

# Chroma variations and perceived quality of color images of natural scenes

**Citation for published version (APA):**

Fedorovskaya, E. A., Ridder, de, H., & Blommaert, F. J. J. (1997). Chroma variations and perceived quality of color images of natural scenes. *Color Research and Application*, 22(2), 96-110.  
[https://doi.org/10.1002/\(SICI\)1520-6378\(199704\)22:2<96::AID-COL5>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1520-6378(199704)22:2<96::AID-COL5>3.0.CO;2-Z)

**DOI:**

[10.1002/\(SICI\)1520-6378\(199704\)22:2<96::AID-COL5>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1520-6378(199704)22:2<96::AID-COL5>3.0.CO;2-Z)

**Document status and date:**

Published: 01/01/1997

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# Chroma Variations and Perceived Quality of Color Images of Natural Scenes

Elena A. Fedorovskaya,<sup>1</sup> Huib de Ridder,<sup>2\*</sup>  
Frans J. J. Blommaert<sup>2</sup>

<sup>1</sup>Department of Psychophysiology, Moscow State University, Mokhovaya Street 8, 103009 Moscow, Russia

<sup>2</sup>Institute for Perception Research (IPO), P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Received 6 October 1995; accepted 7 July 1996

*Abstract: Transformations of natural images in the perceptually uniform CIELUV color space have been investigated with respect to perceptual image quality. To this end, digitized color images of four natural scenes were described on the basis of their color point distributions in the CIELUV color space. A new set of images was created by varying the chroma value of each pixel while the lightness and hue angle were kept constant. The chroma was changed in two different ways: (1) through the addition or subtraction of the same amount of chroma to or from the chroma value of each pixel; (2) through multiplication of the chroma value of each pixel by a constant. In three experiments, subjects judged the perceptual quality, colorfulness, and naturalness of the images on a ten-point numerical category scale. The results indicate that colorfulness is the main perceptual attribute underlying image quality when chroma varies. Colorfulness itself was found to depend on both the average chroma and its variability. In general, the subjects preferred slightly more colorful images to the original ones. The perceptual quality of the images was found to be closely related to the naturalness of the images. © 1997 John Wiley & Sons, Inc. Col Res Appl, 22, 96–110, 1997*

*Key words: natural images; perceptual image quality; colorfulness; chroma; naturalness; numerical category scaling*

## INTRODUCTION

Optimization of color reproduction is a complicated task: per image, some quarter of a million pixels can be manip-

ulated with respect to luminance and chromaticity in order to improve the display of color. Usually, some implicit assumptions are made to limit the almost infinite number of possible manipulations. One of these assumptions is that only global changes need to be considered. This stems from the notion that images are coherent in the sense that different parts of the same image share luminance and spectral properties due to the fact that they belong to the same object or are illuminated by the same light source. This assumption underlies global reproduction techniques in which the same color operation is applied to all pixels, irrespective of the content of the image.<sup>1,2</sup> Another assumption is that the optimum image equals the original (real-life) scene. Hence, all efforts should be directed towards finding ways of approaching this equality as closely as possible.

In this article, a different constraint on color reproduction is proposed, namely, the experienced naturalness of an image. This choice is motivated by the argument that constraints should be formulated not in the physical, but in the psychological domain, the reason being that image quality judgments are based on a comparison between the output of the visual process and some internal reference (memory representation).<sup>3</sup> The basic assumption underlying the naturalness constraint is that, in an appreciation-oriented setting,<sup>4</sup> images of high quality should at least be perceived as “natural.” That is, such an image conforms as much as possible to the ideas and expectations the observers have about the original scene at the time the picture was taken. Accordingly, the observers have to rely on earlier experiences with comparable situations, or, more generally, on their internalized knowledge of the

\* Correspondence to: Huib de Ridder  
© 1997 John Wiley & Sons, Inc.



FIG. 1. Pictures of the original, unprocessed, images used in the experiments: (a) Wanda, (b) Fruit, (c) Terrasgeel, and (d) Stadhuis.

world.<sup>3,5</sup> From this it follows that, for natural images, the impression of naturalness reflects the degree of correspondence to memorized reality. In this way, naturalness can be considered a global perceptual attribute. It should be noted that images that are judged to be natural do not necessarily conform to the physical constraints mentioned above. The perceptually optimum image, for instance, does not necessarily have to equal the optimum physical reproduction of an image. In fact, as will be shown in this article, it does not.

The investigation we are reporting here aimed at exploring the usefulness of the “naturalness” constraint. To this end, we studied the relationship between the perceptual quality and the naturalness of complex images by manipulating the global attribute “colorfulness” by means of chroma variations in the CIELUV color space.<sup>6</sup> To be able to perform the transformations, digitized color images of four natural scenes were described on the basis of their color point distributions in the CIELUV color space. A new set of images was

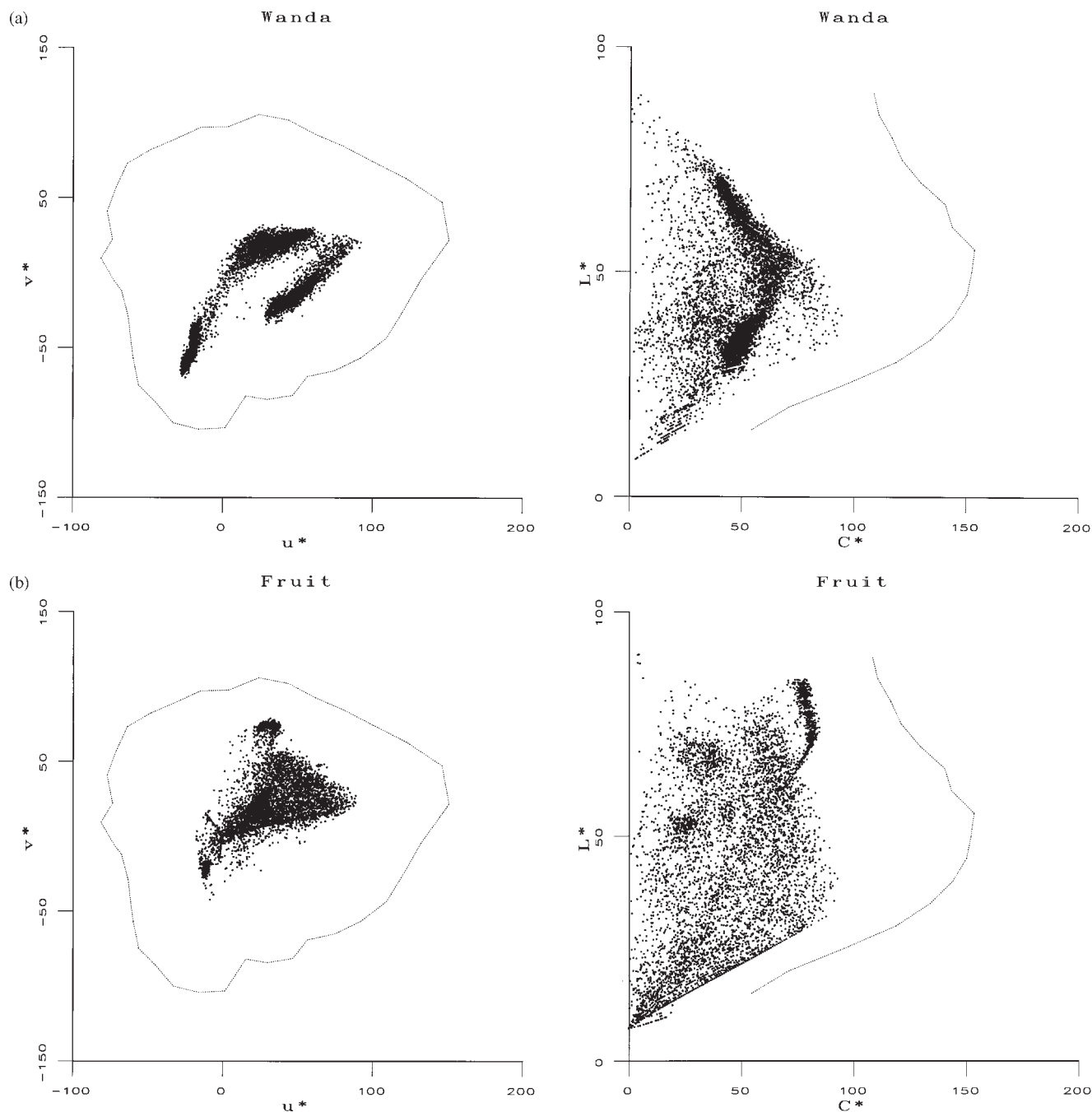


FIG. 2. Left-hand panels: the color point distribution in the CIELUV color space for the images shown in Fig. 1 in the form of the projection on the  $u^*v^*$ -plane. Right-hand panels: lightness values  $L^*$  as a function of chroma  $C^*$ . Every 20th point has been plotted for clarity of representation. Dotted lines denote the maximum gamut for real surface colors measured by Pointer,<sup>14</sup> after a correction for the appropriate white point (D65).

created by varying the chroma value of each pixel while the lightness and hue angle were kept constant. Consequently, the relation between colorfulness and chroma could be established.

