Saliency-Based Sharpness Mismatch Detection For Stereoscopic Omnidirectional Images

Simone Croci
Trinity College Dublin
crocis@scss.tcd.ie

Sebastian Knorr
Trinity College Dublin
knorrs@scss.tcd.ie

Aljosa Smolic
Trinity College Dublin
smolica@scss.tcd.ie

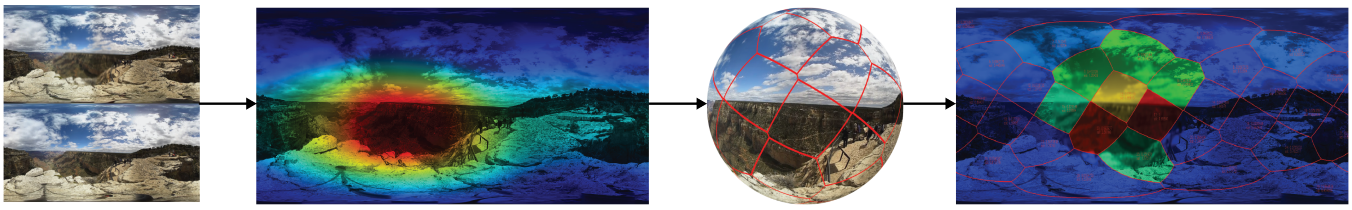


Figure 1: Saliency-based sharpness mismatch detection using Voronoi patches.

ABSTRACT

In this paper, we present a novel sharpness mismatch detection (SMD) approach for stereoscopic omnidirectional images (ODI) for quality control within the post-production workflow, which is the main contribution. In particular, we applied a state of the art SMD approach, which was originally developed for traditional HD images, and extended it to stereoscopic ODIs. A new efficient method for patch extraction from ODIs was developed based on the spherical Voronoi diagram of evenly distributed points on the sphere. The subdivision of the ODI into patches allows an accurate detection and localization of regions with sharpness mismatch. A second contribution of the paper is the integration of saliency into our SMD approach. In this context, we introduce a novel method for the estimation of saliency maps from viewport data of head-mounted displays (HMD). Finally, we demonstrate the performance of our SMD approach with data collected from a subjective test with 17 participants.

CCS CONCEPTS

• **Computing methodologies** → **Image processing**;

KEYWORDS

360 video, omnidirectional images, 3D quality assessment, sharpness mismatch detection, saliency maps, virtual reality

ACM Reference Format:
Simone Croci, Sebastian Knorr, and Aljosa Smolic. 2017. Saliency-Based Sharpness Mismatch Detection For Stereoscopic Omnidirectional Images. In *CVMP 2017: 14th European Conference on Visual Media Production (CVMP 2017), December 11–13, 2017, London, United Kingdom.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3150165.3150168

CVMP 2017, December 11–13, 2017, London, United Kingdom

ACM ISBN 978-1-4503-5329-8/17/12…$15.00
https://doi.org/10.1145/3150165.3150168

1 INTRODUCTION

Virtual Reality (VR) is the next step in consumer electronics providing a more immersive experience than ever before. VR content development has focused on gaming, where the use of headsets to create more immersive experiences is a natural extension of the medium. For VR to evolve though it has to look beyond gaming to attract a wider audience, and Filmmakers are grabbing this opportunity to tell their stories in new and imaginative ways and to draw audiences into their worlds. The challenge outside of gaming is to provide a steady stream of quality content to attract audiences. Shooting a live action immersive 360-degree experience, i.e. omnidirectional images or 360-videos, is not only a technological challenge, but requires the whole production chain to be revolutionized from set construction to acting, from directing to post production.

Capturing 360-videos is not an easy task as there are many technical limitations which need to be overcome, especially for capturing and post-processing in stereoscopic 3D (S3D). In general, such limitations result in artifacts which cause visual discomfort when watching the content with a head-mounted-display. The artifacts or issues can be divided into three categories: binocular rivalry issues (e.g. sharpness and color mismatch, geometrical misalignment, etc.), conflicts of depth cues (e.g. stereopsis vs. perspective, vergence vs. accommodation, etc.) and artifacts which occur in both monocular and stereoscopic 360-degree content production (e.g. stitching and blending artifacts) [17].

With respect to binocular rivalry issues, this paper introduces a novel system and workflow for sharpness mismatch detection in ODIs for quality control within the post-production workflow, i.e. to give automatic feedback to artists and reduce time and efforts in post, which is the main motivation of this paper. In particular, we applied the state of the art SMD approach by Liu et al. [21], which was originally developed for traditional HD images, and extended it

to stereoscopic ODIs, which is the first technical contribution of this paper. The method in [21] estimates a sharpness mismatch score by measuring width deviations of edge pairs in stereoscopic views located at different depth planes. One important component of the extension is a new method for the extraction of patches from the ODIs based on the spherical Voronoi diagram [1]. In our approach, Liu et al.'s method is applied to each patch independently, and then the patches with sharpness mismatch including the SM scores are visualized using a color coded representation as illustrated in Figure 1. Besides the extension to ODI, we also integrated saliency in the quality control workflow in order to weight the appearance of sharpness mismatch dependent on the visual attention of end-users who only see a small portion of the entire ODI using their HMDs, i.e. the viewport.

We first explain how saliency can be used in order to weight regions according to it. Then, we introduce a new method for saliency estimation from viewport data of HMDs, which is the second technical contribution of this paper. A subjective test with 17 subjects was organized in order to collect viewport data while the test subjects were looking at stereoscopic ODIs. With this data the saliency maps of the ODIs were computed. We then analyzed these images with the extension of Liu et al.'s method in order to demonstrate the performance of our proposed approach.

The paper is structured as follows. In Section 2, related work in quality assessment, artifact detection, and saliency estimation is reviewed. Then, in Section 3, we describe the state of the art approach for SMD presented in [21] which we extended to ODIs by applying the spherical Voronoi diagram and by integrating saliency for weighting the importance of the artifacts. In Section 4, we describe the subjective test which we carried out to estimate the saliency maps of ODIs. Results for demonstrating the performance of our proposed method are presented in Section 5. Finally, in Section 6, the paper concludes with a discussion and future work.

