Eye Movements on Assessing Perceptual Image Quality

Cheng-Min Tsai[1](✉), Shing-Sheng Guan[2], and Wang-Chin Tsai[3]

[1] Department of Creative Product Design and Management, Far East University, Tainan, Taiwan
ansel.tsai@gmail.com
[2] School of Design, Fujian University of Technology, Fuzhou, China
[3] Department and Graduate School of Product and Media Design, Fo Guang University, Yilan, Taiwan

Abstract. The purpose of this study is focusing on "Analysis for human's Region of Interest (ROI) on complex images", and collected the subject's eye movement and 'scan path' when assessing perceptual image quality task by using eye tracking method. The participants in this study were 30 students with convenience sampling from design, management and engineering college in National Yunlin University of Science and Technology. 11 stimulus of this experiment was selected from ISO standard image by focus group, and the type of image was divided into still image, portrait, landscape and architecturally image. The two tasks in this experiment are 'assessing perceptual image quality' and 'assessing perceptual color quality'. And analysis the 'Fixation duration' and 'Amplitude of saccade' of ROIs from eye movement when assessing perceptual image quality task. The results show 1. when the 'perceptual color quality' is higher and the 'perceptual image quality' is higher, too; 2. The subjects were tended to browse the still and landscape image widely when assess the perceptual image quality and perceptual color quality task; 3. In contrast to the still image, the subjects were focus on the face of portrait image; 4. The subjects' eye movement was tended to stabilize and homogenize when assess the 'perceptual color quality task'.

Keywords: Image quality · Eye tracking · Visual assessment

1 Introduction

The factors composing image quality are highly complex. From the perspective of research and development, human factors engineering must also be taken into account in addition to the physical attributes of images. The ultimate purpose of displays is to correctly transmit all image content to human visual systems. However, many past studies have emphasized the physical attributes of images as a basis, using algorithms to calculate video compression or video coding to examine the physical quality of images. Studies examining the subjective assessment of perceptual image quality from a psychological perspective are a more recent development. Newell [1] has figure out the three levels for information processing system (IPS). Those were included Physical implementation, Algorithmic manipulation, and Semantic understanding. Maeder and Eckert [2] figure out the three levels for cognition processing system with human's

J. Zhou and G. Salvendy (Eds.): ITAP 2016, Part I, LNCS 9754, pp. 378–388, 2016.
DOI: 10.1007/978-3-319-39943-0_37

cognition. Those were included Mathematical, Psychovisual, and Task oriented. Thus, the aim of this study was focusing on the perceptual image quality base on the information processing system and cognition processing system. As the Fig. 1 showed the inverted triangle is the research scope of perceptual image quality in user-centered approach. The cognition level is widely research approach by implementing human factor research to perform empirical study [1, 2].

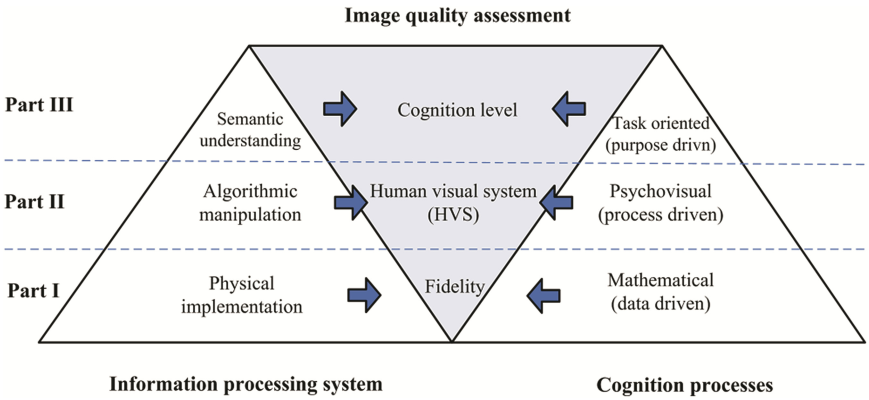


Fig. 1. The research scope of perceptual image quality in user-centered approach (Color figure online)

