Evaluation of Color Saturation and Hue Effects on Image Quality

Ema Popovic, Ivana Zeger, Mislav Grgic, Sonja Grgic
University of Zagreb Faculty of Electrical Engineering and Computing
Zagreb, Croatia
ivana.zeger@fer.hr

***Abstract*—Saturation and hue are image characteristics that significantly affect the perception of a scene. This paper observes the impact of changing saturation and hue on image quality. Concepts of subjective image quality assessment methods are explained. Reasons for developing objective image quality evaluation methods are provided together with a short overview of the methods used in the paper. A brief interpretation of color spaces in which the objective evaluation was carried out is given. The created image database containing images with modified saturation and hue with regard to original images is described. The settings of the performed subjective testing are stated. A subjective and objective evaluation of image quality was executed. Kendall's tau and Spearman's rho correlation coefficients are calculated for the comparison of subjective and objective evaluation results. Efficiency of the methods is discussed.**

***Keywords*—Image Quality; Color Saturation; Hue; Subjective Evaluation; Objective Measures**

I. Introduction

A digital image is displayed as a grid of pixels, each comprising a specific color. When taking photos with a camera, the colors of the real scene are being described with a limited range of digital values. As the end users of images are principally people, adapting the image to the human visual system (HVS) is of great importance. HVS connects the processes that occur in the sensory organ, the eye, and the central nervous system. The eye receives light stimuli from its surroundings. The optic pathway includes the retina, optic nerve and tract, and visual cortex [1]. Two types of photoreceptors are located in the retina: rods, responsible for the perception of lighting levels, and cones, which make distinction of colors possible [2]. The elements of the visual system work together in a complex way to provide the human with the sense of sight and the perception of colors. The feature that a human observer notices first is the color type, hue [3]. In addition to hue, saturation is a feature that affects the human experience of an image. The effect of color change on image quality is a significant part of image quality assessment.

This paper deals with the evaluation of the impact of changing hue and saturation on image quality. In Section II, an overview of the color spaces is given. The procedures of subjective and objective image evaluation are explained in Section III. Section IV describes the created database of images with modified saturation and hue. Details of conducted subjective testing are provided in Section V. Section VI discusses the results. Conclusion ends the paper.

979-8-3503-2512-6/23/$31.00 ©2023 IEEE

II. Color Spaces

Color space is the color description system of a digital image. In RGB model, there are three primary colors – red, green, and blue. The choice of primary colors is related to HVS - cones are most sensitive to one of those three colors. The model is additive; the lowest intensity of the three primary colors results in black, and the highest intensity in white [4]. Different combinations of primary colors produce all the colors that can be displayed.

CIELAB [5] and YUV [6] color spaces are more similar to the way HVS works than RGB color space. Both CIELAB and YUV are described using one luminance and two chrominance channels. Human vision is very sensitive to changes in illumination, while distinguishing colors does not have such an influence on the viewer's experience. Compression can be applied to two chrominance channels. CIELAB is also denoted as L*a*b*, where L*, a* and b* represent the values by which a certain color is measured and calculated. Channel L* indicates the image brightness. Channels a* (green-red) and b* (blue-yellow) are chrominance channels. The space is obtained by non-linear transformation of RGB space. Similarly, YUV color space uses Y*, u* and v* as its channel labels. It was obtained from RGB space by linear transformation. Channel u* represents the difference between blue and the luminance and v* the difference between red and the luminance.

III. Evaluation of Image Quality

Image quality evaluation can be objective or subjective. In subjective image evaluation methods, a selected group of observers shares their opinion about the image quality. To ensure the credibility of the results, the ITU-R BT.500-14 standard [7] was created. In addition to the methods themselves, it describes which conditions should exist during the evaluation (proper room lighting, screen settings, viewers' distance from the screen, viewing angle, number of viewers, duration), how to compile the instructions for the viewers and how to present the results. Subjective evaluation methods are attractive because the obtained values depend on HVS. Nevertheless, the process of subjective evaluation is demanding and time-consuming.

Because of that, objective image evaluation has evolved. Objective methods are faster, easier to implement and more frequently used [8]. However, HVS is difficult for mathematical modeling. Thus, objective methods achieve varying levels of success in approaching human judgment.