2 RELATED WORK

2.1 Quality assessment and artifact detection

Over the last years, binocular rivalry issues and conflicts of depth cues have been investigated in detail for standard S3D content e.g. for cinema screens and 3D-TV [18, 20, 27]. Many publications focused on the assessment of 3D quality in terms of subjective and objective quality metrics. In [16], the authors investigated how viewer annoyance depends on various technical parameters such as vertical disparity, rotation and field-of-view mismatches as well as color and luminance mismatches between the views. In [9], a stereo camera distortion detecting method based on statistical models was presented in order to detect vertical misalignment, camera rotation, unsynchronized zooming, and color misalignment in native S3D content. The authors of [11] introduced a full-reference quality assessment metric for stereoscopic images based on the perceptual binocular characteristics. The proposed metric handles asymmetric distortions of stereoscopic images by incorporating human visual system characteristics. Finally, in [2], a full-reference metric was presented by evaluation of a large variety of measures by taking 2D picture quality, binocular rivalry and depth map degradation into account. The authors maximized the correlation with the mean opinion score (MOS) by using linear regression.

In this paper, however, the focus lies on artifact detection in order to support the artist by giving direct quality feedback during post-production, in particular for sharpness mismatch. Thus, full-reference objective quality metrics can not be applied in this context. In [24], the authors explored the relationship between the perceptual quality of stereoscopic images and visual information, and introduced a model for binocular quality perception. Based on this model, a no-reference quality metric for stereoscopic images was proposed. The proposed metric is a top-down method modeling the binocular quality perception of the human visual system (HVS) in the context of blurriness and blockiness.

A large variety of artifact detection methods, including sharpness mismatch, was introduced in [28] and [4]. For SM the two papers proposed approaches that first apply dense disparity estimation and then analyze high-frequency differences between both views [28] or analyze differences of edges using a gradient-based approach [4]. For measuring in-picture sharpness, different 2D metrics have been developed. In [10], a new perceptual no-reference image sharpness metric based on the notion of just noticeable blur (JNB) was introduced. The proposed metric is able to predict the relative amount of blurriness in images with different content. Furthermore, the authors showed that the HVS masks blurriness around an edge up to a certain threshold. An ideal metric is the cumulative probability of blur detection (CPBD) metric [23], as it outperforms most other no-reference sharpness metrics on Gaussian blur. It was developed based on human blur perception at different contrasts. However, none of the related work focused on 3D artifact detection in ODIs. To our knowledge only the work in [17] focused on 3D quality assessment methods that deal with ODIs. The authors analyzed vertical misalignment and global color mismatch in ODIs using the equirectangular presentation, but no sharpness mismatch was analyzed.

2.2 Saliency

Saliency can be predicted using models of visual attention. In the last 20 years many saliency models have been proposed for traditional 2D images [13]. The focus of these models is on the different features that motivate visual attention towards a particular target location. The visual features can be classified as low level features (color, intensity, orientation) [14, 15] and high level features (face [26] and object [7] detection, image-center prior [15]). The extracted features can be linearly combined [14] or integrated using learned weights [26, 29]. More recently, models using deep neural networks, trained for object recognition, have shown impressive performance in predicting visual attention [12, 19]. However, these models may not suit the specificities of 360-video.

In [5], a saliency model for panoramic images (cylindrical ODIs) was proposed, where conventional saliency models are used on the planar representation of the ODIs. In [3], a spherical saliency model was presented, where different visual features were computed on the sphere. However, these works are based on a model that only considers low level features. Most recently, the authors of [8] estimated saliency maps for ODIs viewed with HMDs, when the use of an eye tracker device is not possible. They collected viewport data of 32 participants for 21 ODIs and proposed a method to transform the gathered data into saliency maps.

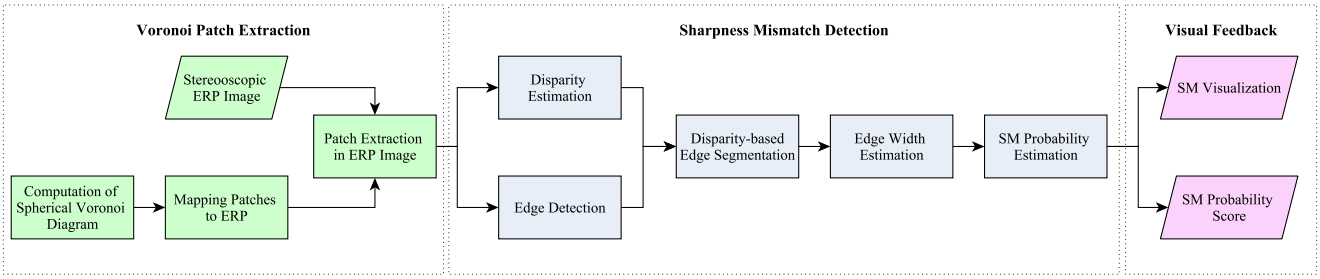
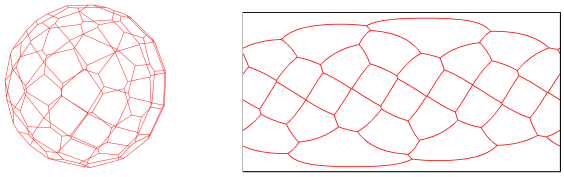


Figure 2: Overview of the sharpness mismatch approach



(a) Spherical Voronoi diagram. **(b) Spherical Voronoi diagram mapped into the ERP format.**

Figure 3: Patch extraction.

3 PROPOSED METHOD FOR SMD

To detect sharpness mismatch in an ODI, we first divide its spherical representation into patches, map the spherical patches into the equirectangular projection (ERP), extract these patches from the stereoscopic ERP image, process each patch independently using the approach proposed by Liu et al. [21], and finally highlight the patches with SM using a color coded representation as illustrated in Figure 1. Figure 2 gives a schematic overview of the proposed SMD approach. The three blocks: Voronoi Patch Extraction, Sharpness Mismatch Detection and Visual Feedback are detailed in the following subsections.

