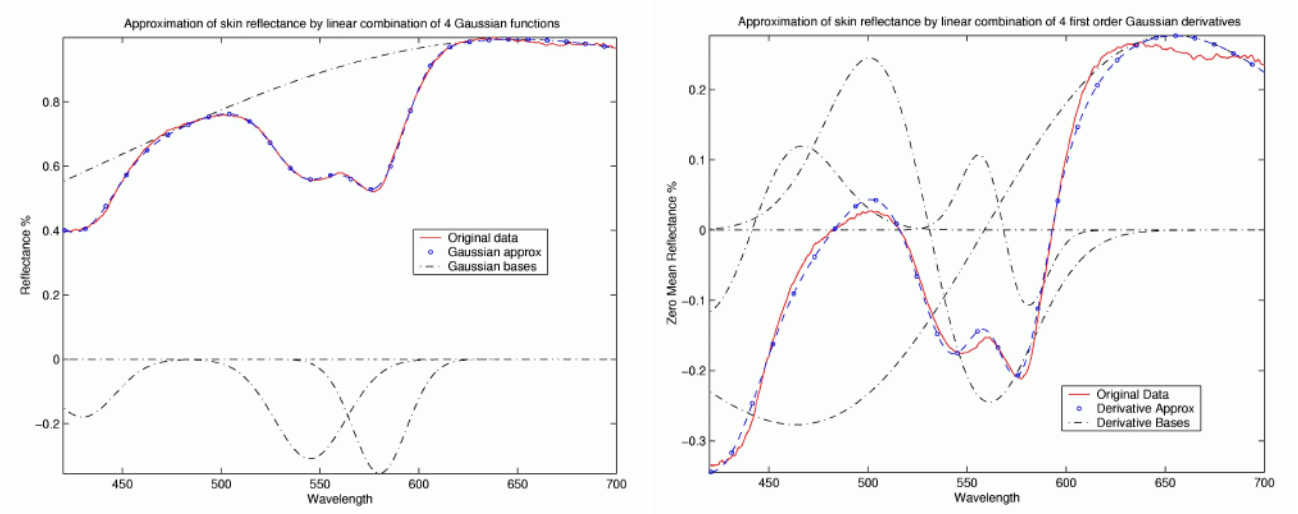


Fig. 6. Approximation of a typical back of hand spectrum as a linear combination of (a) 4 zeroth order Gaussian basis functions and (b) 4 first order Gaussian basis functions. The 4 basis functions are shown as dotted lines.

Gaussians and their first derivatives can be regarded as non-orthogonal wavelets similar to Gabor functions. However, multi-resolution wavelet approximations (even if they use the same Gaussian basis) consider the contributions at every possible offset and scale. Our approximation modeling, though a non-linear process, has the best localization properties and can be optimized with respect to the number of basis functions used. In that sense, it is very similar to Radial Basis Function approximation networks [14], to the approximation of curves by sums of Gaussians used in graphics and image processing [8], and to the Gaussian sum filtering used in estimation theory [16].

Our approach has been motivated by our desire to construct an "N-chromatic" sensor that will best approximate the dense sampling obtained with a spectrophotometer. Assume that we are given a set of m filters, F_j, $j = 1 \ldots m$. The effect of each filter can be modeled as the projection of the reflectance $d(\lambda)$ on the filter's spectral response function $f_j(\lambda)$. The filtered reflectance $d_j(\lambda)$ is:

$$d_j(\lambda) = \int f_j(\lambda)\, d(\lambda)\, d\lambda$$

Replacing $d(\lambda)$ with the proposed approximation yields:

$$d_j(\lambda) = \sum_{k=1}^{N} \gamma_k \int g(\lambda, \mu_k, \sigma_k) F_j(\lambda)\, d\lambda$$

We need at least N filter responses to obtain the N coefficients $\gamma_{k=1 \ldots N}$. In practice, such a requirement is very reasonable. As we will show in see section 7 we get very good skin reflectance approximations, across the 400nm-700nm range with $N = 5$ for Gaussian functions or $N = 4$ for first derivatives of Gaussians.

Note that recovering the $\gamma_{k=1 \ldots N}$ coefficients is a linear process which uses only the basis functions from the modeling step and the filter specifications. Consider the matrix F whose elements F_{jk} are the projections of the filter's spectral response functions on the basis functions. F must not be singular. This implies that recovering the entire spectral distribution is feasible even with just 3 carefully selected filters, though not necessarily RGB.

The optimal solution involves tuning the filter characteristics so that the projection matrix with elements F_{jk} is well-conditioned. For a pre-specified domain like the skin, we can approximate the original distribution $d(\lambda)$ as a linear sum of

the filter responses $F_k(\lambda)$. In that case, each matrix element F_{jk} is the projection of the spectral response of filter F_j on the spectral response of filter F_k .

