A FRAMEWORK FOR QUALITY CONTROL IN CINEMATIC VR BASED ON VORONOI PATCHES AND SALIENCY

Simone Croci, Sebastian Knorr*, Lutz Goldmann**, Aljosa Smolic* **

*Trinity College Dublin, **TU Berlin

ABSTRACT

In this paper, we present a novel framework for quality control in cinematic VR (360-video) based on Voronoi patches and saliency which can be used in post-production workflows. Our approach first extracts patches in stereoscopic omnidirectional images (ODI) using the spherical Voronoi diagram. The subdivision of the ODI into patches allows an accurate detection and localization of regions with artifacts. Further, we introduce saliency in order to weight detected artifacts according to the visual attention of end-users. Then, we propose different artifact detection and analysis methods for sharpness mismatch detection (SMD), color mismatch detection (CMD) and disparity distribution analysis. In particular, we took two state of the art approaches for SMD and CMD, which were originally developed for conventional planar images, and extended them to stereoscopic ODIs. Finally, we evaluated the performance of our framework with a dataset of 18 ODIs for which saliency maps were obtained from a subjective test with 17 participants.

Index Terms— 360 video, omnidirectional images, 3D quality assessment, sharpness mismatch detection, color mismatch detection, depth distribution, saliency, virtual reality

1. INTRODUCTION

Shooting a live action immersive 360-degree experience, i.e. omnidirectional images (ODIs) or 360-videos for cinematic VR, is a technological challenge 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 (HMD). The artifacts or issues can be divided into three categories: binocular rivalry issues, conflicts of depth cues and artifacts which occur in both monoscopic and stereoscopic 360-degree content production [10].

This paper introduces a novel framework for sharpness mismatch detection (SMD), color mismatch detection (CMD) and disparity distribution analysis in ODIs which can be used for quality control in post-production workflows, i.e. to give automatic feedback to artists and supervisors for reduced efforts and improved quality. Our proposed system is divided into a pre-processing step and an artifact detection and analysis step. The pre-processing step consists of three components: patch extraction, saliency estimation and dense disparity estimation. As global methods for artifact detection are often not suitable for ODIs, and to better localize and visualize artifacts, we introduce a novel patch extraction approach which is based on the spherical Voronoi diagram [1]. Besides the patch extraction, we also integrate saliency, here visual attention, in the quality control framework in order to weight artifacts depending on the visual attention of end-users who only see a small portion (viewport) of the entire ODI using their HMDs. A saliency-based artifact weighting may reduce post-production efforts as artifacts in low salient regions (i.e. regions with little attention) might not need to be corrected. Our new saliency estimation method is applied to viewport data of 18 ODIs which we obtained from a subjective test with 17 subjects.

For the actual artifact detection and analysis step, we use the pre-processed data to apply SMD, CMD and disparity distribution analysis of the scene. In particular, we applied the state of the art SMD approach by Liu et al. [13], which was originally developed for planar images, and extended it to stereoscopic ODIs. For the detection of color mismatch, we compare color properties between corresponding patches of the left and right views of the ODI using the color statistics introduced in [17]. Finally, we estimate the disparity distribution and higher order statistics like skewness and kurtosis in order to give feedback regarding the scene depth.

The paper is structured as follows. In Section 2, related work in quality assessment and artifact detection is reviewed. Then, in Section 3, we describe the proposed system for SMD, CMD and disparity distribution analysis. Results, which show the performance and usability of our proposed system with some examples, are presented in Section 4. Finally, in Section 5, the paper concludes with a discussion and future work.

2. RELATED WORK

Over recent years, binocular rivalry issues and conflicts of depth cues have been investigated in detail for planar S3D

