INTERACTIVE SUBJECTIVE STUDY ON PICTURE-LEVEL JUST NOTICEABLE DIFFERENCE OF COMPRESSED STEREOSCOPIC IMAGES

Chunling Fan[1,2], *Yun Zhang*[1], *Raouf Hamzaoui*[3], *Qingshan Jiang*[1]

[1]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences,
[2]Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences,
[3]School of Engineering and Sustainable Development, De Montfort University

ABSTRACT

The Just Noticeable Difference (JND) reveals the minimum distortion that the Human Visual System (HVS) can perceive. Traditional studies on JND mainly focus on background luminance adaptation and contrast masking. However, the HVS does not perceive visual content based on individual pixels or blocks, but on the entire image. In this work, we conduct an interactive subjective visual quality study on the Picture-level JND (PJND) of compressed stereo images. The study, which involves 48 subjects and 10 stereoscopic images compressed with H.265 intra coding and JPEG2000, includes two parts. In the first part, we determine the minimum distortion that the HVS can perceive against a pristine stereo image. In the second part, we explore the minimum distortion that each subject perceives against a distorted stereo image. Modeling the distribution of the PJND samples as Gaussian, we obtain their complementary cumulative distribution functions, which are known as Satisfied User Ratio (SUR) functions. Statistical analysis results demonstrate that the SUR is highly dependent on the image contents. The HVS is more sensitive to distortion in images with more texture details. The compressed stereoscopic images and the PJND samples are collected in a data set called SIAT-JSSI, which we release to the public.

***Index Terms*—** JND, stereoscopic image, image quality assessment, subjective test.

1. INTRODUCTION

The Just Noticeable Difference (JND) reveals the minimum distortion that the Human Visual System (HVS) can perceive. Traditional JND-based models can be divided into two categories: pixel domain [1] [2] and sub-band domain models [3]. Pixel domain models directly calculate the JND on the visual content and mainly focus on background luminance adaptation and contrast masking [4]. Sub-band domain models [5] [6] [7] first transform the image into sub-bands and then calculate the JND thresholds for the sub-bands. These models usually focus on the pixel or block-based local information within an image. However, HVS does not perceive the visual content based on individual pixels or blocks, but on the entire image. One or more pixels beyond the traditional pixel or sub-band range are not necessarily perceived by the human eyes because of the compound HVS visual effects, such as visual attention and masking effects.

Picture-level JND (PJND) reveals the minimum distortion against the whole image, which can better reflect the real perception of the HVS. When the PJND of an image is determined, an appropriate Quantization Parameter (QP) can be chosen to compress it, which can reduce the bit rate while keeping the same perceptual quality. A number of PJND-based subjective tests were carried out recently. Lin *et al.* [8] measured the PJND of compressed images and videos by subjective tests. Jin *et al.* [9] constructed a larger data set, called MCL-JCI. They found that humans are more sensitive to compression artifacts in regions with semantic objects. Wang *et al.* [10] constructed a PJND video data set called MCL-JCV. Videoset [11] was then generated to measure PJND in videos encoded with H.264. The authors built the Satisfied User Ratio (SUR) curve from the JND samples. These works studied the characteristics of PJND for 2D images and videos and found that there exist a number of PJND points for each image and video. In addition, they provided data sets for objective PJND prediction for 2D images and videos.

However, to the best of our knowledge, PJND for stereo images has not been studied before and there is no stereo image data set based on PJND measurements. With this motivation, we propose an interactive subjective study on PJND of compressed stereoscopic images. The main contributions of this work are as follows.

(1) An interactive subjective quality assessment test was

This work was supported in part by the NSFC under Grant 61871372, Guangdong NSF for Distinguished Young Scholar under Grant 2016A030306022, Guangdong Provincial Science and Technology Development under Grant 2017B010110014, Shenzhen International Collaborative Research Project under Grant GJHZ20170314155404913, Shenzhen Science and Technology Program under Grant JCYJ20170811160212033, Guangdong International Science and Technology Cooperative Research Project under Grant 2018A050506063, Membership of Youth Innovation Promotion Association, CAS under Grant 2018392, and Shenzhen Discipline Construction Project for Urban Computing and Data Intelligence.
Corresponding author:Yun Zhang, Email:yun.zhang@siat.ac.cn.

conducted. A fast binary search procedure was adopted in PJND location.