6.2. Wavelets and PCA

We investigated two possible alternatives to the proposed modeling. The first is using an orthogonal wavelet approximation [12]:

$$d(\lambda) = \sum_{j=0}^{J} \sum_{n=0}^{2^L} \langle d, \psi_{j,n} \rangle \psi_{j,n}(\lambda)$$

with

$$\psi_{j,n}(\lambda) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{\lambda - 2^j n}{2^j}\right)$$

We use Wavelab's [3] implementation of orthogonal wavelet multi-resolution analysis. For J resolution levels, Wavelab stores the scaling function projection $\phi(\lambda)$ only for the coarsest level J. For all the other levels, Wavelab stores only wavelet projections. Thus, the spectral distribution can be approximated as follows:

$$d(\lambda) = \sum_{n=0}^{2^L} \langle d, \phi_{J,n} \rangle \psi_{J,n}(\lambda) + \sum_{j=0}^{J-1} \sum_{n=0}^{2^L} \langle d, \psi_{j,n} \rangle \psi_{j,n}(\lambda)$$

where the $\phi_{J,n}(\lambda)$ is the scaled and shifted version of $\phi(\lambda)$. As a basis we use the polynomial spline Battle-Lemarie wavelets which have exponential decay resembling the second Gaussian derivative.

In order to describe the spectral distribution $d(\lambda)$ with few coefficients we sort the wavelet coefficients and select the top M with thresholding. Our experiments indicate that, compared to our Gaussian modeling, we always need a significantly higher number of coefficients in order to achieve the same approximation error. Though this might be reasonable for recognition purposes it does not facilitate the inverse problem: recovering the entire spectral distribution from few filter responses.

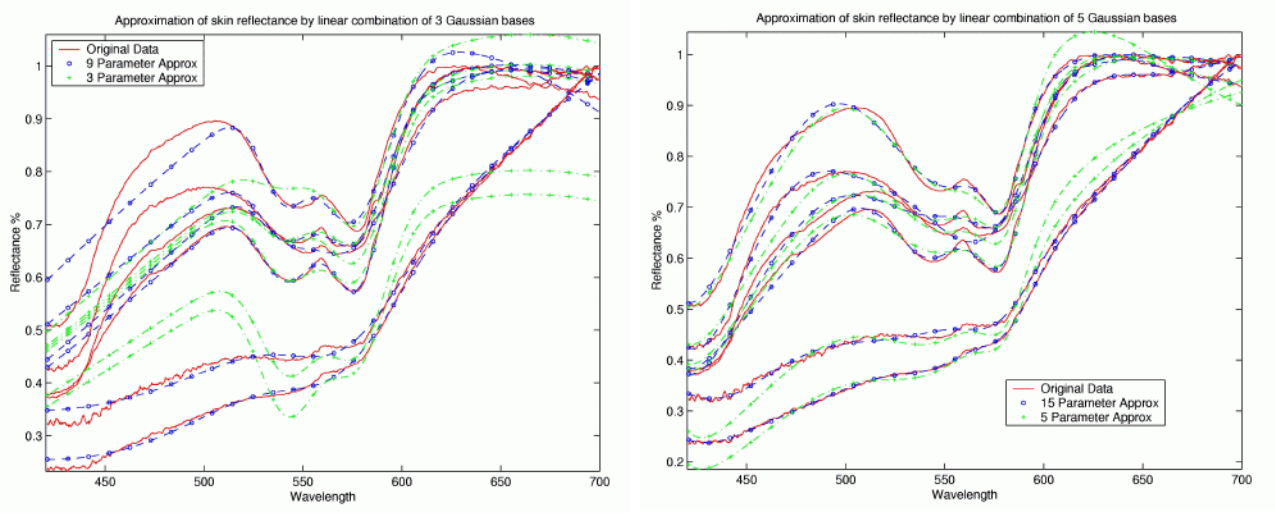
A second alternative to Gaussian modeling is a representation based on principal components. Consider a class of spectral distributions (e.g. skin albedo). Representative samples from that class are grouped into a zero-mean data matrix D with dimensions equal to the length of the function and the number of samples. The class can be represented by the most significant eigenvectors of the covariance $D^T D$. Each new data vector can be projected on these eigenvectors and represented by the coefficients of this projection.

7. Modeling Experiments

7.1. Gaussian modeling

All in all, 5 Gaussian modeling schemes were tested: the sum of $N = 3, 4, 5$ Gaussian functions and the sum of $N = 4, 5$ first derivatives of Gaussian functions. For each scheme, best fits to each of the 46 skin spectral profiles were initially computed on an individual basis with respect to all $3N$ model parameters, over the full 400nm-700nm spectral range. Fig. 7 shows these individually-fitted approximations (plotted as 'o' in blue) for 6 exemplary spectral profiles of skin (plotted as solid lines in red). In fig. 7 (a) we show the approximations obtained by the sums of 3 zeroth order Gaussians, while in fig. 7 (b) we show the results obtained by the sum of 5 zeroth order Gaussian functions. Fig. 8 shows the approximations for the same exemplary skin profiles using the sum of 4 Gaussian derivatives.

