Color-difference assessment and enhancement for driving headlight simulation

Quentin Vidal[1,2], Sylvain Michelin[2], Baptiste Laborie[3] and Andras Kemeny[1,4]

Abstract
Real-time headlight simulation in driving conditions is used by most car manufacturers to assure the quality, cost, and delivery of headlight engineering design. An important parameter judged by the headlight assessment team is color restitution; indeed, this parameter has to meet the standard of "lamps for road vehicles." Therefore, the goal of this study was the color assessment and enhancement of a driving headlight simulator. For this purpose, this study was conducted in two phases: the process of constructing two color acceptability scales that directly reflect the perception of two different populations (experts and "naive"), and the assessment of a method based on the chromatic adaptation transform (CAT) for reducing the color difference between real and virtual environments. In the first phase, we conducted two psychophysical experiments (i.e., one for each population), in which the observers had to report their degree of satisfaction about the color difference. These two experiments enabled the creation of two acceptability scales for headlight simulation. In the second phase, we compared the performance of different chromatic transformations; as a result of this comparison, we advise the use of the CAT02 transformation, in order to reduce the color difference for headlight assessment in driving simulation experiments.

Keywords
Color difference, headlight simulation, psychophysics

1. Introduction

Real-time headlight simulation in driving conditions is used by most of car manufacturers to assure the quality, cost, and delivery of headlight engineering design.[1] In these high-quality driving simulation applications, the colorimetric validity is essential because the headlight specialist uses this information to validate prototypes. For example, if the specialist perceives a slightly reddish orange color instead of the typical orange of a halogen light, he or she could declare that color to be outside the color gamut defined by the standards[2] and reject the prototype. Thus, this situation could lead to an unnecessary increase in headlight development time.

A previous internal study[3] has shown that, even if the luminance of a real headlight and its virtual reproduction differ, the contrast ratio[4] is essentially the same. However, this study also revealed color differences. As previously mentioned, in virtual headlight testing the rendered color fidelity must satisfy industrial assessment. This study, therefore, deals with color-difference perceptibility, which is the ability of an observer to detect a difference between two colors and, more precisely, the acceptability of the perceived color difference.

In this paper, we propose a new method for computing two color-difference acceptability scales,[5] which match the responses of a ''naive'' population (who are not accustomed to the task) and an expert population (i.e., designers). The use of two distinct populations is motivated by the fact that we wanted not only to compare these two populations, but also to provide one kind of certification for the colorimetric rendering of the headlight simulator, such as: ''not perceptible,'' ''acceptable by a color expert,'' or ''acceptable by a naive population.'' We then present a new heuristic for reducing the color difference between the real headlight and its representation in the virtual environment.

[1]Virtual Reality and Immersive Simulation Center, Technocentre Renault, France
[2]Laboratoire Informatique Image et Interaction, University of La Rochelle, France
[3]OKTAL, Meudon, France
[4]Laboratory of Immersive Visualization, Arts et Métiers ParisTech, France

Corresponding author:
Quentin Vidal, 23 Avenue Albert Einstein, La Rochelle 17000, France.
Email: quentin.vidal@percipio.fr

2. Related work

2.1. Color perception

For human color perception, the Commission Internationale de l'Éclairage (CIE) has defined two widely used color spaces: CIELAB and CIELUV.[6] Both spaces are derived from the CIEXYZ color space and are known to be pseudo-uniform, which means that the perceived difference between two colors depends on their locations in that space.

Because of this non-uniformity, the computation of the perceived difference in the CIELAB space has evolved. The first metric ΔE^*_{ab}, released in 1976, is defined as the Euclidean distance between two colors of this space:

$$\Delta E^*_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2} \tag{1}$$

This formula has been succeeded by three other reputed metrics: ΔE^*_{CMC}, ΔE^*_{94},[6] and ΔE^*_{00}.[7] These new metrics introduce application-specific weightings, which are unknown for our simulator. For this reason, when the notion of difference appears in this article, it refers to the first metric.

Using ΔE^*_{ab}, it is often considered that the just-noticeable difference is 1 unit, which means that no difference can be seen between two colors if the difference between them is less than this value.[8,9] Owing to the variety of observers and application fields, the color-difference acceptability is harder to define.[5]

Stokes et al.[10] define, for pictorial images, an acceptability threshold of 2.15 units. Abrardo et al.,[11] in an evaluation of the VASARI scanner, classified a difference of 1–3 as "very good quality," one of 3–6 as "good quality," one of 6–10 as "sufficient," and a difference over 10 as "insufficient." Hardeberg[12] defined a rule of thumb, in which the difference is "acceptable" if it is between 3 and 6. Recently, Thomas[13] extended Hardeberg's rule by taking into account the difference between an expert and a naive population.

