New Figure of Merit for Color Reproduction Ability of Color Imaging Devices using the Metameric Boundary Descriptor

SIMON SCHULZ  
Hamburg University of Technology  
Vision Systems  
Harburger Schlossstrasse 20  
21079 Hamburg, GERMANY

ROLF-RAINER GRIGAT  
Hamburg University of Technology  
Vision Systems  
Harburger Schlossstrasse 20  
21079 Hamburg, GERMANY

Abstract: A measure for the color reproduction ability of a color imaging device is important for comparison purposes as well as for the sensor design process. Several figures of merit which describe this ability in a single measure have been proposed in the past. All these measures have shortcomings in that their error measure is defined in a non-representative color space for human color perception, or they do not consider measurement noise, or they rely on a specific color correction function. We introduce a new figure of merit for color reproduction ability of digital imaging devices. This new approach uses the metameric boundary descriptor (MBD) to quantify the limit of color reproduction for this sensor. It is therefore independent of the color correction method used. The error measure is defined in a perceptually uniform color space (CIELab) and takes measurement noise into account. Only the spectral sensitivities of an imaging device are needed. The advantages over existing quality measures as well as the plausibility of the assumptions required are discussed. Quality measures for 5 different cameras are presented.

Key–Words: Image quality, Optical filters, Spectral Sensitivity, Colorimetry, Quality Measure

December 19, 2006

1 Introduction

Whenever the recorded image of a color imaging device (e.g. a digital color camera) is to be presented to a human, the ability to reconstruct the perceived colors of the scene from the device’s output data is of great interest. A single figure of merit (FOM) would be of great help for choosing the correct device, optimizing the spectral sensitivities in the sensor design process and to be able to compare the performance of different devices for such an application.

The human visual color perception can be described in a tristimulus system. There are three different types of cones in the eye with different spectral sensitivities. Based on these spectral sensitivities the Commission Internationale de l’Eclairage (CIE) has defined the color matching functions for the standard observer and the color spaces CIEXYZ and CIELab. Whereas the CIEXYZ color space describes the human vision from the physical perspective, the CIELab color space was defined to be uniform in terms of perceived color differences. The Euclidean distance of two colors in CIELab closely matches the perceived color difference between these colors. Therefore it is useful to develop a figure of merit in CIELab color space.

2 Related work

Previously published figures of merit may be divided into two groups: the first group tries to measure the goodness as a geometrical difference between the human visual subspace and the subspace of the sensor. The second group measures the goodness as an average color error after an optimized color correction is applied to the sensor’s output.

2.1 Geometric Difference FOM

The figures of merit based on geometrical differences use calculations in the spectral vector space. Each spectrum $x(\lambda)$ is therefore given as a vector of $n$ sample values $\vec{x} \in \mathbb{R}^n$.

The response $\vec{a} = (X, Y, Z)^T$ of the human eye observing a surface with reflectance spectrum $\vec{r}$ under
the illumination \( \vec{b} \) can be expressed as

\[
\vec{a} = A^T D(\vec{b}) \vec{r} = A_b^T \vec{r} \tag{1}
\]

where the color matching functions for the standard observer are stacked into the matrix \( A = (\vec{x}, \vec{y}, \vec{z}) \in \mathbb{R}^{n \times 3} \) and vector \( \vec{b} \) is transformed into a diagonal matrix containing the elements of \( \vec{b} \) by the operator \( D(\bullet) \). Replacing the matrix \( A \) in eqn. 1 with the matrix \( \Omega \in \mathbb{R}^{n \times m} \) of the stacked spectral sensitivities \( \vec{\omega}_i, \ i = 1, \ldots, m \) of an \( m \)-channel sensor, we get the sensor response \( \vec{c} \) of the same stimulus:

\[
\vec{c} = \Omega^T D(\vec{b}) \vec{r} = \Omega_b^T \vec{r} \tag{2}
\]

Using vector space \( V \in \mathbb{R}^n \), we can define a subspace \( \text{HVSS} := \text{Span}(\vec{x}, \vec{y}, \vec{z}) = \text{Span}(A) \), \( \text{HVSS} \subset V \) to be the human visual subspace, which is determined by the columns of \( A \) as basis vectors. Analogously we define the system visual subspace to be \( \text{SVSS} := \text{Span}(\Omega) \).

