EFFECT OF CONTENT ON VISUAL COMFORT IN VIEWING STEREOSCOPIC VIDEOS

Bo Chang*, Fuzheng Yang#, Shuai Wan†
* State key laboratory of ISN, Xidian University.
E-mail: bchang_xidian@163.com Tel: +86-29-88202527
State key laboratory of ISN, Xidian University.
E-mail: fzhyang@ mail.xidian.edu.cn Tel: +86-29-88202527
† School of Electronics and information, Northwestern Polytechnical University.
E-mail: swan@nwpu.edu.cn Tel: +86-29-88431253

Abstract—**Visual discomfort is a serious challenge for stereoscopic videos to become prevalent. However, the assessment of stereoscopic visual discomfort is a complicated issue. In this paper, we have studied four of the major factors which affect visual comfort by conducting extensive subjective assessments. Namely, foreground disparity, size of foreground object, disparity distribution and in-depth motion. Relationships between visual comfort and these four factors are analyzed and four conclusions are drawn according to the experimental results. Firstly, when the influences of disparity magnitude and range of disparity function together and clash, the latter is in the dominant position. Secondly, the degree of visual comfort increases with the increment of the object size in a certain range and then levels off or decreases slightly after a certain threshold. Thirdly, better visual comfort is always obtained when the bottom part was perceived nearer to the viewers than its upper part. Last but not least, the variation of parallax over time might be one of the major factors which affect visual comfort. And its influence is complicated and can be reflected by the object's velocity, motion direction and starting plane.**

I. INTRODUCTION

Recent years, there is an increasing interest in three-dimensional (3-D) images or videos since they are close to the natural way of people to perceive the world. However, feelings of visual fatigue and discomfort are always reported after viewing 3-D content, which significantly disturbs the viewer's experience. The concern about the safety of viewing stereoscopic or 3-D video has been under the spotlight again. And it is believed to be one of the obstacles to the wide spread of 3-D content services. To eliminate the obstacles, a lot of time and efforts have been devoted to the production and post-processing of 3-D content in commercial applications. As a result, the automatic prediction of visual discomfort in viewing stereoscopic 3-D contents has increasingly gained research attention [1], [2].

The underlying causes of visual discomfort are usually known as accommodation-vergence conflict, disparity distribution, cognitive inconsistencies, binocular mismatch, depth cue collision and so on [3]-[5]. It is generally acknowledged that an important factor causing visual discomfort is the accommodation- vergence conflict created by the current type of stereoscopic displays. Accommodation and vergence are normally yoked when viewing objects in a natural scene. However, the normal interaction between these two processes can be disrupted when viewing stereoscopic images [6], [7]. Accommodation is directed at images of objects at the screen distance whereas vergence is directed at the perceived distances of objects. Accommodation-vergence conflict is indeed a traditional issue and has been extensively studied by many researchers. In particular, excessive binocular disparity leads to a high degree of conflict between vergence and accommodation, which increases the stress of the human visual system. To minimize the accommodation-vergence conflict, the disparities in a stereoscopic image should be small enough so that the perceived depths of objects fall within a "comfort zone" which is generally defined as depth of field of the eye.

Not only its magnitude, the type and distribution of the disparity over space and time also seem to affect visual comfort [8]. Nojiri et al. analyze the relationship between the parallax contained in stereoscopic still images and visual comfort [9]. The results show that visual comfort is highly correlated to the overall range and distribution of the parallax. These results are corroborated in a later study which examined the effect of parallax distribution on visual comfort of stereoscopic HD video sequences [10]. The results indicate that stereoscopic scenes are more comfortable to watch when the parallax distribution is such that the bottom portion of the image appeared closer and the top portion of the image appeared farther away.

Specially, authors of [11] reported that visual comfort may also be affected by the object width in stereoscopic 3-D scene contents. In literature, it has been known that the width of visual stimulus substantially influences the binocular fusion limit [12]. On the other hand, the stimulus width is also known to influence the perception of crosstalk [13], [14].

Apparently, visual comfort assessment is a complicated issue. However, previous works usually use only the disparity magnitude as an essential parameter for the prediction of visual discomfort in viewing stereoscopic 3-D contents, such as mean and variance of the disparity magnitude measured in an entire image [1], [10], [15]. Apart from the disparity magnitude, there are actually other factors related to subjective visual discomfort.