2023 International Symposium ELMAR | 979-8-3503-2512-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ELMAR59410.2023.10253901

There are objective methods which require access to a reference image, and those that do not. The methods that need the reference image compare the values of the reference and the test images. The best quality that the test image can achieve is equal to the quality of the reference image [4], these methods cannot assign a better score to the test image in comparison to the reference image. If the subjective quality of test image is higher than subjective quality of reference image, the results of objective measures will be negatively correlated with the results of subjective evaluation. It can happen when the saturation of the test image is increased. Some of the frequently used methods in this category are mean square error (MSE), peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) and chroma error ratio (CER) [9], [10]. SSIM is a combination of luminance (l), contrast (c) and structure (s):

$$SSIM(x,y) = l(x,y)^{\alpha} \cdot c(x,y)^{\beta} \cdot s(x,y)^{\gamma}, \quad (1)$$

where $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ are parameters that adjust the relative importance of l, c and s. Mostly, α, β and γ are set to 1.

Methods that do not require access to a reference image usually detect specific types of distortion. They are used in many applications where no reference image is available [4], such as the evaluation of underwater images, and in assessing non-underwater images as well, because they enable the evaluation of color reproduction [9], [10]. Often used methods are underwater color image quality evaluation metric (UCIQE), underwater image quality measure (UIQM) [9], [10] and CCF (colorfulness, contrast, fog density) [11].

IV. Image Database for Evaluation

For work purposes, 12 original images with a size of 4496x3000 pixels, and color depth 24 bits per pixel, were taken with a Nikon D5600 camera to be sure that camera and sensor characteristics will not affect evaluation process. The images with the corresponding names used are shown in Table I. New database of images containing images with changed color components, saturation and hue, was created. For each original image, one with reduced saturation and one with increased saturation was created. However, only one direction for changing hue was chosen. After selecting direction, hue was changed with two different deviation values: one should produce “just noticeable difference” and the other “very noticeable difference”. Images are not modified with the same deviation from the original. Hue and saturation were changed independently. Each test image contains only one modification (hue or saturation).

The database of test images consists of 48 modified images: 4 modifications of each original image. The characteristics of images were changed using Adobe Photoshop. Test images are shown in Table II. The name of the image gives information about the modification - "h" represents changed hue, and "s" represents changed saturation. The last (signed) number represents the offset (deviation) from the original value.

A vectorscope provides a graphical representation of the color components of the image (color wheel) and is used to determine and analyze hue and saturation. The u* signal is appointed horizontally, and v* vertically. The pixel values determine the position of color vector. A change in hue is shown by circular movement on the vectorscope display, while a change in saturation is shown by a change in the length of the vector [9, 10]. The vectorscope was created using an implementation in MATLAB [12].

A. Change of Color Saturation

Color saturation describes the purity of the color source. The human eye is attracted to bright colors, so saturation plays a big role in the human experience of an image [13]. In CIELAB and YUV color spaces, during saturation modification, the values of the chrominance axes are changed to bring the tones closer or further away from the grayscale representation.

TABLE I. Database of Original Images


Image 0


Image 1

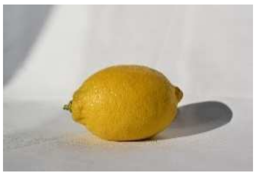
Image 2


Image 3

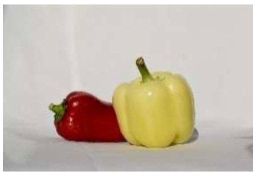
Image 4


Image 5


Image 6


Image 7


Image 8


Image 9


Image 10


Image 11

TABLE II. DATABASE OF TEST IMAGES

Image name	Image	Image name	Image	Image name	Image	Image Name	Image
0-h+10		0-h+23		0-s+30		0-s-30	
1-h-8		1-h-17		1-s+40		1-s-40	
2-h+7		2-h+15		2-s+30		2-s-30	
3-h-9		3-h-18		3-s+30		3-s-30	
4-h+5		4-h+15		4-s+20		4-s-20	
5-h-8		5-h-18		5-s+40		5-s-30	
6-h-8		6-h-18		6-s+30		6-s-30	
7-h-15		7-h-30		7-s+50		7-s-50	
8-h+15		8-h+30		8-s+20		8-s-30	
9-h-8		9-h-15		9-s+30		9-s-30	
10-h-8		10-h-15		10-s+30		10-s-30	
11-h-8		11-h-15		11-s+30		11-s-20	

