Imaging Flesh: Skin-customized Profiling

Mitchell R. Rosen*, Hongqin (Cathy) Zhang**, Robert Velthuizen and Qun (Sam) Sun*******
****Munsell Color Science Laboratory, RIT, Rochester, NY, USA***
*****KLA-Tencor, San Jose, CA, USA***
******Unilever R&D, Trumbull, CT, USA***
*******Micron, San Jose, CA, USA***

Abstract

Choice of camera calibration target influences how well an algorithm can estimate the color of specific scene objects. This research investigates the value of customized targets for estimation of flesh color. The use of a standard color calibration target is compared with a customized skin-based target. A spectral database of measured skin reflectances was available for use in this study. A skin target was simulated and compared to how well a standard MacBeth ColorChecker performed in calibrating a camera for accurate estimation of skin colors. The ColorChecker had a low average of 2 CIEDE2000 error units for the task. A simulated flesh target improved estimation error by 26%. For a second simulation a noise model derived from a commercial clinical photography system was imposed. In this more realistic case, the skin-customized target improved results by a far larger 60%.

Skin Reproduction

In the field of color reproduction, a very common subject is people. Human flesh is often a significant component of scenes. Reproducing skin faithfully is a key requirement of many reproduction systems. If familiar with a person being photographed, an observer can be highly critical of the reproduction accuracy. If not familiar with the subject, an observer will reject an image if skin color is deemed implausible. Skin is a memory color to which humans are highly sensitive even without ever having seen the person who was photographed.

In 1976 McCamy and co-workers published the specifications that became the basis for the widely used camera characterization target, now known as the standard 24-patch ColorChecker[1]. The article described the target as follows: *The 4×6 array of patches … includes spectral simulations of light and dark human skin, foliage, blue sky and a blue flower (chicory). Additive and subtractive primaries and a six-step neutral scale are included for analytical studies and other colours fill a wide gamut.*

It is noted that the first patches mentioned in the description were the spectral simulations of light and dark skin because of their considered importance. But, it is clear that the ColorChecker was appropriately designed for a wide gamut of colors, not just flesh colors. A flower, blue sky, foliage, additive and subtractive primaries and a gray scale all are included to allow the target to stand in for a variety of important colors that would be commonly encountered in typical scenes.

Previously, Rosen and Berns demonstrated the value of having a camera profile built from targets consisting of the same type of material as that being photographed[2]. In that experiment it was shown that patches printed on ink-jet were better served by profiling targets printed on ink-jet than by the traditional ColorChecker target, even though the colorimetry of the ink-jet targets matched the ColorChecker very closely.

There are certain reproduction tasks that are significantly limited or exclusively limited to flesh colors. These would include certain medical and cosmetics applications. In such cases, it is fair to imagine that a ColorChecker-like target with its wide gamut of colors may cause calibration trade-offs that do not necessary result in optimal flesh reproduction. The ultimate target for a skin-only reproduction task should include a large set of patches with the same spectral reflectances as human flesh.

Simulation with Low Noise

An experimental simulation was designed to investigate how much advantage a skin-customized camera calibration target would give to the highly constrained task of only reproducing skin. Such a target should include many patches with spectral reflectances similar to flesh, much like the light and dark skin patches on the standard ColorChecker. A database of skin measurements and photographs, the Sun Skin Database[3], was available for this project. Samples in the database were collected through careful spectroradiometric measurement and photography under the same illumination conditions.

The Sun Skin Database contains multiple skin measurements taken from 34 faces. The population that was measured and photographed for the study came from a variety of ethnic backgrounds. See Table I. RGB photographs of the skin were made under the measurement conditions. Also, a standard ColorChecker target was measured and photographed under the same conditions.

Table I: Sun Skin Database

Ethnicity	Number of participants
Asian	11
White	8
Black	7
Indian (subcontinent)	6
Hispanic	2

A comparison was made between how well skin colorimetry was estimated from a profile made from a simulated target consisting of skin samples from the Sun database and a profile from a standard ColorChecker photographed under the same conditions.