In this paper, we are devoted to the influence of four of the key factors which affect visual comfort in viewing 3-D videos through doing subjective assessments. In traditional work, natural images or videos shot under certain conditions are used as testing data. Sequences obtained this way cannot be precisely described in terms of the factors related to visual discomfort, and the factors may cause cross-coupling effects. In our work, the stereoscopic stimuli are generated by the MATLAB psychtoolbox and only one factor is tested at a time, so that the effect of the single factor can be evaluated. Four factors are considered, i.e., foreground disparity, size of foreground object, disparity distribution and in-depth motion. Through this work, we provide a better understanding of how different factors influences visual discomfort. The corresponding results can be used for stereoscopic image quality assessment, stereoscopic image coding and etc.

The rest of this paper is organized as follows: Section II presents a detailed description of the experiments and related results. The corresponding analysis and conclusions are also given in the same section. Overall concluding remarks are drawn in Section III.

II. EXPERIMENTAL SETUP, RESULTS AND ANALYSIS

In our work, the influences of spatial and temporal characteristics of the disparity on visual comfort are studied by doing extensive subjective assessments. Twenty-five observers, including twelve males and thirteen females, were recruited in the experiments. All of the observers are postgraduate students whose ages vary from 20 to 26 with an average of 23 years. They are not directly concerned with 3-D quality as part of their normal work, and are not experienced assessors. Prior to the session, the observers were screened for normal vision tests, and they have passed test VT-01, VT-02, VT-04, and VT-07 [16].

An ASUS VG 278 active shutter 3-D display with 120 Hz time-sequential and a pair of NVIDIA active shutter 3-D vision wireless glasses were used. The display is 27 inches in size with a resolution of 1920×1080. The subjects' viewing distance was 3H to 3.2H, where H donates the screen height. The assessment was conducted in a dark room, with minimal light levels.

To avoid the influence of other factors on visual discomfort, we used computer-generated stereoscopic sequences for precise control. The stereoscopic sequences consist of a left-view and a right-view image which were generated by the MATLAB psychtoolbox [17], [18]. The resolution of all stereoscopic sequences was 800×450. During the test, the sequences were displayed with the re-scaled size to fit the size of the display.

Observers subjectively evaluated the visual comfort of the stereoscopic sequences on a five–point scale as follows.

1 Extremely uncomfortable.
2 Uncomfortable.
3 Mildly uncomfortable.
4 Comfortable.
5 Very comfortable

Note that to obtain more reliable and stable assessment results, the presentation order of the stimuli was randomized. And each sequence was just displayed for once to avoid long time watching which may induce visual fatigue unexpectedly. The experimental schedule is described in Fig. 1, and Fig. 2 shows the experimental apparatus. During testing, any opinion regarding visual comfort was encouraged.

A. Foreground disparity and range of disparity

To figure out the influence of foreground disparity on visual discomfort, two groups of test sequences were generated by the MATLAB psychtoolbox. Sequences in one group contained both foreground and background. The disparity of the background was set as constant, and there were 7 levels of screen disparity of the foreground: 20, 10, 0, -10,-20,-30,-40 where the unit of the disparity was pixel-based. A resembling Maltese cross with 80×80 pixels was used as the foreground object since it contains both high and low spatial frequency components [19]. The resembling Maltese cross moved along a trajectory which was a circle with center point at the center of the screen, with a radius of 240 pixels. The motion direction was anti-clockwise with the velocity of 36 degrees/s. The reason of using a circle as the trajectory is that it can avoid the step impulse that comes from a sudden change of the motion direction, which may cause unexplained effects of visual discomfort. Another group of sequences was similar to the first one, but there was no background in each sequence. The purpose of this design is to compare the influence of the foreground disparity and that of the disparity range.

As there were 2 groups of sequences and 7 levels of screen disparity, there would be totally 14 stimuli for the experiment. An example of the stimuli is shown in Fig. 3, in which the foreground object is placed in front of the screen with a screen disparity of 10 pixels.

The statistical analysis applied on the subjective data was one-way ANOVA where differences between groups were caused by the level of foreground disparity. The test results of

