A Practical Appearance Model for Dynamic Facial Color

Jorge Jimenez* Timothy Scully† Nuno Barbosa‡ Craig Donner§ Xenxo Alvarez¶
Teresa Vieira‖ Paul Matts** Verónica Orvalho†† Diego Gutierrez* Tim Weyrich†

*Universidad de Zaragoza †University College London ‡Universidade do Porto §Leolux ¶Face In Motion
‖Universidade de Trás os Montes e Alto Douro **Procter & Gamble ††Instituto de Telecomunicações



Figure 1: *Example results of our automatic changes in skin appearance predicted by our method. Changes are due both to mechanical deformations and to involuntary dilation or constriction of blood vessels caused by emotions; all affect the skin's hemoglobin distribution. Our real-time model allows simulation of both, based on* in vivo *measurements of real subjects, and runs in real-time (this scene with five heads runs at 53 frames per second). Our method is easily adopted into existing animation pipelines. From left to right, we show a sad smile, anger, the neutral pose, fear and disgust. The different hemoglobin maps produced by our model are shown in Figure 2.*

Abstract

Facial appearance depends on both the physical and physiological state of the skin. As people move, talk, undergo stress, and change expression, skin appearance is in constant flux. One of the key indicators of these changes is the *color* of skin. Skin color is determined by scattering and absorption of light within the skin layers, caused mostly by concentrations of two chromophores, melanin and hemoglobin. In this paper we present a real-time dynamic appearance model of skin built from *in vivo* measurements of melanin and hemoglobin concentrations. We demonstrate an efficient implementation of our method, and show that it adds negligible overhead to existing animation and rendering pipelines. Additionally, we develop a realistic, intuitive, and automatic control for skin color, which we term a *skin appearance rig*. This rig can easily be coupled with a traditional geometric facial animation rig. We demonstrate our method by augmenting digital facial performance with realistic appearance changes.

ACM Reference Format
Jimenez, J., Scully, T., Barbosa, N., Donner, C., Alvarez, X., Vieira, T., Matts, P., Orvalho, V., Gutierrez, D., Weyrich, T. 2010. A Practical Appearance Model for Dynamic Facial Color. *ACM Trans. Graph. 29*, 6, Article 141 (December 2010), 9 pages. DOI = 10.1145/1866158.1866167 http://doi.acm.org/10.1145/1866158.1866167.

Copyright Notice
Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or direct commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701, fax +1 (212) 869-0481, or permissions@acm.org.
© 2010 ACM 0730-0301/2010/12-ART141 $10.00 DOI 10.1145/1866158.1866167
http://doi.acm.org/10.1145/1866158.1866167

1 Introduction

Facial appearance constantly changes as people talk, change expression, or alter their physical or emotional state. Previous work mainly focuses on the geometric aspects of these changes, such as animating the facial surface (e.g. wrinkle structures, skin stretching). We refer the reader to the survey by Ersotelos and Dong for a comprehensive cross section of the different techniques [2008]. Equally important are changes in skin color caused by differences in hemoglobin concentration [Moretti et al. 1959], which may also occur due to histamine reactions or other skin conditions such as rashes and blushing. Blushing in particular conveys a number of emotions such as shame, arousal, and joy. Figure 1 illustrates several of these physical and emotional states. Despite their ability to transmit emotion, these dynamic changes in skin pigmentation are largely ignored by existing skin appearance models.

The creation of dynamic skin shading in film and game workflows depends mostly on artists, who carefully create all necessary skin textures. In the context of dynamic shading, an *appearance rig* is a structure that defines the details of the skin textures of 3D facial models. As models become more and more complex, it becomes increasingly difficult to define a consistent appearance rig that works well for many different characters; textures for each character must be created individually by hand, a slow and costly process that requires experienced digital artists. (Alternative facial animation techniques circumvent this difficulty by relying on performance capture [Sagar 2006; Bradley et al. 2010] to simultaneously obtain dynamic geometry and appearance, but they are not designed to derive a generic, transferable model.)

In this work, we develop a real-time, *dynamic* facial color appear-

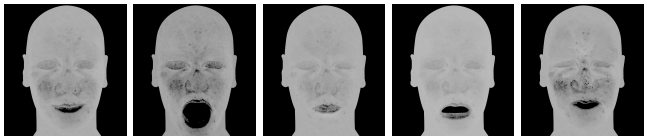


Figure 2: *Hemoglobin maps controlling the skin color in Figure 1. These maps are generated automatically by our method. Darker areas of the maps indicate higher concentrations of hemoglobin.*

ance model, using the common approximation of skin as a two-layered translucent material. Color appearance is defined by the distribution of two chromophores: melanin and hemoglobin. While this model is an approximation, it has proven to well describe a wide range of natural skin appearances [Donner and Jensen 2006; Donner et al. 2008]. We drive this model with *in vivo* measurements of melanin and hemoglobin concentrations for different facial expressions and conditions. The results of our measurements are highly-detailed hemoglobin and melanin maps for a variety of expressions, without the need for artist intervention.

We assume that the thickness of the skin epidermis is not affected by surface deformations during facial expression as it tightly adheres to the underlying deeper tissue [Ryan 1995]. Note that the facial melanin concentration and distribution is static within the skin during animation. This is because melanin is embedded within the keratinocyte cells of the epidermis, and the timeframe of change for melanin concentration is hours to weeks [Park et al. 2002], not the short periods associated with facial movement. This leaves the distribution of *hemoglobin* (the primary carrier of oxygen in the blood) as the only potentially varying factor that affects skin color. We also aim to separate *local phenomena* that change the perfusion of hemoglobin in a characteristic pattern from *global effects* that change the overall perfusion. Accurately modeling the dynamics of blood flow in the face would ultimately require a dynamic model of human blood circulation. A complete model, however, would be prohibitive, as it would require modeling not only the entire network of vessels and capillaries, but also the dynamic flow, defined by the interplay of fluid dynamics with vascular compliance and body mechanics.