Images with saturation changes are shown in Fig. 1. Vectorscopes of these images are shown in Fig. 2. The photographed reference image, viewed through the vectorscope, contains a lot of orange and blue. Saturation is reduced in Fig. 1b. The vectorscope shows that the image still contains mostly orange and blue colors, but not in the brightest shades at the edge of the vectorscope. The range of colors is smaller, and they are closer to the center, to white and shades of gray. Saturation is increased compared to the reference image in Fig. 1c. A larger number of pixels closer to the edge is displayed on the vectorscope, i.e., the colors become brighter and cleaner. The area of used vectors expands to the left and right and the image now uses more different color values, their range is greater.

B. Change of Color Hue

Hue is described by the dominant wavelength of the stimulus [3]. Fig. 3 shows the impact of changing hue on the image, while Fig. 4 shows the impact on the vectorscope. Fig. 3a shows blue and orange glasses, and the vectorscope view in Fig. 4a shows that the most pixels contain tones of exactly these two colors. In Fig. 3b, hue has been changed, the shades of orange have become more yellow, and the shades of blue have become more purple. Such a change is visible on the vectorscope as a rotation of the vector display to the left (Fig. 4b). It remained almost the same shape, but the angle was changed in accordance with the new state of colors. If the vectorscope view of the original image was rotated to the right, the change would be shown as redder shades of orange and greener shades of blue.

V. Conducted Subjective Evaluation Experiment

The subjective method used is the stimulus comparison method from the ITU-R BT.500-14 recommendation [7]. During the evaluation, two images were displayed simultaneously. The viewer was given a scale to describe the relationship between images. All image pairs shown consisted of an original image and a test image. The viewer was unaware which image was which. The subjective evaluation of the images was carried out at the University of Zagreb Faculty of Electrical Engineering and Computing on three identical

Figure 1. (a) Original image, (b) image with reduced saturation, (c) image with increased saturation

Figure 2. (a) Vectorscope of original image, (b) vectorscope of image with reduced saturation, (c) vectorscope of image with increased saturation

Figure 3. (a) Original image, (b) image with changed hue

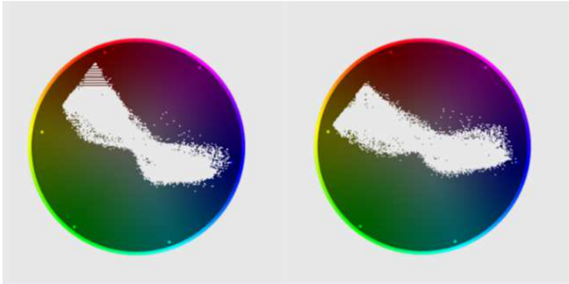

Figure 4. (a) Vectorscope of original image, (b) vectorscope of image with changed hue

monitors (Samsung UE32H6400AK) in the classroom. Large monitors were used to preserve the highest level of details. Viewing conditions were defined and monitor calibration was performed using OHSP-350 portable spectral irradiance colorimeter in accordance with the guidelines for the evaluation quality of the image [7]. Viewing distance was 0.6 m. Features such as color contrast and screen brightness were set to medium values to have as little impact on the test images as possible.

The application used for subjective evaluation was made in accordance with the ITU-R BT.500-14, Annex 4 to Part 2: *Stimulus-comparison method* [7]. The screens with a pair of images (A and B) were shown to the viewers for 15 seconds. The quality rating scale was continuous, and the viewers rated the quality of image B in comparison to image A. The rating scale informs about extent and direction of perceptible differences. 48 pairs were presented to each viewer in random order. The used approach tried to reduce the influence of habituation to the stimulus on the final average rating of the image.

