https://doi.org/10.2352/lim.2024.5.1.13

Investigation of the performance of pixel-domain 2D-JND models for 360-degree imaging

Rivo Andriamanalina [1,2]***, Mohamed-Chaker Larabi*** [1]***, Steven Le Moan*** [2]
[1] ***XLIM Lab, University of Poitiers, France;*** [2] ***Colourlab, NTNU, Norway***

Abstract

Spatial just noticeable difference (JND) refers to the smallest amplitude of variation that can be reliably detected by the Human Visual System (HVS). Several studies tried to define models based on thresholds obtained under controlled experiments for conventional 2D or 3D imaging. While the concept of JND is almost mastered for the latter types of content, it is legitimate to question the validity of the results for Extended Reality (XR) where the observation conditions are significantly different. In this paper, we investigate the performance of well-known 2D-JND models on 360-degree images. These models are integrated into basic quality assessment metrics to study their ability to improve the quality prediction process with regards to the human judgment. Here, the metrics serve as tools to assess the effectiveness of the JND models. In addition, to mimic the 360-deg conditions, the equator bias is used to balance the JND thresholds. Globally, the obtained results suggest that 2D-JND models are not well adapted to the extended range conditions and require in-depth improvement or re-definition to be applicable. The slight improvement obtained using the equator bias demonstrates the potential of taking into account XR characteristics and opens the floor for further investigations.

1 - Introduction

In recent years, the rapid advancement of immersive media has marked a promising trajectory expected to expand further. Despite this progress, democratizing Extended Reality (XR) faces challenges. Rendering content with Head-Mounted Displays (HMDs) is computationally intensive. Overcoming these challenges often involves perceptually-tuned data processing. One way to achieve this requires a deep understanding of visual perception mechanisms such as masking effects. Those are often used to build Just Noticeable Difference (JND) models.

A JND model assesses the maximum amplitude of difference that can go unnoticed by a human observer. This enables adaptive coding that assigns varying degrees of importance to different areas of an image. Numerous models have been developed for 2D [1], and also for 3D [2]. They could be split into two categories: pixel-domain and frequency-domain JND. Building these requires well-controlled psycho-physical experiments involving different targeted conditions (FoV: field of view , viewing distance, degree of freedom). While they work well for 2D or 3D, their integration for XR applications such as 360-degree imaging, Virtual Reality, Augmented and Mixed Reality (VR, AR and MR) could not be effective. The reason lies in the important differences in terms of observation conditions (wide FoV, short viewing distance, active selection of viewports). Nevertheless, the principles exploited in these models remain valid but may require some adaptations.

In general, pixel-domain 2D-JND models consider the sensitivity of the Human Visual System (HVS), as well as visual masking (VM) effects. For example, Luminance Adaptation (LA) addresses the visibility of a stimulus based on the ambient luminance [3]. Most JND models also consider Contrast Masking (CM) effect. This refers to the reduction in the visibility of a stimulus due to the presence in its neighborhood, of other stimuli, called masks, with similar characteristics [3] [4] . In addition to the aforementioned effects, the anisotropy of visual acuity across the Field of view can also be exploited for JND modeling[5] [6].

Different 2D-JND models have been proposed previously. For instance, Chou and Li's model [3] laid the groundwork for pixel-domain models by considering the JND as the maximum of two factors: the threshold due to LA and the one due to CM. Subsequently, this model was improved by various authors. Yang et al. [7], introduced a nonlinear additive relationship between LA and CM, employing the Canny detector to distinguish between edge and non-edge regions. This allows dealing with the co-occurrence of both effects, resulting in higher thresholds when both LA and CM are involved. Liu's model [8] improves upon Yang's by distinguishing textures from edges using the L1-norm total variation (TV-L1) method. Thus, CM becomes a combination of texture masking (TM) and edge masking (EM) allowing to cope with TM underestimation in Yang's model. In [9], Wu proposed a model based on the free energy principle, establishing that the brain actively predicts the content of an image and avoids areas that increase uncertainty. This results in higher thresholds where uncertainty is greater. Later on, Chen and Guillemot [10] and Liu et al [11] used foveation effect, such that the farthest pixels from the fixation point have the highest JND thresholds.

