# JUST NOTICEABLE DIFFERENCE MODEL FOR ASYMMETRICALLY DISTORTED STEREOSCOPIC IMAGES

*Yu Fan* [1,2], *Mohamed-Chaker Larabi* [1], *Faouzi Alaya Cheikh* [2], *Christine Fernandez-Maloigne* [1]

[1] CNRS, Univ. Poitiers, XLIM, UMR 7252, France
[2] Faculty of Computer Science and Media Technology, NTNU, Gjovik, Norway

## ABSTRACT

In this paper, we propose a saliency-weighted stereoscopic JND (SSJND) model constructed based on psychophysical experiments, accounting for binocular disparity and spatial masking effects of the human visual system (HVS). Specifically, a disparity-aware binocular JND model is first developed using psychophysical data, and then is employed to estimate the JND threshold for non-occluded pixel (NOP). In addition, to derive a reliable 3D-JND prediction, we determine the visibility threshold for occluded pixel (OP) by including a robust 2D-JND model. Finally, SSJND thresholds of one view are obtained by weighting the resulting JND for NOP and OP with their visual saliency. Based on subjective experiments, we demonstrate that the proposed model outperforms the other 3D-JND models in terms of perceptual quality at the same noise level.

***Index Terms*—** Just noticeable difference, 3D image/video coding, quality assessment, spatial masking, visual saliency.

## 1. INTRODUCTION

The just noticeable difference (JND) is one of the most important perceptual properties, referring to the minimum visibility threshold below which the pixel intensity variations cannot be perceived by the human visual system (HVS). For decades, the 2D-JND models have been successfully developed and exploited in many applications [1]. However, their use for S3D applications is questionable. They mostly rely on monocular vision, which does not fit with the complexity of our 3D perception requiring specific models accounting for both monocular and binocular depth cues.

Accordingly, it becomes crucial to develop effective 3D-JND models for perceptual improvement of 3D applications. So far, a handful of 3D-JND models can be found in the literature [2–12]. Based on the S3D content format, the existing 3D-JND models are classified into two categories: (1) texture-plus-depth-based models, and (2) stereopair-based models.

The first category estimates the visibility thresholds using either texture-plus-depth content [2,3,11,12], or multi-view video plus depth (MVD) one [6–8]. For instance, De Silva *et al.* [2] propose a JND in depth (JNDD) model which measures the threshold for depth variation that a human can perceive on a 3D display. Similarly, to avoid the impact of the monocular depth cues, Yang *et al.* [11] conduct psychophysical experiments (PEs) based on the dynamic Random Dot Stereogram technique to measure the JNDD thresholds. In a different vein, Lian *et al.* design a JND in multi-view (MJND) model, specially for MVD, by combining spatial and temporal JND with JNDD [6]. Likewise, Zhong *et al.* [8] propose a hybrid JND (HJND) model integrating a 2D-JND model [13] together with depth saliency.

The second category models [4,5,9,10] are developed using left and right views of S3D images. For example, based on PEs, Zhao *et al.* develop a binocular JND (BJND) model [4], which estimates the visibility thresholds in inter-difference between the left and right views, by modeling visual masking effects. Here, the binocular disparity is not taken into account, making the model less reliable for real-world images. To solve this issue, Kim *et al.* conduct PEs to measure JND thresholds by considering both luminance adaptation (LA) and binocular disparity effects [14]. Meanwhile, a joint JND (JJND) model [5] is proposed on top of a 2D-JND model [13], relying on the assumption that the HVS has different perceptions on objects with different depth values. Although JJND accounts for binocular depth cues, its performance is low for S3D images with uniform depth maps. Recently, Xue *et al.* [10] propose a disparity-based JND (DJND) model by combining the JND profile [15] with both depth of focus blur and disparity information. DJND is less efficient for S3D images with a limited depth difference between foreground and background regions. In conclusion, all the above second category models are developed on top of existing 2D-JND models instead of conducting PEs, except BJND and [14].