To obtain a compact model of hemoglobin perfusion yielding realistic results even under real-time constraints, we build base melanin and hemoglobin maps. We then code only hemoglobin *changes* in our model by using a novel *localized* variant of the histogram matching technique. This allows us to treat the characteristic local patterns in hemoglobin distributions independently of the overall effect. Naive computation of this approach would involve prohibitive recalculation and storage of the cumulative distribution functions of the histograms, since it involves analyzing a window around each pixel. Instead, we make the key observation that typical hemoglobin concentrations resemble a Gaussian distribution, which allows us to store only its mean and standard deviation.

To render different expressions, we couple an efficient, yet realistic, skin shader with our appearance rig. This provides automatic pigmentation changes according to changes in facial expression. As the overhead of our model is low, these realistic effects come at essentially no additional cost. While our approach is general enough to complement most existing animation techniques, we focus on blend shapes consisting of the six universal facial emotions: anger, disgust, fear, happiness, sadness and surprise [Ekman 1972], plus the effects of physical exhaustion and alcohol consumption.

2 Related Work

Appearance models for faces have attracted much attention in recent years. In this section we focus on those techniques most closely related to our work. For a more thorough treatment we refer the reader to the survey by Igarashi et al. [2007].

2.1 Physical Appearance Models

The following techniques are all capable of reproducing skin realistically, and can be controlled heuristically to various degrees (e.g. parameter maps, scaling, etc.). None of these techniques, however, is tied to any real measure of mechanical deformation or physiological state of the skin itself.

Through independent component analysis, it is possible to extract hemoglobin and melanin pigmentation from a single skin image [Tsumura et al. 1999]. Adding a pyramid-based texture analysis/synthesis technique allows very realistic effects such as alcohol consumption and tanning [Tsumura et al. 2003]. Weyrich et al. [2006] analyze variations in the reflectance of facial skin under varying *external* conditions of a subject (hot, cold, sweaty, etc.); using a histogram interpolation technique [Matusik et al. 2005], they achieve color transfers and face changes between conditions.

Donner et al. [2008] simulate skin reflectance by accounting for lateral inter-scattering of light between skin layers. Using known chromophore spectra, they derive spatial chromophore distributions from multi-spectral photographs of skin through inverse rendering. Ghosh et al. [2008] use structured light and polarization to determine skin layer properties using an additive multi-layered scattering model. They capture the heterogeneous appearance of the face, and vary it through scaling of their model components.

The multi-image texture representation of skin presented by Cula et al. [2005] focuses on surface microgeometry, and takes into account appearance variations caused by changes in illumination and viewing direction. In the field of dermatology, Cula et al. [2004] use bidirectional imaging to create the Rutgers Skin Texture Database, focusing mainly on skin disorders, for medical applications.

In principle, mechanisms underlying facial perfusion changes have been studied in dermatology. Such reports are based on *ex vivo* histological examination of thin sections of biopsied tissue, or *in vivo* non-invasive *point measures* of blood derivatives (such as reflectance spectrophotometric measures of skin color or laser Doppler velocimetric measurements of blood flow) [Matts 2008]. The novelty of our work lies in the direct *in-vivo* mapping of hemoglobin concentration and distribution across *large areas* of the skin, and how changes correlate with dynamic facial expression. It is precisely this large-scale phenomenon that drives the subtle, yet noticeable, dynamic changes in facial skin color.

2.2 Emotional Appearance Models

Kalra and Magnenat-Thalmann [1994] describe a texture-based model to simulate skin changes from blushing and pallor. Regions in the face are separated by masks in the texture; the user interactively defines different shading functions within each mask, and linearly weights these masks to achieve different looks.

Jung et al. [2006] accumulate pre-designed 2D textures in a 3D stack, from palest to reddest. Each emotion correlates to a given depth, and the fetched texture is then used as the base color for a skin shader. In a follow-up publication [Jung et al. 2009], they parameterize skin changes via a set of fourteen emotional states [Plutchik 1980], along with a high-level description of each (such as *rosy cheeks* or *red blotches in the face*). Melo and Gratch [2009] forgo textures, directly applying a user-defined color change in different areas of the face to simulate blushing.

Yamada and Watanabe [2007] investigate blood flux due to anger and dislike. They simultaneously measure changes in facial skin

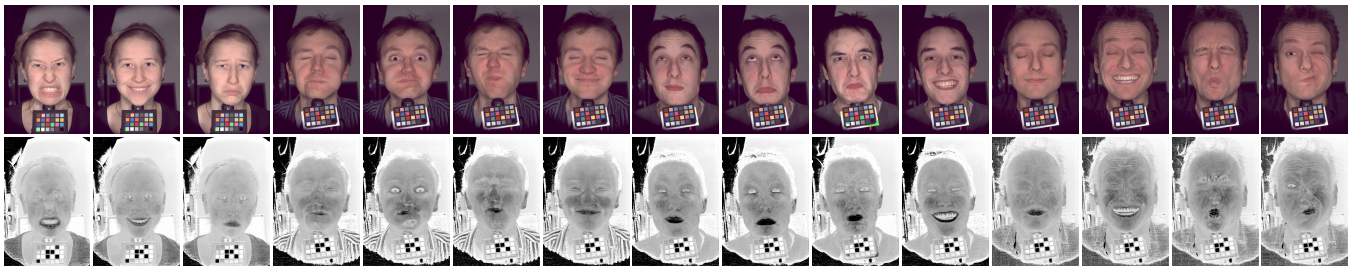


Figure 3: *Samples of images captured to analyze changes in blood perfusion (*top*), with the corresponding hemoglobin maps (*bottom*).*

temperature and color for these emotions, and map them to an average facial color image. Hue and saturation are multiplied by an arbitrary enhancement coefficient to make changes more distinct.