We are particularly interested in modeling visual masking (VM) effects for Extended Reality (XR) applications. The various masking effects exploited in 2D-JND models remain relevant in this context. However, the specific parameters and the proposed methods may require further exploration. Therefore, our objective is to determine whether or not we can reliably employ these JND models in immersive imaging and if they accurately predict VM in this specific scenario. Thus, we conduct a benchmark of representative 2D-JND models, focusing on their performance in 360-degree environments by simulating the use of Head-Mounted Displays (HMDs). Furthermore, we propose a method to incorporate the equator bias, a concept unique to XR, into the estimation of JND thresholds to demonstrate further improvement prospects.

The remainder of this paper is organized as follows: In Section 2, we outline our strategy for applying 2D JND models to XR. An objective method for studying their performance in this specific context is described in Section 3. The results are presented and discussed in Section 4.

2 - Using 2D-JND models with 360-degree images

XR aims to immerse users in digitally modified or generated environments, maximizing sensory engagement [12]. Experiencing XR with HMDs involves an expansive FoV and provides a sense presence. Thus, studying XR requires considering the specific behavior of the HVS while experiencing visual immersion.

For instance, it has been observed in prior studies that human subjects have a strong tendency to focus on elements near the equator. This phenomenon is referred to as the equator bias [13]. This is leveraged in two ways in our study: first, by extracting more viewports along the equator ; and also, by modulating the JND thresholds based on the latitude of the pixels.

2.1 - Viewport extraction

Given that observers do not directly perceive the ERPs (equirectangular projection) and due to significant geometric distortions, the JND should not be calculated on ERPs directly [14]. Instead, we extract multiple viewports to mimic the behavior of a human observer during visual exploration with HMDs. As shown in Fig 1, 8 viewports located along the equator are extracted from the sphere, with camera rotations of -135°, -90°, -45°, 0°, 45°, 90°, 135°, and 180°. This is motivated by the fact that an observer pays closer attention to elements at eye level and the transition from one viewport to another is conditioned by head movements. Only one viewport per polar region is extracted since these areas are often unexplored by viewers.

Since we are interested in what is actually perceived by the observer, we take the characteristics of the HMDs in consideration (field of view and spatial resolution). As the used datasets are from studies utilizing the HTC Vive, with a vertical resolution of 1200 pixels, we extract viewports of 1200x1200 pixels matching a 90° vertical FoV.

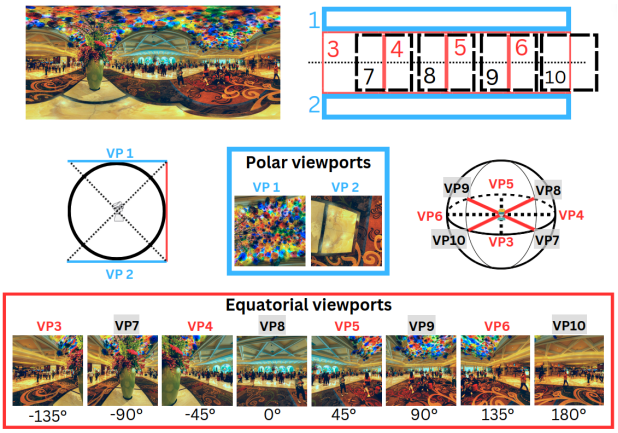


Figure 1: Viewport extraction method based on equator bias

2.2 - Threshold elevation due to equator bias

Given the tendency of observers to focus on regions near the equator, further areas are perceived with lower visual acuity due to foveation effect. Thus, one could argue that JND thresholds may vary based on latitude. To model this, we first need to model the saliency distribution. Next, we deduce a critical foveation latitude range, i.e., the range of latitudes where the majority of observers focus. Finally, a foveation model centered on the boundaries of this range provides the loss of sensitivity as a function of latitude.

