STUDY ON DISTORTION CONSPICUITY IN STEREOSCOPICALLY VIEWED 3D IMAGES

Ming-Jun Chen ,[1,3], Alan C. Bovik [1,3], Lawrence K. Cormack [2,3]

Department of Electrical & Computer Engineering, The University of Texas at Austin, USA[1]
Department of Psychology, The University of Texas at Austin, Austin, TX, USA[2]
Center for Perceptual Systems, The University of Texas at Austin, Austin, TX, USA[3]

ABSTRACT

We describe a study aimed towards increasing our understanding of the perception of distorted stereoscopic 3D images, by analyzing subjects' performance in locating local distortions in stereoscopically viewed images. Nineteen subjects were recruited for this study. The results indicated that contrast and range variations are correlated with the conspicuity of some distortions, but not others.

Index Terms— 3D perception study, binocular suppression, contrast masking, range masking, disparity masking

1. INTRODUCTION

Research on stereoscopic 3D perception has spanned decades, but commercialized digital stereoscopic 3D products have only recently become popular. The current trend in 3D is resulting in an enormous amount of 3D content raising a suddenly urgent question: how can we evaluate the quality of 3D content? To tackle this difficult problem, it is important to first understand and model the human perception of distortions in 3D content. Previous work [1][2][3][4][5][6] has focused on the perception of 3D image quality. For example, in [5], it was claimed that the binocular perception of asymmetric MPEG-2 distorted images is approximately the average of the two views, but that the perception of asymmetric blur distorted images is dominated by the higher quality view. In [6]. It was claimed that the subjective image quality score of a stereo sequence was approximately the average of both views when MPEG-2 distortion was applied. The authors of [2] [3] [4] confirmed previous findings on JPEG compression distortions and also claimed that JPEG encoding has no effect on perceived depth, although Tam *et al.* [1] found that perceived depth is correlated with stereo content quality. Despite these interesting findings, none of this work has linked the statistics of natural stereoscopic content with human 3D judgments of quality.

In this paper, we study 3D distortion perception as a function of distortion types and severity and of low level image content. A specific goal was to better understand the role of masking in stereo perception.

Fig. 1 : Example of stereo image pair.

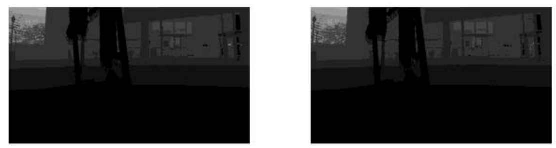

Fig. 2: Range map associated with the stereo image pair in Fig.1.

2. METHOD

2.1. Stereo Image Source

We captured co-registered stereo image and range data with a high-performance range scanner (RIEGL VZ-400) with a Nikon D700 digital camera mounted on top of it. The stereo images pairs were shot with a 65 mm interocular distance. Manual corrections were applied later to deal with translations occurring during capture. The sizes of the images are 640 x 360. Fig. 1 is one image pair used in the study and Fig. 2 shows the ground truth depth map of that image pair. Eight pairs of stereo images were taken on the campus of the University of Texas and a park nearby.

2.2. Display

A 42" Sharp HD television with a 4-mirror stereo rig was chosen as the display for this study and the viewing distance was 30". Although high-quality stereoscopic 3D televisions are available now, the 4-mirror stereo rig was used in order to avoid any question of potential crosstalk

978-1-4577-1286-9/11/$26.00 ©2011 IEEE

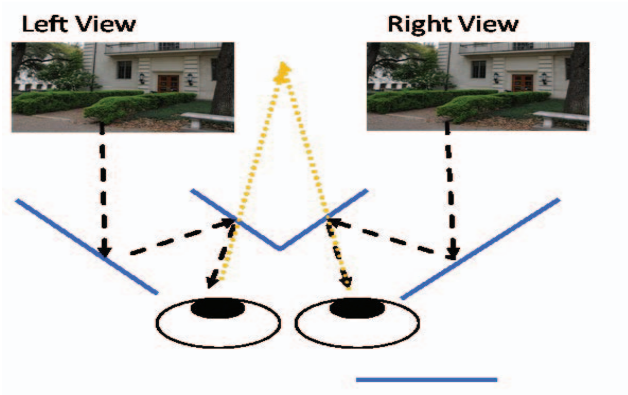

Fig. 3: Illustration of 4-Mirror stereo rig (top view)

or reduced luminance problems. No subject reported discomfort in the study. Fig. 3 shows the setup for a 4-mirror stereo rig and TV.

