Naturalness and Image Quality: Chroma and Hue Variation in Color Images of Natural Scenes

Huib de Ridder and Frans J.J. Blommaert
Institute for Perception Research, Eindhoven, The Netherlands;
Elena A. Fedorovskaya, Department of Psychophysiology, Moscow State University, Mokhovaya street 8, 103009 Moscow, Russia

Abstract

The relation between perceptual image quality and naturalness was investigated by varying the colorfulness and hue of color images of natural scenes. These variations were created by digitizing the images, subsequently determining their color point distributions in the CIELUV color space and finally multiplying either the chroma value or the hue-angle of each pixel by a constant. During the chroma/hue-angle transformation the lightness and hue-angle/chroma value of each pixel were kept constant. Ten subjects rated quality and naturalness on numerical scales. The results show that both quality and naturalness deteriorate as soon as hues start to deviate from the ones in the original image. Chroma variation affected the impression of quality and naturalness to a lesser extent than did hue variation. In general, a linear relation was found between image quality and naturalness. For chroma variation, however, a small but systematic deviation could be observed. This deviation reflects the subjects' preference for more colorful but, at the same time, somewhat unnatural images.

Keywords: natural scenes, perceptual image quality, naturalness, CIELUV, chroma, colorfulness, hue

1. Introduction

The success of any product in the field of electronic imaging depends strongly on how the end-users judge the quality of the final output: the images displayed on the screen. Accordingly, the optimization of imaging systems, processing algorithms etc. will never be accomplished in a satisfactory way unless the psychological and perceptual factors involved in judging images are taken into account. Stated otherwise, finding optimal parameter settings can never be an entirely physical problem for the reason that image quality, loosely defined as the 'degree of excellence of an image',[1] is not directly related to the physical image itself but to its internal or mental representation. This internal representation is sometimes referred to as the sensorial image.[2] Image quality is related to the sensorial image through the perceived strength of global image attributes such as sharpness, brightness, colorfulness, noisiness and visibility of structure. These attributes, in turn, are determined by many physical display parameters. For some of these parameters (e.g. blur, periodic structure, noise) it proved possible to derive explicit expressions for their relation with the corresponding image attributes.[3,4]

Scaling experiments using multiply impaired images have shown that image attributes can be represented by a set of orthogonal vectors in a Euclidean space.[3,5-8] Accordingly, the attributes can be said to be the orthogonal dimensions of a multidimensional psychological space underlying image quality. The sensorial image is represented in this space by a point with the perceived strengths of the attributes as coordinates. Image quality has sometimes been identified as a direction in this space, with the angle to a dimension indicating how relevant that attribute is for the quality judgment.[9,10] In other cases, decrease of image quality has been associated with the distance from the origin (the origin represents the unimpaired image).[2,3,8] Here, the relative importance of the various attributes was expressed by weighting the dimensions differently.

The experimental finding that image quality can be linked to a distance in the psychological space makes it clear that the perceptual quality of an image is judged by comparison with some internal reference. This reference may be thought of as an ideal version of the perceived image, i.e. the image without artifacts such as noise or blur. In general, this ideal image has to be derived from the perceived image itself because observers do not have the original scene at their disposal. Moreover, observers have to judge (natural) scenes that they have probably never seen before. This derivation of the ideal image will undoubtedly benefit from the presence of familiar objects in the scene. Miyake and Haneishi,[11] for example, pointed out that for optimizing color reproduction ". . . the color of the sky, green grass and skin is most important since these colors are familiar in daily living" (p. 224). The underlying, intuitively justifiable assumption is that appreciation-oriented images of high quality should at least be perceived as 'natural', that is, they should be judged to possess a high degree of correspondence to (memorized) reality. But what exactly is the relation between naturalness and image quality? And, does this relation depend on how the sensorial image is transformed? In particular, does it matter whether a transformation is 'natural', in the sense that it often occurs in the real world, or not?

In this paper we report on an experiment in which the relation between perceptual image quality and naturalness

has been investigated by varying the colorfulness[12,13] and hue of color images of natural scenes. These variations were created by digitizing the images, subsequently calculating their color point distributions in the CIELUV color space[14] and finally multiplying either the chroma value or the hue-angle of each pixel by a constant. The changes in colorfulness resulting from this kind of chroma variation may be classified as natural. The hue changes caused by the rotation of the color point distribution in the CIELUV space, on the other hand, must be classified as unnatural.

2. Method

2.1 Subjects

Eight male and two female subjects participated in the experiment. Their age varied between 20 and 25 years. They had normal or corrected-to-normal vision. Their color vision was checked with the H-R-R Pseudoisochromatic Plates.[15] No significant color deficiencies were observed.

