PARALLAX DISTRIBUTION AND VISUAL COMFORT ON STEREOSCOPIC HDTV

Y. Nojiri, H. Yamanoue, S. Ide, S. Yano and F. Okano

Japan Broadcasting Corporation, Japan

ABSTRACT

Watching stereoscopic HDTV sometimes causes visual discomfort or fatigue. In this paper, we report two results of recent research on the relationship between visual comfort and parallax on stereoscopic HDTV. One is the relationship between visual comfort and parallax distribution in a frame, and the other is the dependence of visual comfort on the magnitude of parallax and motions in some scenes of stereoscopic HDTV programs. Therefore, we measure the magnitude of parallax and motion using the block-matching method, and execute two types of test for evaluating visual comfort.

The analysis results indicate that the features of parallax distribution in a frame are strongly related to visual comfort and for greater comfort when viewing stereoscopic images, the upper part of the screen should be located further away from the viewer with less parallax dispersion, and the entire image should be positioned at the back. Further, the magnitude of parallax, motion and these changes affect visual comfort when watching moving scenes, and large parallax causes discomfort even if the motion is small.

INTRODUCTION

In Japan, digital HDTV has already started, using broadcast satellite. Development work is progressing on cameras[1], recording equipment and display devices for HDTV-based stereoscopic HDTV, and stereoscopic HDTV programs are presented at special events and at some permanent public facilities. Many stereoscopic HDTV programs have already been produced, including materials shot at the Salt Lake City Winter Olympics and the Sydney 2000 Olympic Games. Compared with conventional 2-D HDTV programs, these stereoscopic programs provide a powerful sense of presence and immersion, but prolonged viewing can cause fatigue and some scenes are uncomfortable to watch[2] [3]. If stereoscopic HDTV is to be widely accepted as a broadcasting method, the images must be comfortable to view and cause minimal visual fatigue.

NHK has therefore been combining its production of trial programs with research into the causes of fatigue and what it is that makes scenes comfortable or uncomfortable to watch. The underlying causes of visual fatigue and discomfort are thought to be the following:

(1) Geometric distortion in left and right images in the shooting or display system,
(2) Crosstalk or differences in the electrical properties of both systems for left and right images, and
(3) Contradictions of the convergence and accommodation based on binocular parallax.

Recently, we have been studying factor (3), focusing on the dependence of visual comfort and visual fatigue on the image properties, that is, distribution and magnitude of parallax and motion[4]. In this report, we discuss the results of research on two aspects of visual comfort relative to parallax distribution. One is research on the sense of presence and visual comfort relative to parallax distribution for almost still images, and the other is research on the dependence of visual comfort on the amounts of parallax and motion for some programs with moving images.

RELATIONSHIP BETWEEN “SENSE OF PRESENCE” AND “VISUAL COMFORT”, AND PARALLAX DISTRIBUTION

We conducted a subjective evaluation test using 10 stereoscopic HDTV images, measured the parallax distribution for the images used, and analyzed the correlations between the parallax distribution and the results of subjective evaluation.

Evaluated Images

As images to be evaluated, we used 10 scenes included in the standard stereoscopic chart[5)] in which factors such as the shooting conditions and the positional relationships of the objects are known. Each of the scenes used for evaluating images was 15 seconds long, and did not include any large movement.

Evaluated images	10 scenes (10 stereoscopic images and their 2-D images)
Subjects	99 (ages ranging from teens to 60's)
Presentation	Twice repetition of 2D-image and 3D-image
Display system	Stereoscopic HDTV projection system with polarizing glasses
Screen size	120-inch / 16:9
Peak brightness	21.4 cd/m^2
Viewing distance	4.5 m

Table 1 Outline of Experiment 1

Evaluation Test

We conducted a subjective evaluation test using a total of 20 images, consisting of the 10 scenes mentioned above and an additional 2-D image from each scene. Table 1 shows the outline of experiment 1. The display system used was a stereoscopic HDTV projection system with polarizing glasses and screen size of 120 inches. In the test, first the 2-D image was presented as a reference, then the stereoscopic image was presented, and then the process was repeated twice for each image. In the evaluation of the 2-D images, the left-hand image was input to both left and right projectors so that the images could be evaluated while the subjects were still wearing the polarizing glasses.