2.3 Observers

Nineteen naïve observers (four females and fifteen males) were recruited from the University of Texas student population. They were screened for visual acuity and stereo depth perception. Their ages were between 20~34 years old.

2.4 Stimuli

Eight stereo image pairs were used to create all stimuli. In order to explore the relationship between the statistics (texture and range) of the 3D data and the perception of stereoscopic distortions, each pristine stereo pair was distorted only within a local area (128 x 128 square window). To create a locally distorted image, the pristine and global distorted images were blended together within a local patch using a 2D Gaussian weighting function with standard deviation 34 pixels. Fig. 4 is an example of an image with a blending local distortion.

The variables that were controlled to create different stimuli were distortion type, severity of distortion, and the position of the distortion in each view. Four of the distortion types in the LIVE image quality database were used [10]: white noise, blur, JPEG compression distortion, and JPEG2000 compression distortions. The degree of severity of the distortion on each image was randomly chosen within a predefined range, ranging from just noticeable to pretty obvious. In all stimuli where both views were distorted, the distortion severity of both views was made equal.

We created distorted image pairs having random degrees of distortion severity. Varying the degree of severity allows us to probe the distortion conspicuity as a function of both severity and image or range content, possibly revealing insights regarding masking effects under 3D stereoscopic viewing. The locations of the local distortions in the left and right views were defined in two ways, both random. In the first case, the local distortions were inserted at the same position in the left and right views. We will refer to this as ‘’binocular distortion”. In the second case, the distortion was inserted into both images randomly. We will refer to this as “dichoptic masking”. In total, we created 136 stimuli for the study, including 8 pristine stereo image pairs.

2.5 Procedure

We followed the recommendation for a single stimulus continuous quality evaluation (SSCQE) [7] to decide the time spent by observers locating the local distortion and to supply a subjective quality rating for each stereo image pair. For each stimulus, the subject was asked to point out a distortion by moving a mouse cursor. The time subjects spent on the task was recorded. Then the subject was requested to give a subjective 3D image quality rating. A training section was conducted for each subject at the beginning of the study to make sure that subjects were comfortable with our 3D display environment and to help them to be familiar with the user interface and their task in the study. The training content was different from the images in the study and was impaired using the same distortion. In the training, subjects were told there is at least one distortion in every stimulus and they were not informed about the size of the distortion.

Fig. 4 : Image with local white noise distortion. The boundary was blended using a Gaussian blending window. When the image was presented, the subject was requested to point out the distortion by clicking the mouse cursor on the distortion.

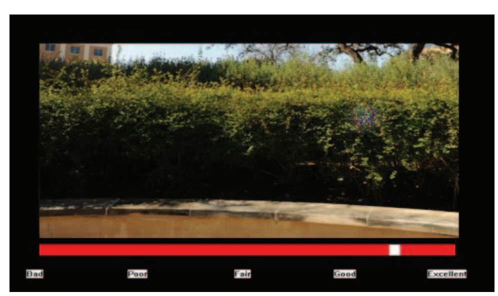

Fig. 5 : When the rating bar showed up, the subject was requested to render a subjective 3D quality opinion on the entire image.

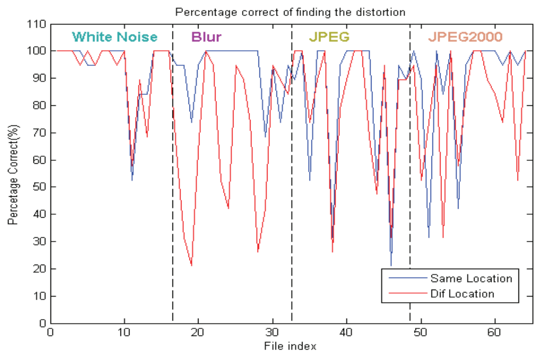

Fig. 6 : Plot of percent correct finding the distortion.