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

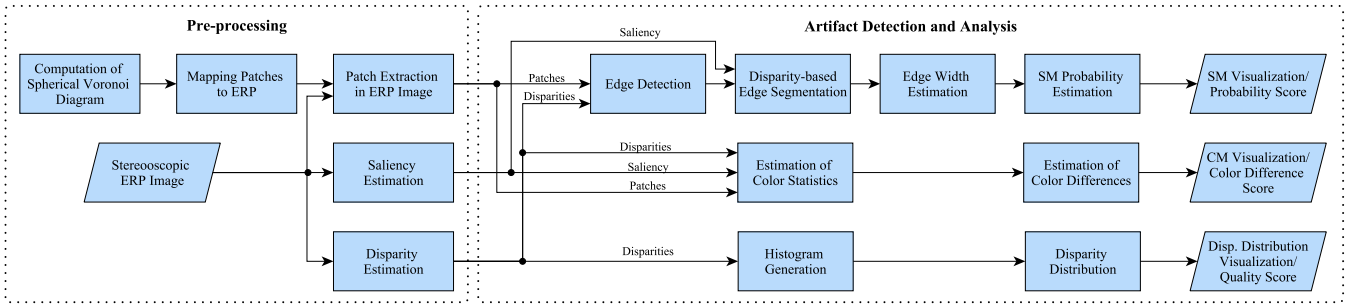


Fig. 1: Overview of the quality control system

content e.g. for cinema screens and 3D-TV [11, 20, 12, 19]. Many publications focused on the assessment of S3D quality in terms of subjective quality tests and objective quality metrics. In [9], the authors investigated how viewer annoyance depends on various technical parameters such as geometrical misalignments as well as color and luminance mismatches between the views. The authors of [4] proposed several objective metrics for luminance mismatch and evaluated their correlation with the results of subjective experiments. In [6], a method for detecting stereo camera distortions based on statistical models was presented in order to evaluate vertical misalignment, camera rotation, unsynchronized zooming, and color mismatch in S3D content. Finally, in [2], a full-reference metric was presented based on a large variety of measures 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 disparity analysis and artifact detection in order to provide direct feedback to the artist regarding the S3D quality during post-production. Thus, full-reference quality metrics can not be applied in this context. In [18], 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 which models 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 methods for the detection of sharpness mismatch (SM) and color mismatch (CM), were introduced in [21] and [3]. For SM the two papers proposed approaches that first apply dense disparity estimation and then analyze high-frequency differences between both views [21] or analyze differences of edges using a gradient-based method [3]. For measuring in-picture sharpness, different 2D metrics have been developed. In [7], 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 [16], as it outperforms most other no-reference sharpness metrics on Gaussian blur. It was developed based on human blur perception at different contrasts. For CM, the authors of [21] use the results of the disparity estimation to reconstruct one view from the other and compare the colors from the original and the reconstructed view based on the mean square error in the RGB color space.

However, none of the related work focused on S3D artifact detection in ODIs. To our knowledge only the work in [10] focused on S3D quality assessment methods that deal with ODIs. The authors analyzed vertical misalignment and global color mismatch in ODIs using the equirectangular projection (ERP), but did not consider sharpness mismatches or the disparity distribution [15] which may also influence the perceived S3D quality.

3. PROPOSED SYSTEM

Figure 1 gives a schematic overview of the proposed quality control system. It consists of a pre-processing step, which includes the extraction of patches, the saliency estimation and the dense disparity estimation, and an actual analysis and artifact detection step, which includes the disparity distribution analysis, the sharpness mismatch detection and the color mismatch detection modules. The following subsections describe all processing steps in detail.

3.1. Pre-processing

3.1.1. Voronoi Patch Extraction

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 2 shows the spherical Voronoi diagram computed from 30 evenly distributed points on the sphere, and its projection into the ERP format.

For the computation of n evenly distributed points $\mathbf{P}_i = (X_i, Y_i, Z_i)$ with $i = 0 \dots n-1$ on the sphere, we implemented a method based on the following equations: $\theta_i = i\pi \cdot \left(3 - \sqrt{5}\right), Z_i = (1 - 1/n) \cdot (1 - 2i/(n-1)), d_i =$

(a) Voronoi diagram. (b) Voronoi diagram mapped into ERP.

Fig. 2: Voronoi patch extraction.

$\sqrt{1-Z_i^2}, X_i = d_i \cdot cos(\theta_i)$ and $Y_i = d_i \cdot sin(\theta_i)$, 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 $\mathbf{P}_j$:

$$VC_i = \{\mathbf{P} \in \Omega_S \mid d_S(\mathbf{P}, \mathbf{P}_i) \leq d_S(\mathbf{P}, \mathbf{P}_j)\ \forall j \neq i\}, \quad (1)$$

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 left and right 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. As sampling can suffer from aliasing, we implemented the random supersampling antialiasing approach to minimize aliasing.