Between the presentation of each two sets of images, a neutral gray stimulus was presented across the entire screen for four seconds. 40 viewers participated in the subjective evaluation, 20 male and 20 female. The average age of the observers was 25 years, and none had color blindness. Image grades from a continuous scale were transformed into a numerical value between -50 (test image has much better quality) and 50 (original image has much better quality). The central value, zero, means that the images have equal subjective quality. Each viewer's data was normalized to the specified range.

VI. Results

The average grades and standard deviations for each test image can be seen in Table III. The highest and lowest scores are highlighted. Results mostly consist of positive average grades, i.e., the viewers prefer original images. Five out of 24 average grades for hue change were negative. In all cases, these were images with slight color change, still looking realistic.

TABLE III. RESULTS OF SUBJECTIVE EVALUATION

Image name	Average grade	Std. dev.
0-h+10	-0.1339	21.5433
0-h+23	**36.5535**	24.6232
0-s+30	-8.3922	19.6136
0-s-30	10.9735	20.6512
1-h-17	18.4798	28.2183
1-h-8	5.6975	24.8762
1-s+40	7.4692	31.5220
1-s-40	27.8985	25.4346
2-h+15	-8.0268	31.8032
2-h+7	-10.4492	25.6877
2-s+30	-4.2712	23.0039
2-s-30	20.0747	20.8353
3-h-18	36.3571	19.9741
3-h-9	23.4630	18.0167
3-s+30	13.1395	21.5287
3-s-30	15.5958	23.4344
4-h+15	3.4265	31.8950
4-h+5	-4.0873	20.9004
4-s+20	**-11.4269**	20.8522
4-s-20	18.1506	20.1373
5-h-18	22.4983	24.6188
5-h-8	7.7232	24.6530
5-s+40	-0.6451	29.7545
5-s-30	13.3179	25.2289
6-h-18	29.1837	24.4424
6-h-8	14.1060	18.9789
6-s+30	-9.3857	25.9766
6-s-30	15.2942	22.4750
7-h-15	5.4823	22.3828
7-h-30	17.6010	26.3448
7-s+50	-10.7078	24.7573
7-s-50	27.1509	18.8270
8-h+15	12.3886	29.2171
8-h+30	8.3078	30.2851
8-s+20	3.2164	21.0035
8-s-30	18.3184	21.8107
9-h-15	4.5274	21.1231
9-h-8	-0.1985	20.9708
9-s+30	-2.6569	25.2649
9-s-30	17.5042	20.8985
10-h-15	14.4033	24.9969
10-h-8	4.3001	24.1741
10-s+30	-4.6445	26.3746
10-s-30	23.2552	18.9366
11-h-15	22.8802	23.2744
11-h-8	6.7313	24.7878
11-s+30	-4.7940	26.1297
11-s-20	17.1402	16.7314

The viewers liked the original image much more in cases where hue was significantly changed and therefore looked less realistic than the original. In Table III, it can be observed that nine out of twelve images with increased saturation are subjectively rated better than the original, while none of twelve images with reduced saturation are. Such an outcome agrees with the results of a previous study which concluded that people prefer more saturated images [14].

The objective evaluation of images includes seven measures used – MSE, PSNR, SSIM, CER, UCIQE, UIQM and CCF. The calculation of the UCIQE, UIQM and CCF method scores was carried out using the Platform for the Evaluation of the Quality of Underwater Images [15], which offers an online program for the evaluation of uploaded images using the mentioned methods. Analysis of data obtained from objective and subjective evaluation was performed using SPSS Statistics [16]. When comparing subjective and objective results, Kendall's tau and Spearman's rho coefficients are used. Both coefficients are non-parametric measures - they do not depend on the distribution of the results they analyze [17], [18]. Kendall's tau studies the dependence between two variables, and Spearman's rho observes deviations in values. Kendall's tau results in a lower coefficient than Spearman's rho and is less sensitive to errors and outliers. In most situations, the measures will result in the same conclusion [17]. However, there are marginal cases with disagreement about correlation significance. By calculating both coefficients, it is possible to get a better insight into the results. Table IV presents Kendall's and Spearman's correlation coefficients of subjective grades and objective quality measures. The mark "*" next to the number indicates that the correlation is statistically significant with a p-value of 0.05, and "**" indicates that the correlation is statistically significant with a p-value of 0.01. All such cases are highlighted in Table IV.

