Color correction judgements of digital images by experts and naive observers

Clotilde Boust [1,4], Hélène Chahine [1], Françoise Viénot [2], Hans Brettel [3], Mohamed Ben Chouikha [4], Georges Alquié [4]

[1] Océ Print Logic Technologies (Créteil, France), [2] Museum National d'Histoire Naturelle (Paris), [3] Ecole Nationale Supérieure des Télécommunications-CNRS (Paris), [4] Université Pierre et Marie Curie (Paris)

Abstract

We have studied the preference judgement of pictorial images with two types of population: image experts and naive observers. We used eight images of indoor and outdoor scenes. The images were first presented to image experts working in pre-press companies. We asked them to improve the images the way they preferred and to comment on their manipulations. They saved the intermediate enhanced versions they judged important. In the second part of this study, we showed those different versions of the same image produced by experts to naive observers. We used a pair comparison protocol in which all possible pairs were presented. The question was **"**which image do you prefer?" Observers told us their criteria for choosing an image. Several results emerged from this study. To enhance an image, an expert divides it into large zones of interest, which mainly correspond to natural colors. Likewise, when judging an image for preference, naive observers principally focus on natural colors like sky, skin or grass when present. Both experts and naives do not focus on parts if no memory colors are associated. The segmentation process into zones permits to first adjust the illuminant and then to correct the other parts with respect to the plausibility and coherence of the whole image. Further studies are necessary to correlate the chromatic signal in the retina to those data.

Introduction

Quality corrections of images often concern experiments with images displayed on screen [3,4,6,9]. Some works, dealing with digital images printed on paper, were reported in the literature [1,2,7,8,]. In general, the attributes that are supposed to influence quality and preference are tested with observers but the variations of those attributes are often incremented systematically and are applied on the whole image[1,2,4,7]. However, professionals working in the image industry, who are responsible for the quality to be approved by the public, do not improve images this way. Image experts first begin to divide the images into large zones. They manipulate each zone separately in order to enhance the final quality of the image. The aim of this paper is to learn from image expert's behavior in image enhancement, and to validate the segmentation procedure by a panel of naive observers.

Experimental procedure

Material

The printer used for this experiment was the Epson Stylus Pro 7600 ink jet printer. It was driven by the "Postershop" Raster Image Processing software (RIP) from Océ. We used Semi gloss Premium Epson Photo Paper 162g and pigmented inks Ultrachrome that are stable to light.
The printer was calibrated and characterized with the RIP's calibration menu and the spectrophotometer Gretag Spectroscan.

We used eight pictorial images of indoor and outdoor scenes chosen from Photodisc image CD n°15 and 60, and from Kodak Photo CD Photo Sampler V2. The images chosen are shown in Figure 1. They were RGB images in JPG or PCD format at 1536 x 2296 pixels or 2048 x 3072 pixels. Upon manipulations in Photoshop software, all images were transformed for printing. The final characteristics of images were 17 cm x 26 cm, with a resolution of 220 dots per inches, in a CMYK Tiff file. To separate the images from RGB to CMYK, we used the color management of Adobe Photoshop and the ICC profile of the Epson 7600 given with the printer. We transformed from the RGB source mode Adobe RGB 1998 to the Epson 7600 destination profile, in the perceptual mode using the Adobe color management engine.

Figure 1. Test images used in the experiments

Expert enhancement of images

We gave the eight images to two image experts. The first expert works in a pre-press company that treats images for travel, television and art magazines. She studied in a fine art school. The second expert works in a major advertising company and went to a photographic school.

The experts did the experiment in their own company, using Photoshop software with a color calibrated workstation. The screen and the software were tuned in order to simulate the printed result on the Epson 7600 seen under daylight. Although screen proofing has limitations and cannot be considered as perfect, it is the way those experts are used to working in their companies.

We asked the experts to improve the images the way they wanted in order to produce the image they prefer. They had to save each intermediate version of the enhanced image, each time they thought it was an important step of quality improvement. It provides several versions of the same image, between two and five versions. The experts were also asked to produce other plausible versions of enhancement. It gives some alternative versions of the same image. The experts commented their manipulations while they were working.

Despite their different background, the two experts had the same methodology. They segmented images with large zones, furthermore they chose the same zones. Thus we used the set of images enhanced by the first expert to run our experiment. We printed all the images on the Epson 7600 in one batch, including calibration charts in each printing page.