In order to get an expression of the saliency distribution, we use saliency maps from Salient360![15] and calculate an average map as shown in Fig 2. It is easy to notice that the average saliency is concentrated around the equator and could be approximated by a Gaussian function as follows:

$$S(\Delta l) = \frac{N}{\sigma * \sqrt{2\Pi}} exp(-\frac{\Delta l^2}{2 * \sigma^2}) + \beta \quad (1)$$

where N is a normalization factor, σ is the standard deviation and β is an offset. In our case, N, σ and β are respectively 29, 14 and 0.17. Δl is the latitude of a pixel P(x,y) in degrees.

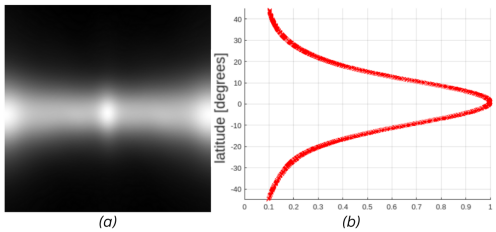


Figure 2: Saliency in function of the Latitude
(a) Average saliency map, (b) average saliency distribution

Since all points near the equator can be foveated, there is a risk of threshold overestimation at those locations if we use the saliency distribution as elevation factor. Therefore, we define a latitude range within which we do not modify the JND thresholds. To achieve this, we consider that the boundary regions of foveation are located at $\Delta l = \{+\sigma, -\sigma\}$ as shown in Fig3. Thus, thresholds calculated within this range are not affected. Then, a foveation model F centered at $\Delta l = \{+\sigma, -\sigma\}$ provides the JND threshold elevation factor. The general model of the equator bias is expressed as follows:

$$\alpha(\Delta l) = min(F(\Delta l) \circledast \delta(\Delta l - \sigma), F(\Delta l) \circledast \delta(\Delta l + \sigma)) \quad (2)$$

where $\alpha(\Delta l)$ is the elevation factor for the JND threshold at a latitude Δl. $F(\Delta l)$ is a foveation model in which the eccentricity is calculated solely in the vertical direction, and δ represents a dirac. The foveation model used here is derived from that of Chen and Guillemot. The latter allows for the modulation of JND thresholds based on eccentricity and local background luminance L.

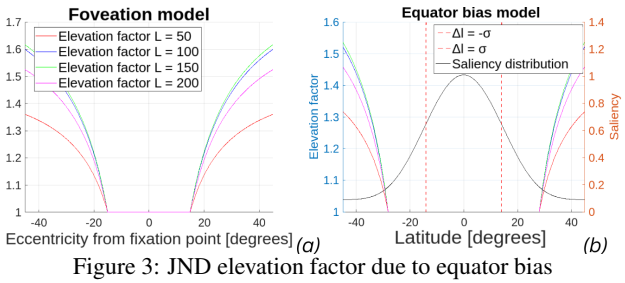


Figure 3: JND elevation factor due to equator bias

(a) represents the foveation model $F(\Delta l)$, (b) represents the resulting elevation factor to be applied to the JND map.

This equator bias function is to be applied on the different 2D-JND maps. Only Chen's model [10] and Liu's VR-JND [11] need to be treated differently. These require predicting fixation points in the images in order to account for foveation effect. For this purpose, a saliency prediction method such as Itti and Koch's [16] is often employed. The performance of these models is therefore directly dependent on the accuracy of the saliency prediction. Thus, to avoid applying the foveation effect twice and to eliminate the bias associated with the saliency prediction, foveation is replaced with the previously described equator bias function. Chen's model is therefore equivalent to the equator-biased version of Chou's. Liu's model for VR is unique because it was established differently from the others by taking depth into account. This modality is out of the scope of this study, so we assigned a uniform depth to all areas of the images. It also has its own foveation computation method which we have used in the equator bias for this model.

