VALIDATION FOR VISUALLY LOSSLESS COMPRESSION OF STEREO IMAGES

Graduate Student: Hsin-Chang Feng[1]
Faculty Advisors: Michael W. Marcellin[1] and Ali Bilgin[2, 1]
[1]Department of Electrical and Computer Engineering, [2]Biomedical Engineering
The University of Arizona, Tucson, AZ 85721

ABSTRACT

This paper described the details of subjective validation for visually lossless compression of stereoscopic 3 dimensional (3D) images. The subjective testing method employed in this work is adapted from methods used previously for visually lossless compression of 2 dimensional (2D) images. Confidence intervals on the correct response rate obtained from the subjective validation of compressed stereo pairs provide reliable evidence to indicate that the compressed stereo pairs are visually lossless.

Keywords: Stereoscopic images; visually lossless coding; JPEG2000; crosstalk; human visual system.

1. INTRODUCTION

The main difference between 2D images and stereoscopic 3D images is depth information. Human viewers can easily perceive the position of each object in a stereoscopic image according to depth information. Stereoscopic 3D images have been applied in many applications such as aerial stereo photography and 3D stereoscopic surgery that requires exact knowledge of depth information. Uncompressed stereo pairs require double the amount of data compared to ordinary 2D images since a left and a right image are contained in a stereoscopic image. Visually lossless compression can be used to reduce the amount of data efficiently without losing information that would be useful to the human visual system (HVS).

One simple approach to visually lossless compression of stereo pairs might be to compress the left and right images independently as ordinary 2D images so that each of them is visually lossless individually. Unfortunately, due to crosstalk caused by leakage between the left and right channels present in current 3D display technologies, initial results obtained in this fashion did not meet expectations. Although the compressed left and right images were visually lossless when each of them was displayed in 2D mode, the compressed stereo pairs were not visually lossless when displayed in 3D mode. Thus, the crosstalk effect must be considered for visually lossless compression of stereoscopic 3D images.

In our previous paper [1], a methodology for the measurement of visibility thresholds (VTs) for 3D displays with active shutter glasses was provided. These measured VTs are then used in the models defined by equation (1) in [1] to find appropriate VTs for visually lossless compression of stereo images. These models are functions of not only the wavelet subband and coefficient variance, but also of gray level I in both the left and right images. The novel visually lossless compression algorithm for monochrome stereo pairs proposed in [5] is then used to compress the left and right images of a stereo pair jointly according to the VT of their wavelet coefficients.

In order to validate the visual losslessness of the resulting compressed images, subjective experiments should be conducted. Two well-known methods, sequential two-alternative forced-choice (2AFC) testing [4] and spatial three-alternative forced-choice (3AFC) testing [2], [3] have been used in the literature to evaluate whether quantization distortions in compressed *2D images* can be detected. In the first method, two copies of the image (one original and one compressed) are displayed sequentially in time. Subjects are forced to make a choice of the one with best quality between these two copies (guessing is permitted if necessary). The second method simultaneously displays three copies of the image. Two of the copies are original and one is compressed. The three copies are presented side-by-side on the screen in random order. Subjects are then asked to choose which one differs from the other two. Again, guessing is permitted if the subject cannot find any difference.

A sequential 3AFC method has been used in subjective *sound* quality evaluations [11]. In this method, three audio clips are presented sequentially to subjects. Two of the clips are original and one of the clips has been compressed. The subjects are forced to choose the clip that is different from the other two.

2. VALIDATION OF VISUAL LOSSLESSNESS

3AFC testing was used in [2] and [3] due to its inherent advantages over 2AFC. Specifically, the 3AFC method can provide more rigorous results than 2AFC method. If the compressed images truly are visually lossless, then subjects will be forced to make their decisions by guessing. In this case, the population proportion of correct answers obtained from 2AFC and 3AFC methods would be 1/2 and 1/3, respectively. The converse also holds for 3AFC. That is, a population proportion of correct answers equal to 1/3 indicates that the images are visually lossless. However, the 2AFC testing method has an inherent problem that can cause an ambiguous interpretation of a population proportion 1/2.