Psychophysical experiments with naive observers

The printed sheets were presented to observers in a Gretag light box with dimensions of 60 x 50 x 45 cm. We measured a color temperature of 5112 K with the Minolta spectrophotometer CS-1000. The average light level in the room was variable due to the presence of a window. The dimensions of the cabin were important enough to consider that external light had no influence on the perception of images.

We worked with ten volunteer observers who were all computer scientists at Océ P.L.T. They were naive to the experiment and to color image processing. There were five men and five women, between 23 and 32 years old. We conducted a preliminary experiment before the major one.

Preliminary experiment

We chose six versions of each image among the original and the five to eight versions produced by the expert to make one set. We usually eliminated the first step of correction and added two alternative versions of corrections when available.

We showed the six versions to naive observers. We used a pair comparison experiment where all the possible pairs were presented. For each pair, we asked them “which image do you prefer?" At the end of each set of images, we asked them “what are your criteria to choose your preferred image?”. In this preliminary experiment, each set of images has been seen at least three times but not all the observers did see all the set of images.

In a pair comparison experiment where all pairs are presented, with six different versions of the same image, fifteen pairs are therefore shown in one set. When presented to naive observers, the six versions of one image were noted randomly A, B, C, D, E or F. In our experiment, images were presented in this order: AB, AC, AD, AE, AF, BC, BD, BE, BF, CD, CE, CF, DE, DF and EF.

We used the results of the preliminary experiment as a selection procedure to include the most preferred image in our experiment. We added some images not shown in the selection test, like the first step of correction or the original image and we eliminated some versions rejected by observers. Consequently, the original image was not always present in the experiment.

Experiment

We used here the same pair comparison technique as described in the preliminary experiment. The ten observers

saw all the sets of images with three repetitions. Observers did six different sessions over ten days where they saw each time four different complete sets of images. The order of appearance of sets in one session was distributed equally and is given in Table 1. The question asked to the observers was "which image do you prefer?" They had no time constraints to give their answer.

Table 1: order of images showed to observers

Session 1	Buddha	Lighthouse	Taj Mahal	Garden
Session 2	Statue	Girl	Hand	Baby
Session 3	Girl	Statue	Buddha	Lighthouse
Session 4	Baby	Hand	Garden	Taj Mahal
Session 5	Hand	Taj Mahal	Statue	Buddha
Session 6	Garden	Baby	Lighthouse	Girl

In this article, the steps of the image corrections made with segmentation are called A1, A2 etc. The steps of alternative corrections are called B1, B2 etc. Other alternative corrections are called C1, C2 etc. The original image is noted O.

Results and Discussion

Expert enhancements strategy

Before modifying an image, the experts first think about the purpose of the image. The first expert may treat differently images that will be used for travel magazines or images that will be used for advertising. She explained for example that she usually increases saturation for travel images, even if this leads to something unnatural. Her methodology confirms that image manipulations should be linked to the purpose of images [4,5]. Here, the expert first treated images the way she would prefer and the alternative versions usually correspond to alternate interpretations of images.

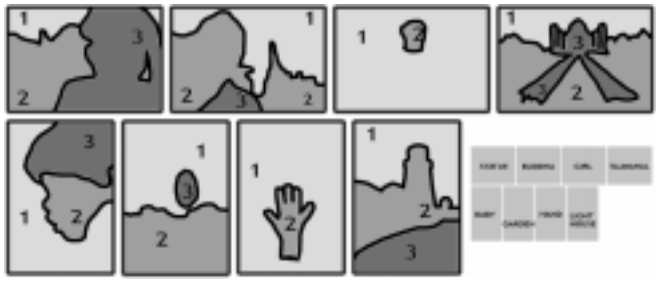

Figure 2. Segmentations made by the first expert

The two experts started the image corrections by segmenting them in a few large zones of interests. They usually divided them into only two or three parts. The segmentation is used to create masks in Photoshop in order to work only on the part they chose. The experts told us to particularly focus on natural colors like sky, greenery and skin tones and to almost ignore some parts. They usually selected and modified parts in image representing one natural element. The first expert told us she did not change the Taj Mahal color for example because she did not know what to correct. Figure 2 shows segmentations made by the first expert.