(2) PJND of stereoscopic images was studied in the subjective test. We not only explored the PJND$_{PRI}$ but also determined the PJND$_{DRI}$. PJND$_{PRI}$ denotes distinguishable quality level against a pristine reference image, and PJND$_{DRI}$ denotes distinguishable quality level against a distorted reference image.

(3) A new PJND data set for stereo images which we call SIAT-JSSI (Shenzhen Institutes of Advanced Technology-JND-based Symmetric Stereoscopic Images) was generated. SIAT-JSSI is released to the public [12].

The findings of PJND in stereo images can be effectively applied to reduce the bit rate in perceptual image and video coding. When compressing with H.265, a smaller QP value often gives a higher quality but increases the bit rate. A larger QP value can reduce the bit rate but provides lower quality. Compressing a pristine image with the PJND$_{PRI}$ can preserve its perceptual quality while reducing the bit rate. When transmitting a distorted image, the compressed version with the minimum bit rate can be chosen while providing the same perceptual quality. Therefore, the PJND$_{DRI}$ can be applied to reduce the bit rate effectively in video streaming.

2. DATA SET CONSTRUCTION

SIAT-JSSI uses ten reference stereoscopic images (left view and right view) in resolution 1920×1080, collected from the SIQA data sets [13] [14]. We used H.265 intra coding [15] with QP ranging from 1 to 51 and JPEG2000 with Compression Ratio (CR) ranging from 1 to 300 (Matlab function "imwrite") to compress the reference stereoscopic images. In addition, FFmpeg-n3.2.8 [16] was used for BMP and YUV color conversion before H.265 intra coding. Therefore, the whole data set contains 3520 stereo image pairs.

To generate a data set which covers diverse image contents, Spatial Information (SI), Colorful Information (CI) [17] [18], and semantic categories were taken into consideration when selecting the reference images. Fig.1 shows the left view of the ten reference images in the proposed data set. Fig. 2 shows the distribution of SI and CF in the source images. We observe that the ten images are dispersed along the horizontal (CF) and vertical (SI) axes. There are seven semantic categories: people, building, plant, fruit, volleyball, computer, and basketball. Five images (People, Treebranches, Flower, Pavilion, and Building) are outdoor and the other five (Basketball, Newsreport, Computer, Volleyball, and Fruitstore) are indoor.

3. SUBJECTIVE TEST ENVIRONMENT AND INTERACTIVE SCORING PROCEDURE

A total of 48 subjects took part in the subjective quality assessment test; most of them were graduate students aged 20 to 35. Among them, five were professionals in the area of image/video quality assessment or coding and the rest were amateurs. Each reference image was assessed by at least 36 subjects. Fig. 3 shows the subjective test environment. All the subjects were seated in a low-light room and wore polarized glasses. According to ITU-R BT.2022 recommendation [19], we set the viewing distance at 1.6 m (twice the image height). The stereo image pairs were showed side by side on a 65' 3D monitor with native resolution 3840×2160. The 3D monitor was used to show stereo image pairs side by side. A 2D monitor was used to show the interface to the subjects.

Before the test, all the subjects were given guidance about the test and those who passed a pre-test were chosen. The subjective test included two parts which aimed to find the PJND$_{PRI}$ and PJND$_{DRI}$, respectively. Each part contained two sessions and the total duration was about 60 minutes with a 5 to 10 minute-break after 30 minutes. A pair of stereo images consists of a reference stereo image and a distorted version. A pair of stereo images were shown side by side for 10 seconds and then the subjects gave their assessment. A "YES" meant a noticeable quality difference and a "NO" meant unnoticeable quality difference. When the image quality continuously degrades, there exists a distortion level (the PJND level) at which the testers change their decisions from "NO" to "YES". When the PJND of a reference image is found, the test proceeds with a new reference image until the PJNDs of all the reference images are found. We adopted the same binary search procedure as in [11] for this PJND location.

4. POST-PROCESSING

Outliers may exist in the subjects and their PJND samples, as some of them may be tired after viewing many images or may have not scored properly. To keep the subjective test results reliable, outlier samples were detected and deleted.