The skin spectral measurement database contained spectral reflectances in the range of 400nm - 700nm for multiple spots on the face for each of the participants. A well-characterized low-noise camera captured the corresponding RGB data, which was subsequently linearized. The database also contained spectral reflectance and linear RGB data for irises and hair of the subjects as well as a standard ColorChecker. The iris and hair data were not used for this investigation.

The spectral reflectance data were integrated with human color matching functions and the spectral power distribution of D50 illumination to calculate CIEXYZ and CIELAB under D50. To transform camera linear RGB to corresponding CIEXYZ, a camera profile was built in MATLAB. The RGB data were previously linearized so a simple 3 by 3 matrix carried out the conversion.

Two profiles were made; one was based on the ColorChecker data and the other one based on a training set consisting of one half of the skin data from the database. The two profiles were then tested on the other, independent, half of the skin data. The prediction errors of each profile on each test data set are shown in the Table II in terms of average color difference CIEDE2000 between the profile predicted color values and the measured values.

Table II: Performance of ColorChecker and Skin-customized Profiles From Sun Skin Database

Photographed target	Profile made from ColorChecker	Profile made from Sun Skin Database training patches
ColorChecker	1.8	4.3
Sun Skin Database independent samples	1.9	1.4

All values in CIEDE2000 units averaged over the test patches

It is not surprising that the profile made from the skin-only samples was poor at predicting the color of the ColorChecker (an average of 4.3 CIEDE2000 units). Also not surprising is the fact that the profile made from skin samples is best at predicting the independent skin samples.

The ColorChecker did relatively well at estimating the flesh colorimetry. On average the ColorChecker profile was off by only 1.9 CIEDE2000. On the other hand, the skin color prediction average for the profile made from imaging the skin-customized calibration target had a 26% improvement bringing the CIEDE2000 average error down to 1.4 units.

Although the profiles made from the skin were not good predictors of the ColorChecker colorimetry, the performance of the profile made from the standard ColorChecker data was very close in absolute terms to the quality of the profile generated from the skin data for predicting the colorimetry of the independent skin test data.

Figure 1 illustrates the histogram of color error for the skin samples through the profile produced from the skin-customized target. The majority of samples are less than 1 CIEDE2000 and almost all are less than 3 units. Figure 2 illustrates the same samples when transformed through the ColorChecker target profile. Although there is a slight shift toward higher color error, the shapes of the two histograms are quite similar.

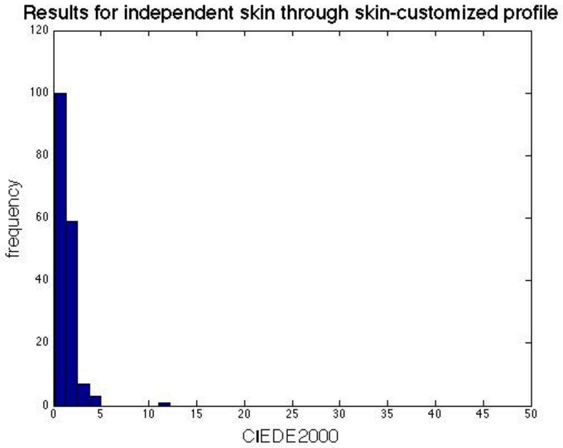

Figure 1: Histogram for Sun Skin Database independent samples through the profile produced from Sun Skin Database training patches.

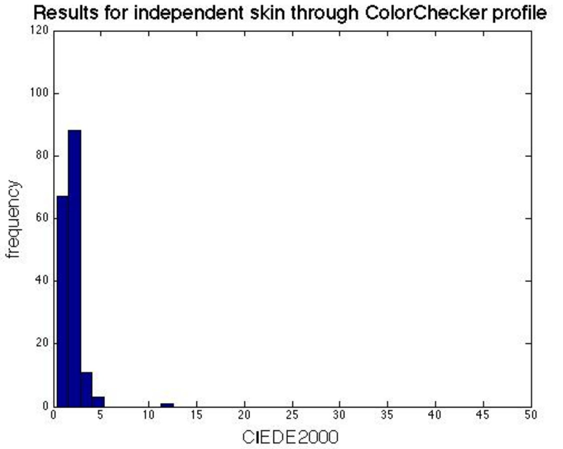

Figure 2: Histogram for Sun Skin Database independent samples through the profile produced from the ColorChecker.

