Influence of Chroma variations on Naturalness and Image Quality of Stereoscopic Images

Andre Kuijsters[a], Wijnand A. IJsselsteijn[a], Marc T. M. Lambooij[a,b], Ingrid E. J. Heynderickx [b,c]

[a] Eindhoven University of Technology, P.O. Box 513, 5600MB, Eindhoven, The Netherlands
[b] Philips Research Laboratories, High Tech Campus 34, 5656AE, Eindhoven, The Netherlands
[c] Delft University of Technology, Mekelweg 4, 2628CD, Delft, The Netherlands

ABSTRACT

The computational view on image quality of Janssen and Blommaert states that the quality of an image is determined by the degree to which the image is both useful (discriminability) and natural (identifiability). This theory is tested by creating two manipulations. Firstly, multiplication of the chroma values of each pixel with a constant in the CIELab color space, i.e., chroma manipulation, is expected to increase only the usefulness by increasing the distances between the individual color points, enhancing the contrast. Secondly, introducing stereoscopic depth by varying the screen disparity, i.e., depth manipulation, is expected to increase both the usefulness and the naturalness. Twenty participants assessed perceived image quality, perceived naturalness and perceived depth of the manipulated versions of two natural scenes. The results revealed a small, yet significant shift between image quality and naturalness as a function of the chroma manipulation. In line with previous research, preference in quality was shifted to higher chroma values in comparison to preference in naturalness. Introducing depth enhanced the naturalness scores, however, in contrast to our expectations, not the image quality scores. It is argued that image quality is not sufficient to evaluate the full experience of 3D. Image quality appears to be only one of the attributes underlying the naturalness of stereoscopic images.

INTRODUCTION

For decades image quality has been regarded as the most important measure for evaluating imaging systems. A lot of effort has been devoted to the development of reliable and widely applicable models that can, directly or indirectly, link perceived image quality to the underlying technological parameters of the imaging system. Such image quality models are important to reduce the need for expensive and time-consuming subjective tests, which have to be repeated for each new parameter setting. The common way to model and predict perceived image quality is to identify and quantify the presence of visual distortions in an image. Examples of common visual distortions are blur, noise, color deviations, non-uniformities, etc.. The amount of deviation between the rendered image and the undistorted original is often used as a measure for image quality. Such an approach, however, cannot describe situations in which the original image does not possess the highest quality, as found by amongst others Hunt (1982)[1] and Siple and Springer (1983).[2] In such situations, image quality should be predicted from the rendered image only. One possible approach is given by the Image Quality Circle model of Engeldrum.[3] He states that all viewers have a kind of common internal reference framework to distinguish good from bad images. Judging a given image with respect to that framework is done by weighting underlying attributes, such as brightness, color, sharpness, etc.. A different view on image quality has defined quality as the degree to which an image can be successfully exploited by the observer.[4] In this view, referred to as computational image quality, two main requirements should be satisfied.[5] Firstly, the internal image representation should be precise, which is defined as the usefulness of an image. Janssen and Blommaert (2000)[6] argued that usefulness is determined by the ability to discriminate between items, and that it does not necessarily imply a precise one-to-one relation between the internal image representation and the physical characteristics of the item. Items that can be discriminated better may provide more useful information to the observer about the outside world. Secondly, the match between the representation and the knowledge of reality as stored in memory should be high, defined as the naturalness of an image. The degree to which an image is both useful and natural determines the quality of the image.

Different experiments revealed the separate effect of naturalness and usefulness on image quality. The main results of these experiments will be discussed shortly, but the authors refer to de Ridder (2001)[7] for a more extensive review. In

Human Vision and Electronic Imaging XIV, edited by Bernice E. Rogowitz,
Thrasyvoulos N. Pappas, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7240, 72401E