2.2 Stimuli

All image material was prepared using a Gould deAnza Image Processing System IPS8400. Pictures of four natural scenes were used: a portrait of a female model (WANDA01), an outdoor scene (TERRASGEEL), fruit displayed in front of a greengrocer's shop (FRUIT) and an abstract sculpture with bushes (STADHUIS). RGB signals obtained by scanning slides of these scenes were digitized with 8 bits/pixel on a grid of 512 by 512 pixels. Before describing these original images by their color point distributions in the CIELUV uniform color space, the size of each image was reduced to 456 by 450 pixels to avoid artifacts on the edges. Each color point corresponded to one pixel. Reference white was D_{65}.[14] For two scenes, Figure 1 shows the resulting color point distributions in the form of projections on the u*v* and u*L* planes. New images were generated by multiplying either the chroma value or the hue-angle of each pixel by a constant. During the chroma/hue-angle transformation the lightness and hue-angle/chroma value of each pixel were kept constant. For each scene this resulted in five images in which chroma decreased (multiplication factors ranging from 0.5 to 0.9), five images in which chroma increased (multiplication factors ranging from 1.2 to 2.0), five images in which the hues were rotated clockwise around the L* (lightness) axis, and five images in which the hues were rotated counter-clockwise. In both

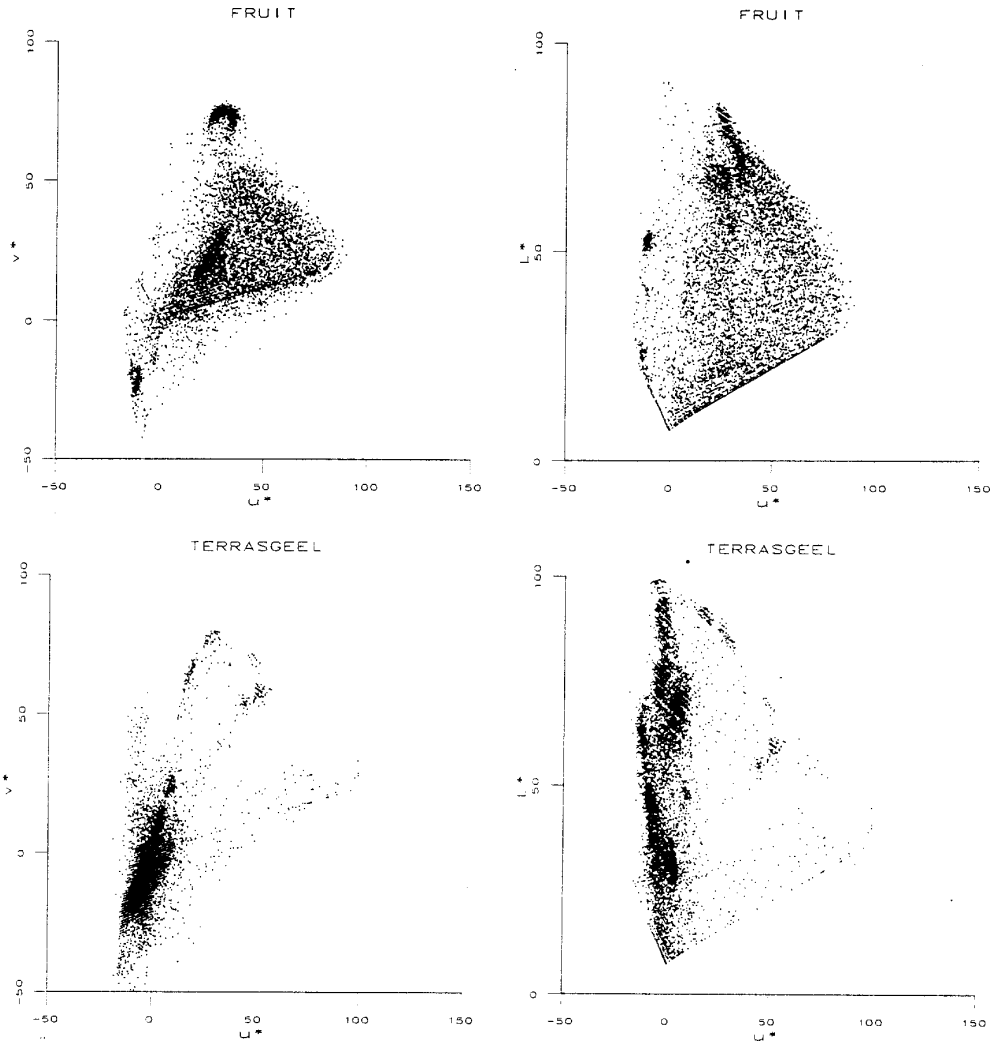

*Figure 1. Color point distribution in CIELUV uniform color space for two scenes: FRUIT and TERRASGEEL. Left-hand panels: projections on u*v* plane. Right-hand panels: projections on u*L* plane. For clarity of representation, only each 20th point of the image has been plotted.*