2.2. Chromatic adaptation transforms

Chromatic adaptation transforms (CATs) are formulas that can predict various chromatic adaptation effects.[14] They are often used for determining corresponding colors under any two different adapting illuminants. A pair of corresponding colors consists of a color observed under one illuminant and another color that has the same appearance when observed under a different illuminant.[15]

Several CATs are described in the literature; most are based on the von Kries model. This model assumed "that although the responses of the three cone types (RGB) are affected differently by chromatic adaptation, the spectral sensitivities of each of the three cone mechanisms remain unchanged."[16] Therefore, this model can be seen as a linear transform by a constant factor for each of the three cone responses. The intensity of this factor will depend on the intensities of the two considered illuminants. Accordingly, the CIEXYZ tristimulus values $[X'Y'Z']^T$ of an object seen under a first illuminant are linearly transformed by a 3×3 matrix M_{CAT}, to represent the physiological responses $[L'M'S']^T$ of the cones:

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = [M_{CAT}] \times \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} \tag{2}$$

The resulting $[L'M'S']^T$ values are then transformed by a diagonal matrix to obtain the physiological cone responses $[L''M''S'']^T$ under the second illuminant. To obtain the CIEXYZ tristimulus values $[X''Y''Z'']^T$ of the object seen under the second illuminant, the $[L''M''S'']^T$ values are then multiplied by the inverse of the matrix M_{CAT}. Equation (3) describes this process[17]:

$$\begin{bmatrix} X'' \\ Y'' \\ Z'' \end{bmatrix} = [M_{CAT}]^{-1} \times \begin{bmatrix} \frac{L''_w}{L'_w} & 0 & 0 \\ 0 & \frac{M''_w}{M'_w} & 0 \\ 0 & 0 & \frac{S''_w}{S'_w} \end{bmatrix} \times \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \tag{3}$$

The matrices$[L'_w M'_w S'_w]$ and $[L''_w M''_w S''_w]$ are computed from the XYZ tristimulus values of the first and second illuminants by multiplying their XYZ tristimulus values $[X'_w Y'_w Z'_w]^T$ and $[X''_w Y''_w Z''_w]^T$ by M_{CAT}.

Süsstrunk et al.,[17] during their evaluation of the different M_{CAT} matrices found that the Sharp CAT, the Bradford CAT and CMCCAT2000 outperform most of the existing transforms when the full adaptation is assumed.

Since their evaluation, new M_{CAT} matrices have been released, such as the modified CMCCAT2000, also known as CAT02, which is used in the famous color adaptation model CIECAM02.[18] Lastly, Bianco and Schettini[15] released the $M_{\mathrm{BS\text{–}PC}}$ CAT, which was found to perform better than the other transformations on 16-color corresponding data sets.

These transforms can also be used to compute the corresponding tristimulus $[RGB]$ of two colors under different illuminants.[19] For this last point, we assumed that the use of a transform that is well adapted to simulate the human eye's perception would reduce the perceptible color difference between a real headlight and its reproduction in a real-time headlight simulator. For this purpose, we implement a shader that applies the CAT to a lightmap under the SCANeR™ studio virtual environment. SCANeR™ studio is a complete software tool meeting all the challenges of driving simulation (traffic generation, visual feedback, etc.) and is used by most car manufacturers. Nevertheless, the existence of different CATs necessitates their evaluation under this virtual environment, to determine the best CAT for this application.

3. Experiments

In this section, we describe three different experiments, which were conducted with two objectives: (1) the construction of two psychophysical scales, which directly reflect the perception of two distinct populations (naive and expert), and (2) a reduction in the color difference between the real headlight and its representation in the virtual environment.

These experiments took place in the lighting simulator at Renault, with all the lights and screens turned off. For the first two experiments, the observers sat in an equipped cab at a distance of 3.5 m from the screen. Following the recommendation of Schanda,[16] the standard 10° observer was used for the color space transformations.

Next, each observer was invited to report the degree of satisfaction with the colors' similarity via a man–machine interface. Observers were required to decide between four semantic categories: ''very satisfied,'' ''satisfied,'' ''not satisfied,'' and ''very unsatisfied.'' To understand this scale, the following instructions were given before the test: ''Very satisfied means that no difference can be seen and very unsatisfied that the difference is too great. For the two other values, imagine that you order a car or a cloth with one of these colors and you get the other one. Would you accept this difference?'' Distances were computed along two axes: the hue and the chroma. For each axis, two directions were considered: positive (clockwise), denoted $+$, and negative (anticlockwise), denoted $-$. ''Hue $+$,'' therefore, indicates that test patch color distance will vary along the hue axis in the clockwise direction.