As in [11], we detected the outlier subjects according to the standard score. A subject was identified as an outlier, if the range R and standard deviation σ of its standard score were both larger than a threshold (set to 3). All the collected samples from outlier subjects were removed. Five subjects from 36 (i.e., 13.89%) were removed for JPEG2000 and H.265. Hence, most of the subjects in our subjective quality assessment test were found to be reliable.

Grubbs' test [20] was then used for the detection of outlier samples [11]. One outlier was detected at a time until there was no outlier. Four samples from 320 (i.e., 1.25%) were deleted for JPEG2000 and fourteen samples from 280 (i.e., 5.00%) for H.265, respectively. Therefore, most of the samples were reliable.

In addition, according to ITU-R BT.500 [21], the β_2 test was used to check whether the distribution of samples is "normal". If the kurtosis K was within the range [2, 4], the distribution of the sample for a reference image was considered to be Gaussian. For PJND$_{PRI}$, the pass percentage of normality

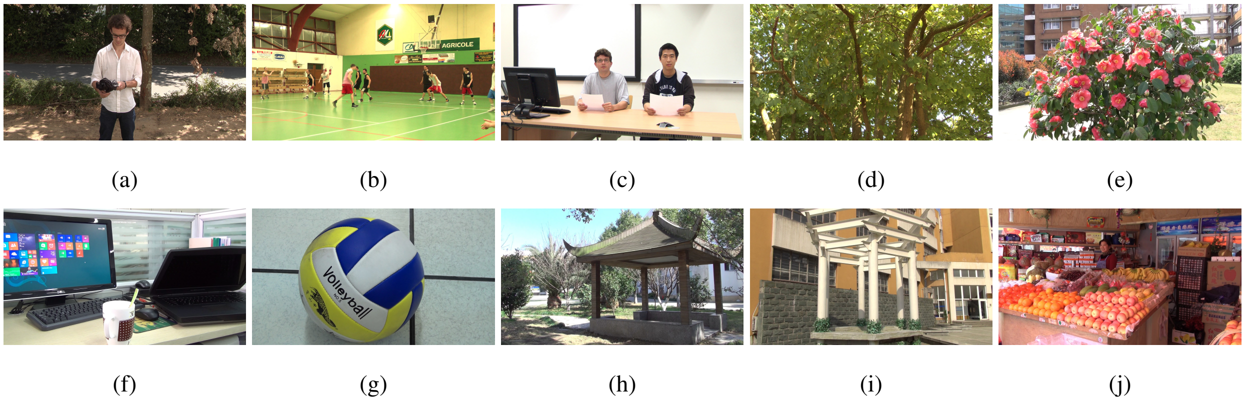


Fig. 1: Left view of source images in the SIAT-JSSI data set. (a) People. (b) Basketball. (c) Newsreport. (d) Treebranches. (e) Flower. (f) Computer. (g) Volleyball. (h) Pavilion. (i) Building. (j) Fruitstore.

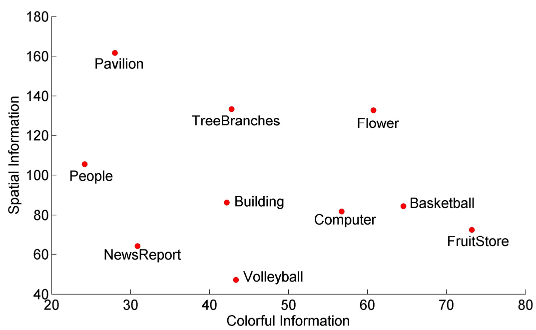


Fig. 2: Distribution of SI and CF in the 10 source images of the SIAT-JSSI data set.

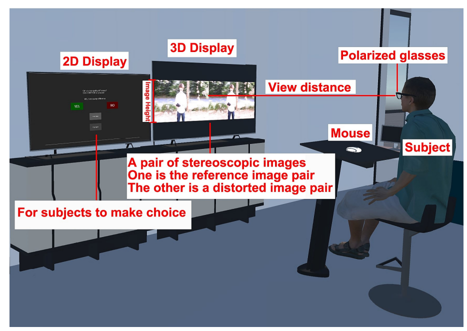


Fig. 3: Interactive subjective assessment environment.

detection was 90% for JPEG2000, and 100% for H.265. For PJND$_{DRI}$, the percentage was 100% for JPEG2000, and 90% for H.265. Most samples passed the normality test.