directions the hue-angle changed in regular steps to a maximum of one radian. If, during the processing of the images, calculated values were out of the possible range for the grey values of the monitor, the nearest possible value of chroma was used (clipping). Figure 2 demonstrates the effect of increasing and decreasing chroma on the color point distribution. During the experiment, the original images were also included, creating an experimental set of 84 images.

2.3 Procedure

Images were displayed on a 70 Hz Barco CCID7351B monitor placed in a dark room in front of a dimly lit 'white' background. The monitor was corrected such that the screen luminance was linearly related to the optical density of the original slides. The images were presented for five seconds after which a 9 cd/m^2 adaptation field appeared on the screen. Viewing conditions were in accordance with CCIR Recommendation 500.[16] The subjects viewed the monitor at a distance of about 1.7 m. At this distance, the pixel size is about 1 min of arc. The experiment consisted of two sessions. In the course of a session all 84 images were displayed four times in a random sequence, except that the same scene never appeared on two consecutive trials. In one session, subjects rated either the perceptual quality or the naturalness of all images on a ten-point numerical category scale ranging from one (lowest quality/naturalness) to ten (highest quality/naturalness). The instructions given to the subjects defined perceptual image quality as 'degree of excellence of the image' and naturalness as 'degree of correspondence between the reproduced image and reality (that is, the original scene as it is according to the viewer)'. Five subjects rated quality in the first session and naturalness in the second session. The reverse holds for the other five subjects. Before starting a session, subjects had to judge a training series of 30 images.

3. Results And Discussion

3.1 Naturalness and Image Quality: General

For the quality as well as naturalness judgments no systematic differences were found between the subjects. Apparently, the order of judging quality and naturalness did not affect the quality and naturalness ratings. We therefore decided to average over the subjects. Figure 3 shows the resulting averaged naturalness ratings as a function of the

