Influence of random-dot textures on perception of suprathreshold color differences

Rafael Huertas, Manuel Melgosa, and Enrique Hita

Departamento de Óptica, Edificio Mecenas, Facultad de Ciencias, Universidad de Granada, 18071-Granada, Spain

Received November 18, 2005; revised February 16, 2006; accepted March 2, 2006; posted March 24, 2006 (Doc. ID 66079)

We have analyzed the way in which simulated textures made of random dots influence visual suprathreshold color tolerances. We considered 32 textures created by a systematic variation of the following variables: size, number, and color of the dots. Each texture was mapped on the five centers recommended by the International Commission on Illumination (CIE) in 1978 [Color Res. Appl. 3, 149 (1978)]. A panel of five experienced observers determined the experimental tolerances using a CRT color monitor by the method of adjustment. At our observation distance, neither small differences in dot size (1 or 4 pixels) nor sparse number of dots (less than 20% of the surface of the sample) changed the tolerances found for homogeneous samples. For the textures that led to statistically significant differences with respect to homogeneous samples, the parametric factors of CIEDE2000 and CIE94 color-difference formulas were fitted. These simulated textures consistently reduced the color difference perceived in a pair or, equivalently, increased the tolerances, mainly (but not only) lightness tolerance. The results demonstrate that it is not simple to provide a unique set of parametric factors for all the potential textures. © 2006 Optical Society of America

1. INTRODUCTION

It is known that viewing conditions appreciably affect perceived color differences. A color-difference tolerance is the magnitude corresponding to the transition between two categories of judgment concerning color differences. The latest color-difference formula proposed by the International Commission on Illumination (CIE), CIEDE2000, is given as

$$
\Delta E_{00} = \left[ \frac{\Delta L'}{S_L k_L} \right]^2 + \left[ \frac{\Delta C'}{S_C k_C} \right]^2 + \left[ \frac{\Delta H'}{S_H k_H} \right]^2 + \frac{R_T \Delta C'}{S_C k_C},
$$

where $\Delta L'$, $\Delta C'$, and $\Delta H'$ are color differences in lightness, chroma, and hue, respectively, transformed from CIELAB coordinates $L^*, C^*, a^*$, and $b^*$. Meanwhile, $S_L$, $S_C$, and $S_H$ are the weighting functions to improve the lack of uniformity of CIELAB. $R_T$ is the rotation term (different from zero only in the blue region). A complete description of the development of the CIEDE2000 formula can be found in Luo et al. The previous CIE-proposed color-difference formula CIE94 (simpler than CIEDE2000) also contained parametric factors, for the first time in a CIE formula, in a way similar to the Color Measurement Committee (CMC) color-difference formula.

The viewing conditions include, among other parameters, the sample surface structure (texture). The influence of texture on color perception is known and has far-reaching industrial relevance. Nevertheless, few reports on quantitative analyses of this influence have been published, most referring to textile samples. The texture of the samples has not been thoroughly studied in color science.

The three parametric factors introduced in CIEDE2000 and CIE94 are equal to 1 under the so-called reference conditions, which include homogeneous samples. For textile samples the CIE has specifically recommended the use of $k_L=2$, $k_C=1$, and $k_H=1$. These values, accepted and used by several authors, assume that texture in textiles affects only lightness tolerances but not chroma or hue tolerances. The accuracy of this recommendation is not yet well understood, and additional research has been called for. In addition, although some recent publications have clearly shown the strong effect of texture in perceived color differences, there are differences in the parametric factors computed, perhaps induced by the dissimilarities in the textures studied. Furthermore, typically only lightness tolerances for textured samples have been studied, although investigation of chroma and hue tolerances has been claimed. Both of the last-mentioned papers analyzed the effects of a full texture scanned from textile samples. Montag and Berns also included a half-texture created from the full texture scanned.

In summary, the effect of the texture on chroma or hue tolerances has not been investigated, and, according to the recommendation for textiles mentioned above, no effect could be expected. In addition, how a gradual texture affects perceived color differences has not been studied, either. We refer to a gradual texture as the one capable of being progressively modified. Thus, we considered gradual textures, analyzing their impact on the visual suprathreshold color tolerances—specifically on lightness, chroma, and hue tolerances corresponding to...
2. METHODS

A. Experimental Device and Color Measurements

The visual experiment was managed by specific software, which we developed in Borland C++ Builder Professional, version 4.0. This software displayed a test in a calibrated Samsung SyncMaster 21 in. (1 in. = 2.54 cm) CRT color monitor, connected to a graphics card NVIDIA RIVA TNT2 Model 64 Pro 32 MB in a compatible-IBM personal computer. The active display of the monitor was $360 \times 270$ mm, and the resolution was $1024 \times 768$ pixels. The software generated the textures, controlled the adaptation times, displayed the tests, recorded the observer’s responses, and also allowed the colorimetric calibration of the CRT monitor. Nowadays, color-information exchanges become more and more important, and CRT devices have been widely used in areas of product design and quality control in many textile companies. In color research, CRT color monitors are widely used, giving results similar to those of experiments using objects.\textsuperscript{16,17} CRT color monitors have also been used in previous works studying texture effects.\textsuperscript{12,15} Furthermore, Lee and Sato\textsuperscript{18} concluded that perception of real textiles and simulated textiles was similar, despite the geometric and space differences between them.