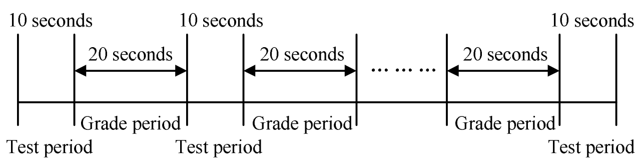

Fig. 1 Experiment schedule

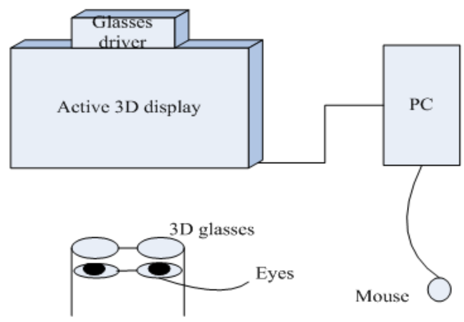

Fig. 2 Experimental apparatus

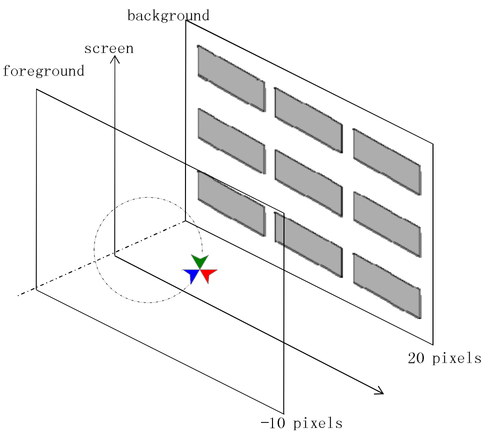

Fig. 3 An example of stimuli used in experiment A.

the ANOVA showed that the population means are significantly different (F= 12.71856, P=0.0001) at the 0.05 level no background was invovled. Namely, disparity magnitude affects visual comfort a lot.

Figure 4 plots the mean and confidence interval of visual comfort scores versus disparity magnitude. The experimental results show that the mean opinion score (MOS) of the test sequence is higher when the foreground is closer to the screen plane when there is no background. A possible reason for this phenomenon is that images with larger parallaxes are more difficult to fuse. However, when background is put in the scene, result is not in accordance with this trend. When background is involved, the same level of screen disparity of the foreground leads to a larger range of the overall disparity. The bigger the range is, the more discomfort we feel. So when the influences of the foreground disparity and the range of disparity function together and clash, the latter is in the dominant position.

For the comfort consideration, the disparity should be as small as possible. However, a limited range of disparity is tied to a weak presence of the stereoscopic scene. Thus it can be seen that visual comfort and pursuit of stereoscopic sense is contradictory. Therefore, 3-D service providers or program producers should find a balance between the two aspects.

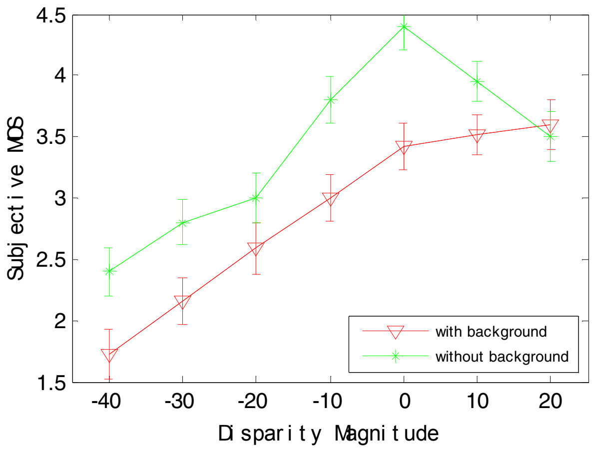

Fig. 4 Relationship between the MOS and disparity magnitude with and without background. Vertical bars denote the confidence interval at the 95% confidence level.

B. *Object size*

To investigate the relationship between the object size and visual comfort, 12 sequences generated by psychotoolbox were employed in this assessment. Here each sequence contained a foreground and background. A resembling Maltese cross with 4 different sizes was used as the foreground object: 40×40, 80×80, 120×120 and 160×160. The disparity of the background remained constant of 20 pixels. Three levels of foreground disparity were set. The resembling Maltese cross moved along a trajectory which was a circle with center point at the center of the screen, and a radius of 240 pixels. The motion direction was anti-clockwise with the velocity of 36 degrees/s. Fig. 5 shows an example of the stimuli, in which the sizes of the foreground object are 40×40 and 80×80 respectively and the screen disparity is -20 pixels.

Similarly, one way-ANOVA was also performed in this section where differences between groups were caused by the factor of distinct object sizes. Test results are depicted in Table I.

Apparently, when foreground disparity of scene objects were farther away from the screen plane (e.g., -50 and -20), the population means are significantly different at the level of 0.05. However, difference within the group became significant while disparity magnitude was relative small (e.g.10). This is because when disparity magnitude was not very large, disparity range in sight remains to be small values with the variation of object size and visual comfort is always quite high.