3.1 Voronoi patch extraction from ODIs

To extract approximately equally sized patches from the ODI, first evenly distributed points are computed on the sphere, and then the spherical Voronoi diagram [1] is computed from them. Each cell of the computed Voronoi diagram corresponds to a patch. Figure 3 shows the spherical Voronoi diagram computed from 30 evenly distributed points on the sphere, and its projection into the equirectangular format.

For the computation of n evenly distributed points $\mathbf{P}_i = (X_i, Y_i, Z_i)$ with $i = 0 \ldots n-1$ on the sphere, we implemented a method based on the following equations:

$$\theta_i = i\pi \cdot \left(3 - \sqrt{5}\right), \tag{1}$$

$$Z_i = \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2i}{n-1}\right), \tag{2}$$

$$d_i = \sqrt{1 - Z_i^2}, \tag{3}$$

$$X_i = d_i \cdot cos(\theta_i) \quad and \tag{4}$$

$$Y_i = d_i \cdot sin(\theta_i), \tag{5}$$

where θ_i is the azimuthal angle and d_i is the distance of the point from the z-axis.

The spherical Voronoi diagram is computed based on the evenly distributed points $\mathbf{P}_i$, and it basically consists of partitioning the surface of the sphere into cells for each point $\mathbf{P}_i$. Each Voronoi cell VC_i defines the region on the surface of the sphere Ω_S containing all points which are closer to the corresponding point $\mathbf{P}_i$ than to any of the other evenly distributed points $\mathbf{P}_j$:

$$VC_i = \{\mathbf{P} \in \Omega_S \mid d_S(\mathbf{P}, \mathbf{P}_i) \leq d_S(\mathbf{P}, \mathbf{P}_j) \; \forall j \neq i\}, \tag{6}$$

where $d_S(\mathbf{P}, \mathbf{P}_i)$ is the spherical distance between the current point $\mathbf{P}$ and the point $\mathbf{P}_i$, i.e., the length of the shortest path on the surface of the sphere connecting these two points.

For each Voronoi cell the centroid that defines the orientation of the patch's image plane is computed, and then the patch is mapped onto the surface of the ERP image. The resolution of each patch is defined by the pixels per visual angle, a parameter that is kept constant for each patch.

During the mapping of the spherical patch of the ODI into the ERP format, the pixel colors are obtained by sampling the ODI in ERP format using bilinear interpolation.

The number of patches and thus the size of each patch has an impact on the localization of SM. The larger the patch size is, the more difficult it is to detect and localize SM if it only appears in small areas of the ODI. We empirically found out that 30 patches is a good number for the localization of SM.

3.2 Sharpness Mismatch Detection

Each of the patches computed in the previous section is now analyzed individually using the SMD approach introduced by Liu et al. [21], which was originally developed for traditional HD images.

The approach first estimates the horizontal disparity maps of a stereoscopic image pair from the left view to the right view, and vice-versa. Since the disparity estimation can be noisy and inaccurate, only disparity values are considered which are consistent in both disparity maps. If DM_{L2R} and DM_{R2L} are the two disparity maps, then the disparity at pixel (x, y) in DM_{L2R} is valid if

$$|DM_{L2R}(x, y) + DM_{R2L}(x - DM_{L2R}(x, y), y)| \leq \delta, \tag{7}$$

where δ is a predefined threshold.

In parallel to the disparity estimation and validation process, edges are extracted in the two views using the Canny edge detector [6]. The analysis is then conducted from the view whose patch contains more edge pixels, i.e., from the sharper view. The edge pixels are then segmented at different disparity levels and only the disparity levels with enough edge pixels are further processed. A disparity level is "edge-significant" if the number of edge pixels on it is larger than 5% of the average amount of edge pixels per disparity level. The edge pixels in the two views are then matched and the edge width is estimated. The edge width is computed perpendicularly to the edge orientation using the method described in [22]. For each matched edge pixel pair p_j^d in disparity level d, a sharpness mismatch criteria $C_{sm}(p_j^d)$ based on the edge width, edge contrast, and edge disparity is evaluated in order to decide whether sharpness mismatch exists or not. For each "edge-significant" disparity level, the percentage of matched edge pixel pairs with sharpness mismatch is estimated as follows:

$$P_{sm}^d = \frac{1}{N_d} \sum_{j=1}^{N_d} \mathbb{1}_{C_{sm}(p_j^d)}, \tag{8}$$

where N_d is the number of matched edge pixel pairs in the same disparity level d. $\mathbb{1}_{C_{sm}}$ is an indicator function based on the SM criteria C_{sm}, which is equal to one if the SM criteria C_{sm} is met, and zero otherwise. The final score for the probability of sharpness mismatch PSM is obtained by averaging the P_{sm}^d of the "edge-significant" disparity levels as follows:

$$PSM = \frac{1}{N} \sum_{k=1}^{N} P_{sm}^{d_k}, \tag{9}$$

where N is the total number of "edge-significant" disparity levels, and d_k is the index of the k-th "edge-significant" disparity level.

3.3 Integration of Saliency

Saliency estimation is the process of analyzing an image to identify regions of high visual attention, i.e., regions that attract the attention of the human visual system. Saliency values for each pixel in an image are usually stored in a so called saliency map. As the end-users are free to navigate within an ODI, the saliency map represents the likelihood of the viewer's viewport location in the ODI. The saliency map is useful in order to identify regions that should have a high visual quality, i.e., regions that should be free from artifacts (e.g. sharpness mismatch) where end-users are actually looking at.

Saliency was integrated in Liu et al.'s method [21] described in Section 3.2 by weighting each pixel p based on its saliency $\psi(p)$. While processing an image, the idea is to give more weight to pixels with high visual attention of the end-users, and less weight to pixels with low visual attention. In order to give more control to the artist, the pixel saliency weight is defined as the function $g'(\psi(p))$, which can be freely chosen. In this way, the artist can decide whether to ignore the saliency by choosing a constant g', or to completely ignore the pixels with a saliency lower than a threshold, by setting the weight of these pixels to zero.