The results of the works discussed above are almost completely user-guided, with no correlation to actual hemoglobin measures. In contrast, we propose a compact, linear model of hemoglobin perfusion based on *in vivo* observations of facial performances, and under different global conditions. This makes our model well suited for real-time renderings of facial appearance.

3 Acquisition

Here we describe our acquisition setup, method, and observations, which will lead to the development of our model. We use a custom reconstruction of the non-contact SIAscope^{TM} system [Cotton et al. 1999], to capture accurate reconstructions of the hemoglobin and melanin distribution within the facial skin. The system uses a Fuji Finepix S2 Pro and two Portaflash 336VM flashes, with the flashes cross-polarized and positioned on each side of the camera. This setup is low-cost, and using the reconstruction method of Cotton et al. [1999] we obtain hemoglobin and melanin maps that, in contrast to other methods, are quantitatively calibrated [Matts et al. 2007].

This method is based on a spectral, multi-layered model of skin coloration [Cotton and Claridge 1996]. A limitation of this model is that it does not account for lateral scattering and thus over-estimates the blur of the underlying chromophore distribution. This feature has little impact on the analysis presented here but in fact allows a valuable optimization for our real-time renderer (see Section 6).

We acquired data from four subjects (one Caucasian female, 33 years old; three Caucasian males, 26, 33, and 35 years old) under a number of different, partially extreme facial expressions, as shown in Figures 3 and 4. To cover a representative range, we included expressions of the six basic emotions [Ekman 1972]. Each expression was acquired multiple times, with other varying expressions in between. In addition, we measured one subject under two "global" conditions: at high exercise level (after descending and climbing 9 flights of a staircase; subsequent measurements for one minute) and under alcohol consumption (500 ml beer, spread over one hour; regular measurements during this period).

3.1 Initial Findings

Throughout the experiments, the changes induced by a given expression were highly repeatable. As expected, melanin concentrations stayed constant across measurements of an individual. Hemoglobin, however, varied greatly. In studying the variations in blood flow, we made the following core observations:

1. Visible blood is mainly contained within the ascending and descending capillaries originating from the relatively superficial sub-papillary plexus. This becomes apparent when studying hemoglobin maps such as the high-resolution examples in Figure 4 and is consistent with the fact that visible light (400–700 nm) cannot penetrate deeply into the optically turbid dermal layers [Anderson and Parrish 1981].



Figure 4: *Facial hemoglobin distribution of a 33-year-old Caucasian female and of a 26-year-old Caucasian male. Dark pixels denote high concentration.*

2. The mechanical deformation of facial expressions may lead not only to drainage of blood in compressed regions, but also to a perfusion increase in other regions. This is most noticeable over the cheek bones, where, for instance, the blood concentration increases during a smile.

3. While the qualitative changes connected with a single expression correspond well across subjects, there are large differences in each individual's spatial pattern of perfusion.

4. Conversely, this subject-specific pattern appears to be very static and only subject to decrease or increase of local blood concentration.

5. Global effects cannot be simulated by simply globally increasing blood concentrations. Figure 5 (top row) plots the temporal perfusion variations due to high exercise level and to alcohol consumption at five representative facial positions, showing two distinctly different spatial distributions of perfusion increase.

A core question of our investigation was whether the expression-induced transitions between different perfusion levels take place at noticeable time scales. Our measurement device, however, does not offer the temporal resolution to measure such effects directly: the acquisition speed is limited by the camera read-out, which amounts to 5.5 seconds per image. To work around this issue, we used a time-multiplexing scheme often used to measure periodic motions (see, e.g., Yu et al. [2007]): we performed two acquisition sequences, where the subject was alternating between a *neutral* and a *smile* expression, synchronized by an acoustic signal of period p_e that prompted an expression change. At the same time, the camera would take an image every p_a seconds. The resulting images for this sequence ($p_e = 5$s, $p_a = 5.5$s) are shown in Figure 6. Reordering the acquired images as shown in the figure allows reconstruction

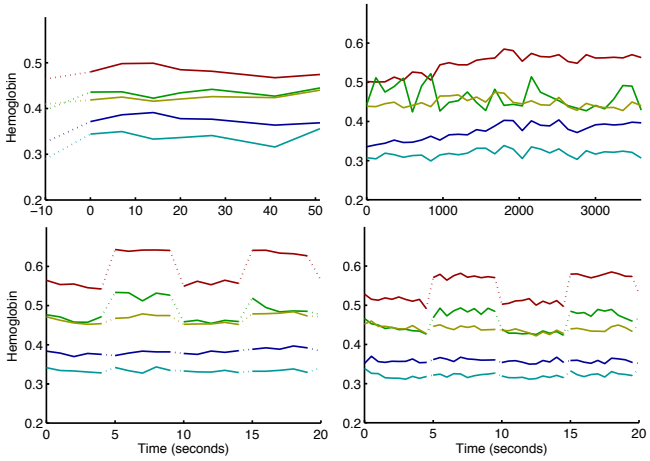


Figure 5: *Perfusion changes of five representative points denoted in Figure 7.* Top Row: *Two global factors (high exercise level and high alcohol level).* Bottom Row: *Concentration changes in the denoted areas during the temporal sequence shown in Figure 6.*

of an up-sampled version of the cyclic expression changes with an effective sampling rate of 0.5 seconds (see Figure 6). We later increased the sampling rate to ($p_e = 5$s, $p_a = 5.25$s), which effectively samples at 0.25-second intervals. Both rearranged sequences observe two full expression cycles, which results in 22 and 44 captures in total, respectively.