Fig. 6 shows the scatter plot of the size of the foreground object versus MOS with 3 distinct screen disparity values, where the vertical bars denote the confidence interval at the 95% confidence level. For a small foreground disparity (e.g., 10 and -20), visual comfort increases rapidly with the increment of the object size when it is smaller than 80×80, but levels off or even slightly decreases when size of the foreground object is bigger than 80×80. One possible reason for this phenomenon is depicted as follow. When the object

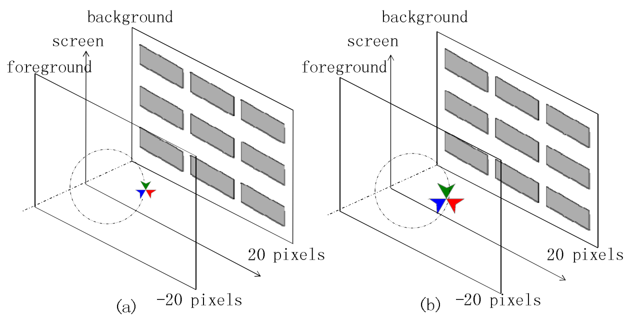

Fig. 5 Examples of stimuli used in experiment B.
(a)Object size 40×40 (b) Object size 80×80

TABLE I
RESULTS OF ANOVA TEST OF THE SUBJECTIVE DATA IN EXPERIMENT B

Foreground Disparity	F Value	P Value
-50	4.77337	0.00576
-20	5.64755	0.00231
10	3.41089	0.04958

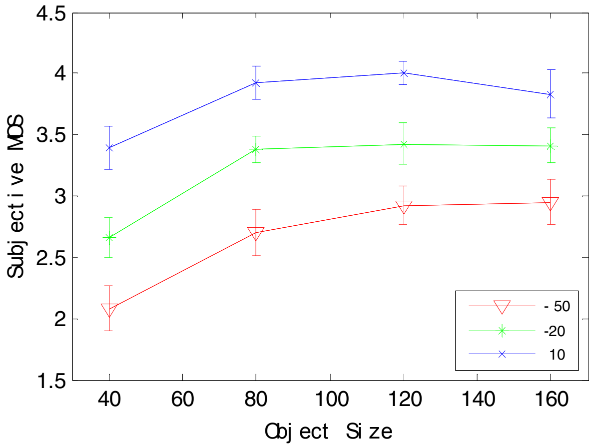

Fig. 6 Relationship between the MOS and size of the foreground object. Vertical bars denote the confidence interval at the 95% confidence level.

size is rather small, the foreground object covers only partially the focus-of-attention region of the eyes. Other part of the scene may also fall into that region, which enlarges the overall range of disparity for the eyes to focus and results in degradation of visual comfort. When the size of the foreground object is big enough to fill the focus-of-attention region of the eyes, visual comfort levels out and an increment of the object size will not improve visual comfort. On the other hand, the MOS of visual comfort decreases slightly when the object size is too big when it is moving at the plane with a small depth. In this case the subjects have to try hard and change eye focuses to catch sight of the whole foreground object, which results in degradation of visual comfort. Thus we can get the following conclusions: subjective visual comfort increases with the enlargement of the object size until it reaches a certain threshold. Then visual comfort levels off or decreases slightly with the increment of the object size.

However, for a large foreground disparity (-50), the degree of visual comfort seems to increase with the object size in a moderate way. This may be explained by the principle of stereoscopic imaging as shown in Fig. 7. When the image is perceived in front of the screen as shown in Fig. 7 (a), the size of the perceived image is smaller than that of the image displayed on a single view. Nevertheless, this size is larger than that of the original one when it is imaging behind the screen. So if the disparity of one object is crossed and quite large, what we perceive is actually a smaller version of the object. Hence, the threshold mentioned above becomes larger.

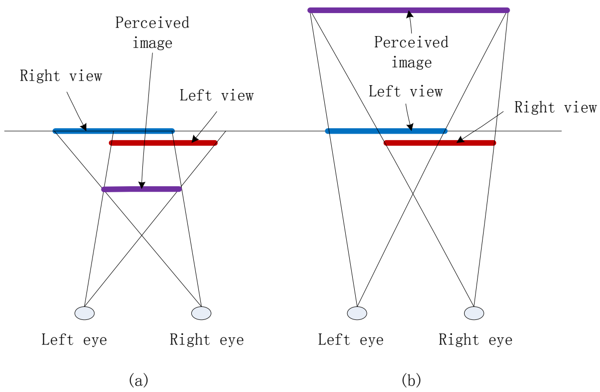

Fig. 7 Principle of stereoscopic imaging. (a)Image perceived in front of the screen. (b)Image perceived behind the screen