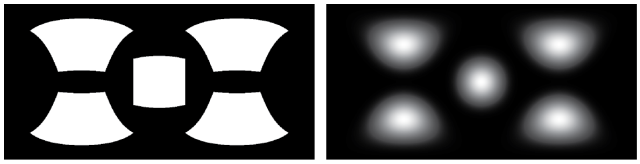


Figure 4: Projections of the viewport (left) and of the Gaussian filter (right) into the ERP format.

Based on saliency, Liu et al.'s method was modified in two components: the selection of the "edge-significant" disparity levels and the computation of the percentage of matched edge pixel pairs with sharpness mismatch defined by Equation 8. In the modified method, a disparity level d is "edge-significant", if the sum of the saliency weights $g'(\psi)$ at the edge pixels in d is larger than 5% of the average amount of saliency weights at the edge pixels per depth plane. The second modification affects Equation 8 as follows:

$$P_{sm}^d = \frac{1}{\sum_{j=1}^{N_d} g'(\psi(p_j^d))} \sum_{j=1}^{N_d} \mathbb{1}_{C_{sm}(p_j^d)}\, g'(\psi(p_j^d)), \tag{10}$$

where $g'(\psi(p_j^d))$ is the weighting function mentioned before with $\psi(p_j^d)$ equal to the saliency of the j-th matched edge pixel pair at the disparity level d.

For ODIs, Liu et al.'s method is applied independently for each patch. In order to compute the global score of the ODI, the patch scores PSM_i can be combined using the average pixel saliency Ψ_i inside each patch i. The ODI global score is then defined as follows:

$$PSM_{global} = \frac{\sum_i g''(\Psi_i)\, PSM_i}{\sum_i g''(\Psi_i)}, \tag{11}$$

where g'' is another function, which similarly to g', can be freely chosen by the artist, giving him the freedom to decide how different saliency values affect the global score.

Another useful global score is the total amount of patches with sharpness mismatch. The patch saliency Ψ_i can also be integrated in this score as follows:

$$\sum_i \mathbb{1}_{g''(\Psi_i) PSM_i \geq \rho}, \tag{12}$$

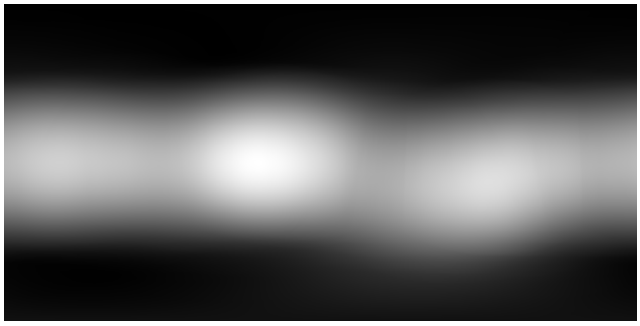
where ρ is a user-defined threshold.

For the generation of the results in Section 5, g' and g'' are piece-wise linear functions, which are defined as

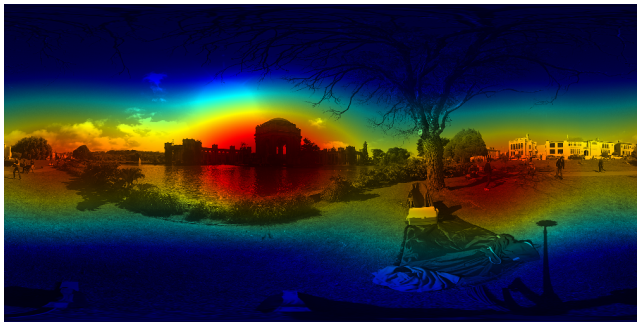
$$g'(x) = g''(x) = \begin{cases} 5x, & x \leq 0.2, \\ 1, & x > 0.2, \end{cases} \tag{13}$$

i.e., pixels and patches with a saliency of more than 20% have a full weight when calculating the scores.

(a) ERP input image.



(b) Saliency map.



(c) Combination of the ERP input image and the saliency map.

Figure 5: Example of a saliency map.

3.4 Computation of the saliency map

For the computation of the saliency maps in the ERP format, we implemented an alternative method to the one by De Abreu et al. [8], which was developed originally for monocular ODIs. The new method computes a saliency map from a sequence of HMD viewport positions recorded while a viewer is freely looking at an ODI. For each viewport position, a filter kernel centered on the viewport and defined by its dimension, is projected onto the ERP image. The projections of the filter kernels are then added in order to obtain the final saliency map.

In our approach, we use the Gaussian filter centered on the viewport, due to the fact that visual acuity is at its maximum at the center of the human visual field, i.e. the viewer tends to look at the center of the viewport rather than at the borders. The Gaussian

Figure 6: Color-based visualization of the patch scores.

filter is defined as follows:

$$h(u,v) = e^{-\frac{1}{2}\left(\frac{u^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right)} \tag{14}$$

where (u, v) are pixel coordinates centered on the HMD viewport, while σ_u and σ_v control the horizontal and vertical filter size and are related to the field of view and the resolution of the HMD viewport. Figure 4 shows five projections of the viewport and of the Gaussian filter defined on the viewport into the ODI in equirectangular format.

Figure 5 shows an example of a saliency map of a stereoscopic ODI obtained from HMD viewport data.

3.5 Visualization of Sharpness Mismatch

Once the sharpness mismatch scores for each patch are computed using Equation 9, the patch scores can be visualized directly on the ERP image. Different colormaps can be applied in order to assign a color to a scalar. In the results presented in this paper we used the jet colormap, which assigns blue to 0, red to the maximum possible score, i.e. 1, and green to 0.5.

Figure 6 illustrates an example of the visualization of patch scores on an ODI with sharpness mismatch using the jet colormap. Additionally, the PSM scores can be displayed directly within each patch to further substantiate the SM analysis for the artist.