3.1.2. Saliency Estimation

For the computation of the saliency maps, we implemented an alternative method to the one introduced by De Abreu et al. [5], which was developed originally for monocular ODIs. Our method computes a saliency map from a sequence of HMD viewport positions on the ODI, obtained while a viewer is freely looking at the 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 filter is defined as follows:

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

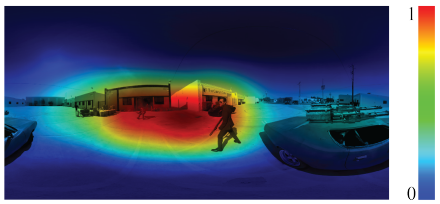


Fig. 3: Combination of the input ODI 5 and the saliency map using a jet color map (blue: low saliency, red: high saliency).

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 (FOV) and the resolution of the HMD viewport.

For our dataset used in Section 4, we performed a subjective test with 17 test subjects similar to the one described in [5]. Figure 3 shows an example of a saliency map computed from a stereoscopic ODI overlaid with the original ERP input image. Red indicates high saliency while blue indicates low saliency, i.e. users tend to look at areas close to the equator of the ODI or at areas which are exciting.

3.1.3. Disparity Estimation

As per pixel disparity information is required for SMD, CMD and disparity distribution analysis, dense disparity estimation is the third pre-processing step of our approach. To estimate disparity maps between the left and right ODI, we apply the Semi-Global Block Matching approach described in [8] which delivers good results at reasonable computational costs.

Since the disparity estimation can be noisy and inaccurate, we apply a consistency check for the disparity values, and only disparity values which are consistent are used for further computations. If DM_{L2R} and DM_{R2L} are the disparity maps from left to right view, and from right to left view, then the disparity at pixel (x, y) in DM_{L2R} is valid if

$$|DM_{L2R}(x,y) + DM_{R2L}(x - DM_{L2R}(x,y), y)| \leq \delta, \quad (3)$$

where δ is a predefined threshold.

3.2. Artifact Detection and Analysis

This section describes the actual analysis and artifact detection part of our system, which is performed after the preprocessing part. We integrate the saliency derived from Section 3.1.2 at two levels: pixel and patch level. At pixel level, we use the pixel saliency $\psi(p)$ in order to weight each pixel p that is processed using a weight equal to $g'(\psi(p))$, where g' is a function that can be freely chosen. The saliency at pixel level is used to compute the local patch scores which are computed by the sharpness and color mismatch detection modules.

At patch level, the patch saliency Ψ, which is equal to the average pixel saliency inside the patch, is used to weight the

local patch scores using a weight equal to $g''(\Psi)$, where g'' is also a function that can be freely chosen. The saliency at patch level is used to compute the global score of the ODI.

We implemented two global scores, the saliency-based weighted sum of local scores, and the number of patches with artifacts, like sharpness or color mismatch. On the one hand, the saliency-based weighted sum is defined by the following equation:

$$S_{global} = \frac{\sum_i g''(\Psi_i) S_i}{\sum_i g''(\Psi_i)}, \tag{4}$$

where S_i is the local artifact patch score, in this paper the SM or CM patch score. On the other hand, the number of patches with artifacts can be computed as follows:

$$\sum_i \mathbb{1}_{g''(\Psi_i) S_i \geq \gamma}, \tag{5}$$

where γ is a user-defined threshold and $\mathbb{1}$ is an indicator function, which is equal to one if the condition $g''(\Psi_i) S_i \geq \gamma$ is true, and zero otherwise.

For the generation of the results in Section 4, g' and g'' are piece-wise linear functions, which are defined as

$$g'(x) = g''(x) = \begin{cases} 5 \cdot x, & x \leq 0.2, \\ 1, & x > 0.2, \end{cases} \tag{6}$$

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

3.2.1. Sharpness Mismatch Detection

For sharpness mismatch detection we implemented a modified version of the SMD method introduced by Liu et al. [13], which was originally developed for planar images. The new version processes each patch independently within the ODI and takes into account the saliency information.

First, edges are extracted in corresponding patches of both views using the Canny edge detector. 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. In the original method, 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. In the modified method, however, 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 disparity level.

The edge pixels in corresponding patches of both 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 [14]. For each matched edge pixel pair p_j^d in the 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. The probability of sharpness mismatch in the disparity level d is estimated as follows:

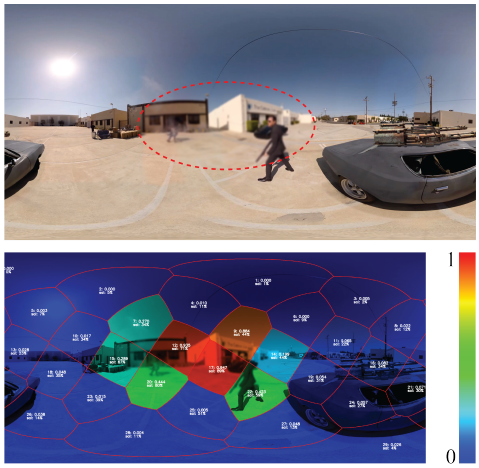


Fig. 4: Example of SM visualization: blurred left ODI 5 using a Gaussian filter (top) and SM visualization (bottom) including a text overlay with the SM patch score and the saliency.