3.1. Experiment no. 1

In the first experiment, we used nine patches from the Natural Color System (NCS). These patches were selected because they meet the specification of the white lamps for road vehicles and they are in the sRGB gamut, which corresponds to the projector's gamut. Six of them correspond to the headlight gamut boundary and three to the generic color coordinates of a LED, a halogen bulb, and a xenon bulb, as used in Renault's headlights.

To obtain psychophysical functions, the constant-stimuli method was used, with the two-alternative forced choice (2AFC) procedure.[20] During the test, a computer program randomly illuminates, using a calibrated Barco Galaxy NW-12 sRGB projector with $\gamma = 2.2$ and a D65 white point, one of the NCS patches and, simultaneously, a virtual color next to it (see Figure 1, which shows the white patch). Usually, a gray background is used to compare two colors.[21] For our application, we used the mean color of the rendered scene because it is in this condition that the headlights are evaluated.

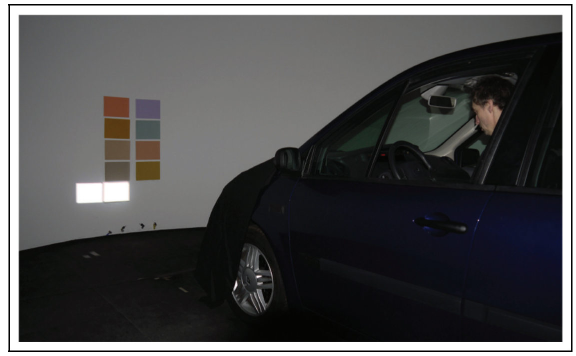

Figure 1. Experimental conditions for the first experiment. The projector illuminates an NCS patch and displays a virtual patch next to it.

To limit the duration of the experiment, the observer has 10 s to make a decision about the color difference. This time was chosen because we assumed that the observer has enough time to see the two patches, to decide whether to accept the difference or not, and, of course, to validate the answer. If the observer is unable to make a decision during this time, the program passes to another patch.

The chosen population for this experiment was composed of 10 women and 27 men, aged 25–50. All participants had normal color vision, as tested using the Ishihara's color deficiencies test, and no one had any experience of the color management system.

3.2. Experiment no. 2

The aim of the second experiment was to evaluate the results of the first experiment under a virtual environment. For this purpose, the experiment was conducted using the SCANeR™ studio virtual environment. In this environment, we virtually reproduced the experimental conditions of the first experiment (NCS color patches, apparent size, distance between patches) and uniformly illuminated the two patches with the car headlights (Figure 2).

Using the simple staircase method,[22] the observer has to accept or not the color difference between two patches. If the observer accepts, the difference is increased; otherwise, it is decreased. At the beginning of the experiment, the two patch colors were widely separated (ΔE^*_{ab} of 20 units), forcing the observer to reject this first value. The initial step value was set at $\Delta E^*_{ab} = 4$ units and progressively reduced to 0.125 unit (to compute a precise threshold, the step is divided by two at each reversal). The threshold was computed as the mean value of eight color differences at which a reversal occurs, starting from the reversal where the step was equal to 0.5 unit.

Figure 2. Experimental conditions for the second experiment.

Figure 3. Real measurement conditions for the Macbeth ColorChecker chart white patch.

In this experiment, the population was composed of three color experts from Renault (design department). Two persons worked on industrial quality validation and the other person worked on color and material assessment. Because of the expert nature of this population, we considered, as a predicate of this experiment, that their results should be highly close among themselves and that the result would not depend on the number of participants.

3.3. Experiment no. 3

The goal of the third experiment was to correlate the color rendering of a headlight simulator with reality. For this purpose, the experiment was divided into two phases: real measurements and virtual ones.

For the measurements in the real condition, we used six different types of headlight, which are currently used in Renault's vehicles (four halogen bulbs, one LED, and one xenon bulb). We placed the headlights at a distance of 25 m from a Macbeth ColorChecker chart and measured, using a CS-1000 spectroradiometer, the XYZ coordinates of the white patch under the six different light sources (Figure 3).