one experiment the color temperature of the reference white was varied in the CIELuv color space between 4650 K and 10300 K, 6500 K being the original. Both image quality and naturalness decreased similarly as soon as the color temperature of the reference white deviated from that of the orginal.[4] Similar results were found for hue manipulated natural images.[8,9] Both image quality and naturalness deteriorated as soon as the hue deviated from its original value. Manipulating the chroma value in the CIELuv colorspace per pixel of the natural image, however, resulted in a systematic deviation between image quality and naturalness judgements.[4,8,9,10,11] Image quality judgements were shifted to higher chroma values relative to the naturalness judgements. In other words, people preferred images that were more colorful than the original although they realized that these images looked somewhat unnatural. Manipulating the brightness contrast of natural black and white images revealed similar results.[4] Participants preferred higher contrast images although they realized that these images looked somewhat unnatural. Increasing the chroma or brightness contrast level resulted in larger distances between individual pixels in the CIELuv color space, and thus, in a higher perceived contrast. This higher contrast provided a more accurate localisation and detection of edges. As a result, the ability to discriminate between items in an image increased, thereby providing a more useful image. Varying the hue or the color temperature of the reference white, however, did not enlarge the overall distance between individual color points. When varying the color temperature, the distance between color points may be locally increased in some parts of the color space, but is compressed in other parts. Hence, these manipulations did not affect the overall ability to discriminate between items in an image, so the usefulness of the image is not affected. As a result, image quality judgements and naturalness judgements were almost identical. Increasing the chroma or the brightness contrast enhanced the usefulness of the images; consequently, image quality deviated from naturalness towards more chromatic or higher contrast images.

Research on the assessment of stereoscopic images revealed a similar shift between image quality and naturalness. Participants evaluated the quality of depth and naturalness of depth of stereoscopic images with different depth levels.[12,13] Both quality of depth and naturalness of depth increased with increasing disparity, yet decreased at the extreme depth levels, resulting in an inverted 'U-shape' function. Additionally, preference in quality was shifted to higher depth levels in comparison to preference in naturalness. Thus, participants preferred (in quality) a reproduction of stereoscopic depth they also judged to be slightly unnatural. The higher depth levels may have increased the ability to discriminate between objects positioned on different depth planes, thereby enhancing the usefulness of the images.

The computational view on image quality suggests that perceived image quality can be enhanced by increasing both the usefulness and the naturalness of an image. We tested this view by introducing stereoscopic depth in natural images. This manipulation was expected to increase both the usefulness and the naturalness of these images, since it was expected to not only increase the discrimination ability between depth planes, but also to provide a more realistic and truthful representation of reality. Seuntiens et al. (2005)[14] already showed that introducing stereoscopic depth significantly enhanced the naturalness judgments. Different studies revealed improved usefulness of 3D images over 2D images.[15,16] By incorporating several depth levels we investigated whether the shift between quality of depth and naturalness of depth as found by IJsselsteijn et al.[12,13], is reproduced when assessing overall quality and naturalness. Besides the depth manipulation, we also varied the chroma value per pixel of the originals, creating more and less colorful images. With this manipulation we tried to replicate the shift between image quality and naturalness, as found on monoscopic images,[4,8,9,10,11] for stereoscopic images. Additionally, we determined whether people use similar considerations when evaluating color manipulated 3D as when assessing color manipulated 2D images.

EXPERIMENTAL SETUP

Design. The experiment had a within-subjects design, with as independent variables: content (2 scenes), chroma manipulation (6 levels) and dimension (i.e., screen disparity, 5 levels). The assessment scores on perceived image quality, perceived naturalness and perceived depth were the dependent variables.

Participants. Twenty people participated in the experiment, six of which were female and fourteen male. Their age ranged from 24 to 46. Nine of them were employees of Philips Research, while eleven were internal graduation students. All observers had a good visual acuity ≥ 1 (as tested with the Landolt-C test), good stereovision (≤ 30 second of arc as tested with a RANDOT stereo test) and no color deficiencies (as tested with the Ishihara test).

Display. The stimuli were displayed on a 42” Philips nine-view auto-stereoscopic display (42-3D6W02). This display had a lenticular placed on top of the pixels of a normal LCD, and as such, allowed stereoscopic vision without the need of wearing glasses.[17,18] The lenses in the lenticular directed the light from different pixels into different directions. By placing the right image information on the right pixels, a slightly different view was presented in different directions, e.g. to the left and right eye. The Philips 3D display used in the experiment provided a total of 9 views, enabling look-around capabilities.[17,18] The optimized viewing distance for the display was three meters.

Fig. 1 Scenes used in the experiment: People (left) and Balloon (right).

