Improving Color-Difference Formulas Using Visual Data

Ingmar Lissner and Philipp Urban;
Institute of Printing Science and Technology, Technische Universität Darmstadt;
Magdalenenstr. 2, 64289 Darmstadt, Germany

Abstract

We propose a method to improve existing color-difference formulas with additional visual data from color discrimination experiments. Color-difference formulas are treated as mean functions of Gaussian processes, and the visual data are considered as observations of these processes. Gaussian process regression is applied to predict unknown color differences. The method was evaluated with a combination of the CIEDE2000 color-difference formula and the RIT-DuPont dataset. The standardized residual sum of squares (STRESS) index between visual and computed color differences was determined for several sets of visual data. The results show a STRESS index of 6.94 (CIEDE2000: 19.47) for the RIT-DuPont dataset. The prediction performance on other visual data (BFD, Leeds, Witt) is not significantly different from CIEDE2000 at a 95% confidence level.

Introduction

A perceptually uniform color space is required especially for quality control and various color technology applications. The CIELAB color space was designed for this purpose in 1976. It is used in many standards of the printing, graphic arts, coating, and automotive industries. Several drawbacks of CIELAB were discovered by now, including its lack of perceptual uniformity and hue linearity. Nevertheless, CIELAB could not be replaced by an improved color space (such as DIN99 [1, 2]) in everyday applications. Since changing common practice is difficult, this is unlikely to happen in the near future.

Unfortunately, for some applications the perceptual uniformity of CIELAB is not sufficient, especially when color tolerances need to be defined. Data derived from color discrimination experiments (e.g., RIT-DuPont [3]) show the disagreement between perceived differences and Euclidean distances in CIELAB. Suprathreshold ellipsoids, which approximately define all colors with similar perceived distance to a color center, are of particular interest. Figure 1 shows four such ellipsoids derived from the RIT-DuPont data [4]. They differ significantly from spheres, which is a good indicator of perceptual non-uniformity of the underlying color space.

To overcome the non-uniformity of CIELAB, various color-difference formulas were created (e.g., CMC [5], CIE94 [6], and CIEDE2000 [7]). Figure 1 shows that CIEDE2000 predictions differ not only from RIT-DuPont suprathreshold ellipsoids (which can be considered reliable [4]), but also from the corresponding color-difference vectors. A comparison of the entire RIT-DuPont dataset with corresponding CIEDE2000 predictions yields a PF/3 measure [8] of 19.56 and a standardized residual sum of squares (STRESS) index [9, 10] of 19.47.

Color-difference formulas are designed by fitting parameters of predefined functions to visual data [11]. There are two main problems with this approach: 1. Visual data from a single experiment are usually sparsely distributed across CIELAB, and combining datasets obtained by different psychophysical methods (e.g., method of constant stimuli or gray-scale method [12]) is highly controversial. 2. The visual data might be overfitted, so that the resulting color-difference formula models noise and loses its generalization ability [13, 14].

The aim of this paper is not to create new color-difference formulas, but to improve existing formulas using visual data. A possible application is to enhance the prediction accuracy around particular color centers, e.g., a company's corporate colors. Visual experiments at these color centers could improve the standardized global color-difference formula, provided that they were conducted under similar viewing conditions [15].

The proposed method is based on *Gaussian process regression* (GPR), a prediction approach often used in a geostatistical context (usually referred to as *Kriging*). It allows to incorporate the uncertainty of the visual data, which is important due to high inter- and intra-observer noise in experimental data [14, 16].

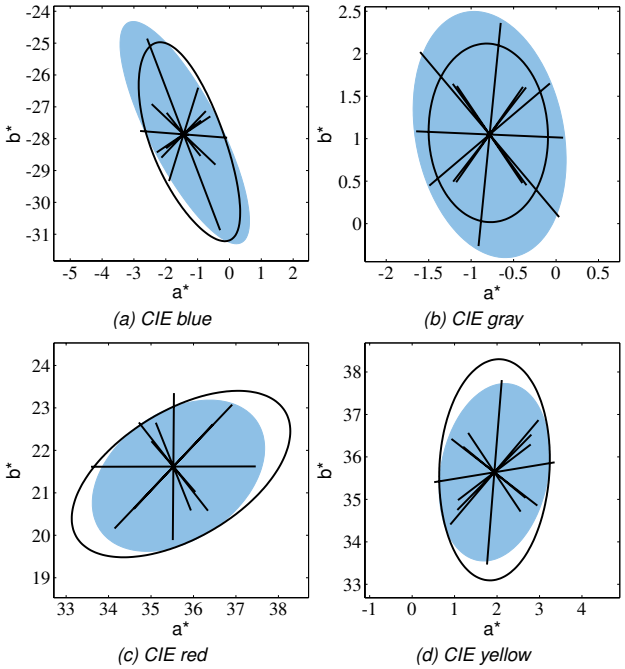

Figure 1. *Four color centers recommended by the CIE for study [17]. RIT-DuPont suprathreshold ellipsoids (blue) with the corresponding T50 and -T50 color-difference vectors in comparison with CIEDE2000 iso-distance contours (black). Each point on a contour has a computed distance of 1 to the respective color center. Projections on the a*b*-plane.*

What Is Gaussian Process Regression?

