Non-Contact Colorimeter for Human Tooth Color Assessment Using A Digital Camera

Du-Yong Ng , Jan P. Allebach,
† Mostofa Analoui and †† Zygmunt Pizlo
School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907-1285
† Pfizer, Inc (USA)
†† Department of Psychological Sciences
Purdue University
West Lafayette, IN 47907-1364

Abstract

Wu et al developed model-based and regression-based colorimeters in their paper entitled *Imaging colorimetry using a digital camera*[1]. They found that the regression-based system worked better than the model-based system. In this work, we modify their regression-based model slightly to improve the system performance. We also present another regression-based model that is capable of recovering the spectral reflectance of teeth. Both models are realized and evaluated clinically.

1. Introduction

The color of teeth probably constitutes one of the most important parts of our first impression of someone. In the past, very little could be done if we did not like the color of our teeth. However, modern technology enables us to change how they look. Tooth whitening chemicals such as carbamide peroxide gel are now available [2]. Besides tooth bleaching, tooth restoration is also gaining popularity in dentistry. Patients will often expect their teeth to be restored to good condition with colors consistent with neighboring teeth. Consequently, the need for an accurate measure of tooth color has become more pressing.

Traditionally, dentists determine the color of a tooth by comparing it to color standards such as a shade guide to obtain a visual match. However, this method is not consistent and does not translate to CIE standard specifications. More accurate and consistent measurements can be made with instruments such as spectroradiometers and spectrophotometers. These devices measure radiometric quantities [3] or ratio of these quantities such as spectral reflectance and transmittance factors [3] and hence are accurate. A tristimulus-filter colorimeter, on the other hand, provides only the tristimulus values of a color [3]. It can be implemented with low cost devices such as a scanner [4] or a digital camera [5]. However, the color spaces of these devices are normally device dependent. Thus, reliable and accurate measurement of the tristimulus values of a color sample is not a trivial task.

Many studies have been done to relate the RGB color space of a device such as a digital camera to the CIE XYZ color space. Two major approaches have been introduced in the past, one from the physics perspective of the device [6,7] and the other from a statistical point of view of the measured data [8,9]. Wu et al [1] attempted both approaches, and found that the statistical approach yielded better results. The error of the measurements was colorimetrically lower.

The most distinct difference between the work of Wu and that of others is in the material property of the color samples used for the calibration. The color samples used by other authors were essentially matte material, which is specifically made to minimize specular reflection in the measured color signal. Wu's samples were human teeth, which have many undesirable properties (from the perspective of colorimetry) such as a smooth surface (introduces specular reflection), an irregular pattern of scattering, unpredictable size and structure, and flourescence. All these properties make calibration even more sensitive to such factors as the measurement setup. In general, the calibration matrix derived from a set of samples used in the calibration process may not work as well for a set of samples that has a different color range. Thus, for a very specific application like Wu's tooth color assessment colorimeter, it is more appropriate to use samples of human teeth to find the calibration matrix. However, it is not clear how the performance of the colorimeter designed using human teeth can be effectively compared to those designed using matte material.

Researchers have studied dental discoloration and bleaching using photometers and spectroradiometers [2, 10, 11]. These devices are expensive, and take only spot measurements. The colorimeter designed by Wu et al overcame these disadvantages. However, Wu's system is not flexible enough to be used in clinical situations because the daylight simulator is not suitable for taking measurements from patients. In addition, the measurement setup can be further improved to minimize specular reflections and errors arising from inconsistent alignment of the spectroradiometer and camera during the calibration process. Finally, the tristimulus values predicted by Wu's system are only valid under the illuminant which was used in the calibration process.

Our current work is a continuation of Wu's work using the statistical approach. We modify the measurement setup and make a small change to the original model to improve accuracy. We also design a system that can be used in clinical situations. A regression-based model that is capable of recovering the spectral reflectance from the camera RGB values is also introduced. The color of a tooth can then be predicted under any arbitrarily chosen illuminant. Imai et al applied a similar technique to human skin [12] and the Macbeth color checker [13]. Finally, we report for the first time the results of tooth color measurements made with human subjects.

This paper is organized as followed: Sec. 2 describes the theoretical derivation of the models. The experimental setup for the measurements is described in Sec. 3. The experimental results will be tabulated in the Sec. 4. Discussion and conclusion are presented subsequently. The notation used in the following sections is consistent with that of Wu et al [1].