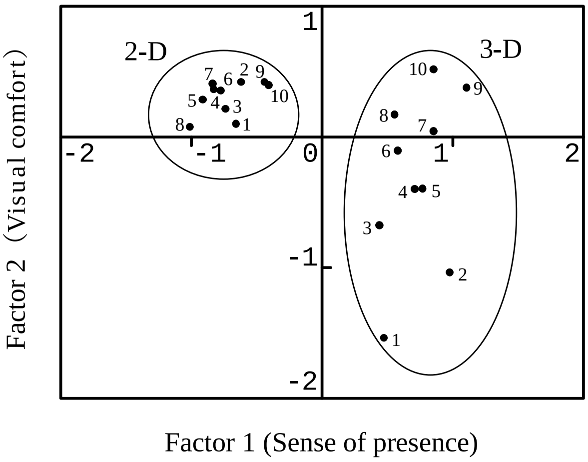


Figure 1 Result of Evaluation Test

Experimental Results

Figure 1 shows the distribution of each of the evaluation images with regard to “sense of presence” and “visual comfort” which are the evaluation terms extracted by the factor analysis. The numbers in the figure 1 refer to the 10 images. The figure clearly shows that, when compared with the 2-D images, the stereoscopic (3-D) images had a strong sense of presence, and that some of the images were comfortable to view but some were not.

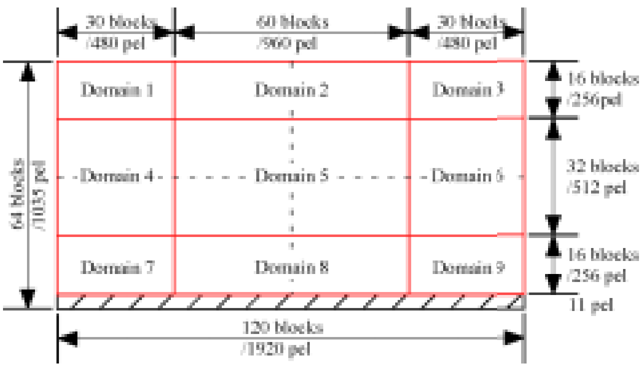


Figure 2 Screen Block Division

Domain 1 $a_1 = 0.85$	Domain 2 $a_2 = 0.89$	Domain 3 $a_3 = 0.87$
Domain 4 $a_4 = 0.93$	Domain 5 $a_5 = 0.97$	Domain 6 $a_6 = 0.96$
Domain 7 $a_7 = 0.86$	Domain 8 $a_8 = 0.93$	Domain 9 $a_9 = 0.92$

Figure 3 Primary Principal Component Loading

Domain 1 $b_1 = 0.24$	Domain 2 $b_2 = 0.37$	Domain 3 $b_3 = 0.46$
Domain 4 $b_4 = -0.15$	Domain 5 $b_5 = 0.06$	Domain 6 $b_6 = 0.12$
Domain 7 $b_7 = -0.41$	Domain 8 $b_8 = -0.34$	Domain 9 $b_9 = -0.33$

Figure 4 Secondary Principal Component Loading

Measurement and Analysis of Parallax Distribution

Next, we focused on parallax distribution as a feature of the scene images and analyzed its correlation to sense of presence and visual comfort. Because the scenes used here do not contain large amounts of motion, we measured and analyzed the parallax distribution in the first frame as a representative sample of the distribution for each scene.

To measure parallax, we used the block matching method on the left and right images. The block size used was 16 pixels x 16 lines. Then, for parallax distribution analysis, the screen was also divided into nine domains, as shown in Figure 2. We assumed that attention would tend to be focused towards the center of the screen and so made that domain larger. The parallax in each screen domain was taken to be the average of the parallax measured in each measurement block in the screen. In our analysis, the parallax value on the screen itself is taken as 0, negative values indicate that the image appears to be in front of the screen, and positive values that the image appears to be beyond the screen. Principal component analysis for parallax in the nine screen domains showed that two principal components were detected. Figure 3 shows the loading of the primary principal component and Figure 4 shows the loading of the secondary principal component.

Results of principal component analysis

The primary principal component score in each image is as follows:

$\Sigma[a_i \times$ (parallax in each domain)]

a_i are primary principal component loadings.