C. Disparity distribution

Apart from the disparity magnitude, stereoscopic visual comfort can also be affected by the disparity distribution. In this experiment, 13 sequences which represent 13 disparity distribution modes generated by psychtoolbox were involved. The whole image was divided into three parts: the top, the middle and the bottom. The top and the bottom were made up by several rectangular and kept still. A resembling Maltese cross moved along a trajectory with the center point at the center of the screen, and the radius of 200 pixels was regarded as the middle part. Relative depth positions of these three parts varied among different modes. Disparity distribution mode was named by the distance of the three parts to the viewers with the order of top-middle-bottom. For example, mode *3-2-1* means the top portion stays furthest to the viewers, the bottom closest and the middle part was between the two. Two of the thirteen sequences are shown in Fig. 8 as examples, where (b) and (d) designate the side views of (a) and (c), respectively. Detailed information of different modes is depicted in Table II.

Also, one-way ANOVA test was performed to validate the reliability of the subjective test data. The F value and the P value were respectively 6.89932 and 0.00004 which reveal the fact that the data was reliable and disparity distribution mode affects visual comfort a lot.

Subjective assessment results of this experiment are summarized in Fig. 9, which indicates that the best four disparity distributions are mode *1_1_1, 3_2_1, 2_2_1* and

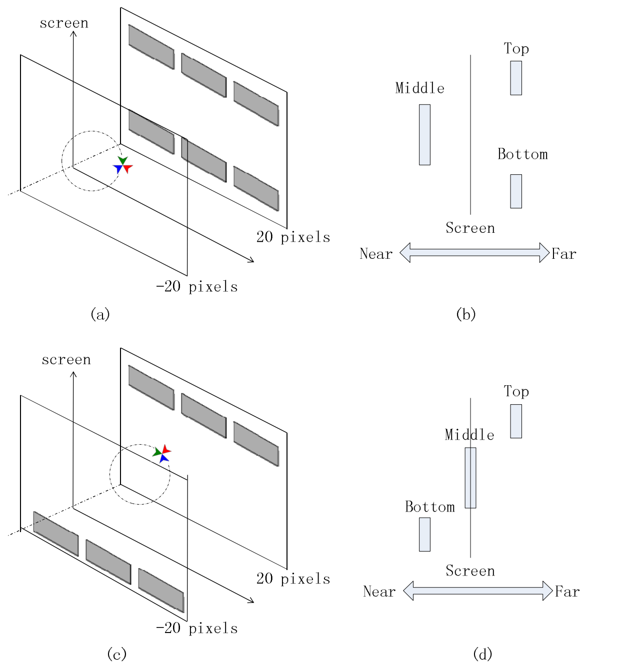

Fig. 8 Examples of stimuli used in experiment C.
(a) 2-1-2 distribution mode. (b) Side view of (a).
(c) 3-2-1 distribution mode. (d) Side view of (c).

TABLE II
MODES OF DISPARITY DISTRIBUTION AND CORRESPONDING MOS

Sequence number	Mode	MOS
1	*1_1_1*	3.38
2	*1_2_1*	3.12
3	*1_2_2*	2.77
4	*1_1_2*	2.53
5	*1_2_3*	2.64
6	*1_3_2*	2.82
7	*2_1_1*	2.77
8	*2_2_1*	3.05
9	*2_1_2*	2.38
10	*2_3_1*	2.75
11	*2_1_3*	2.28
12	*3_2_1*	3.20
13	*3_1_2*	2.49

1_2_1 respectively. Mode *1_1_1* means the top, the middle and the bottom parts are at the same plane. Under this circumstance, the range of parallax is zero, which is the most comfortable range. Nevertheless, this kind of distribution hardly exists in real stereoscopic content services. Mode *3_2_1* represents that the bottom portion of the image appears closer and the top portion of the image appears farther away. Disparity values in the whole image gradually increase from the bottom to the top. If the disparity distributes like this, audiences can get a relatively good viewing experience. This result corroborates the conclusion which is drawn in [11]. Apart from this, it is noticeable that better visual comfort is always obtained when the bottom part was perceived nearer to the viewers than the middle part. This finding is significant in stereoscopic recording, editing, effects, and postproduction.