3 - Performance evaluation method

We define the performance of a JND model as its ability to mimic the flaws of the HVS. However, to our best knowledge, there is currently no standardized objective method available for

Table 1: Effects of Equator bias

Equator bias	Chou/Chen PLCC JND_{SSIM}	Chou/Chen JND energy [dB]	Yang PLCC JND_{SSIM}	Yang JND energy [dB]	Wu PLCC JND_{SSIM}	Wu JND energy [dB]	Liu PLCC JND_{SSIM}	Liu JND energy [dB]	Liu VR - 2D PLCC JND_{SSIM}	Liu VR - 2D JND energy [dB]
Without	0,679	17,56	0,677	17,41	0,657	18,82	0,688	19,83	-	-
With	**0,687**	**18,12**	**0,682**	**17,9**	**0,679**	**20,38**	**0,694**	**21,36**	0,681	20

Table 2: PLCC Correlation of the JND_{IQA} with the MOS and Energy of Just Noticeable Difference: The left part displays the PLCC correlation between the JND-weighted IQA metrics and the Mean Opinion Score (MOS). The right part showcases the energy of JND calculated using Eq 7 for each JND model. In the last row, the correlation scores were computed across the entire dataset, whereas the energy of just noticeable difference represents the average of all individual values.

P	IQA		Yang		Wu		Liu		Chen		LiuVR		Energy of JND profile				
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	Yang	Wu	Liu	Chen	Liu VR
1	0.905	0.920	0.902	0.918	0.907	0.917	**0.924**	**0.934**	0.903	0.920	0.908	0.923	65.2	115.4	**145.5**	41.7	65.2
2	0.745	0.740	0.733	0.736	0.742	0.740	**0.746**	**0.746**	0.738	0.739	0.736	0.738	59.3	104.5	**130.3**	67.1	96.2
3	0.665	0.698	0.667	0.703	0.634	0.684	**0.719**	**0.740**	0.675	0.708	0.668	0.703	64.0	113.3	**142.6**	50.7	71.3
4	0.949	0.937	0.948	0.937	**0.960**	0.928	0.955	**0.945**	0.948	0.938	0.951	0.939	62.5	110.1	**138.4**	53.6	84.1
5	0.705	0.705	0.686	0.696	0.698	0.696	**0.748**	**0.736**	0.699	0.704	0.692	0.698	64.7	114.4	**144.0**	51.3	88.8
6	0.830	0.850	0.825	0.847	0.835	0.844	**0.849**	**0.868**	0.828	0.848	0.829	0.849	63.9	113.2	**142.3**	48.8	72.2
7	0.709	0.725	0.724	0.736	0.650	0.709	**0.761**	**0.763**	0.723	0.736	0.715	0.734	65.7	116.0	**146.2**	35.0	57.2
8	0.961	0.896	0.936	0.887	**0.966**	0.894	0.954	**0.900**	0.942	0.891	0.942	0.890	57.4	101.0	126.0	116.5	**172.3**
9	0.901	0.935	0.907	0.937	0.898	0.944	**0.914**	**0.945**	0.906	0.938	0.908	0.939	62.9	111.1	**139.6**	48.0	71.4
10	0.833	0.803	0.823	0.794	0.828	0.795	**0.844**	**0.813**	0.826	0.797	0.826	0.796	62.5	110.8	**139.3**	61.6	94.7
11	0.906	0.849	0.870	0.834	**0.908**	0.845	0.905	**0.854**	0.880	0.839	0.877	0.837	56.0	99.3	123.5	106.9	**140.2**
12	0.982	0.910	0.970	0.906	**0.981**	0.898	0.978	**0.914**	0.973	0.908	0.973	0.907	58.1	102.1	127.9	107.7	**166.3**
13	0.906	0.933	0.917	0.933	0.887	0.927	**0.927**	**0.942**	0.915	0.934	0.917	0.937	59.3	103.7	130.0	79.3	**149.3**
14	0.790	0.810	0.778	0.809	0.759	0.801	**0.811**	**0.839**	0.781	0.812	0.780	0.814	61.8	108.6	**136.0**	70.1	122.1
15	0.862	0.849	0.856	0.845	0.870	0.845	**0.876**	**0.864**	0.858	0.846	0.860	0.847	63.7	112.9	**142.0**	45.4	65.7
16	0.805	0.796	0.815	0.804	0.795	0.793	**0.847**	**0.830**	0.819	0.809	0.814	0.805	60.1	109.2	**137.2**	54.1	83.6
ALL	0.697	0.718	0.682	0.707	0.679	0.706	**0.694**	**0.74**	0.687	0.713	0.681	0.71	61.7	109.1	**136.9**	64.9	100.0

evaluating the performance of JND models. To quantify the efficiency of the models and conduct a comparative analysis, we follow the workflow outlined in this section.