When changing color saturation, statistically significant correlation of subjective and objective assessment does not appear for almost any measure. Most of the coefficients are close to zero, which indicates no connection. The only method with a significant correlation is SSIM in RGB color space. In this case, SSIM is calculated separately for each color component, what is in accordance with HVS which processes images in similar way. Positive correlation indicates that SSIM and subjective grade change together. SSIM came closest to human answers in this example. One of the features that SSIM focuses on is color contrast, which changes when saturation changes. Methods for evaluating the quality of underwater images also evaluate contrast, but they are not as successful, possibly due to the algorithm adjusted for underwater color vision. Although color information without luminance in YUV and CIELAB color space is influential, it is not sufficient for significant correlation with the human opinion when using SSIM.

When changing hue, different levels of statistical significance can be found in multiple methods. SSIM is one of the most successful measures and achieves significant correlation in YUV and CIELAB color spaces, but not in RGB. The negative correlation indicates that the relationship between SSIM and subjective grade changes in opposing directions. It means that the viewers found test images that are more different from the original, more attractive. These higher rated test images still look natural. SSIM results correlate with subjective grades

TABLE IV. CORRELATION COEFFICIENT RESULTS OF SUBJECTIVE AND OBJECTIVE IMAGE QUALITY EVALUATION

	Saturation		Hue	
	Kendall tau-b	Spearman rho	Kendall tau-b	Spearman rho
	Subj. grade	Subj. grade	Subj. grade	Subj. grade
Subj. grade	1.000	1.000	1.000	1.000
MSE(rgb)	-0.094	-0.162	0.167	0.300
MSE(u*v*)	-0.065	-0.131	0.210	0.372
MSE(a*b*)	-0.065	-0.086	0.239	0.368
PSNR(rgb)	0.094	0.139	-0.283	**-.416***
PSNR(u*v*)	0.014	0.058	-0.210	-0.337
PSNR(a*b*)	0.130	0.176	-0.254	-0.362
SSIM(rgb)	**.449****	**.621****	0.130	0.213
SSIM(u*v*)	0.196	0.286	**-.413****	**-.561****
SSIM(a*b*)	0.159	0.220	**-.449****	**-.616****
CER	0.087	0.154	-0.196	-0.339
UCIQE	-0.145	-0.212	**.341***	**.454***
UIQM	-0.203	-0.272	**.333***	**.449***
CCF	-0.159	-0.217	**.326***	**.465***

* p-value ≤ 0.05 ** p-value ≤ 0.01

when the luminance component is removed. When changing hue, the method is effective when there is no information about luminance, which is always present in RGB space. Other methods that use only chrominance channels (MSE(u*v*), MSE(a*b*), PSNR(u*v*), PSNR(a*b*) and CER) do not achieve as good results as SSIM. SSIM(u*v*) and SSIM(a*b*) are appropriate methods for this type of quality assessment and are closer to human results.

MSE in RGB color space, PSNR in YUV and CER did not show a correlation when examining the images of the changed hue, as well as when changing the saturation. Methods for underwater images, UCIQE, UIQM and CCF, show a significant positive correlation with human opinion when hue is changed. These methods incorporate models that are able to evaluate hue effectively. More detailed research could use a larger number of subjects for subjective evaluation. The image database could be increased using more differently modified images.

VII. CONCLUSION

In this paper, the influence of changing color saturation and hue on image quality was examined. Image quality can be determined by subjective and objective methods. The subjective evaluation requires human observers. The disadvantage of these methods is the complexity of performance. Objective methods therefore use mathematical models to attempt to produce a human-like quality assessment in an automated manner.