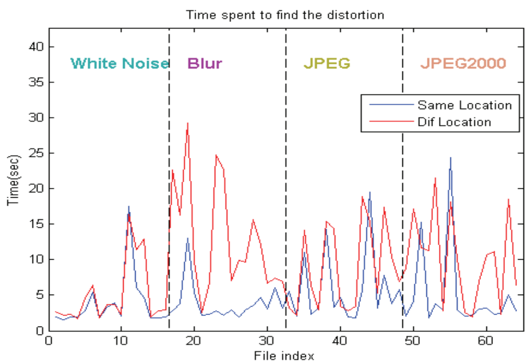

Fig. 7 : Plot of time spent finding the distortion.

3. ANALYSIS AND RESULTS

3.1 Binocular Suppression

This section discusses the effect of binocular suppression on the conspicuity of each distortion type. Fig. 6 shows the Percentage Correct (PC), which is defined as the percentage of subjects who successfully locate a distortion on a single stimulus, for both methods of randomly determining the locations of the distortions in both views. The blue line represents the PC when the distortions were placed at the same location in both views (binocular distortion) and the red line shows the PC when the distortions were inserted at two different random locations in both views (dichoptic distortion). In finding a local distortion on a dichoptic distorted image, locating either one of the local distorted patch is considered as success in the task. Fig. 7 plots the median Time Spent (TS) by all subjects to complete the task of observing each stimulus. We hypothesize that the PC and the TS provide statistical clues regarding the effects of binocular suppression and content masking on distortion conspicuity. If fewer subjects locate a distortion, or more time is spent completing a task, then, on average, this provides evidence that distortion was less visible due to masking or suppression. As shown in Fig. 6 and Fig. 7, a significant difference between the blue line (distortions were inserted into the same location in both view) and the red line (distortions were inserted into two random locations) is apparent. Further, an analysis of variance (ANOVA) was applied on PC and TS and the results given in Table 1. We can see a significant difference in behavior both in the PC (p=0.012) and the TS (p=0.000). The result suggested that binocular suppression [8] plays an important role in the perception of stereoscopic ally viewed 3D distortions.

Furthermore, we examined binocular suppression across distortion types. Table 2 and Table 3 show the results of ANOVA on PC and TS of each distortion type. As shown in Table 2, only blur distortion had a significant influence on the PC. However, Table 3 shows that blur distortion and JP2K distortion both had a significant influence on the TS.

We didn't observe any significant binocular suppression effect for white noise and JPEG compression distortions, whereas binocular suppression appears to have played a significant role in affecting blur and JP2K distortion conspicuity. This accords with [2][5], where the authors pointed out that the binocular suppression effect affects blur distortion, but not MPEG-2 distortion. Our results may provide useful pointers for designing asymmetric codecs. For instance, asymmetric coding may work well on JP2K compression, but not on JPEG compression.

Variables	F ratio	p
PC	6.49	0.0121
TS	20.1	0.0000

Table 1 : The results of the ANOVA on the locations of distortions.

Distortion Type	F ratio	p
White Noise	0.02	0.882
Blur	13.11	**0.0011**
JPEG	0.08	0.7745
JP2K	1.39	0.2485

Table 2 : The results of ANOVA of each distortion type on percent correct.

Distortion Type	F ratio	p
White Noise	0.65	0.4281
Blur	19.59	**0.0001**
JPEG	2.2	0.1487
JP2K	5.43	**0.0267**

Table 3 : The results of ANOVA of each distortion type on time spent to find the distortion.

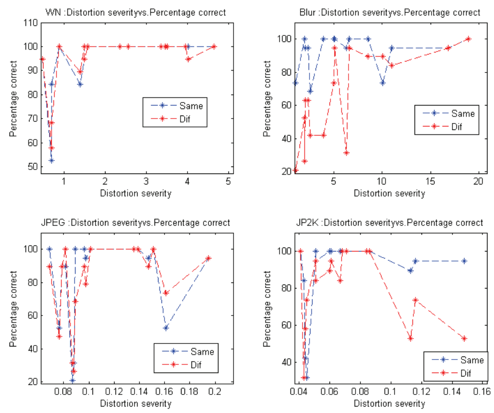

Fig. 8 : Plot of Percent Correct vs. Distortion Strength for 4 types of distortions.