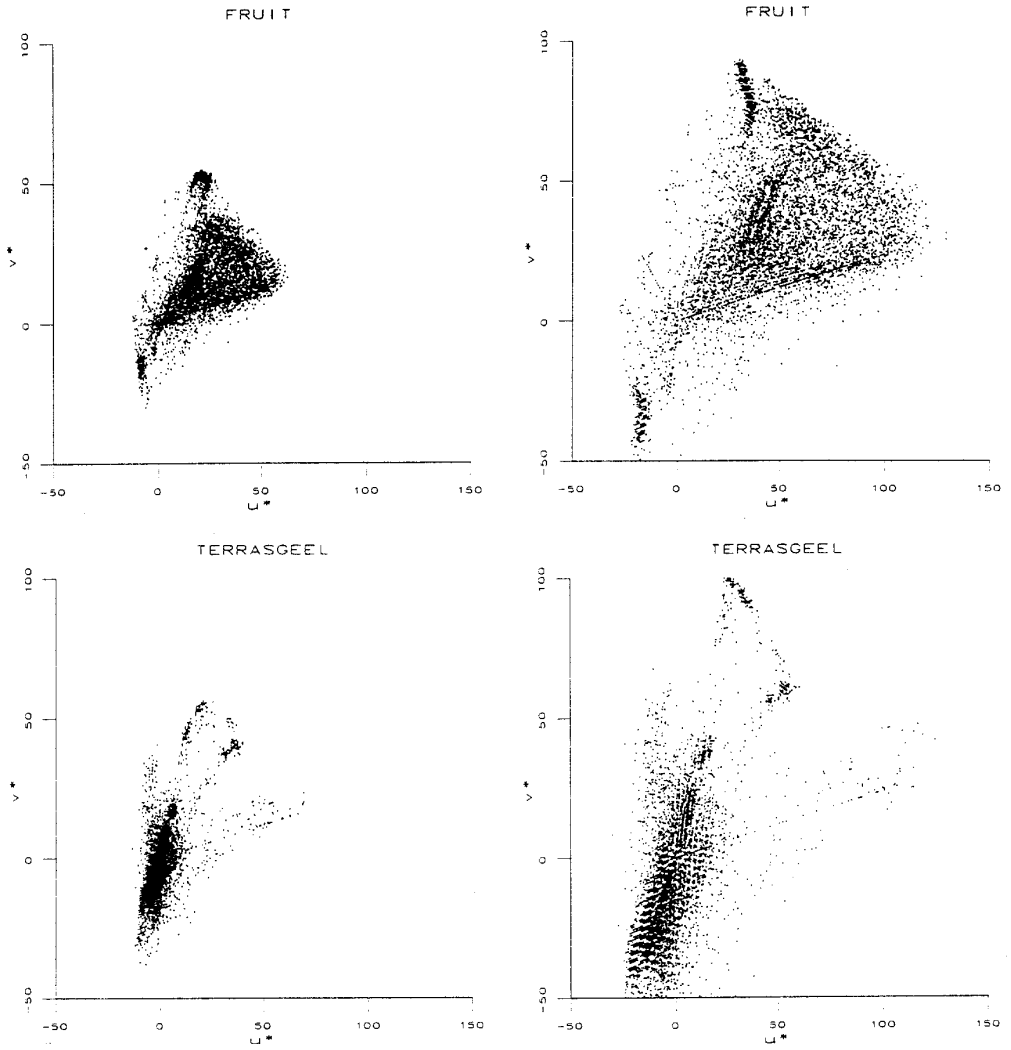

*Figure 2. Projections on u*v* plane of color point distributions in CIELUV space for two scenes: FRUIT and TERRASGEEL. Chroma values have been multiplied by 0.7 (left-hand panels) and 1.6 (right-hand panels). For clarity of representation, only each 20th point of the image has been plotted.*

averaged quality ratings for all 84 images. The first conclusion that can be drawn from these data is that, in general, a strong linear relation exists between scaled quality and naturalness ($r^2 = 0.96$). This agrees with data from Laihanen et al.[17] showing that the impression of increased naturalness of reproduced skin color correlates positively with a quality improvement. Another interesting feature of the data in Figure 3 is that the naturalness and quality scales are almost identical.

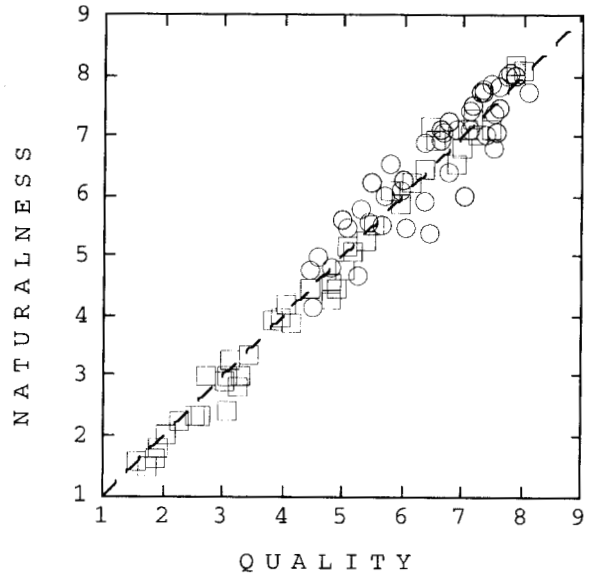

Figure 3. Naturalness estimates as a function of quality judgments for all 84 images. Squares: hue variation. Circles: chroma variation. The average standard error of the mean is 0.19 for the naturalness ratings and 0.20 for the quality ratings.

The second conclusion that can be drawn from Figure 3 is that hue variation influences both quality and naturalness judgments to a much larger extent than does chroma variation. This finding cannot be attributed to differences in the magnitude of the two kinds of color manipulation, since the range of displacement in the u*v* plane is about the same for the hue and chroma variation. The observed difference between hue and chroma variation agrees nicely with the experimental finding that the chromaticity distribution of preferred skin colors forms an ellipse with the principal axis in the chroma direction.[17,18] It also harmonizes with the common practice in gamut mapping of preserving the hues and manipulating chroma and lightness only.[19]

3.2 Naturalness and Image Quality: Hue Variation

In Figure 4 the quality and naturalness ratings are presented for the condition of hue variation only. This figure clearly shows how both quality and naturalness deteriorate as soon as hues start to deviate from the ones in the original image. The rather small effects observed for the STADHUIS scene are due to the fact that this image contains predominantly colors of low chroma. The close relation between naturalness and quality, together with the finding that the qualitatively optimal image is the original, suggests that unnatural image manipulations like rotating hues around the L* axis do not have to be considered as a means for improving image quality.