In this paper, we propose a saliency-weighted stereoscopic JND (SSJND) model that belongs to the second category, based on our findings obtained from PEs. Our model is two-fold: 1) a disparity-aware binocular JND (DBJND) dedicated to non-occluded pixels (NOPs) obtained from LA and contrast masking (CM) experiments accounting for binocular perception, and 2) a 2D-JND model devoted to occluded pixels (OPs) in the stereo pair. A final step of the proposed SSJND model consists of weighting the JND thresholds by the pixel visual saliency to account for its modulator effect. The obtained model is validated thanks to subjective experiments and compared in terms of perceptual 3D image quality to a number of 3D-JND models from the literature.

## 2. PSYCHOPHYSICAL EXPERIMENTS

According to [16], the HVS is able to quickly adjust to the level of the background light in order to distinguish objects. This ability is known as luminance adaptation (LA). Furthermore, contrast masking (CM) describes the masking effect of the HVS in presence of two or more stimuli, if they are of similar contrast/spatial nonuniformity [1]. With the aim to model LA and CM in the S3D context by considering the binocular disparity, we designed two comprehensive PEs.

### 2.1. Stimuli

Fig. 1 illustrates the visual stimuli used in LA and CM experiments, respectively. The difference between $d_l$ and $d_r$ denotes the binocular disparity $d$. The peri-fovea is modeled by a region $R_1$ with a fixed

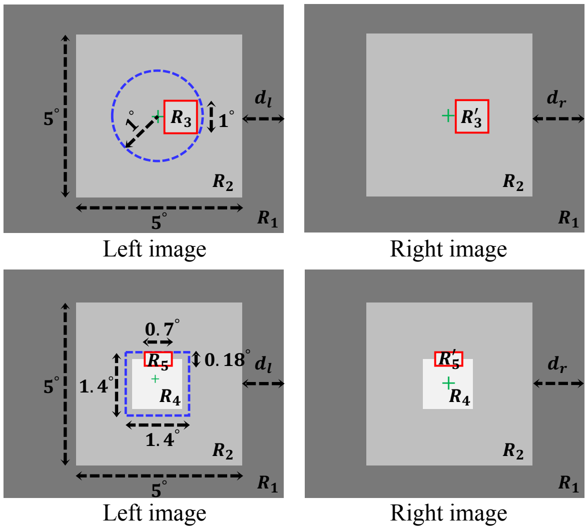


**Fig. 1**: Stereo pair patterns used in psychophysical experiments.

luminance level 72 pixels (px). The human retinal para-fovea and fovea can cover the information within 5° and 2° of visual angles, respectively, around the fixation point [17]. Consequently, our stimuli in LA/CM experiments contain a fixation cross and a square $R_2$ with a visual field of $5^\circ \times 5^\circ$ with a luminance level equal to $L_b$.

#### 2.1.1. *LA experiment*

The fovea-covered region is represented by a dashed circle of 2°. In contrast to [4] and [14], the noise area $R_3/R_3^{'}$ is randomly displayed within the dashed circle so as to avoid the memorization of noise locationwhich may underestimate the JND thresholds. Furthermore, the luminance levels are set to $L_b \pm N_l$ ($R_3$) and $L_b \pm N_r$ ($R_3^{'}$) with $N_{l|r}$ the noise amplitude injected in the left/right view.

#### 2.1.2. *CM experiment*

The fovea-covered region is shown here by a $1.4^\circ \times 1.4^\circ$ dashed square (diagonal of 2°). The noise area $R_5/R_5^{'}$ is located on a randomly chosen side of $R_4/R_4^{'}$ perimeter with an intensity of $N_{l|r}$. Besides, the luminance level of $R_4$ is set to $L_b - \Delta L$, where $\Delta L$ denotes the luminance contrast between $R_2$ and $R_4$.

Considering the Percival's zone of comfort [18] and the experiments' duration, we choose five disparity values (i.e., $0^\circ, \pm 0.5^\circ, \pm 1^\circ$) after several trials. Table 1 describes the attributes values of the stimuli used in LA and CM experiments. We set $N_l = 0$ for LA experiment to obtain the maximum visibility thresholds of the right image. In total, there are 30 stimuli (6 luminance levels × 5 disparities) in LA experiment, and 60 stimuli (3 luminance levels × 5 disparities × 2 contrast values × 2 noise amplitude levels) in CM experiment.

