ESTIMATION OF SUBJECTIVE QUALITY FOR MIXED-RESOLUTION STEREOSCOPIC VIDEO

Payman Aflaki[a], *Miska M. Hannuksela*[b], *Jussi Hakala*[c], *Jukka Häkkinen*[b,c], *Moncef Gabbouj*[a]

[a]Department of Signal Processing, Tampere University of Technology, Tampere, Finland;
[b]Nokia Research Center, Tampere, Finland;
[c]Dept. of Media Technology, Aalto University, School of Science and Technology, Espoo, Finland

ABSTRACT

In mixed-resolution (MR) stereoscopic video, one view is presented with a lower resolution compared with the other one; therefore, a lower bitrate, a reduced computational complexity, and a decrease in memory access bandwidth can be expected in coding. The human visual system is known to fuse left and right views in such a way that the perceptual visual quality is closer to that of the higher-resolution view. In this paper, a subjective assessment of mixed resolution (MR) stereoscopic videos is presented and the results are analyzed and compared with previous subjective tests presented in the literature. Three downsampling ratios 1/2, 3/8, and 1/4 were used to create lower-resolution views. Hence, the lower-resolution view had different spatial resolutions in terms of pixels per degree (PPD) for each downsampling ratio. It was discovered that the subjective viewing experience tended to follow a logarithmic function of the spatial resolution of the lower-resolution view measured in PPD. A similar behavior was also found from the results of an earlier experiment. Thus, the results suggest that the presented logarithmic function characterizes the expected viewing experience of MR stereoscopic video.

Index Terms— Video signal processing, video compression, asymmetric stereoscopic video, mixed resolution, subjective evaluation

1. INTRODUCTION

Mixed resolution (MR) stereoscopic video compression introduced in [1] is a well-known approach in the field of stereoscopic video coding. In MR stereoscopic video, one view is represented with a lower resolution compared to the other one, while, according to the binocular suppression theory [2], it is assumed that the perceived quality by the Human Visual System (HVS) is closer to that of the higher quality view.

A subjective assessment of full- and mixed- resolution stereoscopic video on a 32-inch polarized stereoscopic display and on a 3.5-inch mobile display was presented in [3]. One of the views was downsampled with ratio 1/2 along both coordinate axes. Uncompressed full-resolution (FR) sequences were preferred in 94% and 63% of the test cases for 32-inch and 3.5-inch displays, respectively. While studying different resolutions for the symmetric stereoscopic video and the higher-resolution view of the MR videos, it was found that the higher the resolution, the smaller the subjective difference was between FR and MR stereoscopic video. The lower resolution view had always a downsampling ratio 1/2 vertically and horizontally.

The study presented in [4] included a subjective evaluation of MR sequences with downsampling ratios 1/2 and 1/4 along both coordinate axes. The results revealed that the subjective image quality of the MR image sequences was preserved well but dropped slightly at downsampling ratio 1/2 and 1/4.

In [5], the impact of downsampling ratio in MR stereoscopic video was studied. Downsampling ratios 1/2, 3/8, and 1/4 were applied vertically and horizontally. A 24-inch polarized display was used with a viewing distance of 70 cm. A correlation comparison between the subjective results and the average luma peak-signal-to-noise (PSNR) showed that there might be a breakdown point between downsampling with ratio 1/2 and 3/8, at which the lower-resolution view became more dominant in the subjective quality. Downsampling ratios 1/2 and 3/8 corresponded to 11.2 and 7.6 pixels per degree (PPD) of viewing angle, respectively. Moreover, it was confirmed that the ocular dominance did not affect the subjective ratings regardless of which view was downsampled in the MR sequences.

In this paper, a subjective test for uncompressed MR stereoscopic video is presented using a test setup similar to but not the same as in [5]. The obtained subjective results are compared to the previous subjective test [5] to see if the above-mentioned breakpoint is valid for a different test setup. Moreover, a novel logarithmic estimation of subjective ratings as a function of PPD values of viewing angle is introduced.