4 SUBJECTIVE EXPERIMENT

In order to compute the saliency maps of stereoscopic ODIs using the method described in the previous section, we organized a subjective test similar to the one described in [8]. During the experiment, the participants were asked to freely look at stereoscopic ODIs while wearing an HMD, in our case an Oculus Rift DK2. While the subjects were looking at the images we recorded the viewport center locations on each of the ODIs, assuming that the center point of the viewport corresponds to the visual target location of the user. As explained in [8], this assumption is based on the fact that visual acuity is at its maximum at the center of the human visual field (fovea), and that the head tends to follow the eye movements to preserve the eye resting position (eyes looking straight ahead).

The test was divided into a training session and a test session. During the training session the subjects got familiar with the experiment, while a demo image was displayed. During the test session a set of ODIs were displayed in random order using the software application introduced in [8] which was modified to display stereoscopic

ODIs. For cross-platform compatibility reasons, the application was implemented using the WebVR and the ThreeJS APIs, so that it can run with different HMDs on different web browsers. The application is able to collect viewport information at the refresh rate of the HMD. For Oculus Rift DK2 the maximum refresh rate is 75 Hz, meaning 13.33 ms per frame.

Each image was displayed for 15 seconds, and similarly to [8], the data captured during the first second was discarded, as it adds trivial information on the starting viewing direction. A total of 17 subjects with normal stereo vision took part in the test. In order to keep anonymity we assigned an identifier to each of them. The ODIs used in the test were collected from different sources, mainly from YouTube, with resolutions ranging from 1920x960 to 4640x2320 pixels.

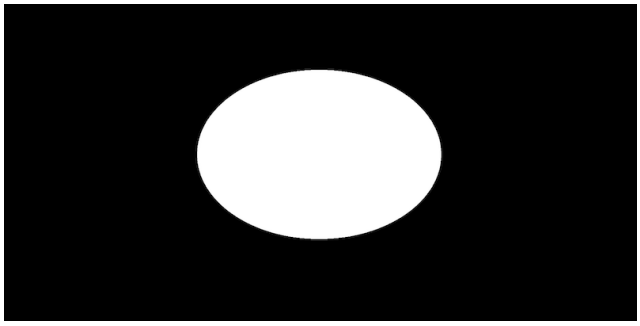
The resulting saliency maps of 10 ODIs analyzed in the subjective test are presented in the Figures 8 and 9.

5 RESULTS

In order to test our proposed method we took a stereoscopic ODI without SM and blurred a selected region of the left ODI with a Gaussian filter as proposed in [25], where it was shown that defocus-based effects of lens aberrations can be modeled with Gaussian blur. Figure 7 shows the blur mask, the blurred left ERP image and the visualization of the patch SM scores using the jet colormap. The figure shows high patch SM scores in the center of the blurred region, which decrease with the distance to the center. Thus, our approach correctly detects and localizes the SM.

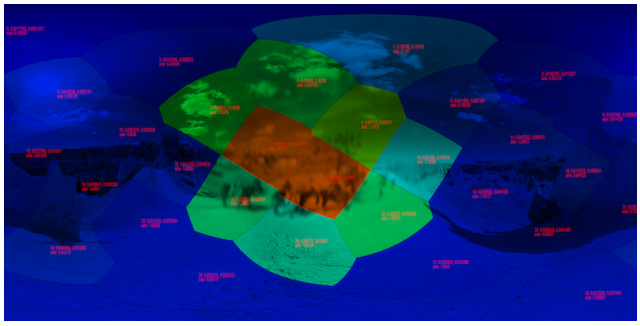
As mentioned in Section 4, we performed a subjective test with ODIs in order to compute their saliency maps. These ODIs were then evaluated with our proposed SMD method. Figures 8 and 9 show 10 ODIs, together with their saliency maps and the results of the SM analysis. The saliency maps show that the test subjects tended to look at the equator of the ODIs rather than at the pole caps. Moreover, high-level features like the bear in ODI 2 and the gun shot in ODI 5 attracted the visual attention of the subjects.

In Figure 8, we illustrate 4 stereoscopic ODIs with sharpness mismatch, while Figure 9 shows the remaining 6 stereoscopic ODIs without detected sharpness mismatch. Furthermore, for each of the 10 ODIs we computed the amount of patches with detected sharpness mismatch and the global SMD scores. The results are illustrated in two diagrams in Figure 10. The first diagram shows the amount of patches with detected sharpness mismatch obtained with Equation 12, where the threshold ρ was set to 0.2, and g'' is defined by Equation 13. As can be seen in this diagram, SM was correctly detected in the first 4 ODIs, while in the remaining ones the analysis didn't detect any patches with SM. The second diagram in Figure 10 shows the global scores computed with Equation 11. ODI 1 and ODI 4, which have the highest number of detected patches with sharpness mismatch, have the highest global SMD scores. In ODI 4 two patches with stitching and blending artifacts introducing SM were correctly identified (see close-up in Figure 8). ODI 2 has also a high global score and it is characterized by one patch with SM. ODI 3 has a lower global score than ODI 2, nevertheless our approach detected sharpness mismatch in a patch which has different glares in the left and right image (see close-up in Figure 8). Although no patch was detected with sharpness mismatch in ODI 8, the global SMD score is above average. This indicates that the entire ODI has some sharpness mismatch, but none of the individual patches passed the defined threshold. The remaining ODIs don't have patches with SM and also have relatively low global scores.



(a) Blur mask.

(b) Blurred left ERP image.



(c) Visualization of SM.

Figure 7: Results of our SMD method using Gaussian blur.

6 CONCLUSIONS

This paper presented a novel sharpness mismatch detection approach for ODIs that analyzes patches derived from the spherical Voronoi diagram. In this context, we took the method proposed by [21] and extended it to ODIs using a patch-based approach. The SM scores of each patch are visualized using a color coded representation directly in the ERP image in order to give automatic feedback of potential sharpness mismatch between the left and right images to artists in post-production.

Besides this, the paper also introduced saliency in SMD. We described how saliency can be applied by weighting each pixel.