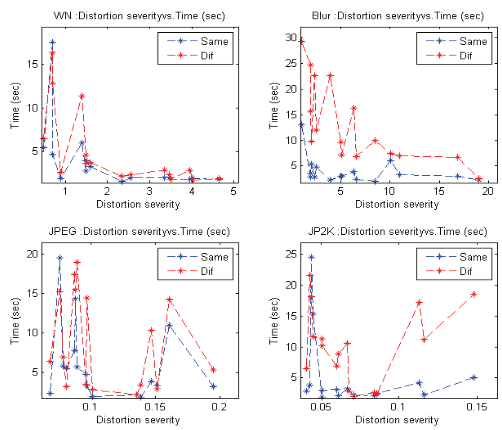

Fig. 9 : Plot of Time Spent to find distortion vs. Distortion Strength for 4 types of distortions.

3.2 Distortion Severity

We also studied the question of whether the conspicuity of distortions was affected by masking effects. This question may be partially addressed by analyzing the relationships between subjects' performance in finding the distortion and the severity of the distortion. Fig. 8 and Fig. 9 plot distortion severity vs. PC and TS. In these two figures, the blue line means the distortions were inserted into the same locations in both views (binocular distortion) and the red lines indicate that the distortions were randomly inserted into each view (dichoptic distortion). In the case where there was no binocular rivalry (blue line), we can see that when the images were severely distorted, the subjects could locate the distortion easily regardless of the location of the distortion. For example, for white noise distortion (the top left plot on Fig. 8), the line was flat if the distortion severity was larger than 2. For JPEG and blur distortion, the same trend can be observed although there are some fluctuations.

When we take binocular rivalry into consideration (the red lines in Fig. 8 and Fig. 9), the perception of severely distorted test stereopairs was not affected by any masking in the cases of white noise and JPEG distortion. However, binocular suppression of the stereopairs may mask some distortions thereby affecting PC and TS, as for blur and JP2K distortions.

3.3 Contrast and Range Masking

3.3.1. Binocular distortion

To analyze the data that were collected without binocular rivalry, we divided all test stereopairs affected by each distortion into two groups; a High percent correct group (High) and a Low percent correct group (Low), then performed a Welch's t test [9] on their PC, TS, range activity, and contrast for each distortion. The range activity was defined as the weighted mean of the gradient values of the range map inside the patches in both views. The contrast was defined as the weighted mean of the gradient values of the image inside the patches in both views. Our assumption is that if there is a contrast masking or range masking effect while the subjects were viewing the stereoscopic 3D image, the subjects' performance (PC and TS) on a test should be correlated with the contrast or the range activity of the local patch. The threshold used to divide the group was 85%, which was set based on the assumption that if more than 15% of our subjects can't find the distortion in a particular test, there should be some masking effect in that test. For instance, in the top left plot on Fig. 8 and Fig. 9, it is clear that the visibility of the local distortion wasn't affected by any masking effect when the distortion severity was larger than 2.

On each plot in Fig. 10, the results of Welch's t test are shown and the significance tests are marked. From the results, it appears that contrast activity affects the visibility of white noise and JP2K distortion. For white noise (top left plot in Fig. 10), we found that there is contrast masking when viewing stereoscopic 3D because significantly higher contrast values occur in the low performance group. Regarding contrast, for JP2K distortion (bottom right plot in Fig. 10), distortions in lower contrast areas tend to be less visible. With respect to range activity, we only saw a significant result on blur distortion conspicuity (top right plot in Fig. 10). As we pointed out in Section 3.2, some of our samples may not have produced any masking effect due to the severe distortion. This may be the reason that we didn't get the same significant results on blur and JP2K distortions. Of course, the distortion produced by JP2K compression is similar to blur distortion. However, from the top right and the bottom right plots in Fig. 10, both had higher range