Stimuli. Two still images were used in the experiment (depicted in Figure 1). The input format for the Philips 3D display was a RGB image (original) + corresponding Z-image (depth map), with a resolution of 940 x 540 each. The depth map was a gray-scale image, in which the gray value indicated the relative depth per pixel (255 is maximal towards the observer, 0 is maximally away from the observer). Five different depth (dimension) levels were created by varying the gain factor in the Z-image. This resulted in a 2D level (zero screen disparity) and four 3D levels with a maximum screen disparity of ±1 mm ($3D_1$), ±2 mm ($3D_2$), ±3 mm ($3D_3$) and ±4 mm ($3D_4$). The original RGB image was used as the middle view of the 9 views, i.e., view 5. The 4 views to the right and to the left were rendered automatically by the display according to the disparities given in the depth map.

Six levels of chroma were created by varying the chroma value per pixel in the CIE 1976 (L* a* b*) color space. The R, G and B values per pixel of the original image, scaled back to [0, 1], were first gamma (γ) corrected with $\gamma = 2.2$. These R, G and B values were then transformed into X, Y and Z tristimulus values, using the monitor characteristics. The chromaticity coordinates for reference white were $(x_w, y_w) = (0.279, 0.306)$, and for the three primaries were $(x_r, y_r) = (0.650, 0.330)$, $(x_g, y_g) = (0.272, 0.603)$ and $(x_b, y_b) = (0.141, 0.306)$. The maximum luminance for white was 500 cd/m^2. The resulting (3 x 3) transformation matrix from RGB to XYZ is shown in Equation 1.

Equation 1:
$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.385 & 0.315 & 0.212 \\ 0.195 & 0.699 & 0.106 \\ 0.012 & 0.151 & 1.194 \end{bmatrix} \bullet \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The XYZ tristimulus values were then translated into the CIELab color coordinates, from which the hue angle (h_{ab}) and the chroma value (C_{ab}) for each pixel were determined. This chroma value was then multiplied with a constant, while the lightness and hue angle were kept constant. The chroma value was either decreased using a constant of 0.4 and 0.8, or increased, using a constant of 1.2, 1.4 and 1.8. Including the original scene (constant of 1.0), six levels of chroma were created. The new R, G and B values were then computed by first transforming the hue angle and chroma value per pixel back to L*a*b* coordinates, then to XYZ tristimulus values, and finally to RGB values using the inverse of the matrix in Equation 1. Pixels with a value larger than one or smaller than 0 were outside the gamut of the display and were clipped. The resulting R, G and B values were corrected for the gamma of the LCD display by applying an inverted γ of 2.69, and scaled back to [0, 255].

Procedure. Participants were requested to score the perceived image quality, perceived naturalness and perceived depth for all 60 stimuli (i.e., 2 different scenes x 6 levels of chroma x 5 dimension levels). Each attribute was assessed in a different session. Participants judged all stimuli using the scale standardized by ITU-BT.500-11,[19] which was labeled with the adjectives [bad]-[poor]-[fair]-[good]-[excellent]. Between these five standard marks four intermediate marks were added to allow the participants to refine their judgments.

Before starting a session, the participants judged in a training session a total of 8 images. The stimuli were displayed until the participants assessed them in terms of one of three evaluation criteria. Between two subsequent stimuli a gray screen was displayed for 3 seconds to minimize adaptation effects. The order of the stimuli as well as the order of the sessions was randomized to counterbalance order effects.

Because unidimensional attributes were rated on an ordinal categorical scale, the data were first transformed using Thurstone scaling. With this scaling it was tested whether equal distances on the assessment scale were perceived as equal, i.e., whether the difference between [bad] and [poor] was perceived similar to the difference between for example [fair] and [good]. The ThurcatD program was used for this analysis.[20] The resulting data indicated that intervals were perceived equally, which allowed the raw data to be analyzed in an ANOVA. The data were analyzed in SPSS 15.0, using a repeated measure ANOVA. The Bonferroni post hoc test was used to specify significant main effects. To specify the importance of an effect, the effect size was calculated based on the omega squared method (ω^2).[21] For interpretation of the effect size, Keppel (1991)[21] gives to following guidelines; 0.01 small effects, 0.06 medium effects and 0.15 large effects.