In two experiments, subjects were asked to judge the perceptual quality and naturalness of the images on a ten-point

numerical category scale. A high degree of correspondence between naturalness and perceptual quality was observed.

In a third experiment, the subjects were asked to scale the colorfulness of the same images. The rationale of this experiment was to determine whether colorfulness was the perceptual attribute underlying per-

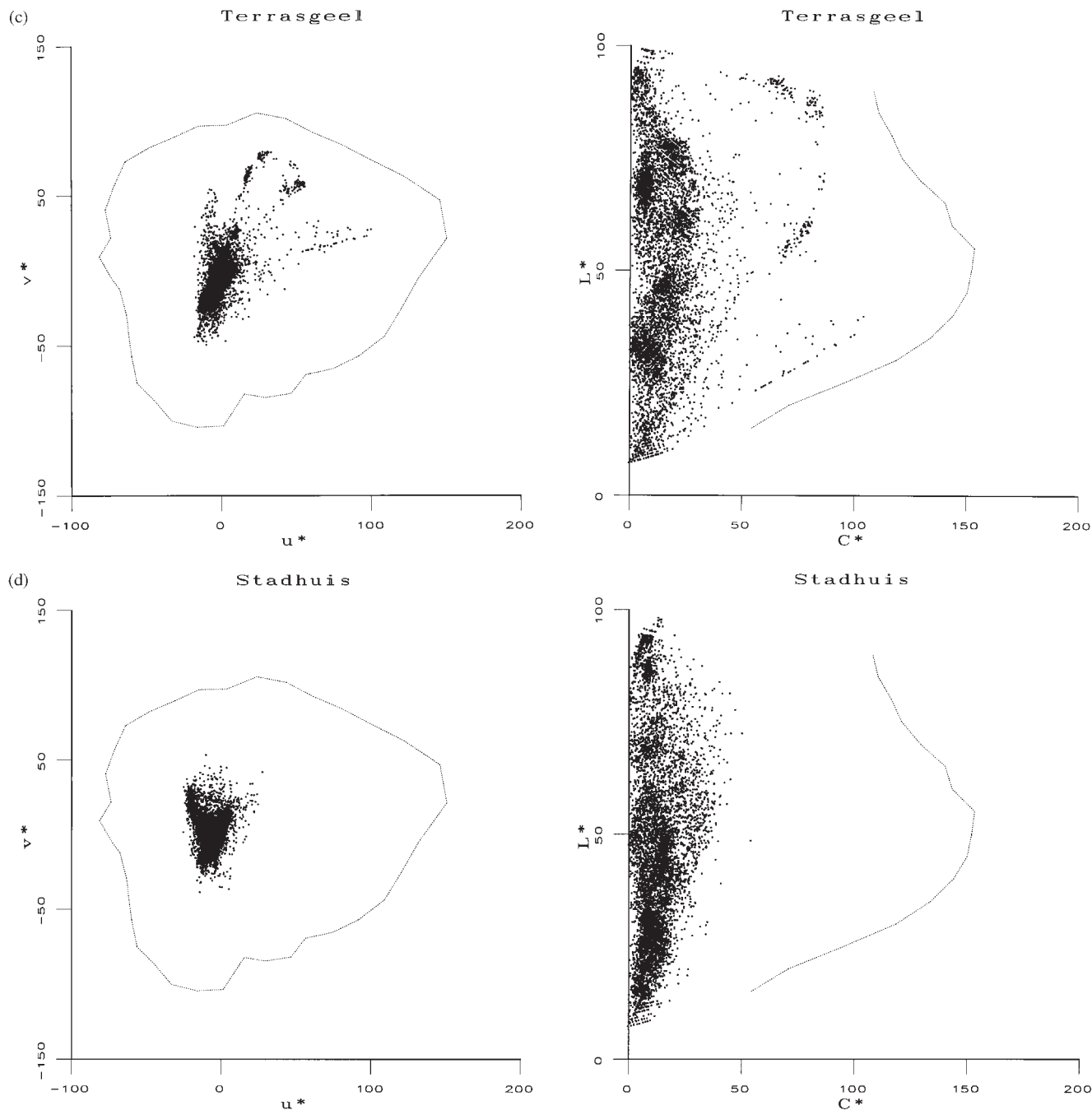


FIG. 2. (Continued)

ceptual quality in the case of these color variations. This indeed proved to be the case.

#### PREPARATION OF THE IMAGES

##### CIELUV Color Space

One way of describing complex color images is by representing them according to their color point distribution in an appropriate color space.<sup>7,8</sup> Nowadays, the CIE-

LUV color space is considered to be one of the most appropriate color spaces for application in television,<sup>9-11</sup> especially because it has an associated uniform chromaticity diagram. We used the CIELUV color space in the current investigation.

In CIELUV color space, every color is described by three values:

$$L^* = 116 \cdot (Y/Y_n)^{1/3} - 16 \quad \text{for } Y/Y_n > 0.008856$$

$$L^* = 903.3 \cdot (Y/Y_n) \quad \text{for } Y/Y_n \leq 0.008856$$

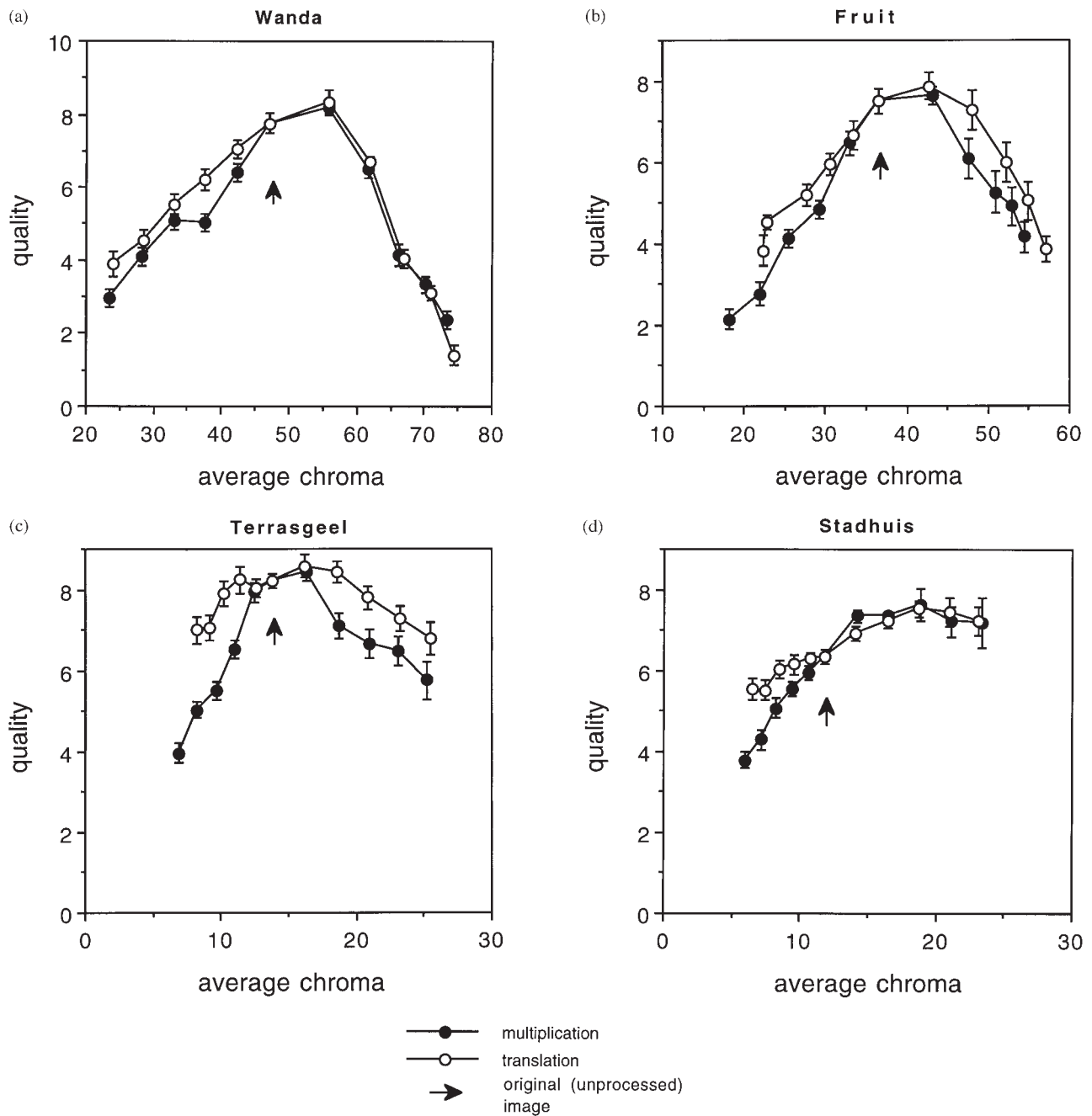


FIG. 3. Averaged scaled quality estimations plotted vs. the average chroma of the images for four scenes: (a) Wanda, (b) Fruit, (c) Terrasgeel, and (d) Stadhuis. Vertical bars indicate twice the standard error of the mean in this and all subsequent figures. Filled symbols: multiplication. Open symbols: translation. The arrow points to the quality judgment of the original image.

$$u^* = 13 * L^* * (u' - u'_n) \quad (1)$$

$$v^* = 13 * L^* * (v' - v'_n),$$

where  $L^*$  is CIELUV Psychometric Lightness and the quantities  $u'$ ,  $v'$ ,  $u'_n$ , and  $v'_n$  are color coordinates in the CIE 1976 uniform chromaticity scale diagram.  $u'_n$ ,  $v'_n$ , and  $Y_n$  are the values of  $u'$ ,  $v'$ , and  $Y$  for the appropriately chosen reference white. The following equations are

used to obtain  $u'$  and  $v'$  from the tristimulus values  $X$ ,  $Y$ , and  $Z$ :

$$u' = 4 * X / (X + 15 * Y + 3 * Z) \quad (2)$$

$$v' = 9 * Y / (X + 15 * Y + 3 * Z).$$

The measure that correlates with hue, CIELUV Psychometric Hue Angle,  $h_w$ :

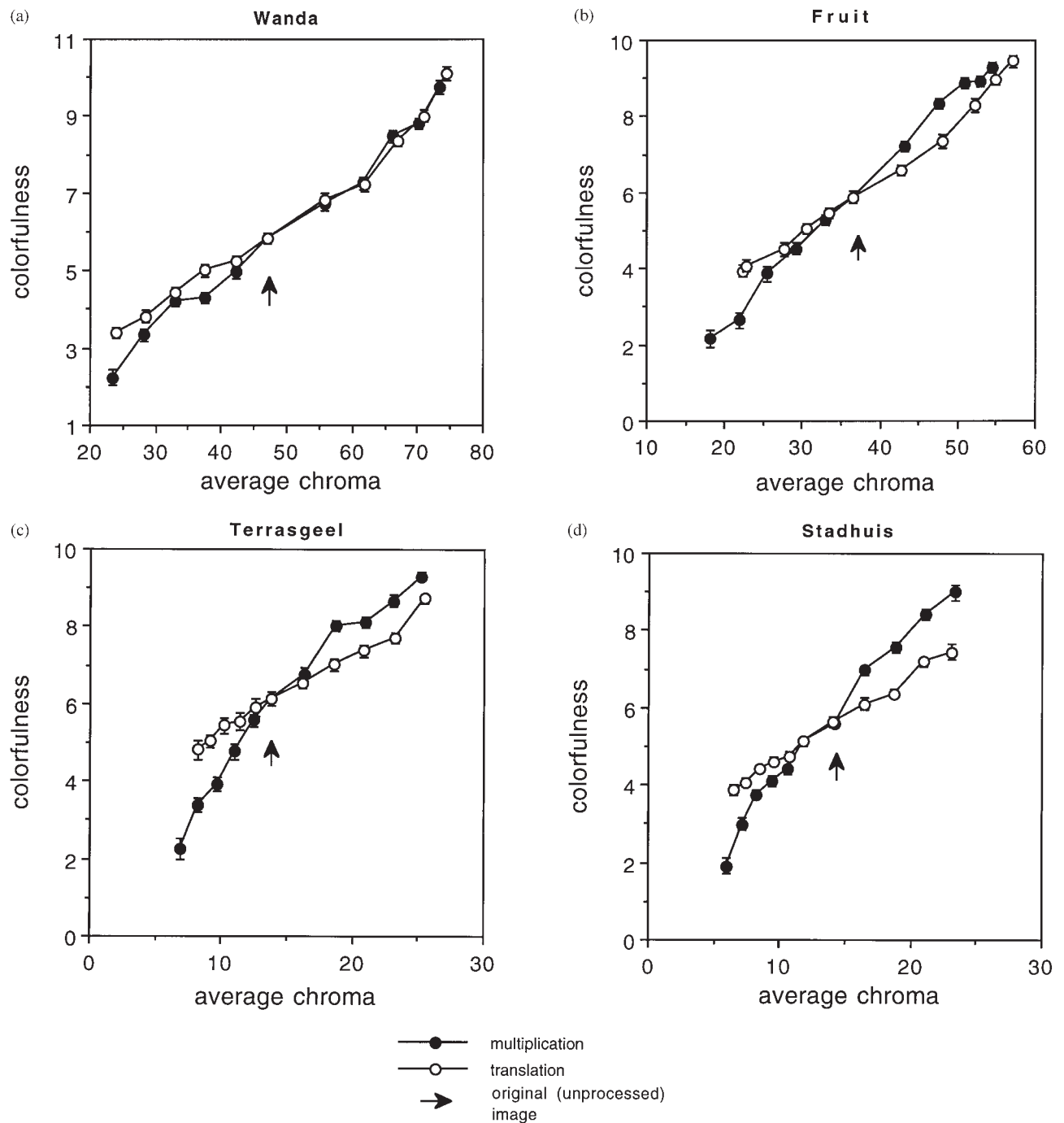


FIG. 4. Averaged scaled colorfulness estimations plotted as a function of the average chroma for four scenes: (a) Wanda, (b) Fruit, (c) Terrasgeel, and (d) Stadhuis. Filled symbols: multiplication. Open symbols: translation. The arrow points to the colorfulness judgment of the original image.

$$h_{uv} = \arctan[(v' - v'_n)/(u' - u'_n)], \quad (3)$$

can also be defined as follows:

$$h_{uv} = \arctan(v^*/u^*). \quad (4)$$

Every color has two perceptual attributes related to its

chromaticness or colorfulness:\* (1) saturation, defined as the colorfulness of an area judged in proportion to its brightness, and (2) chroma, the colorfulness of an area

\* Chromaticness (or colorfulness) is the attribute of a visual sensation according to which the perceived color of an area appears to be more or less chromatic.<sup>12</sup>