For the virtual measurements, we attempted to replicate the same conditions as for the SCANeR™ studio environment. We reproduced, following the recommendation of Pascale,[23] the Macbeth ColorChecker's white patch ($RBG = (243, 243, 242)$) and illuminated it with the virtual reproduction of the headlights used. This reproduction was obtained using .ies files, which are photometric representations of the headlights, and with .xy files, which define the color for the headlights. To ensure that the reproduction and the real measure corresponded to the same point, we assumed that the luminance difference should be a minimum.

To avoid color modification of the entire scene, we applied a fragment shader only to the three-dimensional headlight projection. This fragment shader modifies the color using the von Kries chromatic adaptation model (see Equation (3)), where full adaptation by the observer is assumed. The CAT matrices used in this work are listed in Table 1.

Table 1. Different chromatic adaptation transforms and corresponding matrices used in this study.

CAT	Matrix M_{CAT}
von Kries	$\begin{bmatrix} 0.38970 & 0.68900 & -0.0787 \\ -0.2298 & 1.18340 & 0.04640 \\ 0 & 0 & 1 \end{bmatrix}$
Sharp	$\begin{bmatrix} 1.26940 & -0.0988 & -0.1706 \\ -0.8364 & 1.80060 & 0.03570 \\ 0.02970 & -0.0315 & 1.00180 \end{bmatrix}$
CMCCAT2000	$\begin{bmatrix} 0.79820 & -0.3389 & -0.1371 \\ -0.5918 & 1.55120 & 0.04060 \\ 0.00080 & 0.02390 & 0.97530 \end{bmatrix}$
CAT02	$\begin{bmatrix} 0.73280 & 0.42960 & -0.1624 \\ -0.7036 & 1.69750 & 0.00610 \\ 0.00300 & 0.01360 & 0.98340 \end{bmatrix}$
Bradford	$\begin{bmatrix} 0.89510 & 0.26640 & -0.1614 \\ -0.7502 & 1.71350 & 0.03670 \\ 0.03890 & -0.0686 & 1.02960 \end{bmatrix}$
BS–PC	$\begin{bmatrix} 0.64890 & 0.39150 & -0.0404 \\ -0.3775 & 1.30550 & 0.07200 \\ -0.0271 & 0.08880 & 0.93830 \end{bmatrix}$

4. Results

4.1. Psychophysical function fitting

In the first experiment, we asked a naive population to report their degree of satisfaction on the color difference using a four-point semantic scale. For each color difference, a binarization of the answers was made, in order to compute the percentage of people who judged each difference as acceptable. For this purpose, we combined "very satisfied" with "satisfied" answers and "not satisfied" with "very unsatisfied" answers.

Next, we modeled our data using a psycho-physical function, which is a two-parameter function $F(x; \alpha, \beta)$.

Table 2. First experiment results: psychophysical coefficients α, β, and the coefficient of determination R^2 for each variation axis and each patch.

		1	2	3	4	5	6	7	8	9
Chroma −	α	4.119	4.462	3.393	3.008	4.527	2.817	3.627	2.583	3.145
	β	− 0.589	− 0.920	− 0.719	− 0.352	− 0.574	− 0.657	− 0.933	− 0.818	− 1.317
	R^2	0.992	0.996	0.992	0.984	0.9831	0.911	0.985	0.965	0.976
Chroma +	α	3.116	2.607	2.697	2.503	2.369	2.084	2.018	4.060	3.235
	β	− 0.694	− 0.601	− 0.456	− 0.326	− 0.361	− 0.757	− 0.416	− 0.679	− 1.424
	R^2	0.973	0.976	0.976	0.965	0.968	0.919	0.954	0.989	0.976
Hue −	α	6.964	4.131	3.524	3.832	2.814	3.262	3.954	4.293	7.537
	β	− 2.186	− 0.790	− 0.810	− 0.715	− 0.568	− 0.572	− 0.817	− 0.931	− 1.204
	R^2	0.978	0.899	0.991	0.946	0.988	0.987	0.999	0.913	0.991
Hue +	α	2.052	3.982	4.064	2.457	2.971	4.172	4.449	1.883	2.324
	β	− 0.998	− 0.646	− 0.869	− 0.716	− 0.575	− 0.520	− 1.180	− 0.867	− 0.831
	R^2	0.988	0.961	0.985	0.974	0.912	0.991	0.989	0.969	0.986