Over the past ten years, many studies have begun to examine assessments of image information and visual quality through human visual systems based on perspective of human factors engineering and perception psychology. Some studies have examined the relationship between image information and regions of interest (ROI) [3–5]. Other studies have established predictive models or derived algorithms based on the focuses of visual attention during the viewing of images [6, 7]. In the past five years, a number of studies involving visual assessment methods have attempted to examine the primary factors in static perceptual image quality based on human visual systems [3, 4, 8, 9]. However, these studies have yet to reach a consensus. Over the past 20 years, many studies on complex images have emphasized the physical measurement of image quality, most commonly using peak signal-noise ratio (PSNR) and root mean square error (RMSE) statistical analysis techniques to assess differences in image quality. However, these image quality measurement methods are primarily intended to calculate the fidelity of processed images to original images. Consequently, it may be that calculated image quality cannot be directly analogized or applied to the perceptual image quality judgment standards of the human eye. In recent years, many visual psychology studies have attempted to use mathematical algorithms to predict the location or positions of interest within a specified space. Privitera et al. [6] used a series of studies to predict the scan-paths of eye movement. His studies involved the use of geometrical spatial kernels and linear filter models for analysis, calculating algorithms of regions of interest (aROIs) to predict the regions focused on by humans when they viewing the images (human ROIs, hROIs). Nguyen et al. [10] established and grouped hROIs based on analysis of scan-paths and sequences of fixation occurring when humans viewing grayscale images; they further performed local image compression based on the ROIs. However, the images normally viewed by humans through their visual systems are all color images. Whether the results of the relevant research described above can be generalized to all visual focuses remains to be clarified.

As the eye is the first element of the visual system to receive visual information, it is also the only means by which the brain obtains external images. Consequently, the positioning and movements of the eye are an important index for observation when people focus on or view visual information. The emphasis of this study differs from previous image quality assessment methods and attempts to use a different research approach by using human factors research to perform empirical study. The purposes of this study are twofold: (1) to understand the "perceptual image quality" and "perceptual color quality" subjectively perceived by humans; (2) to understand the correlation of eye movement information when humans assess "perceptual image quality" and "perceptual color quality." The results obtained through this study can provide producers and designers of LCD television color chips with a subjective standard for understanding human judgments of perceptual image quality.

2 Research Method

2.1 Experiment Material and Environment

The LCD-TV was adopted as the sample-displaying monitor. Moreover, GretagMacbeth Eye-One was adopted to conduct the characterization correction and establish the ICC profile. The study was conducted in a laboratory with fixed luminance control as 233 lx and the color temperature set as 6,500 k.

2.2 Eye Tracking Technique

The video-based, pupil/corneal reflection eye tracking apparatus used was an infra-red eye movement recording system (EyeLink II) manufactured by SR Research Ltd, Canada. The subjects were seated facing a calibrated 30 in. Sharp LCD-TV (30 cm high and 40 cm wide) which was 120 cm away. The visual angle of the whole screen is 36.8 degrees wide, 28.1 degrees high. The visual angle of each single stimulus is 6 degrees wide and 6 degrees high. The monitor has a vertical scan frequency of 60 Hz, and a resolution of 1280 × 768 pixels. Subjects wore a headset containing a camera which monitored and recorded their eye movements and fixation locations (Fig. 2).

2.3 Image Stimuli

There were 11 stimuli have been selected from ISO standard image by focus group that were study in our color and image lab, and those image including still image, portrait, landscape and architecturally image (Table 1). All the images were be selected from

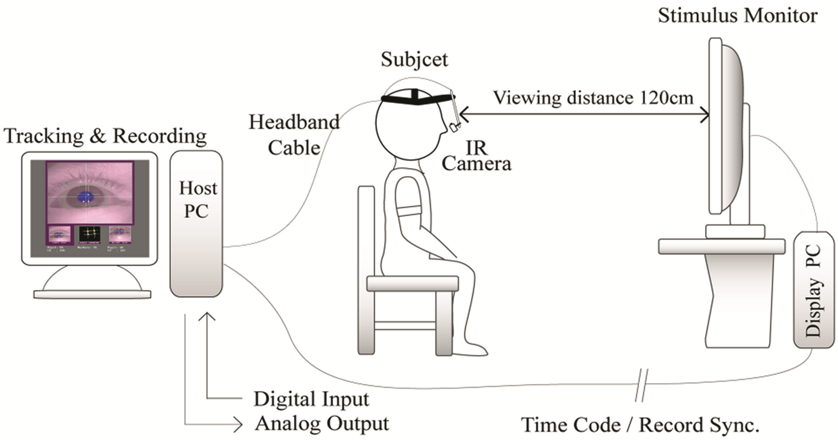
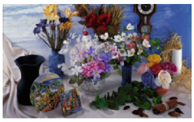
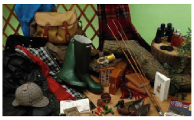
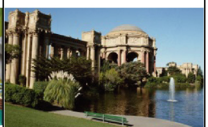
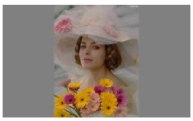
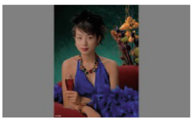