2. Regression-based Colorimeters

In this section, we describe the design of our multi-exposure single-illuminant colorimeters in three separate subsections. In the first subsection, the regression framework that is common to both models are described. The first and the second models differs in the form of data used in computing the model parameters. In the first model, we use the camera data and the corresponding tristimulus values to obtain the model parameters. In the second model, the tristimulus values are replaced with the coordinates of teeth in the eigenspace of the spectral reflectance. Thus, we describe the unique characteristics of each model separately in the other two subsections.

2.1. Common Regression Framework

A typical color digital camera has three measurement channels; red (R), green (G) and blue (B). External filters are used with this camera to increase the number of measurement channels. The resulting combination is a multi-exposure system. For each sample, one exposure will be taken for every filter. Let N_f be the number of filters in the system and N_r be the total number of samples. Let S be the stack of outputs of the camera; that is

$$S = \begin{bmatrix} R_1^1 & G_1^1 & B_1^1 & \dots & R_1^{N_f} & G_1^{N_f} & B_1^{N_f} & 1 \\ R_2^1 & G_2^1 & B_2^1 & \dots & R_2^{N_f} & G_2^{N_f} & B_2^{N_f} & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots \\ R_{N_r}^1 & G_{N_r}^1 & B_{N_r}^1 & \dots & R_{N_r}^{N_f} & G_{N_r}^{N_f} & B_{N_r}^{N_f} & 1 \end{bmatrix}.$$

Instead of the linear homogeneous model introduced in [1], we use an affine model. Let T and T' (both are $N_r \times p$ matrices, where p denotes the total number of columns) be the true and predicted outputs of the system, respectively. Suppose that M, an $(3N_f + 1) \times p$ matrix, is a linear mapping that maps S to T', then S is related to T' through the equation

$$T' = SM; \tag{1}$$

and M can be obtained such that the total least squared error between the true and the predicted outputs is minimized, i.e

$$\hat{M} = \arg\min_{M} \|T - T'\|_F^2, \tag{2}$$

where F denotes the Frobenious norm. The least square solution is given by

$$\hat{M} = (S^T S)^{-1} S^T T. \tag{3}$$

Since both T and S depend on illumination conditions, the model obtained will only predict correctly under the same illumination conditions. Thus, in each of our models, the calibration matrix is illuminant dependent. However, there is a difference in the nature of the outputs. The first model, our reference model, derived by least square regression in the CIE XYZ color space, predicts outputs that can only be matched under the same illuminant. The calibration matrix of our second model is obtained in a similar way but in the eigenspace of the tooth spectral reflectance. Consequently, the output of the second model is the coordinates of the spectral reflectance of the tooth. Thus, the spectral reflectance can be reconstructed and the color of the tooth can be estimated under any other illuminants. From this point onward, we will refer to the first and second models as the illuminant-dependent (IDC) and illuminant-independent colorimeters (IIC), respectively.

2.2. Illuminant-Dependent Colorimeter (IDC)

The IDC and IIC differ in the content of matrix T. With the IDC, T consists of the stack of CIE XYZ values of the tooth samples. Let r_i, an 31×1 vector, denote the spectral reflectance of the i^{th} sample. Let $A = \begin{bmatrix} \bar{x} & \bar{y} & \bar{z} \end{bmatrix}$ where $\bar{x}$, $\bar{y}$, and $\bar{z}$ are the CIE 1931 2 degree color matching function vectors. The CIE XYZ tristimulus values of the i^{th} sample can be calculated as follows:

$$\begin{bmatrix} X_i & Y_i & Z_i \end{bmatrix} = r_i^T LA. \quad (4)$$

where L is a 31 × 31 diagonal illuminant matrix. In Wu's regression model [1], the camera data and the spectral reflectance was collected under daylight. Then, the T matrix was calculated with a known L and Eqs. 4 and 5. Several different illuminants were used in his work and each produced a different T matrix. Then for each of the $\{S, T\}$ pairs of matrices, the corresponding M matrix was calculated using Eq. 3. The M matrix associated with each pair of $\{S, T\}$ matrices can only be used for prediction provided two conditions are met. Firstly, the illuminant under which the RGB data are obtained must match that of matrix S. Secondly, the predicted results, namely the XYZ values, can only be matched under the illuminant that was used to calculate the T matrix. In our model, we only consider the case where both the T and S matrices are obtained under the same illuminant. We interpret this condition as "calibrate, measure, and match under the same illuminant". For this model, T is defined as

$$T = \begin{bmatrix} X_1 & Y_1 & Z_1 \\ X_2 & Y2 & Z_2 \\ \vdots & \vdots & \vdots \\ X_{N_r} & Y_{N_r} & Y_{N_r} \end{bmatrix}. \quad (5)$$