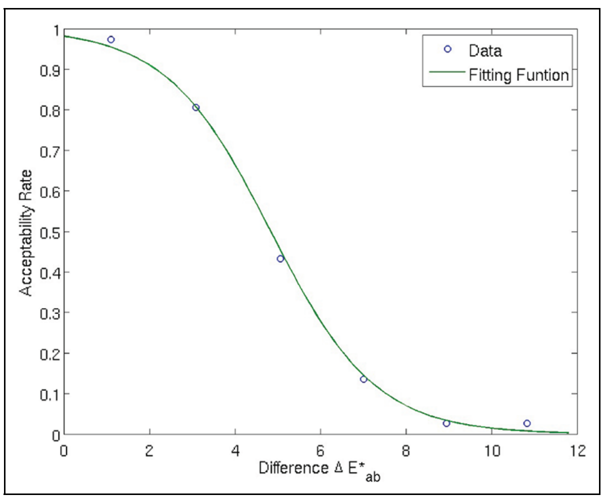

Figure 4. Fitting of the first experiment's data to a logistic function.

This function is typically a sigmoid function, such as the Weibull, logistic, cumulative Gaussian, or Gumbel distribution.[24] This kind of shape is explained by the fact that the closer a stimulus x (in our experiment, a color difference) is to a reference, the more people do not see any difference and accept it. In our case, the function that best describes the data from Experiment no. 1 is a logistic one (see Figure 4):

$$F(x;\alpha,\beta) = \frac{1}{1 + \exp(\beta x + \alpha)} \tag{4}$$

Thus, α represents the displacement along the abscissa and β represents the slope of the function $F(x)$. These coefficients were computed using the generalized linear model regression glmfit of Matlab's statistics toolbox.

The overall results for the function fitting are presented in Table 2; as expected, the parameters α and β are different for each patch and for each axis. This result is explained by the fact that the CIELAB space is non-uniform.[16] This means that the perception of color difference will be dependent on the position of the color in CIELAB space. This is true when comparing two different patches but it is also true when comparing two variation axes of a patch because the shape that includes all the colors that are visually identical to a reference is more like an ellipse than a circle. The use of advanced metrics, such as ΔE^*_{CMC} or ΔE^*_{00}, could reduce these differences.

Despite these differences, the function fitting is strongly correlated to the real data, with only 6 of the 36 values under 0.95, a mean coefficient of determination of 0.97, a standard deviation of 0.02, and a minimum value of 0.8988.

Although the data were highly correlated to the real data, some psychological functions had to be remove from the set. This is the case for patch no. 6, where its acceptability percentage does not go below 25% and moves back up at the maximal difference. This result could have been predicted by considering the patch position on the sRGB gamut. Indeed, this patch lies on the border of the gamut; computation of the new color generates a color that cannot be displayed by the projectors.

From this result and the knowledge that a threshold measured with the method of constant stimuli is defined as the intensity value that elicits perceived responses on 50% of the trials,[22] it is possible to reverse the function $F(x;\alpha,\beta)$, to obtain the acceptable difference x from the function of the acceptability rate:

$$x = \frac{\ln\left(\frac{1}{F(x;\alpha,\beta)} - 1\right) - \alpha}{\beta} \tag{5}$$

Thus, by replacing $F(x;\alpha,\beta)$ with its classical value of 50%, we have a mean acceptability scale for the naive population of 4.8 units.

Table 3. Results of experiment no. 2: ΔE^*_{ab} is the computed acceptability threshold; E/N is the acceptability rate of the naive population as a function of the expert population's acceptable difference.

		1	2	3	4	5	7	8	9
Chroma −	ΔE^*_{ab}	1.24	2.39	1.90	5.07	6.14	2.06	1.08	1.55
	E/N	96.74%	90.59%	88.35%	77.25%	73.18%	84.67%	84.56%	75.06%
Chroma +	ΔE^*_{ab}	1.59	2.67	3.21	4.4	4.12	2.92	2.63	2.18
	E/N	88.23%	73.19%	77.43%	74.42%	70.66%	69.07%	90.67%	53.21%
Hue −	ΔE^*_{ab}	2.66	2.39	3.74	3.17	7.5	4.49	1.68	3.4
	E/N	75.79%	90.43%	62.11%	82.74%	19.06%	56.94%	93.86%	96.90%
Hue +	ΔE^*_{ab}	1.36	2.42	3.5	3.09	8.79	4.63	1.71	2.75
	E/N	66.68%	91.84%	73.43%	56.02%	11.04%	26.64%	59.80%	51.00%

4.2. Expert validation

As expected from an expert population that is accustomed to this kind of experiment, the responses for the color-difference acceptability test are closely connected, with a mean standard variation of 0.49, which is less than the just-noticeable difference of the color perception.[8] This result shows that experts are agreed among themselves, validating our predicate for this experiment.