A detailed description of Gaussian processes and Gaussian process regression (GPR) goes far beyond the scope of this paper. A good introduction to this topic including the mathematical background is given by Rasmussen and Williams [18]. In the following we will discuss the central equations to allow a quick implementation of the method. To illustrate the basic concept, let

us consider a simple example.

An Example from Mining

Imagine we are planning a gold mine and want to find out about gold concentrations in the area of interest. Drilling is expensive and should be avoided as far as possible. The idea is to treat the whole area as a Gaussian process [18] of gold concentrations. The concentration at each location is modeled as a normally distributed random variable. The joint probability of any finite number of these random variables is again considered as normally distributed. By drilling we can determine the gold concentration at a specific location with some uncertainty. It is unlikely that the gold concentration changes drastically in the vicinity of the borehole. Ideally, we have an idea of the correlation of concentrations between two locations. This correlation can be modeled by a *covariance function*. If typical gold concentrations of the region are known, we can also include this knowledge as a *mean function*. The resulting Gaussian process is completely defined by its mean and covariance functions. Using a Gaussian process model and a few boreholes, GPR allows us to predict the gold concentration at any location in the region. Based on our model we can not only compute the most likely concentration at a point of interest, but also the uncertainty of this prediction.

GPR for Improving Color-Difference Formulas

Interestingly, color-difference research and mining have a lot in common. Instead of gold deposits we are interested in perceived differences between colors. Color discrimination experiments at selected color centers are our boreholes from the mining example. The result of such an experiment is a set of perceived distances ΔV^i for color pairs $\mathbf{x}^i$, where $\mathbf{x}^i = (x_1^i, x_2^i)$ and $x_1^i, x_2^i \in$ CIELAB, $i = 1,\ldots,n$. These observed distances are affected by noise and usually sparsely distributed across CIELAB.

A perceptually optimal color-difference formula is a function that maps each color pair to its perceived color difference, i.e.:

$$\Delta E_V : \begin{cases} \text{CIELAB} \times \text{CIELAB} \longmapsto \mathbb{R}_0^+ \\ \mathbf{x} \to \Delta V(x_1, x_2) \end{cases} . \tag{1}$$

Standardized color-difference equations such as CIE94 or CIEDE2000 are rough approximations of ΔE_V. Unfortunately, we only have limited knowledge of the disagreement between a given color-difference formula and the ideal ΔE_V. This is why we use a Gaussian process model to improve color-difference formulas with data from visual experiments. We assume that each color difference is a normally distributed random variable, and that the joint distribution of any number of color differences is again normally distributed. As previously mentioned, our Gaussian process is completely defined by its mean and covariance functions. Any given color-difference formula can be used as a mean function, provided that it returns an initial guess for ΔE_V. In case we use CIEDE2000, this mean function becomes

$$m(\mathbf{x}) = \Delta E_{00}(x_1, x_2), \tag{2}$$

where $\mathbf{x} = (x_1, x_2)$ is a pair of CIELAB colors x_1 and x_2.

The crucial part of Gaussian process regression is the selection of the covariance function, which describes the correlation between color pairs. A difference vector $\mathbf{x}_d = x_1 - x_2$ is associated with each color pair $\mathbf{x} = (x_1, x_2)$. We assume that the correlation between two color pairs depends on their distance in color space as well as the angle and the length difference between their difference vectors. This means that close color pairs with similarly aligned difference vectors of similar length are highly correlated. For two color pairs $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ and the corresponding difference vectors $\mathbf{x}_d = x_1 - x_2$ and $\mathbf{y}_d = y_1 - y_2$ we define the following covariance function:

$$k(\mathbf{x}, \mathbf{y}) = c(\mathbf{x}, \mathbf{y}) \cdot \exp\left[-\left(\frac{\phi_{x,y}^2}{l_1} + \frac{\alpha_{x,y}}{l_2} + \frac{\rho_{x,y}}{l_3} \right) \right], \tag{3}$$

where

$$\phi_{x,y} = \left\| \left(\frac{x_1 + x_2}{2} \right) - \left(\frac{y_1 + y_2}{2} \right) \right\|_2 \tag{4}$$