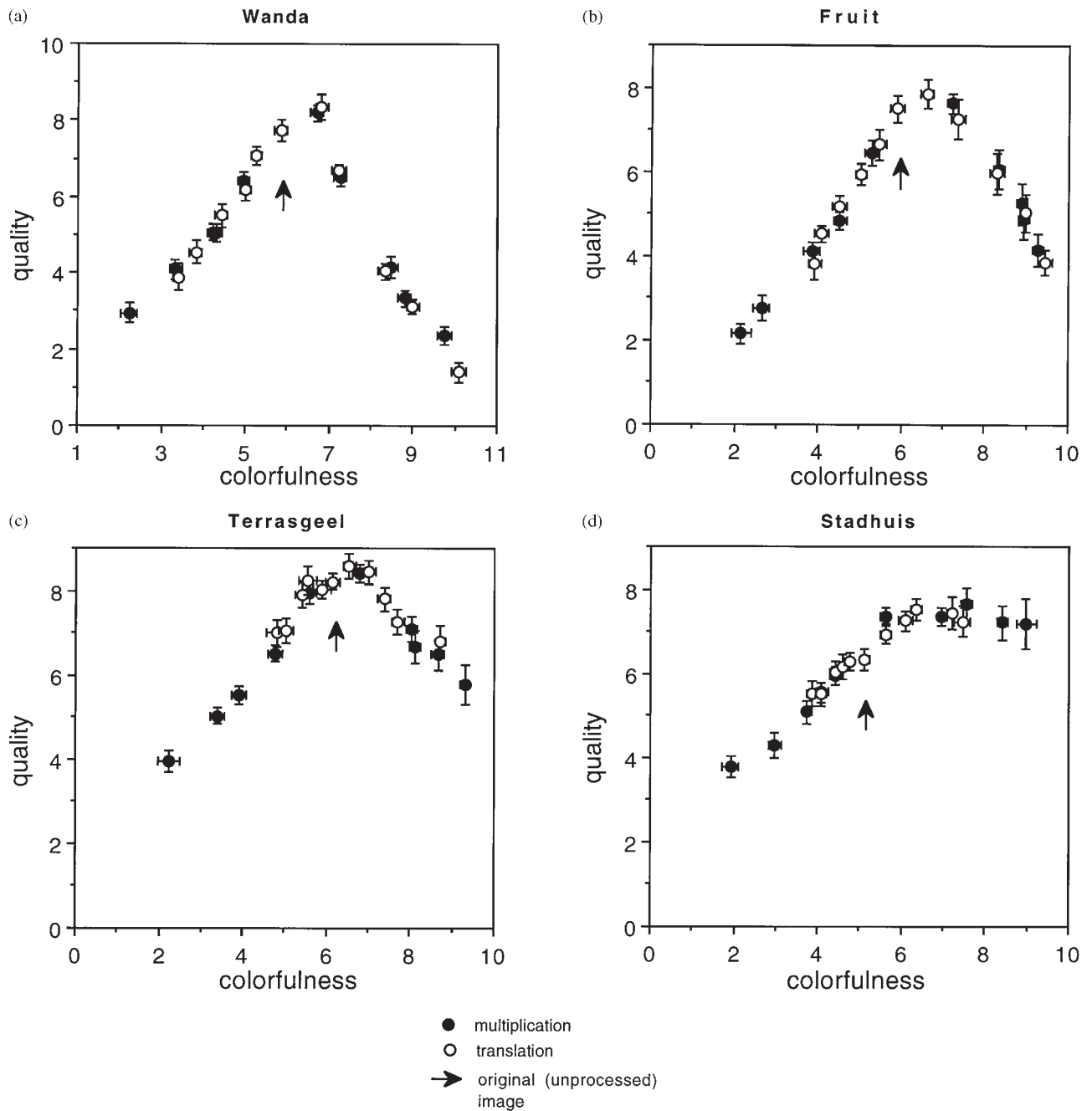


FIG. 5. Averaged scaled quality estimations plotted as a function of the averaged colorfulness for four scenes: (a) Wanda, (b) Fruit, (c) Terrasgeel, and (d) Stadhuis. Filled symbols: multiplication. Open symbols: translation. The arrow points to the judgments of the original image.

judged as a proportion of the brightness of a similarly illuminated white or highly transmitting area. In the CIELUV space two different measures that correlate with chromaticness are defined. They are: CIELUV Psychometric Saturation,  $s_{uv}$ ,

$$s_{uv} = 13 * [(u' - u'_n)^2 + (v' - v'_n)^2]^{1/2}, \quad (5)$$

and CIELUV Psychometric Chroma,  $C_{uv}^*$ ,

$$C_{uv}^* = (u^{*2} + v^{*2})^{1/2} = L^* * s_{uv}. \quad (6)$$

Although both characteristics could be relevant in the case of related colors,<sup>11</sup> chroma was chosen as a parameter for manipulation in our experiments for two reasons. First, several different color areas present in a natural scene are similarly illuminated and, therefore, they refer to the same reference white. Secondly, if the lightness and hue of two colors are the same, then the shift in chroma corresponds



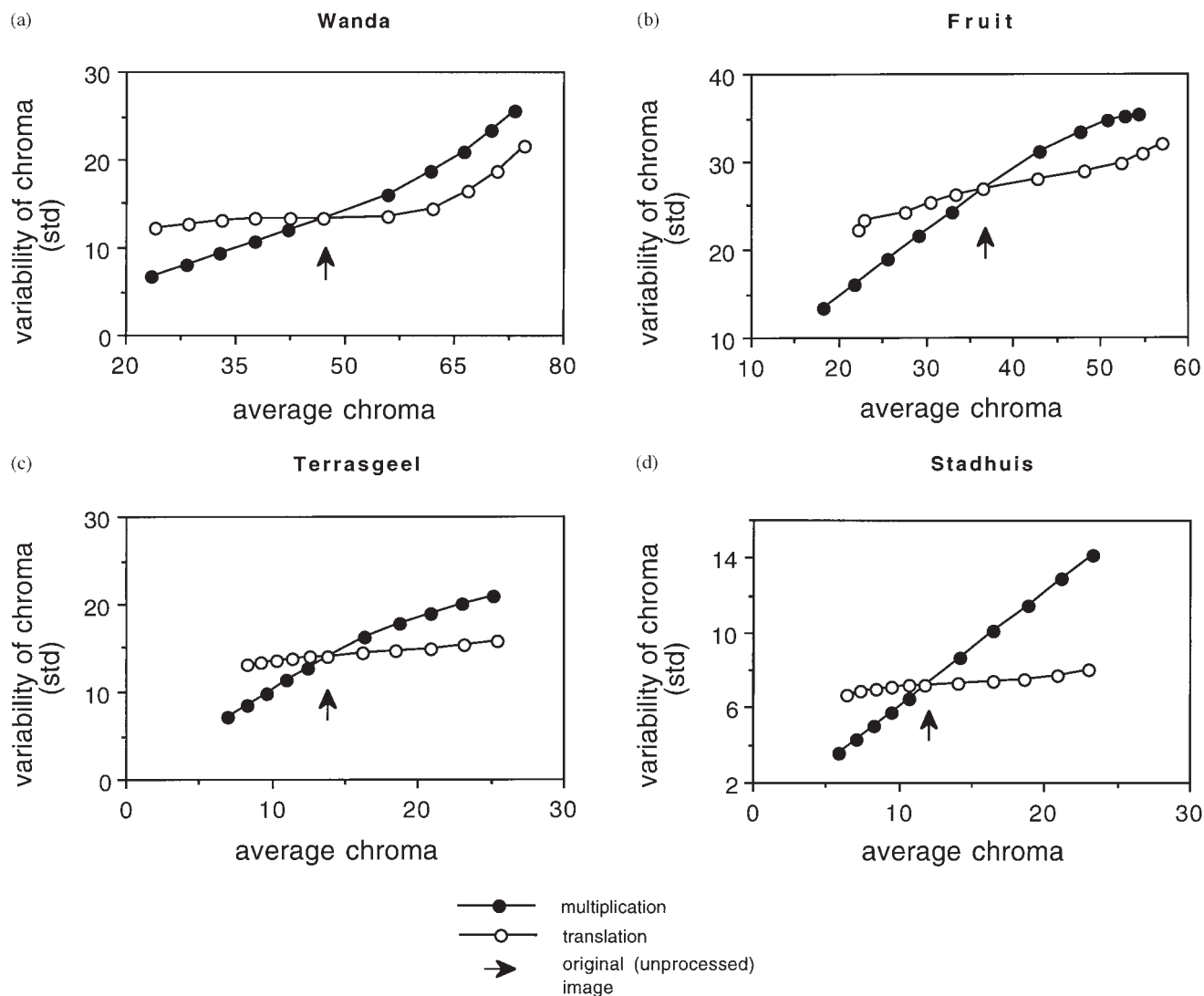


FIG. 6. Variability of chroma plotted vs. the average chroma of the images for four scenes: (a) Wanda, (b) Fruit, (c) Terrasgeel, and (d) Stadhuis. Filled symbols: multiplication. Open symbols: translation. The arrow points to the original image.

directly to the color difference that follows from the color difference formula for the CIELUV space, expressed in terms of differences in lightness ( $\Delta L^*$ ) and chroma ( $\Delta C^*$ ), and in terms of a measure that correlates with difference in hue,  $\Delta H^*$ :

$$\Delta E_{uv}^* = [(\Delta L^*)^2 + (\Delta H^*)^2 + (\Delta C^*)^2]^{1/2}, \quad (7)$$

where  $\Delta E_{uv}^*$  is the difference between two colors in the CIELUV space. In our experiments we assumed that the loci of constant hue were straight lines. In fact, it was experimentally shown that they are slightly curved in chromaticity diagrams.<sup>11,13</sup>