Figure 5 (bottom row) shows that despite the long duration of the experiment, the expression-induced blood variations were highly reproducible; our subjects were able to consistently produce the same expressions multiple times. In addition, neither sequence exhibits transitional phases between the neutral and the smile expression. This leads us to propose that:

6. Dynamic effects in the transition of expression-induced perfusion changes are of lesser visual importance and may hence be ignored in a practical model.

In the next section we use the above six observations to develop a simple, practical model for blood variation.

4 Color Appearance Model

The principal quantity controlled by our color appearance model is the spatially-varying concentration of hemoglobin. As any changes in blood concentration are bound to the location of vessels and capillaries (observations 1 and 4 from the previous section), we express perfusion variations by dynamically modifying a static, "neutral" hemoglobin texture map H_n. This map may be hand-painted, or acquired from real subjects. We use our measurements of different facial expressions and conditions to derive the *relative* changes that must to be applied to H_n to produce a desired appearance. Thus, our model is data-derived, not data-driven; we do not directly apply the measured maps when rendering.

Our approach is to use a localized variant of histogram matching (HM) to modify the neutral hemoglobin map H_n. To transform H_n to the blood distribution H of a joyful smile, for example, we must alter each pixel value in H_n, such that *locally*, within a radius r region, that pixel would contribute to a new distribution of concentrations which matches the histogram of the corresponding region in H. Traditional HM achieves this *globally* by determining the percentile q of each pixel's intensity in a source image's histogram, and mapping q to a new intensity that corresponds to the q^{th} percentile in a destination image's histogram. This involves looking up the cumulative distribution function (CDF) of the source histogram and



Figure 6: *Our time multiplexing scheme.* Top: *We sample the expressions at varying time offsets from the expression transition.* Bottom: *The images were taken in column-major order (as numbered), with a short phase shift between acquisition and expression changes. This gives a consistent 5-image temporal sequence of four expressions (neutral/smile/neutral/smile) when arranged in row-major order.*

the inverse CDF of the destination histogram [Heeger and Bergen 1995]. To localize this approach, we use histogram CDFs within a sliding window around each pixel in H_n. This would normally require recomputation (or storage) of two CDFs per pixel, which would be prohibitive for real-time applications.

4.1 Approximate Local Histogram Matching

In order to turn localized HM into a tractable problem, we leverage the observation that for a wide range of skin appearances, hemoglobin distribution is near-Gaussian (see Figure 7). This allows us to approximate hemoglobin histograms with Gaussian distributions, thus compressing each local histogram to a two-valued description of mean μ and standard deviation σ. For a given radius r, these descriptors are easily extracted for every pixel $\boldsymbol{x}$ in a hemoglobin map H, as:

$$\mu(\boldsymbol{x}) = \big(H * f\big)(\boldsymbol{x}) \,,$$
$$\sigma(\boldsymbol{x}) = \sqrt{\big(H^2 * f\big)(\boldsymbol{x}) - \mu(\boldsymbol{x})^2} \,, \qquad (1)$$

with the averaging convolution kernel $f(\boldsymbol{u}) = \frac{1}{\pi r^2}\big(\|\boldsymbol{u}\| \le r\big)$.

Adjusting a pixel's concentration $H_n(\boldsymbol{x})$ from a distribution $(\mu_n(\boldsymbol{x}), \sigma_n(\boldsymbol{x}))$ to a target distribution $(\mu(\boldsymbol{x}), \sigma(\boldsymbol{x}))$ now involves

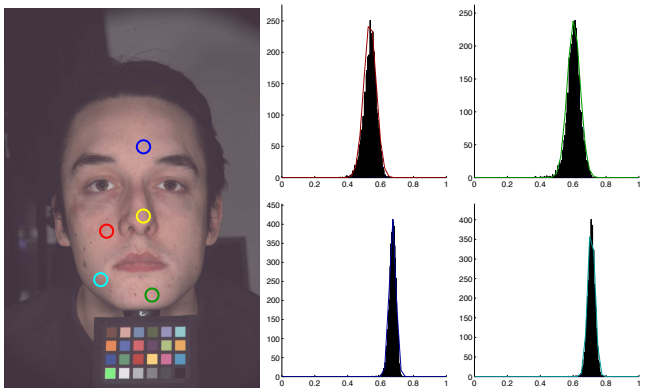


Figure 7: *Example hemoglobin distributions and their Gaussian approximation for representative areas of the face. (The origin of the histograms and their approximations are color-coded; yellow is referred to by Figure 5, its histogram is omitted for space reasons.)*

the CDFs of two Gaussian distribution:

$$H(\boldsymbol{x}) = CDF^{-1}_{\sigma(\boldsymbol{x}),\mu(\boldsymbol{x})}\left[CDF_{\sigma_n(\boldsymbol{x}),\mu_n(\boldsymbol{x})}\left(H_n(\boldsymbol{x})\right)\right] . \tag{2}$$

Due to the affine similarity of Gaussians, this mapping simplifies to a scale and bias:

$$H(\boldsymbol{x}) = \mu(\boldsymbol{x}) + \frac{\sigma(\boldsymbol{x})}{\sigma_n(\boldsymbol{x})}\left(H_n(\boldsymbol{x}) - \mu_n(\boldsymbol{x})\right) . \tag{3}$$

This simplification represents a physiologically plausible yet computationally efficient model to blend between perfusion distributions, yet maintaining the structural characteristics of the underlying vascular system. In the following subsection we show how to control this mechanism in the context of facial animation.