is a measure of distance between $\mathbf{x}$ and $\mathbf{y}$,

$$\alpha_{x,y} = \arccos\left[\frac{\text{abs}(\mathbf{x}_d^{\mathrm{T}} \mathbf{y}_d)}{\|\mathbf{x}_d\|_2 \cdot \|\mathbf{y}_d\|_2} \right] \tag{5}$$

represents the angle between $\mathbf{x}_d$ and $\mathbf{y}_d$,

$$\rho_{x,y} = \text{abs}\left(\|\mathbf{x}_d\|_2 - \|\mathbf{y}_d\|_2 \right) \tag{6}$$

is the absolute length difference of $\mathbf{x}_d$ and $\mathbf{y}_d$, and

$$c(\mathbf{x}, \mathbf{y}) = 1 - \exp\left(-\|\mathbf{x}_d\|_2 \cdot \|\mathbf{y}_d\|_2 \right) \tag{7}$$

is a normalization term that forces the correlation to be zero if $\mathbf{x}_d = 0$ or $\mathbf{y}_d = 0$. An illustration of the three influence factors $\phi_{x,y}$, $\alpha_{x,y}$, and $\rho_{x,y}$ is provided in figure 2. The covariance function is symmetric by construction: $k(\mathbf{x}, \mathbf{y}) = k(\mathbf{y}, \mathbf{x})$. It is a normalized product of three functions: one squared-exponential type covariance function and two modified exponential type covariance functions [18].

For $\mathbf{x} = \mathbf{y}$ all three terms are zero, so that $k(\mathbf{x}, \mathbf{y}) = c(\mathbf{x}, \mathbf{y})$. With increasing distance, angle, and length difference the covariance function approaches its minimum $k(\mathbf{x}, \mathbf{y}) = 0$. The scaling parameters l_1, l_2, and l_3 control the correlation decay and can be adjusted to a particular problem. A small parameter l_i means that the corresponding term has a high impact on the resulting correlation, whereas a large l_i means that the correlation function does not strongly depend on this influence factor.

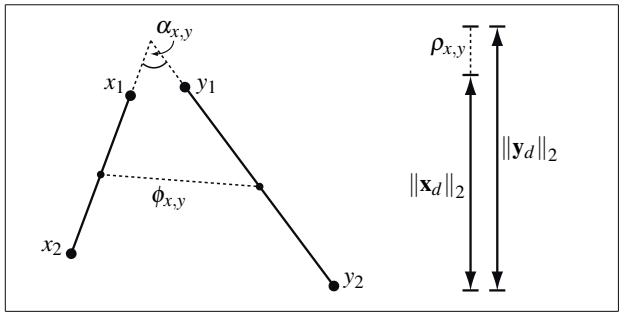

Figure 2. *Influence factors $\phi_{x,y}$ (distance), $\alpha_{x,y}$ (angle), and $\rho_{x,y}$ (length difference) on the correlation of two color pairs* $\mathbf{x}$ *and* $\mathbf{y}$.

So far we have defined the Gaussian process by its mean and covariance functions (eqns. (2) and (3)). In order to improve a color-difference formula with the visual data $(\mathbf{x}^i, \Delta V^i)$, $i = 1,\ldots,n$, we use Gaussian process regression to calculate the predicted color difference ΔV^* for any color pair $\mathbf{x}^*$ [18]:

$$\Delta V^* = m^* + \underbrace{\mathbf{k}^*(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}(\mathbf{v} - \mathbf{m})}_{\text{correction term}}, \tag{8}$$

where ΔV^* is the most likely color difference given the observations, $m^* = m(\mathbf{x}^*)$ is the value of the mean function at $\mathbf{x}^*$, $\mathbf{k}^* = (k(\mathbf{x}^*, \mathbf{x}^1), \ldots, k(\mathbf{x}^*, \mathbf{x}^n))$ is a vector of covariances between $\mathbf{x}^*$ and all observations,

$$\mathbf{K} = \{k(\mathbf{x}^i, \mathbf{x}^j)\}_{n \times n} = \begin{bmatrix} k(\mathbf{x}^1, \mathbf{x}^1) & \cdots & k(\mathbf{x}^1, \mathbf{x}^n) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}^n, \mathbf{x}^1) & \cdots & k(\mathbf{x}^n, \mathbf{x}^n) \end{bmatrix} \quad (9)$$

is a symmetric, positive semidefinite covariance matrix, $\sigma_\varepsilon^2 \mathbf{I}$ is an $n \times n$ dimensional noise matrix with noise variance σ_ε^2, $\mathbf{v} = (\Delta V^1, \ldots, \Delta V^n)^{\mathrm{T}}$ are the observed color differences, and $\mathbf{m} = (m(\mathbf{x}^1), \ldots, m(\mathbf{x}^n))^{\mathrm{T}}$ are the values of the mean function at the observations. Note that the mean function represents the color-difference equation to be improved (e.g., CIEDE2000). Its prediction m^* is corrected by $\mathbf{k}^*(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}(\mathbf{v} - \mathbf{m})$ to incorporate the additional visual data as shown in eq. (8).