Because all the results for principal component loading a_i are positive and roughly equivalent (between 0.85 and 0.97), as Figure 3 shows, the greater the parallax in the positive direction, the higher the scores for the primary principal component. Accordingly, the results show that the first primary component tends to increase when the overall image is further toward the back.

The secondary principal component score in each image is as follows:

$\Sigma[b_i \times$ (parallax in each domain)]

b_i are secondary principal component loadings.

Figure 4 shows that b_i is larger for domains at the top of the screen, and also that b_i is positive for domains at the top of the screen and negative for domains at the bottom. As a result, we can see that images tend to score higher when they have a depth distribution (parallax distribution) such that the top of the image is beyond the screen and the bottom of the image is in front of the screen.

Distribution of parallax between domains (dispersion)

The dispersion of the amount of parallax in each of the domains shown in Figure 2 was also analyzed as one of the parallax distribution characteristics.

Relationship Between Subjective Evaluation Results and Parallax Distribution

We analyzed the correlation between the two factors obtained from the subjective evaluation test and the characteristics of images, namely the parallax dispersion and the scores for the primary and secondary principal components obtained by principal component analysis of the parallax distribution for each image. The results show that there is no clear relationship between "sense of presence" and the parallax distribution, that is, primary and secondary principal components and parallax dispersion. On the other hand, there was a clear correlation between "visual comfort" and these factors, for which we obtained the following regression equation:

$$\begin{aligned}\text{Factor (visual comfort)} &= 0.36 \times \text{(primary principal component score)} \\ &\quad + 1.00 \times \text{(secondary principal component score)} \\ &\quad - 0.90 \times \text{(parallax dispersion)} \\ &\quad + \text{(constant)}\end{aligned}$$

From this equation, we can see that stereoscopic images are more comfortable to watch when the scores for the primary and secondary principal components are higher and parallax dispersion is lower. In other words, this suggests that stereoscopic images are more comfortable to watch when the bottom of the image appears in front of the screen, the top of the image appears beyond the screen, and there is little or no variation in the parallax over the image as a whole.

DEPENDENCE OF VISUAL COMFORT ON MAGNITUDE OF MOTION AND PARALLAX

In the previous sections, we discussed the relationship between the visual comfort and the parallax distribution within the image for short scenes lasting 15 seconds without large motion. In this section, we discuss the results of continuous subjective evaluation tests of the visual comfort relative to the amount of motion and amount of parallax for relatively long programs with motion.

Evaluated images	Waffen (15 min): Toed-in camera configuration Africa (10 min): Parallel camera configuration
Subjects	5
Evaluation method	Single stimulus continuous quality evaluation
Display system	Stereoscopic HDTV projection system with polarizing glasses
Screen size	120-inch / 16:9
Peak brightness	20 cd/m^2
Viewing distance	4.5 m