$$P_{sm}^d = \frac{1}{\sum_{j=1}^{N_d} g'(\psi_i(p_j^d))} \sum_{j=1}^{N_d} \mathbb{1}_{C_{sm}(p_j^d)} g'(\psi_i(p_j^d)), \tag{7}$$

where N_d is the number of matched edge pixel pairs in the disparity level d and $g'(\psi_i(p_j^d))$ is the weighting function based on the saliency as introduced in Equation 6. $\mathbb{1}_{C_{sm}}$ 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 patch score for the probability of a sharpness mismatch PSM_i in a patch i is obtained by averaging P_{sm}^d of the "edge-significant" disparity levels as follows:

$$PSM_i = \frac{1}{N} \sum_{k=1}^{N} P_{sm}^{d_k}, \tag{8}$$

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. The global SM score of the ODI is then determined with Equation 4 by substituting S_i with PSM_i.

Figure 4 shows the left ODI blurred with a Gaussian filter in the center, and the visualization of the patch SM scores using a jet color map. As can be noticed, our approach correctly detects and localizes the SM.

3.2.2. Color Mismatch Detection

The color mismatch detection module compares the color properties between corresponding patches of left and right views. In the first step, pixels which are present in both corresponding patches of the two views are detected using the disparity maps validated by Equation 3. Let's assume that Ω_L and Ω_R

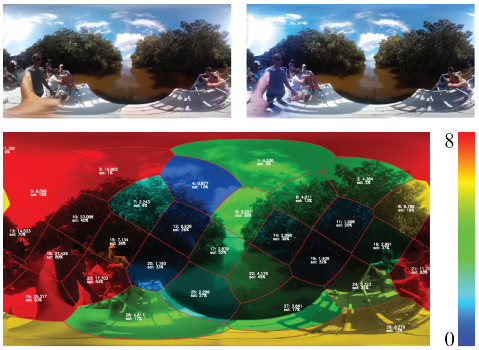


Fig. 5: Example of CM visualization: left and right ODI 11 (top) and CM visualization (bottom) including a text overlay with the CM patch score and the saliency.

are common regions in the left and right view. Then, two corresponding pixels $p(x, y) \in \Omega_L$ and $p(x', y) \in \Omega_R$ belong to the same region, if $x' = x - DM_{L2R}(x, y)$. For all pixels belonging to Ω_L and Ω_R, the color properties, i.e. mean and standard deviation of the color channels, are computed as introduced by [17]. Instead of using the $l\alpha\beta$ color space, as proposed by the authors, we extract the same statistics from the Lab color space.

Assuming that $I_L(p)$ and $I_R(p)$ are the colors at the pixel p defined in the Lab color space in the left I_L and right I_R view, then the saliency-based means for each color channel are defined as

$$\mu_X = \frac{1}{\sum_{p \in \Omega_X} g'(\psi_i(p))} \sum_{p \in \Omega_X} I_X(p) \cdot g'(\psi_i(p)), \quad (9)$$

where $X \in \{L, R\}$, and $g'(\psi_i(p))$ is the weighting function based on the saliency as introduced in Equation 6. The saliency-based standard deviations for each color channel are defined as

$$\sigma_X = \frac{1}{\sum_{p \in \Omega_X} g'(\psi_i(p))} \sum_{p \in \Omega_X} (I_X(p) - \mu_X)^2 g'(\psi_i(p)). \quad (10)$$

Finally, the color mismatch score for patch i is computed with:

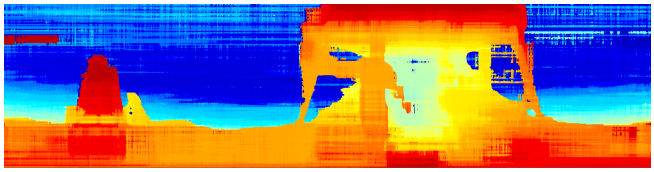
$$CMS_i = \sqrt{\|\mu_L - \mu_R\|^2 + \lambda \|\sigma_L - \sigma_R\|^2}, \quad (11)$$

where λ is a tuning parameter and set to one for the generation of the results. The global CM score of the ODI is then determined with Equation 4 by substituting S_i with CMS_i.