A scenario that can cause ambiguous interpretation of results obtained via 2AFC method arises when the compressed images are very high quality, but subjects still can perceive differences between the original and the compressed images (the compressed images are only *nearly* visually lossless). In this situation, the resulting population proportion could also be 1/2, since these two images may be distinguishable, yet neither looks clearly better than the other. The 3AFC testing method can avoid this problem because two uncompressed copies are contained in each trial. The subjects can easily make a correct choice if the differences between the compressed and the uncompressed images can be perceived.

In [2] and [3], spatial 3AFC testing was performed. Accordingly, three copies of a 2D image were displayed side-by-side on the screen. Unfortunately, the spatial 3AFC method is not appropriate for visual testing of stereoscopic 3D images. This is due to the fact that three full resolution copies of stereo pairs do not fit side-by-side on the screen. Visibility of artifacts, as well as 3D perception of stereoscopic images can be adversely impacted if the dimensions of the images are decreased by cropping or subsampling. Thus, such techniques should be avoided during perceptual testing. For this reason, the sequential 3AFC method was adopted for this work.

The 3D vision system used to obtain the results presented below includes the Nvidia Quadro FX 3800 graphics card and active shutter glasses. A USB IR transmitter is used to synchronize the glasses with a Samsung SyncMaster 2233 RZ display (22” with 1680 × 1050 resolution, 120 Hz refresh-rate, 300 cd/m^2 brightness, 1,000:1 typical contrast ratio, and $170^o/160^o$ Viewing Angle). The display resolution d for the SyncMaster display is 35.45 pixels/cm. When the viewing distance υ is 60 cm, the resulting visual resolution r is 37.12 pixels/degree, which is derived via equation (1) in [9]. All visual experiments were conducted with normal office lighting conditions.