### Stimuli

All image material was prepared using a Gould deAnza

Image Processing System IPS8400. Slides of four natural scenes (Fig. 1) were scanned using a Leaf System Leaf-scan 35-SCSI after which the resulting  $r$ ,  $g$ , and  $b$  video signals were digitized with 8 bits/pixel on a grid of 512 by 512 pixels. In order to avoid the influence of artifacts at the edges, only a central region of 456 by 450 pixels was used in the experiments. The images were displayed on a 70-Hz interlaced Barco CCID7351B monitor. The monitor was corrected by means of a color Look-Up-Table such that the screen luminance was linearly related to the optical density of the original slides. The white point was set to D65 with the chromaticity coordinates  $(x_w, y_w) = (0.313, 0.329)$ . The chromaticity coordinates for the three phosphors of our monitor were:  $(x_r, y_r) = (0.625, 0.337)$ ,  $(x_g, y_g) = (0.307, 0.609)$ ,  $(x_b, y_b) = (0.144, 0.077)$ . The maximum luminance for white was  $60 \text{ cd/m}^2$ . All chromaticity and luminance mea-

TABLE I. Results of linear regression analysis for colorfulness estimations.

Wanda			Standard error of estimate: 0.293	
Adjusted squared multiple R: 0.984				
Variable	Coefficient	Std Error	T	P (2 Tail)
Constant	-0.520	0.208	-2.495	0.023
Chroma	0.105	0.008	13.279	0.000
Std chroma	0.100	0.028	3.586	0.002
Fruit			Standard error of estimate: 0.187	
Adjusted squared multiple R: 0.993				
Variable	Coefficient	Std error	T	P (2 Tail)
Constant	-1.816	0.211	-8.600	0.000
Chroma	0.121	0.008	15.298	0.000
Std chroma	0.126	0.017	7.607	0.000
Terrasgeel			Standard error of estimate: 0.283	
Adjusted squared multiple R: 0.977				
Variable	Coefficient	Std Error	T	P (2 Tail)
Constant	-0.205	0.286	-0.717	0.483
Chroma	0.162	0.019	8.728	0.000
Std chroma	0.276	0.032	8.522	0.000
Stadhuis			Standard error of estimate: 0.253	
Adjusted squared multiple: 0.981				
Variable	Coefficient	Std error	T	P (2 Tail)
Constant	0.423	0.178	2.372	0.029
Chroma	0.204	0.016	13.172	0.000
Std chroma	0.293	0.035	8.456	0.000

measurements were made using a Spectroradiometer SpectraScan PR-650 and an LMT L1003 Luminance meter, respectively.

In order to check how accurate the colors of the original scene were reproduced on the screen, we analyzed the chromaticities of the patches of a Kodak color control chart photographed at the original scene immediately after the original picture was taken. Comparing the chromaticity coordinates of the patches displayed on the screen with those of the original patches, we found that, for the saturated patches, the color points measured on the screen were lying at or close to the boundaries of the color gamut of the monitor. The differences between reproduced and real chromaticities were small enough to assume a faithful reproduction of the real colors.

To obtain the CIELUV color coordinates from  $r$ ,  $g$ , and  $b$  gray values for each pixel, the corresponding  $R$ ,  $G$ , and  $B$  values were calculated and then the following linear transformation was used to convert  $R$ ,  $G$ , and  $B$  into  $X$ ,  $Y$ , and  $Z$  tristimulus values:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.432 & 0.338 & 0.179 \\ 0.233 & 0.671 & 0.096 \\ 0.026 & 0.092 & 0.971 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (8)$$

where the  $3 \times 3$  matrix was calculated on the basis of the monitor measurements and calibration. Finally, with the aid of the set of Eqs. (1) and (2), the CIELUV color coordinates for each pixel were obtained.

The left-hand panels of Fig. 2 show the color point distributions obtained in the form of the projection on the  $u^*v^*$  plane for all four scenes. The right-hand panels present the distribution of  $L^*$  values as a function of  $C^*$ . For every original image, a set of new images was computed by changing the chroma value for each pixel while the lightness and hue angle were kept constant. The chroma was changed in two different ways: (1) the chroma value of each point was increased or decreased by adding/subtracting the same number of chroma units to or from the chroma value of each pixel (translation), and (2) the chroma value of each pixel was multiplied by a constant (multiplication). After that, the new  $r$ ,  $g$ , and  $b$  gray values were computed using the inverse of the matrix in Eq. (8). If, during the processing of the images, for some pixels the calculated values were outside the color gamut of the monitor, the maximum possible value of chroma was used for those pixels (clipping).

For each scene twenty images were prepared using ten different constants of translation and ten of multiplication. The multiplication constants we used were 0.5, 0.6, 0.7,

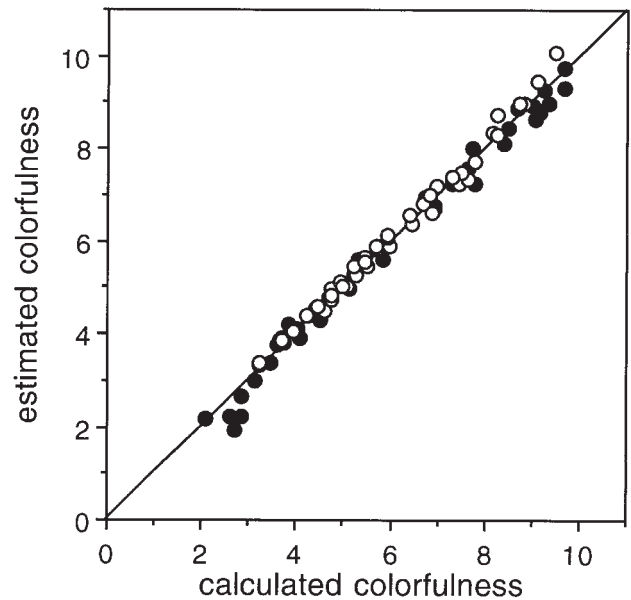


FIG. 7. Colorfulness calculated as the linear combination of the average chroma and the variability of chroma in terms of standard deviation plotted vs. the experimentally obtained colorfulness judgments, averaged over all the subjects, for the images of all four scenes. Filled symbols: multiplication. Open symbols: translation.