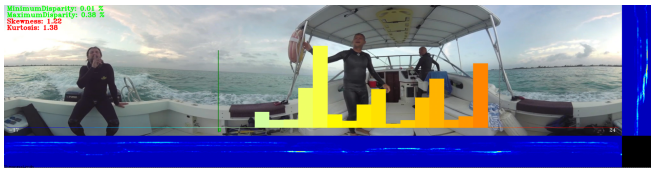
Figure 5 shows the left and right input ODIs and the visualization of the patch CM scores using a jet color map. As can be noticed, our approach correctly detects and localizes the CM.

3.2.3. Disparity Distribution

While color and sharpness mismatch may impact the perceived 3D quality, it is also largely influenced by the proper



(a) Disparity map (red: foreground, blue: background)



(b) Disparity histograms: global (center), horizontal (bottom), vertical (right)

Fig. 6: Example of disparity distribution analysis for ODI 12.

use of the available disparity budget, i.e. the range between minimum and maximum disparity. The optimal disparity range depends mainly on the screen size and the viewing distance. Violating it may cause hyper-convergence or hyper-divergence and increase the vergence-accommodation conflict, and lead to discomfort and fatigue [19]. Within this range, the distribution of the disparity values has a large impact on the 3D effect including the perceived depth within (intra) and between (inter) objects.

The disparity distribution module analyzes the distribution of the disparities for the whole frame or selected regions. For the visualization, global as well as local (horizontal and vertical) histograms are extracted by quantizing the disparities and counting the amount of pixels per disparity level across the whole frame or individual rows and columns, respectively.

For further evaluation, the disparity distribution is described with different measures including the disparity range as well as skewness and kurtosis. The skewness describes the symmetry of the distribution and is defined as the ratio of the 3rd moment m_3 and the standard deviation σ as $s = m_3/\sigma^3$. A skewness of 0 indicates a balanced (middleground) disparity distribution while a negative/positive skewness corresponds to a disparity distribution focused on the background/foreground. The (excess) kurtosis describes the uniformity of the distribution and is defined as the ratio of the 4th moment m_4 and the standard deviation σ as $s = m_4/\sigma^4 - 3$. A kurtosis of 0 indicates a disparity distribution that follows a normal distribution while a negative/positive kurtosis corresponds to a more uniform/peaky distribution, respectively. The normalized statistical moments are defined as $m_k = 1/n \sum_i (x_i - \mu_x)^k$ with the disparity values x_i and the mean disparity μ_x.

Figure 6 shows an example of the estimated disparity map and the different disparity histograms together with the statistical measures as overlay. Note that only 90° of the vertical FOV of the ERP images were used for the disparity analysis due to the large distortions of the pole caps and because pole caps were not captured for some of the ODIs (e.g. Panocam POD 3D and Odyssey).

Analysis	Measurements	ODI																	
		1[a]	2[b]	3[c]	4[d]	5[e]	6[c]	7[e]	8[a]	9[f]	10[g]	11[h]	12[g]	13[c]	14[c]	15[c]	16[h]	17[a]	18[f]
SMD	*No. of patches with SM*	3	1	1	2	0	0	0	0	0	0	0	0	0	1	0	0	0	0
	Global SM Score	0.132	0.090	0.062	0.091	0.046	0.058	0.053	0.068	0.042	0.041	0.081	0.045	0.083	0.059	0.058	0.078	0.069	0.107
CMD	*No. of patches with CM*	15	6	10	13	4	2	2	7	0	0	20	0	6	4	4	0	22	3
	Global CM Score	2.463	1.438	2.115	2.401	1.365	1.117	1.145	1.766	0.910	0.065	7.583	0.096	1.639	1.411	1.218	0.785	3.519	1.038
Disparity	*Min. disparity in %*	-1.89	-1.53	-0.16	-1.40	-1.23	-1.01	-1.62	-1.70	-0.70	-0.48	-3.99	-0.01	-1.53	-0.65	-0.55	-3.33	-0.67	-4.09
	Max. disparity in %	2.58	0.04	0.09	1.06	0.09	-0.01	0.48	0.74	0.57	-0.04	0.16	0.43	0.01	-0.01	-0.01	0.31	0.50	0.31
	Skewness	-1.14	-2.47	1.35	0.41	-1.28	-1.94	-1.23	1.12	-1.79	-1.06	-2.52	1.27	-1.57	-1.72	-1.43	-4.28	1.32	-2.45
	Kurtosis	1.93	6.15	1.31	2.82	4.43	4.58	2.33	4.05	2.61	2.42	9.51	1.39	3.11	2.36	1.95	39.02	3.59	8.59

a) Panocam POD 3D, b) Ozo, c) Odyssey, d) unknown, e) Jaunt, f) GoPro, g) 3D conversion, h) Vuze VR

Table 1: Results of our system for SMD, CMD and disparity analysis (min. and max. values in % relative to the image width).