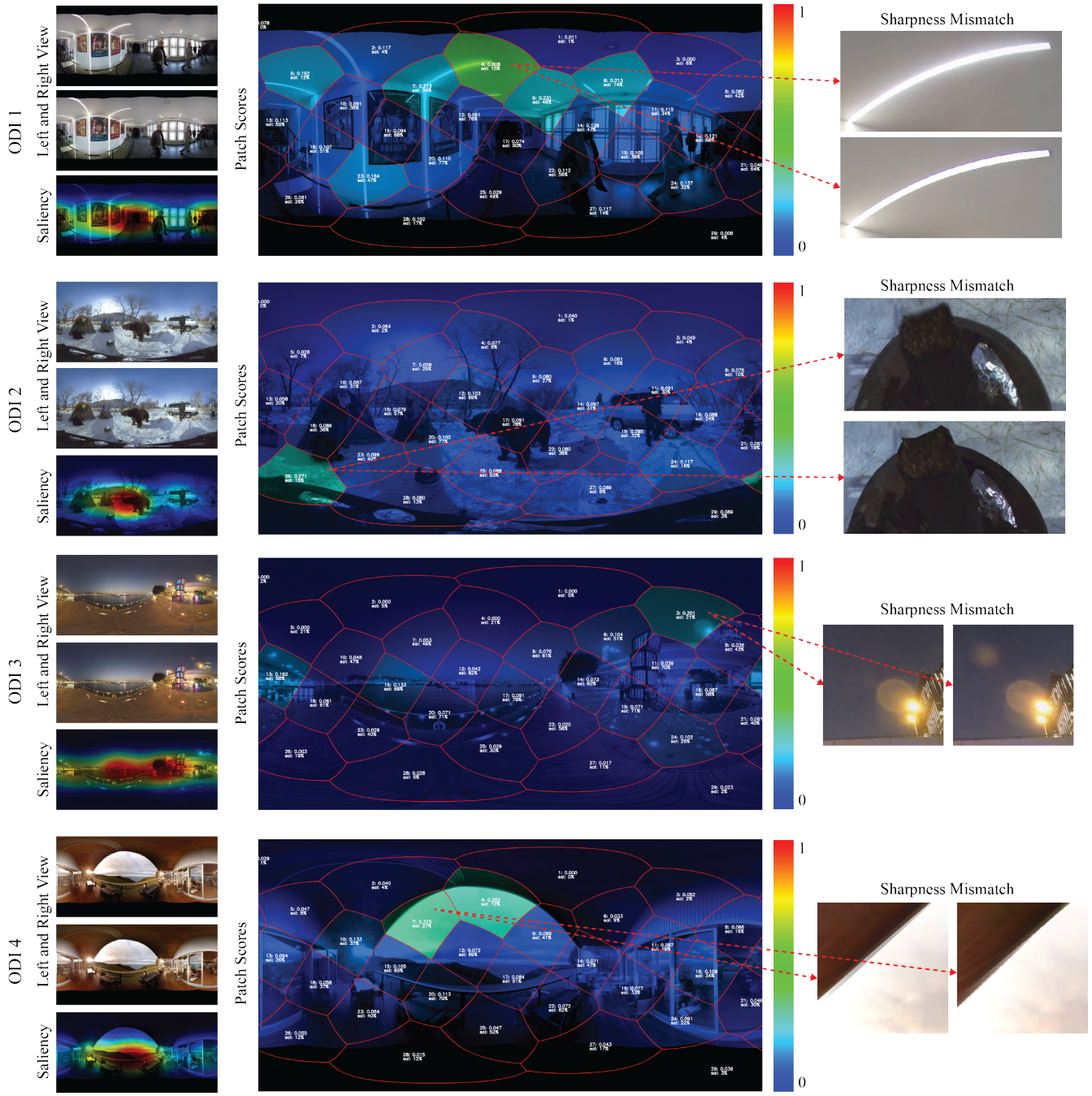


Figure 8: Examples of ODIs with sharpness mismatch.

For the computation of the saliency maps of stereoscopic ODIs, we presented an alternative to De Abreu et al.'s approach [8]. Our alternative approach is based on filtering viewport data with a Gaussian filter, and was then applied to 10 ODIs tested during a subjective experiment with 17 participants.

In order to demonstrate the performance of the proposed SMD approach, we first added blur to a region of the left ODI of a stereoscopic ODI without SM. The results demonstrated that SM could be properly detected with our method. Then, we analyzed 10 ODIs as presented in Section 5. The results demonstrated that existing artifacts could be detected in four of these ODIs. However, only