RESULTS

Because the repeated measure ANOVA revealed no significant effect of content on all three evaluation criteria (e.g., image quality $F(1,19)=2.243$ $p=.151$, naturalness $F(1,19)=3.778$ $p=.067$ and depth $F(1,19)=1.534$ $p=.231$), only the averaged results are presented in the figures.

Fig. 2 represents the averaged naturalness and image quality scores as a function of chroma for both the 2D images as well as their 3D counterparts. Obviously, the assessment scores of image quality and naturalness are highly affected by chroma; they both decrease at low chroma values (i.e., 0.4 and 0.8) and also when the chroma is too high (i.e., at 1.8), resulting in an inverted ‘U-shape’ behavior. Stimuli with a slightly higher chroma than the original (i.e., 1.2) receive the highest scores in terms of image quality, while the original (i.,e., chroma of 1.0) is perceived as most natural. This small shift between the image quality and naturalness judgments is observed for both scenes and in each dimension level. It is more pronounced for the stereoscopic images than for the monoscopic images. The ANOVA results confirm this finding: the effect of chroma is significant on the shift (defined as the image quality assessment score minus the naturalness assessment score) for the stereoscopic stimuli ($F(5,954)=4.161$, $p=.001$, $\omega^2=.001$), however, not for the ‘2D’ dimension level ($F(5,234)=0.336$, $p=.891$, $\omega^2<.001$). Although the effect is highly significant, the size of the effect is very small. As expected, preference in quality is shifted to higher chroma values with respect to preference in naturalness.

Fig. 3 represents the averaged naturalness and image quality scores as function of dimension. Again, there is the tendency of an inverted ‘U-shape’ behavior, albeit far less pronounced than compared to the chroma manipulation. Especially the image quality scores hardly depend on dimension level at low values; only at a higher disparity the quality scores drop. The ANOVA analysis reveals a significant effect of dimension on image quality ($F(4,76)=12.402$, $p<.001$, $\omega^2=.046$). The highest dimension level is scored significantly lower than the other four dimension levels (all $p<0.001$); without the highest dimension level no significant effect of dimension on the image quality assessment is found ($p=0.095$). As expected, the perceived naturalness scores are enhanced by the introduction of stereoscopic depth, although too much depth decreases the naturalness score again. The ANOVA analysis reveals a significant effect of dimension on naturalness ($F(4,76)=4.068$, $p=.005$, $\omega^2=.022$). All four stereoscopic dimension levels are scored significantly higher than the '2D' dimension level (all $p\leq0.019$). Because of the absence of a clear maximum in the preference in image quality, it is difficult to evaluate any shift in preference between image quality and naturalness as a function of dimension. At least, we can state that the results of Fig. 3 do not at all tend to show a shift in preference in quality to higher depth values with respect to preference in naturalness.