2.3. Illuminant-Independent Colorimeter (IIC)

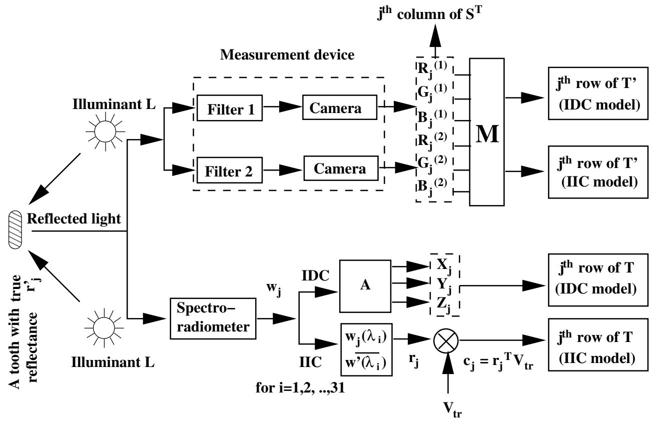
The T matrix for this model consists of the stack of the coordinates of the teeth in the eigenspace of the tooth reflectances:

$$T = \begin{bmatrix} c_1^1 & c_1^2 & \dots & c_1^{N_e} \\ c_2^1 & c_2^2 & \dots & c_2^{N_e} \\ \vdots & \vdots & \dots & \vdots \\ c_{N_r}^1 & c_{N_r}^2 & \dots & c_{N_r}^{N_e} \end{bmatrix}, \quad (6)$$

where N_e is the number of principal eigenvectors selected and c_i^j denotes the coordinates of the i^{th} sample along the j^{th} principal axis. These coordinates are obtained by principal component analysis (PCA). Conventional PCA can be done by performing spectral decomposition of the covariance matrix of the data [14]. Nevertheless, we use the autocorrelation matrix instead because the mean of the data set is not required.

Define autocorrelation matrix $\hat{R}_{SS} = E[S^T S]$, where E[.] denotes the expectation operation. In practice, since the true probability distribution is unknown, $\hat{R}_{SS}$ is estimated using the following formula:

$$\hat{R}_{ss} = \frac{1}{N_r - 1} \sum_{i=1}^{N_r} r_i r_i^T. \quad (7)$$



c_j : coordinates of r_j in the eigenspace of the training samples
λ_i : ith component of the 31–point wavelength vector
r_j : 31×1 estimated spectral reflectance
D : 31×3 camera sensitivity matrix
M : calibration matrix, 7×3 for IDC model and 7×6 for IIC model
A : 31×3 CIE XYZ color matching functions
w_j : light reflected from the tooth measured by the spectroradiometer
w' : light reflected from a standard white measured by the spectroradiometer
T' : predicted output matrix
T : true output matrix
IDC : Illuminant Dependent Colorimeter
IIC : Illuminant Independent Colorimeter
S : as defined in Sec. 2.1
V_{tr} : as defined in Sec. 2.3
$(.)^T$: matrix transposition

Figure 1: Summary of Illuminant-Dependent and Illuminant-Independent Colorimeters.

The autocorrelation matrix $\hat{R}_{ss}$ is then decomposed using singular value decomposition such that $\hat{R}_{ss} = V\Lambda^{\frac{1}{2}}U$ where V and U are the left and the right eigenvector matrices. The matrix Λ is a diagonal matrix which contains the singular values associated with each pair of left and right eigenvectors. The elements of Λ are arranged in descending order; and the first N_e left eigenvectors associated with the N_e largest singular values are selected. The parameter N_e is picked so that this number of eigenvectors produces a good reconstruction of all the r_i's. Let $V_{tr} = \begin{bmatrix} v_1 & v_2 & \dots & v_{N_e} \end{bmatrix}$ be the matrix containing the N_e most significant left eigenvectors, then

$$c_i^j = r_i^T v_j. \quad (8)$$

The matrix $\hat{M}$ for the model is computed from Eq. 3. The i^{th} reflectance is reconstructed as

$$\hat{r}_i = \sum_{j=1}^{N_e} c_i^j v_j. \quad (9)$$

In this model, we calibrate and measure under the same illuminant but match under any arbitrarily chosen illuminant. Figure 1 summarizes both of the models described in this section.