To correct images the experts changed attributes like hue, lightness, contrast or saturation. They mainly used Photoshop's menus called "tonal curves", "brightness / contrast" and "hue / saturation". When they treated the image, the experts changed attributes of segmented parts until they thought the whole image would be the one they

Table 2. Images transformations by first expert, with segmented parts and the preference rank for naives.

	Image	Segmented part	Transformations	Rank
Hand	**O**	-	-	**4**
	A1	**1**	Change background color	**2**
	A2	**2**	Change skin color	**1**
	B1	**2**	Increase hand contrast from A2	**3**
	C1	**all**	Change global color from O	**5**
	D1	**all**	Change global color from O	**1**
Garden	**A1**	**1, 3**	Change green color	**5**
	A2	**3**	Change hose pipe color	**4**
	A3	**1**	Change red wall color	**3**
	A4	**all**	Increase global contrast	**2**
	A5	**2**	Change grass color	**1**
	B1	**all**	Change global contrast from O	**6**
Baby	**O**	-	-	**3**
	A1	**1**	Remove color from blanket	**5**
	A2	**2**	Change skin color	**2**
	A3	**3**	Change hair color	**1**
	B1	**1**	Modify blanket color from A3	**6**
	C1	**2**	Change skin color from A3	**4**
Girl	**O**	-	-	**3**
	A1	**2**	Change skin color of face	**1**
	A2	**2**	Change color of teeth	**2**
	A3	**1**	Change background saturation	**4**
	B1	**all**	Change color from A3	**6**
	C1	**all**	Change global color from O	**5**

	Image	Segmented part	Transformations	Rank
Lighthouse	**A1**	**all**	Change white point of image	**4**
	A2	**1**	Change sky color	**2**
	A3	**all**	Change global contrast	**1**
	A4	**2**	Change lighthouse color	**3**
	B1	**all**	Change global contrast from O	**6**
	C1	**3**	Change grass color from A4	**5**
Buddha	**A1**	**2**	Remove color from stone	**2**
	A2	**1**	Change sky color	**1**
	A3	**all**	Change global contrast	**4**
	B1	**1**	Change sky color from O	**6**
	B2	**3**	Change clothe color	**5**
	B3	**2**	Change shadow color	**3**
Statue	**O**	-	-	**1**
	A1	**1**	Change sky color	**2**
	A2	**2**	Change background contrast	**3**
	A3	**2**	Change background saturation	**4**
	A4	**2**	Change green color of tree	**5**
	B1	**all**	Change global color from O	**6**
Taj Mahal	**O**		-	**2**
	A1	**all**	Change global contrast	**3**
	A2	**1**	Change sky color	**6**
	A3	**2**	Change grass and water color	**4**
	B1	**all**	Change color contrast from O	**5**
	C1	**2**	Change grass and sky from O	**1**

preferred. It could be an iterative process; the experts know the influence of adjacent color zones. For example, the expert told us that she first changed background color to integrate its influence on the skin color for the baby image. For the alternative corrections, the expert usually operated on the whole image without segmentation. She changed attributes applied to the entire image. The different steps of corrections are summarized in Table 2.

Naives observer preference

Criteria for preference

Naive observers generally chose their preferred image without difficulties, it seemed to be quite an easy task. The naives told us they fixed their attention on some precise elements in the image and classed their importance. They mainly focused on natural colors like sky, greenery and skin tones. They did not tell us they concentrated on parts with unknown color objects like the Taj Mahal color or the clothes of Girl.

Scales of image preference

In order to classify the images according to a preference scale, we transformed the percentages of choice into z-scores. Z transformation is usually used for unidimensional factors. However, it yields to readable results with multidimensional studies as well [3]. Scales of image preference were calculated following the Engeldrum's solution for Thurstone's case V solution [12]. The different z-scores for all images are plotted in Figure 3.

When we study the naives' preferences with the expert's manipulations, we distinguish several cases. In three images, Garden, Hand and Baby, naive observers prefer the final expert correction obtained after segmentation. For Baby and Hand, the step correction of skin tone dramatically increases the preference. The expert first changed the background of the skin and then the skin tone. She fixed the background color before focusing on the most important part of the image. Only in one image, Hand, a version made with global correction obtained the higher z score value but on equality with a segmented version. For the Garden image, the four steps of correction increase a lot the preference. The changes concerned the grass, then the red of the wall, the global contrast and the grass again. The grass was first modified as an important part of image preference and re-modified at the end of the process to ensure that the colors are exactly the wanted ones.