Fig. 2. The eye tracking apparatus set for experiment condition (Color figure online)

ISO 12640-1 (1997), ISO 12640-2 (1997), ISO 12640-3 (2004), ISO 12640-3 (2007) and Kodak standard image database by focus group. Each stimulus was displayed for 5 s, and then a calibration point popped up for drift correction to avoid systematic shift and after-image effects.

Table 1. The image stimuli set for experiment

No.	1	2	3	4
Image				
No.	5	6	7	8
Image				
No.	9	10	11	
Image				

2.4 Image Stimuli Set for Experiment

The interface of experiment window was design by Visual Basic software. The size of single image was 1280 by 768 (pixels). The images were randomly displayed at each experiment where the background was middle grey (RGB were set to 128) each time. Constructed from random noises, a noise frame showed at the screen after subjects' completion of image quality assessment to avoid the after-image effect generated from the previous stimulus image.

2.5 Subjects

The subjects were 30 undergraduates from the design, management and engineering college at National Yunlin University of Science and Technology in Taiwan. Those with less than perfect vision had had their vision corrected using glasses or contact lenses, so that all subjects were able to see normally. All of them had normal color vision according to the Ishihara color vision test.

2.6 Experimental Procedure

Two stage were be set for experimental procedure. First, after subjects reported to the laboratory and passed the color vision test, they should read the instructions of the experiment and practiced participating in it. A process to calibrate and validate the eye tracker was performed by having each subject fix the location of 9 points on the calibration screen. Once the calibration and validation was done, the subject started to view each of the displayed frames. Second, each participant should to assess 11 images with 7 scale of Likert scale on two tasks those were "perceptual image quality" and "perceptual color quality".

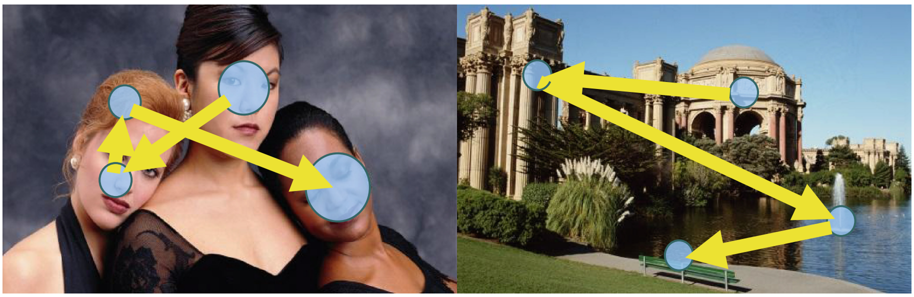


Fig. 3. Sample of fixation duration (left); Sample of amplitude of saccade (right) (Color figure online)

2.7 Data Analysis of Eye Movement

The serial order in which the images were projected was randomized by the computer so that the image order changed every trial. The data of eye position were collected into host PC with eye tracking system. Two raw data from eye position were “Fixation Duration (FD)” and “Saccade Amplitude (SA)”. The FD is mean that “The time that subject gaze the point of image, such as regions of interest of image (as left image of Fig. 3, the blue point was count for fixation duration)”; The SA is mean that “The amplitude of saccade when subject viewing the image (as right image of Fig. 2, the yellow arrow line was the amplitude between two gaze point)”.