The noise term $\sigma_\varepsilon^2 \mathbf{I}$ expresses our assumption that the observed color differences ΔV^i are affected by Gaussian noise that is uncorrelated between color pairs. The unknown variance σ_ε^2 has to be adjusted to the observations along with the decay parameters l_1, l_2, and l_3.

Possible Improvements of the Model

Please note that negative color-difference predictions ΔV^* arise if $\mathbf{k}^*(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}(\mathbf{v} - \mathbf{m}) < -m^*$. Although this was only observed for extremely small color differences, it needs to be intercepted in an implementation of the method. For instance, we can force the corrected color-difference prediction to be non-negative by introducing a function f as follows:

$$\Delta V^* = m^* + f\left[\mathbf{k}^*(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}(\mathbf{v} - \mathbf{m}),\, m^*\right], \quad (10)$$

where f satisfies

$$f\left[\mathbf{k}^*(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}(\mathbf{v} - \mathbf{m}),\, m^*\right] \geq -m^*. \quad (11)$$

The function f should be strictly monotonically increasing, continuously differentiable, and it should leave the correction term unchanged for $m^* \gg \left|\mathbf{k}^*(\mathbf{K} + \sigma_\varepsilon^2 \mathbf{I})^{-1}(\mathbf{v} - \mathbf{m})\right|$.

Another issue is the positive semidefiniteness of the covariance function k, which ensures the positive semidefiniteness of the covariance matrix $\mathbf{K}$. One possible way to create a valid covariance function is to combine several covariance functions by summation or multiplication [18]. This could be a guideline for further investigations regarding the validity of our covariance function. It should be mentioned that we did not encounter any problems during the evaluation of our method: the covariance matrix $\mathbf{K}$ was positive semidefinite for all sets of visual data and parameters l_1, l_2, l_3, and σ_ε^2 we used.

An in-depth analysis of color discrimination data should be performed to design a covariance function that optimally reflects the properties of the visual data. This may lead to an improved prediction performance especially on unknown data.

Adjusting the Parameters

There are different ways to adjust the parameters of the Gaussian process model to a particular problem. One is to maximize the *marginal likelihood* $p\left(\mathbf{v} \,|\, \hat{\mathbf{x}}, \{l_1, l_2, l_3, \sigma_\varepsilon^2\}\right)$ [18], i.e., the probability of the observations $\mathbf{v}$ given the vector of color pairs $\hat{\mathbf{x}} = \{\mathbf{x}^i\}_{i=1}^n$ and a parameter set $\{l_1, l_2, l_3, \sigma_\varepsilon^2\}$.

Another option is *cross-validation*, which is a technique to evaluate the prediction performance of a method on unknown data. In our case the visual data could be randomly split into a larger training set and a smaller test set. The Gaussian process, more precisely its covariance function k, would be adjusted to the training set data. The color differences of the test set would then be predicted with different combinations of parameters. This could be performed iteratively for many randomly chosen training and test sets. The parameter set yielding the highest prediction accuracy would eventually be considered optimal.

Experience shows that local optima are likely in both cases. The parameter sets corresponding to these optima can be seen as "particular interpretations of the data" [18]. This is especially inconvenient if the amount of data is large (resulting in a large covariance matrix $\mathbf{K}$), and if there are many parameters to optimize. In some cases the computational effort may even be too high to find a global optimum. In other cases, there may be many local optima of almost equal optimality corresponding to vastly different parameter sets. In these situations one is forced to manually select parameters that represent a desirable interpretation of the data.

For the evaluation of the method the parameters were manually chosen as $l_1 = 20$, $l_2 = 7.5$, $l_3 = 10$, and $\sigma_\varepsilon^2 = 0.25$. These parameters were roughly adjusted to the orders of magnitude of the corresponding influence factors $\phi_{x,y}$, $\alpha_{x,y}$, and $\rho_{x,y}$. They were also found to allow for a good generalization performance of the method. Note that this parameter set may not be ideal and efficient parameter optimization should be subject to further research.

Results and Discussion

The performance of several color-difference equations was evaluated on four sets of visual data: RIT-DuPont (312 color differences) [3], BFD-D65 (2028) [19], Leeds pair comparison (104) [20], and Witt (418) [21]. Our method was trained with the RIT-DuPont data, meaning that the covariance matrix $\mathbf{K}$ was created from the 312 RIT-DuPont color pairs. We chose this dataset because of its rather small size and high reliability. The CIEDE2000 color-difference formula ΔE_{00} was used as a mean function. As already mentioned, the parameters were manually chosen as $l_1 = 20$, $l_2 = 7.5$, $l_3 = 10$, and $\sigma_\varepsilon^2 = 0.25$.