In three cases, for Buddha, Statue and Taj Mahal, the corrections were not chosen by naive observers. The preference decreases with manipulations. Only an alternative version of Taj Mahal equals the preference with the original one.

In the last two cases, Girl and Lighthouse, naive observers prefer an intermediate version of enhancement with segmentation. For Lighthouse, the preferred image is the third one of four corrections. The expert changed the image white point, the sky color, global contrast and sky and then contrast. The other correction is global on contrast and is rejected. For Girl, the preferred image is the first step of correction out of three. In this step, the expert changed only the skin tone of the girl's face. The second one is a little less preferred, she put white on the girl's teeth. The decrease of the saturation of the background made the preference lower than under the original image. As with the other images containing skin tones, Hand and Baby, we can notice the very important increase of preference after the correction of skin tones.

Final discussion

When we look at the parts isolated by the expert, we see that they correspond mainly to elements like the sky, the grass and the skin when they were present. In all the images where the sky is present, it became a segmented part. On three of the four images where grass or greenery is present, it became a segmented part. The expert worked first on colors associated to well known objects. It mainly corresponds to natural objects that we are used to seeing, with associated memory colors.

When we asked the naive observers about the criteria they used to choose images, they answered with almost no hesitations. They indicated the natural parts such as sky, skin and grass/ tree colors to be the most important parts, with a priority on sky when it was present. They said they do not look much at parts when they don't know the original color. In the Taj Mahal image, the Taj Mahal seems to be the principal subject but they don't really focus on it because they don't know its real color. Observers seemed to often have a precise idea of how the color parts should be. It

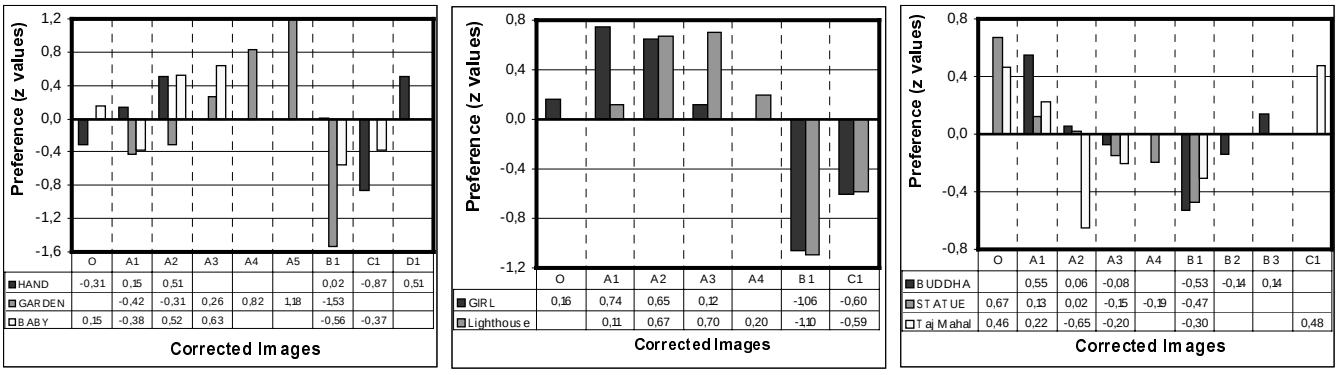

Figure 3. Z scores of image preference by naive observers

seems that they compare the shown color to their existing internal representations. This could be related to works on memory color from Bodrogi and Tarczali [9] or from Yendrikhovskij et al [6].

While correcting the colors of images, the expert usually begins by modifying the sky. By doing so, she corrects the color temperature of the illuminant. Once the illuminant is fixed, all the colors should be coherent with this choice. Then, all corrections will aim at constructing the color coherence of the scene. In other words, the scene should appear the same color as in a natural situation. When the expert changed the color of the blanket behind the baby's face, she was changing the highlight that reflects the illuminant.