Table 2 Outline of Experiment 2

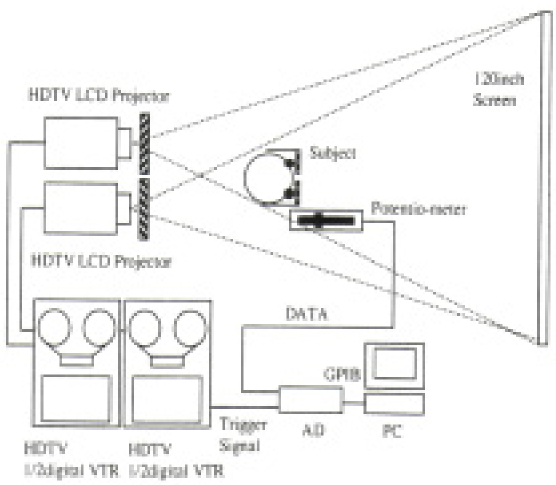


Figure 5 Experimental Method

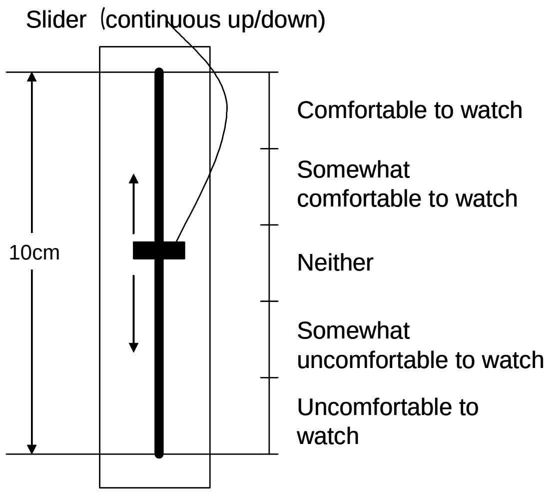


Figure 6 Potentiometer

Evaluated Images

We used two stereoscopic HDTV programs: "Waffen" and "Africa." "Waffen" is a dramatized short story shot using two cameras in a "toed-in" arrangement such that their optical axes intersected. In each scene, the distance between the cameras and their axial intersection points varied. "Africa" is a documentary-style program about animals in the wild and was shot using two cameras set permanently 65 mm apart in a parallel camera configuration. "Waffen" is 15 minutes long, "Africa" is 10 minutes long.

Evaluation Test

Table 2 summarizes the conditions of experiment 2 and Figure 5 shows the experimental method. As in the previous experiment, when assessing the 2-D images, left-eye images were shown to both eyes so that the viewers could evaluate the 2-D images while still wearing polarizing glasses.

To provide continuous evaluation of the visual comfort for the 2-D and 3-D (stereoscopic) images, we conducted single-stimulus continuous quality evaluation tests. We provided a potentiometer to allow the subjects to continuously input their evaluation scores. Voltage changes in the potentiometer were run through an A/D converter and recorded on a PC. The potentiometers were as shown in Figure 6, with a movement distance of 10 cm divided into 5 sections of 2 cm each. The data downloading frequency was set at 0.5 Hz. The use of single-stimulus continuous quality evaluation allowed the subjects to focus on the evaluation images even over extended evaluation tests.

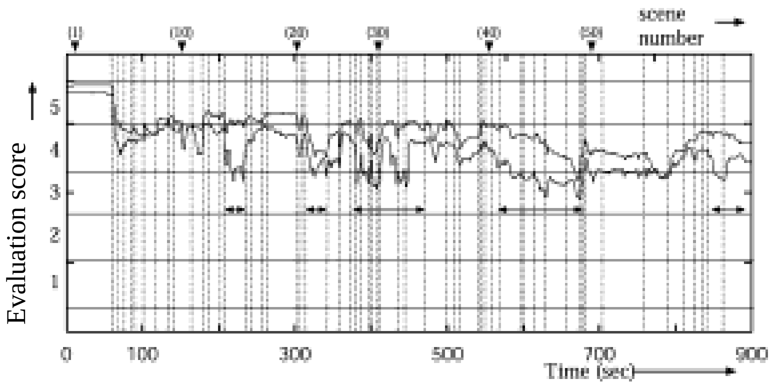


Image: Stereoscopic & 2D image, upperline 2D and lower line 3D
Vertical scale (Visual comfort): 5=good, 1=bad

Figure 7 Result of single stimulus continuous quality evaluation (SSCQE) for "Stereoscopic & 2D image: Waffen"

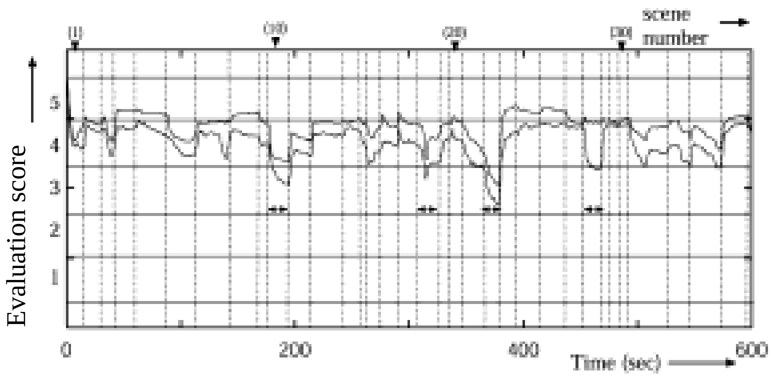


Image: Stereoscopic & 2D image, upperline 2D and lower line 3D
Vertical scale(Visual comfort): 5=good, 1=bad