Neugebauer’s \( q \)-factor [1] uses a projection \( \mathcal{P}_{A_b}(\vec{\omega}_i) \) of a single spectral sensitivity \( \vec{\omega}_i \) of a sensor onto the \( \text{SVSS}_b \) under the illumination \( \vec{b} \). If we consider the length of the projection \( \mathcal{P}(\bullet) \) in proportion to the length of the spectral sensitivity \( \vec{\omega}_i \), we get a representation of the part of the color energy that is recoverable from such a sensor channels response:

\[
q(\vec{\omega}_i) = \frac{\|\mathcal{P}_{A_b}(\vec{\omega}_i)\|^2}{\|\vec{\omega}_i\|^2} \tag{3}
\]

The projection \( \mathcal{P}_{A_b}(\bullet) \) may be expressed by the pseudo inverse \( (A_b^T)^\odot \).

\[
\mathcal{P}_{A_b}(\vec{\omega}_i) = (A_b^T)^\odot A_b^T \vec{\omega}_i = (A_b^T)^\odot (A_b^T \vec{\omega}_i) \tag{4}
\]

Note that on the other hand using eqn. 4 we could try to reconstruct the reflectance spectrum \( \vec{r} \) from the sensor responses \( \vec{c} \)

\[
\vec{r} = A_b (\Omega_b^T \Omega_b)^{-1} \Omega_b^T \vec{c} + \text{Kern}(\Omega_b^T) = (\Omega_b^T)^\odot \vec{c} + \text{Kern}(\Omega_b^T) \tag{5}
\]

As long as \( n > m \), \( \text{Kern}(\Omega_b^T) \neq 0 \) holds true. This means that the reflectance spectrum \( \vec{r} \) cannot be reconstructed from the sensor responses \( \vec{c} \) without error.

The major drawback of Neugebauer’s FOM is it’s restriction to a single sensor channel only. Even the separate calculation of the \( q \)-factors for all sensor channels doesn’t help, as the joined performance of the sensor channels is of interest. Therefore Vora [2] extended Neugebauer’s \( q \)-factor to be used for multiple channel sensors. Vora calculates a measure for the overlap of the \( \text{SVSS}_b \) and the \( \text{HVSS}_b \) using an orthonormal basis \( O \) and \( N \) respectively for both. Under the assumption of reflectance spectra \( \vec{r}_i \) being composed of statistically independent, identically distributed random variables, he defines \( \nu(A_b, \Omega_b) \) to be his FOM:

\[
\nu(A_b, \Omega_b) = \frac{1}{\alpha} \sum_{i=1}^{\alpha} \gamma_i^2 (O^T N) \tag{6}
\]

where \( \gamma_i(O^T N) \) denotes the \( i \)-th singular value of \( O^T N \).

The Color Quality Factor (CQF) proposed by Sharma [6] uses Neugebauer’s FOM but in the opposite direction, calculating projections (e.g. \( \mathcal{P}_{\Omega_b}(\vec{x}) \)) of the human visual “channels” onto the SVSS.

### 2.2 Average Error Color FOM

The second class of FOMs use the mean color error on a set of reflectance spectra in a CIE color space. The sensor responses \( \vec{c} \) are mapped to \( CIEXYZ \) using a linear transformation. The resulting color differences between reference colors and corrected colors are taken as measurement of quality. Shimano’s \( Q_{st} \) and \( Q_{sf} \) metrics [3, 4] consider the minimized color error in \( CIEXYZ \) space. Tajima’s indexes [5] assume object color spectral characteristics to be composed of few principle components only. Hung’s CRI and Sharma-Trussel’s FOM [6] rate the color reproduction ability using an error measure in a perceptually uniform color space using local or global linearization techniques. They even take signal independent recording noise into account, which none of the previously mentioned methods does. Quan [7] further extends this FOM to the so called Unified measure of goodness (UMG) taking signal dependent recording noise into account.