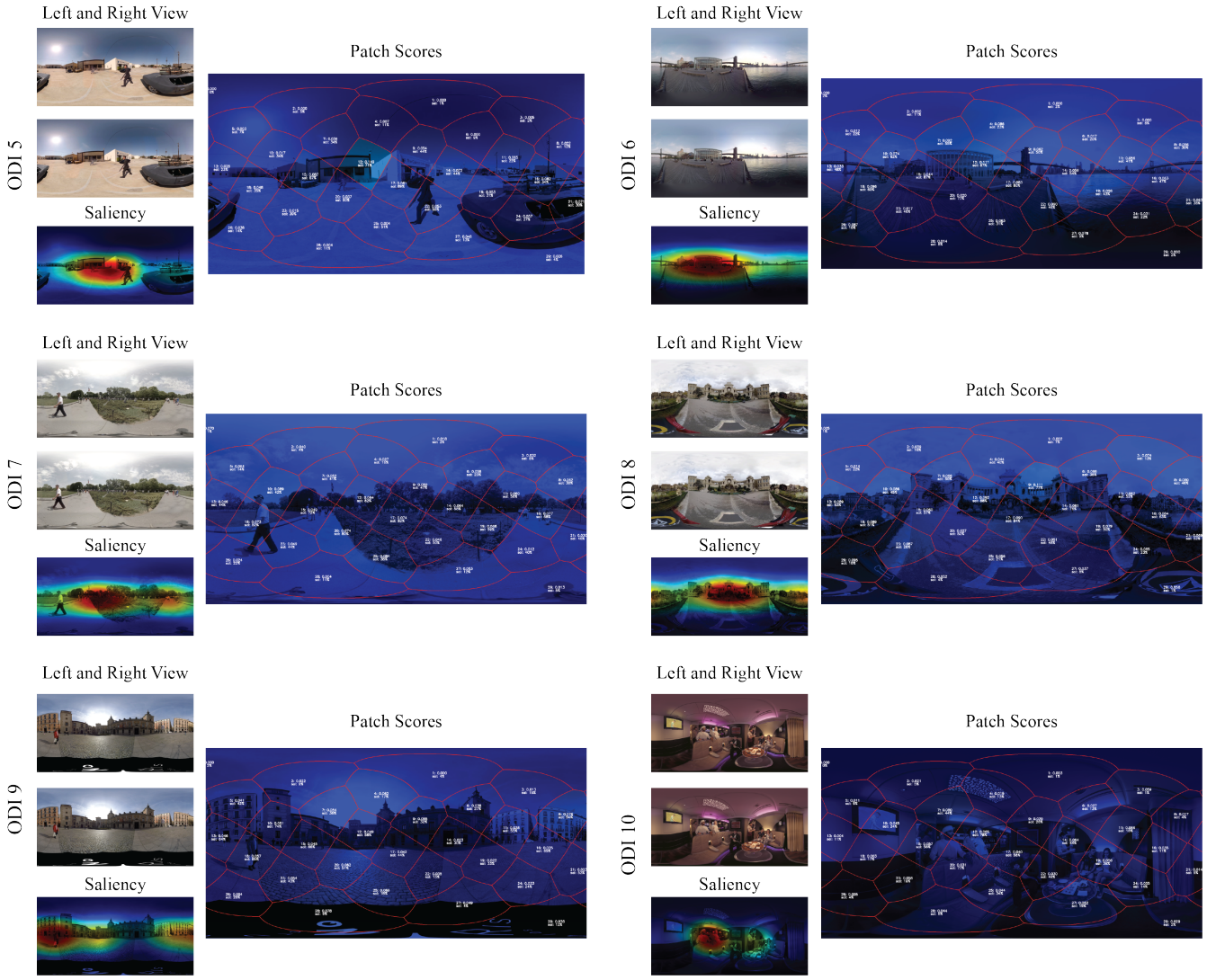
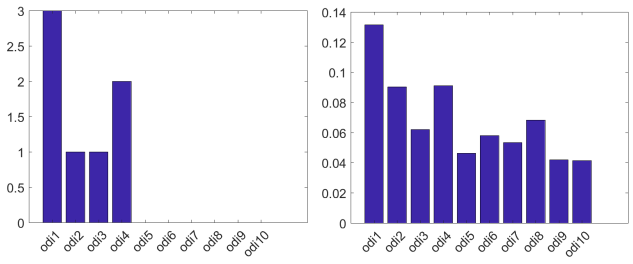


Figure 9: Examples of ODIs without sharpness mismatch.



(a) Number of patches with SM. **(b) Global SMD scores.**

Figure 10: SM patch detection and global SM scores PSM_{global}.

two of them contained SM while the other two had other binocular rivalry issues (glares and stitching artifacts), i.e., the proposed method for SMD is sensitive to artifacts which can be interpreted as SM. Moreover, the proposed method increases the efficiency in post-production workflows, by drawing the artist's attention to binocular artifacts, which was one of the motivations of this work.

In our future work, we will analyze a larger dataset of ODIs by computing new saliency maps and by applying the proposed SMD method. Furthermore, we will present a new method for SMD and compare it to Liu et al.'s method. Finally, we will extend our system to detect and, if possible correct, SM and other artifacts like e.g. color mismatch using our patch- and saliency-based approach as introduced in this paper.

ACKNOWLEDGMENTS

The authors would like to thank Ana De Abreu and Cagri Ozcinar for their support in the organization of the subjective test.

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under the Grant Number 15/RP/2776.

REFERENCES

[1] Franz Aurenhammer. 1991. Voronoi Diagrams - A Survey of a Fundamental Data Structure. *Comput. Surveys* 23, 3 (1991), 345–405. https://doi.org/10.1145/116873.116880

[2] Federica Battisti, Marco Carli, Alessio Stramacci, Atanas Boev, and Atanas Gotchev. 2015. A perceptual quality metric for high-definition stereoscopic 3D video. In *Image Processing: Algorithms and Systems XIII, 939916.* https://doi.org/10.1117/12.2086901

[3] Iva Bogdanova, Alexandre Bur, and Heinz Hügli. 2008. Visual attention on the sphere. *IEEE Transactions on Image Processing* 17, 11 (2008), 2000–2014. https://doi.org/10.1109/TIP.2008.2003415

[4] Alexander Bokov, Dmitriy Vatolin, Anton Zachesov, Alexander Belous, and Mikhail Erofeev. 2014. Automatic detection of artifacts in converted S3D video. In *Proc. of the SPIE*, Vol. 9011. 901112–901114. https://doi.org/10.1117/12.2054330

[5] Alexandre Bur, Adriana Tapus, Nabil Ouerhani, Roland Siegwart, and Heinz Hügli. 2006. Robot navigation by panoramic vision and attention guided features. In *Proceedings - International Conference on Pattern Recognition*, Vol. 1. 695–698. https://doi.org/10.1109/ICPR.2006.998

[6] J Canny. 1986. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 8, 6 (June 1986), 679–698. https://doi.org/10.1109/TPAMI.1986.4767851

[7] Kai Yueh Chang, Tyng Luh Liu, Hwann Tzong Chen, and Shang Hong Lai. 2011. Fusing generic objectness and visual saliency for salient object detection. In *Proceedings of the IEEE International Conference on Computer Vision*. 914–921. https://doi.org/10.1109/ICCV.2011.6126333

[8] Ana De Abreu, Cagri Ozcinar, and Aljosa Smolic. 2017. Look around you: saliency maps for omnidirectional images in VR applictions. *9th International Conference on Quality of Multimedia Experience (QoMEX)* May (2017). https://doi.org/10.1109/QoMEX.2017.7965634