From the computed expert acceptance threshold and the function giving the percentage of color-difference acceptability in the naive population, it is possible to determine how the expert population is situated compared with the naive population (Table 3).

This data set shows that Renault's color experts do not accept a color difference when 76% of the naive population continues to accept it. The 76% value was computed after cutting off the highly influential values (gray values in Table 3). The outlier suppression was performed using the Cook's distance with a threshold of $4/n$ (with n the number of observations).[25] The suppression was made after a measurement session for which it appears that, for the deleted data, we were not able to reproduce the right color.

Commonly, a difference between two stimuli is considered unacceptable when 50%[22] of the population begins not to accept it. For the headlight simulation, this result reveals that the common threshold is not optimal if one wishes to satisfy a trained eye, such as those of the Renault design experts.

Thus, to obtain expert certification, the virtual headlight color should not differ from the real headlight color by more than 3.1 units. This leads to the creation of the proposed double scale, presented in Table 4.

4.3. Color enhancement

For the transformation of the tristimulus XYZ into the CIELAB space, we used, as the white reference $X_w Y_w Z_w$, the XYZ values of the projector's white. The measured differences between the real and virtual environments are listed in Table 5.

Table 4. Proposed color-difference acceptability scales of the naive and expert population.

	Expert	Naive
$\Delta E^*_{ab} \leq 1$	Good	Very good
$\Delta E^*_{ab} \leq 3.1$	Acceptable	Good
$\Delta E^*_{ab} \leq 4.8$	Unacceptable	Acceptable
$\Delta E^*_{ab} > 4.8$	Unacceptable	Unacceptable

In this table, the column with the symbol $\varnothing$ refers to the simulator performance without any post-processing (i.e., no chromatic adaptation transform for the headlights). As expected, we can see a noticeable difference between real and virtual reproduction with a mean value of 5.87 units. Such a difference can be explained by three main factors:

1. The projector is not perfect and cannot fit exactly to its related color space (generally the sRGB). Moreover, with the natural aging of the technology, color reproduction is not constant over time.
2. The computation of the interaction between the light and the object in the virtual environment does not work with spectral data. Therefore, owing to its sampled nature, the use of a "restricted" color space can create a shift.
3. The last point comes from the color management system. Indeed, the video projector works in a specific color space (the sRGB gamut in our case). Therefore, it is not able to generate a color that is outside its gamut; even the colors that fit in the gamut can be shifted because of the use of three stimuli instead of a complete light distribution.[26]

With the use of the chromatic adaptation transform, in 25 of the 36 samples, the shader reduces the headlight color difference by an average value of 2.29 units, which is a noticeable difference.[8] Taken overall, three of the six tested transforms enhanced the headlight reproduction system (von Kries, CAT02, and BS-PC).

Table 5. Color difference of the Macbeth ColorChecker's white patch between real and simulated environments. Bold indicates that the color difference is reduced using the chromatic adaptation transform.

M_{CAT}	∅	von Kries	Bradford	CMCCAT2000	Sharp	CAT02	BS–PC
Hal 1	1.77	5.59	8.32	8.95	9.01	**0.96**	3.64
Hal 2	7.12	**4.99**	**5.59**	**5.46**	**6.14**	**4.78**	**5.05**
Hal 3	8.01	**4.97**	**4.26**	**4.05**	**3.95**	**6.26**	**5.72**
Hal 4	7.11	**4.81**	**5.82**	**5.79**	**6.03**	**4.57**	**5.01**
LED	3.52	6.76	7.99	7.94	7.86	**1.68**	6.49
Xenon	7.66	**4.33**	**5.28**	13.74	**5.12**	**4.56**	**4.79**
Mean	5.87	**5.24**	6.25	7.65	6.35	**3.8**	**5.12**

However, when the difference between the real and initial chromaticity reproduction is small ($\Delta E^*_{ab} \leqslant 3.52$ units), the chromatic adaptation transform seems to increase the gap instead of reducing it. Therefore, a check must be performed prior to an industrial use of the shader.

5. Discussion

5.1. Use of the double scale

The purpose of the first experiment was to find psychophysical curves that represent the acceptability of the color difference of a naive population in a driving simulator. With this experiment, we computed logistics curves that strongly correlated with the experimental data, with a mean coefficient of determination of 0.97. With these curves, we proposed to use the common threshold of 50% acceptability rate[22] and obtained a first threshold of 4.8 units. However, for the headlight quality assessment, the headlight specialists would like to change this 50% to a higher rate. This brings us to the second experiment and thus to the question of the maximum value of the acceptability rate.