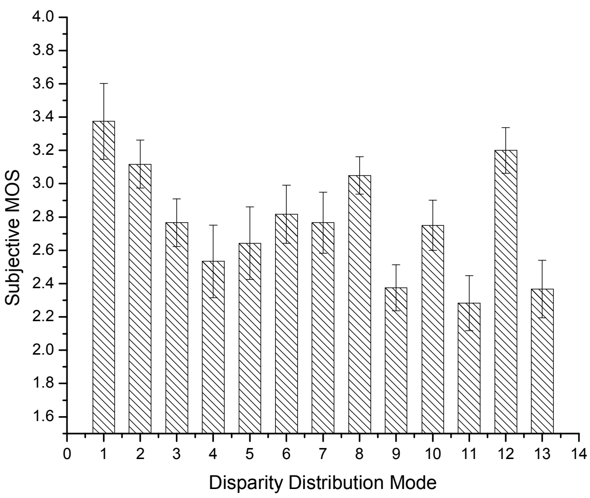

Fig. 9 Relationship between the MOS and disparity distribution. Vertical bars denote the confidence interval at the 95% confidence level

D. *In-depth motion*

Variation of the disparity over time significantly influences the viewing comfort. In this experiment, the specific influence of in-depth motion on visual comfort is evaluated. The subjective assessment was conducted with 32 stereoscopic computer generated moving pictures as test sequences. All sequences were 5 seconds long. These sequences were classified into two groups according to their starting plane. In half of the sequences, the object began to move in depth from the depth plane of -10 pixels (in front of the screen). The object in the other sequences moved along the depth plane from 10 pixels behind the screen. There were 8 levels of in-depth velocity: 2, 4, 6, 8, 10, 12, 14, and16 pixels per second. The direction of motion included forward and backward, respectively. Here, if the object moved forward to the viewers, the velocity was defined as negative. The direction and velocity remained constant in one sequence. Examples of the stimuli used in this assessment are shown in Fig. 10.

One-way ANOVA test results (F=6.89, P=0.00004) indicate that velocity of in-depth motion was a main factor which affect visual comfort. The scatter plot of in-depth velocity versus the MOS and the confidence intervals is depicted in Fig. 11 with the starting-plane as a parameter. Evidently, visual comfort increased with the decrease of the absolute velocity in the depth direction. This is because if objects travel with high speed in depth, audiences have to adjust their eyes frequently to fuse the image, which is one of the major sources of discomfort. On the other hand, equal speed and direction with different starting planes result in different levels of discomfort. It is observed that when the object is

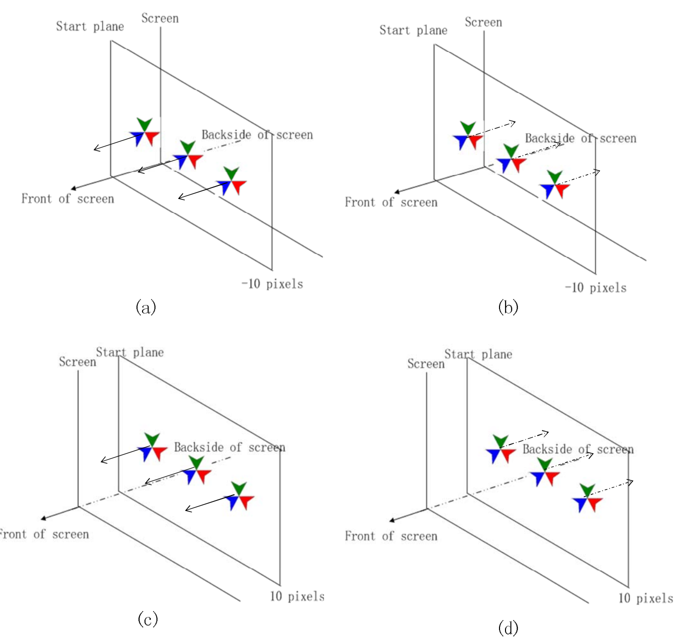

Fig. 10 Examples of stimuli used in experiment D. (a) Starting plane is in front of the screen and moves forward. (b) Starting plane is in front of the screen and moves backward. (c) Starting plane is behind the screen and moves forward. (d) Starting plane is behind the screen and moves backward

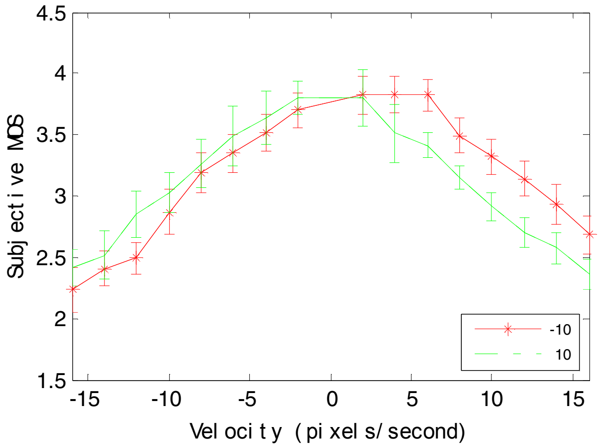

Fig. 11 Relationship between in-depth motion and the MOS. Vertical bars denote the confidence interval at the 95% confidence level

moving forward, better visual comfort is produced when the starting plane is behind the screen than in front of the screen. A contrary result can be seen when moving backward. That is, when the object moves backward, visual comfort is better when the starting plane is in front of the screen compared to behind the screen.