[9] Quanwu Dong, Tong Zhou, Zongming Guo, and Jianguo Xiao. 2013. A stereo camera distortion detecting method for 3DTV video quality assessment. In *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*. 1–4. https://doi.org/10.1109/APSIPA.2013.6694209

[10] Rony Ferzli and Lina J. Karam. 2009. A no-reference objective image sharpness metric based on the notion of Just Noticeable Blur (JNB). *IEEE Transactions on Image Processing* 18, 4 (2009), 717–728. https://doi.org/10.1109/TIP.2008.2011760

[11] Sid Ahmed Fezza, Mohamed-chaker Larabi, and Kamel Mohamed Faraoun. 2014. Stereoscopic Image Quality Metric Based on Local Entropy and Binocular Just Noticeable Difference. *International Conference on Image Processing (ICIP)* (2014), 2002–2006.

[12] Xun Huang, Chengyao Shen, Xavier Boix, and Qi Zhao. 2015. SALICON: Reducing the semantic gap in saliency prediction by adapting deep neural networks. In *Proceedings of the IEEE International Conference on Computer Vision*. 262–270. https://doi.org/10.1109/ICCV.2015.38

[13] Laurent Itti and Ali Borji. 2015. Computational models: Bottom-up and top-down aspects. *The Oxford Handbook of Attention* (2015), 1–20. arXiv:1510.07748 http://arxiv.org/abs/1510.07748

[14] Laurent Itti, Christof Koch, and Ernst Niebur. 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20, 11 (1998), 1254–1259. https://doi.org/10.1109/34.730558 arXiv:arXiv:1011.1669v3

[15] Huaizu Jiang, Jingdong Wang, Zejian Yuan, Tie Liu, Nanning Zheng, and Shipeng Li. 2011. Automatic Salient Object Segmentation Based on Context and Shape Prior. *British Machine Vision Conference* (2011), 1–12. https://doi.org/10.5244/C.25.110

[16] Darya Khaustova, Jérôme Fournier, Emmanuel Wyckens, and Olivier Le Meur. 2015. An objective method for 3D quality prediction using visual annoyance and acceptability level. (2015). https://doi.org/10.1117/12.2076949

[17] Sebastian Knorr, Simone Croci, and Aljosa Smolic. 2017. A Modular Scheme for Artifact Detection in Stereoscopic Omni-Directional Images. In *Proceedings of the Irish Machine Vision and Image Processing Conference.*

[18] Sebastian Knorr, Kai Ide, Matthias Kunter, and Thomas Sikora. 2012. The Avoidance of Visual Discomfort and Basic Rules for Producing "Good 3D" Pictures. *SMPTE Motion Imaging Journal* 121, 7 (2012), 72–79. https://doi.org/10.5594/j18236

[19] Matthias Kümmerer, Thomas S. A. Wallis, and Matthias Bethge. 2016. DeepGaze II: Reading fixations from deep features trained on object recognition. *ArXiv e-prints* (2016), 1–16. arXiv:1610.01563 http://arxiv.org/abs/1610.01563

[20] Marc Lambooij, Wijnand IJsselsteijn, Marten Fortuin, and Ingrid Heynderickx. 2009. Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review. *Journal of Imaging Science and Technology* 53, 3 (2009), 030201. https://doi.org/10.2352/J.ImagingSci.Technol.2009.53.3.030201

[21] Mohan Liu, Karsten Müller, and Alexander Raake. 2016. Efficient no-reference metric for sharpness mismatch artifact between stereoscopic views. *Journal of Visual Communication and Image Representation* 39 (2016), 132–141. https://doi.org/10.1016/j.jvcir.2016.05.010

[22] Pina Marziliano, Frédéric Dufaux, Stefan Winkler, and Touradj Ebrahimi. 2004. Perceptual blur and ringing metrics: application to JPEG2000. *Sig. Proc.: Image Comm.* 19 (2004), 163–172.

[23] Niranjan D. Narvekar and Lina J. Karam. 2009. A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection. *2009 International Workshop on Quality of Multimedia Experience, QoMEx 2009* 20, 9 (2009), 87–91. https://doi.org/10.1109/QOMEX.2009.5246972

[24] Seungchul Ryu and Kwanghoon Sohn. 2014. No-reference quality assessment for stereoscopic images based on binocular quality perception. *IEEE Transactions on Circuits and Systems for Video Technology* 24, 4 (2014), 591–602. https://doi.org/10.1109/TCSVT.2013.2279971

[25] T Sakamoto. 1984. Model for spherical aberration in a single radial gradient-rod lens. *Applied optics* 23, 11 (1984), 1707. https://doi.org/10.1364/AO.23.001707

[26] Juddk Tilke, Krista Ehinger, Fredo Durand, and Antonio Torralba. 2009. Learning to predict where humans look. In *Proceedings of the IEEE International Conference on Computer Vision*. 2106–2113. https://doi.org/10.1109/ICCV.2009.5459462

[27] Dmitriy Vatolin, Alexander Bokov, Mikhail Erofeev, and Vyacheslav Napadovsky. 2016. Trends in S3D-Movie Quality Evaluated on 105 Films Using 10 Metrics. *Proceedings of the SPIE, Stereoscopic Displays and Applications XXVII* 2016, 5 (2016), 1–10.

[28] Alexander Voronov, Dmitriy Vatolin, Denis Sumin, Vyacheslav Napadovsky, and Alexey Borisov. 2013. Methodology for stereoscopic motion-picture quality assessment. In *Proceedings of the SPIE, Stereoscopic Displays and Applications XXIV*, Vol. 8648. https://doi.org/10.1117/12.2008485

[29] Jingwei Wang, Ali Borji, C. C.Jay Kuo, and Laurent Itti. 2016. Learning a Combined Model of Visual Saliency for Fixation Prediction. *IEEE Transactions on Image Processing* 25, 4 (2016), 1566–1579. https://doi.org/10.1109/TIP.2016.2522380