This paper is organized as follows. Section 2 explains the subjective test setup and test procedure. The subjective results are presented and discussed in Section 3. Finally, the paper is concluded in Section 4.

2. TEST SETUP

2.1 Preparation of the Test Stimuli

Four sequences were used: Pantomime, Dog, Newspaper, and Kendo. They are all common test sequences in the 3D Video (3DV) ad-hoc group of the Moving Picture Expert Group (MPEG). No audio track was used.

For each sequence, we had the possibility to choose between several camera separations or view selections. This was studied first in a pilot test of 9 subjects. The test procedure of the pilot test was similar to that of the actual test presented in Section 2.2. Several camera views were available for each sequence in the pilot test, and based on the subjective scores achieved, the 5 cm camera separation was chosen for all test sequences.

The test clips were prepared as follows. Both left and right view image sequences were first downsampled from their original

978-1-61284-162-5/11/$26.00 © 2011 IEEE

Table 1. Spatial resolution of sequences

	Full	*1/2*	*3/8*	*1/4*
All sequences	768x576	384x288	288x216	192x144

Table 2. Visual angle (in pixels per degree) of the two test setups

Downsampling ratio	Test setup presented in this paper	Test setup presented in [5]
1	30.2	22.8
1/2	15.1	11.4
3/8	11.3	7.6
1/4	7.5	5.7

resolution to the "full" resolution mentioned in Table 1. The "full" resolution was selected to occupy as large an area as possible on the used monitor with a reasonable downsampling ratio from the original resolution. As eye dominance was shown to have no impact on which view is provided with a better quality [5], only one set of MR sequences was prepared. The right view was kept in "full" resolution while the left view was downsampled and subsequently upsampled to the "full" resolution. Downsampling ratios 1/2, 3/8, and 1/4 were selected and symmetrically applied along both coordinate axes in order to keep the results easily comparable with those presented in [5]. The filters of the JSVM reference software of the Scalable Video Coding standard were used in the downsampling and upsampling operations [5].

2.2 Test Procedure

The same 24" polarizing stereoscopic screen as in [5] was used for subjective experiments. It has width and height of 515 and 322 mm, respectively, a total resolution of 1920x1200 pixels, and a resolution of 1920x600 per view when used in stereoscopic mode.

22 subjects attended this experiment of which 7 were female and 15 were male. The average age of the subjects was 23.5 years. The test viewing distance was changed from 70 cm used in [5] to 93 cm which is 3 times the height of the image, as used in some subjective test standards [7]. Hence, the visual angle differed from that in [5]. Table 2 reports the visual angle in PPD for both test setups.

Prior to the experiment, the candidates were subject to thorough vision screening. Two candidates did not pass the criterion of 20/40 visual acuity with each eye and were thus rejected. All participants had a stereoscopic acuity of 60 arc sec or better. The viewing conditions were kept constant throughout the experiment and in accordance with the sRGB standard [8] ambient white point of D50 and illuminance level of about 200 lux.

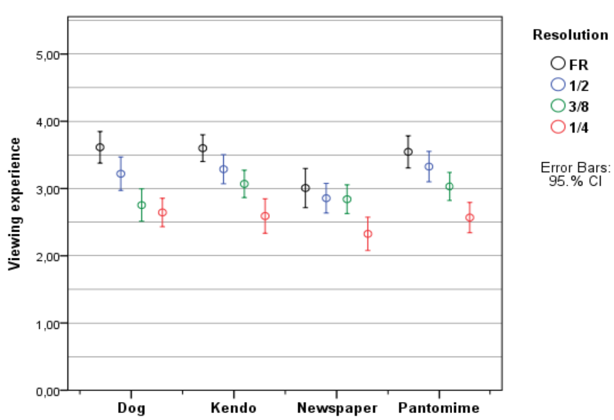


Fig. 1. Average of viewing experience ratings and the 95% CI

3. RESULTS AND DISCUSSION

3.1 Viewing Experience

The average viewing experience ratings and the 95% confidence interval (CI) are presented in Fig. 1. The subjective ratings tend to have less variation in this test than in the test presented in [5]. We observed that 18% and 69% of the total rating interval were covered by the average subjective scores of the sequences in this experiment and in [5], respectively. This result was expected, because increasing the viewing distance diminishes the subjective quality difference among MR stereoscopic videos with different downsampling ratios.