The STRESS index [9, 10] was computed for the visual and the computed color differences. The results are shown in table 1.

	RIT-DuPont	BFD-D65	Leeds-PC	Witt
CIELAB	33.42	40.98	36.58	51.71
CMC	27.44	26.60	26.77	35.04
CIE94	20.31	32.88	32.69	31.94
CIEDE2000	19.47	24.09	20.98	30.22
GPR	6.94*	24.55	17.64	32.87

Table 1. STRESS index comparison. An asterisk indicates that the GPR-based predictions are significantly different from the CIEDE2000 predictions on the respective dataset (according to the F-test with a 95% confidence interval).

The STRESS index allows a statistical judgment of the prediction performance of color-difference formulas. To determine whether two formulas ΔE_a and ΔE_b are *significantly* different, one has to take the squared ratio of the corresponding STRESS values S_a and S_b on the same test data [9, 10]:

$$S_r = \frac{S_a^2}{S_b^2}, \quad S_a > S_b. \quad (12)$$

Let F_c be the value of the corresponding F-distribution with $n-1$ degrees of freedom (number of test color pairs minus one) and the desired confidence level, e.g., 95%. If S_r lies outside the confidence interval $[F_c, 1/F_c]$, the predictions of the two color-difference formulas are significantly different. For example, to compare CIEDE2000 and our GPR approach on the RIT-DuPont data with a 95% confidence level, one would compute $S_r = (19.47)^2/(6.94)^2 = 7.87$. This squared ratio is (a) greater than one, which indicates that the GPR approach performs better than CIEDE2000, and lies (b) outside the 95% confidence interval $[F_c, 1/F_c] = [0.729, 1.372]$, meaning that this difference is statistically significant [10]. Note that the number of color pairs n was set to 156 for the RIT-DuPont dataset, because the T50 and -T50 color-difference vectors are not really independent [3, 10].

Additional F-tests show that the prediction performance of our method on the remaining three datasets is not significantly different from CIEDE2000 at a 95% confidence level.

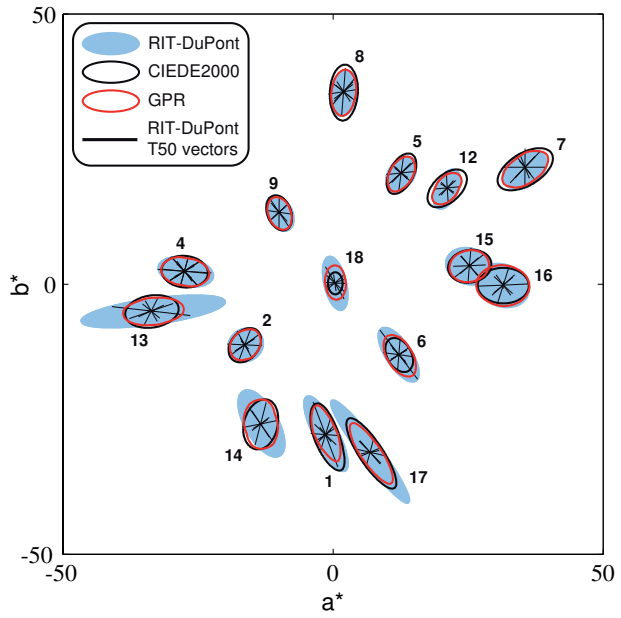

Figure 3. *Suprathreshold ellipsoids (blue) fitted to the RIT-DuPont T50 and -T50 vectors (black) compared to CIEDE2000 iso-distance contours (black) and GPR iso-distance contours (red). 15 out of 19 RIT-DuPont color centers are shown. All ellipsoids and contours are twice their original size and are projected on the CIELAB a*b*-plane.*

The GPR-based method was expected to achieve a better STRESS value on the RIT-DuPont dataset than all other color-difference equations (see table 1), because this dataset was used to train the method. Note that the STRESS value of the GPR-based predictions on the RIT-DuPont data is not zero. This follows directly from our assumption that the observed color differences are affected by additive noise with variance σ_ε^2 (eq. (8)). As a consequence, GPR-predicted color differences differ from the actual observations, even for color pairs included in the (noisy) training data.

Interestingly, the STRESS difference between CIEDE2000 and the GPR-based method is statistically not significant for all other investigated datasets according to the F-test. This indicates that the underlying CIEDE2000 mean function has a stabilizing effect on the method. In case the test data show little correlation with the training data, the method can fall back on its mean function.