4.2 Model Parameters

Equipped with Equation (3), it is now possible to span an appearance space defined by a palette of target hemoglobin distributions H_i, $i = 1 \ldots k$. We define a mean-free base concentration:

$$H_0 = H_n - \mu_n , \tag{4}$$

and a set of scales and biases:

$$\begin{aligned} \text{bias}_0 &= \mu_n, \qquad \text{bias}_i = \mu_i, \quad i > 0 , \\ \text{scale}_0 &= 1, \qquad \text{scale}_i = \frac{\sigma_i}{\sigma_n} , \end{aligned} \tag{5}$$

the argument "$(\boldsymbol{x})$" being ommitted for readability. These allow the seamless transformation of the neutral hemoglobin distribution to any desired blend of texture characteristics of a set of basis distributions H_i, by applying an affine combination of scales and biases,

$$H = \sum_{i=0}^{k} w_i \text{bias}_i + H_0 \sum_{i=0}^{k} w_i \text{scale}_i , \tag{6}$$

where w_i are relative weights that sum up to one. The main benefit of this representation is storage efficiency, as the low-frequency nature of these biases and scales can be exploited (see Section 6). Following existing terminology in geometric animation, we refer to the set of w_i as parameters of an *appearance rig*.

In order to acquire a given subject's basis set of hemoglobin distributions, we gather the respective perfusion maps using our measurement setup and manually warp them to align with a photograph of the neutral expression. The complex deformations during expression changes make it prohibitive to align the acquired maps at pixel accuracy. This precludes simply interpolating between precaptured distributions. Instead, using Equation (1), we determine the statistics used to derive the parameters in (5). Thus, even when image alignments are not perfect, we still achieve good results.

The non-local support of the histograms implicitly leads to a natural separation of low-frequency changes from high-frequency details (as opposed to frequency-based approaches to separate global changes from local structure [Guo and Sim 2009]). This has two additional advantages: First, it becomes possible to apply relative changes derived from an existing subject's native hemoglobin map to other (for instance hand-painted) maps. This allows for *appearance transfer* between characters. Second, as the hemoglobin scale s_i and bias b_i textures store relatively low-frequency information (see Figure 8), we can downsample them to minimize their memory footprint, without significant loss of information. This makes our approach practical for real-time applications.

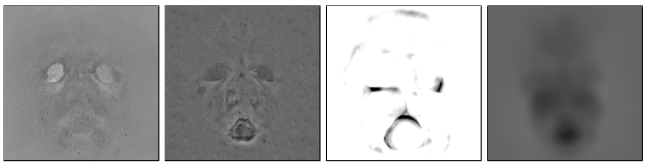


Figure 8: *Different maps used by our algorithm. The hemoglobin scale and bias maps shown here belong to the* smile *expression (Hemoglobin Scale shown at 90x scale for visualization purposes).*

5 Geometry-dependent Rigging

Our appearance model considers hemoglobin changes that are either tightly bound to the geometric deformations of the skin, or controlled by a global state that incorporates emotional and physiological parameters. Accordingly, we use a common rig for geometry and appearance control where these are correlated, and introduce independent parameters for the global state. As global effects are unique to facial expressions, we model both expression-related and global effects using a basis of hemoglobin statistics, as described in the previous section.

There are a number of different geometric animation approaches that either globally or locally control geometry [Ward 2004]. The three most common approaches create animation rigs based on blend shapes, bones, or a combination of both. Many of these approaches use either global or local weights of influence; our weight-based, linear model can be adapted to most existing techniques. While there are no intrinsic limitations to any of these approaches, our method lends itself particularly well to global blend shapes [Deng et al. 2006], an approach commonly used in production environments [Richie et al. 2005]. We demonstrate the applicability of our model in a production pipeline in Section 7, with an implementation in Autodesk® Maya® 2010.

6 Implementation

Our real-time facial color animation rig has four components (in execution order):

1. interpolation of facial shapes
2. obtaining hemoglobin change from the scale and bias textures
3. computing the base color of the skin
4. simulation of subsurface scattering in the skin

To compute the animation as efficiently as possible, we use graphics hardware and stream out blend shapes using DirectX® 10. This circumvents the limitation that only four blend shapes can be packed into per-vertex attributes at once. We store a set of transformed vertices into a buffer, which allows us to apply an unlimited number

of blend shapes using multiple passes [Lorach 2007]. Wrinkles are rendered using the recent approach of Jimenez et al. [2011], based on masking wrinkle zones and efficiently adding the influence of multiple normals coming from different zones using partial derivative normal maps.

We precompute the actual color of skin using the spectral model of Donner et al. [2008]. This model predicts spectral absorption σ_a and reduced scattering σ_s' coefficients based on the volume concentrations of hemoglobin in the dermis and epidermis, and the volume concentration and type of melanin in the epidermis. Table 1 summarizes these parameters to this model.