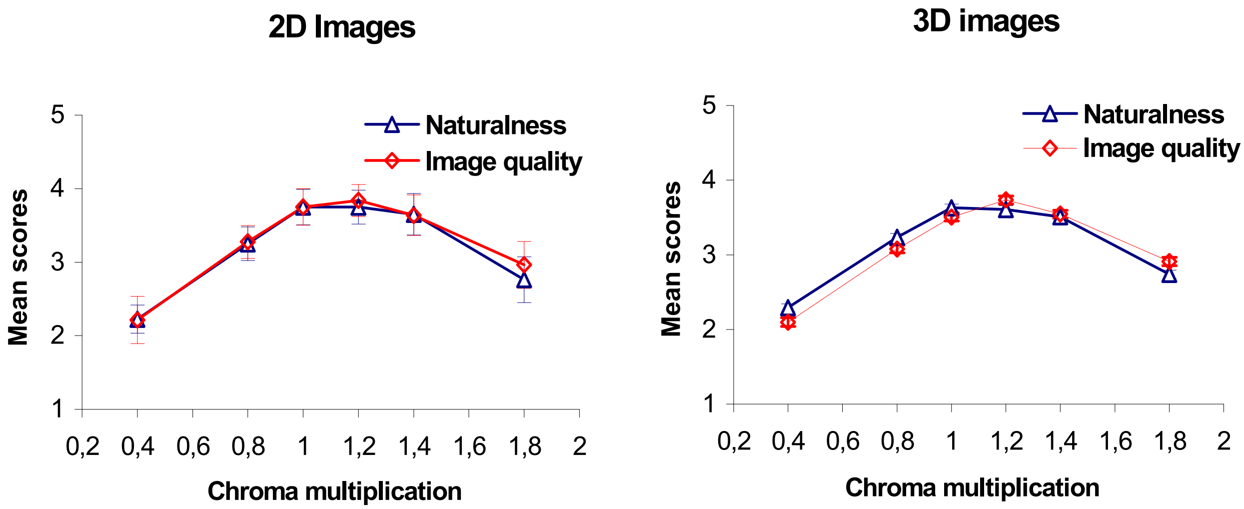

Fig. 2 Averaged naturalness (triangles) and image quality (diamonds) scores as function of chroma manipulation for the 2D stimuli (left) and the 3D stimuli (right). Error bars show the 95% confidence intervals.

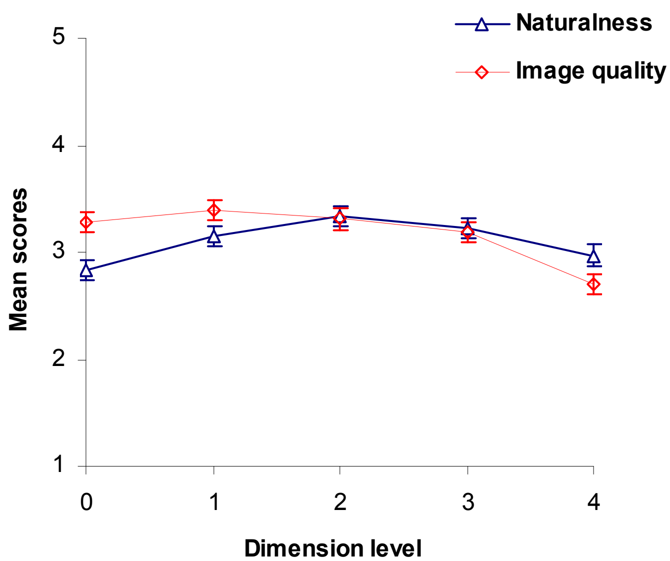

Fig. 3 Averaged naturalness (triangles) and image quality (diamonds) scores as function of dimension level. Error bars show the 95% confidence intervals. The numbers 0-4 on the horizontal axis denote the dimension levels '2D', '3D$_1$', '3D$_2$', '3D$_3$' and '3D$_4$', respectively.

Fig. 4 shows the mean naturalness, image quality and perceived depth scores as function of chroma; the different lines represent the different dimension levels. The graphs containing the naturalness and image quality scores illustrate again the conclusions already drawn from Figs. 2 and 3, but in more detail. It clearly demonstrates that the shift in preference in image quality to higher chroma values with respect to preference in naturalness is rather independent of dimension level. The graph containing perceived depth scores shows a clear distinction between the five dimension levels, i.e., depth perception increases with the amount of screen disparity introduced. There is also a small, yet highly significant effect of chroma on perceived depth ($F(5,95)=7.124$, $p<.001$, $\omega^2=.006$). The perceived depth scores increase with increasing chroma. Although this effect is very small, it is consistent over all dimension levels, including the '2D' dimension level.

CONCLUSION AND DISCUSSION

A deviation between image quality and naturalness judgments as a result of chroma manipulation is found. The highest image quality scores are shifted to higher chroma levels in comparison to the highest naturalness scores. This small shift, already known for monoscopic images[4,8,9,10,11] is reproduced for stereoscopic images, in both scenes and in all dimension levels. This result suggests that participants use similar considerations, when judging color manipulation of 2D and 3D images in terms of perceived naturalness and perceived image quality.

As expected, a small increase in chroma only increases the image quality, but not the naturalness of an image. The preference for a more colorful image than the original refers to a need for a higher color contrast. The manipulation resulted in larger distances between the individual pixels in the CIELab color space, increasing as such the color contrast of the images. The higher color contrast thereby increases the usefulness of the image. This result supports the computational view on image quality, i.e., for quality judgements both usefulness and naturalness are weighted.[5] Also the perceived depth increases with increasing chroma levels for all dimension levels. This suggests that people are better able to distinguish different depth layers at increased chroma values, and so that chroma adds to the usefulness of an image.