3.3 Naturalness and Image Quality: Chroma Variation

Figure 5 presents the naturalness ratings as a function of scaled image quality for chroma variation only. A small but systematic deviation from linearity can be seen for all scenes but is most pronounced for STADHUIS and WANDA01. Comparable results were obtained in a previous study employing chroma variation only.[12,13] In that study, the observed deviation was interpreted as follows. With increasing chroma, both naturalness and quality increase linearly as a function of colorfulness. Above a certain chroma value, however, naturalness starts to drop while quality remains relatively high.

An alternative interpretation is presented in Figure 6, showing that the inverted-U-shaped function for the quality judgments has been 'shifted' to higher average chroma values relative to the function for the naturalness judgments. The maxima of second-order polynomials, fitted to the quality as well as naturalness ratings ($0.96 \leq r^2 \leq 0.99$), confirmed this trend. The average chroma values at which these maxima were reached lie at 14.9 and 17.1 (STADHUIS), 14.6 and 15.7 (TERRASGEEL), 36.3 and 37.6 (FRUIT) and 46.7 and 51.5 (WANDA01) for naturalness and quality, respectively. Apparently, the subjects used different criteria to assess naturalness and image quality; the subjective preference in quality was biased towards more colorful images, although the subjects realized that these images looked somewhat unnatural. The finding that the qualitatively optimal images are slightly more colorful than the original ones highlights the importance of natural image manipulations such as chroma variation as a means for improving image quality. Finally, Figure 7 demonstrates how the quality and naturalness judgments can be fitted by the same second-order polynomial after the chroma values belonging to the naturalness function have been multiplied by a constant. This constant varied from 1.06 (FRUIT) to 1.15 (STADHUIS).

5. Conclusions

In the present study, the relation between perceptual image quality and naturalness was investigated by varying the colorfulness and hue of color images of natural scenes. The main conclusions are:

- The perceptual quality of appreciation-oriented images is strongly related to naturalness.
- Quality and naturalness deteriorate as soon as hues start to deviate from the ones in the original image.
- Chroma variation affects image quality and naturalness to a lesser extent than does hue variation.
- Subjects tend to prefer more colorful images, although they realize that these images look some what unnatural.
- Only image manipulations producing natural transformations of the sensorial image should be considered as a means for improving image quality.

6. Acknowledgments

The authors wish to thank Johan Vereijken for generating the images and Chris Meijers for carrying out the experiment.