The sum of 5 Gaussians was found to give the lowest mean average root mean squared (RMS) reconstruction error (1.0%, calculated over all 46 spectra, over the entire fitting range), closely followed by the sum of 4 Gaussian derivatives (1.2%). Although, the sum of 3 Gaussians shows poorer performance over the full spectral range (mean average RMS reconstruction error of 2.6%), it is locally very accurate for the majority of skin profiles, with an RMS error of 0.7% over the 500-620nm range - 20% better than that of the other schemes. Qualitatively, it leads to a superior fit over the "**W**" pattern (see Table 2).



(a) (b)

Fig. 7. Approximation of the original spectral data of six individuals using a linear combination of (a) 3 zeroth order Gaussian basis functions and (b) 5 zeroth order Gaussian basis functions.

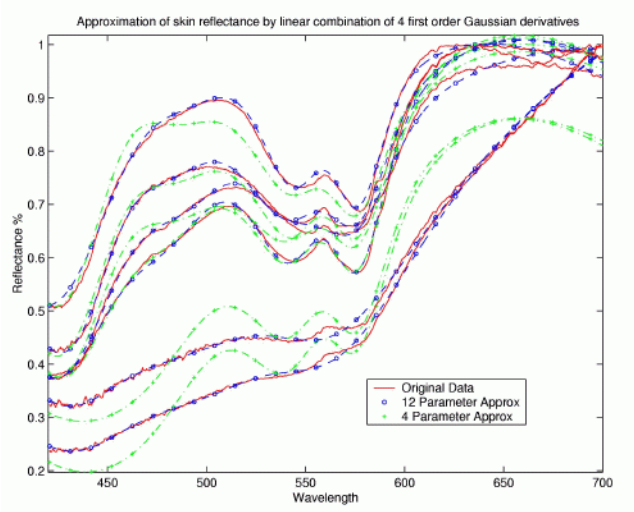


Fig. 8. Approximation of the same original data using a linear combination of 4 first order Gaussian basis functions.

Table 2: Performance of 4 tested Gaussian schemes

Scheme	Sum of 3 Gaussians	Sum of 4 Gaussians	Sum of 5 Gaussians	Sum of 4 Gaussian Derivatives
Mean RMS Error	2.6%	1.2%	1.0%	1.5%

For the 3 illustrated schemes (e.g. sum of 3 Gaussians, sum of 5 Gaussians and sum of 4 Gaussian derivatives), the μ_k and σ_k parameters estimated in the $3N$ parameter modelling were very tightly distributed over the 46 human skin spectra. The μ_k in particular, tends to show less than 1% variation across the 46 individual fits. With a view to skin recognition and also full spectrum recovery from N chromatic filters, we estimated a global set of $2N$ values for the μ_k and σ_k that seem to characterize human skin, for each of these 3 schemes, by optimizing the fit to all 46 of our skin spectral profiles. The result was a set of N Gaussian basis functions that are optimally placed to characterize skin, and fitting them to individual skin profiles was reduced to a linear optimization with respect to the N γ parameters. These approximations are also presented (plotted as '+' in green) for the same 6 spectral profiles in fig. 7 and fig. 8. The sum of 5 Gaussians scheme achieves the least RMS errors in reconstruction, both across the full spectrum (1.7%) and over the range centered on the "**W**" pattern (1.3%). In fact, if we choose successively smaller ranges, the RMS penalty of fitting 4 optimally placed Gaussians, or even 3 Gaussians is small. However, 5 parameters is not a high price to pay for coverage of the full spectrum, and so we consider this as our scheme of choice for filter selection.

A particular computer vision application for multispectral skin modeling is the design of a sensor capable of extracting the detailed skin spectra through the use of only a small number of bandpass filters, specifically chosen for skin spectra. The sums of Gaussian functions is very well-suited for this application, since many bandpass filters can be modelled as zeroth order Gaussian functions with central wavelength equal to μ and half-bandwidth approximately equal to 2σ. Thus, in order to best approximate the optimal bandpass filters we computed the means μ and standard deviations σ for each of the $N = 3, 4, 5$ Gaussian function schemes. We used the following methodology in estimating the best filter parameters for each scheme.

1. Find the best fit for each individual skin spectrum w.r.t. all 3 parameters, μ, σ and γ
2. Set the μ parameters at the mean of the 46 μ parameter values found in step 1, disregarding 'outlier' μ values.
3. With μ values fixed according to step 2, find the best fit to each individual skin spectrum w.r.t. the remaining 2 parameters σ and γ.
4. Set the σ parameters at the mean of the 46 σ parameters values found in step 3, disregarding 'outlier' σ values.

The sums of specific $N = 3, 4, 5$ Gaussian filters that best approximate the detailed spectrum of human skin are shown in Table 3, Table 4 and Table 5 respectively.