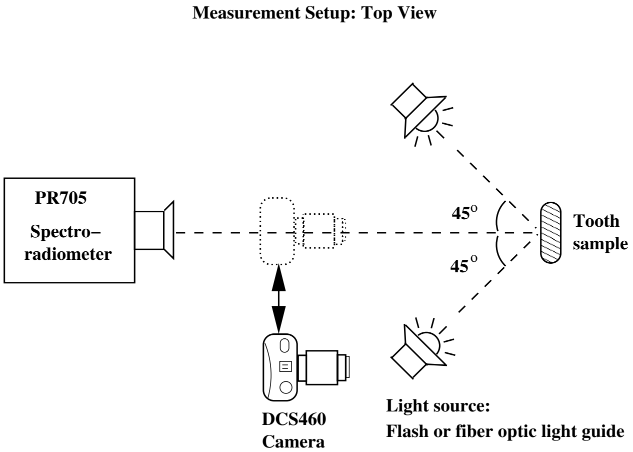


Figure 2: *Measurement setup for both models.*

3. Experiment

A Kodak DCS460 color digital camera, equipped with Wratten filter sets WR11+WR85N6 and WR38A+WR80B, was used as in the Wu et al setup [1]. Two exposures, one for each filter set, were taken for each sample. This produces 6 different channels of measurements for each sample. We chose 6 channels because a conventional PCA revealed that six eigenvectors were sufficient to account for more than 99% of variance of the tooth samples. A Photo Research PR705 spectroradiometer[1] was used to measure the radiance of the samples. Individually, we tried three different light sources in the experiment: a MacBeth daylight simulator[2], two Vivitar 285 camera flash units (AC flash units)[3], and a PL900-B DC regulated Halogen lamp with a bifurcated fiber optic light guide[4]. The measurement setups for the flash-based and halogen-based systems are as shown in Fig. 2. When measurements were performed with daylight simulator, the tooth samples were placed in the viewing booth. All the measurements were performed in a dark room. In the first experiment, 85 extracted teeth were measured. Each of them was placed in a holder when the measurement was performed. All the teeth were prepared as described by Papakosta[15]. In the second experiment, the holder was replaced with a chin rest. A total of 21 subjects were recruited. Two exposures of the top incisors of each subject were taken followed by the spectroradiometer measurements. Each subject put on a disposable cheek retractor during the experiment.

[1] Photo Research, Inc, 9731 Topanga Canyon Place, Chatsworth, CA, 91311-4135

[2] GretagMacbeth, 617 Little Britain Road, New Windsor, NY 12553

[3] Vivitar Corporation, Santa Monica, CA 90406

[4] Dolan-Jenner Industries Inc., 678 Andover Street, Laurence, MA 01843-1060

Table 1: *Experiment 1 (IDC): Extracted Teeth Measured With 3 Different Light Sources*

Illuminant	ΔE_{avg}			
	Mean	StdDev	Max	% < 1.5*
D65	1.101	0.169	1.885	98.5
Flash	0.965	0.198	1.775	99.4
Halogen	0.843	0.115	1.362	100

* percentage of random selections with average $\Delta E_{a*b*} < 1.5$.