The results of these FOMs depend on the set of reflectance spectra used. Therefore these sets should be choosen with care. The reflectance spectra of the GretagMacbeth ColorChecker are commonly used due to their representativeness for much larger sets [8] as well as the widespread use of this target in the color imaging area.
3 Metameric Boundary Descriptor FOM

An resonsable FOM should meet the following requirements: The figure of merit should rate the quality of color reproduction ability in a perceptual relevant error measure as color reproduction for a human observer is the focus. Recording noise is an issue in all practical use of imaging sensors and should therefore be considered in the FOM calculation. The FOM should not rely on a specific color correction technique. Incorporating a specific color correction technique into the FOM would result in a quality measure that rates the joint performance of the device with this specific correction technique. A statement for color reproduction ability of the specific device alone could therefore not be given.

The motivation to introduce a new FOM is the limitation of the previously presented FOMs:

The geometric difference FOMs all suffer from their error measure not being defined in a perceptual uniform color space. The error is quantified in the spectral vector space which represents the physical layer of human vision. For our figure of merit to be relevant we rather need to define our error in a perceptual uniform color space.

Some average color error FOMs use error measures defined in linearized perceptual uniform color spaces. Sharma-Trussell’s FOM [6] and Quan’s UMG [7] additionally consider recording noise. But all previous color error FOMs rely on an optimized linear color correction function in order to transform the sensor output values \( \tilde{c} \) into the estimate \( \tilde{a} \) of perceived colors.

Using the Metameric Boundary Descriptor (MBD) suggested by Urban [9, 10] we may define a color error measure in the perceptually mostly uniform CIELab color space that is completely independent of the color correction function used. It only describes the limitation of color reproduction itself using a theoretical, best possible color correction function. For a given sensor output \( \tilde{c} \) we can interpret the MBD as a measure of uncertainty with respect to the perceived color that led to this sensor output. Recording noise can easily be taken into account.

The MBD describes the metameric subspace \( M_{XYZ}^\tilde{c} \) of a sensor output \( \tilde{c} \). The metameric subspace \( M_{XYZ}^\tilde{c} \) contains all perceived colors \( \tilde{a} = A_{ba}^T \tilde{r} \) under viewing illumination \( \tilde{b}_a \) that may lead to the sensor output \( \tilde{c}_{ba} = \Omega_{ba}^T \tilde{r} \) under recording illumination \( \tilde{b}_a \):

\[
M_{XYZ}^\tilde{c} := \{ A_{ba}^T \tilde{r} \mid \forall \tilde{r} : \tilde{c}_{ba} = \Omega_{ba}^T \tilde{r} \} \tag{7}
\]

If we use the well known tranformation \( L : \text{CIEXYZ} \mapsto \text{CIELab} \) [11], we can express this metameric subspace in a perceptually uniform color space:

\[
M_{\text{CIELab}}^\tilde{c} := \{ L(A_{ba}^T \tilde{r}) \mid \forall \tilde{r} : \tilde{c}_{ba} = \Omega_{ba}^T \tilde{r} \} \tag{8}
\]

Not all reflectances \( \tilde{r} \) that comply eq. (8) are necessarily physically reasonable. Assuming the reflectances \( \tilde{r} \) to be smooth, non negative \((r_i \geq 0, i = 1, \ldots, n)\) and bounded \((r_i \leq 1, i = 1, \ldots, n)\), we can reduce the volume of the metameric subspace to realistic description of the color reproduction ability for the camera for a given sensor output \( \tilde{c} \).

In order to extract an illustrative error measure from the MBD method, we consider the average volume \( \bar{V} = \frac{1}{k} \sum_{i=0}^{k} V(M_{\text{CIELab}}^\tilde{c}) \) of the metameric subspaces over the sensor responses \( \tilde{c} \) to a set of \( k \) representative reflectance spectra. The error measure \( \tau_{\text{CIELab}} \) is defined as the radius of a globe with equal volume \( \bar{V} \):

\[
\tau_{\text{CIELab}} = \sqrt[3]{\frac{3}{4\pi} \bar{V}} = \sqrt[3]{\frac{3}{4\pi} \frac{1}{k} \sum_{i=0}^{k} V(M_{\text{CIELab}}^\tilde{c})} \tag{9}
\]

\( \tau_{\text{CIELab}} \) may be interpreted as the mean color error \( \Delta E_{ab} \) that is to be expected using the sensor with an optimal color correction. \( \tau_{\text{CIELab}} \) is an illustrative FOM due to \( \Delta E_{ab} \) being a well known error measure in the color imaging community.