Table 3: Recommended set of 3 Gaussians

Parameters	1st Gaussian	2nd Gaussian	3rd Gaussian
μ	543	578	664
σ	14.93	14.85	198.05
mean γ values for fitting to skin	-0.20	-0.31	0.99
	mean RMS	4.9%	

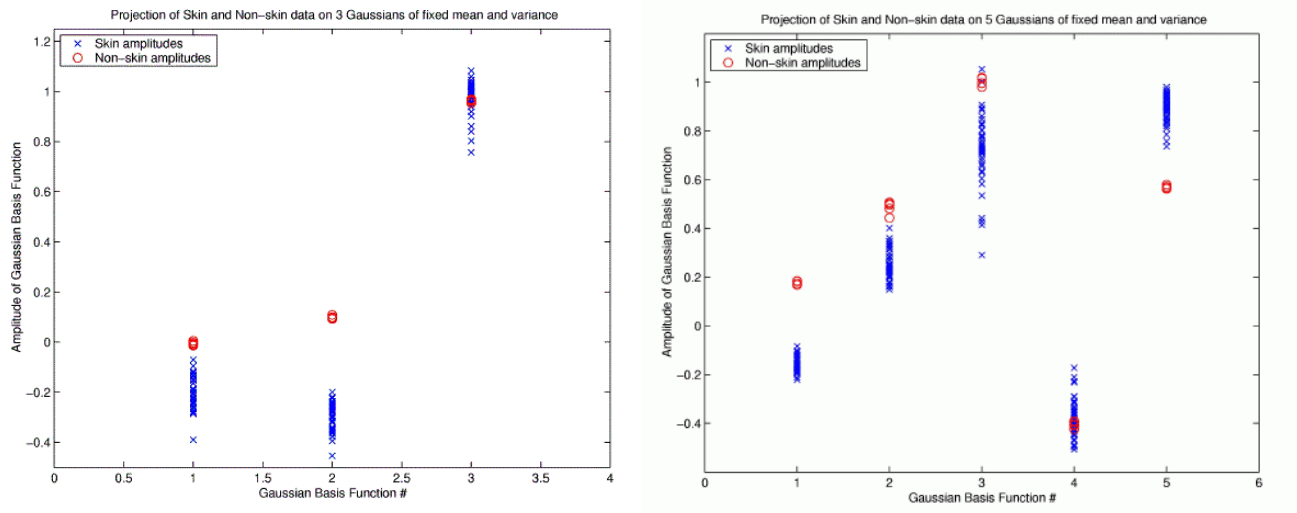
Table 4: Recommended set of 4 Gaussians

Parameters	1st Gaussian	2nd Gaussian	3rd Gaussian	4th Gaussian
μ	431	547	580	676
σ	24.32	25.07	12.54	277.32
mean γ values for fitting to skin	-0.24	-0.26	-0.23	0.99
	mean RMS	3.8%		

Table 5: Recommended set of 5 Gaussians

Parameters	1st Gaussian	2nd Gaussian	3rd Gaussian	4th Gaussian	5th Gaussian
μ	395	538	554	581	747
σ	14.42	89.78	33.08	10.54	123.55
mean γ values for fitting to skin	0.25	0.75	-0.39	-0.16	0.90
		mean RMS	1.7%		

Regarding skin recognition, one would hope that the amplitudes, γ_k, of the optimally placed Gaussian basis functions, can be used to discriminate between skin and non-skin profiles. Fitting the basis functions to the 5 spectra captured for the mannequin showed that this was indeed the case. The γ coefficients for skin were found to be tightly clustered for most of the schemes, and in all 5 schemes there was at least one coefficient that would enable 100% correct classification of all 51 spectra by simple thresholding. Results are shown for two of the schemes in fig. 9. Evidently, more examples of non-skin data are required to validate this as a skin detection scheme, however the typically tight within class co-variance of the γ coefficients suggests that this technique could be viable for only 4 or 5 basis functions at most.

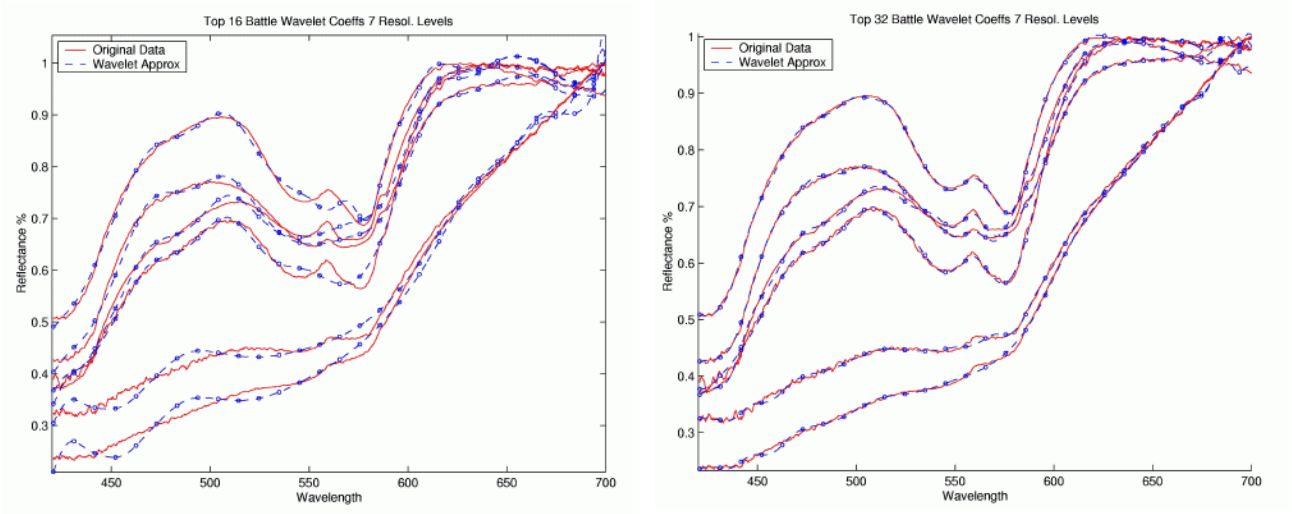


(a) (b)

Fig. 9. (a) The 3 amplitude parameters γ_k of the sum of 3 zeroth order Gaussians for 46 human skin (x) and 5 mannequin (o) data. Separation by vertical thresholding is possible for basis functions 1and 2. (b) The 5 amplitude parameters γ_k of the sun of 5 zeroth order Gaussians for the same data. Separation by vertical thresholding is possible for basis functions 1, 2 and 5.

7.2. Wavelets and PCA

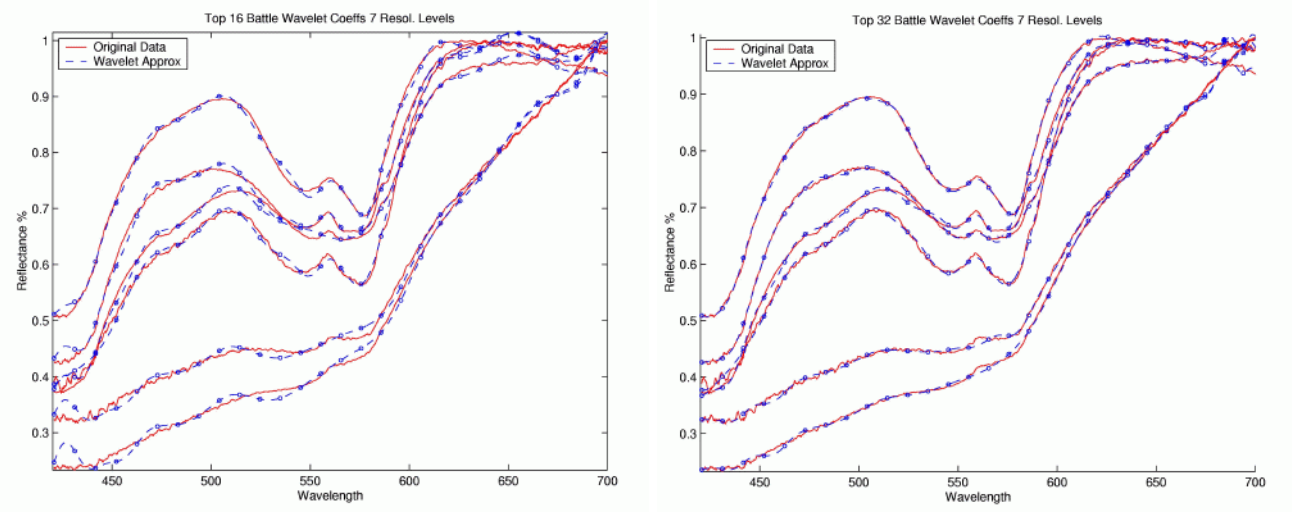
Our parameter choice in wavelets and PCA modeling was driven by maintaining similar RMS error levels with the Gaussian modeling. In both techniques we compute the best fit to each of the 46 skin spectral profiles over the full 400nm-700nm spectral range. In our wavelet modeling, we used seven resolution levels obtaining 8 scaling function coefficients for a spectrum length of 1024 and 1024-8 wavelet coefficients. Fig. 10 shows the individually-fitted wavelet approximations (plotted as 'o' in blue) for the same exemplary spectral profiles (plotted as solid lines in red) that were used in the Gaussian modeling figures. In fig. 10 (a) we show the approximations obtained if we keep only the 16 coefficients with the highest absolute values. In fig. 10 (b) we show the approximations we obtained by keeping only the 32 coefficients with the highest absolute values. Although the mean RMS error over all 46 skin spectra is comparable to the Gaussian modeling (2.03% for 16 coefficients and 0.61% for 32 coefficients), we observed an inability of both wavelet techniques to accurately fit the left and right ends of the spectrum.



(a) (b)

Fig. 10. Approximations of the original data for skins of the same six exemplary persons using (a) the highest 16 wavelet coefficients and (b) the highest 32 wavelet coefficients.