**Table 1**: Stimulus attributes for LA and CM experiments.

| Attribute | LA | CM |
|---|---|---|
| Noise amplitude $N_l$ (px) | 0 | 0, 2 |
| Luminance contrast $\Delta L$ (px) | – | 16, 48 |
| Background luminance $L_b$ (px) | 22, 32, 48<br>96, 144, 192 | 96, 144, 192 |
| Disparity $d$ (degree) | -1, -0.5, 0, 0.5, 1 | |

### 2.2. Subjects

Twenty-two subjects (ages ranging from 20 to 33) are invited for both LA and CM experiments. Before the experiments, each subject undergoes a visual acuity check based on the Freiburg Vision Test, in addition to the Randot stereo test.

### 2.3. Apparatus

The experiments are conducted in the XLIM psychophysical test room that is isolated from the outside diffuse light and noise. The ambient illumination is adjusted to 65 lux measured by an illuminance-meter. To display the 3D test images, we use a calibrated 46" Hyundai TriDef S465D monitor having HD (1920 × 1080) resolution with a brightness set to 250 $cd/m^2$. Polarized 3D glasses are used. According to the ITU-R BT.2021-1 recommendations [19], the viewing distance between the subject and the monitor is set to 1.7 m (approx. 3× the height of the display).

### 2.4. Procedure

The experiments are designed using the Psychtoolbox of Matlab [20]. Each subject is informed about the purpose of the experiments, and instructed on how to report the results by using the keyboard thanks to a training sequence before the actual experiments. The JND threshold of the right view is obtained in two steps according to [21]. Step 1 determines the just noticeable noise of the right view $A_{JNN}$, whereas step 2 measures the just unnoticeable noise $A_{JUN}$. The noise amplitude of the right view is varying, while the left view remains constant in order to generate an asymmetric noise.

In step 1, for a stimulus, the noise amplitude of the right image $N_r$ is initially set to 0 to make it invisible to subjects. Then, $N_r$ is increased with a step of $A_s$ until it becomes just noticeable, and the final value is saved as the subject's $A_{JNN}$. $A_s$ was set to 0.0083 and 0.1 for LA and CM experiments, respectively. Subsequently, $N_r$ is increased to $A_{JNN} + A$ immediately to ensure that subjects can easily detect the noise. $A$ is set to 1.7 and 2.0 for LA and CM, respectively.

In step 2, the subjects follow a reversed procedure. Initially, the noise area is visible to subjects. Then, $N_r$ is gradually decreased from $A_{JNN} + A$ by a level of $A_s$ until the noise becomes just unnoticeable. The corresponding value is saved as the subject's $A_{JUN}$. The JND threshold of the right view is finally obtained as the average of $A_{JNN}$ and $A_{JUN}$. The procedure is repeated for the whole set of stimuli and subjects are asked to take a rest every 15 minutes.

## 3. PSYCHOPHYSICAL DATA ANALYSIS AND MODELING

### 3.1. Data analysis

To derive a reliable 3D-JND model, we perform an outlier detection [22]. To do so, subject's responses screening is performed following the ITU-R BT 1788 recommendations [23]. The decision criterion is based on the correlation level between subject's values and the mean observations. Consequently, four subjects for LA and three for CM are identified as outliers, and discarded for the further analysis.

With the aim to obtain consistent data for each subject, we proceed to the rejection of outlier observations for each subject [24]. The median-absolute-deviation method is used for LA experimental data, because the distribution for each subject is approximately symmetric. At the opposite, the samples distribution for CM experimental data is mostly asymmetric for which the Tukey's-fences method is preferred. In addition, to confirm the reliability of the JND data after outliers' rejection, we adopt the Jarque-Bera test [25] to verify that all JND values of each stimulus follow a normal distribution