3.2 Limit of Downsampling Ratio

With the test setup presented in [5], we found that the downsampling ratio that could be applied before the lower resolution view became dominant in subjective results was between 1/2 and 3/8, i.e., between 7.6 and 11.4 PPD of viewing angle as indicated in Table 2. We studied whether the same PPD ratio threshold appeared in this experiment too. Therefore, as also done in [5], we analyzed the correlation of subjective viewing experience ratings of the presented study with PSNR of the lower resolution view upsampled to the full resolution. Unlike in [5], practically no correlation was found between the subjective viewing experience rating and the average luma PSNR of the lower resolution view for any downsampling ratio. Consequently, the analysis did not reveal the limit of the downsampling ratio for the lower-resolution view in the presented study. We suspect that the lack of correlation could have been caused by the selection of the test sequences and the smaller variation in subjective viewing experience ratings in general. It has also been discovered that the greater the angular size of the display, the more contrast sensitivity the human visual system has [9]. Thus, the threshold angular resolution for mixed-resolution stereoscopic video may also depend on the angular size of the display. As the correlation analysis of the average luma PSNR of the lower resolution view did not lead to conclusions in this test, we explored another approach for discovering the limits of the downsampling ratio in MR stereoscopic video, as presented in the next sub-section.

3.3 Logarithmic Estimation of Subjective Ratings

We analyzed ratings achieved in these experiments and those included in [5] against the PPD values of each test setup. A logarithmic relationship was observed between the subjective viewing experience ratings and the corresponding PPD values of each downsampling ratio. The fitting model used to generate the curves in Fig. 2 under the mean square error criterion is as follows:

$$y = C_1 * log(ppd - k) + C_2 \quad (1)$$

where:

ppd = pixels per degree (PPD) of viewing angle
C_1, C_2 = coefficients calculated for each sequence separately
k = fixed offset for each test setup
y = estimated subjective rating