Parameter	Description	Range
C_{m}	Melanin fraction	$0 - 0.5$
β_{m}	Melanin type blend	$0 - 1$
C_{he}	Hemoglobin fraction (epi)	$0 - 0.1$
C_{hd}	Hemoglobin fraction (dermis)	$0 - 0.32$

Table 1: *Physiological parameters describing skin spectral absorption and scattering.*

For the simulation of subsurface scattering, we use the method by Jimenez et al. [2009; 2010]. We perform the sum-of-gaussians texture-space diffusion [d'Eon et al. 2007] as a postprocess in screen-space, by modifying the width of the convolution kernel according to depth gradient information. The method scales well with geometric detail, yet still retains high visual fidelity. Since lateral color bleeding due to subsurface scattering is already explicitly captured as part of the acquisition process, we use post-scattering texturing [d'Eon et al. 2007] to avoid blurring the albedo twice. We describe this in more detail below. As the screen-space approximation depends on pre-scattering blur, we separate the diffuse illumination, the albedo, and the specular reflectance into different render targets, and then selectively apply subsurface scattering only to the diffuse component. We then composite these components back together to produce the final image.

Recall that the parameter maps obtained as part of the acquisition process are blurred due to light scattering in the skin layers (see Section 3). This intrinsic blurring of the maps allows us to further optimize the rendering process: we use a precomputed skin color lookup table that is indexed by the local value in the melanin and hemoglobin parameter maps. Note that this is a departure from previous work, where these maps are used as inputs to a heterogeneous subsurface scattering simulation, as in [Donner et al. 2008]. Were we to use the maps in this fashion, there would be excessive blurring of the chromophore contributions.

Our skin color lookup table contains RGB color values which are pre-computed across the space of parameters shown in Table 1. We compute each entry using the total diffuse reflectance predicted by the two-layered translucent skin model, with $C_{\mathrm{he}} = 0.25 C_{\mathrm{hd}}$. As the skin model is spectral, and is highly non-linear with respect to the effects of chromophores on RGB color, we sample the space of melanin and hemoglobin cubicly to help maximize the use of the texture space. This allows us to use a smaller table, while still spanning the space of useful parameters. To find the index (u, v) in the space of the texture given a melanin and hemoglobin volume fraction, we apply:

$$u = \sqrt[3]{C_{\mathrm{m}}}\,,\; v = \sqrt[3]{C_{\mathrm{hd}}} \tag{7}$$

where melanin varies along the u axis, and hemoglobin varies along the v axis. Figure 9 shows this lookup table.

The hemoglobin scale s_i and bias b_i textures are downsampled to 256x256 resolution, as justified in Section 4. During rendering, we perform a standard subsurface scattering simulation using parameters for colorless skin. We then modulate the resulting value with the color from the lookup table, indexed by the melanin and hemoglobin values at each point. We use the specular model of Kelemen and Szirmay-Kalos [2001] with spatially-varying parameters for roughness and specular intensity [Weyrich et al. 2006].

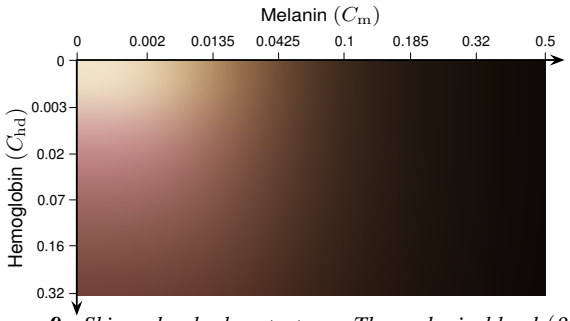


Figure 9: *Skin color lookup texture. The melanin blend (β_m) is 61% eumelanin and 39% pheomelanin, and epidermal hemoglobin (C_{he}) is 25% of dermal hemoglobin (C_{hd}).*

7 Results

Here we demonstrate our novel skin appearance rig under various expressions and conditions. All of the images in this section use hemoglobin maps of the 26-year-old male subject and were rendered in real-time. In many of our examples, the hemoglobin maps are inset into the lower-left corner of the images. Note that due to the delicacy of skin color and the widely-varying gamut of display devices we suggest viewing the PDF version of this paper on a color-calibrated monitor.

Figure 1 shows five emotional basis states, a sad smile, anger, the neutral pose, fear, and disgust. The hemoglobin maps generated by our rig are shown in Figure 2. Note the wide variation in appearance across the expressions. Some of the color is caused by blood perfusion due to deformation, while some is due to capillary dilation. In reality, extreme emotions like anger or fear trigger a very strong dilation or constriction of blood vessels. This causes significant involuntary blushing or pallor [Goldstein 2006], which is difficult to measure experimentally. With our model, these effects are easily simulated by scaling the global hemoglobin with respect to the neutral pose, as shown for the anger and fear poses. The global hemoglobin was scaled by 127% and 81% respectively in those cases, adding more expressiveness and realism to the images.

Figure 10 shows two additional physiological states, exercise and alcohol consumption, applied to a neutral pose. Although these changes may appear subtle, they increase realism and convey emotions, and free animators from the cumbersome tweaking of skin textures. Figure 13 shows close-up comparisons of the neutral and disgust poses, where the automatic changes predicted by our model are clearly shown. Our approach is general enough that any expression, emotion or state may be added, and allows for multiple combinations of poses (see Figure 11).



Figure 10: *The neutral geometric pose of a face rendered with a neutral hemoglobin map (left), and predicted hemoglobin perfusion after exercise (middle) and after alcohol consumption (right).*



Figure 11: *Example of blending poses: from left to right, the character transitions from the neutral pose to full anger (third image), to a combination of full anger and full surprise (fifth image). The last image adds the changes after physical exercise.*



Figure 12: *Various emotional states predicted by our model. The top row uses our method to dynamically predict the hemoglobin perfusion. The bottom row uses the neutral state hemoglobin map (best viewed on a calibrated monitor).*



Figure 13: *Close-up of the neutral and the disgust poses. The changes in hemoglobin distribution are generated by our model.*

As shown in this section, the encoded hemoglobin changes are directly transferrable to different characters. They can also be applied to a different neutral map, either captured from a real subject or created manually by an artist. We believe this feature can greatly improve the workflow in production environments, as it hides the details of facial color animation, but exposes useful controls. As a proof-of-concept integration, we have implemented it as a shader in mental ray® within Autodesk® Maya® 2010, and used it to render a small cartoonish animation (see video in the supplementary material). The artist has immediate feedback on the color variation of the skin, and can render final results off-line.