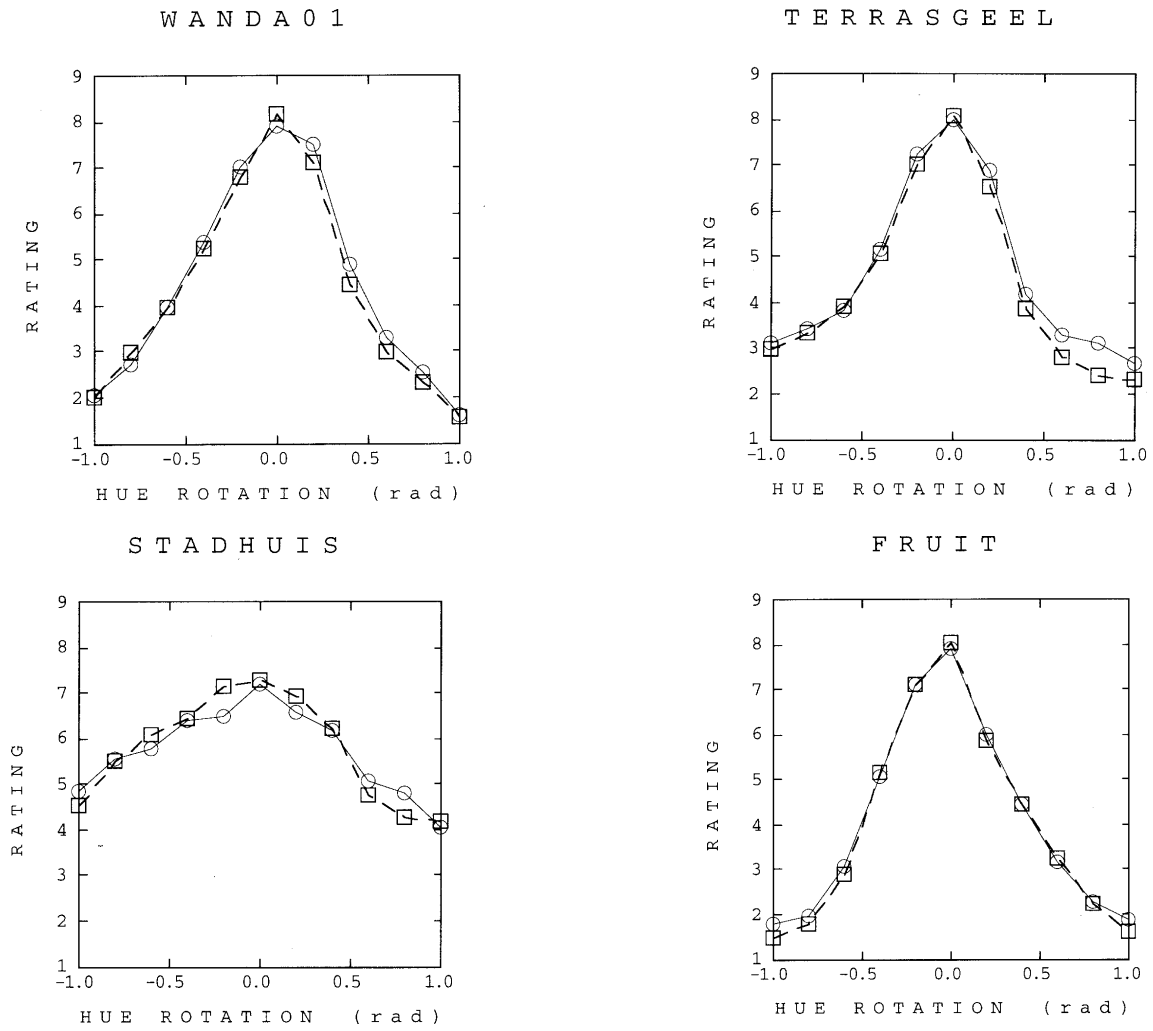

Figure 4. Naturalness (squares) and quality (circles) ratings as a function of change in hue-angle for four scenes. The hues were rotated clockwise (negative sign) as well as counter-clockwise (positive sign).

7. References

1. J. A. J. Roufs, "Perceptual image quality: Concept and measurement", *Philips Journal of Research*, vol. **47**, pp. 35-62, 1992.
2. M.R.M. Nijenhuis and F.J.J. Blommaert, "Perceptual error measure for sampled and interpolated imagery", *Eurodisplay '93, The 13th International Display Research Conference*, Strasbourg, France pp. 135-138, 1993.
3. M.R.M. Nijenhuis, Sampling and interpolation of static images: a perceptual views *PhD thesis*, Eindhoven University of Technology, Eindhoven, The Netherlands, 1993.
4. V. Kayargadde, Feature extraction for image quality prediction. *PhD thesis*, Eindhoven University of Technology, Eindhoven, The Netherlands, 1995.
5. H. de Ridder and G.M.M. Majoor, "Numerical category scaling: an efficient method for assessing digital image coding impairments", *Human Vision and Electronic Imaging: Models, Methods, and Applications*, B.E. Rogowitz and J.P. Allebach, Eds, *Proc. SPIE* **1249**, pp. 65-77, 1990.
6. H. de Ridder, "Subjective evaluation of scale-space image - coding", *Human Vision Visual Processing, and Digital Display II*, B.E. Rogowitz, M.H. Brill and J.P. Allebach, Eds, *Proc. SPIE* **1453**, pp. 31-42, 1991.
7. H. de Ridder, "Comparison of combination rules for digital-image-coding impairments", *IPO Annual Progress Report*, vo1. **26**, pp. 70-79, 1991.
8. H. de Ridder, "Minkowski-metrics as a combination rule for digital-image-coding impairments", *Human Vision, Visual Processing, and Digital Display III*, B.E. Rogowitz, J.P. Allebach and S.A. Klein, Eds, *Proc. SPIE* **1666**, pp. 16-26, 1992.
9. B. Escalante Ramirez, H. de Ridder and J.B.O.S. Martens, "Perceptual quality of noise-reduced computed tomography images: a multidimensional approach", *IPO Annual Progress Report*, vo1. **27**, pp. 63-71, 1992.
10. B. Escalante Ramirez, H. de Ridder and J.B.O.S. Martens, "Multidimensional perceptual quality of noise-reduced Computed-Tomography images", *Medical Imaging 1194: Image Perception*, H.L. Kundel, Ed, *Proc. SPIE* **2166**, pp. 119-130, 1994.
11. Y. Miyake and H. Haneishi, "The evolution of nonimpact print—impact on image quality", *Proceedings of the First IS&T/SID Color Imaging Conference: Transforms & Transportability of Color*, Scottsdale, Arizona, USA, pp. 222-227, 1993.