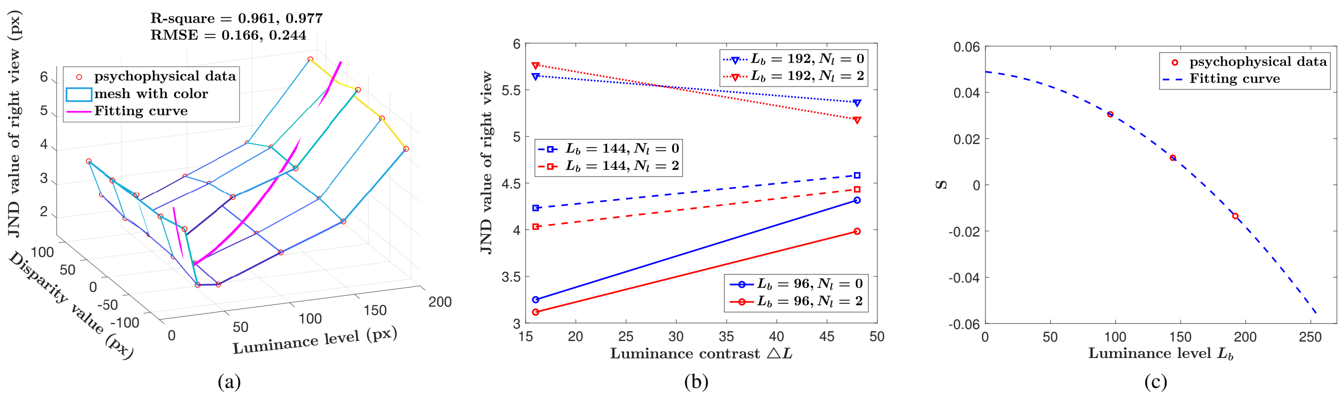


**Fig. 2**: (a) JND thresholds for difference background luminance levels $L_b$ and disparities $d$ from LA experiment, (b) JND thresholds for difference $L_b$ and noise amplitudes of the left view $N_l$ from CM experiment. (c) Average slopes of the two curves in (b) for each $L_b$.

($p - value > 0.05$). Finally, the mean JND threshold is obtained for each stimulus using the post-processed JND data.

To further investigate the effects of background luminance $L_b$ and disparity $d$ on the JND values, we conduct a two-way analysis of variance (ANOVA) with the null hypothesis of no statistical significant difference between JND thresholds for different $L_b$ and $d$. It is worth noting that the effects of $\Delta L$ and $N_l$ are not exploited, because both of them have only two values (see Table 1). Before ANOVA, we first validate the normality of the distributions with the Shapiro-Wilk test [26] and the homogeneity of variances with the Levene's test [27].

The resulting $F(1,6) = 290.26$, $p < 0.001$ for LA, and $F(1,3) = 90.01$, $p < 0.001$ for CM demonstrate that there is a significant difference between the luminance levels in terms of JND thresholds. Furthermore, for the binocular disparity, the analysis indicates a significant effect for LA ($F(1,4) = 2.95$, $p = 0.04$) and no effect for CM ($F(1,4) = 0.56$, $p = 0.69$). This is probably caused by the influence of the luminance contrast and the left view noise on JND threshold than by disparity in the complicated CM experiment patterns.

### 3.2. 3D-JND modeling

In this section, the post-processed JND data from the conducted experiments are used to derive a 3D-JND model by considering both LA and CM effects, as well as the disparity. Based on the study in [4], the BJND model serves as a framework for our proposed model. Therefore, using $L_b$, $\Delta L$, $N_l$ and $d$ (*cf.* Table 1), we define a disparity-aware binocular JND threshold of the right image DBJND$_r$ as:

$$DBJND_r = T_{r_{max}}(L_b, \Delta L, d)\left[1 - (\frac{N_l}{T_{r_{max}}(L_b, \Delta L, d)})^{\lambda}\right]^{\frac{1}{\lambda}}, \tag{1}$$

with $\lambda$ a parameter that controls the influence of $N_l$, and its estimation will be discussed later. In additions, $T_{r_{max}}$ denotes the maximum JND threshold of the right image by considering both LA and CM effects, and is calculated as follows:

$$T_{r_{max}} = S(L_b)\Delta L + T^{'}_{r_{max}}(L_b, d), \tag{2}$$

where $T^{'}_{r_{max}}$ is the LA JND threshold for $N_l = 0$. Fitting the data of Fig. 2a requires a curve having two distinct intervals: one for $L_b \leq 48$ and the other for $L_b \geq 48$. $L_c$ represents the intersection point between these two curves, and is equal to 33. As presented in the top of Fig. 2a, the values of R-square and the root mean square error (RMSE) indicate a good fitting. Hence, for different $L_b$ and $d$, $T^{'}_{r_{max}}$ can be expressed as:

$$T^{'}_{r_{max}} = \begin{cases} c_1(L_b^2 + c_2 L_b + c_3 d) + c_4, & L_b \in [0, L_c[ \\ c_5(L_b^2 + c_6 L_b + c_7 d) + c_8, & L_b \in ]L_c, 255] \end{cases} \tag{3}$$

where the damped least-square fitting method [28] used on LA experimental data allows to identify the different constants as $c_1 = 0.0043$, $c_2 = 83.939$, $c_3 = 0.344$, $c_4 = 9.611$, $c_5 = 0.0001$, $c_6 = 57.884$, $c_7 = 2.333$, and $c_8 = 2.536$.