Figure 8 Result of single stimulus continuous quality evaluation (SSCQE) for "Stereoscopic & 2D image: Africa"

Evaluation Results

The evaluation results are shown in Figures 7 and 8. The results in the figures are simple averages of the subject responses. Figure 7 shows the evaluation results for the 3-D and 2-D images in "Waffen." The lower horizontal axis shows the time, and the upper horizontal axis shows the scene number. The vertical axis shows the subjective evaluation scores. The dotted vertical lines in the figure indicate the timing of the scene changes. In the case of "Waffen," there are 58 scene changes in total. The line in the figure with high evaluation scores indicates those for 2-D images, while the low evaluation scores indicate 3-D images. Figure 8 shows the evaluation results for "Africa" the same as in Figure 7. There are 35 scene changes in "Africa." In this case also, the evaluation scores were higher for the 2-D images than for the 3-D images. The pairs of arrows in the figure indicate examples of scenes where the evaluation scores for 2-D images are markedly higher than those for 3-D images.

Measurement of the Amounts of Parallax and Motion

The block matching method was used to measure the amounts of parallax and motion. However, because the HDTV movies used for the evaluation contain a large amount of data compared with previous measurements, the block matching method generates a huge calculation load. Consequently, we used an HDTV-to-NTSC down-converter to convert to 720 pixels x 480 lines of RGB signals per frame and then downloaded the data to a computer and converted it to Y signals. We then used the block matching method to calculate the amount of parallax between the left and right images and the amount of motion between adjacent frames for the left image. In each case, the block size was set to 8 pixels x 8 lines. Next, we used the amount of parallax and the amount of motion measured for each block to define the values which represent the magnitude of the parallax and motion in the frame, as shown below.

Number of blocks in frame N with a measured parallax of k: S(k)
Threshold values for amount of parallax: $\alpha 1$, $\alpha 2$

$$I(N) = \sum_{\alpha 1 < |k| < \alpha 2} (k \times S(k)) \,/\, \sum_{0 \leqq |k| \leqq \alpha 1} (k \times S(k))$$

I(N) defines the magnitude of parallax in frame N.
However, in this case $\alpha 1 = 3$ and $\alpha 2 = 50$.

Similarly, the amount of motion in the frame was defined as follows:

Number of blocks in frame N with the amount of motion m: P(m)
However, because the motion in 2-D frames is only in two dimensions, m was set to an absolute value for the vector sum of the horizontal and vertical motions.
Threshold values for amount of motion: $\beta 1$, $\beta 2$

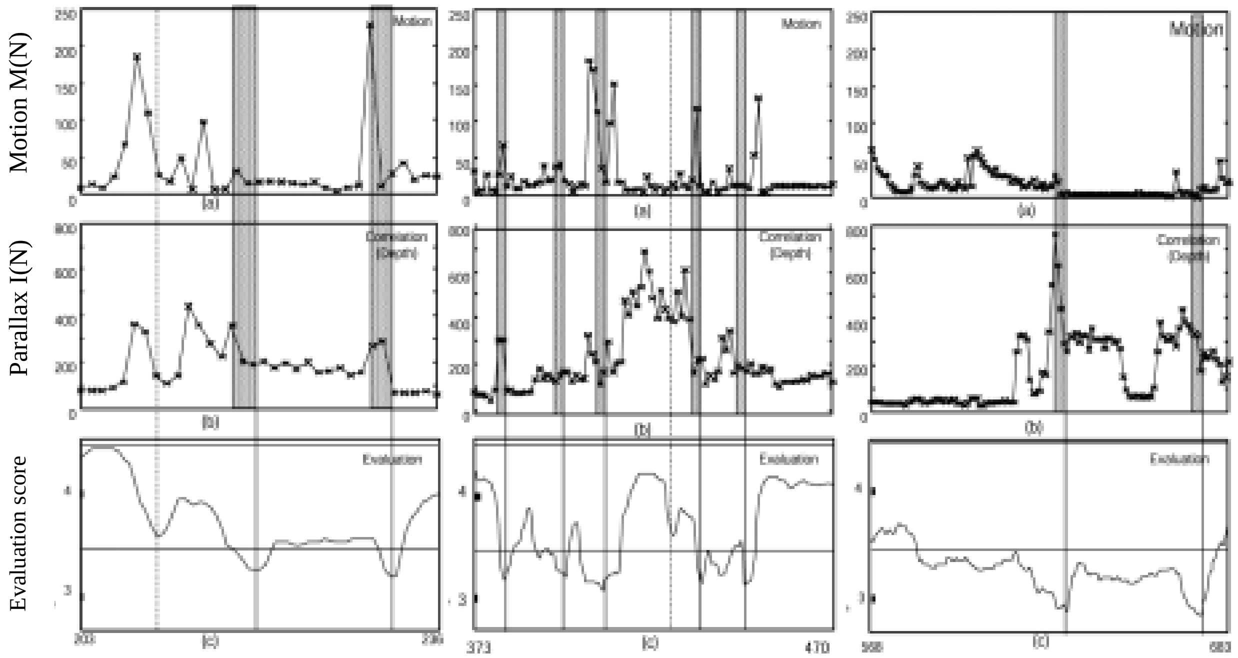
$$M(N) = \sum_{\beta 1 < |m| < \beta 2} (m \times P(m)) \,/\, \sum_{0 \leqq |m| <\leqq \beta 1} (m \times P(m))$$