5. EXPERIMENTAL RESULTS AND DISCUSSION

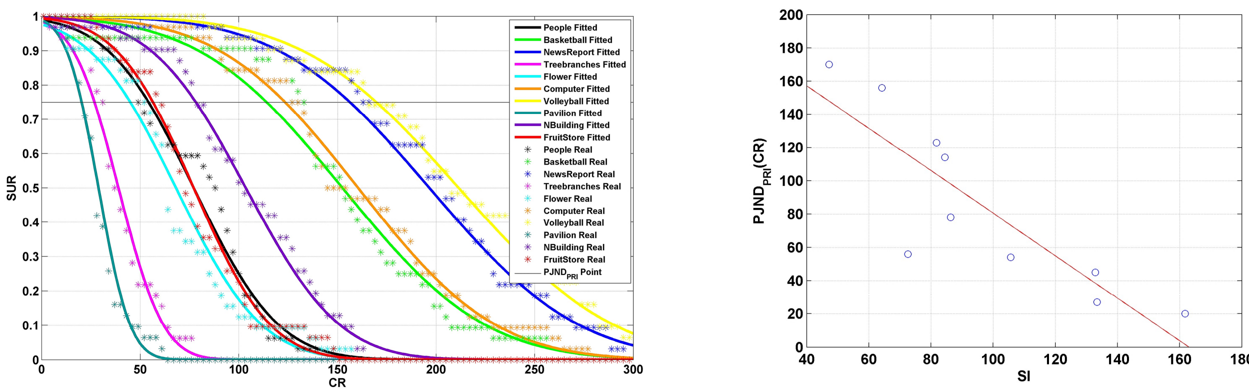
A subject that cannot perceive the distortion is said to be satisfied with the quality. The SUR S_n for a distorted image d_n of a source image can be obtained by calculating

$$S_n = 1 - \frac{1}{M} \sum_{m=1}^{M} \Phi_m(d_n), \tag{1}$$

where M is the number of subjects and $\Phi_m(d_n) = 1$ if subject m can perceive a quality difference between the reference image and the distorted image d_n and $\Phi_m(d_n) = 0$; otherwise.

Because the collected samples passed the normality test (see Section 3), we modeled them according to a normal distribution. Consequently, the SUR can be obtained as the complementary cumulative distribution function of a normal distribution. Fig.4a shows the SUR curves of all the reference images in the data set for JPEG2000 compression. The dots are real data and the curves are fitted based on the normal distribution. The intersection points between the fitting curves and the horizontal line 75% is the PJND point for the reference image. It gives the distortion level at which 75% of the subjects are satisfied with the corresponding visual quality. When the SUR is larger than 75% more than 75% of the subjects are satisfied with the quality of the compressed stereoscopic image. When the SUR is smaller than 75% fewer than 75% of the subjects are satisfied with the quality. Actually, we can set different thresholds, e.g., 85%, for the SUR curve based on the quality requirements of the specific application. We can observe that 1) the SUR decreases as the CR of JPEG2000 increases, which means increased distortion will degrade users' satisfaction with the image quality; 2) the real data fits well the normal distribution; 3) the PJND points and the slopes of the SUR curve vary with the images, which indicates the SUR is highly dependent on the image contents. Similar results were obtained for H.265 intra coding.

Table 1 presents the PJND (QP/CR) and the corresponding Peak Signal-to-Noise-Ratio (PSNR) (for both the left view and right view) for each source stereoscopic image under H.265 intra coding and JPEG2000 compression. PSNR, QP, and CR were used to assess the quality of the left view



(a)

(b)

Fig. 4: SUR curves and relationship between PJND$_{PRI}$ and SI. (a) SUR curves of PJND$_{PRI}$ for JPEG2000.(b) Relationship between PJND$_{PRI}$ and SI of source images for JPEG2000, R^2= 16.4258.