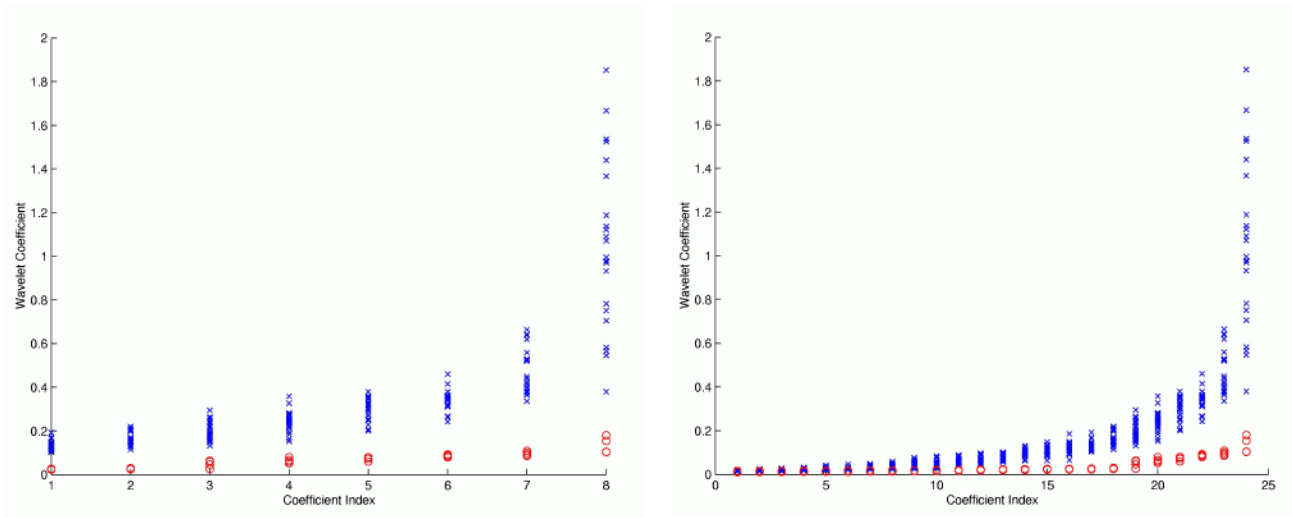
To alleviate the latter problem we selected the highest coefficients without considering coefficients with offsets close to the boundaries. As can be seen in fig. 11, the results for both 16 and 32 coefficients are better, resulting in an approximately 10% decrease in RMS error. The performance of the wavelet schemes are summarized in Table 6.



(a) (b)

Fig. 11. Approximations of the same data without considering coefficients with offsets close to the boundaries using (a) the highest 16 wavelet coefficients and (b) the highest 32 wavelet coefficients and

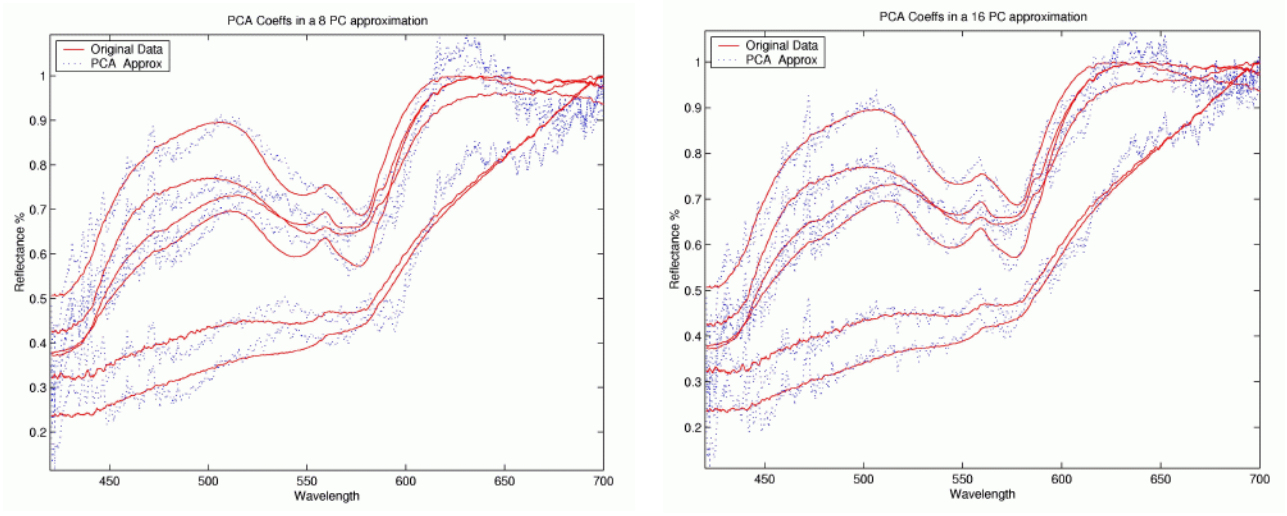
In skin recognition, one would expect that spectral distribution curves with different shapes (as in the case of human skin versus mannequin data) would generate wavelets with distinct coefficient values. We approximated each of the 23 back of the hand skin curves and 4 of our mannequin curves using the 16 highest and the 32 highest coefficients, and recorded the computed coefficients for each fit. Fig. 12 shows the highest wavelet coefficients excluding the scaling function coefficients of the coarsest level. The wavelet coefficients of non-skin data can be easily separated from real skin data, even without a high-dimensional classifier.