M(N) defines the magnitude of motion in frame N.
However, in this case, $\beta 1 = 3$ and $\beta 2 = \mathrm{SQRT}(20^2 + 10^2)$.

From the above definitions, I(N) is larger where the parallax in the frame is greater, in other words, where the displayed image appears further in front of or further beyond the screen. Similarly, M(N) is larger where the motion in the image is larger.

Discussion

Figures 9 and 10 show the results of some of the measurements taken during the program for the amounts of parallax and motion using the definitions in the previous section. The scenes in the figure are the scenes indicated by arrows in Figures 7 and 8. Figure 9 (1 to 3) shows the results for "Waffen" and Figure 10 (1 to 3) those for "Africa." The top level (a) in Figures 9 and 10 shows the chronological changes in motion for each second, level (b) shows the chronological changes in the amount of parallax, and level (c) shows subjective evaluation results. Areas with localized low evaluations in the evaluation results are marked by solid lines and compared with the amounts of motion and parallax. Cross-hatching is used to indicate two-second sections that precede locations with low evaluations. Also, where areas of localized decline continued over time, we assumed that the portions that were latest and had the lowest evaluations were areas where the evaluations were stable. In Figures 9 and 10, the features of locations where the evaluation results are locally small are as follows:

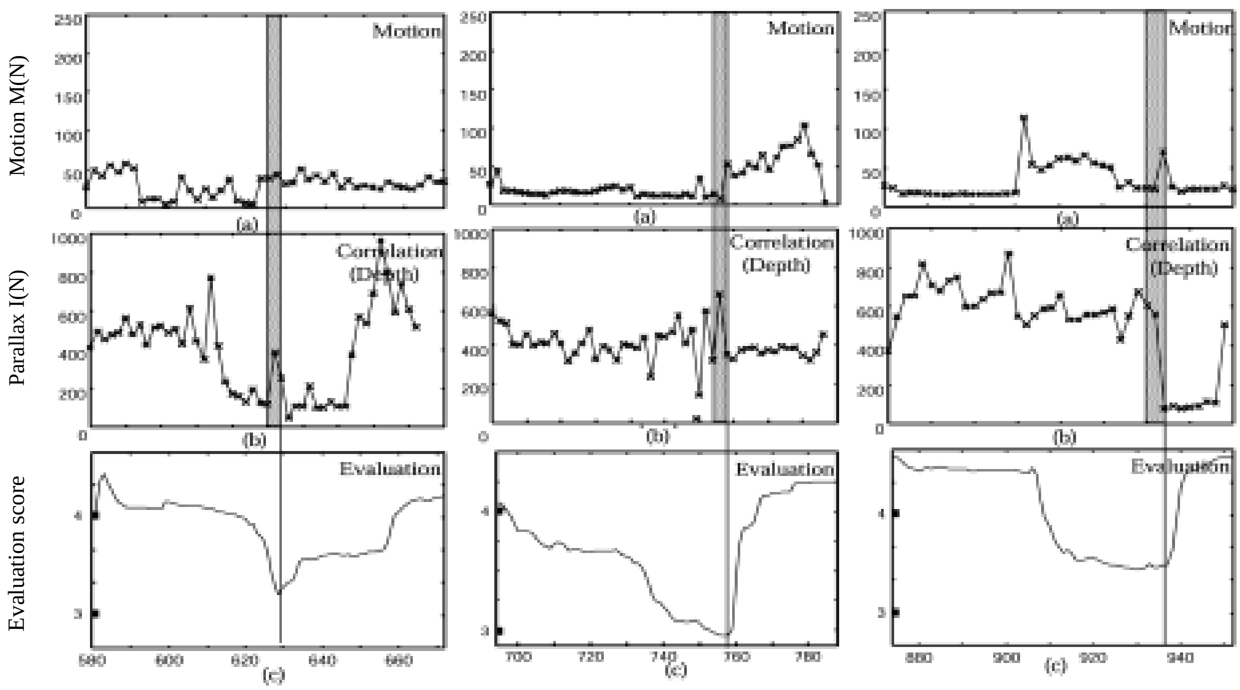


Time [sec] → (1) scene 14 to 18

Time [sec] → (2) scene 25 to 33

Time [sec] → (3) scene 42 to 49

Figure 9 Relationship among evaluation, motion and parallax: “Waffen”



Time [sec] → (1) scene 17 to 19

Time [sec] → (2) scene 21 to 23

Time [sec] → (3) scene 26 to 28

Figure 10 Relationship among evaluation, motion and parallax: “Africa”

* Locations with locally low evaluations with scores of 3 or less
  - Figure 9 (3): Motion is small and parallax and its variations are large.
  - Figure 10 (2): Motion is small and parallax and its variations are large.
* Locations with locally low evaluations but with scores higher than 3
  - Figure 9 (1): Motion is small and parallax and its variations are relatively large.
  - Figure 9 (2): The variations in motion are large, and parallax and its variations are relatively large.
  - Figure 10 (1): Motion is large, and parallax and its variations are relatively large.