4. RESULTS

In order to demonstrate the performance and usability of our proposed system, we evaluated the quality of a dataset of 18 ODIs by measuring SM, CM and disparity distribution as described in the previous section. Table 1 shows the computed global scores for SM and CM and the number of patches with detected artifacts according to Equation 4 and Equation 5, respectively, for the entire dataset taking saliency into account. The values with the highest scores are highlighted in red while the ones with the lowest scores are highlighted in green. ODI 10, which was converted from 2D to 3D in post, has the lowest scores and no patch with SM or CM was detected. SM and CM are very unlikely for post-converted images as the stereoscopic views are generated using depth-image-based rendering (DIBR). ODI 1 has the highest SM score and also a high CM score. This ODI was captured with Panocam's POD 3D which uses 9 stereo camera pairs capturing the left and right view of the ODI independently (see Figure 7 for visualization of the SMD in ODI 1). ODI 11 was captured with the Vuze VR camera which uses 4 stereo camera pairs. Here the SM scores are average while the CM global score is the highest of all ODIs under evaluation (see Figure 5 for visualization of the CMD in ODI 11). ODI 11 and ODI 17 are the two ODIs which are characterized by more than 20 patches with CM.

In Figure 7, we exemplary show three ODIs with SM detected by our method, together with their saliency maps and close-ups of detected regions. As shown in the figure, SM was correctly detected and highlighted in ODI 1 and ODI 2. In ODI 4, a blending artifact was detected as SM. Figure 8 exemplary shows two ODIs with detected CM, together with their saliency maps and close-ups of detected regions. Similar to SM, our proposed system detects and highlights patches with CM correctly. Furthermore, the patch scores and the saliency values of the patches are displayed in each patch for both SMD and CMD.

Finally, Table 1 also shows the results of the disparity analysis module, i.e. the minimum and maximum disparities in percentage of the ERP image width as well as the skewness and kurtosis for each of the ODIs. At this point, we would like to note that the percentage values of the disparities need to be based on the width of the viewport in order to compare them with disparity ranges of 3DTV or 3D cinema content. As an example, the disparity budget of -4.09% to 0.31% for ODI 18 would be -16.36% to 1.24% for the viewport of an HMD with 90° horizontal FOV. This results in hyper-convergence and causes high visual discomfort for the end-user.

Figure 9 illustrates the disparity distribution of ODI 16 which has a high positive kurtosis, i.e. Laplacian (peaky) distribution. Here, the hand of the person (with a negative disparity of -3.33%, i.e. -13.23% within the viewport) and the person itself are very close to the camera rig while the rest of the scene is at far distance. Some of the test subjects mentioned that this close object was quite annoying.

5. CONCLUSION

With the increasing number of HMDs and omnidirectional capture systems, cinematic VR is becoming a popular trend in immersive media. However, creating error-free stereoscopic 360-videos is still a challenging task.

In this paper we presented a novel framework and system for quality control in cinematic VR applications, i.e. of ODIs and 360-videos. The developed tools can be used to monitor the quality and detect artifacts within post-production workflows.

The ODIs are preprocessed in order to ease the subsequent quality analysis and artifact detection. This includes the partitioning into spherical Voronoi patches for accurate detection and localization of artifacts, the extraction of saliency information to weight artifacts according to the visual attention, and the estimation of dense disparity maps.

The current framework contains three modules for further quality evaluation. The sharpness and color mismatch detection modules detect local deviations between the two views of the stereoscopic ODI. The disparity distribution module analyzes the estimated disparity values and calculates the disparity distribution as well as its skewness and kurtosis.

The developed modules were tested and evaluated on a comprehensive dataset of 18 ODIs originating from different omnidirectional capture systems. The evaluation shows that meaningful and reliable results can be achieved by applying the patch and saliency based approaches for the different analysis modules.