Table 1: PJND$_{PRI}$ and PJND$_{DRI}$ values for H.265 and JPEG2000 compression in SIAT-JSSI data set.

Image	H.265				JPEG2000			
	PJND$_{PRI}$		PJND$_{DRI}$		PJND$_{PRI}$		PJND$_{DRI}$	
	QP	PSNR(L/R)	QP	PSNR(L/R)	CR	PSNR(L/R)	CR	PSNR(L/R)
People	28	35.48/35.46	36	30.86/30.87	54	28.46/28.56	148	27.96/28.10
Basketball	28	40.37/40.28	37	35.34/35.17	114	33.66/33.28	180	31.41/31.17
NewsReport	33	38.78/38.60	35	37.96/37.74	156	38.45/37.96	193	36.78/36.42
Treebranches	27	36.20/36.21	36	30.51/30.50	27	26.25/26.27	137	25.31/25.52
Flower	33	33.84/34.76	37	31.21/31.24	45	26.77/28.62	145	26.11/27.84
Computer	28	39.64/40.80	36	35.72/36.72	123	34.60/36.35	174	33.25/34.86
Volleyball	30	38.93/38.34	36	36.43/35.06	170	37.78/35.83	201	36.89/35.28
Pavilion	26	37.89/38.10	34	31.91/30.09	20	25.64/26.25	143	25.34/25.97
Building	30	38.11/38.69	35	35.14/35.70	78	33.43/34.39	150	32.81/33.80
FruitStore	25	40.32/40.66	35	34.38/34.70	56	32.27/32.89	146	31.30/31.84

and right view respectively. PJND$_{PRI}$ was obtained when using the pristine stereo image as a reference and PJND$_{DRI}$ was acquired when using a distorted stereo image as a reference. We observe that the PSNR of the distorted image corresponding to the PJND$_{PRI}$ for H.265 ranges from 33.84 dB to 40.80 dB, with QP ranging from 25 to 33. This means that most people (more than 75%) cannot perceive the quality difference compared to the pristine image if the PSNR is higher or QP is lower than PJND$_{PRI}$. Thus, the PJND$_{PRI}$ can be used in image communication to minimize the bit rate for the same perceptual quality as the source image. The PJND$_{DRI}$ means over 75% of the subjects cannot perceive the quality difference between PJND$_{PRI}$ and PJND$_{DRI}$. Similar results were obtained for JPEG2000 compression.

To further study the relationship between PJND and image contents, Fig.4b shows the relationship between PJND (CR) and SI. SI is an indicator of texture complexity, and larger SI means more texture details. CR denotes the compression ratio, and an image with smaller CR has higher quality with higher bit rate. Fig.4b shows the relationship between SI and PJND$_{PRI}$ under JPEG2000 compression, where the Y-axis is the PJND$_{PRI}$ measured with CR. We observe that the PJND$_{PRI}$ (CR) decreases as the SI increases, which indicates that human eyes are more sensitive to compression distortions in images with more texture details.

6. CONCLUSIONS

We proposed an interactive subjective visual quality assessment test to generate a PJND-based stereoscopic image data set. We called the data set SIAT-JSSI and released it to the public. SIAT-JSSI gives the PJND$_{PRI}$ and PJND$_{DRI}$ for a set of 10 stereoscopic images compressed with H.265-based intra coding and JPEG2000. Statistical analysis demonstrated that the PJND points and the slopes of the SUR curve change with the images, which indicates that the SUR depends on the image contents. The human visual system is more sensitive to distortion in images with more texture details. Our results can be used to reduce the bit rate for stereoscopic images without affecting the perceptual quality.

7. REFERENCES

[1] A. Liu, W. Lin, M. Paul, C. Deng, and F. Zhang, "Just noticeable difference for images with decomposition model for separating edge and textured regions," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 20, no. 11, pp. 1648–1652, 2010.

[2] Y. Zhao, Z. Chen, C. Zhu, Y. Tan, and L. Yu, "Binocular just-noticeable-difference model for stereoscopic images," *IEEE Signal Process. Lett.*, vol. 18, no. 1, pp. 19–22, 2011.