A database of images with changes in saturation and hue was created. 40 viewers participated in the subjective examination. Images were objectively evaluated using MSE, PSNR, SSIM, CER, UCIQE, UIQM and CCF methods implemented in three color spaces – RGB, YUV and CIELAB. When changing the saturation, the SSIM(rgb) method showed significant correlation with the results of the subjective evaluation. When changing hue, the results are significant when using SSIM(u*v*) and SSIM(a*b*). However, not even SSIM shows the same level of correlation in all color spaces. Its success also depends on the type of image modification. In general, the objective quality measures show no significant success in representing human color evaluation. The human visual system is very complex and some of its parts cannot be described effectively by objective methods. Already with two different modifications of image color (hue and saturation), the inability of objective methods to accurately assess the impact of the change on the human experience of quality is noticeable.

REFERENCES

[1] WebMD Editorial Contributors, W. Seltman. (2021). *Vision Basics: How Does Your Eye Work?* Accessed: March 2023. [Online]. Available: https://www.webmd.com/eye-health/amazing-human-eye.

[2] R Nave. *Rods and Cones*. Accessed: April 2023. [Online]. Available: http://hyperphysics.phy-astr.gsu.edu/hbase/vision/rodcone.html.

[3] *Hue, Value, Saturation.* Accessed: April 2023. [Online]. Available: http://learn.leighcotnoir.com/artspeak/elements-color/hue-value-saturation/.

[4] J. Preiss, Color-Image Quality Assessment: From Metric to Application, PhD thesis, Technischen Universität Darmstadt, Darmstadt, 2015.

[5] K. Phillips. (2022). *What is CIELAB Color Space?* Accessed: May 2023. [Online]. Available: https://www.hunterlab.com/blog/what-is-cielab-color-space/.

[6] M. Podpora, G. P. Korbaś and A. Kawala-Janik, "YUV vs RGB – Choosing a color space for human-machine interaction," *FedCSIS*, 2014, pp. 29–34.

[7] Recommendation ITU-R BT.500-14 (10/2019). Methodologies for the subjective assessment of the quality of television images.

[8] P. Mohammadi, A. Ebrahimi-Moghadam and S. Shirani, "Subjective and objective quality assessment of image: A survey." https://arxiv.org/abs/1406.7799, 2014.

[9] I. Žeger, N. Bilanović, G. Šišul and S. Grgić, "Comparison of Metrics for Colorized Image Quality Evaluation," *International Symposium ELMAR*, 2022, pp. 209-214.

[10] N. Bilanović, Objective image quality assessment methods for colorized images, MSc thesis, Faculty of Electrical Engineering and Computing, Zagreb, 2022.

[11] Y. Wang, N. Li, Z. Li, Z. Gu, H. Zheng, B. Zheng, M. Sun, "An imaging-inspired no-reference underwater color image quality assessment metric," in *Computers & Electrical Engineering*, vol. 70, pp. 904-913, 2018.

[12] S. Laguerre. (2016). *Vectorscope Image Analyzer.* Accessed: May 2023. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/56546-vectorscope-image-analyzer.

[13] M. Manickam. (2019). *What is Saturation? And How to Get Optimal Saturation*. Accessed: February 2023. [Online]. Available: https://photographylife.com/what-is-saturation-and-how-to-get-optimal-saturation.

[14] E. Kumakura, K. Schmid, K. Yokosawa and A. Werner, "Subjective evaluation of natural high saturated images on a wide gamut display," in Color Research and Application, vol. 44, no. 6, pp. 886-893, 2019.

[15] C. Y. Li, R. Mazzon, A. Cavallaro. (2020). *Underwater image filtering: methods, datasets and evaluation, Platform for Underwater Image Quality Evaluation*. Accessed: February 2023. [Online]. Available: https://arxiv.org/abs/2012.12258, https://puiqe.eecs.qmul.ac.uk/.

[16] IBM SPSS Software. Accessed: February 2023. [Online]. Available: https://www.ibm.com/spss.

[17] *Kendall's Tau and Spearman's Rank Correlation Coefficient.* Accessed: May 2023. [Online]. Available: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/kendalls-tau-and-spearmans-rank-correlation-coefficient.

[18] G. A. Fredricks and Roger B. Nelsen, "On the relationship between Spearman's rho and Kendall's tau for pairs of continuous random variables, " in *Journal of Statistical Planning and Inference*, vol. 137, pp. 2143 – 2150, 2006.