3 Data Analysis and Results

3.1 Subjective Assessment Analysis

Data analysis was according to participants’ subjective assessment the perceptual image quality and perceptual color quality. One-way ANOVA results demonstrated that was no significant variation between the assessment results for “perceptual image quality” and “perceptual color quality” ($F_{(1, 660)} = 3.331, p > .05$). On the other hand, correlation analysis performed between the two revealed that the assessments of “perceptual image quality” and “perceptual color quality” exhibited significant direct correlation ($r = .684$, $p < .001$). These subjective assessment results suggest that higher color quality will lead to correspondingly higher overall image quality. Particularly, for the four images of “Images 5, 7, 8, and 2,” subjects gave these images relatively high scores whether they were performing the assessment tasks for “perceptual image quality” or for “perceptual color quality.” Conversely, when subjects performed “perceptual image quality” and “perceptual color quality” assessment tasks for the two portraits in “Images 11 and 10,” the images were assessed as being of lower quality.

3.2 Eye Movement Analysis

This study also analyzed the scan-paths exhibited by subjects as they viewed images in the image assessment process; the scan-path analysis included the two indices of ‘fixation duration’ (FD) and ‘saccade amplitude’ (SA). The former refers to the continuous time for which the eyes fix upon a specific location during repeated visual search processes (ex. Scan-paths) when viewing images. Scan-paths also include other important information such as SA. SA can reflect whether image viewing consists of fixation on localized regions or global browsing.

SA Variance Analysis. As eye movement information is massive, outlier tests and SA standardization must be taken into account prior to statistical analysis in order to ensure the accuracy of the information. Analysis was performed for the SA produced in scan-paths when subjects viewed images of different qualities. One-way ANOVA revealed that significant differences did exist in SA ($F_{(10, 9771)} = 57.288$, $p < .001$); at the same time, the Duncan test showed that when subjects assessed perceptual image quality, the SA was greatest when they viewed “image 5” (see Fig. 4). In other words, when subjects assessed the perceptual image quality of “image 5,” they tended to browse the image in a global fashion. In addition, in perceptual color quality assessment, One-way ANOVA showed that there were significant differences in the SA for different images ($F_{(10,}$

$_{7090)} = 37.921$, $p < .001$). The Duncan test revealed that when subjects assessed perceptual color quality, SA was the greatest when "image 4" was viewed (see Fig. 5). In other words, when subjects assessed the perceptual color quality of "image 5," they tended to view the image more globally and also spent more time evaluating the image. These results are consistent with those produced when subjects assessed the perceptual image quality of "Image 5."

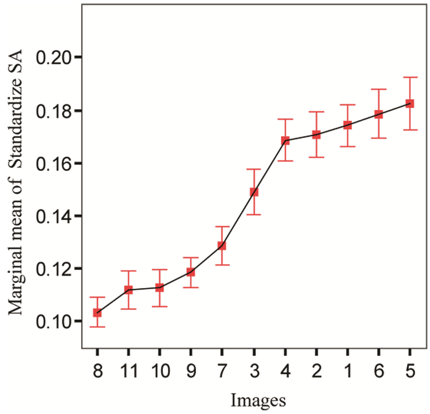


Fig. 4. SA variance analysis for "perceptual image quality" assessment of different images

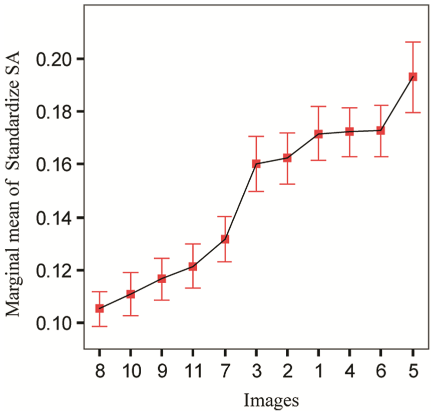


Fig. 5. SA variance analysis for "perceptual color quality" assessment of different images

FD Variance Analysis. Analysis of the visual fixation duration exhibited by users when viewing images of different qualities reveals that there are significant differences in FD between different images ($F_{(10, 9806)} = 3.366$, $p < .001$). At the same time, Duncan testing showed that when subjects assess perceptual image quality, FD was shortest for "images 5 and 4." In contrast, the greatest FD was associated with "image 11" (308 ms on average) (see Fig. 6). In addition, in assessment of perceptual color quality, One-way ANOVA showed that there are significant differences in FD between different images

($F_{(10, 7108)} = 3.167$, $p < .001$). Duncan testing showed that when subjects assess perceptual color quality, “image 7” had the shortest FD (see Fig. 7); by contrast, “image 11” had the longest FD (307 ms). This result is consistent with the result obtained from subjects evaluating the perceptual image quality of “image 11.”