To compare the influence of the motion direction, the MOS distribution obtained by the condition of equal absolute disparity but different motion direction is shown in Fig. 12. Apparently, objects with an equal speed with different directions lead to different levels of discomfort when objects start to move in front of the screen (F=14.932, P=0.00015) according to Fig.12 (a). While, the level of comfort was not sensitive to motion direction when starting plane was behind the screen (F=0.88925, P=0.34688) as shown in Fig.12 (b). Actually, better visual comfort is always obtained when object moves toward the screen.

From the above analysis, it can be concluded that the influence of in-depth motion on visual comfort is quite complicated. Visual comfort can be affected not only by the velocity of the moving object, but also the motion direction, starting plane, and so on.

III. Conclusions

Visual comfort of viewing stereoscopic contents can be affected by many factors. In this paper, we examined four of the major factors by conducting subjective assessment, including the foreground disparity, size of foreground object, disparity distribution and in-depth motion. Conclusions are drawn according to the experimental results. Firstly, the range of disparity between the foreground and the background is more significant in determining the visual discomfort than the binocular disparity of the foreground. Secondly, when the absolute disparity of the object is small, visual comfort increases with the increment of the object size at the beginning and then levels off or decreases slightly when the object size reaches a certain threshold. However, for relative large disparities, the degree of visual comfort goes up continuously possibly owing to the enlargement of the threshold. The reason for the enlargement of the threshold can be explained by the principle of stereoscopic imaging. Thirdly, the best visual comfort is experienced when the parallax distribution is such that the bottom portion of the image appeared closer, the top portion of the image appeared farther away and the middle was between the above two portions. Last but not least, influences of in-depth motion on visual comfort can be reflected by many aspects, not only its velocity, but also the motion direction, starting plane and so on.

Stereoscopic visual comfort assessment is a complicated issue and further research is currently under investigation.

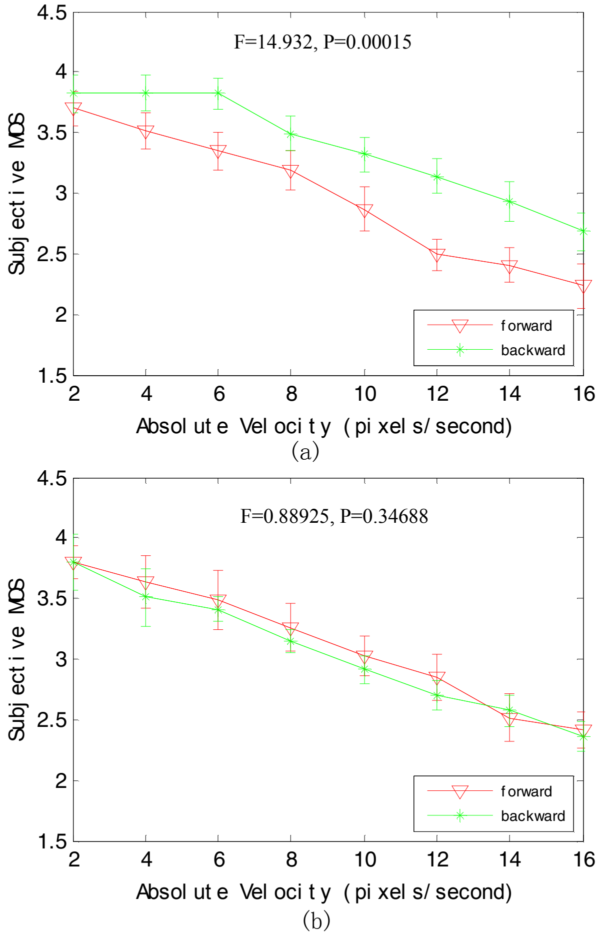

Fig. 12 Relationship between absolute velocity and the MOS. (a)Starting plane is -10 pixels. (b) Starting plane is 10 pixels. Vertical bars denote the confidence interval at the 95% confidence level

References

[1] M. Lambooij, W. A. IJsselsteijn, and I. Heynderickx, “Visual discomfort of 3-D TV: Assessment methods and modeling,” *Displays,* vol. 32, no.4, pp. 209–218, Oct. 2011.