  - Figure 10 (3): Motion is large and the variations in parallax are large.

In the methodology for the subjective assessment used here, it is difficult to comment on the absolute accuracy of the evaluation results, and so we must focus on relative changes in the evaluations. Accordingly, from the above results we consider that the evaluations become lower when the parallax is relatively large or the variation in parallax is relatively large.

CONCLUSION

We studied sense of presence and visual comfort for stereoscopic HDTV images based on parallax distribution. In the study, we performed subjective evaluation tests using relatively short scenes of HDTV and stereoscopic HDTV images. The results showed that there was no correlation between sense of presence and parallax distribution in the screen. However, there was a clear correlation between parallax distribution and visual comfort, suggesting that images were more comfortable to watch when the bottom of the image appeared closer and the top of the image appeared further away.

We then used a single-stimulus continuous quality evaluation to evaluate HDTV and stereoscopic HDTV programs with a relatively long duration, and studied the relationship between the evaluations, and the amounts of parallax and motion in the programs. The results showed that scenes with low evaluations had large amount of parallax or large variations in this amount.

In terms of parallax distribution and visual comfort, further study of these factors in combination with visual functions in stereoscopic viewing is needed for scenes where there is little motion but large amounts of parallax and where visual comfort deteriorates, and for scenes that are evaluated highly despite a certain amount of parallax and motion.

REFERENCES

1) H. Yamanoue, F. Okano: Hivision Stereoscopic Harp Camera, NHK R&D, 31, 1994, pp. 58 to 67.
2) M. Wopkin: Visual comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth focus, J. of SID, 3, 1995, pp. 101 to 103.
3) J. Konrad: Enhancement of viewer comfort in stereoscopic viewing: parallax adjustment, Proc. of 1999 SPIE, pp. 179 to 190.
4) S. Ide, H. Yamanoue, M. Okui, F. Okano, M. Bitou and N. Terashima: Parallax distribution for ease of viewing in stereoscopic HDTV, Proc. of 2002 SPIE, pp. 38 to 45.
5) M. Okui, H. Yamanoue, M. Emoto, S. Yano and T. Yoshida: Test materials for evaluating stereoscopic television systems, Proc. of 1999 IBC, pp. 565 to 570.