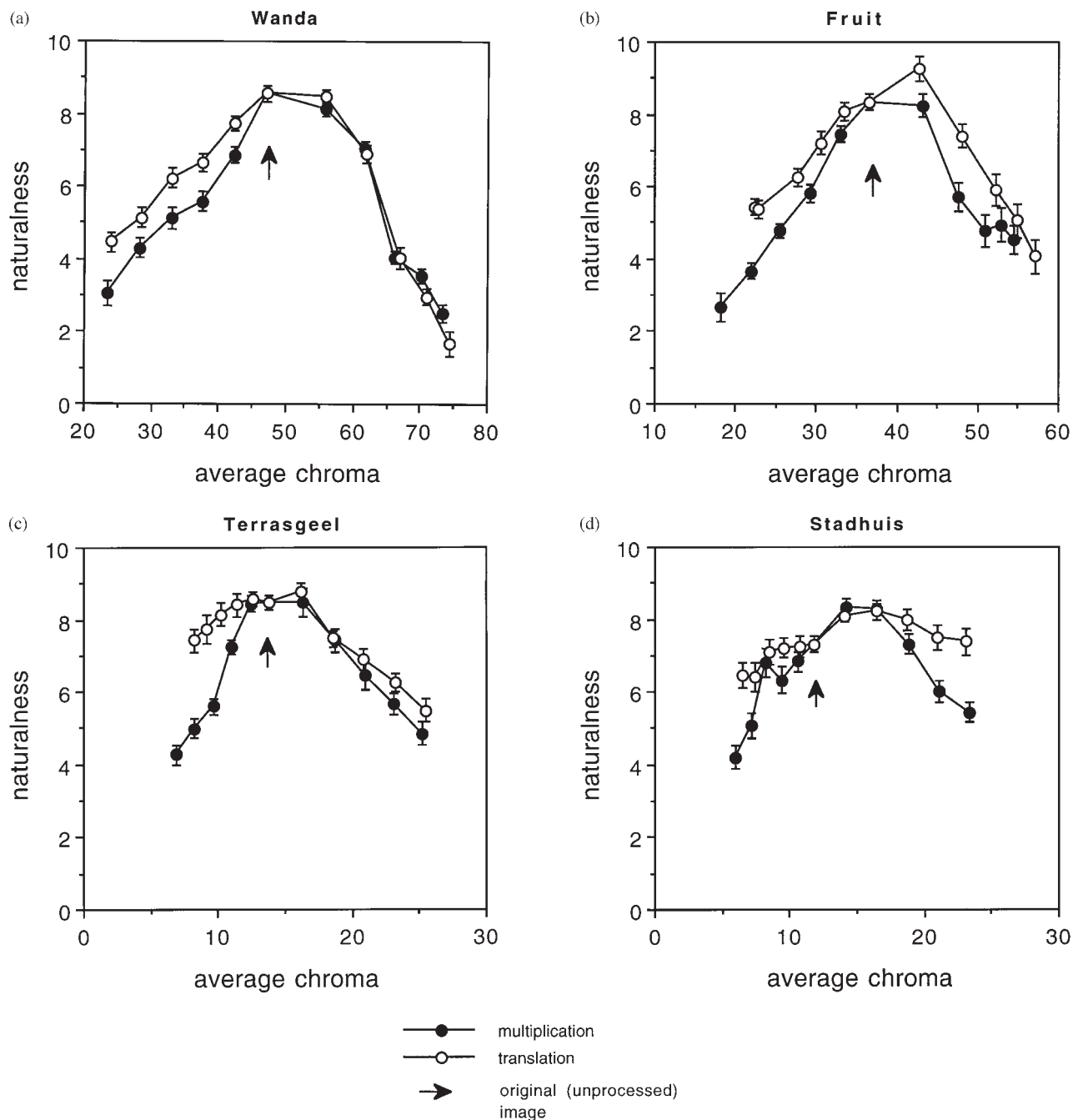


FIG. 8. Averaged scaled naturalness estimations plotted as a function of the average chroma for four scenes: (a) Wanda, (b) Fruit, (c) Terrasgeel, and (d) Stadhuis. Filled symbols: multiplication. Open symbols: translation. The arrow points to the naturalness judgment of the original image.

0.8, 0.9, 1.2, 1.4, 1.6, 1.8, and 2.0. As for translation, in order to make the two methods comparable, we subtracted the values of chroma that amounted to five tenth, four tenth, three tenth of the average chroma of the original image, and so on, or added two tenth, four tenth, etc., in accordance with the constants used for multiplication. Thus, the whole stimulus set including the original images consisted of 84 different images. The average chroma

and its variability (standard deviation of chroma) were computed for each image. The mean of the differences in lightness and hue angle between corresponding pixels before and after transformation of an image was always less than 0.013 expressed in the appropriate units. Visual inspection of the prepared images confirmed the absence of noticeable distortions due to deviations in hue and lightness.

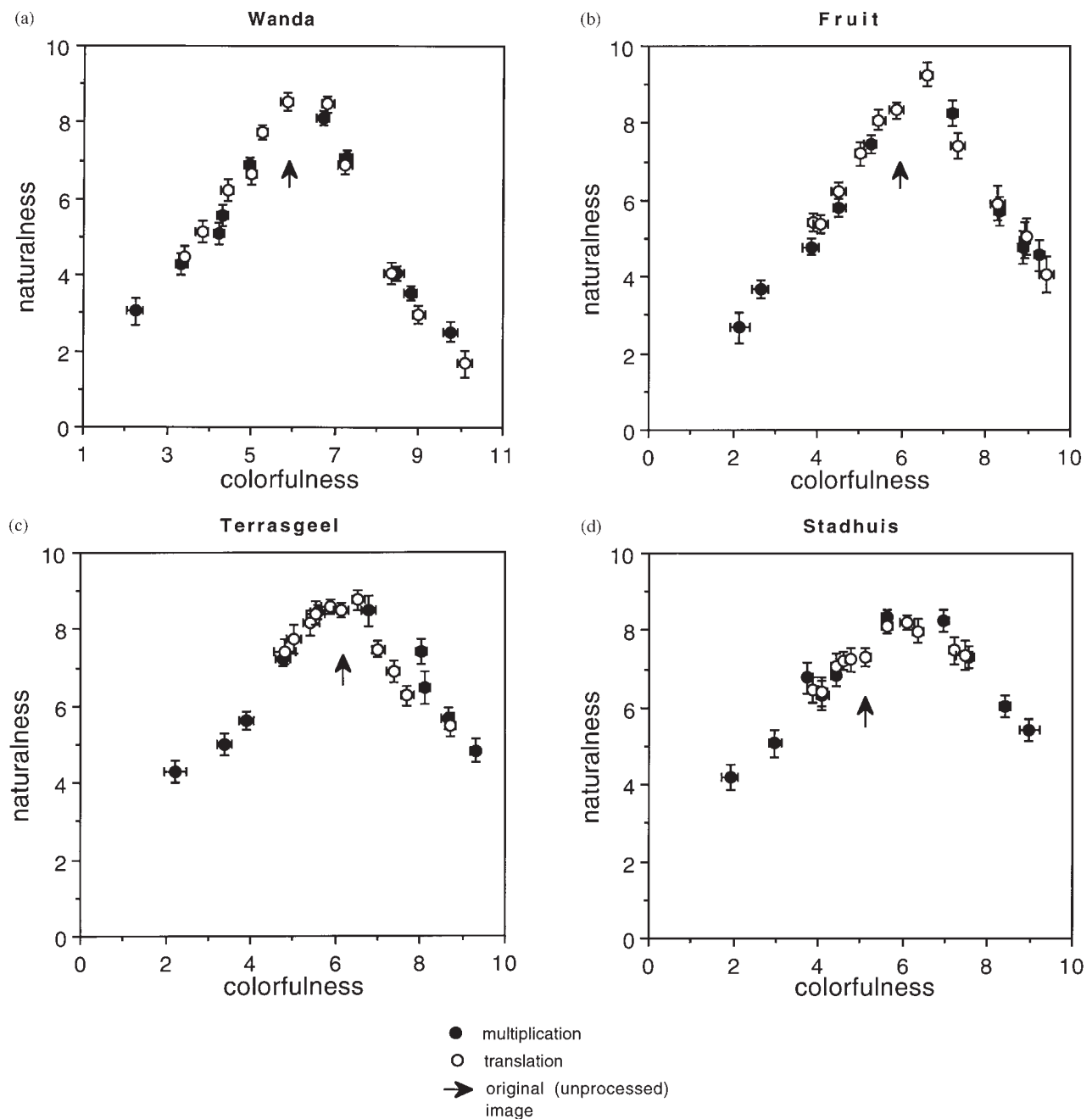


FIG. 9. Averaged scaled naturalness estimations plotted as a function of the colorfulness for the images of four scenes: (a) Wanda, (b) Fruit, (c) Terrasgeel, and (d) Stadhuis. Filled symbols: multiplication. Open symbols: translation. The arrow points to the judgments of the original image.

### EXPERIMENTAL PROCEDURE

Five subjects participated in the experiments. They had normal or corrected-to-normal visual acuity. Their color vision was checked with the aid of the H-R-R Pseudoisochromatic Plates.<sup>15</sup> No color deficiencies were observed.

The subjects viewed the monitor placed in a darkened room at a distance of about 1.7 m. At this distance, the pixel size is about 1 min of arc. In three separate experi-

ments, the subjects rated the perceptual quality, the colorfulness, and the naturalness of the images on a ten-point numerical category scale ranging from one (lowest attribute strength) to ten (highest attribute strength). The following instructions were read to the subjects.

#### Experiment 1: Perceptual Quality Estimation

“You will be presented with a sequence of images. Every time, you will see one image. Your task is to

judge the perceived quality of that image using an integer number from one to ten. One corresponds to the lowest perceptual quality, ten to the highest perceptual quality, according to your opinion. Perceptual image quality is defined as the degree of excellence of the image. Before starting the experiment, you will see a training series, which helps you to understand the task and to get an idea about the range of perceptual quality for the images employed.”

### Experiment 2: Colorfulness Estimation

“You will be presented with a sequence of images. Every time, you will see one image. Your task is to judge the perceived colorfulness of that image using an integer number from one to ten. One corresponds to the lowest degree of colorfulness, ten to the highest degree of colorfulness, according to your opinion. Colorfulness is defined as presence and vividness of colors in the whole picture. Before starting the experiment, you will see a training series, which helps you to understand the task and to get an idea about the range of colorfulness for the images employed.”