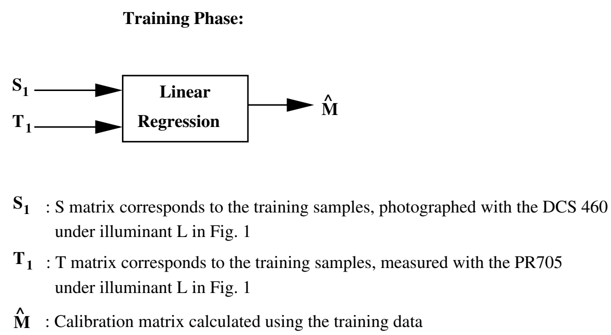


Figure 3: *Training phase of Illuminant-Independent Colorimeter.*

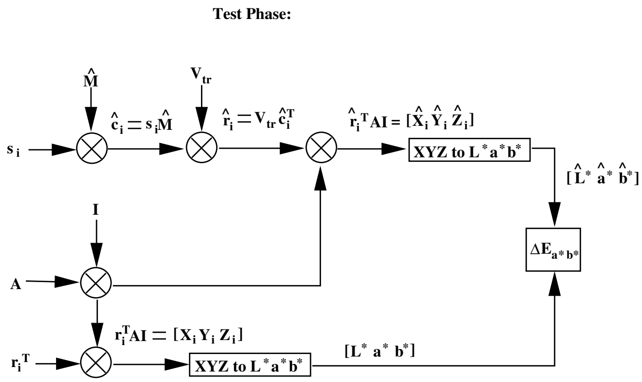
4. Results

4.1. Experiment 1

In Experiment 1, 70 extracted teeth were randomly selected as a training set to derive the model parameters, i.e the calibration matrix. The remaining 15 teeth were used to test the performance of the system. Since certain choices of training and test sets outperformed others, we ran 10,000 such random selections to provide a more accurate estimate of performance. The average ΔE_{a*b*} for the test set in each selection was computed and recorded. The statistics of this average for 10,000 random selections for the IDC are shown in Table 1. The measurements were taken under each of the three light sources mentioned in the previous section. For the IIC, the calibration matrix is computed as shown in Figure 3, with the halogen light; and the performance test was conducted according to the procedures shown in Figure 4. The spectral reflectance of each test tooth was computed. The average ΔE_{a*b*} of the reconstructed spectral reflectances under each of the three light sources was then calculated. For comparison purposes, the ΔE_{a*b*} for the IDC for the same test set was also calculated and included in Table 2.

4.2. Experiment 2

In Experiment 2 conducted with the subjects' teeth, the measurements were made under the halogen light. The gamuts of the extracted teeth and subjects' teeth were compared. Our data indicated that both gamuts overlap very little, and the subjects' teeth are redder than the extracted



s_i : i^{th} row of S matrix for the test sample set
$\hat{c}_i$: predicted c_i (as defined in Figure 1)
I : 31×31 diagonal illuminant matrix, may be different from L in Fig. 1
$\hat{r}_i$: reconstructed spectral reflectance of the i^{th} Test Tooth
A : as defined in Fig. 1
V_{tr} : as defined in Fig. 1
$\hat{M}$: calibration matrix from Fig. 3a
r_i : as defined in Fig. 1
$(\,.\,)^T$: matrix transposition

Figure 4: *Test phase of Illuminant-Independent Colorimeter.*

Table 2: *Experiment 1 (IIC): Extracted teeth measured only with halogen light.*

Illuminant	ΔE_{avg}			
	Mean	StdDev	Max	% < 1.5*
D65 (IIC)	0.833	0.109	1.354	100
Flash (IIC)	0.841	0.110	1.364	100
Halogen (IIC)	0.843	0.115	1.366	100
Halogen(IDC)	0.843	0.115	1.362	100

* percentage of random selections with average $\Delta E_{a*b*} < 1.5$. The IDC results are included for cross reference. The training and test samples in each random selection for Halogen(IDC), Halogen(IIC), D65 (IIC), and Flash (IIC) were the same.

teeth. Therefore, it is more appropriate to use the subjects' teeth to calibrate our system for clinical applications. To evaluate the system performance, 34 out of 42 of the subjects' teeth were selected randomly as a training set and the remaining 8 comprised the test set. 10,000 such selections were performed. The results obtained using both the IDC and the IIC are tabulated in Table 3.

Table 3: *Experiment 2 (IIC and IDC): Subjects' teeth measured only with halogen light.*

Illuminant	ΔE_{avg}			
	Mean	StdDev	Max	% < 1.5*
D65 (IIC)	1.080	0.118	2.03	98.2
Flash (IIC)	1.051	0.176	1.966	98.8
Halogen (IIC)	1.048	0.177	1.972	98.7
Halogen (IDC)	1.049	0.177	1.980	98.7

* percentage of random selections with average $\Delta E_{a*b*} < 1.5$.

Figure 5 and 6 show the best and the worst cases out of 15 reconstructions for the extracted teeth in a random selection. The ΔE_{a*b*} values were calculated for halogen light for both figures.