The results show that GPR improves the CIEDE2000 predictions on the Leeds-PC (pair comparison) data (STRESS 17.64 vs. 20.98), whereas the predictions on the gray-scale-based datasets BFD-D65 (24.55 vs. 24.09) and Witt (32.87 vs. 30.22) are less accurate. Even though these differences are statistically not significant, the prediction performance of the GPR-based method seems to depend on the psychophysical method used to obtain the visual data — the training data, RIT-DuPont, are based on pair comparison experiments. Further investigations are required to validate this assumption.

Figure 3 provides an overview of the CIEDE2000 and the GPR-based predictions at selected RIT-DuPont color centers. The difference is especially evident at color centers 6, 8, 12, and 18, where the GPR-based predictions represent the RIT-DuPont color-difference vectors more accurately.

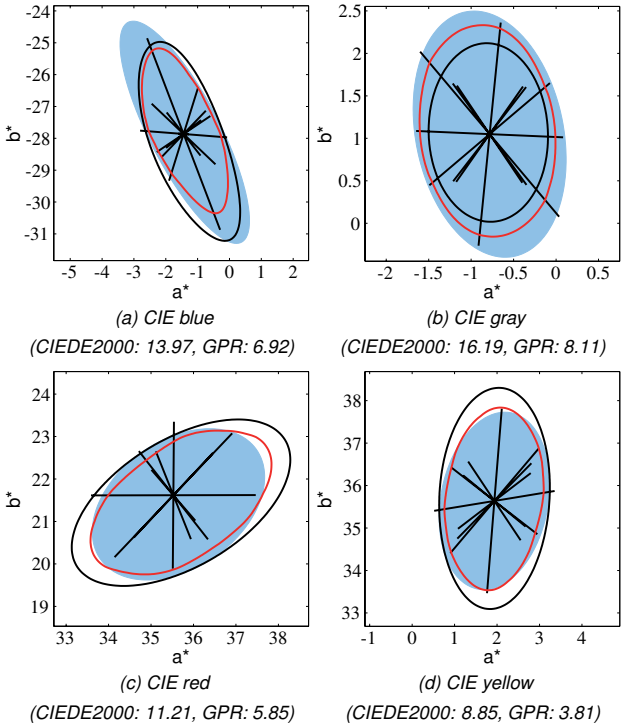

Figure 4. *Four color centers recommended by the CIE for study [17], equivalent to RIT-DuPont color centers 1, 3, 7, and 8. RIT-DuPont suprathreshold ellipsoids (blue) and the corresponding T50 and -T50 color-difference vectors in comparison with CIEDE2000 iso-distance contours (black) and GPR-based iso-distance contours (red). Each point on a contour has a computed distance of 1 to the respective color center. A comparison of STRESS values is shown under each figure. Projections on the a*b*-plane.*

Figure 4 shows RIT-DuPont suprathreshold ellipsoids [4] in comparison with CIEDE2000 and GPR-based iso-distance contours at four color centers. In figure 4 (a) an improvement may not be obvious from the plot — the STRESS index, however, indicates that the GPR approach is more accurate. Figure 4 (b) shows that CIEDE2000 underestimates the perceived distances represented by the longest color-difference vectors, which is reflected by a higher STRESS index for CIEDE2000. Figures 4 (c) and (d) show that CIEDE2000 overestimates the color-differences along the major principal axes of the RIT-DuPont ellipsoids, whereas the GPR-based predictions are closer to the color-difference vectors. This is confirmed by the STRESS values: the GPR-based predictions are more accurate in both cases.

Figure 5 provides another comparison of CIEDE2000 and

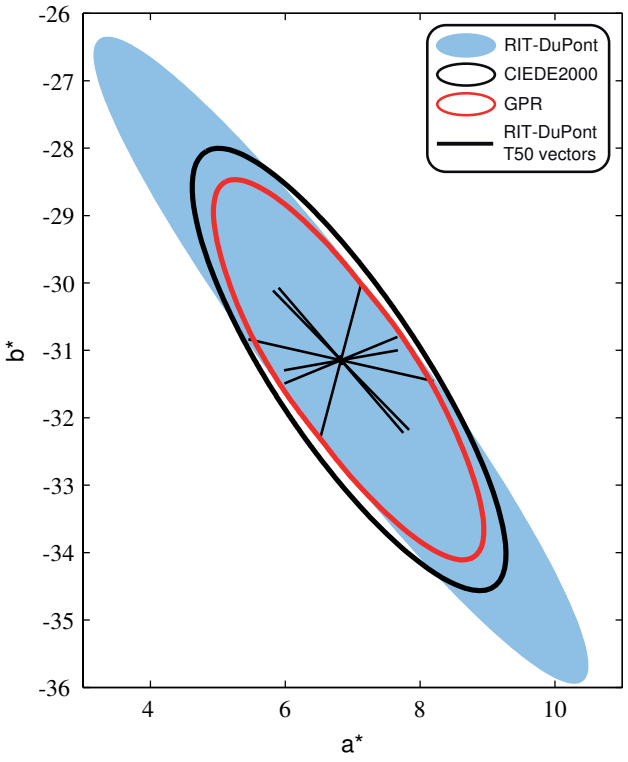

Figure 5. *RIT-DuPont color center 17 projected on the CIELAB a*b*-plane. Suprathreshold ellipsoid (blue) fitted to the RIT-DuPont T50 and -T50 vectors compared to CIEDE2000 iso-distance contour (black) and GPR-based iso-distance contour (red). STRESS values: CIEDE2000 4.50, GPR 2.41.*