In the future, we will further evaluate the overall framework and the individual modules through subjective tests with potential end-users in order to find reliable thresholds for assessing the quality. We will also develop additional modules

to detect other issues commonly found in 360-video such as stitching and blending artifacts, and compare our modules with state of the art approaches for conventional S3D content. Finally, we will apply saliency prediction for automatic saliency map generation as saliency maps optained from subjective tests are not practical in post-production workflows.

6. REFERENCES

[1] Franz Aurenhammer. Voronoi Diagrams - A Survey of a Fundamental Data Structure. *ACM Computing Surveys*, 23(3):345–405, 1991.

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

[3] Alexander Bokov, Dmitriy Vatolin, Anton Zachesov, Alexander Belous, and Mikhail Erofeev. Automatic detection of artifacts in converted S3D video. In *Proc. of the SPIE*, volume 9011, pages 901112–901114, 2014.

[4] J Chen, J Zhou, J Sun, and A C Bovik. Binocular mismatch induced by luminance discrepancies on stereoscopic images. In *IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, 2014.

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

[6] Quanwu Dong, Tong Zhou, Zongming Guo, and Jianguo Xiao. A stereo camera distortion detecting method for 3DTV video quality assessment. In *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pages 1–4, 2013.

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

[8] Heiko Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(2):328–341, February 2008.

[9] Darya Khaustova, Jérôme Fournier, Emmanuel Wyckens, and Olivier Le Meur. An objective method for 3D quality prediction using visual annoyance and acceptability level. In *Proc of the SPIE*, volume 9391, 2015.

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

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

[12] Marc Lambooij, Wijnand IJsselsteijn, Marten Fortuin, and Ingrid Heynderickx. Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review. *Journal of Imaging Science and Technology*, 53(3):030201, 2009.

[13] Mohan Liu, Karsten Müller, and Alexander Raake. Efficient no-reference metric for sharpness mismatch artifact between stereoscopic views. *Journal of Visual Communication and Image Representation*, 39:132–141, 2016.

[14] Pina Marziliano, Frédéric Dufaux, Stefan Winkler, and Touradj Ebrahimi. Perceptual blur and ringing metrics: application to jpeg2000. *Sig. Proc.: Image Comm.*, 19:163–172, 2004.

[15] Bernard Mendiburu. *3D Movie Making*. Tayler & Francis, 2009.

[16] Niranjan D. Narvekar and Lina J. Karam. 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):87–91, 2009.

[17] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Comput. Graph. Appl.*, 21(5):34–41, September 2001.

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

[19] Takashi Shibata, Joohwan Kim, David M Hoffman, and Martin S Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision*, 11(8):1–29, 2011.

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

[21] Alexander Voronov, Dmitriy Vatolin, Denis Sumin, Vyacheslav Napadovsky, and Alexey Borisov. Methodology for stereoscopic motion-picture quality assessment. In *Proceedings of the SPIE, Stereoscopic Displays and Applications XXIV*, volume 8648, 2013.