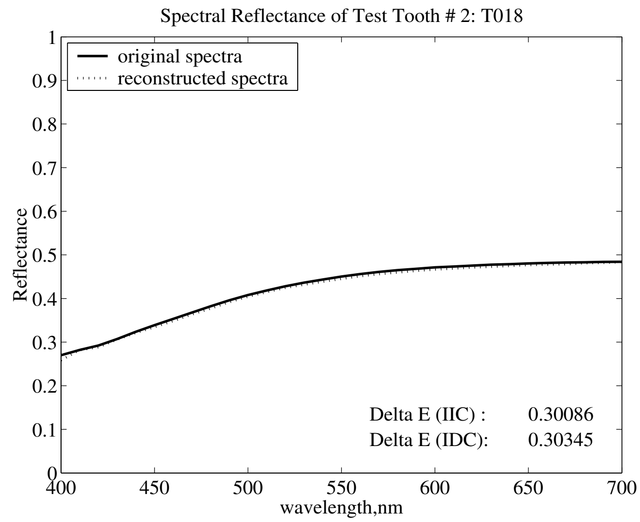


Figure 5: *Spectra for extracted teeth with IIC: The best among 15 reconstructions in a random selection. The ΔE_{a*b*} values were computed for the halogen light.*

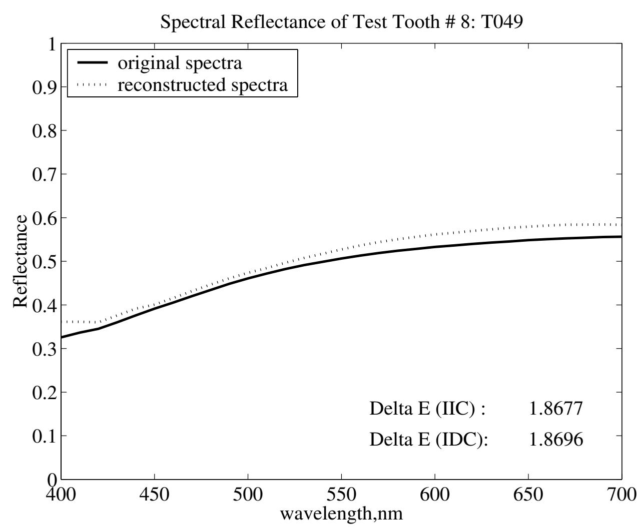


Figure 6: *Spectra for extracted teeth with IIC: The worst among 15 reconstructions in a random selection. The ΔE_{a*b*} values were computed for the halogen light.*

5. Discussion

The results displayed in Table 1 indicate that the illuminant-dependent colorimeter performed the best under the halogen light source. The ΔE_{a*b*} data set had the lowest mean value; and the average ΔE_{a*b*} for all selections was lower than 1.5. The system with daylight illumination did not perform as well as those of halogen and flash but the performance under all three light sources were similar. In addition to its somewhat inferior performance, the physical design of the daylight simulator is not suitable for clinical use. On the other hand, the flash lights are small, highly portable, and relatively cheap compared to the daylight simulator. With output characteristics similar to that of daylight in the visible spectrum, flash units offer more flexibility in the experiment setup design. Thus, the flashbased system seems to be a natural choice for clinical applications. However, the flash units have large temporal fluctuations in intensity. Meticulous work is needed to correct for the intensity fluctuation during the experiment. Moreover, it is difficult to estimate the position of the specular reflection before taking measurements with the spectroradiometer. This leads to increased measurement time which is undesirable when human subjects are involved. In addition, both the flash and the daylight illuminants have significantly higher output in the ultraviolet region (compared to the halogen light) which amplify the fluorescence in teeth. Consequently, the halogen-based system was chosen for the clinical test.

The results in Experiment 2 were slightly worse. This is probably due to subjects' movement during the measurement. It is very difficult for a subject not to move at all during the measurements. Since the color of a tooth is not uniform throughout the surface, any movement of the subject will cause the camera and the spectradiometer to measure slightly different regions of the tooth surface in subsequent measurements. Such movement also changes the scatter angle of the illumination on the tooth.