Moreover, to determine $S(L_b)$ in (2), we first depict the average JND values (for five disparity values) according to $\Delta L$ under different $L_b$ and $N_l$ in Fig. 2b. It illustrates that the JND threshold of the right image increases as the luminance level increases. Furthermore, the JND threshold is inversely proportional to the amplitude of the noise injected in the left image under the same $L_b$, except for the case where $L_b = 192$. This is because high luminance intensity in CM experiment may result in subjects' misjudgment on the visibility thresholds. The slopes of the two curves for each $L_b$ are determined, and are averaged as $S$ in Eq.2. Fig. 2c shows the relation between $S$ and $L_b$ based on the obtained CM data, and its corresponding fitting function is modeled by:

$$S = c_9(L_b^2 + c_{10} L_b) + c_{11}, \tag{4}$$

where the fitting parameters $c_9$, $c_{10}$ and $c_{11}$ are equal to $-1.389 \times 10^{-6}$, 30.238 and 0.049, respectively. The disparity $d$ in Eq. 4 is not considered because of the lack of effect on CM JND values (see Section 3.1). As a result, we estimate $\lambda$ described in Eq. 1 by fitting the JND values for $N_l = 0$ and $N_l = 2$, and obtain $\lambda = 3.76$ with RMSE $= 0.421$.

In addition to the above effects, we consider the occlusions for 3D-JND modeling. To this end, image pixels are classified into non-occluded (NOP) and occluded (OP) pixels based on [29]. Then, DBJND (Eq.1) is applied to NOP and a robust 2D-JND model [30] is applied to OP. Besides, the studies in [31, 32] demonstrate that JND thresholds are affected by the visual importance of objects in the image, *i.e.*, visual saliency (VS). Specifically, the salient regions, which attract more visual attention, have lower visibility thresholds than the non-salient ones. Thereby, we propose to employ a VS map to weight different JND estimates for NOPs and OPs. The VS of the S3D image is estimated using a promising 3D saliency detection algorithm [33]. Finally, the proposed saliency-weighted stereo JND

(SSJND) model is defined as:

$$SSJND_{l|r}(k) = \begin{cases} T_{l|r}(k)(1+\alpha(T_s - \bar{S}_{l|r}(k))), \bar{S}_{l|r}(k) \in [0, T_s] \\ T_{l|r}(k)(1-\alpha(\bar{S}_{l|r}(k) - T_s)), \bar{S}_{l|r}(k) \in ]T_s, 1] \end{cases} \quad (5)$$

where $l \mid r$ refer to the left or right image, $k$ is the $k^{th}$ pixel of the image. $T_{l|r}$ respectively corresponds to $\text{DBJND}_{l|r}$ for NOPs and $JND_{l|r}$ for OPs. $\bar{S}$ represents the visual saliency normalized in the range of $[0, 1]$. In addition, the parameters $Ts$ and $\alpha$, bounded in [0, 1], control the impact of VS on $SSJND$. For the next section, we set $Ts = 0.5$, and $\alpha = 0.6$.

## 4. EXPERIMENTAL VALIDATION

In this section, we validate the performance of the proposed SSJND model by comparing with three very recent 3D-JND models, *i.e.*, BJND [4], JJND [5] and DJND [10], as well as the SSJND model without considering saliency (DBJND).