Our visual experiment consisted basically of a pair-comparison test [namely, the method of adjustment (MOA)], similar to the constant-stimuli method used by other authors,\textsuperscript{19,20} with the arrangement shown in Fig. 1. The names and sizes (in millimeters) in Fig. 1 were not included in the test. The MOA is basically the constant-stimuli method, but the observer changes one of the samples of the comparison pair until its difference is accepted as equal to or just higher than the perceived one in the anchor pair. Visual assessment of color differences is usually performed by two methods: the constant-stimuli method\textsuperscript{21} and the gray-scale comparison method.\textsuperscript{22} Montag and Wilber\textsuperscript{23} investigated the validity and precision of these two methods and concluded that, although the results followed the same trends, differences resulted with the two techniques. These authors concluded that, on the basis of comparisons of the validity and precision of the results, the constant-stimuli method was preferable. A variation of this method has been employed in the current work.

In our case the comparison pair was made of textured color samples, and the anchor pair was made of homogeneous gray samples. Except for the texture of the comparison pair, the reference conditions given by the CIE\textsuperscript{5} were followed as closely as possible, so that we investigated only the effect of texture on visual tolerances and computed the appropriate parametric factors for these textures. A black outline 1 pixel wide framed each sample, which subtended 6.2° from the position of the observer at 770 mm. Both pairs were displayed over a neutral background and a white surround subtending 1°. Table 1 shows the coordinates measured and their uncertainties (estimated from the standard deviation of three independent measurements) for the anchor pair, the background, and the surround under the CIE 1964 Standard Supplementary Observer. Therefore, the anchor pair has a fixed color difference of 1.57 CIELAB units, a suprathreshold color difference. The white surround corresponds to the white of the monitor (maximum values of the three guns).

We computed the tolerances as the color differences in the comparison pair, which were established by observers as just above the one perceived in the anchor pair. These

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig1}
\caption{Arrangement of sample pairs in the MOA on a CRT monitor as used in our experiment (distances are in millimeters).}
\end{figure}
tolerances were worked out by the CIELAB color-difference formula\textsuperscript{24} for the statistical analyses and by the CIEDE2000 and CIE94 formulas for the computation of the parametric factors of CIEDE2000 and CIE94, respectively.

It is well known that the reliability of the results is critical in psychophysical experiments. Observer accuracy and repeatability were tested to check the reliability of the results. The performance factor (PF/3)\textsuperscript{25} has been widely used as an indicator for the observer’s accuracy and also for the performance of color-difference formulas in comparison with visual results.\textsuperscript{26} PF/3 allows a statistical comparison of two data sets by means of combining three measures of fit: gamma factor $\gamma$, $CV$, and $V_{AB}$.

The factor $\gamma$ is a statistical measure proposed by Coates \textit{et al.}\textsuperscript{27} and represents the antilogarithm of the root-mean-square value of the decimal logarithm of the ratios ($\Delta E_i/\Delta V_i$):

$$
\log_{10}(\gamma) = \left[ \frac{1}{N} \sum_{i=1}^{N} \left( \log_{10}(\Delta E_i/\Delta V_i) - \log_{10}(\Delta E_i/\Delta V_i) \right)^2 \right]^{1/2}.
$$

\(V_{AB}\), proposed by Schultze,\textsuperscript{28} and \(CV\), proposed by Coates \textit{et al.}\textsuperscript{27} are coefficients of variation with appropriate overall scale factors:

$$
V_{AB} = \left[ \frac{1}{N} \sum_{i=1}^{N} \left( \frac{\Delta E_i - F \Delta V_i}{\Delta E_i F \Delta V_i} \right)^2 \right]^{1/2},
$$

with

$$
F = \left[ \frac{\sum_{i=1}^{N} \Delta E_i/\Delta V_i}{\sum_{i=1}^{N} \Delta V_i/\Delta E_i} \right]^{1/2},
$$

$$
CV = 100 \left[ \frac{1}{N} \sum_{i=1}^{N} \left( \frac{\Delta E_i - f \Delta V_i}{(\Delta E_i)^2} \right)^2 \right]^{1/2},
$$

$$
f = \frac{\sum_{i=1}^{N} \Delta E_i \Delta V_i}{\sum_{i=1}^{N} \Delta V_i^2}.
$$

The computation of PF/3 is given as

$$
PF/3 = \frac{100}{3} \left[ (\gamma - 1) + V_{AB} + CV \right].
$$

PF/3 has low values for minor differences between the two data sets. For perfect agreement between the two data sets, the PF/3 should be zero (that is, $\gamma = 1$, $V_{AB} = 0$, and $CV = 0$). Observer accuracy, or consensus, represents the deviation between each individual and the mean visual result of a panel. The observer repeatability, also called consistency, represents the variation of the visual assessments of a particular observer and is computed as the deviation between each individual result and the mean of his/her replications. On the other hand, the performance of a color-difference formula is obtained as the deviation between visual results and the corresponding computed color differences.