We adopt a similar method to the one developed for 3D-JND models comparison in [17]. In this approach, well-known quality metrics are modified to create JND-weighted versions in which the weight of a pixel depends on its JND threshold. Therefore, the performance of the JND model is linked to the performance of the weighted metrics. Given that we only consider luminance in this study, we use SSIM (Structural Similarity Index), as described in [18], and PSNR (Peak Signal-to-Noise Ratio). Embedding JND within these two metrics is straightforward. We thus define two JND_{IQA}, namely JND_{SSIM} and JND_{PSNR}, formulated as follows:

$$QualMap(x,y) = \frac{SSIM(I(x,y),I'(x,y)).min(JND)}{JND(x,y)} \tag{3}$$

$$JND_{SSIM}(I,I') = \frac{\sum_{x=1}^{H}\sum_{y=1}^{W} QualMap(x,y)}{\sum_{x=1}^{H}\sum_{y=1}^{W} \frac{min(JND)}{JND(x,y)}} \tag{4}$$

where ***I*** and ***I'*** are respectively the pristine image and the impaired image. ***QualMap*** represents the quality map of **I'** weighted by the JND thresholds. $\boldsymbol{JND_{SSIM}(I,I')}$ is the final predicted score of the impaired image ***I'***.

Similarly, we modify the calculation of Mean Squared Error (MSE) to give more importance to pixels with lower JND thresholds. The resulting Weighted MSE (***WMSE***) is then used in the calculation of $\boldsymbol{JND_{PSNR}}$ as follows:

$$WMSE = \frac{\sum_{x=1}^{H}\sum_{y=1}^{W} \frac{min(JND)}{JND(x,y)}.(I(x,y)-I'(x,y))^2}{\sum_{x=1}^{H}\sum_{y=1}^{W} \frac{min(JND)}{JND(x,y)}} \tag{5}$$

$$JND_{PSNR}(I,I') = 10 * log_{10}(\frac{255^2}{WMSE}) \tag{6}$$

A score is predicted for each viewport of the degraded images using Eq 4 and 6. Subsequently, a global score for the entire image is computed by averaging the scores of the viewports. Finally, the correlation between the predicted scores and the Mean Opinion Score (MOS) is examined. The higher the correlation, the closer the predicted quality is to the human judgement, and thus, the more efficient the JND model is. However, correlation alone is not sufficient to compare the models. It is possible for two models to produce very close correlation scores. In such cases, to distinguish between them, one can compare the amount of luminance variation they tolerate. This can be approximated with the mean energy of the JND (θ) map such as:

$$\theta = \frac{1}{H.W}\sum_{x=1}^{H}\sum_{y=1}^{W} JND(x,y)^2 \tag{7}$$

We believe that performance should be a combination of the correlation scores calculated with Eq 4 and 6 and the amount of imperceptible difference. We define a relative performance index *PI* as follows:

$$PI(JND_k) = corr(MOS, JND_{IQA}).\frac{\theta_k}{max(\theta)} \tag{8}$$

where JND_k is the k-th model considered in this study and θ_k its mean energy. Thus, when the correlation score and the mean energy of JND profile are simultaneously high, the performance index is high. Conversely, if one or both are low, the performance is decreased. This index is used only for inter-model comparison because it evaluates only the performance of each model relative to the others.