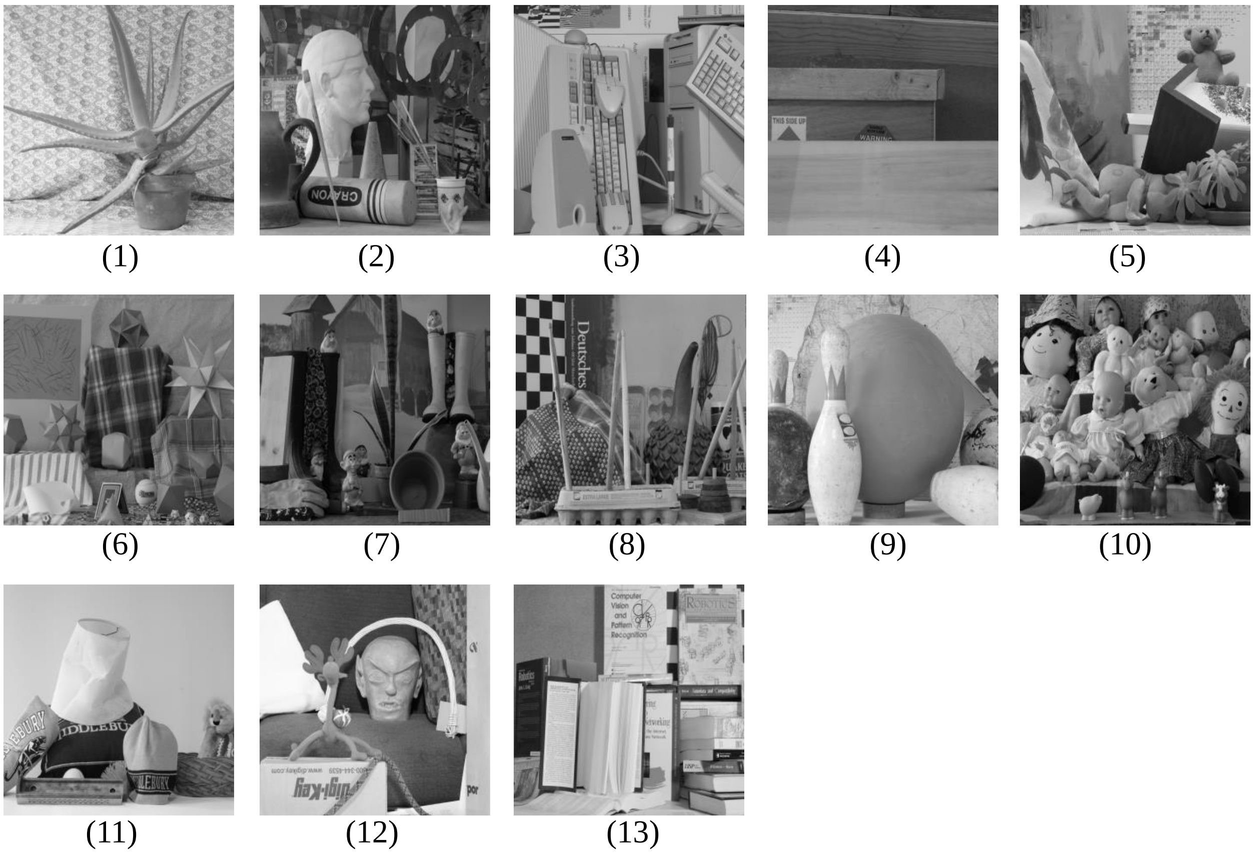

Fig. 1. The thirteen monochrome stereo images used in the validation of visually lossless stereoscopic compression (only the left image of each stereo pair is shown). (1) Aloe, (2) Art, (3) Computer, (4) Wood, (5) Teddy, (6) Moebius, (7) Dwarves, (8) Drumsticks, (9) Bowling, (10) Dolls, (11) Midd, (12) Reindeer, and (13) Books.

Thirteen stereo pairs were used in the subjective experiments described below. These images were obtained from the Middlebury stereo datasets [6], [7], [8]. These thirteen stereo pairs are shown in Figure 1. The compressed stereo pairs were obtained from the visually lossless coding method proposed in [5] according to VTs derived by the models provided in [1]. Fifteen subjects with normal or corrected-to-normal vision participated in the validation. Each subject viewed the full sequence of thirteen stereo pairs six times. The order of thirteen stereo pairs was randomized in each sequence. The total number of subjective trials was then $15 \times 6 \times 13 = 1170$. In each trial, one compressed copy and two original copies of a stereo image were displayed in random order. These tree copies were displayed sequentially. The images were displayed in stereo mode, with the subjects wearing 3D shutter glasses. According to a number shown in the upper left corner of the screen, the subjects were asked to use the keyboard to indicate which copy was different from the other two. No feedback was provided to the subjects to indicate whether their choice was correct.

No time limit was placed on each trial so that the subjects could observe each copy of the image as many times and for however long they wanted. The subjects could move backward/forward between the three copies of the image via the left/right key on the keyboard. To avoid the phenomenon of persistence of vision, a black screen was displayed for 0.5 seconds while switching between two copies. Despite the fact that the visually lossless compression method [5] was designed at a typical viewing distance υ = 60 cm between subject and display, the subjects were allowed to approach to screen as closely as they wanted during the validation.

3. RESULTS

Figure 2 shows the resulting sample proportion of correct choices obtained from the subjective validation of the compressed monochrome stereo pairs. As mentioned before, under the assumption that the images are visually lossless, the population percentage of correct choices would be 1/3. This value is represented by the red dotted line in the figure. The blue rectangle bars are used to indicate the obtained sample proportion of correct choices for each individual compressed stereo image.