A Photo Research PR704 spectroradiometer was used to measure the color coordinates of the samples and also to calibrate the monitor following the method given by Díaz \textit{et al.}\textsuperscript{29} and Jiménez \textit{et al.}\textsuperscript{30} The distance between the spectroradiometer and the displayed image was 770 mm, with 0° from the perpendicular to the sample, the same geometry as visual observations. In the spectroradiometric measurements, 1° field size and the CIE 1964 Standard Supplementary Observer were used. Three independent measurements of the tristimulus values were performed for each center, the background, and the white of the monitor (the surround), used as reference white for transformations to CIELAB.\textsuperscript{31} The tristimulus values were computed as the average of these three measurements, and their uncertainties were computed as the standard deviations. Subsequently, we computed the CIELAB coordinates and their corresponding uncertainties by quadratic propagation of random uncertainties from tristimulus values to CIELAB coordinates, following the method proposed in a previous work.\textsuperscript{32} Specifically, we computed the uncertainties (termed $\sigma$) in the coordinates $L^*$, $a^*$, and $b^*$ from the uncertainties in the tristimulus values, following the next general formula:

$$
\sigma^2(V) = \left( \frac{\partial V}{\partial \alpha} \right)^2 \sigma^2(\alpha) + \left( \frac{\partial V}{\partial \beta} \right)^2 \sigma^2(\beta) + 2 r_{\alpha\beta} \left( \frac{\partial V}{\partial \alpha} \right) \left( \frac{\partial V}{\partial \beta} \right) \sigma(\alpha) \sigma(\beta),
$$


<table>
<thead>
<tr>
<th>Stimulus</th>
<th>$L_{10}$</th>
<th>$a_{10}$</th>
<th>$b_{10}$</th>
<th>$X_{10}$</th>
<th>$Z_{10}$</th>
<th>$Y_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anchor pair</td>
<td>Sample 1</td>
<td>52.95 ± 0.03</td>
<td>-0.15 ± 0.19</td>
<td>-0.57 ± 0.15</td>
<td>22.39 ± 0.02</td>
<td>33.14 ± 0.04</td>
</tr>
<tr>
<td></td>
<td>Sample 2</td>
<td>51.43 ± 0.03</td>
<td>-0.52 ± 0.19</td>
<td>-0.65 ± 0.14</td>
<td>20.86 ± 0.03</td>
<td>31.07 ± 0.04</td>
</tr>
<tr>
<td>Surround</td>
<td>100.00 ± 0.11</td>
<td>0.0 ± 0.04</td>
<td>0.0 ± 0.3</td>
<td>106.8 ± 0.3</td>
<td>155.6 ± 0.4</td>
<td>115.2 ± 0.3</td>
</tr>
<tr>
<td>Background</td>
<td>54.42 ± 0.02</td>
<td>0.19 ± 0.18</td>
<td>-0.35 ± 0.15</td>
<td>23.93 ± 0.01</td>
<td>35.12 ± 0.04</td>
<td>25.767 ± 0.012</td>
</tr>
</tbody>
</table>

*Uncertainties were obtained from standard deviations of three independent measurements.*

Table 1. Color Coordinates of the Anchor Pair, Background, and Surround*
where $V$ is the dependent variable ($L^*, a^*, b^*$ in our case), $\alpha$ and $\beta$ are the independent variables ($X, Y, \text{and } Z$ in our case), and $r_{\alpha\beta}$ is the correlation coefficient between $\alpha$ and $\beta$. In general, the correlation coefficients between tristimulus values differ from zero and can be computed following the method proposed by Gardner and Frenkel.³³ In the previous work, it was concluded that the random uncertainties, estimated by quadratic propagation, may be negligible in most of the applications but depend strongly on the instruments used to measure the color and the samples measured. In the present work, we estimated the uncertainties in the color measurement of the anchor pair, background, and surround to give an idea of the precision in our experimental device and instruments.

B. Textures

As we used a CRT color monitor to perform the visual experiments, the texture of the samples was simulated (real textured samples were not studied), as has been done in previous works.¹²,¹⁵ However, in our work, considering textures capable of being gradually changed, we undertook a more detailed analysis of their effects over the perceived color differences. In particular, our simulated textures were made of randomly distributed dots over a homogeneous sample displayed in a CRT color monitor. Below, we call the homogeneous sample the “background sample,” while “texture” refers to the dots mapped over the background sample.

It is controversial how to compute the color difference in the case of textured samples and even more so in the case of images. In our case, one possibility would be to average the color coordinates ($L^*, a^*, b^*$) of texture and background but in a proper way, considering the variables of the texture. In this manner, the effect of the texture would be included in the total color coordinates of the samples, and then the color difference would be computed. However, in our work, we have treated the texture as a parametric effect, which must be included in the computation of the color differences through the parametric factors but not in the color coordinates. Thus, in the calculation of the color differences, only the color coordinates of the background samples were taken into account.

The experiment was carried out with the five centers recommended by the CIE in 1978 for the study of color differences¹⁰ as the background samples. Table 2 provides the proposed CIELAB values of the five CIE centers and the measured CIELAB values with their corresponding uncertainties. The differences between the proposed and the measured CIELAB values are due to the calibration of the monitor, on which it was not possible to display exactly the desired color.

As in previous publications, two main aspects in the texture were considered: the color difference between texture and background sample and the percentage of background sample covered by texture (this latter aspect controlled by dot size and number of dots). In our work the simulated texture was composed of dots of 1 or 4 pixels (the 4 pixels forming a square). Depending on dot size, different percentages of surface coverage were possible by changing the number of dots. We considered the cases of 5%, 20%, and 50% coverage of the background sample. The random dots differentiate the background sample by color in three ways: black dots; dots with different lightness and same chroma and hue as the background sample; or dots with different chroma and same lightness and hue as the background sample. These three cases led to five types of texture (termed A to E) differentiated by dot color. Both lightness and chroma are important in differentiation of background samples and textures.³⁶ Figure 2 shows some examples of simulated textures mapped over different CIE centers.