(a) (b)

Fig. 12. (a) The highest 8 wavelet coefficients for 23 human (x) and 4 mannequin (o) skins. Separation is easy even with just vertical thresholding. (b) The highest 24 wavelet coefficients for 23 human (x) and 4 mannequin (o) skins. Separation with vertical thresholding becomes even easier.

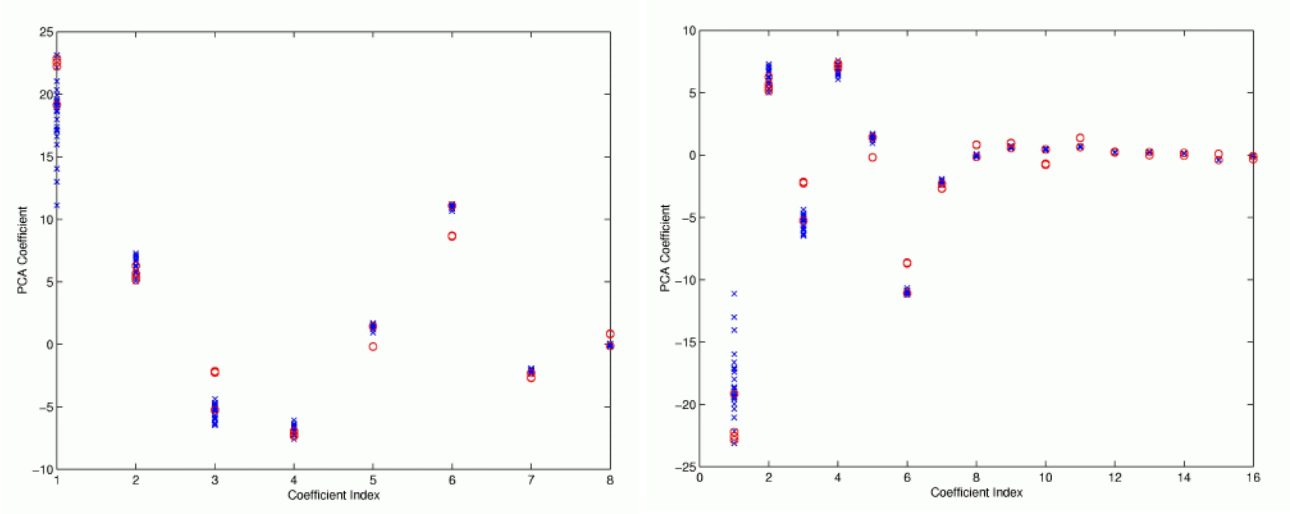
In the Principal Component Analysis approach we create a zero-mean data matrix using 13 representative skin samples. We use the eigenvectors which correspond to the *N* highest eigenvalues as the basis of our multispectral skin model. Each of the skin spectra is projected on this basis. In our approximations we used the 8 and 16 most significant eigenvectors. The results of these 2 approximations (plotted in blue) on the same exemplary profiles (plotted in red) that were used in the previous subsection are shown in fig. 13(a) and (b) respectively. In contrast to the wavelet approaches, the approximation via eigenvectors is very noisy with mean RMS errors sometimes an order of magnitude higher than in the wavelet approximation (4.65% for 8 eigenvectors and 3.27% for 16 eigenvectors). The mean RMS error of the all the modeling schemes are summarized in Table 6.



(a) (b)

Fig. 13. Approximations of the original data for skins of six exemplary persons using (a) the 8 highest ranked eigenvectors and (b) the 16 highest ranked eigenvectors.

When we analyzed the projection coefficients for both human skin and mannequin, we saw that projections to some of the eigenvectors can clearly build a discernible feature vector for skin. Please note, that the PCA approaches need to provide not only the coefficients but also the non-closed form basis functions: the 1024-dimensional eigenvectors. Our PCA approach, as well as our wavelet approximations used closed-form a-priori given basis functions. Another interesting observation is that none of the PCA basis functions could be directly associated with the reflectance spectra of the main skin chromophores, i.e. melanin and oxygenated hemoglobin.



(a) (b)

Fig. 14. (a) Projection coefficients on 8 eigenvectors characterizing the human skin.The 3rd, 5th and 8th eigenvectors encode information particular to the human skin. (b) Projection coefficients on 16 eigenvectors characterizing the human skin. The 3rd, 5th, 8th, 10th and 11th eigenvectors encode information particular to the human skin.

Overall, the majority of the modeling methods produced a mean RMS error of about 2.1% (see Table 6). If one was to choose a model based solely on accuracy and discernibility power, the clear choice would be the 32 coefficient wavelet model. However, according to Occam's principle (*Occam's Razor: one should not increase, beyond what is necessary, the number of entities required to explain anything*) [9], given a set of roughly equivalent models, i.e. the 32 coefficient wavelet model and the sum of 5 zeroth order Gaussians, one should choose the simplest one, which is the sum of 5 Gaussians. Furthermore, the sum of Gaussians fits naturally in both vision applications (filters for skin sensor) and graphics rendering (spectral convolutions).