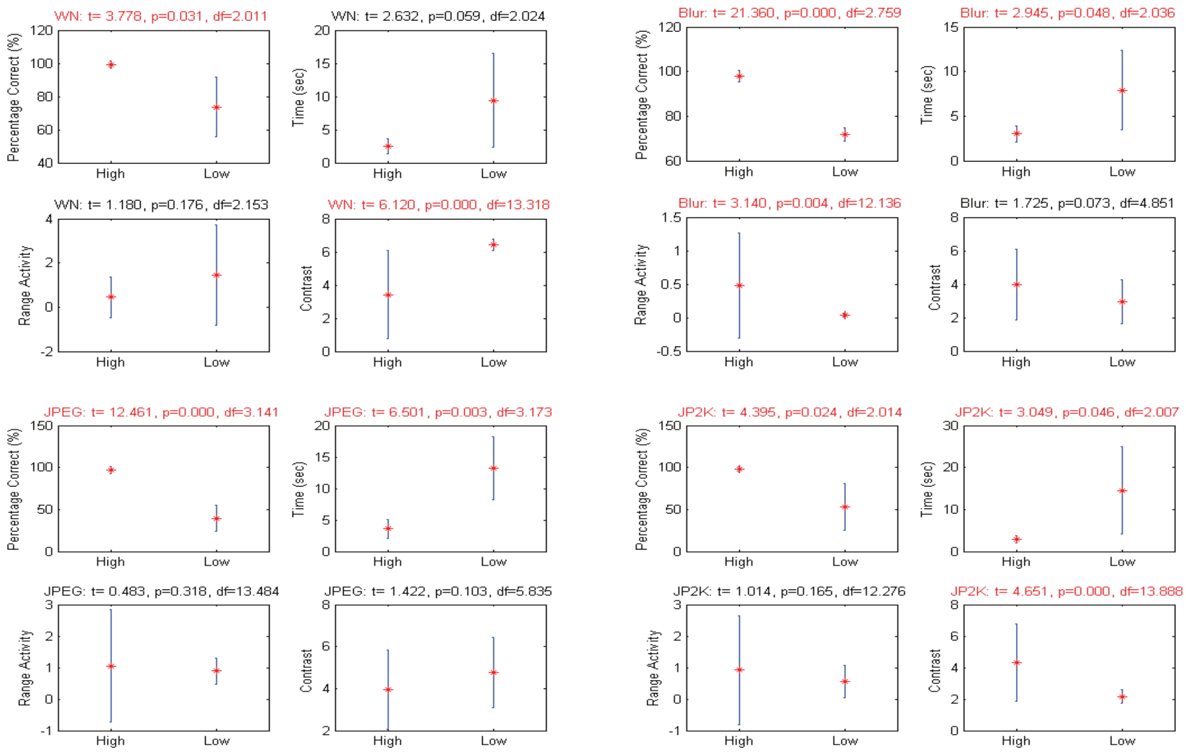

Fig. 10 : The results of Welch's *t* test on white noise (top left), blur (top right), JPEG (bottom left) and JP2K (bottom right) compression distortion. Red dot indicates mean and blue bar represents standard deviation.

activity and higher contrast in the high performance group. This suggests that both blur and JP2K distortion are correlated with contrast and range activity in stereoscopically viewed 3D images.

3.3.2. Dichoptic Distortion

For the data that were collected with binocular rivalry, we used the chi-square test to verify whether contrast or range activity had a significant impact on the distortion conspicuity. In these data, different subjects chose different patches in a single test and the contrast or range activity inside these two patches (left and right) may vary. Hence, the analysis method used in section 3.3.1 to analyze the correlation between local image statistics and subjects' performance was not applied. The null hypothesis of the chi-square test is that the contrast or range activity in left or right view is not related to the subjects' selections. Namely, the subjects' selections on left view or right view are decided by chance. We first performed a *t*-test on the overall selections on left and right view. The result was shown in Fig. 11 which failed to reject the null hypothesis (p=0.751 > 0.05). Therefore, we concluded that the subjects' choices on left view and right view are decided by chance.

Next, we ran a chi-square test to see if there is any correlation between contrast or range activity with subjects' selection. Table 4 is an example of the chi-square test setup. Table 5 shows the results of all tests on contrast masking. From it we can see that the contrasts of local patches are correlated with the visibility of blur (p=0.003), JPEG (p=0.004), and JP2K (p=0.03) distortions. As for the influence of range activity, Table 6 shows that local range energy of local patches is correlated with the visibility of blur (p=0.000) and JP2K (p=0.016) distortions. Thus, both contrast and range activities have a significant influence on the perception of stereoscopically viewed blur and JPEG distortions, which corresponds to our argument in Section 3.3.1.