The three variables considered in the texture—size, type (or color of the dots), and covered surface—were systematically modified and combined, providing 32 different simulated textures, as shown in Table 3. For simplicity, this table does not list the size and therefore really shows 16 textures, but each one is possible with the two sizes. Type E was called absolute texture, while the other types (A, B, C, and D) were termed relative textures because the color of the dots depended on the color of the background sample (labeled with the subscript BS in Table 3). Coverage by the dots of 80% of the surface was also considered only in the case of absolute texture. For relative textures, more than 50% of the surface covered is meaningless because it is equivalent to interchanging the terms background sample and texture. These simple textures enabled a rapid investigation of how texture affects color tolerances and also how the influence was changed by gradually modifying the textures, through changing the texture variables defined above.

To evaluate how textures influence color differences, a reference experiment using solid samples was also conducted under the same viewing conditions, with the same anchor pair and by the same observers. In this experiment the comparison pairs consisted of homogenous samples with the same centers used as background.

| Table 2. Proposed and Measured Color Coordinates of the Five CIE 1978 Centers |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Center | $L_{10}$ | $a_{10}$ | $b_{10}$ | $C_{ab,10}$ | $h_{ab,10}$ | Measured Coordinates |
| $L_{10}$ | $a_{10}$ | $b_{10}$ | $C_{ab,10}$ | $h_{ab,10}$ |
| 1—Gray | 62 | 0 | 0 | 0 | 0 | 65.06±0.04 | −0.1±0.2 | −0.49±0.17 | 0.50±0.18 | −1.8±0.4 |
| 2—Red | 44 | 37 | 23 | 43.57 | 31.87 | 46.66±0.03 | 35.26±0.18 | 20.71±0.12 | 40.89±0.17 | 0.53±0.003 |
| 3—Yellow | 87 | −7 | 47 | 47.52 | 98.47 | 89.73±0.08 | −6.6±0.3 | 45.0±0.2 | 45.52±0.2 | 1.71±0.007 |
| 4—Green | 56 | −32 | 0 | 32 | 180 | 59.79±0.03 | −29.42±0.19 | −0.4±0.15 | 29.42±0.19 | −3.12±0.005 |
| 5—Blue | 36 | 5 | −31 | 31.40 | 279.16 | 39.73±0.02 | 3.84±0.17 | −30.03±0.14 | 30.27±0.14 | −1.44±0.005 |
samples (Table 2). The tolerances found were considered reference tolerances for the calculation of the parametric factors.

C. Observations

Visual observations combined the 32 textures with the five CIE centers used as background samples to the extent possible, since it was not possible to combine all the textures with all the centers. For example, for the gray center it was not possible to map the texture type D because dots with chroma less than 0 are meaningless. A few other combinations were also not possible because either the colors needed were beyond the gamut of the monitor or they were not within the required accuracy owing to the calibration of our CRT.

For each achievable combination of texture and center, a set of 40 samples was prepared for the comparison pair. These samples differed progressively only in one CIELAB coordinate of the background: \( L^* \), \( C_{ab}^* \), or \( h_{ab} \). The progression was from no difference in the backgrounds of the comparison pair to differences of about 16 CIELAB units in small steps of about 0.4 CIELAB unit. The steps were exactly a mean of 0.30 unit in \( L^* \) (with a minimum of 0.19 and maximum of 0.42), a mean of 0.51 unit in \( C_{ab}^* \) (with a minimum of 0.47 and maximum of 0.57), and 0.52 unit (in degrees) in \( h_{ab} \) (with a minimum of 0.39 and maximum of 0.64). Then the texture was mapped to the 40 background samples. In the case of relative textures (the color of the dots was linked to the background), a different color of the dots corresponds to each of the 40 samples prepared.

Because of the calibration, it was not always possible to have a color difference in the comparison pair exclusively in one CIELAB coordinate, but, in the worst case, 75% of the total CIELAB color difference was in one coordinate: \( \Delta L^* \), \( \Delta C_{ab}^* \), or \( \Delta h_{ab} \). On the average, 92.85% of the color differences were only in \( \Delta L^* \), \( \Delta C_{ab}^* \), or \( \Delta h_{ab} \), with a standard deviation of 4.56%. Thus, the tolerances found may be considered only in lightness, chroma, or hue, below called an attribute of the tolerance.

![Fig. 2. Examples of some simulated textures mapped over different background samples. The upper row corresponds to absolute textures (type E), while the lower row corresponds to relative textures (types A, B, C, and D).](image-url)

<table>
<thead>
<tr>
<th>Texture Type</th>
<th>Dots’ Color</th>
<th>Sample Covered (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A(^a)</td>
<td>( L_{10} = L_{10,BS} + 10 )</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>( C_{ab,10} = C_{ab,10,BS} )</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>( h_{ab,10} = h_{ab,10,BS} )</td>
<td>50</td>
</tr>
<tr>
<td>B</td>
<td>( L_{10} = L_{10,BS} - 10 )</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>( C_{ab,10} = C_{ab,10,BS} )</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>( h_{ab,10} = h_{ab,10,BS} )</td>
<td>50</td>
</tr>
<tr>
<td>C</td>
<td>( L_{10} = L_{10,BS} )</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>( C_{ab,10} = C_{ab,10,BS} + 15 )</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>( h_{ab,10} = h_{ab,10,BS} )</td>
<td>50</td>
</tr>
<tr>
<td>D</td>
<td>( L_{10} = L_{10,BS} )</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>( C_{ab,10} = C_{ab,10,BS} - 15 )</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>( h_{ab,10} = h_{ab,10,BS} )</td>
<td>50</td>
</tr>
<tr>
<td>E(^c)</td>
<td>Black</td>
<td>5</td>
</tr>
</tbody>
</table>

\(^a\)The two sizes (1 or 4 pixels) of the random dots are not considered. Color of the background sample is labeled with the subscript BS.