[3] Z. Wei and K. Ngan, "Spatio-temporal just noticeable distortion profile for grey scale image/video in dct domain," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 19, no. 3, pp. 337–346, 2009.

[4] S. Wang, L. Ma, Y. Fang, W. Lin, S. Ma, and W. Gao, "Just noticeable difference estimation for screen content images," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3838–3851, 2016.

[5] J. Wu, G. Shi, W. Lin, A. Liu, and F. Qi, "Just noticeable difference estimation for images with free-energy principle," *IEEE Trans. Multi.*, vol. 15, no. 7, pp. 1705–1710, 2013.

[6] J. Wu, L. Li, W. Dong, G. Shi, W. Lin, and C.-C. J. Kuo, "Enhanced just noticeable difference model for images with pattern complexity," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2682–2693, 2017.

[7] D.V.S.X. De Silva., W. Fernando, S. Worrall, S. Yasakethu, and A. Kondoz, "Just noticeable difference in depth model for stereoscopic 3d displays," in *Multimedia and Expo (ICME), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1219–1224.

[8] J. Lin, L. Jin, S. Hu, I. Katsavounidis, Z. Li, A. Aaron, and C.-C. J. Kuo, "Experimental design and analysis of jnd test on coded image/video," in *Applications of Digital Image Processing XXXVIII*. International Society for Optics and Photonics, 2015, vol. 9599, p. 95990Z.

[9] L. Jin, J. Lin, S. Hu, H. Wang, P. Wang, I. Katsavounidis, A. Aaron, and C.-C. J. Kuo, "Statistical study on perceived jpeg image quality via mcl-jci dataset construction and analysis," *J. Electronic Imaging*, vol. 2016, no. 13, pp. 1–9, 2016.

[10] H. Wang, W. Gan, S. Hu, J. Lin, L. Jin, L. Song, P. Wang, I. Katsavounidis, A. Aaron, and C.-C. J. Kuo, "Mcl-jcv: a jnd-based h. 264/avc video quality assessment dataset," in *Image Processing (ICIP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1509–1513.

[11] H. Wang, I. Katsavounidis, J. Zhou, J. Park, S. Lei, X. Zhou, M. Pun, X. Jin, R. Wang, X. Wang, Y. Zhang, J. Huang, S. Kwong, and C.-C. J. Kuo, "Videoset: A large-scale compressed video quality dataset based on jnd measurement," *J. Visual Commun. Image Represent.*, vol. 46, pp. 292–302, 2017.

[12] "SIAT-JSSI," http://codec.siat.ac.cn/SIAT-JSSI/index.html, 2018.

[13] M. Urvoy, M. Barkowsky, R. Cousseau, Y. Koudota, V. Ricorde, P. Le Callet, J. Gutierrez, and N. Garcia, "Nama3ds1-cospad1: Subjective video quality assessment database on coding conditions introducing freely available high quality 3d stereoscopic sequences," in *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*. IEEE, 2012, pp. 109–114.

[14] Q. Jiang, F. Shao, G. Jiang, M. Yu, and Z. Peng, "Three-dimensional visual comfort assessment via preference learning," *J. Electron. Imaging*, vol. 24, no. 4, pp. 043002, 2015.

[15] F. Bossen, D. Flynn, and K. Suhring, "Hm software manual," *Document: JCTVC-M1010*, 2013.

[16] "FFmpeg-n3.2.8," http://ffmpeg.org/ffmpeg-all.html.

[17] S. Winkler, "Analysis of public image and video databases for quality assessment," *IEEE J. Sel. Top. Signa. Proc.*, vol. 6, no. 6, pp. 616–625, 2012.

[18] ITU-T RECOMMENDATION P.910, "Subjective video quality assessment methods for multimedia applications," 2008.

[19] RECOMMENDATION ITU-R BT.2022, "General viewing conditions for subjective assessment of quality of sdtv and hdtv television pictures on flat panel displays," 2012.

[20] F. Grubbs, "Sample criteria for testing outlying observations," *The Annals of Mathematical Statistics*, vol. 21, no. 1, pp. 27–58, 1950.

[21] RECOMMENDATION ITU-R BT.500-11, "Methodology for the subjective assessment of the quality of television pictures," 2002.