To achieve this, we use twelve stereo pairs from the Middlebury stereo datasets [34]. Similar to [35] and [36], we compare the perceptual quality between the noise-injected S3D images relying on different 3D-JND models under the same noise level. Note that the noise is injected only in the right image of the stereo pair in order to simulate an asymmetric distortion. The S3D image $I^*$ contaminated by the JND-based noise is calculated as: $I^*(k) = I(k) + C_n \cdot N_{rand}(k) \cdot JND(k)$, where $I$ denotes the original image. $C_n$ is a control parameter that makes the same noise level for different 3D-JND models leading to the same peak signal-to-noise ratio (PSNR), *i.e.* PSNR $\in$ [28dB, 29dB].

To subjectively compare our model to the state-of-the-art, we use the same experimental setup as for previous PEs. The room ambient illumination and the viewing distance are set to 100 lux and 1.8 m, respectively. Furthermore, eighteen subjects are invited to participate the test. Note that two subjects (side-by-side) participate to the test simultaneously while the influence of viewing direction on the quality judgment will be investigated later. we opted for the stimulus-comparison method described in the ITU-R BT.2021-1 [19]. Firstly, a mid-grey image with zero disparity, containing the image sequence number, is presented to the subjects for 2s. Then, a couple of JND-based distorted 3D images (SSJND and SOTA model) are shown with random position on a mid-gray background for 10s. Subsequently, subjects are asked to provide a score depending on the preference: 0 (the same), 1 (slightly better), 2 (better), 3 (much better). These scores are then used to compute the mean opinion score over all subjects for each S3D image. In addition, we use the Pearson's chi-squared test [37] to verify the statistical significance of the comparative scores. The adopted null hypothesis of this test is: "there is no preference between the proposed SSJND model and the other 3D-JND models".

Table 2 shows the quality comparison results in terms of mean opinion scores and p-values for each image. $p-value < 0.05$ for all pair comparison cases rejects the null hypothesis, and thus validates the statistical significant preference between the proposed model and the other 3D-JND models. Overall, SSJND outperforms all the other models on almost all the used images. Complex scenes may lead to difficulties in VS estimation where SSJND may overestimate the JND thresholds for smooth regions with high luminance intensity when the latter regions are considered as non-salient.

Compared to the BJND, the proposed SSJND model considers occlusion effect, and thus globally provides better estimation for S3D image containing large number of occluded pixels. In the same vein, our model performs quite better than the JJND and DJND models in terms of average scores, because they are both developed based

**Table 2**: Quality comparison between our SSJND and state-of-the-art models using 12 images from the Middlebury stereo datasets.

| S3D | vs.DBJND | | vs. BJND [4] | | vs. JJND [5] | | vs. DJND [10] | |
|---|---|---|---|---|---|---|---|---|
| image | $\bar{M}$ | p-value | $\bar{M}$ | p-value | $\bar{M}$ | p-value | $\bar{M}$ | p-value |
| Art | 0.39 | 0.0001 | 0.06 | 0.0001 | 1.44 | 0.0058 | 1.61 | 0.0015 |
| Reindeer | 0.72 | 0.0001 | 0.33 | 0.0001 | 2.56 | 0.0001 | 2.89 | 0.0001 |
| Moebius | 0.39 | 0.0001 | -0.17 | 0.0001 | 2.06 | 0.0001 | 2.22 | 0.0001 |
| Dolls | 0.72 | 0.0002 | 0.50 | 0.0001 | 1.72 | 0.0001 | 0.94 | 0.0001 |
| Aloe | 0.39 | 0.0001 | 0.83 | 0.0006 | 0.78 | 0.0016 | 1.22 | 0.0027 |
| Baby2 | 0.17 | 0.0034 | 0.11 | 0.0004 | -0.50 | 0.0131 | 0.78 | 0.0001 |
| Midd2 | 0.56 | 0.0001 | 0.22 | 0.0007 | -0.94 | 0.0001 | 0.83 | 0.0002 |
| Plastic | 0.56 | 0.0001 | 0.28 | 0.0001 | 0.89 | 0.0045 | 0.44 | 0.0013 |
| Motorcycle | -0.11 | 0.0001 | -0.17 | 0.0001 | 1.06 | 0.0001 | 1.94 | 0.0001 |
| Piano | -0.22 | 0.0001 | 0.28 | 0.0002 | 2.56 | 0.0001 | 2.33 | 0.0001 |
| Playroom | 0.44 | 0.0001 | 0.22 | 0.0006 | 1.89 | 0.0001 | 1.44 | 0.0001 |
| Playtable | 0.22 | 0.0001 | 0.56 | 0.0001 | 1.00 | 0.0052 | 1.44 | 0.0007 |
| **Average** | **0.32** | **0.0003** | **0.24** | **0.0002** | **1.12** | **0.0023** | **1.39** | **0.0005** |

on 2D-JND, which makes them less reliable than the 3D-JND model based on PEs. As a conclusion, our SSJND model performs better for almost the whole dataset except for some rare cases, where it should be noticed that the difference is close to 0.