\(^b\)Types A, B, C, and D are termed relative textures.

\(^c\)Type E is termed absolute texture.

There were two possibilities to increase the color difference in the comparison pair for the three attributes of tolerances: to increase or reduce the values of the selected CIELAB coordinate with regard to the center. We differentiated these two cases by means of the variable sense of the tolerance.

The combinations of centers, textures, attributes of the tolerance, and senses of the tolerance were scheduled within different experimental sessions, which prior to the visual assessments were created and saved as files of the software mentioned above. Each experimental session grouped five visual assessments ready to be performed. The five assessments prepared were randomly blended. Again, because of the CRT calibration, all the combinations of centers, textures, and attributes and senses of the
tolerance were not possible, and eventually 645 assessments were prepared for each observer.

The observations, conducted in a completely dark room to eliminate the influence of ambient illumination, were preceded by 2 min of darkness adaptation just after the observer opened any file with the software to begin a session. Before each visual assessment of the session the observer was adapted for another 2 min to the background and the surround mentioned above, and no more than five assessments were performed in one session.

The observer’s task was to establish a color difference in the comparison pair just above the one perceived in the anchor pair, which was constant during all the experiments. The observers were instructed to judge an overall impression of the textured samples, not only the backgrounds. One of the samples of the comparison pair was fixed while the other, randomly selected up or down by software in each session (to avoid orientation effects), was changed to establish the color difference. The observer could gradually increase or reduce (by clicking the left or right button of the mouse, respectively) the color difference in the comparison pair until accepting it (by clicking the central button of the mouse). When observers clicked the left or right button, one sample in the comparison pair was replaced by one of the sets of 40 samples prepared previously, which progressively differed from the other sample of the comparison pair only in one coordinate. This method of change of the comparison pair and comparison with an anchor pair is termed the MOA. For most assessments, three replications were performed by a panel of five experienced observers (three females and two males, ranging from 29–45 yr of age) with normal color vision, carrying out about 1600 visual assessments each. Overall, 7706 suprathreshold visual tolerances were recorded.

D. Computations of the Parametric Factors

To fit the parametric model for CIEDE2000 and CIE94, the visual tolerances were computed as the color differences in the comparison pairs (established by the observer) with CIEDE2000 and CIE94 formulas, respectively. All the visual tolerances obtained correspond to the same perceived color difference because the anchor pair was fixed throughout the experiment. This perceived color difference is also the same for the tolerances of the homogeneous samples in the reference experiment. The reference tolerances of the homogeneous samples were computed using \( k_L = 1 \), \( k_C = 1 \), and \( k_H = 1 \) in both color-difference formulas.

Any color-difference formula tries to provide a computed value as near as possible to the visual perception. Therefore the color differences computed with CIEDE2000 and CIE94 must be equal for textured as well as homogeneous samples, since both resulted from a comparison with the same anchor pair. This is achieved by means of the parametric factors, which in this way take into account the influence of the textures on the perceived color differences. We used the lightness tolerances to compute \( k_C \), chroma tolerances to compute \( k_C \), and hue tolerances to compute \( k_H \). The factors were worked out so that the differences between the tolerances corresponding to textured samples and the reference tolerances were minimum. The minimization of the differences through the adequate parametric factor was made by numeric computation with the modified simplex method.

3. RESULTS AND DISCUSSION

The set of all tolerances, computed in CIELAB units, constituted our experimental data set. These data were the observations (in a statistical sense) or the dependent variable in each analysis. Through the SPSS 9.0.1 software at 95% confidence limit, we statistically analyzed the tolerances as a function of the independent variables or factors: observer, center, attribute of the tolerance, sense of the tolerance, type of texture, size of dots, and percentage of covered surface. The variable’s residuals fulfilled the hypothesis of normality, independence, and homocedasticity. Thus, we applied a statistical factorial multivariate model with replication (we had three replications), fixed effects (the factors had fixed levels), and no balanced design (different number of observations for each level of the factors). Statistically, the levels of a factor are the different values that this factor (or independent variable) can reach, and in our experiment these values were fixed, not randomly chosen.

We checked the reliability of the results by computing the observer accuracy and repeatability using the performance factor PF/3. In our experiment the observer accuracy was 42.3, while the mean of the observer repeatability was 18.9 with a standard deviation of 3.0. As might be expected, the repeatability was somewhat lower than accuracy. Similar results were reported in previous studies using the gray-scale method, where the values in PF/3 units for observer accuracy were 40.025 and 24.039 and for observer repeatability 19.0. On the other hand, analyzing the standard deviation between observers’ tolerances (interobserver variability), we found that it was higher for the textured samples (1.87, 2.75, and 1.74 for lightness, chroma, and hue tolerances, respectively, for the textures with statistically significant differences with regard to the reference experiment, which will be discussed below) than in the reference experiment (0.882, 2.01, and 1.11 for lightness, chroma, and hue tolerances, respectively). This means that the agreement among the observers was lower when the samples had texture. Also, to determine whether the results showed observers’ learning effects, we studied the differences among the three replications. The Mann–Whitney test indicated there were no statistical differences among replications (\( p \) value of 0.760), and thus we gave the same weight to the three replications.