Table 6: Performance of the all the different modeling schemes

Scheme	Mean RMS Error
Sum of 3 Gaussians	2.6%
Sum of 4 Gaussians	1.2%
Sum of 5 Gaussians	1.0%
Sum of 4 Gaussian Derivatives	1.5%
16 coef. Wavelets	2.03%
32 coef. Wavelets	0.61%
8 eigenvectors PCA	4.65%
16 eigenvectors PCA	3.27%

8. Conclusions

We showed that, although prevalent, RGB color representation introduces color distortions and loss of information. An improvement on skin color modeling would be to depart from the traditional tri-stimulus paradigm. Using our own skin spectrograph data, we showed that spectrophotometric information reveals skin color details that help explain the uniqueness of skin color and its variations. Despite differences in the spectral distribution of skin samples, 95% of them exhibit a certain pattern which is directly tied to human skin composition.

Given the empirical spectrograph data we investigated different sets of basis functions for developing a model that closely approximates skin spectral distribution. The performance of wavelets and linear combinations of Gaussians was comparable. We opted for the Gaussian representation because it provides us with a description of filters that are optimal for the detailed reconstruction of the skin's spectral distribution.

9. References

[1] E. Angelopoulou, "Understanding the Color of Human Skin," *SPIE Conference on Human Vision and Electronic Imaging VI*, SPIE Vol. 4299, 2001, pp. 243-251.

[2] L.A. Brunstig and C. Sheard, "The Color of the Skin as Analyzed by Spectrophotometric Methods," *The Journal of Clinical Investigation*, Vol. 7, 1929, pp 559-613.

[3] J. Buckheit and D. Donoho, *Stanford University Technical Report 474: Wavelab and Reproducible Research*, Stanford University, 1995.

[4] C.D. Clement, *Gray's Anatomy of the Human Body*, 13th ed., Williams & Wilkins, 1985.

[5] E.A. Edwards and S.Q. Duntley, "The Pigments and Color of Living Human Skin," *American Journal of Anatomy*, Vol. 65, No. 1, July 1939, pp. 1-33.

[6] R.A. Hall, "Comparing Spectral Color Computations," *IEEE Computer Graphics and Applications*, Vol. 19, No. 4, July 1999, pp. 36-45.

[7] R.A. Hall and D.P. Greenberg, "A testbed for realistic image synthesis," *IEEE Computer Graphics and Applications*, Vol. 3, No. 8, November 1983, pp. 10-19.

[8] A. Goshtasby and W.D. O'Neill, "Curve Fitting by a Sum of Gaussians," *CVGIP: Graphical Models and Image Processing*, Vol. 56, No. 4, July 1994. pp. 281-288.

[9] F. Heylighen, "Occam's Razor," *Principia Cybernetica Web*, http://pespmc1vub.ac.be/OCCAMRAZ.html.

[10] S.L. Jacques, "Monthly News and Articles on Biomedical Optics and Medical Lasers," *Web-page of Oregon Medical Laser Center*, http://omlc.ogi.edu/news/index.html.

[11] D.L. MacAdam, *Color Measurement: Theme and Variations*, 2nd ed., Springer Verlag, 1985.

[12] S. Mallat, *A Wavelet Tour of Signal Processing*, Chapter 7, Academic Press, 1998.

[13] T. Ohtsuki and G. Healey, "Using Color and Geometric Models for Extracting Facial Features," *Journal of Imaging Science and Technology*, Vol 42, No. 6, December 1998. pp. 554-561.

[14] T. Poggio and F. Girosi, "Networks for Approximation and Learning," *Proceedings of IEEE*, Vol. 78, No. 9, September 1990, pp. 1481-1497.

[15] J. Schombert, *Web-page of J. Schombert's astronomy glossary at the University of Oregon*, http://zebu.uoregon.edu/~js/albedo.html.

[16] H.W. Sorenson and D.L. Alspach, "Recursive Bayesian Estimation using Gaussian Sums," *Automatica*, Vol. 7, 1971, pp. 465-479.

[17] M. Stoerring, H.J. Andersen and E. Granum, "Skin Colour Detection under Changing Lighting Conditions," *Symposium on Intelligent Robotics Systems*, July 1999. pp. 187-195.

[18] M. Stoerring, H.J. Andersen and E. Granum, "Estimation of the Illuminant Colour from Human Skin Colour," *IEEE Conference on Automatic Face and Gesture Recognition*, March 2000. pp. 26-30.

[19] M.J.C. Van Gemert, S.L. Jacques, H.J.C.M. Sterenborg, and W.M. Star, "Skin Optics," *IEEE Transactions on Biomedical Engineering*, Vol. 36, No. 12, 1989, pp. 1146-1154.

[20] G.Wyszeski and W.S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., John Wiley & Sons, 1982.

[21] W.G. Zijlstra, A. Buursma and W.P. Meeuwsen-van der Roest, "Absorption Spectra of Human Fetal and Adult Oxyhemoglobin, De-oxyhemoglobin, Carboxyhemoglobin, and Methemoglobin," *Clinical Chemistry*, Vol. 37, No. 9, 1991, pp. 1633-1638.