Measuring color differences in automotive samples with lightness flop: A test of the AUDI2000 color-difference formula

Manuel Melgosa,[1,*] Juan Martínez-García,[2] Luis Gómez-Robledo,[1] Esther Perales,[3] Francisco M. Martínez-Verdú,[3] and Thomas Dauser[4]

[1]*Department of Optics, University of Granada, 18071 Granada, Spain*
[2]*Laboratory Hubert Curien, UMR CNRS 5516, University Jean Monnet, Saint-Etienne, France*
[3]*Color and Vision Group, Department of Optics, Pharmacology and Anatomy, University of Alicante, Alicante, Spain*
[4]*AUDI AG, I/PG-C41, 85045 Ingolstadt, Germany*
[*]*mmelgosa@ugr.es*

Abstract: From a set of gonioapparent automotive samples from different manufacturers we selected 28 low-chroma color pairs with relatively small color differences predominantly in lightness. These color pairs were visually assessed with a gray scale at six different viewing angles by a panel of 10 observers. Using the Standardized Residual Sum of Squares (*STRESS*) index, the results of our visual experiment were tested against predictions made by 12 modern color-difference formulas. From a weighted *STRESS* index accounting for the uncertainty in visual assessments, the best prediction of our whole experiment was achieved using AUDI2000, CAM02-SCD, CAM02-UCS and OSA-GP-Euclidean color-difference formulas, which were no statistically significant different among them. A two-step optimization of the original AUDI2000 color-difference formula resulted in a modified AUDI2000 formula which performed both, significantly better than the original formula and below the experimental inter-observer variability. Nevertheless the proposal of a new revised AUDI2000 color-difference formula requires additional experimental data.

OCIS codes: (330.1690) Color; (330.1710) Color, measurement; (330.1730) Colorimetry.

References and links

1. M. Melgosa, A. Trémeau, and G. Cui, *Advanced Color Image Processing and Analysis* (Springer, 2013), Chap. 3.
2. W. V. Longley, "Automotive color certification," Color Res. Appl. **20**(1), 50–54 (1995).
3. E. J. J. Kirchner and J. Ravi, "Setting tolerances on color and texture for automotive coatings," Color Res. Appl., **39**(1), 88–98 (2012).</jrn>
4. M. Melgosa, "Testing CIELAB-based color-difference formulas," Color Res. Appl. **25**(1), 49–55 (2000).
5. Commission Internationale de l'Éclairage (CIE), *Improvement to Industrial Colour-Difference Evaluation*, CIE Publication 142–2001 (CIE Central Bureau, Vienna, 2001).
6. Commission Internationale de l'Éclairage (CIE), *Colorimetry - Part 6: CIEDE2000 Colour-Difference Formula*, CIE S 014–6 E:2013 (CIE Central Bureau, Vienna, 2013).
7. R. G. Kuehni, "CIEDE2000, milestone or final answer?" Color Res. Appl. **27**(2), 126–127 (2002).
8. S. Wen, "A color difference metric based on the chromaticity discrimination ellipses," Opt. Express **20**(24), 26441–26447 (2012), http://www.opticsinfobase.org/oe/abstract.cfm?uri=oe-20-24-26441.
9. F. J. J. Clarke, R. McDonald, and B. Rigg, "Modification to the JPC79 colour-difference formula," J. Soc. Dyers Colour. **100**(4), 128–132 (1984).
10. Commission Internationale de l'Éclairage (CIE), *Industrial Colour-Difference Evaluation*, CIE Publication 116–1995 (CIE Central Bureau, Vienna, 1995).
11. G. Cui, M. R. Luo, B. Rigg, G. Roesler, and K. Witt, "Uniform colour spaces based on the DIN99 colour-difference formula," Color Res. Appl. **27**(4), 282–290 (2002).
12. M. R. Luo, G. Cui, and C. Li, "Uniform colour spaces based on CIECAM02 colour appearance model," Color Res. Appl. **31**(4), 320–330 (2006).
13. C. Oleari, M. Melgosa, and R. Huertas, "Euclidean color-difference formula for small-medium color differences in log-compressed OSA-UCS space," J. Opt. Soc. Am. A **26**(1), 121–134 (2009).

14. M. Melgosa, R. Huertas, and R. S. Berns, "Performance of recent advanced color-difference formulae using the Standardized Residual Sum of Squares index," J. Opt. Soc. Am. A **25**(7), 1828–1834 (2008).
15. Z. Huang, H. Xu, M. R. Luo, G. Cui, and H. Feng, "Assessing total differences for effective samples having variations in color coarseness and glint," Chin. Opt. Lett. **8**, 717–720 (2010).
16. Z. Huang, H. Xu, and M. R. Luo, "Camera-based model to predict the total difference between effect coatings under directional illumination," Chin. Opt. Lett. **9**, 093301 (2011).
17. Commission Internationale de l'Éclairage (CIE), *Methods for Evaluating Colour Differences in Images,* CIE Publication 199–2011 (CIE Central Bureau, Vienna, 2011).
18. DIN 6175–2, "Farbtoleranzen für Automobillackierungen – Teil 2: Effektlackierungen," Deutsches Institut für Normung eV, Berlin, 2001–03.
19. T. Dauser, AUDI AG, I/PG-C41, 85045 Ingolstadt, Germany (personal communication, 2012). See *http://www.detroitcc.org/files/Color%20Management%20at%20AUDI%20(DCC%20March%202012).pdf*
20. American Society for Testing and Materials (ASTM), "Standard terminology of appearance," ASTM E **284**, 95a (1995).
21. M. Melgosa, L. Gómez-Robledo, J. Martínez-García, E. Perales, F. M. Martínez-Verdú, and T. Dauser, "Testing a colour-difference formula for the automotive industry using the experimental visual datasets employed in CIEDE2000 development," in *Proceedings CIE Centenary Conference* (CIE, Paris, 2013), CIE x038:2013 Publication, pp. 465–469.
22. J. Martínez-García, M. Melgosa, L. Gómez-Robledo, C. Li, M. Huang, H. Liu, G. Cui, M. R. Luo, and T. Dauser, "Testing the AUDI2000 colour-difference formula for solid colours using some visual datasets with usefulness to automotive industry," in Proceedings 8thIberoamerican Optics Meeting and 11th Latin American Meeting on Optics, Lasers and Applications (Proc. SPIE Vol. 8785, http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1782351).
23. M. Melgosa, "Request for existing experimental datasets on color differences," Color Res. Appl. **32**(2), 159 (2007).
24. O. Gómez, F. J. Burgos, E. Perales, E. Chorro, V. Viqueira, F. M. Martínez-Verdú, and J. Pujol, "Preliminary comparative performance of the AUDI2000 and CIEDE2000 color difference formulas by visual assessments in directional lighting booth," in *Proceedings AIC Colour 2013* (Newcastle, UK, 2013), **4**, pp. 1545–1548.
25. American Society for Testing and Materials (ASTM), *Standard Practice for Specifying the Geometry of Multiangle Spectrophotometers,* Technical Report ASTM E2194.01 (2011).
26. Commission Internationale de l'Éclairage (CIE), *Colorimetry, 3rd Edition*, CIE Publication 15:2004 (CIE Central Bureau, Vienna, 2004).
27. S. Guan and M. R. Luo, "Investigation of parametric effects using small colour-differences," Color Res. Appl. **24**(5), 331–343 (1999).
28. *http://www.sdcenterprises.co.uk/products/testing.asp* See also ISO 105–A02:1993, "Test for colour fastness – Part A02: Gray scale for assessing change in colour," International Organization for Standardization (Geneva, Switzerland).
29. F. J. Burgos, E. Perales, O. Gómez, E. Chorro, V. Viqueira, F. M. Martínez-Verdú, and J. Pujol, "Instrumental and visual correlation between a multiangle spectrophotometer and a directional lighting booth," in *Proceedings AIC Colour 2013* (Newcastle, UK, 2013), **4**, pp. 1373–1376.
30. A. Ferrero, A. M. Rabal, J. Campos, A. Pons, and M. L. Hernanz, "Spectral and geometrical variation of the bidirectional reflectance distribution function of diffuse reflectance standards," Appl. Opt. **51**(36), 8535–8540 (2012).
31. M. Huang, H. Liu, G. Cui, M. R. Luo, and M. Melgosa, "Evaluation of threshold color differences using printed samples," J. Opt. Soc. Am. A **29**(6), 883–891 (2012).
32. Commission Internationale de l'Éclairage (CIE), *CIE Colorimetry – Part 5: CIE 1976 L*u*v* Colour Space and u',v' Uniform Chromaticity Diagram*, ISO 11664–5:2009(E) / CIE S 014–5 E:2009 (CIE Central Bureau, Vienna, 2009).
33. Commission Internationale de l'Éclairage (CIE), *CIE Colorimetry – Part 4: 1976 L*a*b* Colour Space*, ISO 11664–4:2008(E) / CIE S 014–4 E:2007 (CIE Central Bureau, Vienna, 2007).
34. M. R. Luo and B. Rigg, "BFD(l:c) colour-difference formula. Part 1 – Development of the formula," J. Soc. Dyers Colour. **103**(2), 86–94 (1987).
35. P. A. García, R. Huertas, M. Melgosa, and G. Cui, "Measurement of the relationship between perceived and computed color differences," J. Opt. Soc. Am. A **24**(7), 1823–1829 (2007).
36. M. Melgosa, P. A. García, L. Gómez-Robledo, R. Shamey, D. Hinks, G. Cui, and M. R. Luo, "Notes on the application of the standardized residual sum of square index for the assessment of intra- and inter-observer variability in color-difference experiments," J. Opt. Soc. Am. A **28**(5), 949–953 (2011).
37. R. S. Berns and B. Hou, "RIT-DuPont supra-threshold color-tolerance individual color-difference pair dataset," Color Res. Appl. **35**(4), 274–283 (2010).
38. G. G. Attridge and M. R. Pointer, "Some aspects of the visual scaling of large colour differences – II," Color Res. Appl. **25**(2), 116–122 (2000).

39. M. Huang, H. Liu, G. Cui, M. R. Luo, N. Liao, M. Melgosa, Y. Zhang, and C. Zheng, "Assessing color differences in a wide range of magnitudes," *Proceedings AIC Colour 2013* (Newcastle, UK, 2013), **4**, pp. 1533–1536.

1. Introduction

Color-difference formulas try to predict visually-perceived color differences between two stimuli for an average observer under specified illumination and viewing conditions [1]. A color-difference formula provides a number ΔE from instrumental color specifications of two given color stimuli having a visually-perceived color difference ΔV. Note that while ΔV is the subjective answer of our complex human visual system, ΔE is a simple objective instrumental measurement. In industrial applications related to quality control it is important to use appropriate color-difference formulas to accurately predict subjective assessments of color differences performed by colorists or final users. For example, in the automotive industry, the different parts of cars are produced by different manufactures using different materials, and reliable color-difference formulas are necessary for automatic pass/fail color decisions [2, 3].

During the last 40 years advances have been made in the development of successful color-difference formulas for industrial applications [4], and the last CIE-recommended color-difference formula CIEDE2000 [5] has been recently adopted as a joint CIE/ISO standard [6]. Nevertheless it must be recognized that CIEDE2000 is not a completely successful formula [7, 8], and other recent color-difference formulas like CMC [9], CIE94 [10], DIN99d [11], CAM02 [12], or OSA-GP-Euclidean [13] perform similarly to CIEDE2000 for different experimental data sets [14]. In addition to the simple situation of samples with homogeneous colors (also called solid colors), research on color difference evaluation has also faced the most important problem of non-homogeneous samples with additional appearance attributes to color (e.g. gloss, texture, etc.), as well as color differences in complex images [15–17].

In the middle of the 90's AUDI developed a tolerance formula for the approval of effect paint batches when colors with strong flop effects were important in the automotive sector [18]. Later this formula was modified to predict color tolerances for solids as well as for effect colors, leading to the AUDI2000 color-difference formula [19], currently employed by different manufacturers in the automotive industry. Gonioapparent materials or materials with flop effects represent a big challenge for color-difference evaluation in the automotive industry, and nowadays AUDI2000 is the only available color-difference formula considering such effects. According to ASTM [20], "appearance" can be defined as "the aspect of visual experience by which things are recognized", "goniochromatism" as the "change in any of all attributes of color of a specimen on change in angular illuminating-viewing conditions but without change in light source or observer", and "flop" as "a difference in appearance of a material viewed over two widely different aspecular angles". While preliminary tests of AUDI2000 with homogeneous (solid) color pairs provided satisfactory results [21, 22], the goal of the current paper is to perform a color-difference experiment using automotive samples with lightness flop effects, analyzing the performance of AUDI2000 and other advanced color-difference formulas with respect to our experimental results. As requested by CIE [23], it is desirable to have new reliable experimental color-difference data sets to test and improve current color-difference formulas, and color-difference data sets involving gonioapparent materials are very scarce in current literature [24].

2. Materials and method

2.1 Selection of color pairs

Color measurements of 277 samples from different automotive suppliers were carried out with a BYK-mac 23 mm multi-angle spectrophotometer, assuming D65 illuminant and CIE 1964 standard colorimetric observer. This instrument provides color measurements considering a light source placed 45° with respect to the perpendicular to the sample, and detection at six different angles [25]: −15°, + 15°, + 25°, + 45°, + 75° and + 110°. The negative/positive signs of these six angles indicate clockwise/counterclockwise rotation angles with respect to the specular reflection of the incident light (see Fig. 1). These six illumination-detection geometries are also designed by CIE [26] as 45°x:-60°, 45°x:-30°, 45°x:-20°, 45°x:0°, 45°x:30° and 45°x:65°, respectively.

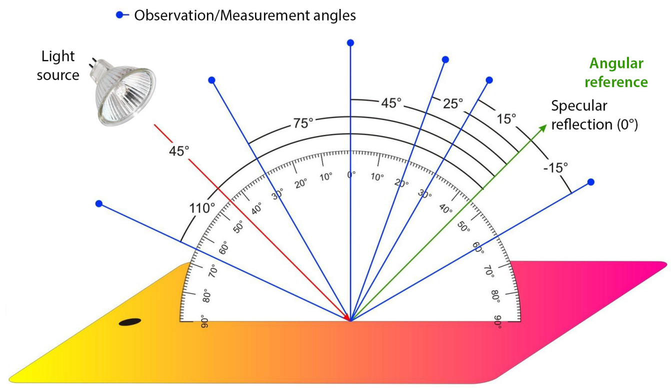

Fig. 1. Illumination and detection (viewing) geometries [25] considered in the current work.

From measurements at + 45° (i.e. the direction perpendicular to the sample), a set of 28 color pairs were selected for visual assessment. These color pairs met the next requirements: 1) the two samples in the pair were in the achromatic region (C^*_{ab}<10); 2) their color differences were relatively small (ΔE^*_{ab}<5); 3) the pairs had predominant lightness differences ($\Delta L^*<2[|\Delta C^*_{ab}| + |\Delta H^*_{ab}|]$); 4) the two samples in the pair came from the same supplier and showed very similar texture (non-perceptible in most cases). These requirements were adopted according to our main goal: to test the performance of most recent color-difference formulas disregarding texture effects, and analyze the lightness flop term proposed by the AUDI2000 color-difference formula. The selected 28 color pairs (Fig. 2) were well distributed in CIELAB color space, particularly with respect to the presence of pairs with small, medium and high L^* values.

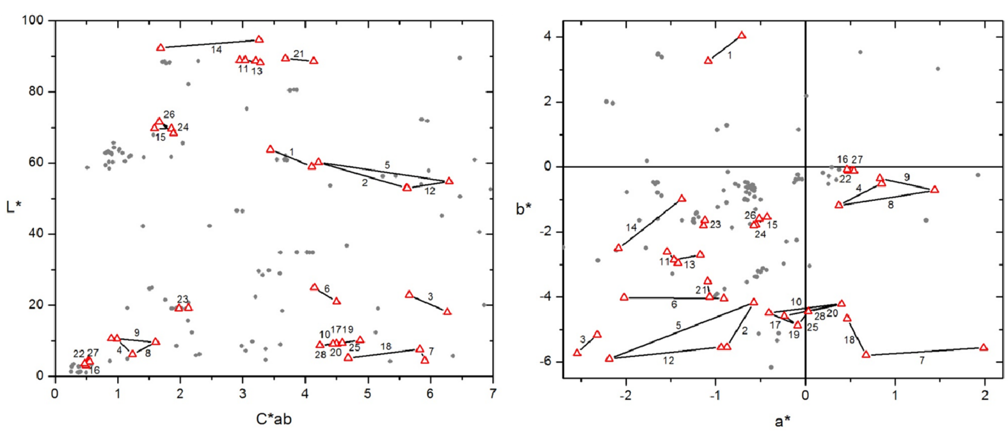

Fig. 2. CIELAB coordinates (BYK-mac 23 mm color measurements at + 45°) of the 28 color pairs selected for our visual experiment. The dots are available samples which were not used in our visual experiment.

2.2 Visual assessments

Using a gray scale method [27], each one of the selected 28 color pairs was visually assessed in a multi-angle byko-spectra effect light booth at the same six viewing angles previously mentioned: −15°, + 15°, + 25°, + 45°, + 75° and + 110° (see Fig. 1). Two issues of the SDC Grey Scale for Change in Colour [28] were used to provide the observer a set of 9 gray color

#199287 - $15.00 USD Received 10 Oct 2013; revised 23 Dec 2013; accepted 3 Jan 2014; published 6 Feb 2014

pairs with increasing color differences. These 9 color pairs were placed on the same plane than the test color pair, the distance between the center of the test pair and the central gray color pair #3 being 7.5 cm. The illuminance at the center of the test pair was 1800 lx. Bearing in mind that our selected samples had different sizes, looking also for more accurate visual assessments, we used a black gray mask providing a rectangular 3.5 cm x 3.7 cm test pair, with the same size than the color pairs in the gray scale. The color and texture of this mask (L^* = 18.6, a^* = 0.2, b^* = 0.1) were very similar to the ones of the background in the light booth (L^* = 24.2, a^* = 0.5, b^* = –0.6).

The observer's task was to assess the magnitude of the color difference in the test pairs in comparison with the color differences shown in the different gray pairs. Observers were encouraged to report intermediate assessments values between contiguous gray pairs. For a given test color pair, the six viewing angles were presented to the observer in random order, and the different test pairs were also presented in random order. Observers have a fixed viewing position in the light booth, in such a way that the distance from the observer to the test pair was 45 cm, and the viewing angle subtended by the test pair at + 45° was 4.5 degrees, allowing us to use the CIE 1964 standard colorimetric observer. Each observer performed a total of 504 visual assessments corresponding to 28 color pairs x 6 viewing angles x 3 replications, which were distributed across sessions of around 30 minutes, for a total of 10 sessions. Experiments were held in a dark room, and each observer adapted to the background of the light booth for 2 min before each experimental session. A panel of 10 observers (4 males and 6 females) with normal color vision (Ishihara test) participated in our experiment. They were staff and graduate students at the Department of Optics of the University of Granada, 7 of them with previous experience in color-difference assessments. All observers were subjected to a training round before their first observation session.

2.3 Spectroradiometric color measurements

It must be noted [29] that the light source in the byko-spectra effect light booth is close to TL84 while measurements with the BYK-mac spectrophotometer assumed D65, and also that the illumination in this light booth is not uniform, in particular at the extreme color pairs #1 and #5 of the gray scale. Considering that it is convenient to measure just what observers see, we decided to measure the color of all our samples in the byko-spectra effect light booth, using a Konica Minolta CS2000 spectroradiometer placed at the same position than the head of the observer, and use these measurements in all our next computations of performance of color-difference formulas. Therefore, while the selection of the color pairs used in our visual experiment was based on measurements made with the BYK-mac spectrophotometer at + 45°, as explained in section 2.1, the analyses of all current experimental results will be based only on spectroradiometric measurements performed under the same experimental conditions than visual assessments. The spectroradiometer was fixed on a robust tripod and it was slightly tilted to measure each one of the samples in both the test pairs and the gray scale pairs inside the light booth. A PTFE reference white placed at the same position than the samples was also measured with the spectroradiometer to allow transformations to CIELAB. On average, color differences measured with the Konica Minolta spectroradiometer and the BYK-mac spectrophotometer differed in 3.4 CIELAB units, the maximum differences being in L^* values and for the −15° angle.

From our spectroradiometric measurements at the six viewing angles we found that color differences in our 28 color pairs were below 6.0 CIELAB units in more than 83% of the cases [Fig. 3(left)], which is an appropriate range considering that the maximum color differences in our gray scale were in the range 10-15 CIELAB units, depending on the viewing angle we consider. We also found from our spectroradiometric measurements that the average CIELAB lightness difference in our 28 color pairs was high (71% of the total color difference), as desired in our experiment. Finally, from spectroradiometric measurements in the light booth, we measured the lightness flop (defined as the difference between maximum and minimum L^* values in the six measured angles) of the color samples involved in our selected 28 color pairs [Fig. 3(right)], obtaining an average value of 33.4 with a high standard deviation of

32.2. The samples with moderate lightness flop were predominant, and this flop range was considered appropriate to test the performance of the lightness flop term in AUDI2000.

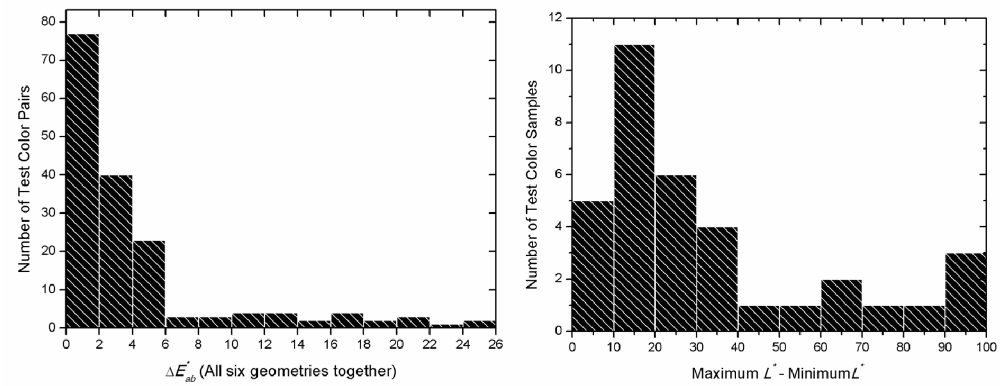

Fig. 3. Histograms with magnitudes of color differences (left) and lightness flop (right) for the pairs/samples in our visual experiment (CS 2000 spectroradiometric measurements in the byko-spectra effect light booth).

The raw gray scale values reported by the observers were transformed into true ΔV values, as usually made in the literature [27], using a fourth polynomial function fitted from spectroradiometric CIELAB color differences measured in the gray scale color pairs at each one of the six observation angles (Fig. 4). The color pairs of the SDC gray scale [28] had different color differences at different viewing angles. This is not surprising bearing in mind that gonioapparent effects have been also reported for other reflectance standards [30].

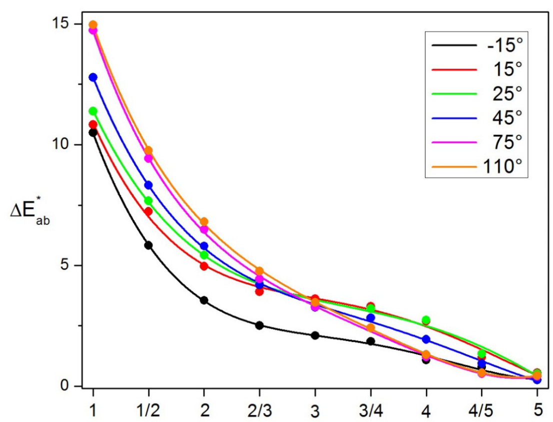

Fig. 4. CIELAB color differences (CS 2000 spectroradiometric measurements in the byko-spectra effect light booth) for each one of the nine color pairs in the SDC gray scale [28] at each one of the six viewing geometries (−15°, + 15°, + 25°, + 45°, + 75° and + 110°). Lines connecting points correspond to fitted fourth degree polynomial functions, used for transformations from raw gray scale values to true ΔV values.

2.4 Performance of color-difference formulas

The results of our visual experiment were tested against predictions made by 12 modern color-difference formulas, which have been also tested with other experimental data sets in recent color-difference research [14, 31]: CIELUV [32], CIELAB [33], CMC [9], BFD [34], CIE94 [10], AUDI2000 [19], CIEDE2000 [5], CAM02-UCS [11], CAM02-SCD [11], DIN99d [12], DIN99b [12] and OSA-GP-Euclidean [13]. Parametric factors in all these

#199287 - $15.00 USD Received 10 Oct 2013; revised 23 Dec 2013; accepted 3 Jan 2014; published 6 Feb 2014

formulas were kept as $k_L = k_C = k_H = 1.0$ in our current analyses. We paid special attention to the performance of the AUDI2000 color-difference formula, defined as follows [19]:

$$\Delta E_{AUDI2000,\gamma_i} = \sqrt{\left(\frac{\Delta L^*_{\gamma_i}}{k_L S_{L,\gamma_i}}\right)^2 + \left(\frac{\Delta C^*_{ab,\gamma_i}}{k_C S_{C,\gamma_i}}\right)^2 + \left(\frac{\Delta H^*_{ab,\gamma_i}}{k_H S_{H,\gamma_i}}\right)^2} \tag{1}$$

where $\Delta L^*_{\gamma i}$, $\Delta C^*_{ab,\gamma i}$ and $\Delta H^*_{ab,\gamma i}$ are the CIELAB lightness, chroma and hue differences between the two samples in the color pair at viewing angle γ_i; k_L, k_C, k_H are parametric factors kept as 1.0 in the current paper, and $S_{L,\gamma i}$, $S_{C,\gamma i}$, $S_{H,\gamma i}$, are the weighting functions for lightness, chroma and hue at viewing angle γ_i, respectively. When γ_i is equal to + 15°, + 25°, + 45° and + 75° these weighting functions are defined by the next equations:

$$S_{L,\gamma_i} = 1.0\left(\frac{\left|L^*_{\gamma_i+1} - L^*_{\gamma_i}\right|}{\gamma_{i+1} - \gamma_i}\right)^{2/3} + 0.002C^*_{ab,45^o} + 0.33 = a_1\left(\frac{\left|L^*_{\gamma_i+1} - L^*_{\gamma_i}\right|}{\gamma_{i+1} - \gamma_i}\right)^{a_2} + a_3 C^*_{ab,45^o} + a_4 \tag{2}$$

$$S_{C,\gamma_i} = 1.478\left(\frac{\left|C^*_{ab,\gamma_i+1} - C^*_{ab,\gamma_i}\right|}{\gamma_{i+1} - \gamma_i}\right) + 0.014C^*_{ab,45^o} + 0.27 \tag{3}$$

$$S_{H,\gamma_i} = 0.800\left(\frac{\left|C^*_{ab,\gamma_i+1} - C^*_{ab,\gamma_i}\right|}{\gamma_{i+1} - \gamma_i}\right) + 0.004C^*_{ab,45^o} + 0.30 \tag{4}$$

If $S_{C,\gamma i} < S_{H\,\gamma i,}$ the values from Eqs. (3) and (4) are averaged and this average value replaces both previous $S_{C,\gamma i}$ and $S_{C,\gamma i}$ values. In previous equations, $\gamma_{i\,+\,1}$ represents the next higher angle in the sequence + 15°, + 25°, + 45°, + 75° (e.g. $\gamma_{i\,+\,1} =$ + 25° when $\gamma_i =$ + 15°, or $\gamma_{i\,+\,1} =$ + 110° when $\gamma_i =$ + 75°), and L^* and C^*_{ab} correspond to CIELAB lightness and chroma, respectively. Here the first term in Eq. (2) will be called lightness flop term, and the first terms in Eqs. (3) and (4) will be called chroma flop terms.

When $\gamma_i = -15°$, the $S_{L,-15°}$, $S_{C,-15°}$ and $S_{H,-15°}$ functions are equal to $S_{L,\,+\,15°}$, $S_{C,\,+\,15°}$ and $S_{H,\,+\,15°}$ respectively, multiplying the slopes of all flop terms by a factor 1.2. That is, for −15° the slopes of each one of the flop terms are changed as follows: 1.0 x 1.2 = 1.2 = a_5, 1.478 x 1.2 = 1.774 and 0.800 x 1.2 = 0.960. Analogously, when $\gamma_i =$ + 110°, the $S_{L,110°}$, $S_{C,110°}$ and $S_{H,110°}$ functions are equal to $S_{L,\,+\,75°}$, $S_{C,\,+\,75°}$ and $S_{H,\,+\,75°}$ respectively, multiplying the slopes of each one of the flop terms by a factor 0.5. That is, for + 110° the slopes of each one of the flop terms are changed as follows: 1.0 x 0.5 = 0.5 = a_6, 1.478 x 0.5 = 0.739 and 0.800 x 0.5 = 0.400.

In summary, coefficients a_1 and a_2 are the slope and exponent in the lightness flop term, respectively, coefficients a_3 and a_4 are not in the lightness flop term but in the weighting function for lightness [Eq. (2)], and coefficients a_5 and $a_{6,}$ are the slopes of the lightness flop terms for angles −15° and + 110°, respectively. The values of the six a_i (i = 1,…,6) coefficients in the original AUDI2000 color-difference formula will be optimized later.

In usual automotive practice there is a master or reference sample, and Eqs. (2) to (4) are computed using color measurements from this master sample. However, in our current experiment we have two different samples and there is no reason to establish one of them as master. Therefore, here we computed Eqs. (2) to (4) for the two samples in each color pair and the corresponding arithmetical means were adopted as the final weighting functions in computations using Eq. (1). Note also that Eq. (1) provides a different value for each one of the six γ_i angles: the average and maximum of these six values are usually employed to establish pass/fail tolerances in the automotive industry.

The *STRESS* index [35] has been employed to measure the goodness of the predictions made by each tested color-difference formulas with respect to the whole set of results in our

#199287 - $15.00 USD Received 10 Oct 2013; revised 23 Dec 2013; accepted 3 Jan 2014; published 6 Feb 2014

experiment (i.e. 28 color pairs x 6 geometries = 168 color pairs). The ΔV_i (i = 1,…,168) values in our *STRESS* computations were the average of true ΔV values reported by our 10 observers. Low *STRESS* values (always in the range 0-100) indicate better performance of a color-difference formula. F-tests can be used with the *STRESS* index in order to know whether two color-difference formulas are or not statistically significant different [35]: the square of the ratio of the *STRESS* values from two color-difference formulas is compared with a specific confidence interval, which in our case was [0.74 ; 1.36], assuming a 95% confidence interval and taking into account that we have 168 color pairs. We have also used the *STRESS* index to compute intra- and inter-observer variability in our experiment [36]. For each observer, intra-observer variability was computed as the average of *STRESS* values of each one of the 3 replications made by this observer with respect to the average of the 3 replications, while inter-observer variability was computed as the *STRESS* value considering the average result of the 3 replications of this observer with respect to the average results of all 10 observers. Final intra- and inter-observer variability in our experiment were defined as the average intra- and inter-observer variability *STRESS* values from the 10 observers, respectively. A weighted *STRESS, WSTRESS* [37], has been also computed here bearing in mind that it is convenient to take into account the consistency of visual assessments made by the observers, giving a higher weight to color pairs where visual assessments have low standard deviations. Specifically, we used as appropriate weight the ratio $\Delta V_i\ /4\sigma_i$ (i = 1,…,168), where ΔV_i and σ_i are the average and standard deviation of true ΔV values from the 10 observers in each color pair, respectively. In our experiment the average weight was 0.90 with a standard deviation of 0.50.

3. Results and discussion

STRESS values corresponding to intra- and inter-observer variability in our whole experiment were 25.2 and 23.2, respectively. Similar values can be found in recent literature [30].

Table 1 shows *STRESS* and *WSTRESS* values for the whole results of our experiment and different tested color-difference formulas. F-tests from *STRESS* values shown in Table 1 indicate that the best color-difference formulas were AUDI2000 and CAM02-SCD (in this order), and they were not statistically significant differences among them. Analogously, F-tests from *WSTRESS* values shown in Table 1 indicate that the best color-difference formulas were AUDI2000, CAM02-SCD, CAM02-UCS and OSA-GP-Euclidean (in this order), and they were not statistically significant differences among them. While low performance (high *STRESS* or *WSTRESS* values) was expected for the CIELUV and CIELAB formulas, it was surprising to find also low performance for the CIE94 and CIEDE2000 formulas, which may be attributed to the fact that some conditions in our experiment (e.g. light source different than D65 illuminant, background behind the color pairs with L^* very different to 50, etc.) were not those recommended for these two color-difference formulas [5, 10]. The best performance of AUDI2000 may be attributed to the fact that this is the only color-difference formula considering flop effects and different lightness weighting functions in each viewing angle.

Table 1. *STRESS* and *WSTRESS* values for 12 color-difference formulas and our whole visual results

	CIELUV	CIELAB	CMC	BFD	CIE94	AUDI2000	CIEDE2000	CAM02-UCS	CAM02-SCD	DIN99d	DIN99b	OSA-GP-Euc.
STRESS	44.6	43.7	33.9	37.3	44.6	25.1	41.0	29.6	28.6	37.7	37.2	32.0
WSTRESS	42.8	42.2	33.2	37.2	43.3	26.6	39.9	29.0	28.2	36.4	36.0	30.8

#199287 - $15.00 USD Received 10 Oct 2013; revised 23 Dec 2013; accepted 3 Jan 2014; published 6 Feb 2014

Bearing in mind that results in Table 1 show *STRESS* and *WSTRESS* values for the AUDI2000 color-difference formula higher than our inter-observer variability (23.2 *STRESS* units), the six a_i (i = 1,…,6) coefficients of the weighting function for lightness in AUDI2000 were optimized to achieve the lowest *STRESS* and *WSTRESS* values (i.e. the best performance). The results found are shown in Table 2. Optimized coefficients in Table 2 reduced *STRESS* value from 25.1 to 21.8 and *WSTRESS* value from 26.6 to 24.0. As we can see, Table 2 shows enough similar values for coefficients a_1 and a_2 but different values for coefficient a_3, although in fact this is not very relevant because samples in our experiment have very low $C^*_{ab,45}$ values and the term including the coefficient a_3 is almost negligible. The values for coefficients a_4 and a_6 in Table 2 are also enough divergent in the two optimizations. In particular, adoption of the value a_4 = 0.01 found in *STRESS* optimization may be problematic because in the case of very achromatic samples with no flop (solid colors), the value of the weighting function S_L may be very small, producing meaningless high values in AUDI2000 computations [Eq. (1)]. It can be also noted that values for optimized coefficients a_5 and a_6 in Table 2 are enough different to those in the original AUDI2000 formula. If an unique set of coefficients a_i (i = 1,…,6) describing the results of our current experiment is required, we'd suggest values found minimizing *WSTRESS* (i.e. values in last row of Table 2). Using a_i values in the last (fourth) row of Table 2 the next advantages can be mentioned: 1) the potential abnormal high values derived from very low a_4 values in solid achromatic color pairs may be avoided; 2) *WSTRESS* value achieves the lowest value of 24.0 units indicating best performance; 3) *STRESS* value for the whole results of our experiment is 22.1 units, which is a good result, very similar to 21.8 units found as the lowest *STRESS* value using coefficients in the third row of Table 2.

Table 2. Coefficients of the weighting function for lightness in the original AUDI2000 color-difference formula (second row) and in two AUDI2000 modified formulas which were optimized to achieve lowest *STRESS* and *WSTRESS* values for the results of our current experiment (third and fourth rows)

Color-difference formulas	a_1	a_2	a_3	a_4	a_5	a_6
AUDI2000 original	1.0	0.67	0.002	0.33	1.2	0.5
AUDI2000 modified to minimize *STRESS*	1.1	0.50	0.010	0.01	1.0	1.0
AUDI2000 modified to minimize *WSTRESS*	0.9	0.67	0.005	0.15	0.8	0.9

Taking into account that color differences in our experiment covered a relatively wide range [Fig. 3(left)], we considered that an exponential function $(\Delta E_{AUDI2000})^{\alpha}$ may reduce the *WSTRESS* found with the AUDI2000 formula with optimized coefficients (i.e. coefficients in fourth row of Table 2), as suggested in previous literature [38,39]. A lowest *WSTRESS* value (best performance) of 21.8 units was found with α = 0.82, improving the earlier *WSTRESS* value of 24.0 found without the exponential function (i.e. assuming α = 1.0). It is interesting to note that this 21.8 *WSTRESS* value is lower than the inter-observer variability in our experiment (23.2 *STRESS* units), which means that this optimized non-linear formula predicts our visual results as close as an individual observer, as desired by useful color-difference formulas employed in industrial applications. Also it is worth to mention that F-tests [35] indicate that this modified AUDI2000 formula with optimized coefficients plus exponential function describes satisfactorily the whole visual results in the current experiment, being statistically significant better than the original AUDI2000 formula.

While the results of our current experiment can be fitted with a modified AUDI2000 formula, by the moment we are not proposing to replace current original AUDI2000 formula [19] before additional experiments are performed. Different modifications in our current experiment can be suggested leading to slightly different results to those shown in Table 1 and therefore to different values of the optimized coefficients. For example, we selected color pairs from spectrophotometric measurements at 45° but our spectroradiometric measurements showed that for 6 color pairs at specific angles (mainly −15° and 15°) the color differences were above the range provided by the gray scale used in our experiment. Missing these 6 color pairs in our analyses is unacceptable because they showed the medium-high lightness flop values we wanted to analyze. Removing just only some visual assessments in these 6 color pairs is against our experimental design considering all six standard visual angles. In a few words, we consider that the design of our experiment was correct and adapted to the available samples, but we recognize that it can be improved. In fact, the current experiment is now being replicated using a more appropriate set of color pairs, and also color pairs with greater size to allow the flop effects can be more evident. Hopefully our next results together with current ones can be used to propose an optimized AUDI2000 color-difference formula. By the moment, the results in this paper only indicate that the structure of the AUDI2000 formula is good to describe a visual color-different experiment using automotive solid and gonioapparent samples, while the optimal values of some of the coefficients in the weighting function for lightness in the AUDI2000 color-difference formula may be different to the original ones.

4. Conclusion

A visual data set with automotive samples showing flop effects has been developed. Detailed results in this data set (i.e. color coordinates of samples and visual differences) are available from the authors. While AUDI2000 was the best formula predicting our visual results, it seems that this formula can be improved. Further tests of AUDI2000 or other potential future color-difference formulas including flop effects are encouraged.

Acknowledgments

To all volunteer observers participating in our experiment. We are also very grateful to BYK Additive and Instruments (https://www.byk.com/en/instruments/products/color-measurement.html), for the loan of the byko-spectra effect light booth used in our experiments. Thanks also to the Color in Informatics and Media Technology (CIMET) Erasmus-Mundus Master (http://www.master-erasmusmundus-color.eu/); the results of this paper and Refs. [21] and [22], constituted Juan Martínez-Garcia's CIMET Master Thesis. This research was supported by the Ministry of Economy and Competitiveness (Spain), research projects DPI2011-30090-C02 and FIS2010-19839, with European Regional Development Fund.