It was also found that the two senses of the tolerance did not statistically differ at a \( p \) value of 0.695. That is, the observers did not distinguish between two color differences obtained by either increasing or decreasing one CIELAB coordinate with regard to the center. The same result was found for dot size (1 or 4 pixels), at a \( p \) value of 0.230 with Tamhane’s test. Therefore, these two variables (tolerance sense and dot size) were not further considered.

Observers differed significantly in tolerances at a \( p \) value lower than 0.001. However, the important fact was that there were no interactions between the variable observer and the other variables. This means that all the
observers followed the same trend for the different centers, textures, and attributes of tolerances. For example, if one observer had the lowest tolerances, this happened for all centers, with all textures, and for tolerances in lightness, chroma, or hue. Because there were no interactions, our results can be applied theoretically to any of our observers, and henceforth the variable observer was disregarded. Five observers are not many in order to generalize our results to any given person; however, the good agreement among them all is an important reason to regard our results as an approximation to the behavior of the human visual system in the influence of random-dot textures on perception of suprathreshold color differences.

In the case of the five textures studied, the same result was also found: statistically significant differences among centers (p value lower than 0.001) and no interaction with the rest of the variables, with one exception. One important interaction between centers and attributes of the tolerance was found. Specifically, for the red CIE center the tolerances in lightness (ΔL) were lower than expected, as shown in Fig. 3. This result may be explained bearing in mind that the lightness of the red center and that of the background were quite similar (see Tables 1 and 2), and by the crispening effect the perceived color differences in lightness were overestimated; thus the visual tolerance was lower for this center. However, surprisingly, the same effect did not appear with respect to the green CIE center. No further interactions between centers and the rest of the variables allowed us to generalize our results theoretically to any place of the color space. The independent variable center was also eliminated further on.

Analyzing the percentage of covered surface, we found that for absolute textures the tolerances strongly increased with the percentage of covered surface, as might be expected, but for relative texture this is not so clear (Fig. 4). Only the case of texture covering 5% of the surface of the sample statistically did not differ from the homogeneous samples (p value 0.362 with Tamhane’s test).

With respect to the type of texture, Tamhane’s test revealed no statistical differences between types A and B (p value 0.117) or between types C and D (p value 1.000).

For the relative textures the tolerances differed slightly more from reference tolerances when dot differentiation increased in lightness (type A) or chroma (type C) over the background sample than when reduced in lightness (type B) or chroma (type D). The more the tolerances differed from reference tolerances, the more effect the texture had on the perceived color differences. Thus, we considered only the relative texture types A and C, as well as the absolute texture E, in the forthcoming computation of parametric factors.

In short, the textures for which tolerances statistically differ from those of the homogeneous samples were types A and C with dots covering 20% and 50% and type E with dots covering 20%, 50%, and 80% of the surface. For each of these textures we calculated the parametric factors of CIEDE2000 and CIE94 color-difference formulas.

The visual tolerances from the reference experiment, performed with the homogeneous five CIE centers, are shown in Table 4. Listed in rows are lightness, chroma, and hue tolerances (CIELAB units) as well as chroma and hue tolerances with chroma correction (CIE94 units). The values correspond to the average for observers (standard deviation in parentheses). The last column shows the average and standard deviation for centers. These tolerances were taken as reference tolerances to compare with the tolerances of textured samples and to compute the parametric factors corresponding to each texture.

For the set of all tolerances, statistically significant differences (p value lower than 0.001 with Tamhane’s test) existed between chroma and lightness or chroma and hue tolerances. Between lightness and hue tolerances, Tamhane’s test revealed no statistically significant differences at a p value of 0.292. These results correspond to the tolerances computed in CIELAB units. The corrections to chroma and hue tolerances with chroma introduced in CIE94 and CIEDE2000 are the most important corrections to the CIELAB formula. These corrections reduce the differences between the tolerances (now in CIE94 units), mainly in chroma tolerances, as reflected in Table
Table 4. Visual Tolerances for the Homogeneous Five CIE Centers

<table>
<thead>
<tr>
<th>Tolerance</th>
<th>1—Gray</th>
<th>2—Red</th>
<th>3—Yellow</th>
<th>4—Green</th>
<th>5—Blue</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta L_{10}^*$</td>
<td>3.10 (0.58)</td>
<td>2.70 (0.96)</td>
<td>3.43 (1.14)</td>
<td>3.15 (1.23)</td>
<td>3.35 (1.12)</td>
<td>3.15 (0.28)</td>
</tr>
<tr>
<td>$\Delta C_{a,10}^*$</td>
<td>—</td>
<td>5.42 (2.25)</td>
<td>4.93 (1.71)</td>
<td>5.50 (1.89)</td>
<td>4.62 (2.20)</td>
<td>5.12 (0.41)</td>
</tr>
<tr>
<td>$\Delta H_{b,10}^*$</td>
<td>—</td>
<td>3.73 (1.60)</td>
<td>3.32 (1.04)</td>
<td>3.08 (0.76)</td>
<td>2.65 (0.98)</td>
<td>3.20 (0.45)</td>
</tr>
<tr>
<td>$\Delta C_{a,b}^* / S_{C}^*$</td>
<td>—</td>
<td>1.83 (0.76)</td>
<td>1.59 (0.55)</td>
<td>2.49 (0.91)</td>
<td>1.93 (0.94)</td>
<td>1.96 (0.38)</td>
</tr>
<tr>
<td>$\Delta H_{a,b}^* / S_{H}$</td>
<td>—</td>
<td>2.26 (0.97)</td>
<td>1.94 (0.61)</td>
<td>2.13 (0.52)</td>
<td>1.80 (0.67)</td>
<td>2.03 (0.20)</td>
</tr>
</tbody>
</table>

$^a$Average tolerances for the observers with the standard deviation in parenthesis. For the achromatic center (gray), only lightness tolerances were considered. The last column corresponds to the average and standard deviation over centers.