In the second experiment, we replicated the color patches used in the first experiment under the SCANeR™ studio virtual environment and asked color experts to adjust, using a 2AFC staircase, a second virtual patch, to produce an acceptable color difference. The protocols of these two experiments were not the same but we made the assumption that we could compare the results. Thus, the acceptability rates of the naive population as a function of the experts' acceptable differences were computed. We expected that the rates would be constant and higher than 50% because the two populations worked in the same color space. As expected, the corresponding rate of the expert population is higher than 50% (μ = 76%) but the standard deviation is not so small (σ = 12%). We assumed that this deviation is mostly due to the protocol differences; we consider that it does not affect the final result because the shifts caused by the protocol differences should have equally affected the experts' thresholds (i.e., exaggerating some of them and reducing others).

With these two experiments, we have proposed the two scales shown in Table 4, which are linked with different levels of certification. The first level corresponds to a color difference that is not visible, which the two populations find "good" or "very good." The second level corresponds to expert validation where the color difference is below 3.1 units and where a naive population rate the difference as "good." The third level of certification is the naive validation. This level is computed using the classical 50% rate with the psychophysical coefficients given in Table 2. If the headlight specialists consider that this value is not appropriate for its application, they could increase or decrease this value using the given coefficients. However, because the mean corresponding value between the expert and the naive population equals 76%, we would not recommend increasing the rates above this value.

5.2. White color correction

In the final experiment, we proposed a method to evaluate the correlation between a real headlight and its virtual reproduction. During this evaluation, we found that the initial color difference of the simulator was, on average, 5.87. Considering the state of the art[11,12] in color-difference assessment, such a difference is often considered acceptable, but this method did not validate the computed scales of the naive population.

With the use of the chromatic adaptation transform, the virtual headlight chromaticity was enhanced in three of the six transforms we tested. Among them, the CAT02 has always performed better than the reference environment (i.e., the environment without the shader). Furthermore, its mean color difference of 3.8 units is under the naive-population threshold and close to the expert-population threshold; this makes it reliable for industrial use.

With the aging of the video projector, its color reproduction will vary and it will be necessary to recalibrate the parameters of the chromatic adaptation transform to increase the photorealism of the environment. Therefore, to ensure that our shader maintains its performance over time, we suggest that the simulator be equipped with a spectrophotometer, to be used each time a headlight is loaded in the virtual environment. In a first step, the spectrophotometer will measure the *XYZ* value of the virtual patch without the white color correction and, in a second step, it will measure the *XYZ* value of the virtual patch with the CAT. If the color difference is effectively reduced with the CAT, the system will validate the transformation and the shader can be used by the virtual environment.

In addition to checking the performance of the white correction shader, the spectrophotometer can also be used to provide some feedback to headlight specialists. In fact, by using the psychophysical curves from the first experiment, a headlight specialist can determine the proportion of a naive population that would accept this color difference. Furthermore, if the difference is below the specialist threshold, the specialist will know that this color would be validated by a color expert.

6. Conclusions

In this paper, we present a method of assessing the acceptability of a color difference for a driving car simulator. For this purpose, we conducted two psychophysical experiments; first one with a naive population and then one with color experts from Renault's design team. These experiments enabled the construction of two color-difference acceptability scales, which directly reflect the perception of the two populations.

This proposed color-difference acceptability scale for the headlight simulation could be improved with the use of more colors in the psychophysical experiments. Indeed, nine colors are enough for the evaluation of the headlights' color reproduction but in a more complex scene, there are more colors, which lead to the evaluation of a larger palette.

By comparing real and corrected virtual environments, we have shown that the chromatic adaptation transform based on the von Kries model improves, in real-time, the color rendering of the driving headlight simulator. Indeed, when the initial color difference is not small (i.e., is less than the naive-population threshold) this transform reduces the color difference by 2.24 units. Furthermore, the use of the CAT only on the headlight projection does not damage the rest of the image and maintains the relationship between the different scene colors.

On the six headlights that we have tested, the only transformation that always reduces the color difference is the CAT02, with a mean reduction of 2.07 units. Therefore, owing to its reliability, the proposed solution for reducing the color difference will be integrated in the validation process of the car's headlight simulation.

Funding

This research received a grant from the ANRT (Association Nationale de a Recherche et de la Technologie) and Renault.