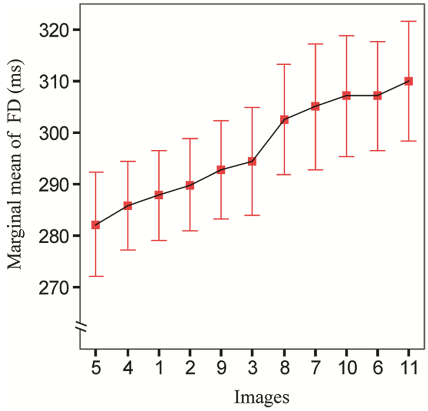


Fig. 6. FD variance analyses for “perceptual image quality” of different images

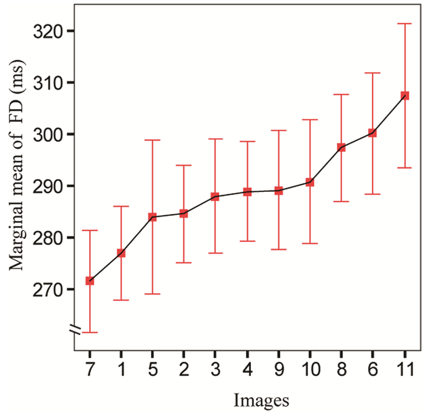


Fig. 7. FD variance analyses for “perceptual color quality” of different images

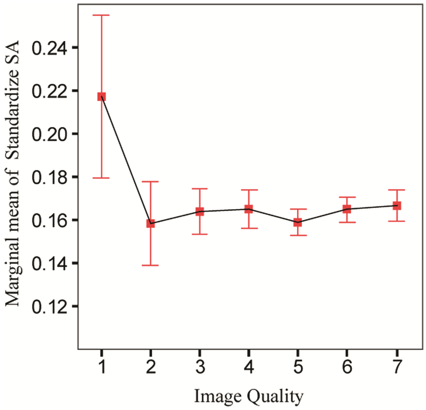


Fig. 8. SA variance analysis for assessments of “perceptual image quality”

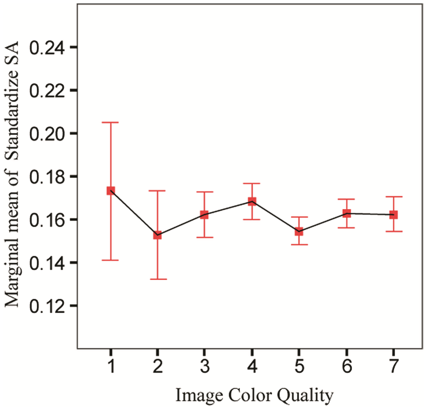


Fig. 9. SA variance analysis for assessments of “perceptual color quality

Correlation Analysis Between “Perceptual Image Quality” and “Perceptual Color Quality” in SA. Variance analysis was performed for scan-path in “perceptual image quality” and “perceptual color quality” assessments by users. Results showed that there was no significant variation between SA of scan-paths when subjects assessed “perceptual image quality” and “perceptual color quality” ($F_{(1,\ 16861)} = .910$, $p > .05$). In other words, correlation potentially existed between the SA of scan-paths in “perceptual image quality” and “perceptual color quality.” Consequently, correlation analysis was performed for the relationship between those two; results showed that there was significant direct correlation between the SA of “perceptual image quality” and “perceptual color quality” ($r = .447$, $p < .001$). These results demonstrate that when subjects assessed “perceptual image quality” and “perceptual color quality,” a certain degree of consistency existed in the browsing behaviors associated with the two types of assessments. In addition, One-way ANOVA showed that when subjects gave lower “perceptual image quality” assessments, SA was significantly higher than in other assessment results ($F_{(6,}$

$_{6344)} = 1.881, p < .001$) (See Fig. 8). In other words, speaking generally, when subjects believed that an image had lower quality, their scan-paths browsed images with large amplitudes of saccade; by contrast, when subjects assessed “perceptual color quality,” image browsing strategies were not affected by the assessment of “perceptual color quality” ($F_{(6, 5597)} = 1.385, p = .217$) (See Fig. 9).