Simulation with Typical Noise

The experimenters were interested in how this difference in color prediction changed as noise from a typical clinical photography system was introduced. An imaging environment, Visia-CR, advertised as "high quality, reproducible facial imaging for clinical research"[4] was selected. This imaging module is designed for facial photography. It has head support and built-in lighting.

Over a period of three and a half months, ColorChecker SG targets and uniform gray targets were regularly imaged in the Visia-CR. The gray targets allowed for flat-fielding of the images. The camera-delivered RGB values were normalized to between 0 and 1. A linearization function was optimized in the form of:

$$RGB' = 1.1RGB^{2.6} \tag{1}$$

A Gaussian noise function was derived from the linearized data. It was highly stable throughout the data collection. The noise was seen to have a mean of 0.1% and a variance of 0.09%. Noise sources included both lighting and camera response variations.

In MATLAB, the linear RGB values from the Sun Skin Database were modified by applying normally distributed random offsets that exhibited the mean and variance of the clinical setting. With these augmented RGB values, the previous experiment was repeated. Again, two profiles were produced, one based on the simulated noisy ColorChecker data and the other on half of the noisy skin data. Table III shows the performance when passing the noisy ColorChecker data and the noisy independent skin data through the two profiles.

Table III: Performance of ColorChecker and Skin-customized Profiles From Sun Skin Database with Noise-augmentation

Photographed target	Profile made from ColorChecker with noise	Profile made from noisy Sun Skin Database training patches with noise
ColorChecker with noise	6.4	17.2
Sun Skin Database independent samples with noise	8.5	3.4

All values in CIEDE2000 units averaged over the test patches

In this case, the profile made from the ColorChecker was no longer relatively good at estimating the flesh colorimetry. On average the ColorChecker profile was now off by 8.5 CIEDE2000 units. The skin color prediction average for the ColorChecker-produced profile is two and a half times worse than the 3.4 CIEDE2000 average error that the skin-customized calibration target produced.

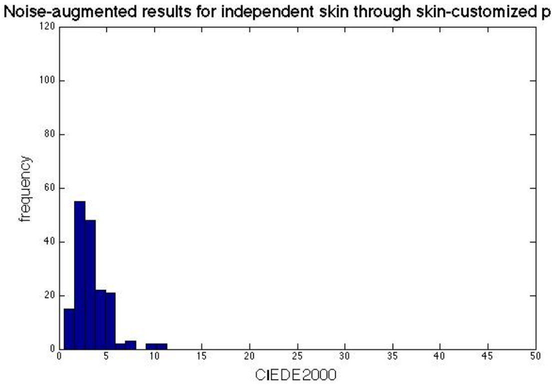

Figure 3: Histogram for noise-augmented Sun Skin Database independent samples through the profile produced from noise-augmented Sun Skin Database training patches.

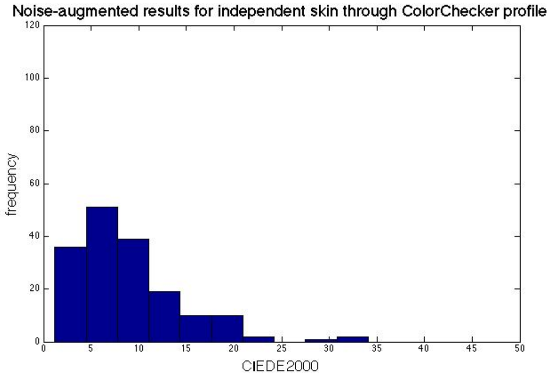

Figure 4: Histogram for noise-augmented Sun Skin Database independent samples through the profile produced from the noise-augmented ColorChecker.

Figure 3 shows the histogram of CIEDE2000 for independent skin samples transformed through the skin-customized profile. The shift from Figure 1 to Figure 3 is quite significant. This shows the impact of noise on the making of skin-customized profiles.

What is far more dramatic, though, is the tremendous shift toward high error exemplified between Figure 3 and Figure 4. With noise in the simulation, the ColorChecker-produced profile is no longer able to predict skin-colors with reasonable accuracy.

Typical noise has clearly differentiated among the standard and customized profiling approaches.