The question remains why the shift between image quality and naturalness is more pronounced for the stereoscopic images than for their 2D counterparts. It could be argued that a higher contrast is especially useful in stereoscopic images. The higher contrast may improve the ability to discriminate between items on different depth planes, thereby enhancing depth perception. The perceived depth scores, however, do not confirm this argument; increasing the chroma level has a similar effect on all dimension levels, including the '2D' dimension level.

Preference in quality is not shifted to higher depth levels in comparison to preference in naturalness. In contrast to the findings of IJsselsteijn et al.[12,13], the image quality scores are negatively affected at higher disparity even before the naturalness scores start to decrease. The lower image quality scores at the higher dimension level are most likely caused by a technological limitation of the display used; parts of the content displayed at high disparities become blurred due to crosstalk between the views.[22] The display used in the experiment is characterized by a large viewing zone for comfortable viewing, but has the drawback that it is limited in displaying extreme disparities without the perception of blur. The displays used in the experiments of IJsselsteijn et al.[12,13] did not have such a drawback. As a result, extreme depth levels could be displayed that were only decreased in quality because the extreme depth was perceived as unrealistic.[12,13]

Our results reveal that the computational view on image quality of Janssen and Blommaert[5] is not applicable to stereoscopic images. Introducing stereoscopic depth enhances naturalness, however, not image quality. The stereoscopic images are scored significantly higher in terms of naturalness than their 2D counterparts. Apparently, stereoscopic images are a more realistic and truthful representation of reality. This is also found earlier with the results of other studies using different 3D displays.[14,23,24] The image quality scores are not increased by the stereoscopic depth, and even decreased at the higher depth levels. The negative influence of introducing too much depth, i.e., the perception of blur due to crosstalk between the views, is incorporated in the image quality scores, while the positive effect of depth, namely, a more natural and useful image, is not observed in the image quality scores. Additional research exists that supports the finding that image quality does not take the positive effect of stereoscopic depth into account,[23,24,25,26,27] while only negative effects (due to crosstalk[28] and blur[24]) are incorporated in the image quality judgment. These results suggest that quality assessment of stereoscopic images is not based on the naturalness and the usefulness of these images. Indeed, an increase in naturalness at constant quality with increasing depth can only be explained with a decrease in usefulness with increasing depth.