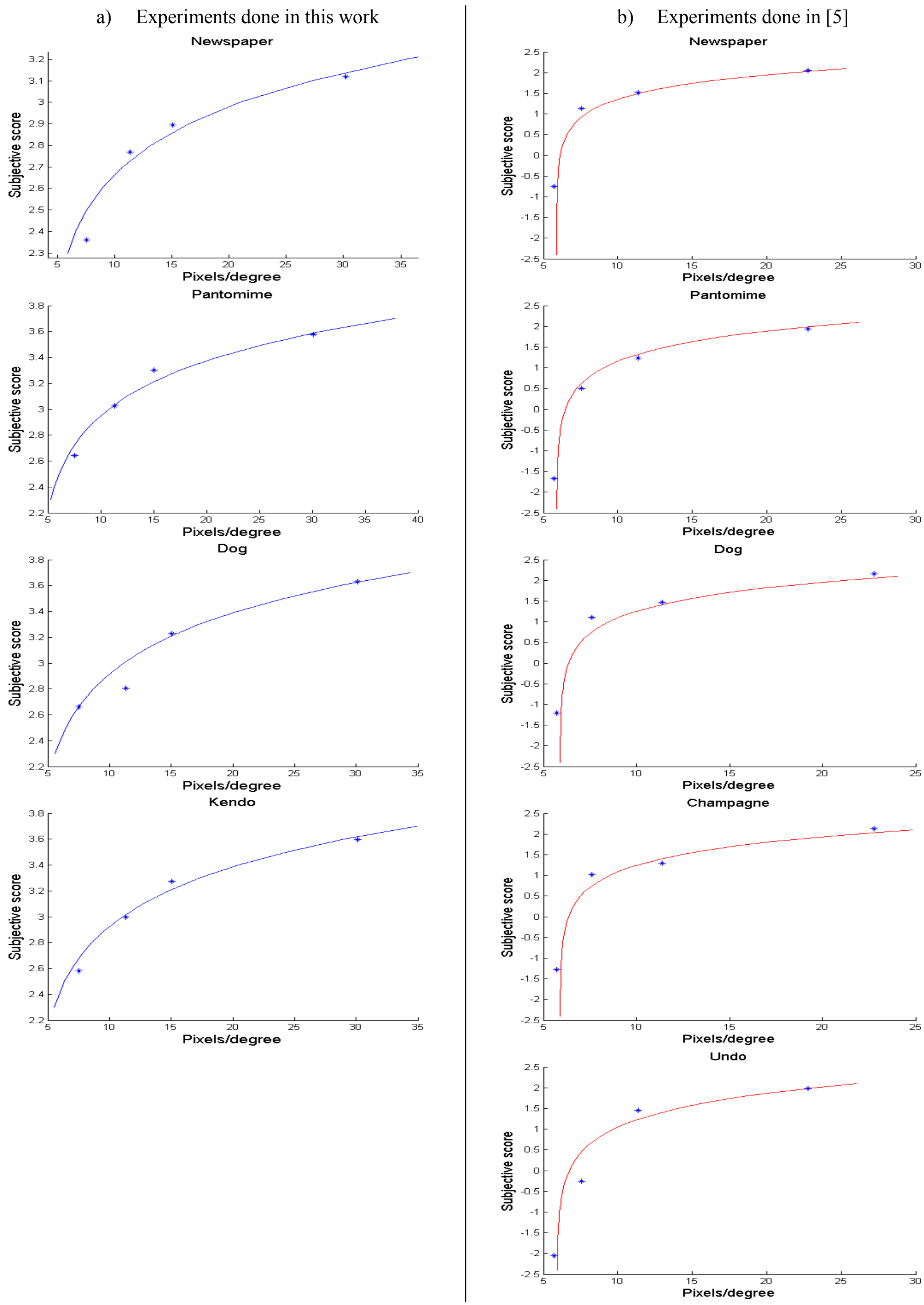


Fig. 2. Relation of the subjective average viewing experience ratings and PPD values

Fig. 2 shows the estimated curves for each of the sequences used in this work and in [5]. The subjectively obvious correlation of the data points and the logarithmic estimates were confirmed by deriving the Pearson correlation coefficients presented in Table 3. Note that as the Pearson correlation measures the linear dependence between two variables, the x-axis of the plots in Fig. 2

should be modified to be log(ppd - k) in order to reflect a correct geometric interpretation of the correlation coefficients in Table 3. On average, the Pearson correlation coefficient between all data points and estimated values among all sequences was 0.97 and 0.98 for tests held in this experiment and [5], respectively.

As the estimation curves turned out to be similar for each test setup, Fig. 3 presents the logarithmic relations estimated in the mean square error sense for all the sequences except Newspaper, whose data points differed significantly from the data points of the other sequences. The other eight test cases fitted the logarithmic estimation very well. The Pearson correlation coefficient between all data points and the joint logarithmic estimation equation is 0.96 for both tests setups in this work and also in [5].

The presented logarithmic equation provided a good estimation of the subjective viewing experience ratings of two different test setups; hence, one could conclude that there might be always a high correlation between MR stereoscopic video subjective scores and the angular resolution of the lower-resolution view measured in PPD. This conclusion should be confirmed by more intensive subjective experiments.

4. CONCLUSIONS

In this work, a set of subjective tests on four asymmetric resolution stereoscopic video sequences was performed. Three different downsampling ratios were applied to the sequences to produce the lower-resolution views. We observed a logarithmic relationship between the subjective viewing experience rating and the angular resolution of the lower-resolution view measured in pixels per degree of viewing angle. The results of the subjective tests presented in this paper and in an earlier work were used to derive two sets of coefficient values for the logarithmic relationship. While the coefficients were remarkably different between the test presented in this paper and the earlier paper, the logarithmic relation provided good estimates of the subjective ratings across all test sequences. Thus, the results suggest that when some subjective evaluations for a few mixed-resolution sequences are available for particular viewing conditions, the proposed logarithmic relation can be used to estimate the subjective rating for other video sequences and downsampling ratios for the lower-resolution view under the same viewing conditions. It is acknowledged that the results should be verified with other video clips and test conditions.