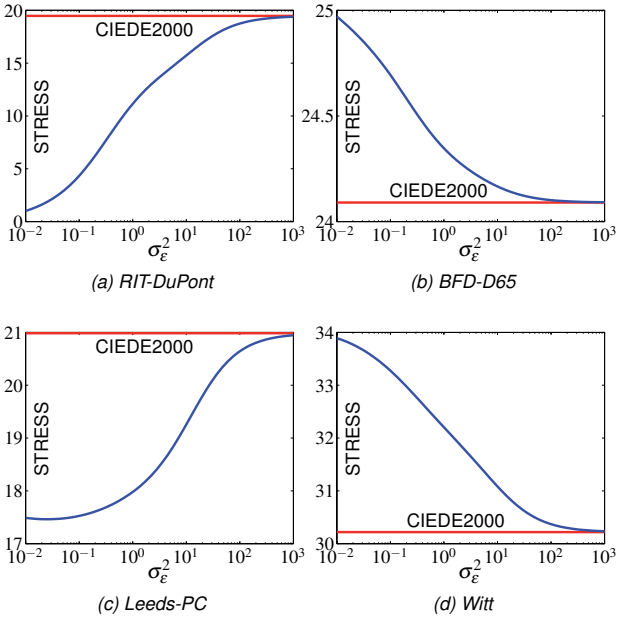

Figure 6. *STRESS index (blue curve) under varying noise variance σ_ε^2 for different sets of visual data. The CIEDE2000 STRESS index for the respective dataset is shown in red. Note the differently scaled y-axes.*

the corresponding GPR-based predictions at color center 17 of the RIT-DuPont data. We use this particular example because there is no color-difference vector with a similar direction as the longest principle axis of this suprathreshold ellipsoid [14] (note that the figure shows a projection on the a*b*-plane). Consequently, even though the GPR-based iso-distance contour agrees less with the RIT-DuPont ellipsoid, this does not mean that it agrees less with the visual data. In fact, along the major principle axis the method computes a trade-off between the CIEDE2000 mean function and the RIT-DuPont color-difference vectors. Along the minor principal axis the GPR-based predictions agree with the color-difference vectors as well as the suprathreshold ellipsoid, whereas the CIEDE2000 predictions differ somewhat from the visual data.

To illustrate the fact that there are many reasonable parameter combinations (see section "Adjusting the Parameters"), figure 6 shows the prediction performance of our method on different sets of visual data under increasing noise variance σ_ε^2. It is evident that a higher noise variance decreases the prediction accuracy on the RIT-DuPont training data, because the data are considered less reliable with increasing noise. In all cases the STRESS index converges against that of the mean function (CIEDE2000) for the respective dataset, because the influence of the correction term (eq. (8)) drops with increasing noise. This causes the STRESS index to decrease for BFD-D65 and Witt, and to increase for RIT-DuPont and Leeds-PC. Choosing a particular σ_ε^2 can be seen as choosing a trade-off between high accuracy on the training data and good generalization ability.

A drawback of the GPR-based method is the computational effort required to determine a single color difference. It increases with increasing number n of observations, because the covariance function needs to be evaluated n times to calculate the vector $\mathbf{k}^*$ from eq. (8). Fortunately, the vector $(\mathbf{K}+\sigma_\varepsilon^2\mathbf{I})^{-1}(\mathbf{v}-\mathbf{m})$ can be precalculated — what remains is the evaluation of an inner product. In summary, the computational complexity should be sufficiently low to improve color-difference formulas at specific color centers with a few additional observations.

Conclusions

With the proposed Gaussian process approach for the improvement of color-difference formulas we could significantly increase the prediction accuracy of CIEDE2000 on the RIT-DuPont dataset. The predictions for the BFD-D65, Leeds-PC, and Witt datasets, which were not used to train our Gaussian process model, did not differ from CIEDE2000 on a statistically significant level. It should be investigated whether the prediction accuracy decreases if the training and the test data were obtained by different psychophysical methods (pair comparison or gray-scale method).

Although the computational complexity increases rapidly with the amount of training data, it seems feasible to systematically improve color-difference formulas at particularly important color centers.