The results of ANOVA with the null hypothesis of no significant difference of the subject position in terms of subjective scores, give $p - value = 0.28, 0.89, 0.78$, and $0.99$ respectively for the DB-JND, BJND, JJND and DJND models, and indicate that the viewing direction has not significant influence on subjective scores.

## 5. CONCLUSION

In this paper, we propose a saliency-weighted stereoscopic JND (SSJND) model. To this end, we first conduct psychophysical experiments in which we measure the visibility thresholds of the asymmetric noise. The psychophysical data is used to develop a disparity-aware binocular JND (DBJND) model allowing to estimate the JND thresholds for non-occluded pixels. The SSJND profile is build on top of DBJND by including a 2D-JND model for occluded-pixels and accounting for visual saliency. The experimental validation shows that the proposed model outperforms the other 3D-JND models in terms of perceptual quality at the same noise level. A more reliable VS detection approach and an effective VS-map-based weighting function will be investigated in the future to improve the effectiveness of the proposed 3D-JND model.

## 6. ACKNOWLEDGMENT

This research has been funded by both the Research Council of Norway through project no. 221073 "HyPerCept Colour and quality in higher dimensions" and the Region "Nouvelle Aquitaine".

## 7. REFERENCES

[1] E.-L. Tan and W.-S. Gan, "Computational models for just-noticeable differences," in *Perceptual Image Coding with Discrete Cosine Transform*. Springer, Singapore, 2015, pp. 3–19.

[2] D. V. S. X. De Silva, E. Ekmekcioglu, W. A. C. Fernando, and S. T. Worrall, "Display dependent preprocessing of depth maps based on just noticeable depth difference modeling," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 2, pp. 335–351, 2011.

[3] S.-W. Jung, "A modified model of the just noticeable depth difference and its application to depth sensation enhancement," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3892–3903, 2013.

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

[5] X. Li, Y. Wang, D. Zhao, T. Jiang, and N. Zhang, "Joint just noticeable difference model based on depth perception for stereoscopic images," in *2011 IEEE Int. Conf. Vis. Commun. Image Process.*, Tainan, Taiwan, Nov. 2011, pp. 1–4.

[6] F. Lian, S. Liu, X. Fan, D. Zhao, and W. Gao, "A new just-noticeable-distortion model combined with the depth information and its application in multi-view video coding," in *The Era of Interactive Media*. Springer New York, 2013, pp. 229–240.

[7] C. Liu, P. An, Y. Zuo, and Z. Zhang, "Applications of just-noticeable depth difference model in joint multiview video plus depth coding," in *Proc. SPIE*, vol. 9273, 2014, pp. 9273–11.

[8] R. Zhong, R. Hu, Z. Wang, and S. Wang, "3D hybrid just noticeable distortion modeling for depth image-based rendering," *Multimed. Tools Appl.*, vol. 74, no. 23, pp. 457–478, 2015.

[9] F. Qi, D. Zhao, X. Fan, and T. Jiang, "Stereoscopic video quality assessment based on visual attention and just-noticeable difference models," *Signal, Image Video Process.*, vol. 10, no. 4, pp. 737–744, 2016.

[10] F. Xue, C. Jung, and J. Kim, "Disparity-based just-noticeable-difference model for perceptual stereoscopic video coding using depth of focus blur effect," *Displays*, vol. 42, pp. 43–50, 2016.

[11] Y. Yang, C. Lu, J. Li, and H. Yao, "Just noticeable depth difference of human during viewing of dynamic random dot stereograms," in *9th Inter. Symp. on Computational Intelligence and Design*, vol. 1, China, Dec. 2016, pp. 422–424.