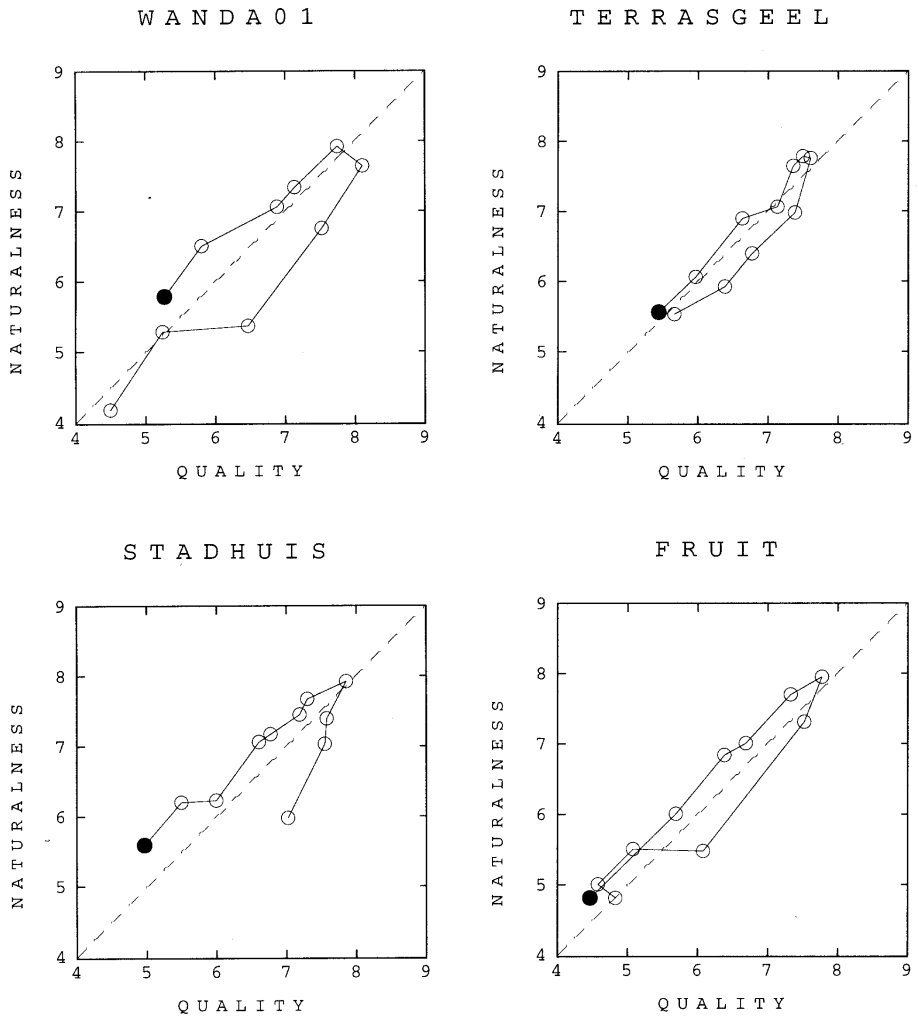

Figure 5. Naturalness estimates as a function of quality judgments for the condition of chroma variation. The filled symbols denote the images with the lowest average chroma.

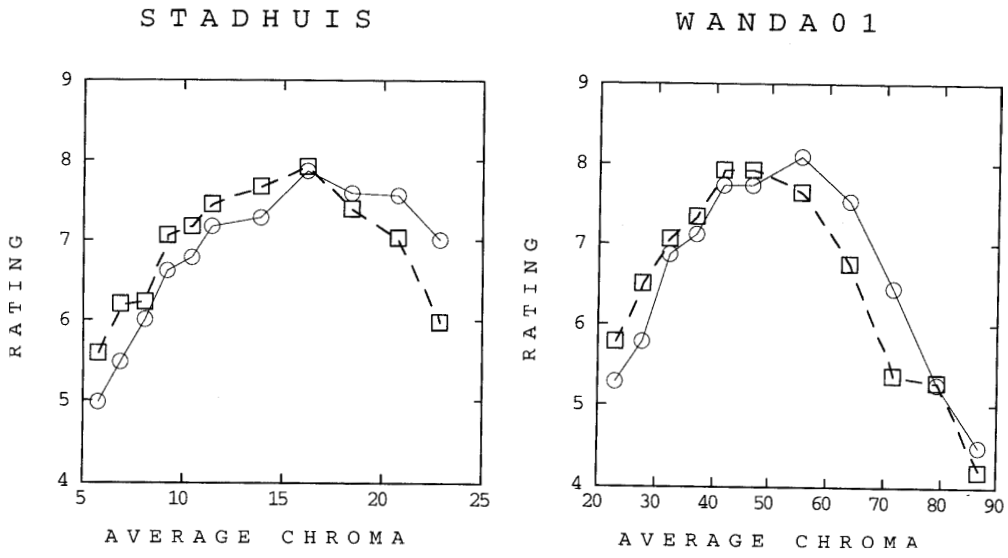

Figure 6. Naturalness (squares) and quality (circles) ratings as a function of average chroma for two scenes: STADHUIS and WANDA01.