References

1. Lecocq KK, Lecocq P, Kelada JM, et al. Interactive headlight simulation. In *Proceedings of the DSC'99 driving simulation conference*, Paris, France, 1999, pp.173–180.
2. NF EN 60809/A4. Lampes pour véhicules routiers – Prescriptions dimensionnelles, électriques et lumineuses.
3. Laborie B. Expertise corrélation photo-colorimétrique – différences colorimétriques entre SCANeR 1.2 et SCANeR 14. Technical report, Renault, 2013.
4. Michelson AA. *Studies in optics*. New York: Dover Publications, 1995.
5. Laborie B, Viénot F and Langlois S. Methodology for constructing a colour-difference acceptability scale. *Ophthal Physl Opt* 2010; 30(5): 568–577.
6. Carter EC. Colorimetry. Technical report, International Commission on Illumination, 2004.
7. Sharma G, Wu W and Dalal EN. The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations. *Color Res Application* 2005; 30(1): 21–30.
8. Henry R. Kang. *Color technology for electronic imaging devices*. Bellingham, WA: SPIE Press, 1997.
9. Mahy M, Van Eycken L and Oosterlinck O. Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV, *Color Res Appl* 1994; 19(2): 105–121.
10. Stokes M, Fairchild MD and Berns RS. Precision requirements for digital color reproduction. *ACM T Graphic* 1992; 11 (4): 406–422.
11. Abrardo A, Cappellini V, Cappellini M, et al. Art-works colour calibration by using the VASARI scanner. In: *4th color imaging conference: color science, systems, and applications*, Scottsdale, AZ, November 1996, pp.94–96.
12. Hardeberg JY. *Acquisition and reproduction of color images: Colorimetric and multispectral approaches*. Boca Raton, FL: Universal-Publishers, 2001.
13. Thomas J-B. *Colorimetric characterization of displays and multi-display systems*. PhD Thesis, Université de Bourgogne, France, 2009.
14. Nayatani Y. Development of chromatic adaptation transforms and concept for their classification. *Color Res Appl* 2006; 31(3): 205–217.
15. Bianco S and Schettini R. Two new von Kries based chromatic adaptation transforms found by numerical optimization. *Color Res Appl* 2010; 350(3): 184–192.
16. Schanda J. *Colorimetry: Understanding the CIE system*. Hoboken, NJ: John Wiley & Sons, 2007.
17. Süsstrunk SE, Holm JM and Finlayson GD. Chromatic adaptation performance of different RGB sensors. *SPIE* 2000; 4300: 172–183.
18. Moroney N, Fairchild MD, Hunt RWG, et al. The CIECAM02 color appearance model. *IST/SID 10th color imaging conference*, Scottsdale, AZ, January 2002, pp.23–27.
19. Viggiano JAS. Comparison of the accuracy of different white-balancing options as quantified by their color constancy. *SPIE* 2004; 5301: 323–333.
20. Kingdom FAA and Prins N. *Psychophysics: A practical introduction*. London: Academic Press, 2009.

21. International Commission on Illumination. Parametric effects in colour-difference evaluation. Report, Central Bureau of the CIE. CIE 101-1993.
22. Ehrenstein WH and Ehrenstein A. Psychophysical methods. In: U Windhorst and H Johansson (eds) *Modern techniques in neuroscience research*. Berlin: Springer, 1999, pp.1211–1241.
23. Pascale D. RGB coordinates of the Macbeth ColorChecker. Technical report, The BabelColor Company, 2006.
24. Wichmann FA and Hill NJ. The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept Psychophys* 2001; 63(8): 1293–1313.
25. Bollen KA and Jackman RW. Regression diagnostics: an expository treatment of outliers and influential cases. *Sociol Method Res*, 1985; 13 (4): 510–542.
26. Stone MC, Cowan WB and Beatty JC. Color gamut mapping and the printing of digital color images. *ACM T Graphic* 1988; 7(4): 249–292.

Author biographies

Quentin Vidal is a PhD student at the University of Rochelle and CNRS in collaboration with Arts et Métiers ParisTech, with a research grant from ANRT and Renault.

Sylvain Michelin has been a professor of computer science since 2005. His main domain of research concerns computer graphics and image synthesis. He is also the head of a Master's degree program in Quality Health Safety and Environment.

Baptiste Laborie graduated from ENS Louis Lumière and holds a PhD in vision science and colorimetry. He is currently an engineer in research and development for vehicle simulation at OKTAL.

Andras Kemeny is an associate professor at Arts et Métiers ParisTech, France and Director of LIV, Arts et Métiers ParisTech–Renault. He was also director of LPCMV, CNRS–Renault and Associate Research Director at College de France, Paris.