$^b$Chroma tolerance with chroma correction (CIE94 units).

$^c$Hue tolerance with chroma correction (CIE94 units).

4 for the reference tolerances in CIE94 units. In any case, the differences in CIE94 units continued to be statistically significant at a p value lower than 0.001.

Somewhat surprising, the average of the reference tolerances was 3.82 CIELAB units or 2.38 CIE94 units, when the color difference in the anchor pair was 1.57 in both CIELAB and CIE94 units. However, this difference cannot be attributed to texture because the reference experiment used homogeneous samples. CIELAB is not a uniform space, and the color of the five CIE centers must be considered while the anchor pair is achromatic. Therefore, in CIE94 units the difference is lower. In addition, many other considerations affect the perception of achromatic color differences, as demonstrated by Carter. Among all the effects, crispening appears to be important, leading observers to overestimate the color difference in the anchor pair because its lightness approximates the lightness of the background. On the other hand, the observers tried to achieve a color difference in the comparison pair just above the one perceived in the anchor pair. The difference between reference tolerances and the anchor-pair color difference is probably due to a combination of these reasons. A similar result was reported in previous works.

As explained in Section 2, the parametric factors were fitted in such a way as to minimize the average difference between the reference tolerances (given in Table 4) and the tolerances of the textured samples. For each type of texture and percentage of surface covered, the differences were computed by subtracting tolerances of the same observer, center, and attribute of tolerance and then averaging over observers and centers. Table 5 lists the tolerances in CIELAB units for the textures mentioned above. These tolerances are the mean of the two tolerance senses, the two dot sizes, the five observers, and the five centers. The standard deviation of the average is also shown in parentheses. Comparing Tables 4 and 5, we find that textures increase the tolerances, as might be expected. However, the tolerances in Table 5 were not employed to compute the parametric factors, which are not defined in the CIELAB color-difference formula. Figure 5 shows the computed parametric factors for CIEDE2000 and CIE94 for the texture types A and C with the surfaces 20% and 50% and type E with the surfaces 20%, 50%, and 80%.

Minor differences were found between the parametric factors for CIEDE2000 and CIE94. In both formulas, parametric factors greater than 1.0 indicate that texture increased the tolerances or, equivalently, reduced the perceived color differences. Figure 5 shows that texture consistently reduces the color difference perceived in a pair. The human visual system is known to be more sensitive to luminance contrast than to chromatic contrast. However, Li and Lennie reported that, for textured surfaces similar to the ones used in this work, chromatic and luminance differences in the texture resulted in similar perception but at the threshold level. Our results showed differences in the perception of luminance and chromaticity in suprathreshold differences. The greater sensitivity to luminance may explain why the influence of texture type A (which had dots with different lightness) was stronger than the influence of texture type C (which had dots with different chroma) or, equivalently, why for texture type A the parametric factors were higher than the ones for texture type C. The same reason might explain why $k_L$ was usually higher than $k_C$ or $k_H$.

As in the recommendation for textile samples, $k_L$ increased up to 2.1 in the case of texture type E (absolute). However, the $k_C$ and $k_H$ factors were also higher than 1.0 for some textures, indicating that texture affects not only lightness tolerances but chroma and hue tolerances, too. Even $k_L$ had different values, depending on the texture, and not a fixed value of 2.0. This implies that the recommendation of $k_L=2.0$, $k_C=1.0$, and $k_H=1.0$ for textiles samples is not completely exact.

The absolute texture (type E) induced the highest values of parametric factors; that is, absolute texture strongly influences the perceived color differences, as might be expected. For this texture the value of the parametric factors increased with the amount of texture, considering the percentage of covered surface as a measurement of the quantity of texture.

The influence of the relative textures A and C was more important than lightness and chroma tolerances, respectively. That is, the $k_L$ and $k_C$ factors are the highest for textures A and C, respectively. It appears that the attribute that differentiates the dots (lightness for texture A and chroma for texture C) was the attribute of the tolerance most influenced by the texture.

The parametric factors $k_L$, $k_C$, and $k_H$ shown in Fig. 5 correspond to color differences exclusively in lightness, chroma, or hue. That is, the parametric factors have been computed independently. In the general case, a color difference includes simultaneously lightness, chroma, and hue differences; thus the three factors must be combined, and possible interactions should be investigated.
Finally, we computed how different color-difference formulas performed with our experimental results in comparison with CIEDE2000 and CIE94 with the appropriate parametric factors. Table 6 shows the PF/3 values computed between the tolerances of the textured samples (only the textures that statistically differed from homogeneous samples) and the corresponding reference tolerances. The second column shows the value of the 1296 pairs corresponding to the textures that statistically differed from homogeneous samples, while the third column shows the PF/3 value of the 658 pairs corresponding to absolute texture E_in the case of formulas with parametric factors [CIEDE2000, CIE94, CMC, or BFD (Bradford University), the value 1 was always used for the computation of the reference tolerances and the values recommended for textile samples for the tolerances of textured samples. As should be expected, the use of the parametric factors computed in this work for each simulated texture gave the best results. For our simulated textures, the choice of \( k_L=2, k_C=1, k_H=1 \) for CIEDE2000 and CIE94 and \( l=2, c=1 \) for CMC and BFD is inappropriate for the computations with all textures or only with the strongest texture (absolute texture). BFD (1:1), CMC (1:1),