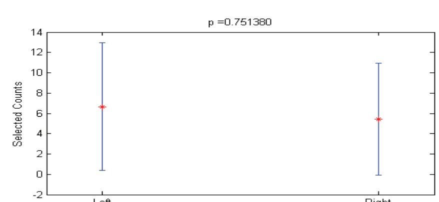

Fig. 11 : The results of *t* test on subjects' selection on left and right views.

	Subjects Selection		
Contrast on blur	Left	Right	Total
Left > Right	113 (107.72)	33(38.28)	146
Left < Right	8(13.28)	10 (4.72)	18
Total	121	43	164

Table 4 : Example of a Chi-square test setup. The number inside brackets is the expected value.

Chi-square test on Contrast				
Distortion type	WN	Blur	JPEG	JP2K
Chi value	2.551	8.995	8.400	4.733
p	0.110	**0.003**	**0.004**	**0.030**

Table 5 : The results of Chi-square test on contrast

Chi-square test on Range				
Distortion type	WN	Blur	JPEG	JP2K
Chi value	1.576	18.389	1.532	5.770
p	0.209	**0.000**	0.216	**0.016**

Table 6 : The results of Chi-square test on range

4. CONCLUSION AND DISCUSSION

This study aimed at understanding the perception of distortions in stereoscopic 3D images by analyzing subjects' performance (PC and TS) in finding the local distortions. From the results, we found that the contrast and range variations are correlated with human perception when viewing stereoscopic 3D images with blur and JP2K distortions and the blur and JP2K distortions are more visible in regions containing higher local measured contrast or range energy. In addition, the contrast masking effect in white noise and JPEG compression distortion was observed, but there wasn't a significant masking effect detected arising from range variation. However, in some tests of viewing distorted stereoscopic 3D images with white noise and JPEG compression, subjects behaved differently although the contrast and distortion severity in those samples were similar.

ACKNOWLEDGEMENT

This research was supported by the National Science Foundation under Grant NSF IIS-0917175.

5. REFERENCES

[1] W. J. Tam, L. B. Stelmach, and P. J. Corriveau, "Psychovisual aspects of viewing stereoscopic video sequences", *Stereoscopic Displays and Virtual Reality Systems V,* pp.226-235

[2] P. Seuntiëns, L. Meesters, and W. Ijsselsteijn, "Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation," *ACM Trans. Appl. Percept.,* vol.3, no.2, pp. 95 – 109, 2006.

[3] P. Seuntiëns, L. Meesters, and W. Ijsselsteijn, "Perceptual evaluation of JPEG coded stereoscopic images,", *Proc. SPIE,* vol. 5006, pp. 215-226, 2003

[4] L. Meesters, W. Ijsselsteijn, and P. Seuntiëns, "A survey of perceptual evaluation and requirements of threedimensional TV," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 14, no. 3, pp. 381-391, Mar. 2004.

[5] D. V. Meegan, L. B. Stelmach, and W. J. Tam, Unequal weighting of monocular inputs in binocular combination: implications for the compression of stereoscopic imagery.", *Journal of Experimental Psychology: Applied 7,* 143-153,2001

[6] L. B. Stelmach and W. J. Tam, "Stereoscopic image coding: effect of disparity image quality in left- and right-eye views.", *Signal Processing: Image Communication 14,* 111-117, 1998.

[7] ITU-R Recommendation BT.500-11, "Methodology for the subjective assessment of the quality of television pictures," *International Telecommunication Union, Geneva, Switzerland,* 2002.

[8] W. Levelt, "On Binocular Rivalry." Royal VanGorcum, Assen, The Netherlands, 1965.

[9] B. L. Welch, (1947), "The generalization of "Student's" problem when several different population variances are involved", *Biometrika* 34 (1–2): 28–35

[10] H. R. Sheikh, Z. Wang, L. K. Cormack and A. C. Bovik, "LIVE Image Quality Assessment Database," Release 2, [online]: Available at: http://live.ece.utexas.edu/research/quality/subjective.htm