[12] C. Li, P. An, L. Shen, K. Li, and J. Ma, "A modified just noticeable depth difference model for 3d displays," in *Int. Forum of Digital TV and Wireless Multimedia Commun.* Springer Singapore, 2017, pp. 63–71.

[13] X. K. Yang, W. S. Ling, Z. K. Lu, E. P. Ong, and S. S. Yao, "Just noticeable distortion model and its applications in video coding," *Signal Process. Image Commun.*, vol. 20, no. 7, pp. 662–680, 2005.

[14] H. G. Kim, S.-i. Lee, and Y. M. Ro, "Experimental investigation of the effect of binocular disparity on the visibility threshold of asymmetric noise in stereoscopic viewing," *Opt. Express*, vol. 24, no. 17, pp. 19 607–19 615, 2016.

[15] C.-H. Chou and Y.-C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Trans. circuits Syst. video Technol.*, vol. 5, no. 6, pp. 467–476, 1995.

[16] K. Gelatt and D. Esson, "Physiology of the eye," *Veterinary Ophthalmology. 4th ed. Iowa: Blackwell Publishing Ltd*, pp. 149–182, 2007.

[17] E. R. Schotter, B. Angele, and K. Rayner, "Parafoveal processing in reading," *Attention, Perception, Psychophys.*, vol. 74, no. 1, pp. 5–35, 2012.

[18] F. Devernay and P. A. Beardsley, "Stereoscopic Cinema." *Image Geom. Process. 3-D Cinematogr.*, vol. 5, pp. 11–51, 2010.

[19] "Subjective methods for the assessment of stereoscopic 3DTV systems," Int. Telecommun. Union, Geneva Switzerland, Tech. Rep. ITU-R BT.2021-1, 2015.

[20] D. H. Brainard and S. Vision, "The psychophysics toolbox," *Spat. Vis.*, vol. 10, pp. 433–436, 1997.

[21] Y. Li, H. Liu, and Z. Chen, "Perceptually-lossless image coding based on foveated-JND and H. 265/HEVC," *J. Vis. Commun. Image Represent.*, vol. 40, pp. 600–610, 2016.

[22] H. Wang, I. Katsavounidis, J. Zhou, J. Park, S. Lei, X. Zhou, M.-O. Pun, X. Jin, R. Wang, X. Wang, and Others, "VideoSet: A large-scale compressed video quality dataset based on JND measurement," *J. Vis. Commun. Image Represent.*, vol. 46, pp. 292–302, 2017.

[23] "Methodology for the subjective assessment of video quality in multimedia applications," Int. Telecommun. Union, Geneva Switzerland, Tech. Rep. ITU-R BT.1708-0, 2007.

[24] P. R. Jones, "A note on detecting statistical outliers in psychophysical data," *bioRxiv*, 2016.

[25] C. M. Jarque and A. K. Bera, "A test for normality of observations and regression residuals," *Int. Stat. Rev.*, pp. 163–172, 1987.

[26] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, 1965.

[27] H. Levene and Others, "Robust tests for equality of variances," *Contrib. to Probab. Stat.*, vol. 1, pp. 278–292, 1960.

[28] P. E. Gill and W. Murray, "Algorithms for the solution of the nonlinear least-squares problem," *SIAM J. Numer. Anal.*, vol. 15, no. 5, pp. 977–992, 1978.

[29] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, no. 1-3, pp. 7–42, 2002.

[30] 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. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1648–1652, 2010.

[31] H. Hadizadeh, "A saliency-modulated just-noticeable-distortion model with non-linear saliency modulation functions," *Pattern Recognit. Lett.*, vol. 84, pp. 49–55, 2016.

[32] H. Hadizadeh, A. Rajati, and I. V. Bajić, "Saliency-guided just noticeable distortion estimation using the normalized laplacian pyramid," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1218–1222, 2017.

[33] Y. Fang, J. Wang, M. Narwaria, P. Le Callet, and W. Lin, "Saliency detection for stereoscopic images," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2625–2636, 2014.

[34] D. Scharstein and R. Szeliski, "Middlebury stereo datasets," *Online at http://vision.middlebury.edu/stereo/data/*, 2002.

[35] 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.

[36] 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.

[37] N. Balakrishnan, V. Voinov, and M. S. Nikulin, *Chi-squared goodness of fit tests with applications*. Academic Press, 2013.