As the covariance function k represents our estimated model of the visual data, we hope that the prediction accuracy on unknown data can be improved with an optimized covariance function in the future. The same applies to the parameters of the covariance function. Further research is necessary to find an effective optimization approach to adjust these parameters to the visual data.

Acknowledgments

We thank Roy Berns and Shizhe Shen for their support and the Deutsche Forschungsgemeinschaft (German Research Foundation) for financing our research.

References

[1] DIN6176. Farbmetrische Bestimmung von Farbabständen bei Körperfarben nach der DIN-99-Formel. Technical report, DIN, Deutsches Institut für Normung e.V., 2001.

[2] G. Cui, M. R. Luo, B. Rigg, G. Roesler, and K. Witt. Uniform colour spaces based on the DIN99 colour-difference formula. *Color Research and Application*, 27(4):282–290, 2002.

[3] R. S. Berns, D. H. Alman, L. Reniff, G. D. Snyder, and M. R. Balonon-Rosen. Visual determination of suprathreshold color-difference tolerances using probit analysis. *Color Research and Application*, 16(5):297–316, 1991.

[4] M. Melgosa, E. Hita, A. J. Poza, D. H. Alman, and R. S. Berns. Suprathreshold color-difference ellipsoids for surface colors. *Color Research and Application*, 22(3):148–155, 1997.

[5] Committee of the Society of Dyers and Colorists. BS 6923: Method for calculation of small colour differences. Technical report, British Standards Institution, London, UK, 1988.

[6] CIE Publication No. 116. Industrial Colour-Difference Evaluation. Technical report, Central Bureau of the CIE, Vienna, Austria, 1995.

[7] CIE Publication No. 142. Improvement to Industrial Colour-Difference Evaluation. Technical report, Central Bureau of the CIE, Vienna, Austria, 2001.

[8] S. S. Guan and M. R. Luo. Investigation of parametric effects using small colour differences. *Color Research and Application*, 24(5):331–343, 1999.

[9] P. A. García, R. Huertas, M. Melgosa, and G. Cui. Measurement of the relationship between perceived and computed color differences. *Journal of the Optical Society of America A*, 24(7):1823–1829, 2007.

[10] M. Melgosa, R. Huertas, and R. S. Berns. Performance of recent advanced color-difference formulas using the standardized residual sum of squares index. *Journal of the Optical Society of America A*, 25(7):1828–1834, 2008.

[11] M. R. Luo, G. Cui, and B. Rigg. The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research and Application*, 26(5):340–350, 2001.

[12] E. D. Montag and D. C. Wilber. A comparison of constant stimuli and gray-scale methods of color difference scaling. *Color Research and Application*, 28(1):36–44, 2003.

[13] M. Melgosa, R. Huertas, and R. S. Berns. Relative significance of the terms in the CIEDE2000 and CIE94 color-difference formulas. *Journal of the Optical Society of America A*, 21(12):2269–2275, 2004.

[14] S. Shen and R. S. Berns. Evaluating color difference equation performance incorporating visual uncertainty. *Color Research and Application*, 34(5):375–390, 2009.

[15] CIE Publication No. 101. Parametric Effects in Colour-Difference Evaluation. Technical report, Central Bureau of the CIE, Vienna, Austria, 1993.

[16] R. G. Kuehni. Variability in estimation of suprathreshold small color differences. *Color Research and Application*, 34(5):367–374, 2009.

[17] A. R. Robertson. CIE guidelines for coordinated research on colour-difference evaluation. *Color Research and Application*, 3:149–151, 1978.

[18] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

[19] M. R. Luo and B. Rigg. Chromaticity-discrimination ellipses for surface colours. *Color Research and Application*, 11(1):25–42, 1986.

[20] D. H. Kim and J. H. Nobbs. New weighting functions for the weighted CIELAB colour difference formula. In *Proceedings of AIC Colour*, pages 446–449, 1997.

[21] K. Witt. Geometric relations between scales of small colour differences. *Color Research and Application*, 24(2):78–92, 1999.

Author Biography

Ingmar Lissner received his engineering degree in Computer Science and Engineering from the Hamburg University of Technology (Germany) in 2009. He currently is a research assistant and doctoral candidate at the Institute of Printing Science and Technology of the Technische Universität Darmstadt (Germany). His research interests include color perception, color reproduction, and high dynamic range imaging.

Philipp Urban has been head of an Emmy-Noether research group at the Technische Universität Darmstadt (Germany) since 2009. His research focuses on spectral-based acquisition, processing, and reproduction of color images considering the limited metameric and spectral gamut and low dynamic range of output devices. From 2006-2008 he was a visiting scientist at the RIT Munsell Color Science Laboratory. He holds a MS in mathematics from the University of Hamburg and a PhD from the Hamburg University of Technology (Germany).