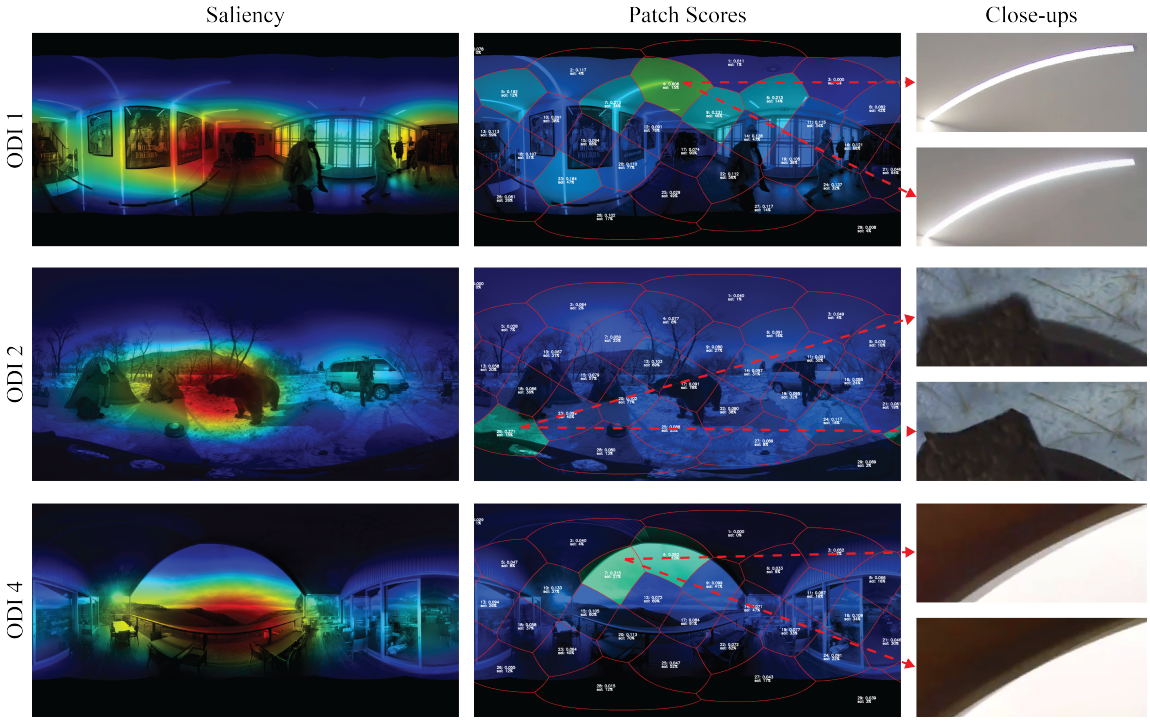


Fig. 7: Examples of ODIs with sharpness mismatch.

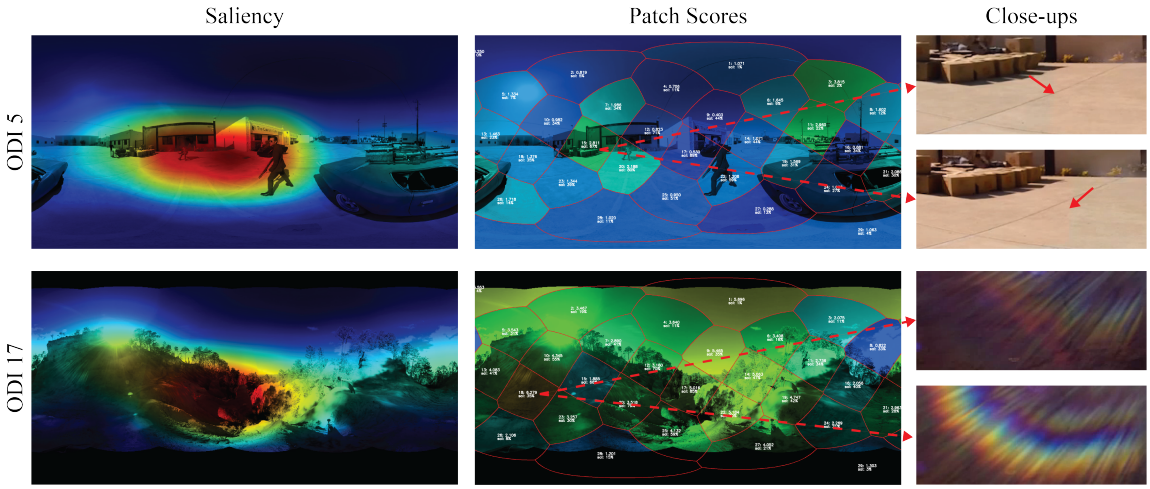


Fig. 8: Examples of ODIs with color mismatch.

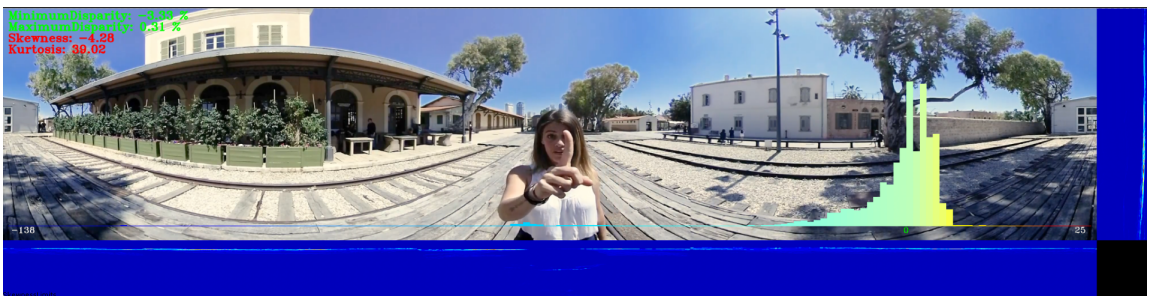


Fig. 9: Disparity distribution of ODI 16.