Finally, we make neutral hemoglobin and melanin maps available at ACM Portal, together with the skin color lookup table. The maps can be directly used for rendering or taken as a starting point to create different variations.

8 Conclusions

We have demonstrated an efficient, low-overhead skin appearance rig for predicting skin color changes during facial animation. Our dynamic model uses a novel local histogram matching technique, allowing efficient calculation of blood perfusion in skin layers. The model is anatomically motivated, and is based on statistical information gathered from *in vivo* measurements of hemoglobin perfusion. We have validated our model using these measurements, and shown potential applications, such as appearance transfer, and the synthesis of novel poses with realistic variations.

Due to the intrinsic chromophore blurring in our measurements, our rendering method adds minimal overhead, and lends itself to a real-time GPU implementation. While we have used blend shapes, our model integrates well with any standard geometric rig; we have demonstrated it in a pre-existing workflow pipeline of a common commercial software package (please refer to the video).

The current model focuses on skin color only. There exist other equally important aspects to skin appearance, such as wrinkles and pores. These, however, present geometric and topographic as well as rendering challenges. Coupling these to chromophore changes would be one direction for future work. Using a more standardized expression classification system, such as FACS [Sagar 2006], would simplify the use of our appearance rig in other fields, such as the behavioral and cognitive sciences. Our method could also be used in an inverse fashion, to detect emotions based on deformation and color cues.

One limitations lies in the use of the SIAscope algorithms. They are able to detect overall concentrations of melanin and hemoglobin, but not depth of deposition, or other features, such as hemoglobin oxygenation. We plan to develop a more robust system capable of providing more detailed and accurate physiological information.

9 Acknowledgements

We thank the reviewers for their keen insight and useful suggestions. Thanks also to our test subjects Gabriel Brostow and Kirstin Rautzenberg. This research has been partially funded by a Marie Curie grant from the Seventh Framework Programme (grant agreement no.: 251415), the Spanish Ministry of Science and Technology (TIN2010-21543) and the Gobierno de Aragón (projects OTRI 2009/0411 and CTPP05/09, plus a grant for Jorge Jimenez). We also thank XYZRGB Inc. for the high-quality head scan.

References

ANDERSON, R. R., AND PARRISH, J. A. 1981. The optics of human skin. *Investigative Dermatology 77*, 13–19.

BRADLEY, D., HEIDRICH, W., POPA, T., AND SHEFFER, A. 2010. High resolution passive facial performance capture. *ACM Trans. on Graphics (Proc. SIGGRAPH) 29*, 3.

COTTON, S. D., AND CLARIDGE, E. 1996. Developing a predictive model of human skin colouring. In *Proceedings of the SPIE Medical Imaging 1996*, vol. 2708, 814–825.

COTTON, S. D., CLARIDGE, E., AND HALL, P. N. 1999. 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 '99*, 49–52.

CULA, O., DANA, K., MURPHY, F., AND RAO, B. 2004. Bidirectional imaging and modeling of skin texture. *IEEE Trans. on Biomedical Engineering 51*, 12 (Dec.), 2148–2159.

CULA, O., DANA, K., MURPHY, F., AND RAO, B. 2005. Skin texture modeling. *International Journal of Computer Vision 62*, 1–2 (April–May), 97–119.

DENG, Z., CHIANG, P., FOX, P., AND NEUMANN, U. 2006. Animating blendshape faces by cross-mapping motion capture data. In *I3D '06: Proceedings of the 2006 symposium on Interactive 3D graphics and games*, ACM, New York, NY, USA, 43–48.

D'EON, E., LUEBKE, D., AND ENDERTON, E. 2007. Efficient rendering of human skin. In *Rendering Techniques (Proc. EGSR)*, 147–157.

DONNER, C., AND JENSEN, H. W. 2006. A spectral BSSRDF for shading human skin. In *Rendering Techniques (Proc. EGSR)*, 409–417.

DONNER, C., WEYRICH, T., D'EON, E., RAMAMOORTHI, R., AND RUSINKIEWICZ, S. 2008. A layered, heterogeneous reflectance model for acquiring and rendering human skin. *In Trans. on Graphics (Proc. SIGGRAPH Asia) 27*, 10:1–10:12.

EKMAN, P. 1972. Universal and cultural differences in facial expression of emotion. *Proc. Nebraska Symposium on Motivation 1971 19*, 207–282.

ERSOTELOS, N., AND DONG, F. 2008. Building highly realistic facial modelling and animation: a survey. *The Visual Computer 24*, 1, 13–30.

GHOSH, A., HAWKINS, T., PEERS, P., FREDERIKSEN, S., AND DEBEVEC, P. 2008. Practical modeling and acquisition of layered facial reflectance. *In Transactions on Graphics (Proc. SIGGRAPH Asia) 27*, 9:1–9:10.

GOLDSTEIN, D. 2006. *Adrenaline and the Inner World: an Introduction to Scientific Integrative Medicine*. The John Hopkins University Press.