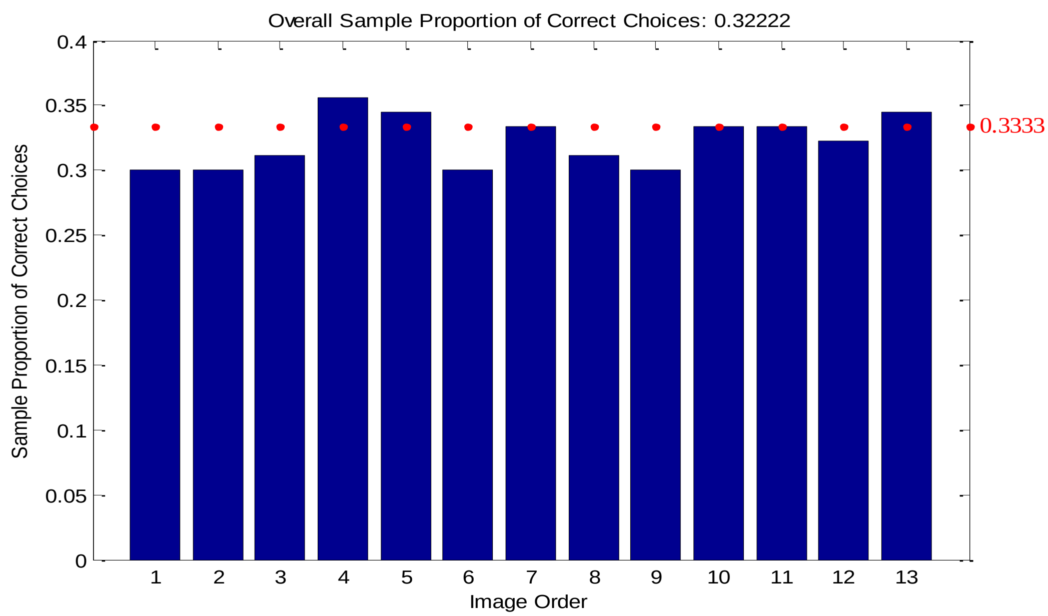

Fig. 2. Validation results. The blue rectangle bars represent the individual sample proportions of correct choices for each image individually. The assumed population proportion of 1/3 is

represented by the red dotted line.

The overall sample proportion (averaged over all 13 images) was calculated to be $\hat{p}$ = 0.32222. A confidence interval for the population percentage can be calculated [10] via

$\hat{p} \pm Z^* * \sqrt{\hat{p}(1-\hat{p})/n}$. (1)

Here n is the sample size and the value of Z* is determined according to the desired confidence. For a 95% confidence interval, the value of Z* is 1.96. For n = 1170 and $\hat{p}$ = 0.32222, the upper bound and lower bound of the 95% confidence interval are 0.3490 and 0.2954, respectively. Since the hypothesized population mean 1/3 is well within this 95% confidence interval, it is claimed that the compressed monochrome stereo pairs are visually lossless under the viewing conditions employed.

4. CONCLUSIONS

In order to validate visual losslessness of compressed stereo pairs, subjective visual tests were conducted via the sequential 3AFC testing method. The sequential 3AFC testing method provides convincing evidence that the compressed stereo pairs are no different than the original stereo images for a human viewer. Since the hypothesized population proportion of correct choices is contained within the 95% confidence interval of the sample proportion, as obtained from the subjective validation, visual losslessness is claimed for the monochrome stereo pairs compressed using the compression method proposed in [5].

REFERENCES

[1] H. Feng, M. W. Marcellin, and A. Bilgin, "Measurement of visibility thresholds for compression of stereo images," in *Proceedings of International Telemetering Conference*, San Diego, Oct. 2012.

[2] H. Oh, A. Bilgin, and M. W. Marcellin, "Visually lossless encoding for JPEG2000," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 189-201, Jan. 2013.

[3] D. M. Chandler, N. L. Dykes, and S. S. Hemami, "Visually lossless compression of digitized radiographs based on contrast sensitivity and visual masking", in *Proceeding of SPIE Medical Imaging*, vol. 5749, pp. 359 -372, 2005

[4] R. M. Slone, D. H. Foos, B. R. Whiting, E. Muka, D. A. Rubin, T. K. Pilgram, K. S. Kohm, S.S. Young, P. Ho, and D. D. Hendrickson, "Assessment of visually lossless irreversible

image compression: comparison of three methods by using an image-comparison workstation," *Radiology-Computer Applications*, vol. 215, no. 2, pp. 543 -553, 2000.

[5] H. Feng, M. W. Marcellin, and A. Bilgin, "Visually lossless compression of stereo images," in *Proceedings of Data Compression Conference*, Salt Lake, March 2013.

[6] D. Scharstein and R. Szeliski, Middlebury stereo datasets. [Online]. Available: http://vision.middlebury.edu/stereo/

[7] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007),* Minneapolis, MN, June 2007.

[8] H. Hirschmüller and D. Scharstein, "Evaluation of cost functions for stereo matching," In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007),* Minneapolis, MN, June 2007.

[9] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1164-1175, 1997.

[10] D. J. Rumsey, *Statistics workbook for dummies*. Hoboken: Wiley Publishing, 2005.

[11] M. Frank, A. Sontacchi, T. Lindenbauer, and M. Opitz, "Subjective sound quality evaluation of a codec for digital wireless transmission," *Audio Engineering Society 132nd Convention*, Budapest, Hungary, April 2012.