Table 3. Pearson correlation coefficient between actual ratings and estimated values, for all sequences of both test cases

Experiment held in this paper		Experiment held in [5]	
Sequence	Pearson Coef.	Sequence	Pearson Coef.
Dog	0.96	Dog	0.97
Newspaper	0.90	Newspaper	0.98
Pantomime	0.97	Pantomime	0.99
Kendo	0.99	Champagne	0.97
		Undo dancer	0.99

5. REFERENCES

[1] M. G. Perkins, “Data compression of stereopairs,” IEEE Transactions on Communications, vol. 40, no. 4, pp. 684-696, Apr. 1992.

[2] R. Blake, “Threshold conditions for binocular rivalry,” Journal of Experimental Psychology: Human Perception and Performance, vol. 3(2), pp. 251-257, 2001.

[3] H. Brust, A. Smolic, K. Mueller, G. Tech, and T. Wiegand, “Mixed resolution coding of stereoscopic video for mobile devices,” Proc. of 3DTV Conference, May 2009.

[4] L. Stelmach, W. J. Tam, D. Meegan, and A. Vincent, “Stereo image quality: effects of mixed spatio-temporal resolution,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 10, no. 2, pp. 188-193, Mar. 2000.

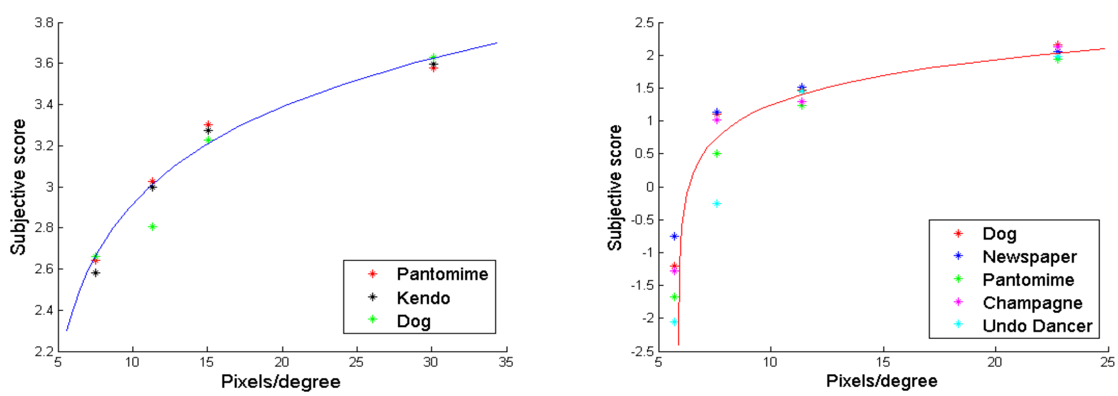
[5] P. Aflaki, M. M. Hannuksela, J. Häkkinen, P. Lindroos, M. Gabbouj , ” Impact of downsampling ratio in mixed-resolution stereoscopic video”, Proc. of 3DTV Conference, June 2010.

[6] JSVM Software, http://ip.hhi.de/imagecom_G1/savce/downloads/SVCReference-Software.htm, Accessed at March 2011.

[7] ITU-T, Subjective assessment methods for image quality in high-definition television, ITU-R BT.710-4.

[8] M. Anderson, R. Motta, S. Chandrasekar, and M. Stokes, “Proposal for a standard default color space for the Internet – sRGB,” Proc. 4th IS and T/SID Color Imaging: Color Science, Systems and Applications, pp. 238-246, Nov. 1996.

[9] P. G. J. Barten, “The effects of picture size and definition on perceived image quality,” IEEE Transactions on Electron Devices, vol. 36, no. 9, pp. 1865-1869, Sep. 1989.



(a) (b)

Fig. 3. Logarithmic relation for (a) tests done in this work (b) tests done in [5]