Discussion

In previous work, Rosen and Berns[2] showed that color patch targets produce profiles that perform better when the targets and scene objects share similar spectral characteristics. The results from the current study are compatible with those earlier findings. Here we find that a skin-customized target does a superior job at producing a profile that well predicts colorimetry of independent skin samples than a ColorChecker produced profile.

It might be straightforward to predict that a profiling target made of spectral components similar to the objects being imaged should out-perform a standard target. The question then becomes, how much better will it do? For the first simulation with low noise reported above, it was surprising how little difference in absolute terms there was when switching between the ColorChecker target and the skin-customized target. An average improvement of half a CIEDE2000 might not be significant enough for more than the most demanding application.

The second simulation reported here brings out the power of using the appropriate target. By adding typical acquisition noise to the simulation, the difference between the profiling results derived from a standard target and a customized target become far clearer. Applying to the skin capture database a noise model derived from analysis of a clinical research facial capture apparatus, the use of a profiling target with patches spectrally similar to the objects being imaged is shown to be very important.

For the noisy case, predicting the colorimetry of skin using a profile produced by a ColorChecker has an average CIEDE2000 of 8.5. There is a 60% improvement when building a profile from patches of skin reflectances. The average CIEDE2000 error falls to 3.4 units. An application would not require extremely tight tolerances to enjoy the benefits of the customized target.

Conclusions

It has been shown that color reproduction errors increase as camera calibration targets become spectrally different from the materials that are to be imaged. A hypothesis followed from this that for reproduction of skin colors, a target made up of skin reflectances should improve reproduction over the use of a standard target. This proved to be true, although to only 0.5 CIEDE2000 on average. The standard target did as well as it did because its design consciously includes flesh even while it is designed to do well on a wide gamut of colors. For ultimate flesh performance, though, it was shown that a skin-customized target was better because the standard target was shown to do on average 36% worse.

In a capture situation where typical noise is present, the advantages to profiling using a target made up of skin reflectances was far more pronounced. The average deficit for using the ColorChecker for predicting skin was 250% of the CIEDE2000 found when using a profile made from skin-like patches. Unfortunately, the converse is also true, that the profile made from a customized target can be a very poor predictor of colors that are not spectrally similar to skin. In the noisy simulation, the average CIEDE2000 for predicting the colors of the ColorChecker from a profile produced from skin colors was a whopping 17.2 units.

In many different situations, skin-customized targets should improve the results for producing profiles for skin color reproduction. When predicting non-skin colors is of very little value, the use of customized target should be considered, as the improvement for the colors of interest is significant.

Acknowledgements

Portions of this research were supported by Unilever Research and Development. The authors would like to acknowledge Jomer Delacruz of Unilever who did much of the data gathering for producing the noise model in the second simulation.

References

1. C.S. McCamy, H. Marcus, J.G. Davidson, A Color-Rendition Chart, J. Appl. Phot. Eng., 2, pp. 95-99 (1976).
2. MR Rosen and RS Berns, Spectral Reproduction Research for Museums at the Munsell Color Science Laboratory, Proc. NIP, pp. 73 - 77 (2005).
3. Q. Sun, Spectral imaging of human portraits and image quality, Ph.D. Dissertation, R.I.T., Rochester (2003).
4. Canfield Imaging Systems, "Visia-CR", www.canfieldsci.com/FileLibrary/BRCH-VISIA-CR_0408-29.pdf

Primary Author Biography

Mitchell Rosen is a Research Professor in RIT's Munsell Color Science Laboratory and directs the Infinite Pixel Liberation Laboratory (iPixLab), a teaching and research laboratory dedicated to the emerging concept of liberated media. Such content depends upon non-traditional display formats such as immersive, transient, wide-view or environmentally-reactive projection. Rosen's published research includes the areas of color management, spectral imaging systems, museum imaging and eye movement analysis. He teaches graduate courses on color systems and tutorials on color management, color reproduction and spectral imaging. He was Color Imaging Editor of IS&T's Journal of Imaging Science and Technology for 5 years. Rosen is active in organizing international conferences on spectral imaging. His website is http://www.cis.rit.edu/rosen and he can be reached at rosen@cis.rit.edu.

.