[2] F. Speranza, W. J. Tam, R. Renaud, and N. Hurb, “Effect of disparity and motion on visual comfort of stereoscopic images,” in Proc. 13th SPIE Stereoscopic Displays Virtual Reality Syst, vol. 6055,pp. 60550B-2–60550B-10,Feb. 2006.

[3] W. J. Tam, F. Speranza, S. Yano, K. Shimono, and H. Ono, “Stereoscopic 3-D-TV: Visual comfort,” IEEE Transactions on

Broadcasting, vol. 57, no. 2, pp.335-346, Jun. 2011.

[4] M. T. M. Lambooij, W. A. IJsselsteijn, and I. Heynderickx, “Visual discomfort in stereoscopic displays: A review,” Stereoscopic Displays Virtual Reality Syst. XIV, vol. 6490, pp. 649001–649013, 2007.

[5] H. Yamanoue, M. Okui, and F. Okana, “Geometrical analysis of puppet theater and cardboard effects in stereoscopic HDTV images,” IEEE Trans. on Circuits and Systems for Video Technology, vol. 16, no. 6, pp. 744-752, 2006.

[6] N. Hiruma and T. Fukuda, “Accommodation response to binocular stereoscopic TV images and their viewing conditions,” SMPTE J., vol. 102, pp. 1137–1144, 1993.

[7] J. P. Wann, S. Ruston, and M. Mon-Williams, “Natural problems for stereoscopic depth perception in virtual environments,” Vis. Res, vol. 35, pp. 2731–2736, 1995.

[8] S. Ide, H. Yamanoue, M. Okui, and F. Okano, “Parallax distribution for ease of viewing in stereoscopic HDTV,” in Proc. Stereoscopic Displays Virtual Reality Syst. IX, vol. 4660, pp. 38–45 ,2002.

[9] Y. Nojiri, H. Yamanoue, A. Hanazato, and F. Okano, “Measurement of parallax distribution and its application to the analysis of visual comfort for stereoscopic HDTV,” in Proc. Stereoscopic Displays Virtual Reality Syst. X, vol. 5006, pp. 195–205 ,2003.

[10] Y. Nojiri, H. Yamanoue, S. Ide, S. Yano, and F. Okana, “Parallax distribution and visual comfort on stereoscopic HDTV,” in Proc. IBC, pp. 373–380, 2006.

[11] Hosik Sohn , Yong Ju Jung, Seong-il Lee, Yong Man Ro.” Predicting visual discomfort using object size and disparity information in stereoscopic images,” IEEE Transactions on Broadcasting, vol. 59,no. 1, Mar. 2013.

[12] W. Richards and M. G. Kaye, “Local versus global stereopsis: Two mechanisms,” Vis. Res., vol. 14, no. 12, pp. 1345–1347, Dec. 1974.

[13] I. Tsirlin, L. M. Wilcox, and R. S. Allison, “The effect of crosstalk on depth magnitude in thin structures,” in Proc. 22nd SPIE Stereoscopic Displays Appl, vol. 7863, pp. 786313-1–786313-10, Feb. 2011.

[14] I. Tsirlin, L. M. Wilcox, and R. S. Allison, “The effect of crosstalk on the perceived depth from disparity and monocular occlusions,” IEEE Transactions on Broadcasting, vol. 57, no. 2, pp. 445–453, Jun. 2011.

[15] D. Kim and K. Sohn, “Visual fatigue prediction for stereoscopic image,” IEEE Trans. on Circuits and Systems for Video Technology, vol. 21, no. 2, pp. 231–236,Feb. 2011.

[16] ITU-R, “Subjective assessment of stereoscopic television pictures,” ITU-R Recommendation BT.1438, 2000.

[17] D.H. Brainard, ”The psychophysics toolbox,” Spatial Vision,vol. 10, pp. 443-446, 1997.

[18] D.G.Pelli, “The videotoolbox software for visual psycho-physics: Transforming numbers into movies,” Spatial Vision, vol. 10, pp. 437-442, 1997.

[19] Jing Li; Barkowsky, M.; Wang, J.; Le Callet, P., "Study on visual discomfort induced by stimulus movement at fixed depth on stereoscopic displays using shutter glasses,",17th International Conference on Digital Signal Processing (DSP),pp.1-8,Jul. 2011.