Within this scheme, the second operation tends to adjust the color of the present objects, like grass, to the color of the illuminant. The result could be pre-known from the expert's experience in the past. As the grass gets its color from chlorophyll and has a typical spectral reflectance, not all corrections are allowed. A similar scheme shows up for skin correction. Once the expert has fixed the illuminant, either directly correcting the sky, or indirectly correcting the highlights of the bed blanket, she can make the corrections within tolerances allowed by a typical healthy skin reflection. The spectral reflectance of hemoglobin is extractable from the skin reflectance [10]. The weight of this component depends upon the degree of oxygenation of the skin tissue as well as the age of the subject. We are probably well trained to the control of the color of the skin. It may be inherited from the evolution of mankind [11].

The segmentations made by the experts permit to control the coherence of the whole image. The expert stopped her corrections when she found plausibility or coherence in the image, founded on her experience.

For the images called Buddha, Statue and Taj Mahal, the expert correction failed to improve the image for the naives' judgements. It could be because the only well known color is the sky, we know those images were taken in an exotic country we never saw. There may be a lack of memory colors in those images, so it is a difficult image to correct and to judge. The expert judged those corrections difficult to do. Or, may be these images are already coherent and do not need improvements.

In order to further understand the concept of coherence, we need the results of the colorimetric study of showed images.

Conclusion

To enhance an image, an expert divides it into zones of interest. They mainly correspond to natural colors. To judge an image for preference, naive observers principally focus on natural colors like sky, skin tones or grass when presented. Both experts and naives do not focus on parts if no memory colors are associated with them. When we look at images preferred by naives with the corrections made by experts, we see that naive observers preferred five corrected images with segmentation out of eight. One image is preferred equally with segmented correction and with a global correction. For two images, all the corrections are rejected and the preferred image is the original one. The segmentation process first permits to adjust the illuminant and then to correct the other part with respect to the coherence of the whole image. Further studies are necessary to correlate the chromatic signal in the retina to those data.

Acknowledgements

The authors would like to thank the experts, Marie Raymond-Bertrand and Carine Cottereau, and all the observers for their participation.

References

1. S. R. Fernandez, M. D. Fairchild, Preferred Color Reproduction of Images with unknown Colorimetry. *Proc. 9th Color Imaging Conference*, p. 274-279. (2001).
2. S. R. Fernandez, M. D. Fairchild, Observer Preferences and Cultural Differences in Color Reproduction of Scenic Images. *Proc. 10th Color Imaging Conference*, p. 66-72. (2002).
3. G. M. Johnson and M.D. Fairchild, Sharpness Rules, *Proc. 8th Color Imaging Conference*, p 24-30. (2000).
4. J. A. J. Roufs, Perceptual image quality: Concept and measurement. *Philips J. Res*. **47** n°1, 35-62, (1992).
5. S. Yendrikhovskij, Colour Categorisation and Colour Imaging, *Colour Imaging Science 2000*, Derby, (2000).
6. S. Yendrikhovskij, F. J. J. Blommaert, H. de Ridder, Color Reproduction and the Naturalness Constraint, *Color Research and Application*, 24/1, (1999).
7. J. Uroz, M.R. Luo and J. Morovic, Colour Difference Perceptibility for Large-Size Printed Images, *Proc. Colour Imaging Science Conference*, pg. 138-151, (2000).
8. J. Morovic, P. Sun, Visual Differences In Colour Reproduction And Their Colorimetric Correlates, *Proc. 10th Color Imaging Conference*, p. 292-297. (2002).
9. P. Bodrogi, T. Tarczali, Colour memory for various sky, skin, and plant colours: effect of the image context, *Color Res. Appl*. **26/4**, 278-289, (2001).
10. S. J. Dain, J. W. Hood, S. Montano, C. Arali, A method for evaluating the acceptability of light sources for clinical visual evaluation of cyanosis, *Color Res. Appl*. **23/1**, 4-17, (1998).
11. J. S. Werner, M. Webster, Color vision is form and object vision, *Proc. 9th Congress of the international Colour association*, pg. 10-15. (2001).
12. P.G. Engeldrum, *Psychometric Scaling*, Imcotek Press, Winchester, MA, 2000, pg. 93

Biography

Clotilde Boust received her engineering degree in photography from the Ecole Nationale Supérieure Louis Lumière, France in 1998. After working for two years as color consultant in the press industry and one year as researcher in the Vision laboratory of the Museum National d'Histoire Naturelle, she began a Ph.D. in image quality with Océ Print Logic Technologies and Paris VI University.