Recording noise can easily be integrated in eq. (2) by adding a vector of random variables \( \tilde{c} \): \( \tilde{c} = \tilde{c} + \epsilon \). This leads to a modified description of the metameric subspace:

\[
M_{\text{CIELab}}^{\tilde{c} + \epsilon} := \{ L(A_{ba}^T \tilde{r}) \mid \forall \tilde{r} : \tilde{c}_{ba} = \Omega_{ba}^T \tilde{r} + \epsilon \} \tag{10}
\]

Whereas the noise \( \epsilon_i \) is only bounded \(-\Delta \epsilon < \epsilon_i < \Delta \epsilon \) but a priori unknown.

The metameric subspace is described in the MBD by storing points in the CIELab color space that are equally distributed over the surface of the metameric subspace. These points are calculated by solving linear optimization problems. We refer to the publications of Urban [9, 10] for futher details.

We can calculate the MBD FOM for several combinations of recording and viewing illuminations \((\tilde{b}_{i1}, \ldots, \tilde{b}_{ip})\) and \((\tilde{b}_{a1}, \ldots, \tilde{b}_{aq})\) respectively. The results may be presented in a quality matrix \( T \) as proposed in [7]:
The results of the MBD FOM calculation for the five sensors are shown in figure 1. The different spectral sensitivities of the sensors used are shown in figure 2. For comparison purposes we have included an “ideal” 3-channel sensor with Gaussian shaped spectral sensitivities as well as an experimental 6 channel sensor [12] denoted as K6. An example of a metameric subspace for the 8th field of the GretagMacbeth ColorChecker recorded and viewed under D50 illumination is shown in figure 3. The K6 sensor was used to generate this MBD.

The results of the MBD FOM calculation for the five sensors are shown in figure 1. The different spectral sensitivities of the sensors used are shown in figure 2. For comparison purposes we have included an “ideal” 3-channel sensor with Gaussian shaped spectral sensitivities as well as an experimental 6 channel sensor [12] denoted as K6. An example of a metameric subspace for the 8th field of the GretagMacbeth ColorChecker recorded and viewed under D50 illumination is shown in figure 3. The K6 sensor was used to generate this MBD.

The results of the MBD FOM calculation for the five sensors are shown in figure 1. The different

$$T = \begin{bmatrix} \tau_{11} & \tau_{12} & \cdots & \tau_{1q} \\ \tau_{21} & \tau_{22} & \cdots & \tau_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{p1} & \tau_{p2} & \cdots & \tau_{pq} \end{bmatrix}$$  \hspace{1cm} (11)

where \(\tau_{ij}\) is the quality measure defined in eq. (9) for recording illumination \(\tilde{b}_{ij}\) and viewing illumination \(\tilde{b}_{oij}\). We can even formulate a single quality factor using a weighted sum of the matrix elements:

$$\tau = \sum_{i=1}^{p} \sum_{j=1}^{q} w_{ij} \tau_{ij} \text{ while } \sum_{i=1}^{p} \sum_{j=1}^{q} w_{ij} = 1$$  \hspace{1cm} (12)

The weights \(w_{ij}\) may be chosen according to the importance of this specific combination of recording and viewing illumination for the application.

4 Results

In order to evaluate the newly proposed FOM, we calculated the MBDS for five different sensors using the reflectances of the GretagMacbeth ColorChecker. The reflectances of the ColorChecker were used due to their well known representativeness of much larger spectral sets [8].

The spectral sensitivities of the sensors used are shown in figure 2. For comparison purposes we have included an “ideal” 3-channel sensor with Gaussian shaped spectral sensitivities as well as an experimental 6 channel sensor [12] denoted as K6. An example of a metameric subspace for the 8th field of the GretagMacbeth ColorChecker recorded and viewed under D50 illumination is shown in figure 3. The K6 sensor was used to generate this MBD.

The results of the MBD FOM calculation for the five sensors are shown in figure 1. The different
5 Conclusion

We presented a new figure of merit for color reproduction ability of digital imaging devices. This FOM is defined in a perceptual uniform color space, takes recording noise into consideration and quantifies the ability to reproduce colors independently of the color correction method used. The resulting $\tau_{CIELab}$ values are illustratively given in $\Delta E_{ab}$ units of the expected mean color error using a specific sensor with the best possible color correction function.

The results presented in this paper correspond to the experiences in color reproduction ability with the sensors used.

References:


