https://doi.org/10.2352/CIC.2022.30.1.04

Testing the performance of color difference formula for a mixed display technology setup

Pooshpanjan Roy Biswas[1,2], Dominique Dumortier[2], Sophie Jost[2], Herve Drezet[1], Marie-Laure Avenel[1]
[1]Technocentre Renault, 1 Av. du Golf, 78280 Guyancourt, France
[2]ENTPE, l'école de l'aménagement durable des territoires. 3 rue Maurice Audin, 69518 Vaulx-en-Velin, France

Abstract

Automotive cockpits are becoming more and more digital day by day and this is evident by the increase in the number of displays inside the cockpit. With the advent of OLED displays, automotive cockpits not only have LCD but have started having mixed display technologies. A recent example of such a cockpit display is the MBUX Hyperscreen in Mercedes EQS [1].

A general situation is most mid ranged cars from Renault includes several displays inside the cockpit. These displays are placed at different locations inside the cockpit and go through various changes in external environmental condition during day and night. A vehicle cockpit experiences a wide range of illumination change during the day, from 35 kLux in bright sunlight to a few Lux during the night [2]. *The displays are generally used in high luminance range during the day and very low luminance during the night.*

Traditional color difference formulae like the CIEDE76 or CIEDE2000, the latter being the current industry standard, are defined for a specific set of evaluation conditions. For the myriad of different conditions that the displays undergo inside the cockpit, there isn't a recommended color difference formula which can be used to quantify the color difference between any two chosen displays inside the cockpit. An attempt has been made to include various real time parameters involving color difference evaluation between two displays, one of which is an LCD and other OLED. The motive of this study is to find out which color difference formula is the most representative of the perceived color difference between two mixed technology displays for a group of observers. The metric used for this the study is the CIE recommended STRESS index. The outcome of this research is to serve as a reference regarding the choice of color difference metric by display manufactures and OEM suppliers.

Introduction

A lot of work has been done to establish the relation between perceived color difference and calculated color difference using various metrics. A review about the performance of such metrics on a set of simulated data can be found in [3]. The CIE has laid down clear guidelines about how to quantify this relation on the basis of the STRESS index [4]. The STRESS index's performance on already available dataset and color difference formulae can be found in [5], [4]. It is evident from literature that STRESS index outperforms other metrics like PF/3 or CV in establishing a relationship between perceived and measured color difference.

It is worth noting that experiments that are conducted to establish such a relationship rely upon a particular experimental situation, and it is practically impossible to consider every possible factor that can influence a color difference evaluation. [6], [7] have evaluated conditions like distance between samples (up to 3 inches), size of the samples, outlining border etc in color difference evaluation, whereas [8] have evaluated a scenario where there is no separation between the samples. The former experiments dealt with CRT color display , while the latter dealt with surface colors. [9], [10] evaluated the performance of widely used color spaces and color difference formula in a standard no separation scenario using the STRESS index on WCG display. All these experiments used the Greyscale method [11] as a tool to quantify the agreement of perceived and measured color difference.

Cross-media based color matching experiments have also been conducted in the past [12] [13]. In such experiments, it becomes important to consider the intrinsic difference in technology of the different media. From the literature reviewed, a greyscale experiment using a cross media approach has not yet been studied. Apart from this aspect, certain extreme experimental conditions, such as a considerable distance between media (more than the displayed patch sizes), has not been evaluated. Experiments including displays are generally done without an external illumination source and with a high luminance (for example approximately 310 cd/m^2 in [9], [10]). Display luminance of this range puts stress on the observer's eyes when colors are evaluated in a dark environment [Renault internal study, confidential]. A greyscale experiment where an external illumination is also present has yet not been done. Using external illumination leads to other challenges as well because even if different media (displays) are maintained at the same white point chromaticity/luminance, difference in technology of the displays inside the cockpit lead to different color perception once external illumination comes into play. This happens because the illuminance of the external illumination increases the resultant luminance of the automotive displays. This results in reduction in contrast ratio, gamut or can impair gray-level differentiation [14]. Differences in the nature of display primaries across technologies (LCD v/s OLED) also results in difference in white point chromaticities.

Displays used in the automotive context go thru variety of changes during day and journey (from day till night). As there are multiple displays in the cockpit, the color difference evaluation might be done between large separation/distances. Also, as discussed above, the displays used can have different intrinsic technologies. In this paper, an attempt has been made to include all these aspects so that a real case scenario study can be established to quantify the performance of widely used color difference formulae with respect to perceived color differences, using the STRESS index.

Performance of eight different color difference formulae has been evaluated in this study. Most primitive of them is the CIELAB color space. Much work has been done since the introduction of CIELAB color space and its built-in color difference formula (Euclidean distance in the CIELAB color space) to improve the perceptual uniformity. CIEDE2000 color difference formula, which is also based on CIELAB is currently the recommended color difference formula for industrial color difference evaluation [15]. The automotive display industry is also well acquainted with the formula. With the introduction of CIECAM97s, researchers working on several aspects of color appearance could now focus on improvement on a single color appearance model adopted by the CIE. After the successful adoption of CIECAM02 in 2004 [16], various other CAMs have been introduced such as CIECIECAM16 [17], CIECAM16u, CAM20u etc, the first being the current recommendation by the CIE for evaluating color differences. The ultimate goal of color appearance models is to derive uniform color spaces (UCS). UCS is defined as a color space where equal perceived color differences are represented by equally calculated color difference in the derived color space. Several UCSs have been derived by CAMs, for example, CIECAM02-UCS from CIECAM02 [18], CIECAM16-UCS from CIECAM16 [19], DIN99d from DIN99 [20].

The goal of the present study was to find out which color difference formula or UCS based on color appearance models has the best agreement with visually perceived color differences for a real case automotive cockpit scenario. The parameters used to simulate this scenario was (i) a large distance between the evaluated displays. (ii) two displays belonging to different technologies (OLED and LCD), (iii) surface treatment as present in automotive displays (anti-reflection, anti-gloss etc) and (iv) a presence of simulated light source. The study was also done for two scenarios (i) day-light with higher luminance of displays and (ii) night-time with low luminance of displays illuminated by 2600 lux and 60 lux projector lightings respectively.

Apparatus Used

Two displays are selected for this study, one OLED and one LCD. The OLED display was 18 cms long diagonally and had a peak luminance of 437 cd/m2. The luminance of the display could be modified as per requirement using the supplied electronics. An anti-reflective layer was used on this display, as is the case with a real automotive multimedia display.

Figure 1: *Experiment setup with the car mockup and external illumination simulating the daylight scenario.*

The LCD display had a higher peak luminance but it was set at 437 cd/m2 with the help of a rheostat connected to the display electronics. The size of the display was also bigger than the OLED display, but to have an equal evaluation, the excess area beyond 18 cms diagonally was covered with a matte black paper. The peak white of the two displays did not have the same chromaticity but the same gray background (L*=50) was used for both the displays during the grayscale experiment. The two displays were connected to an HP laptop via an HP Thunderbolt Dock 120W G2 so that the number of display output ports could be increased to connect the two displays in Windows^R extended projection mode. The nomenclature followed for the displays was according to real world vehicle cockpit. The LCD display, placed on the left was referred as the Cluster and the OLED display was referred to as the CID (Central Information Display). The two displays were maintained at the same luminance level of 437 cd/m2 by using an in-contact colorimeter and using the driving electronics of the respective displays.

Experimental Setup

The two displays were placed at a distance of 24 cms (center to center) with their bases aligned. This was done using a manual rig to adjust and set the displays rigidly. Both the displays had their own dedicated electronics to enable HDMI connection to the controlling computer. A mock-up of car cockpit was created to mimic the effect of space and environment as is experienced inside an actual car cockpit. For this a rigid structure using wood and textile was used, all painted in black. Figure 1 below shows the actual experimental setup. A chin-rest was placed at a distance of 60cms from the center of the display setup, so that the head of the observer is fixed. However, the observer can move his eyes during the experiment. For the day condition, the Desire D60 from ETC Lighting was used to create an illuminance of approximately 2600 lux at the center of the display setup, while a Source 4 LED (ETC Lighting) was used to create the night-time driving situation of approximately 60 lux. These illuminances levels were chosen as they represented a real case scenario of illuminances inside a car cockpit [Renault internal, confidential]. The projector systems' illuminance was very stable in nature.

ICC Profiling

To characterize the displays, the ICC workflow was used. For usual display evaluation under a dark scenario, a contact-based colorimeter is generally used to create ICC profiles but for the displays' setup of this study, external illumination also had to be taken into account. For this reason, a JETI Specbos 1211UV spetroradiometer was used along with the software DisplayCAL [21] running on the open source libraries of ArgyllCMS [22]. ICC profiles for both the day and night conditions were created using 175 patches at default settings (no white point adjustment, white luminance level of 437 cd/m2 and no gamma curve forced on the displays). Using these settings, we were able to quantify the default relation between RGB triplets and the color displayed (CIEXYZ tristimulus values which we will exploit in the later part of the study).

Dataset Generation

14 color centers were used for this study, including 3 recommended by the CIE [23]. A uniformly distributed dataset scattered around these 14 color centers was first generated in the CIELAB color space. The criteria used for the scattered color centers was a CIEDE76 of either 3 or 6 units from the respective color center. For example, for the Red color center, the color centers in the a* and b* axes can be seen in Fig 2.

For every color center, 16 surrounding colors were calculated such that each surrounding color had a CIELAB difference of either 3 or 6 units. For the a* dimension, 4 surrounding colors having a difference of +/-3 and +/- 6 purely in the a* direction were calculated. Similarly, for the b* dimension, 4 surrounding colors having a difference of +/-3 and +/- 6 purely in the b* direction and for the L* dimension, 2 surrounding colors were chosen to have +/-3 difference purely in the L* direction were calculated. Similarly, to combine the dimensions, 6 surrounding colors having a difference of 3 units were calculated in the L*a*, a*b* and L*b* directions. The projection of these points in the 2D planes can be seen in Fig 2.

Table 1: CIELAB coordinates of the color centers

Color Center	L*	a*	b*
Red	53.41	92.62	72.49
Green 1	85.76	-83.82	93.11
Blue	36.62	42.21	-103.05
Yellow	96.33	-11.17	106.85
Cyan	89.96	-58.34	-14.18
Magenta	62.20	93.36	-60.21
Orange	68.96	14.11	81.63
Purple	39.12	58.65	-68.29
CIE Red	44.38	36.91	23.33
CIE Blue	35.60	4.83	-30.18
CIE Gray	61.65	0.11	0.04
Green 2	60.92	-42.65	8.88
Blue	56.66	-4.91	-46.53
Yellow 2	90.59	-9.19	89.73

Once the 224 CIELAB coordinates (14 color centers * 16 surrounding colors) were calculated, the RGB triplets that could produce these CIELABS had to be generated. For this the AtoB tag of the ICC profiles created before was used. The AtoB tag of an ICC profile captures the forward transform of a device, which related the device space to tristimulus values. A complete grid of the RGB space was converted to CIEXYZ using the A2B tag of the ICC profiles. The background grays of the GUI windows of both the displays had similar CIEXYZ tristimulus values [approximately CIEXYZ: 82, 88 and 102]. Because of this, the white point corresponding to this background gray (CIEXYZ × Peak White Luminance (437 cd/m2)) was chosen as the white point of the entire scene. The CIEXYZ values of both the displays were converted to CIELAB values using this whitepoint.

A MATLAB function was used to iterate over the entire RGB->CIEXYZ->CIELAB Look-up-Table to search for each of the target 224 CIELABS. The RGB triplet having the least CIEDE2000 was chosen as the RGB triplet that would produce the required CIELAB for the two displays.

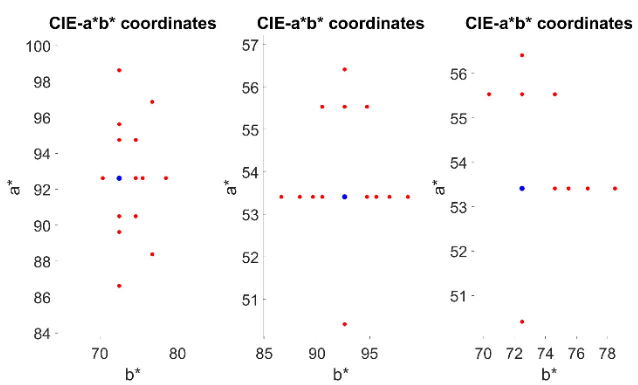

Figure 2: *Projection of the thresholded colors around the RED (see table 1) color centers in L*a*, a*b* and L*b* 2D plane. Color center is in blue and surrounding colors in red.*

The GUI is shown below. The RGB triplets were displayed on both the displays using a Graphical Observer Interface (GUI) created in MATLAB. The GUI had two windows, the first being displayed in the cluster and the second at the CID. For both the displays, the RGB patches were displayed at the bottom of the GUI while the Grayscale was displayed at the top of the GUI windows. For every instance of the 224 patches, the observer can use the left and right click buttons to change the Grayscale patch displayed on the cluster. The darkest Grayscale reference patch on the CID always remained the same.

The instructions given to the observer were:

"*You will use the two squares of the greyscale to express your judgement on the difference between the two coloured squares. In other words, the difference between the two grey squares that you will have selected will have to correspond to the difference that you perceived between the two coloured squares.*

When you click on the arrows on the screen, you have to make the difference between the top row of "Grayscale" patches, most similar to the difference that you perceive between the "Test Pair" in the bottom row patches."

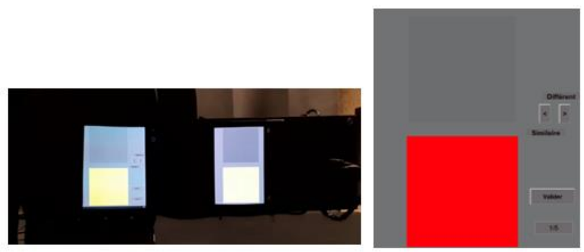

Figure 3*: Grayscale experiment GUI with two displays. Figure on the right shows the buttons to change the grayscale and validate the choice.*

The detailed instructions for the experiment were explained to the observer in French before the commencement of the experiment. After the instructions were explained, the observer had to do a demo of 5 patches (Fig 3 above), and if he/she gets accustomed to the experiment protocols, he/she would proceed to the real experiment. In general, each observer took 45 minutes to finish one situation (day or night). As explained earlier, the darkest grays between the displays had a CIEDE2000 of 0.42 and 0.48 for the day and night situation.

25 observers participated in the Day situation experiment and 19 for the night situation. Out of these 44 observer instances, 15 observers did both the day and the night situation. All the observers were in the age group of 18-30 years. None of the observers were presented with the same order of the patches as the patches' order being displayed were randomized using Latin squares [24]. In total every observer was shown 246 patches. The first 224 patches were the ones described above. The next 12 patches were patches repeated randomly from the 224 patches. These 12 patches were used to calculated intra observer variability. The last ten patches were the grayscale samples, so that the sanctity of the observer understanding of the experimental protocols can be established.

If an observer's choices are repeatable enough, the response for the patches 225 to 236 should be similar the first occurrence of these patches. Nevertheless, this repeatability would reflect in the intra-observer being discussed later. Also, for the last ten grayscale patches, the observer should select a similar grayscale displayed at the top row of the GUI.

Grayscale Experiment

The grayscale method [11] was used to quantify the perceived color difference seen between a color center (on the CID) and its surrounding color (on the Cluster). The grayscale was shown as the top two squares of our displays setup. In previous studies non-linear grayscales have been used [8], [9], [10], but some studies have proposed using a linear grayscale [25]. For this paper, a linear grayscale was developed as well. Using the ICC profiles for two displays, RGB coordinates were calculated using the method explained above so that a linear change in CIE L* values can be produced on the two displays. The grayscale ranged from 42 to 52 L* units with the darkest grayscale always being displayed on the OLED CID. In this way, there were 10 discrete points on the L* grayscale. In our observation, it was very difficult to create a pure L* scale (a*=0, b*=0), therefore there was always a very small chromatic a* or b* component in the Grayscale, thus it was not purely neutral in the true sense. This can be considered as a limitation of display systems in general, and not specific to our case. The CIEDE2000 values for the grayscale pair values can be found in table 2 below.

Using the left and right button placed besides the gray patch on the LCD, the observer can choose any grayscale patch out of the 10 values on the scale. Clicking the left button would make the two gray patches displayed respectively on two displays similar to the extent that the last patch would be identical to the darkest patch displayed on the CID, while the right button would make the two gray patches more and more different (see Fig 3). As the scale was linear, the grayscale can be converted to perceived visual difference linearly, contrary to previous literature.

Table 2: CIEDE2000 of the grayscale patches

Gray Scale	Day CIEDE2000	Night CIEDE2000
1	0.42	0.48
2	0.72	1.26
3	1.71	2.3
4	2.74	3.38
5	3.81	4.48
6	4.98	5.55
7	6.27	6.59
8	7.64	7.61
9	8.92	8.66
10	9.25	9.69

STRESS Index

To establish the relation between perceived color difference and calculated color difference using a formula, various metrics have been developed in the past [3] out of which, the CIE currently recommends the usage of the STRESS index [4]. STRESS is defined as follows:

$$STRESS = \left(\frac{\sum_{i=1}^{n}(A-FB_i)^2}{\sum_{i=1}^{n}F^2{B_i}^2}\right)^2 \times 100 \qquad (1)$$

$$where\ F = \sum_{i=1}^{n}A_i^2 \Big/ \sum_{i=1}^{n}A_iB_i \qquad (2)$$

STRESS index can also be used to calculate inter and intra observer variations [10][9][8]. To establish if the observer is coherent across repetitions, and also if the observers are coherent among each other, Intra and Inter observer STRESS were calculated for both the day and night situations. For the Intra observer STRESS, the response for the 12 repeated patches was used (patches 225 to 236).
STRESS values of 8 color spaces or color difference formulae, (CIEDE2000, CIEDE76, IC_tC_p, CIECAM02, CIECAM16-UCS, DIN99d, $J_zA_zB_z$ and OSA-GP), was evaluated.

Results

For the day situation, Inter and Intra observer STRESS for all the observers was found to be 39 and 42 units respectively, while for the night situation, the values are 43 and 36 respectively. These values are more than what has been found in grayscale studies in surface colors [8] but similar to displays [10][9]. A better intra observers STRESS for the night situation means that responses from the same observer were more repeatable in the night situation. A reason for this could be that observers' eyes were less strained for the low luminance condition during the night experiment as compared to the high luminance condition during the daytime experiment.

The mean visual color difference was calculated by averaging the grayscale values for all the observers for the 224 colors. STRESS was calculated between this average grayscale value and the mathematical color difference for the 224 colors using equation (1) described above. Eight color difference formulae (CIEDE2000, CIEDE76, IC_tC_p, CIECAM02-UCS, CIECAM16-UCS, DIN99d, $J_zA_zB_z$ and OSA-GP) were evaluated for their agreement with the grayscale responses using the STRESS index. As can be clearly observed in table 3, CIEDE2000 has the least STRESS value for both day and night conditions, followed closely by CIECAM16-UCS and DIN99d. To understand if the color difference formula having similar values are statistically significantly different, statistical F-test was conducted on every pair of color difference formula under discussion. For two formulae, F value can be calculated as:

$$F = \frac{STRESS_{DE1}^2}{STRESS_{DE2}^2} \qquad (3)$$

The F_c value for the present dataset was 0.8, with $1/F_c$ value as 1.23 for the 224 colors and α = 0.05 (95 % confidence level). Cells in table 4 whose value is less than 0.8 means that they are statistically significantly better than the other values in the same row, while value more than 1.23 mean that they performed statistically worse. Cells filled with green and blue color signify statistically significant improvement over other formulae for the day and night situation respectively. Similarly, orange and red font color signify statistically worse performance compared to other color difference formulae in the same row.

It can be seen that CIEDE2000 is statistically significantly better than most of the other color difference formulae considered for this study, except CIECAM16-UCS and DIN99d. A similar trend is observed for CIECAM16-UCS and DIN99d when they are compared to other formulae (except that they are not better than IC_tC_p while CIEDE2000 is). It is worth noting that that CIEDE2000, CIECAM16-UCS or DIN99d are not significantly better than each other. They should be considered as alternatives, rather than better choices. In terms of STRESS values, CIEDE2000 is only marginally better than CIECAM16-UCS or DIN99d. The worse performance is shown by OSA-GP as can be seen by the red and orange font color in table 4. This is similar to what was found in a previous study where OSA-GP in its default settings had similar worse STRESS performance [9].

STRESS was also calculated for each of the color centers so that areas in color space having a better agreement to visual color difference for various color difference formulae can be identified (see table 5). It can be seen that for color centers having good

Table 3: STRESS values for all the color difference formulae. Values underlined in bold and underlined have the best results in day and night situation

	CIEDE2000	CIEDE76	ICtCp	CIECAM02 UCS	CIECAM16-UCS	DDIN99d	JzAzBz	OSA-GP
Day	**28.79**	34.29	42.13	42.77	29.03	29.94	43.26	47.02
Night	**30.80**	37.86	35.26	38.02	32.95	31.80	40.43	42.36

Table 4: F-Test results for all the formulae. Refer to the text for the color scheme.

	CIEDE2000		CIEDE76		ICtCp		CIECAM02 UCS		CIECAM16-UCS		DDIN99d		JzAzBz		OSA-GP	
CIEDE2000	1	1	0.71	0.66	0.47	0.76	0.45	0.66	0.98	0.9	0.9	0.9	0.4	0.58	0.4	0.53
CIEDE76	1.42	**1.5**	1	1	0.66	**1.15**	0.64	0.99	1.4	**1.3**	1.3	**1.4**	0.6	0.88	0.5	0.8
ICtCp	2.14	**1.3**	1.51	0.87	1	1	0.97	0.86	2.11	1.2	2	1.2	1	0.76	0.8	0.69
CIECAM02 UCS	2.21	**1.5**	1.56	1.01	1.03	1.16	1	1	2.17	**1.3**	2	**1.4**	1	0.88	0.8	0.81
CIECAM16	1.02	1.1	0.72	0.76	0.47	0.87	0.46	0.75	1	1	0.9	1.1	0.5	0.66	0.4	0.6
DDIN99d	1.08	1.1	0.76	0.71	0.5	0.81	0.49	0.7	1.06	0.9	1	1	0.5	0.62	0.4	0.56
JAzBz	2.26	**1.7**	1.59	1.14	1.05	**1.31**	1.02	1.13	2.22	**1.5**	2.1	**1.6**	1	1	0.9	0.91
OSA-GP	2.67	**1.9**	1.88	**1.25**	1.25	**1.44**	1.21	**1.24**	2.62	**1.7**	2.5	**1.8**	1.2	1.1	1	1

Table 5: STRESS values for all the individual color centers

	CIEDE2000		CIEDE76		ICtCp		CIECAM02 UCS		CIECAM16-UCS		DIN99d		JzAzBz		OSA-GP	
	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night	Day	Night
Red	29.39	27.53	**20.45**	24.78	43.04	25.58	70.47	42.00	34.57	36.03	29.57	27.53	78.67	33.26	76.74	44.77
Green 1	21.92	16.92	29.54	21.98	22.17	16.85	38.26	22.63	23.38	20.69	**21.61**	17.19	31.12	28.42	28.59	30.27
Blue	18.24	27.73	22.87	30.00	35.22	31.71	43.93	44.95	30.85	40.46	**13.23**	24.56	48.50	46.89	43.30	44.66
Yellow	12.51	24.84	15.27	20.37	**10.48**	17.15	47.07	23.70	13.82	23.73	10.74	15.12	50.46	23.93	50.84	20.83
Cyan	12.79	20.99	13.25	23.25	13.63	16.30	26.20	32.88	**12.33**	28.04	13.01	19.72	29.47	40.76	23.71	33.72
Magenta	**11.60**	14.95	17.04	17.34	22.38	14.50	29.00	33.72	15.40	21.75	11.70	14.97	41.29	42.54	31.85	33.09
Orange	19.46	21.06	22.30	23.94	21.32	23.39	20.34	30.97	**19.34**	27.09	21.76	23.47	22.87	33.68	20.64	32.07
Purple	31.86	30.58	**26.75**	27.42	29.70	29.83	36.13	34.53	29.72	30.18	34.57	33.30	32.06	33.64	27.88	24.48
CIE Red	21.52	22.29	22.17	25.71	19.79	23.17	18.53	30.99	**16.98**	25.14	20.71	21.74	25.19	35.09	22.79	34.22
CIE Blue	30.08	29.38	29.37	27.97	**27.48**	26.08	27.58	28.91	29.68	27.84	29.28	30.41	36.34	34.95	31.52	31.18
CIE Gray	19.14	21.52	19.65	16.55	30.91	27.42	19.91	24.15	**18.68**	27.82	20.88	19.06	28.13	25.17	23.55	19.88
Green 2	27.76	25.26	32.39	28.25	31.29	27.14	30.05	29.25	**25.83**	26.03	27.88	25.39	32.69	33.14	28.30	28.45
Blue	30.42	26.45	27.48	29.67	**24.46**	22.07	28.43	34.58	26.76	31.95	29.44	27.77	34.20	40.35	28.58	33.02
Yellow 2	27.14	23.05	28.98	28.70	29.24	25.11	31.98	29.85	**27.56**	30.31	30.54	24.93	31.58	32.02	32.79	30.91
Mean	**22.42**	23.75	23.39	24.71	25.79	23.31	33.42	31.65	23.21	28.36	22.49	23.23	37.33	34.56	33.65	31.54

Table 6: STRESS values with optimized color difference formulae with the k_L parameter

	CIEDE2000	CIEDE76	ICtCp	CIECAM02 UCS	CIECAM16-UCS	DDIN99d	$J_zA_zB_z$	OSA-GP
Day Default	**28.79**	34.29	42.13	42.77	29.03	29.94	43.26	47.02
Day Optimized	28.76	33.84	39.71	42.13	**27.65**	29.92	41.20	45.83
k_L (optimized)	0.96	0.82	0.57	0.79	0.76	1.03	0.56	0.69
Night Default	**30.80**	37.86	35.26	38.02	32.95	31.80	40.43	42.36
Night Optimized	**30.80**	36.81	35.00	34.65	30.83	31.77	36.93	39.55
k_L (optimized)	0.99	0.71	0.86	0.62	0.72	1.06	0.53	0.62

STRESS numbers as compared to other centers for the day situation, might have a different pattern for the night situation. This can be observed for all the color difference formulae. The mean STRESS value was calculated by averaging the STRESS values of the individual color centers. Considering these values, CIEDE2000 outperformed other color difference formula with a STRESS value of 22. CIECAM16-UCS also performed very well for 6 out of 14 color centers, having least STRESS value for them. Average STRESS values calculated by averaging color centric data are expected to be lower than the total data. The latter should be considered to have real representation of the agreement of the overall color space with visually observed color differences.

Every color difference formula was optimized using the k_L parameter (see equation 4) in order to minimize the STRESS between the perceived color difference and the calculated color difference.

$$\Delta E = \sqrt{\left(\frac{\Delta L}{k_L}\right)^2 + (\Delta C)^2 + (\Delta H)^2} \qquad (4)$$

It was found that optimized CIECAM16-UCS gave the best STRESS performance (27.65, see table 6) for the day situation (marked in bold), while optimization did not result in a better result for the night situation. The level of improvement that has been reported in previous literature by the optimization technique was not found in the present study. Also, from an industrial usage perspective, it is also not easy to propose changes to default parameters for a widely accepted color difference formula (for example CIEDE2000). This that can lead to disagreements and changes to already accepted color difference values calculated by these formulae which have had been industry standards since a long time.

Conclusions

Eight color difference formulae have been tested in this study in an automotive context. The context involved using different display technologies (LCD with OLED) to conduct the grayscale experiment, a large distance between the displays, high luminance of displays and with external illumination to simulate two situations, day and night. For the 44 observers who participated in the two experiments, it was found that CIEDE2000 had the best agreement with the visually perceived color differences, followed by CIECAM16-UCS and DIN99d. The improvement was found to be statistically significant when compared to most of the formulae considered for this study. Different color centers performed differently. For some colors, the agreement to visual differences is better (magenta for CIDE2000) than others (red for $J_zA_zB_z$). Therefore, it is important to look at the total data (table 3), rather than color centric values (table 5). CIEDE2000 also performed statistically better than most of the other color difference formula. The formulae were also optimized with the k_L parameter with the expectation of improving the STRESS values, but this was not achieved with our dataset.

The motive of this study is to serve as a reference for the performance of color difference formulae in the automotive display industry so that display manufacturers and OEM suppliers can be confident about their choice of color difference metric.

References

[1] "MBUX Hyperscreen: High-Tech Highlight | Mercedes-Benz Group." https://group.mercedes-benz.com/magazine/technology-innovation/hyperscreen.html (accessed Jun. 19, 2022).

[2] T. Akhavan, H. Yoo, and A. Chubarau, "Solving challenges and improving the performance of automotive displays," *Inf. Disp. (1975).*, vol. 35, no. 1, pp. 13–27, 2019, doi: 10.1002/msid.1006.

[3] 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*, vol. 24, no. 7, p. 1823, 2007, doi: 10.1364/josaa.24.001823.

[4] CIE, "RECOMMENDED METHOD FOR EVALUATING THE PERFORMANCE OF COLOUR-DIFFERENCE FORMULAE," 2016. [Online]. Available: www.cie.co.at.

[5] M. Melgosa, R. Huertas, and R. S. Berns, "Performance of recent advanced color-difference formulas using the standardized residual sum of squares index," *J. Opt. Soc. Am. A*, vol. 25, no. 7, p. 1828, 2008, doi: 10.1364/josaa.25.001828.

[6] S. S. Guan and M. R. Luo, "Investigation of parametric effects using large colour differences," *Color Res. Appl.*, vol. 24, no. 5, pp. 356–368, 1999, doi: 10.1002/(SICI)1520-6378(199910)24:5<356::AID-COL7>3.0.CO;2-N.

[7] J. H. Xin, C. C. Lam, and M. R. Luo, "Investigation of parametric effects using medium colour-difference pairs," *Color Res. Appl.*, vol. 26, no. 5, pp. 376–383, 2001, doi: 10.1002/col.1053.

[8] F. Mirjalili, M. R. Luo, G. Cui, and J. Morovic, "A color-difference formula for evaluating color pairs with no separation-ENS," 2018.

[9] Q. Xu, B. Zhao, G. Cui, and M. R. Luo, "Testing uniform colour spaces using colour differences of a wide colour gamut," *Opt. Express*, vol. 29, no. 5, p. 7778, 2021, doi: 10.1364/oe.413985.

[10] B. Zhao, Q. Xu, and M. R. Luo, "Color difference evaluation for wide-color-gamut displays," *J. Opt. Soc. Am. A*, vol. 37, no. 8, p. 1257, Aug. 2020, doi: 10.1364/josaa.394132.

[11] "ISO_15008.pdf." .

[12] A. Sarkar, L. Blondé, P. Le Callet, F. Autrusseau, P. Morvan, and J. Stauder, "A color matching experiment using two displays: design considerations and pilot test results."

[13] M. Huang *et al.*, "Influence of viewing conditions on cross-media color matching," *Opt. Appl.*, vol. 49, no. 4, pp. 627–640, 2019, doi: 10.37190/OA190408.

[14] J. Bauer and M. Kreuzer, "Understanding the requirements for automotive displays in ambient light conditions," *Inf. Disp.*

(1975)., vol. 32, no. 3, pp. 14–22, 2016, doi: 10.1002/j.2637-496x.2016.tb00902.x.
[15] "CIE142-2001_ImprovementIndustrialColourDifferenceEvaluation.pdf." .
[16] "CIE159-2004_ColourAppearanceModelCIECAM02.pdf." .
[17] "A new colour appearance model for colour management systems: CIECAM16 | CIE." https://cie.co.at/technicalcommittees/new-colour-appearance-model-colour-management-systems-ciecam16 (accessed Apr. 21, 2022).
[18] R. M. Lou, G. Cui, and C. Li, "Uniform colour spaces based on CIECAM02 colour apperance model," *Color Res. Appl.*, vol. 31, no. 4, pp. 320–330, Aug. 2006, doi: 10.1002/col.20227.
[19] C. Li *et al.*, "Comprehensive color solutions: CAM16, CAT16, and CAM16-UCS," *Color Res. Appl.*, vol. 42, no. 6, pp. 703–718, 2017, doi: 10.1002/col.22131.
[20] G. Cui, M. R. Luo, B. Rigg, G. Roesler, and K. Witt, "Uniform Colour Spaces Based on the DIN99 Colour-Difference Formula," 2001, doi: 10.1002/col.10066.
[21] F. Höch, "DisplayCAL: Display Calibration and Characterization powered by ArgyllCMS," [Online]. Available: https://displaycal.net/#about.
[22] "Argyll Documentation Top." https://www.argyllcms.com/doc/ArgyllDoc.html (accessed Apr. 21, 2022).
[23] K. WITT, "CIE Guidelines for Coordinated Future Work on Industrial Colour-Diff erence Evaluation," vol. 20, no. 6, pp. 399–403, 1995.
[24] J. T. E. Richardson, "The use of Latin-square designs in educational and psychological research," *Educ. Res. Rev.*, vol. 24, pp. 84–97, 2018, doi: 10.1016/j.edurev.2018.03.003.
[25] L. M. CÁRDENAS, "Evaluation of Variability in Visual Assessment of Small Color Differences," 2019.

Author Biographies

Pooshpanjan Roy Biswas *received his masters in color science from the Erasmus Mundus masters COSI (Colour in Science and Industry) in 2017. He spent 4 years at HP R&D Labs in Barcelona as an R&D Engineer (Color). Currently, he is pursuing his PhD at the Renault-Nissan-Mitsubishi alliance and ENTPE (University of Lyon) in the domain of attaining colorimetric homogeneity in automotive displays. He has filed for 4 patents till date in the field of color science.*

Dominique Dumortier *received his PhD in Civil Engineering from the University of Savoie, France (1995).He has worked at ENTPE, Vaulx-en-Velin, France, since then. He is a member of CIE France. His research focuses on the influence of both artificial light and daylight on vision and biological rhythms.*

Sophie Jost *received her BS (2006) and PhD (2010) in Building Science and Indoor quality from ENTPE - University of Lyon FRANCE. Since 2012 she is a researcher in the Lighting group of ENTPE. Her work has focused on visual perception and lighting quality particularly colour rendering, colour appearance and psychophysics. She is the French representative for CIE division 1.*

Marie-Laure Avenel *received her engineering degree in photonics from E.N.S.S.A.T – Université Rennes 1 (2008) and PhD from Université Paris Sud (2012). Since 2017, she is working as the Display Innovation Leader in EE & Systems Engineering Department at Groupe Renault. Currently, she is working on new display technologies, new features and new integration for future cockpit concepts.*

Herve Drezet *is the Team leader of HMI components (clusters, central panel, HUD) at Renault Group. He has been an Engineer in the Renault Engineering division since 2000 as an expert in displays technologies*