4 - Experimental results

For the benchmark, we used the OIQA dataset proposed in [19] . This dataset consists of a set of 16 pristine images impaired with JPEG and JPEG2k compression, white Gaussian noise, and Gaussian blur. Each type of degradation was applied

with five intensity levels. Since most IQA metrics perform poorly for the images with WGN in the dataset [19], we chose not to include these images to reduce potential biases due to the metric. Therefore, we rely on a set of 15 degraded images per pristine, resulting in a total set of 240 images.

First, we successively calculated the JND maps for all viewports. Then, the correlation scores of the JND_{IQA} with the MOS over the whole dataset, and the energy of tolerable difference were computed. We compared the performance of JND models with and without elevation by the equator bias function and the results are presented in Table 1. We observe that the equator bias function allows for a slight improvement in terms of correlation with the human judgement. We also tolerate more luminance variation, with an average gain of 1.035 dB. The simultaneous improvement of both factors shows that the equator bias can be effectively utilized in modeling visual masking for the XR context. This function is exploited throughout the remainder of this study.

The second part of our study focused on the overall behavior of JND models with 360-degree images. To investigate this, the correlation of the weighted metrics with the MOS was calculated, on each pristine image and its degraded versions, and also on the entire dataset. Even if our work was not on improving the IQA metrics, a performance enhancement of the SSIM and PSNR was expected if the JND models were adapted to XR. This has been achieved by Fezza et al. with SSIM in 3D imaging [18]. The results of our experiment are presented in Table 2. We observe that the overall PLCC and SROC values are generally low for all JND models on the entire dataset even if we have some very good scores for individual images. However, we are more interested in the differences provoked by the JND models rather than the correlation values. We see in the last row of the table that, across the entire dataset, the correlation values obtained with the baseline metrics are slightly higher than those embedding the JND, except for Liu's model with PSNR. At first glance, the 2D-JND models do not improve the performance of the metrics but we do notice some increased values for certain models on each pristine image individually. Thus, we attempted to verify if these differences brought by the JND could be significant. Statistical tests between the scores from each JND-IQA and the baseline metrics were conducted. As given in Table 3 , the results of the T-Test indicates that the changes are not statistically significant. The fact that the 2D-JND models do not have significantly positive impact on the quality metrics suggests that they may not be well-suited for our experimental context. It is thought to be related to a misestimation of thresholds in different areas of the image. Some imperceptible differences in 2D conditions become noticeable in immersive conditions and vice-versa. This prevents the 2D-JND models from improving the quality metrics in this context.

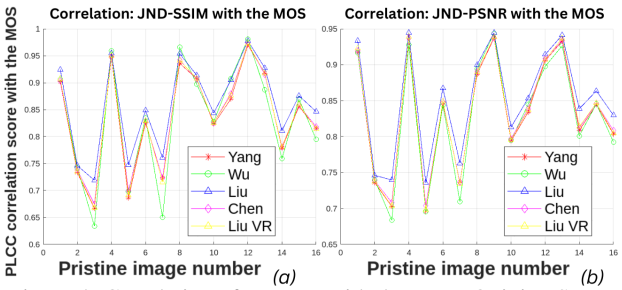
Table 3: P-values of the T-Test between JND_{IQA} and baseline IQA

p-value	Yang	Wu	Liu	Chou / Chen	LiuVR
PSNR	0.43	0.409	0.278	0.47	0.458
SSIM	0.499	0.473	0.2	0.455	0.469

The third observation from this study concerns inter-model comparison. Despite using different JND_{IQA} metrics, we observe in Table 2 that they demonstrate roughly equivalent correlations with the MOS on the whole dataset. This suggests that the JND models are similarly correlated with human judgment. Nonetheless, as shown in Fig 4 (a) and (b), Liu's JND_{IQA} seems to correlate slightly better with the MOS than the other JND-weighted metrics. However, as correlation alone does not provide sufficient information to rank the JND models, we complemented the comparison with the amount of difference tolerated by each of them as reported in Table 2. In order to establish a ranking of 2D-JND models in the XR context, we calculated the performance index of each using Eq 8 . The findings are presented in Table 4 and reveal that when considering both the correlation of JND_{IQA} with human judgement and the ability of the JND models to tolerate luminance variation, Liu's outperforms the others. As the JND models studied here rely on the same masking effects, we can deduce that it is the use of the TV-L1 method for CM that has made the difference. This demonstrates the importance of differentiating texture from edges, thus indicating an increased sensitivity of the HVS to these characteristics in our context.