Table 5. Visual Tolerances for Textures with Statistically Significant Differences regarding Reference Tolerances

<table>
<thead>
<tr>
<th>Texture (%)</th>
<th>( \Delta L_{10} )</th>
<th>( \Delta C_{ab,10} )</th>
<th>( \Delta H_{ab,10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (20)</td>
<td>3.51 (1.66)</td>
<td>5.63 (2.62)</td>
<td>3.27 (1.60)</td>
</tr>
<tr>
<td>A (50)</td>
<td>3.59 (1.64)</td>
<td>5.53 (2.63)</td>
<td>3.32 (1.64)</td>
</tr>
<tr>
<td>C (20)</td>
<td>3.04 (1.25)</td>
<td>5.92 (2.37)</td>
<td>3.50 (1.60)</td>
</tr>
<tr>
<td>C (50)</td>
<td>3.35 (1.56)</td>
<td>5.88 (2.37)</td>
<td>3.56 (1.61)</td>
</tr>
<tr>
<td>E (20)</td>
<td>4.30 (2.12)</td>
<td>5.96 (2.72)</td>
<td>3.91 (2.16)</td>
</tr>
<tr>
<td>E (50)</td>
<td>5.29 (2.79)</td>
<td>6.94 (3.36)</td>
<td>4.48 (2.37)</td>
</tr>
<tr>
<td>E (80)</td>
<td>6.19 (2.97)</td>
<td>8.19 (3.93)</td>
<td>5.29 (2.65)</td>
</tr>
</tbody>
</table>

*Tolerances averaged for sense of the tolerance, size of the dots, observers, and centers. Standard deviation in parenthesis.

Type and percentage of covered surface in parenthesis.

Table 6. Performance Factor PF/3 for Some of the Most Recent Color-Difference Formulas

<table>
<thead>
<tr>
<th>Formula</th>
<th>PF/3(^a)</th>
<th>PF/3(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIEDE2000 ( (k_L:k_C:k_H)^f )</td>
<td>26.0</td>
<td>28.0</td>
</tr>
<tr>
<td>CIEDE2000 (1:1:1)</td>
<td>35.0</td>
<td>33.7</td>
</tr>
<tr>
<td>CIEDE2000 (2:1:1)</td>
<td>47.4</td>
<td>43.8</td>
</tr>
<tr>
<td>CIE94 ( (k_L:k_C:k_H)^f )</td>
<td>26.3</td>
<td>28.5</td>
</tr>
<tr>
<td>CIE94 (1:1:1)</td>
<td>35.7</td>
<td>34.5</td>
</tr>
<tr>
<td>CIE94 (2:1:1)</td>
<td>47.8</td>
<td>44.0</td>
</tr>
<tr>
<td>CIELAB</td>
<td>34.1</td>
<td>33.5</td>
</tr>
<tr>
<td>BFD (1:1)</td>
<td>34.1</td>
<td>32.9</td>
</tr>
<tr>
<td>BFD (2:1)</td>
<td>46.9</td>
<td>43.4</td>
</tr>
<tr>
<td>CMC (1:1)</td>
<td>34.7</td>
<td>33.8</td>
</tr>
<tr>
<td>CMC (2:1)</td>
<td>46.9</td>
<td>43.5</td>
</tr>
<tr>
<td>LCD(^{kLc})</td>
<td>35.6</td>
<td>34.3</td>
</tr>
<tr>
<td>Volz(^{kLc})</td>
<td>35.7</td>
<td>34.4</td>
</tr>
<tr>
<td>Thomsen(^{kLc})</td>
<td>35.9</td>
<td>34.6</td>
</tr>
<tr>
<td>DIN99d(^{kLc})</td>
<td>35.1</td>
<td>34.1</td>
</tr>
<tr>
<td>( \Delta E_{GP} )</td>
<td>34.6</td>
<td>33.7</td>
</tr>
</tbody>
</table>

*Textures A (20%), A (50%), C (20%), C (50%), E (20%), E (50%), and E (80%).
*Textures E (20%), E (50%), and E (80%).

Parametric factors fitted for each texture (shown in Fig. 5).
Leeds Colour Difference.

CIELAB, and the formula \( \Delta E_{GP} \) performed acceptably well for the textured samples, even when only the texture E was considered in the computation.

4. CONCLUSIONS

The present study clearly shows that visual color-difference evaluation is influenced by the texture of the samples and confirms texture as an important parametric effect. However, this influence differs for each texture, and thus a simple set of parametric factors cannot be supplied for all potential textures available in industrial applications. Rather, each texture or sort of texture should be studied separately.

In general, random-dot textures increase the tolerances on lightness, chroma, and hue, each in a different way. Usually, lightness tolerances are increased by our simulated textures more than chroma or hue ones. However, the influence of the texture on chroma and hue tolerances is not negligible in general. Lightness textures increase mainly lightness tolerances, whereas chroma textures increase primarily chroma tolerances. In the case of other simulated textures or even real textures, similar results could be expected, but much more additional research is necessary.

ACKNOWLEDGMENTS

Special thanks to the observers who took part in the experiment and to Pedro A. García (University of Granada, Spain) for his useful advice on statistical analyses. The authors acknowledge funding support from the research project FIS2004-05537, Ministerio de Educación y Ciencia, Spain.

Address correspondence to Rafael Huertas at the address on the title page or by e-mail, ruertas@ugr.es.
REFERENCES