The results of Table 2 and Table 3 suggest that both the IDC and IIC are colorimetrically equivalent. The difference in the average E_{a*b*} was less than 0.1. In fact, the difference in the E_{a*b*} value between the illuminant-dependent and illuminant-independent systems is less than 0.1 for all the reconstructions in all selections. Fig. 5 and 6 show examples of reconstructed spectral reflectances for the extracted teeth in a randomly chosen selection. The reconstructed reflectance, when viewed under D65 and flash light, has E_{a*b*} values close to those of halogen light. The statistics in both tables also show that the predicted spectral reflectances from the halogen-based system, when evaluated under D65 or flash illumination, yield E_{a*b*} values close to those of the halogen-based system.

6. Conclusion

Both the illuminant-dependent and illuminant-independent colorimeters worked equally well under the same illumination condition. However, colors predicted by the illuminant-dependent colorimeter could only be matched visually under the same illuminant with which they were measured. The illuminant-independent colorimeter eliminates this constraint. The results of our clinical test also show that this system works in clinical situation. Nevertheless, a set of training samples, which represents the teeth gamut adequately, must be carefully picked to ensure good prediction.

References

1. Wencheng Wu and Jan P. Allebach and Mostafa Analoui, "Imaging colorimetry using a digital camera," *Journal of Imaging Science & Technology*, vol. 44, no. 4, pp. 267–79, 2000.
2. Bruce A. Matis and Michael A. Cochran and Timonthy J. Carlson, "The efficacy and safety of a 10% carbamide peroxide bleaching gel," *Quintessence International*, vol. 29, no. 9, pp. 555–63, 1998.
3. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley & Sons, Inc, 2nd ed., 1982.
4. J. Farrell and D. Sherman and B. Wandell,15 "How to turn your scanner into a colorimeter," *Proceedings of the IS&T's Tenth International Congress on Advances in Non-impact Printing Technologies*, pp. 579–81, 1994.
5. S. Tominaga, "Spectral imaging by a multi-channel camera," *SPIE*, vol. 3648, pp. 38–47, 1999.
6. Laurence Daligault and Jean Glasser, "Colorimeteric characterization of CCD sensors by spectrophotometry," *SPIE*, vol. 1512, pp. 124–30, 1991.
7. Shuxue Quan and Noboru Ohta and Naoya Katoh, "Optimization of camera spectral sensitivities," *The Eighth IS&T/SID Color Imaging Conference*, pp. 273–78, 2000, Scottdale, Arizona.
8. Costas Balas, "An imaging colorimeter for noncontact tissue color mapping," *IEEE Trans. on Biomedical Engineering*, vol. 44, no. 6, pp. 468–74, 1997.
9. Graham D. Finlayson and Mark S. Drew, "Constrained leastsquare regression in color spaces," *Journal of Electronic Imaging*, vol. 6, no. 4, pp. 484–93, 1997.
10. Markus Lenhard, "Assessing tooth color change after repeating bleaching in vitro with a 10 percent carbamide peroxide gel," *JADA*, vol. 127, pp. 1618–24, 1996.
11. T. P. van der Burgt and J. J. ten Bosch and P. C. F. Borsboom and W. J. P. M Kortsmit, "A comparison of new and conventional methods for quantification of tooth color," *The Journal of Prosthetic Dentistry*, vol. 63, no. 2, pp. 155–62, 1990.
12. Francisco H. Imai and Norimichi Tsumura and Hideaki Haneishi and Yoichi Miyake, "Principal component analysis of skin color and its application to colorimetric color reproduction on CRT display and hardcopy," *Journal of Imaging Science & Technology*, vol. 40, pp. 422–30, 1996.
13. Francisco H. Imai and Norimichi Tsumura and Hideaki Haneishi and Yoichi Miyake, "A comparative Analysis of spectral Reflectance Estimated in Various Spaces Using a Trichromatic Camera System," *Journal of Imaging Science & Technology*, vol. 44, no. 4, pp. 280–87, 2000.
14. Bernhard Flury, Common Principal Components and Related Multivariate Models. New York: Wiley, 1st ed., 1988.

15. E. Papakosta, *Master Thesis: Quantitative evaluation of the color of extracted human teeth and three shade guides with the use of a non-contacting technique*. Indiana University School of Dentistry, 1998.

Biography

Du-Yong Ng received his B.S. and M.S. degrees in electrical engineering from Purdue University in 1998 and 2001, respectively. He is currently pursuing his Ph.D degree in School of Electrical and Computer Engineering at Purdue University. His research interests include colorimetry, image capture and rendering, and color image processing.