Table 4: Performance index for ranking

	Liu	Wu	Liu VR - 2D	Chou / Chen	Yang
Perf Idx	0.694	0.541	0.497	0.326	0.310



(a) (b)

Figure 4: Correlation of JND_{IQA} with the Mean Opinion Score

Illustration of the correlation between JND_{IQA} metrics and the MOS. (a) JND_{SSIM} (b) JND_{PSNR}

However, these results should be nuanced as several visual modalities were left aside in this study. Notably, we did not take into account the sense of depth provided by visual immersion which may explain why Liu's model for VR did not stand out from the others. Finally, we realize that the effect of JND models on quality metrics is subtle. This limits our ability to deduce an exact performance from them. We could have raised the JNDs to a certain power in Eq 4 and 6, but that would have overestimated the performance gaps between the different models and biased our study. We acknowledge that this benchmarking method is debatable, particularly the notion of assessing JND models through quality metrics. However, it allows us to obtain results that we can interpret quantitatively.

Conclusion

The performance analysis of various 2D-JND models within an XR framework poses a complex challenge since there is no standardized objective method to do so. In this study, we conducted a benchmark of 2D JND models specifically on 360-degree images by integrating them into quality metrics and studying their correlation with human judgment. However, the aim was not to propose new IQA metrics. The findings of this study suggest that original 2D-JND models exhibit limited suitability for immersive environments as none of them can be reliably used for quality assessment purpose. Although the models are roughly equivalent in this regard, the final inter-model comparison showed that Liu's 2D-JND model performs relatively better than the others. This underscores the significance of distinguishing between texture and edge within our specific context. We also found that accounting for equator bias led to slight improvements in the models' performance, thereby showcasing the potential for improvement tailored to the specificities of XR. This approach will undergo subjective cross-validation in future work. A strategy to integrate more sophisticated phenomena in XR will also be explored in our forthcoming research.

References

[1] Zhenzhong Chen and Hongyi Liu. JND modeling: Approaches and applications. In *19th International Conference on Digital Signal Processing*, pages 827–830, Hong Kong, August 2014. IEEE.

[2] Yu Fan, Mohamed-Chaker Larabi, Faouzi Alaya Cheikh, and Christine Fernandez-Maloigne. A Survey of Stereoscopic 3D Just Noticeable Difference Models. *IEEE Access*, 7:8621–8645, 2019.

[3] Chun-Hsien Chou and Yun-Chin Li. A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile. *IEEE Transactions on circuits and systems for video technology*, 5(6):467–476, 1995.

[4] Xiaokang Yang, Weisi Lin, Zhongkhang Lu, EePing Ong, and Susu Yao. Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile. *IEEE Transactions on Circuits and Systems for Video Technology*, 15(6):742–752, June 2005.

[5] Peiyao Guo, Qiu Shen, Mingkai Huang, Rongbing Zhou, Xun Cao, and Zhan Ma. Modeling peripheral vision impact on perceptual quality of immersive images. In *IEEE Visual Communications and Image Processing (VCIP)*, pages 1–4, St. Petersburg, FL, December 2017.

[6] Zhou Wang, Ligang Lu, and Alan C Bovik. Foveation scalable video coding with automatic fixation selection. *IEEE Transactions on Image Processing*, 12(2):243–254, 2003.

[7] 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 Processing: Image Communication*, 20(7):662–680, August 2005.

[8] Anmin Liu, Weisi Lin, Manoranjan Paul, Chenwei Deng, and Fan Zhang. Just Noticeable Difference for Images With Decomposition Model for Separating Edge and Textured Regions. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(11):1648–1652, November 2010.