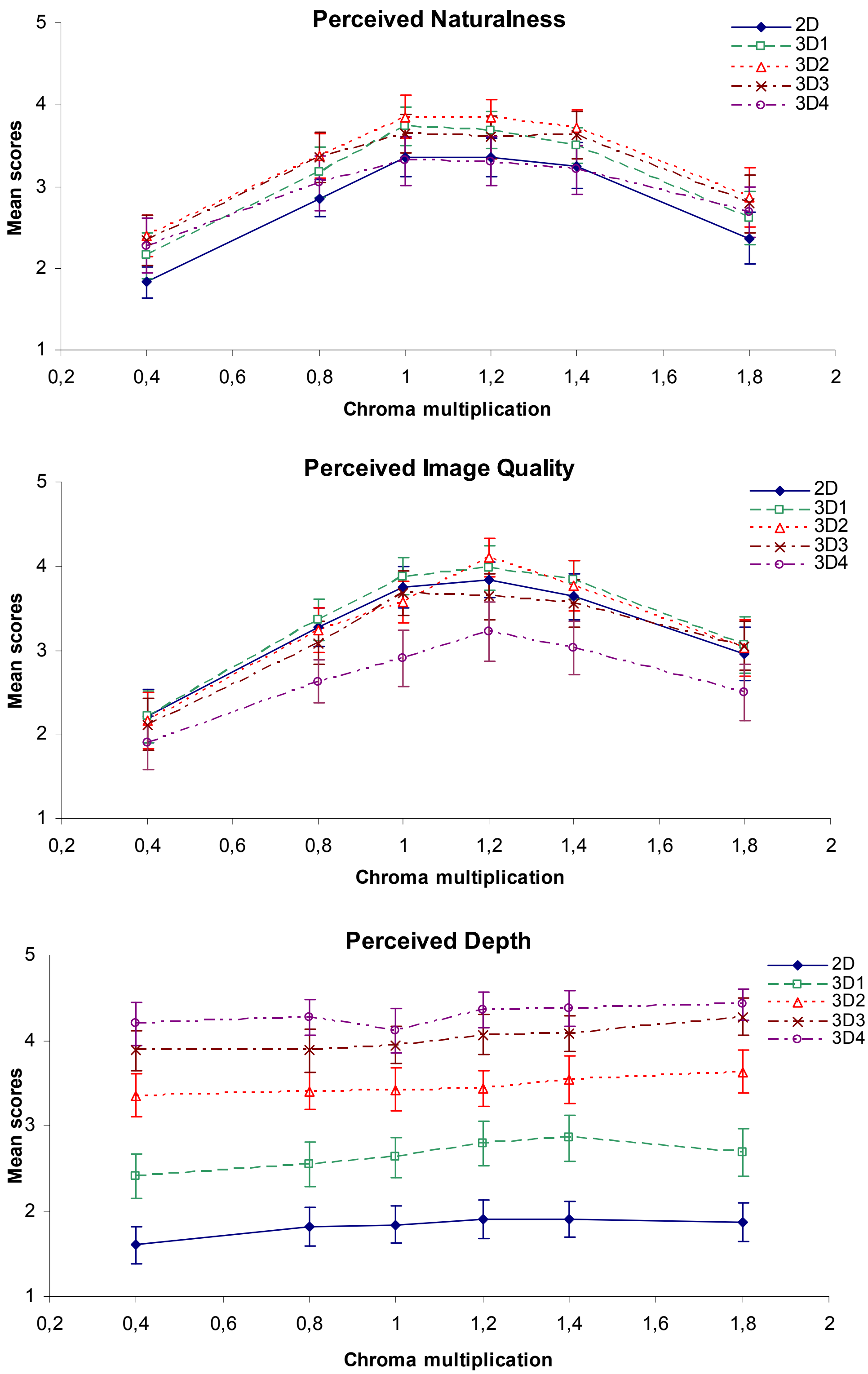

Fig. 4 Mean assessment scores for Naturalness, Image Quality and Depth as function of chroma. The separate lines represent the different dimension levels. Error bars show the 95% confidence intervals.

This is counterintuitive unless the occurrence of artifacts dominates the usefulness of an image. This might have been the case at high disparity for this particular display, but definitely is not an appropriate explanation for low disparities in each of the displays used in the various studies.

Hence, as already suggested earlier, the perception of stereoscopic depth goes beyond image quality.[29] Since naturalness does incorporate both the positive and negative effect of stereoscopic depth,[14,24,25,28] it seems to be a more appropriate attribute for evaluating the experience of stereoscopic images. Research of Seuntiëns (2006)[25] and Kaptein et al. (2008)[24] even suggest that image quality is only one of the attributes, which determines the naturalness of stereoscopic images.

ACKNOWLEDGEMENTS

The work reported here is supported by funding from the EC FP6 MUTED project (Multi-User 3D Television Display).

REFERENCES

1. Hunt, R. W. G., "Chromatic adaptation in image reproduction," Color: Research and Application, 7, 46-49 (1982).

2. Siple, P. and Springer, R. M., "Memory and preference for the colors of objects," Perception & Psychophysics, 34, 363-370 (1983).

3. Engeldrum, P. G., [Psychometric Scaling: a Toolkit for Imaging Systems Development], Imcotek Press, Winchester (2000).

4. Janssen, T. J. W. M. and Blommaert, F. J. J. "Image quality semantics," Journal of Imaging Science and Technology 41(5), 555–560 (1997).

5. Janssen, T. J. W. M. and Blommaert, F. J. J., "A computational approach to image quality," Displays 21(4), 129-142 (2000).

6. Janssen, T. J. W. M. and Blommaert, F. J. J., "Predicting the usefulness and naturalness of colour reproductions," Journal of Imaging Science and Technology, 44(2), 93-104 (2000).

7. de Ridder, H., "Naturalness and image quality: towards perceptually optimal color reproduction of natural scenes," SID international symposium digest of technical papers, 496-499 (2001).

8. de Ridder, H., "Naturalness and image quality: hue, saturation and lightness variation in color images of natural scenes," Proceedings of the 8th Congress of the International Colour Association, 658-661. (1997).