4 Conclusion

4.1 Results of Subjective Assessment for “Perceptual Image Quality” and “Perceptual Color Quality”

A summary of the experimental results from this study reveals that “perceptual image quality” and “perceptual color quality” were directly correlated. In other words, in subjective assessments, overall image quality increases if color quality increases. This result also begins to substantiate the claim of Trémeau and Charrier [11] that “the color appearances of images are related to perceptual image quality”. This study primarily emphasized “examination of human’s regions of interest in complex images.” Whether or not regions or objects with relatively high frequency domain attracted greater attention remains to be tested by in future research.

4.2 Results of Correlation Analysis for “Perceptual Image Quality” and “Perceptual Color Quality” on Eye Movement

Generally, when subjects assessed the “perceptual image quality” and “perceptual color quality” of static object and landscape images, they tended towards broader browsing; for human images, visual attention was typically focused on faces. These results are consistent with the research of Nguyen et al. [10]. However, this study found that subjects in general did not approve of the “perceptual image quality” or “perceptual color quality” of human images. We posit that because experimental tasks did not require subjects to assess the skin tones of human images, the majority of subjects focused their visual attention on facial regions when they assessed images. Consequently, the facial features of the subjects photographed in the images may have interfered with the assessment of image quality of subjects. In addition, the time required to assess perceptual image quality was significantly higher than the time needed to assess perceptual color quality. This result suggests that the concept of perceptual image quality is broader and vaguer in the cognitive processes of subjects compared to the concept of perceptual color quality. At the same time, in addition to subjective assessment results showing that high perceptual color quality were consistent with high overall image quality, eye movement information for perceptual color quality assessments were also directly correlated with eye movements in overall image assessments. Furthermore, the tasks for evaluating “perceptual color quality” were significantly clearer and more specific than tasks for evaluating “overall image quality.” At the same time, when individuals assessed “perceptual color quality,” their eye movement information was also more stable and consistent than in “perceptual image quality” assessments. As a result, we suggest that future studies regarding “perceived image quality” also consider “perceptual color quality" assessment criteria in order to reduce cognitive differences between subjects regarding "perceptual image quality" which affect the accuracy of assessment results.

Acknowledgments. The authors would like to thank the participants and TTLA (Taiwan TFT LCD Association) for supporting this research and providing insightful comments.

References

1. Newell, A.: Unified Theories of Cognition. Harvard University Press, Cambridge (1990)
2. Maeder, A.J., Eckert, M.: Medical image compression: quality and performance issues. In: Pham, B., Braun, M., Maeder, A.J, Eckert, M.P (eds.) New Approaches in Medical Image Analysis, Proceedings of SPIE, vol. 3747, pp. 93–101 (1999)
3. Sheikh, H.R., Bovik, A.C.: Image information and visual quality. IEEE Trans. Image Process. **15**(2), 430–444 (2006)
4. Ginesu, G., Massidda, F., Giusto, D.D.: A multi-factors approach for image quality assessment based on a human visual system model. Sig. Process. Image Commun. **21**, 316–333 (2006)
5. Fedorovskaya, E.A., Ridder, H., Blommaert, F.J.: Chroma variants and perceived quality of color images of natural scenes. Color Res. Appl. **22**(2), 96–110 (1996)
6. Privitera, C.M., Fujita, T., Chernyak, D., Stark, L.W.: On the discriminability of hROIs, human visually selected regions of interest. Biol. Cybern. **93**(2), 141–152 (2005)
7. Santella, A., DeCarlo, D.: Robust clustering of eye movement recordings for quantification of visual interest, In: The proceeding of Eye Tracking Research and Applications (ETRA) Symposium (2004)
8. Calabria, A.J., Fairchild, M.D.: Perceived image contrast and observer preference I: the effects of lightness, chroma, and sharpness, manipulations on contrast perception, Munsell Color Science Laboratory, Rochester Institute of Technology (2003)
9. Sheedy, J.E., Smith, R., Hayes, J.: Visual effects of the luminance surrounding a computer display. Ergonomics **48**(9), 1114–1128 (2005)
10. Nguyen, A., Chandran, V., Sridharan, S.: Gaze tracking for region of interest coding in JPEG 2000. Sig. Process. Image Commun. **21**, 356–377 (2006)
11. Trémeau, A., Charrier, C.: Influence of chromatic changes on the perception of color image quality. Color Res. Appl. **25**(3), 200–213 (1999)