[9] Jinjian Wu, Weisi Lin, Guangming Shi, and Anmin Liu. Perceptual Quality Metric With Internal Generative Mechanism. *IEEE Transactions on Image Processing*, 22(1):43–54, January 2013.

[10] Zhenzhong Chen and Christine Guillemot. Perceptually-Friendly H.264/AVC Video Coding Based on Foveated Just-Noticeable-Distortion Model. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(6):806–819, June 2010.

[11] Di Liu, Yingbin Wang, and Zhenzhong Chen. Joint foveation-depth just-noticeable-difference model for virtual reality environment. *Journal of Visual Communication and Image Representation*, 56:73–82, October 2018.

[12] Andrew Perkis and Christian Timmerer. White Paper on Definitions of Immersive Media Experience (IMEx).

[13] Yucheng Zhu, Guangtao Zhai, and Xiongkuo Min. The prediction of head and eye movement for 360 degree images. *Signal Processing: Image Communication*, 69:15–25, 2018.

[14] Sami Jaballah, Amegh Bhavsar, and Mohamed-Chaker Larabi. A comprehensive framework for 2D-JND extension to 360-deg images. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2752–2756, 2020.

[15] Jesús Gutiérrez, Erwan J David, Antoine Coutrot, Matthieu Perreira Da Silva, and Patrick Le Callet. Introducing un salient360! benchmark: A platform for evaluating visual attention models for 360 contents. In *IEEE Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–3, 2018.

[16] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, November 1998.

[17] Yu Fan, Mohamed-Chaker Larabi, Faouzi Alaya Cheikh, and Christine Fernandez-Maloigne. On the performance of 3d just noticeable difference models. In *IEEE International Conference on Image Processing (ICIP)*, pages 1017–1021, 2016.

[18] Sid Ahmed Fezza, Mohamed-Chaker Larabi, and Kamel Mohamed Faraoun. Stereoscopic image quality metric based on local entropy and binocular just noticeable difference. In *IEEE International Conference on Image Processing (ICIP)*, pages 2002–2006, 2014.

[19] Huiyu Duan, Guangtao Zhai, Xiongkuo Min, Yucheng Zhu, Yi Fang, and Xiaokang Yang. Perceptual quality assessment of omnidirectional images. In *IEEE international symposium on circuits and systems (ISCAS)*, pages 1–5, 2018.

Author Biography

Rivo Andriamanalina *received his BSc in Science and Technology, specializing in audiovisual techniques and digital media (2021), as well as his MSc in Media Technology Engineering (2023) from the National Institute of Applied Science, Hauts-de-France, University of Valenciennes. He is currently pursuing his Ph.D. in Signal and Image Processing with the University of Poitiers and the Norwegian University of Science and Technology. His research focuses on visual masking effects in the context of extended reality.*

Mohamed-Chaker Larabi *(Senior Member, IEEE) received the Ph.D. degree from Université de Poitiers, in 2002. He is currently full Professor with Université de Poitiers. He is also the Deputy Scientific Director of the GdR-ISIS (French Research Group on Signal and Image Processing). He has participated in several national and international projects. He supervised more than 20 Ph.D. students and published more than 200 papers. His research interests include quality of experience and bio-inspired processing/coding/optimization of images and videos, such as 2-D, 3-D, HDR, and 360/VR/AR/MR. He has been elected to serve as a member of the IEEE SPS IVMSP, MMSP, and EURASIP TAC-VIP Technical Committees. He is also a member of the CIE, IS&T, and the MPEG and JPEG Committees.*

Steven Le Moan *(Member, IEEE) received the Ph.D. degree in image processing from the University of Burgundy, France, in 2012. He then worked as a Postdoctoral Researcher at the Technical University of Darmstadt, Germany, and the Gjøvik University College, Norway. Until 2021, he was a Senior Lecturer in electronics, information and communication systems with Massey University, New Zealand. He is currently an Associate Professor of colour imaging at the Norwegian University of Science and Technology, Gjøvik, Norway. His research interests include multi/hyperspectral image analysis, color science, and visual perception.*