9. de Ridder, H., Blommaert, F. J. J. and Fedorovskaya, E. A., "Naturalness and image quality: Chroma and hue variations in color images of natural scenes," Proc. SPIE 2411, 51–61 (1995).

10. de Ridder, H., "Naturalness and image quality: saturation and lightness variation in color images of natural scenes," Journal of Imaging Science and Technology 40(6), 487–493 (1996).

11. Fedorovskaya, E. A., de Ridder, H. and Blommaert, F. J. J., "Chroma variations and perceived quality of color images of natural scenes," Color: Research and Application 22(2), 96–110 (1997).

12. IJsselsteijn, W.A., de Ridder, H. and Hamberg, R., "Perceptual factors in stereoscopic displays: The effect of filming parameters on perceived quality and reported eye-strain," Proc. SPIE, 3299, 282–291 (1998).

13. IJsselsteijn, W.A., de Ridder, H. and Vliegen, J., “Subjective evaluation of stereoscopic images: Effects of camera parameters and display duration,” IEEE Transactions on Circuits and Systems for Video Technology, 10, 225–233 (2000).

14. Seuntiëns, P. J. H., Heynderickx, I. E. J. and IJsselsteijn, W. A., “Viewing experience and naturalness of 3D images,” Proc. SPIE, 6016, 34-49 (2005).

15. Hubona, G. S., Shirah, G. W. and Fout, D. G., “The effects of motion and stereopsis on three-dimensional visualization, International Journal of Human-Computer Studies,” 47(5), 609-627 (1997).

16. Sollenberger, R. and Milgram, P., “Effects of stereoscopic and rotational displays in a three-dimensional path-tracing task,” Human Factors, 35(3), 483-499 (1993).

17. van Berkel, C. and Clarke, J. A., "Characterization and optimization of 3D-LCD module design," Proc. SPIE 3012, 179-187 (1997).

18. Willemsen, O. H., Zwart, S. T. D. and Hiddink, M. G. H. "2-D/3-D switchable displays," Journal of the Society for Information Display 14, 715-722 (2006).

19. ITU, “Methodology for the subjective assessment of the quality of television pictures,” Recommendation BT.500-10 (2000).

20. Boschman, M. C., "ThurCatD: a tool for analyzing ratings on an ordinal category scale," Behavior Research Methods, Instruments, & Computers, 32(3), 379–388 (2000).

21. Keppel, G., [Design and analysis: A researcher's handbook], Englewood Cliffs, (3rd edition) Prentice-Hall, Engelwood Cliffs NJ, Chapter 6 (1991).

22. Kaptein, R .G. and Heynderickx, I. E. J., "Effect of Crosstalk in Multi-View Autostereoscopic 3D Displays on Perceived Image Quality," SID Symposium Digest of Technical Papers, 38(1), 1220-1223 (2007).

23. Seuntiëns, P. J. H., Meesters, L. M. J. and IJsselsteijn, W. A., “Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation,” ACM Transactions on Applied Perception, 3(2), 95-109 (2006).

24. Kaptein, R. G., Kuijsters, A., Lambooij, M. T. M., IJsselsteijn, W. A., and Heynderickx, I. E. J., “Performance evaluation of 3D-TV systems,” Proc. SPIE, 6808, 680819 (2008).

25. Seuntiëns, P. J. H., “Visual experience of 3D TV,” Ph.D. thesis, Eindhoven University of Technology, The Netherlands. (2006).

26. Stelmach, L. B. and Tam, W. J., “Stereoscopic image coding: effect of disparate image-quality in left- and right-eye views,” Signal Processing: Image Communications, 14, 111–117 (1998).

27. Tam, W. J., Stelmach, L. B. and Corriveau, P. J., “Psychovisual aspects of viewing stereoscopic video sequences,” Proceedings of the SPIE, 3295, 226–235 (1998).

28. Seuntiëns, P. J. H., Meesters, L. M. J. and IJsselsteijn, W. A., “Perceptual attributes of crosstalk in 3D images,” Displays, 26, 177–183 (2005).

29. de Ridder, H., Rozendale, M. C., “Beyond image quality: designing engaging interactions with digital products,” Proc. SPIE, 6806, 68060F (2008).