### Experiment 3: Naturalness Estimation

“You will be presented with a sequence of images. Every time, you will see one image. Your task is to judge the perceived naturalness of that image using an integer number from one to ten. One corresponds to the lowest degree of naturalness, ten to the highest degree of naturalness, according to your opinion. Naturalness is defined as the degree of correspondence between the image reproduced on the screen and reality, that is, the original scene as it is according to you at the time

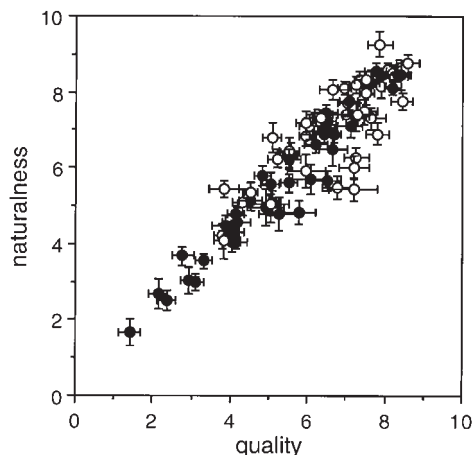


FIG. 10. Averaged scaled naturalness estimations plotted vs. estimations of the perceptual quality for the images of all four scenes. Filled symbols: multiplication. Open symbols: translation.

the picture was taken. Before starting the experiment, you will see a training series, which helps you to understand the task and to get an idea about the range of naturalness for the images employed.”

Although the original images were included, it should be emphasized that they were not explicitly identified to the subjects as such.

All subjects first participated in Experiment 1. Then three subjects took part in Experiment 2 followed by Experiment 3. The reverse holds for the other two subjects. Each experiment consisted of one session in which all 84 images were presented four times in a random sequence. Within each experiment, this sequence was the same for every subject. However, the sequence varied between experiments. The images were displayed on the black screen for 5 s, while the interval between two image exposures was at least 3 s, during which a 4.6 cd/m<sup>2</sup> gray adaptation field with the chromaticity coordinates of D65 appeared on the screen. The images and adaptation field had the same size. Before starting a session, the subjects judged a training series of 30 images.

## RESULTS

### Quality Estimation

In the first experiment, we obtained four estimations of the perceptual quality of each image for each subject. In order to correct for possible nonlinearities in the response scale, the data obtained for each subject were transformed to a subjective interval scale on the basis of Thurstone’s model of categorical judgment (for details, see Torgerston<sup>16</sup>). The computer program we used<sup>17</sup> performed the transformation via a maximum likelihood estimation. After the transformation no systematic differences between the subjects were observed. The quality estimations of the four series of images, averaged over the subjects, are presented in Fig. 3, plotted as a function of their computed average chroma. Four main conclusions can be derived from these graphs. First, in every series of images there is one image with a maximum quality value. Secondly, this image has an average chroma that is slightly higher than that of the original (unprocessed) image. Thirdly, there is a difference between the scenes: the quality estimations for “Stadhuis” differ from those for the other scenes in that the perceived quality of “Stadhuis” increased in almost all the steps in which the chroma was increased. This we tend to attribute to the low value of chroma for all pixels of the original image.

The observed difference is in line with the results obtained by Stokes et al.<sup>18</sup> showing that scene content significantly influences acceptability but not perceptibility tolerances for chroma manipulations of pictorial images in the CIELAB color space. Finally, there is a difference between the quality estimations of the images prepared via multiplication and those of the images prepared via

translation: the quality estimations of the latter images are almost always higher for the same average chroma. This was particularly evident in the case of the scenes ‘‘Fruit’’ and ‘‘Terrasgeel,’’ and also in the case of all the scenes in the part of the graph where the chroma was less than in the original image. The systematic difference in the results obtained for the two different methods implies that, in our experiment, the average chroma was not the only variable determining the perceptual image quality. This other factor is not likely to be the result of the clipping procedure. Although the color point distributions showed a similar amount of clipping when chroma is increased, translated images were in general judged to be of higher quality than the ones transformed by multiplication. Further, unlike the multiplication procedure, decreasing chroma by means of translation leads to clipping at zero value. Yet the quality ratings for the translated images are systematically higher for all four scenes. None of the subjects reported local distortions due to clipping.

To understand the observed discrepancy, we looked for a global perceptual attribute of the pictures underlying the quality estimations. Since we had changed the chroma of each pixel of the image, it was proposed that such an attribute might be the colorfulness of the whole image. The next section describes the results of the second experiment, in which the subjects judged the colorfulness of the same set of images.

### Colorfulness Estimation

In the second experiment, we obtained estimations of colorfulness, which we then transformed in the same way as in the previous experiment. Because of the great similarity of the data of the different subjects, we averaged the estimations over all the subjects for every scene. The averaged subjective colorfulness estimations are presented as a function of the average chroma in Fig. 4. Subsequently, we used the values of colorfulness as the magnitudes of the perceptual attribute that determines the quality estimations. The results of this substitution are presented in Fig. 5. The discrepancy between the two sets of images (obtained via multiplication and translation), observed when chroma was used as the independent variable determining the quality estimations, has disappeared. The resulting coincidence of the quality judgments demonstrates that, in our experiments, colorfulness was indeed the main perceptual attribute determining the quality. There is an optimum value of colorfulness (Fig. 5) that provokes a subjective preference in quality. When the colorfulness decreases, the perceptual quality decreases very rapidly. The opposite was observed when the colorfulness increased. Only within the small region around the original images were the quality estimations high. The colorfulness of the images of the highest perceptual quality is somewhat higher than that of the originals. The same, systematic trend was observed for all four scenes. The

difference in quality between the original images and the images of the highest perceptual quality is significant as shown by Student’s *t*-test ( $t = 3.267$ ,  $p < 0.001$ ).

### What Determines Colorfulness?

One of the intriguing questions is how colorfulness depends on the distribution of the colors in color space. Intuitively, we propose that colorfulness depends on (1) the distance of image colors from neutral gray, and (2) the distances between individual image colors. We model the first factor as average chroma. The second factor is in first order incorporated as the standard deviation of the distribution of the chroma values of all individual image colors.

From Fig. 4, it is indeed evident that besides the average chroma there is another factor that systematically influences subjective colorfulness: the colorfulness of the images prepared via multiplication is lower than the colorfulness of those prepared via translation when the average chroma is less than in the unprocessed images. The opposite relation is observed when the average chroma is higher than in the original images. We suggest that this other factor might be the average distance between individual image colors modeled as the variability of the average chroma expressed in, for instance, the standard deviation of chroma. The fact that these values differ in the cases of multiplication and translation follows from the definition of these computations. In Fig. 6, the computed standard deviation of chroma is plotted vs. the average chroma. In an ideal situation, we would have two straight lines for each scene: an oblique line for the images prepared via multiplication, where the slope corresponds to the ratio of the standard deviation and the average of the distribution of chroma values, and a horizontal line for the images produced via translation. However, because of the restrictions of the monitor, the values of chroma expected for some pixels could not be achieved when the chroma was increased, or some pixels had zero value and their chroma could not be decreased as for the other pixels. These limitations led to more complicated relations between the average chroma and its spread.

If we compare Figs. 4 and 6, we observe a strong dependency of the subjective value of colorfulness on both the average chroma and its variability represented by the standard deviation of the distribution of chroma values. This dependency of colorfulness on these two variables can be efficiently described by their linear combination constituting the total value of the perceptual attribute considered:

$$\text{Colorfulness} = A * (\text{Average Chroma}) + B * (\text{Standard deviation of Chroma}) + C. \quad (9)$$

The results of the regression analysis are given in Table I. Both variables, i.e., average chroma and its variability,

significantly contribute to the impression of colorfulness in the case of all four scenes. Note, however, that the relative contribution of the two variables differs between the scenes with a high average chroma (Wanda, Fruit) and those with a low average chroma (Stadhuis, Terrasgeel). We can rewrite the equation as follows:

$$\text{Colorfulness} = A * (\text{Average Chroma} + D * \text{Standard deviation of Chroma}) + C, \quad (10)$$

where  $A$  and  $C$  are constants depending on the scaling procedure, and  $D$ , being  $B/A$ , reflects the relative contribution of the two variables. The fact that the value of the adjusted squared multiple correlation coefficient exceeds 0.977 in the case of all scenes supports the validity of the model for these experiments. The close correspondence between the experimentally determined colorfulness and the computed colorfulness is illustrated in Fig. 7.