GUO, D., AND SIM, T. 2009. Digital face makeup by example. In *IEEE Computer Vision and Pattern Recognition*, IEEE Computer Society, Los Alamitos, CA, USA, 73–79.

HEEGER, D. J., AND BERGEN, J. R. 1995. Pyramid-based texture analysis/synthesis. In *Proc. SIGGRAPH*, Computer Graphics Proceedings, Annual Conference Series, 229–238.

IGARASHI, T., NISHINO, K., AND NAYAR, S. K. 2007. The appearance of human skin: A survey. *Foundations and Trends in Computer Graphics and Vision 3*, 1, 1–95.

JIMENEZ, J., SUNDSTEDT, V., AND GUTIERREZ, D. 2009. Screen-space perceptual rendering of human skin. *ACM Trans. Appl. Percept. 6*, 4, 1–15.

JIMENEZ, J., WHELAN, D., SUNDSTEDT, V., AND GUTIERREZ, D. 2010. Real-time realistic skin translucency. *IEEE Computer Graphics & Applications 30*, 4, 50–59.

JIMENEZ, J., ECHEVARRIA, J. I., OAT, C., AND GUTIERREZ, D. 2011. *To appear in GPU Pro 2*. AK Peters Ltd., ch. Practical and Realistic Facial Wrinkles Animation.

JUNG, AND KNÖPFLE. 2006. Dynamic aspects of real-time face-rendering. In *Proc. ACM Symposium on Virtual Reality Software and Technology (VRST)*, 1–4.

JUNG, Y., WEBER, C., KEIL, J., AND FRANKE, T. 2009. Real-time rendering of skin changes caused by emotions. In *Proc. of the 9th International Conference on Intelligent Virtual Agents (IVA)*, Springer-Verlag, Berlin, Heidelberg, 504–505.

KALRA, P., AND MAGNENAT-THALMANN, N. 1994. Modelling of vascular expressions in facial animation. In *Proc. of Computer Animation*, 50–58.

KELEMEN, C., AND SZIRMAY-KALOS, L. 2001. A microfacet based coupled specular-matte brdf model with importance sampling. In *Eurographics '01 (short presentations)*, 1–11.

LORACH, T. 2007. Directx 10 blend shapes: Breaking the limits. In *GPU Gems 3*, H. Nguyen, Ed. Addison Wesley, ch. 3, 53–67.

MATTS, P. J., DYKES, P. J., AND MARKS, R. 2007. The distribution of melanin in skin determined in vivo. *British Journal of Dermatology 156*, 4, 620–628.

MATTS, P. J. 2008. New insights into skin appearance and measurement. *J. Investig. Dermatol. Symp. Proc. 13*, 6–9.

MATUSIK, W., ZWICKER, M., AND DURAND, F. 2005. Texture design using a simplicial complex of morphable textures. *ACM Trans. Graph. (Proc. SIGGRAPH) 24*, 3, 787–794.

MELO, C., AND GRATCH, J. 2009. Expression of emotions using wrinkles, blushing, sweating and tears. In *Intelligent Virtual Agents: 9th International Conference*, Springer, Ed., 188–200.

MORETTI, G., ELLIS, R., AND MESCON, H. 1959. Vascular patterns in the skin of the face. *The Journal of Investigative Dermatology 33* (Sep), 103–112.

PARK, S.-B., HUH, C.-H., CHOE, Y.-B., AND YOUN, J.-I. 2002. Time course of ultraviolet-induced skin reactions evaluated by two different reflectance spectrophotometers: Dermaspectrophotometer® and minolta spectrophotometer cm-2002®. *Photodermatology, Photoimmunology & Photomedicine 18*, 23–28.

PLUTCHIK, R. 1980. A general psychoevolutionary theory of emotion. *Emotion Theory, Research, And Experience 1*.

RICHIE, K., ALEXANDER, O., AND BIRI, K. 2005. *The Art of Rigging*. CG Toolkit.

RYAN, T. 1995. Mechanical resilience of skin: A function for blood supply and lymphatic drainage. *Clinics in Dermatology 13*, 5, 429–342.

SAGAR, M. 2006. Facial performance capture and expressive translation for king kong. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Courses*, ACM, New York, NY, USA, 26:1–26:1.

TSUMURA, N., HANEISHI, H., AND MIYAKE, Y. 1999. Independent-component analysis of skin color image. *J. Opt. Soc. Am. A 16*, 9, 2169–2176.

TSUMURA, N., OJIMA, N., SATO, K., SHIRAISHI, M., SHIMIZU, H., NABESHIMA, H., AKAZAKI, S., HORI, K., AND MIYAKE, Y. 2003. Image-based skin color and texture analysis synthesis by extracting hemoglobin and melanin information in the skin. *Trans. on Graphics (Proc. SIGGRAPH) 22*, 3, 770–779.

WARD, A. 2004. *Game Character Development with Maya*. New Riders Publishing.

WEYRICH, T., MATUSIK, W., PFISTER, H., BICKEL, B., DONNER, C., TU, C., MCANDLESS, J., LEE, J., NGAN, A., JENSEN, H. W., AND GROSS, M. 2006. Analysis of human faces using a measurement-based skin reflectance model. *Trans. on Graphics (Proc. SIGGRAPH) 25*, 1013–1024.

XU, Y., AND ALIAGA, D. G. 2007. Dense depth and color acquisition of repetitive motions. In *Proc. of International Conference on 3-D Digital Imaging and Modeling (3DIM)*, 141–148.

YAMADA, T., AND WATANABE, T. 2007. Virtual facial image synthesis with facial color enhancement and expression under emotional change ofanger. In *16th IEEE International Conference on Robot & Human Interactive Communication*, 49–54.