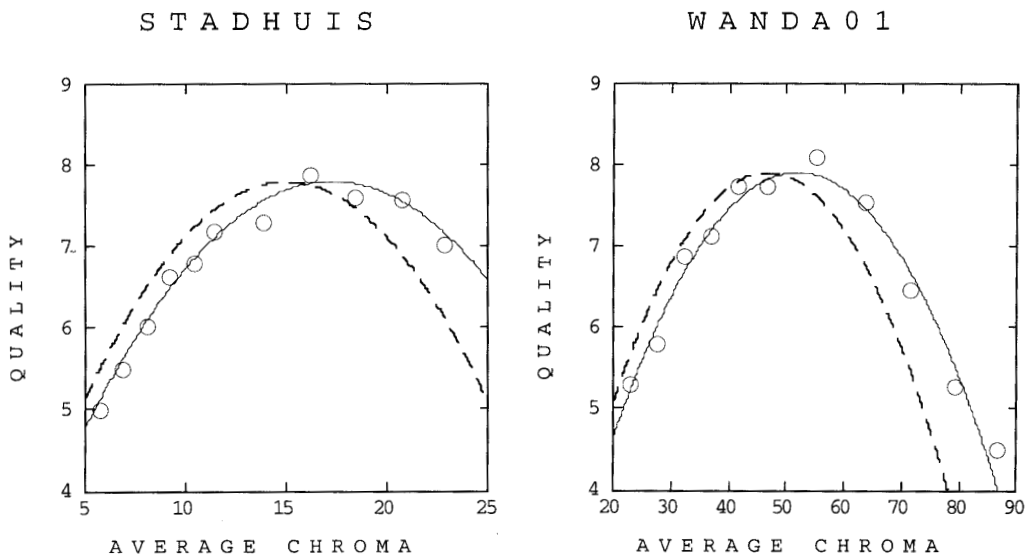

Figure 7. Quality judgments as a function of average chroma for two scenes: STADHUIS and WANDA01. Dashed line: second-order polynomial fitted to the naturalness ratings of Figure 6. Solid line: same polynomial after the chroma values belonging to the naturalness function have been multiplied by 1.15 (STADHUIS) and 1.11 (WANDA01).

12. E. A. Fedorovskaya, F.J.J. Blommaert and H. de Ridder, "Perceptual quality of color images of natural scenes transformed in CIELUV color space", *Proceedings of the First IS&T/SID Color Imaging Conference: Transforms & Transportability of Color*, Scottsdale, Arizona, USA, pp. 37-41, 1993.(See page 42, this publication)
13. H. de Ridder, E.A. Fedorovskaya and F.J.J. Blommaert, "Naturalness and image quality: chroma variation in colour images of natural scenes", *IPO Annual Progress Report,* vol. **28**, pp. 89-95, 1993.
14. G. Wyszecki and W.S. Stiles*, Color Science: Concepts and Methods, Quantitative Data and Formulae*. John Wiley and Sons, New York, 1982.
15. L.H. Hardy, G. Rand and M.C. Ritter, H-R-R *Pseudoisochromatic Plates for Detecting, Classifying, and Estimating the Degree of Defective Color Vision*. American Optical Company, New York, 1957.
16. CCIR, "Method for the subjective assessment of the quality of television pictures. Recommendation 500-3", In: *Recommendations and Reports of the CCIR*, International Telecommunication Union, Geneva, 1986.
17. P. Laihanen, M. Lindholm, S. Rouhiainen, H. Saarelma and L. Tuuteri, "Automatic colour correction", *Proceedings of the Second IS&T/SID Color Imaging Conference: Color Science, Systems and Applications*, Scottsdale, Arizona, USA, pp. 97-102, 1994.
18. D. Sanger, T. Asada, H. Haneishi and Y. Miyake, "Facial pattern detection and its preferred color reproduction", *Proceedings of the Second IS&T/SID Color Imaging Conference. Color Science, Systems and Applications*, Scottsdale, Arizona, USA, pp. 149-153, 1994.
19. M. Wolski, J.P. Allebach and C.A. Bouman, "Gamut mapping: Squeezing the most out of your color system", *Proceedings of the Second IS&T/SID Color Imaging Conference: Color Science, Systems and Applications*, Scottsdale, Arizona, USA, pp. 89-92, 1994.

published previously in SPIE, Vol. 2411, page 51