### Why the Quality Decreases

To answer the question of why the perceptual quality of the images first increases and then decreases with increasing colorfulness, we asked our subjects to judge the naturalness of the images in the third experiment. The results of that experiment are very similar to those obtained in the quality estimation experiment. As one can see in Fig. 8, there is a difference between the values of naturalness plotted as a function of the average chroma obtained for the images produced via multiplication and those obtained for the images produced via translation. At the same time, the arrangement of the naturalness estimations based on the colorfulness of the images irrespective of the transformation method applied demonstrates the key role of subjective colorfulness in determining the naturalness of the images (Fig. 9).

The correspondence between the quality and naturalness estimations obtained for all the images is presented in Fig. 10. There appears to be a strong relation between those attributes. In fact, the correlation coefficient between the quality and naturalness estimations, averaged over the four scenes, amounts to 0.929. Such a significant coincidence implies that perceptual quality and naturalness lie very close to one another.<sup>19</sup> This result stresses the importance of naturalness as a decisive constituent of image quality.

### DISCUSSION

The experimental results show that the perceptually uniform CIELUV color space is well suited for manipulating color images in a way that can be easily interpreted. This follows from the fact that the transformation of the images in this space led to an increase in their quality. But most importantly, such a global perceptual attribute as the colorfulness of the whole picture could be expressed in terms of the statistics of the color point distribution in the

CIELUV color space. Using two different types of chroma variation, we were able to identify the variability of chroma as the factor that, together with average chroma, influences perceived colorfulness. Hence, both factors should be considered when optimizing image quality.

The fact that the perceptual quality of the original image can be improved emphasizes the subjective preference for more saturated, indeed, more colorful images. A similar effect has been reported by Hunt,<sup>20</sup> who found that an observer may sometimes prefer a rendering of a scene that looks different from the original, with “. . . skies that are bluer, skin that is more saturated and sunsets that are redder” (p. 48). The same preference for more highly saturated colors was measured by Siple and Springer.<sup>21</sup> The consistency of subjective preference that was found in our experiments supports the approach of systematic transformation of images on the basis of different attributes of color perception for optimization of color reproduction.

The strong link between image quality and naturalness that was found in the experiments suggests that naturalness is an important perceptual constraint in the color reproduction of images of natural scenes. Naturalness, however, is not identical to image quality. This can be appreciated, if one compares the quality and naturalness estimations in the case of the “Stadhuis” scene [Figs. 5(d), 9(d)]. For this scene, the perceived quality remained high while naturalness decreased for the most colorful images. This phenomenon is less pronounced but still present in the case of the other scenes.<sup>19,22</sup> This difference between quality and naturalness indicates that subjects sometimes used different criteria to assess naturalness and quality: the subjective preference in quality was biased towards more colorful images although the observers realized that these images looked somewhat unnatural. The strong relation between quality and naturalness in the case of the woman’s portrait highlights the great importance of the presence of familiar objects in a scene, in this case, for example, the skin.<sup>23</sup> As a result, subjective uncertainty in estimating image quality and naturalness was found to be minimal.

In our experiments, all images contained living and natural objects, e.g., a person, fruit, plants, which might strengthen the naturalness constraint on quality judgments. Further research will determine whether naturalness criteria can also be applied to images lacking such natural objects.

### CONCLUSIONS

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

1. The colorfulness of an image is the main perceptual

- attribute underlying image quality and naturalness when chroma varies.
- The colorfulness of an image can be modeled as a function of the averaged chroma and the variability of chroma as derived from the color point distribution in the CIELUV color space.
  - Colorfulness enhancement enables higher perceptual quality. The further increase in colorfulness, however, leads to a reduction in quality due to the resulting decrease in naturalness.
  - The perceptual quality of appreciation-oriented images is strongly related to naturalness, which demonstrates the importance of naturalness as a decisive constituent of image quality.

#### ACKNOWLEDGMENTS

The authors wish to thank Prof. J. A. J. Roufs, Prof. E. N. Sokolov, and Dr. J. B. Martens for their useful advice and help. The research of E. A. Fedorovskaya was made possible by a fellowship from the Eindhoven University of Technology, the Netherlands.

- J. C. Dalton, Visually optimized color-image enhancement, in *Perceiving, Measuring, and Using Color*, M. H. Brill, Ed., Proc. SPIE 1250, 1990, pp. 177–189.
- J.-Y. Kim and Y.-H. Ha, Pseudo-linearly modified IHS color model and its application to color image enhancement, in *Proceedings of the First IS&T/SID Color Imaging Conference: Transforms & Transportability of Color*, Scottsdale, Arizona, USA, 1993, pp. 23–27.
- R. W. G. Hunt, *The Reproduction of Colour in Photography, Printing & Television*, Fountain Press, Tolworth, U.K., 1987, pp. 36–48.
- J. A. J. Roufs, Perceptual image quality: concept and measurement. *Philips J. Res.* **47**, 35–62 (1992).
- M. A. Arbib and A. R. Hanson, Vision, brain, and cooperative computation: an overview, in *Vision, Brain, and Cooperative Computation*, M. A. Arbib, A. R. Hanson, Eds., A Bradford Book, The MIT Press, Cambridge, MA, 1987, pp. 1–83.
- G. Derefeldt, Colour appearance systems, in *Vision and Visual Dysfunction: Vol. 6: The Perception of Colour*, P. Gouras, Ed., Macmillan Press, New York, 1991, pp. 218–261.
- S. Tominaga, Color classification of natural images. *Color Res. Appl.* **17**, 230–239 (1992).
- Y. Ohta, *Knowledge-Based Interpretation of Outdoor Natural Color Scenes*, Pitman Advanced Publishing Program, Boston, 1985, pp. 136.
- W. De Corte, Finding appropriate colors for color displays. *Color Res. Appl.* **11**, 56–61 (1986).
- R. C. Carter and E. C. Carter, High-contrast sets of colors. *Appl. Opt.* **21**, 2936–2939 (1982).
- R. W. G. Hunt, *Measuring Colour*, 2nd ed., Ellis Horwood, Chichester, 1992.
- International Lighting Vocabulary, Publication CIE No. 17.4, Geneva, 1987.
- G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed., John Wiley and Sons, New York, 1982.
- M. R. Pointer, The gamut of real surface colours. *Color Res. Appl.* **5**, 145–155 (1980).
- L. H. Hardy, G. Rand, and M. C. Rittler, *H-R-R-Pseudoisochromatic Plates for Detecting, Classifying, and Estimating the Degree of Defective Color Vision*, American Optical Company, New York, 1957.
- W. S. Torgerson, *Theory and Methods of Scaling*, Wiley, New York, 1958.
- M. C. Boschman, THURCATD: A program for processing results of numerical category scaling: calculation of parameters in the Thurstone model, condition D. *IPO report 108*, 1991.
- M. Stokes, M. D. Fairchild, and R. S. Berns, Colorimetrically quantified visual tolerances for pictorial images, in *Proceedings of the TAGA & ISCC Conference on Comparison of Color Images Presented in Different Media*, Williamsburg, Virginia, February 23–26, 1992, pp. 757–777.
- H. de Ridder, E. A. Fedorovskaya, and F. J. J. Blommaert, Naturalness and image quality: chroma variation in color images of natural scenes. *IPO Annual Progress Report* **28**, 89–95 (1993).
- R. W. G. Hunt, Chromatic adaptation in image reproduction. *Color Res. Appl.* **7**, 46–49 (1982).
- P. Siple and R. M. Springer, Memory and preference for the colors of objects. *Percept. & Psychophys.* **34**, 363–370 (1983).
- H. de Ridder, F. J. J. Blommaert, and E. A. Fedorovskaya, Naturalness and image quality: chroma and hue variation in color images of natural scenes, in *Human Vision, Visual Processing, and Digital Display VI*, B. E. Rogowitz and J. P. Allebach, Eds., Proc. SPIE 2411, 1995, pp. 51–61.
- Y. Miyake and H. Haneishi, The evolution of nonimpact print—impact on image quality, in *Proceedings of the First IS&T/SID Color Imaging Conference: Transforms & Transportability of Color*, Scottsdale, Arizona, USA, 1993, pp. 222–227.