https://doi.org/10.2352/ISSN.2470-1173.2020.11.HVEI-128

Quality Evaluation of 3D Objects in Mixed Reality For Different Lighting Conditions

Jesús Gutiérrez, Toinon Vigier, and Patrick Le Callet
Image, Interaction, Perception Group, Laboratoire des Sciences du Numérique de Nantes (LS2N), Université de Nantes, France

Abstract

This paper presents a study on Quality of Experience (QoE) evaluation of 3D objects in Mixed Reality (MR) scenarios. In particular, a subjective test was performed with Microsoft HoloLens, considering different degradations affecting the geometry and texture of the content. Apart from the analysis of the perceptual effects of these artifacts, given the need for recommendations for subjective assessment of immersive media, this study was also aimed at: 1) checking the appropriateness of a single stimulus methodology (ACR-HR) for these scenarios where observers have less references than with traditional media, and 2) analyzing the possible impact of environment lighting conditions on the quality evaluation of 3D objects in mixed reality (MR), and 3) benchmark state-of-the-art objective metrics in this context. The subjective results provide insights for recommendations for subjective testing in MR/AR, showing that ACR-HR can be used in similar QoE tests and reflecting the influence among the lighting conditions, the content characteristics, and the type of degradations. The objective results show an acceptable performance of perceptual metrics for geometry quantization artifacts and point out the need of further research on metrics covering both geometry and texture compression degradations.

Introduction

The recent emergence of consumer devices for visualizing 3D objects (e.g., meshes and point clouds) has boosted the applications of immersive technologies like mixed reality (MR) and augmented reality (AR). As with any other audiovisual technology, to assure a successful development, Quality of Experience (QoE) evaluation is crucial, especially considering the perceptual factors related to immersiveness and interactivity provided by emerging immersive technologies. In this sense, the research on QoE evaluation requires the development of new subjective test methodologies and new properly characterized datasets. In fact, the existing studies on subjective quality evaluation with immersive media generally make use of traditional methodologies originally developed for image/video quality. Thus, the research community is actively working towards the definition of proper testing methodologies, taking into account the new technical and perceptual factors offered by these technologies. Regarding QoE evaluation of 3D objects in MR/AR scenarios, together with methodological factors, test environment aspects, like lighting conditions, play an important role, given the use of see-through displays.

Lately, some works have been already presented studying the quality effects of typical degradations that affect the geometry and texture of 3D meshes and point clouds (e.g., noise, compression, etc.) [1–3]. In this sense, subjective tests have been mainly conducted considering classical double-stimulus methodologies, given that the observers are less used to this type of content than to traditional media. Also, these tests have been performed using normal TV displays showing videos with pre-calculated trajectories (e.g., rotation around the objects) [1–5]. Since these techniques are not representative of real use cases, other approaches offer the observers some interaction, although limited and not totally natural for MR/AR scenarios, such as in the study of Mekuria *et al.* [6] where users were represented avatars visualized on a TV screen, or in the work of Torlig *et al.* [7], where a rendering software was used to show 2D projections of the 3D point clouds. Recently, thanks to the arrival of affordable Head Mounted Displays (HMDs) for virtual and augmented reality applications, some studies start to consider the evaluation of QoE visualizing 3D objects with these systems in more natural scenarios. For instance, Alexiou *et al.* [8] used a MR HMD system for quality evaluation of point clouds, but only considering geometry degradations, as Yu *et al.* [9] did with meshes. Also, Zhang *et al.* [10] proposed a study towards a QoE model to evaluate holographic AR devices using the Microsoft HoloLens and two AR applications, focused on technical and perceptual factors related to the usability (and not to quality) of these systems. Furthermore, some efforts have been done to evaluate the quality of 3D meshes and point clouds using objective metrics [1, 5, 11, 12], but generally the results do not correlate well with subjective scores and the performance is highly content-dependent [8].

Taking this into account, this paper aims at providing insights for guidelines and recommendations for subjective testing of QoE in MR/AR scenarios, using HMD displays and considering both texture and geometry distortions by: 1) using the single-stimulus methodology ACR-HR (Absolute Categorical Rating with Hidden Reference) [13], 2) checking the impact of ambient lighting conditions in this context, and 3) checking the performance of state-of-the-art metrics in this scenario. Thus, the rest of the paper is organized as follows. Firstly, the details of the subjective test are provided. Secondly, the obtained results are reported and discussed, together with the benchmark of the objective metrics. Finally, the conclusion and future work are presented.

Subjective experiment

Test Material

The 3D meshes used for this subjective test were drawn from the LIRIS Textured Mesh Database [1], which is online available. After a pre-test with experts visualizing the different test

Figure 1: Screenshots of the four models used.

Content	Geometry Quantization (bits)			Texture - JPEG Compression (%)		
	Q1	Q2	Q3	T1	T2	T3
Hulk	9	8	7	14	10	8
HeadStatue	9	8	7	16	12	8
Car	9	8	7	5	3	1
Dwarf	10	9	8	10	8	6

Table 1: Properties of the test stimuli

stimuli of the dataset, and considering the aim of the present experiment and the limited duration of the whole test described in this section, four source (SRC) models (among the five published in the dataset) were selected ("Hulk", "HeadStatue", "Car" and "Dwarf", see screenshots in Fig. 1), covering a variety of texture and geometry properties. Also, two types of degradations and three quality levels for each one were chosen among those included in the dataset: 1) compression of the geometry using uniform geometric quantization, and 2) JPEG compression of the texture. This resulted in seven different quality versions (or HRCs) of each model, including the reference (R), and a total 28 test stimuli. The main details of the selected HRCs are indicated in Table 1, where the three levels of geometry quantization (Q1-Q3, from best to worst quality) are indicated in bits, and the three levels of texture compression (T1-T3, from best to worst quality) are reported in terms of the JPEG quality factor.

Environment and Equipment

The 3D objects were visualized using Microsoft HoloLens in a test room following ITU standardized conditions [13]. The objects were rendered thanks to Unity3D software (version 2017.3.1f1). They were illuminated with a directional light from the upper-right part of the observer at 45 degrees (position: X=2, Y=2, Z=0; rotation: X=45,Y=-45, Z=0; intensity=1, and rendering with soft shadows) and with scene ambient light (Skybox with intensity multiple set to 1.5). Also, default processing settings from Unity3D were deactivated, such as anti-aliasing. All 3D objects were static and they were displayed at a distance of 2 meters, in front of the observer (X=0, Y=0.1, Z=2). Observers could physically move around and approach the 3D objects, up to the near clipping plane at 0.85 meters as it is recommended by HoloLens and represented in Fig. 2.

Three different background light conditions were considered, which were obtained with dimmable ceiling lights. The three levels were measured on the starting position of the observers at an average height with the sensor pointing towards the location of the 3D objects, obtaining 285 cd/m^2 for the brightest level, 70 cd/m^2 for the medium and 10 cd/m^2 for the darkest (they were also measured on the floor being 75, 25 and 3 cd/m^2, respectively). Also, the default settings of Hololens for display brightness were used.

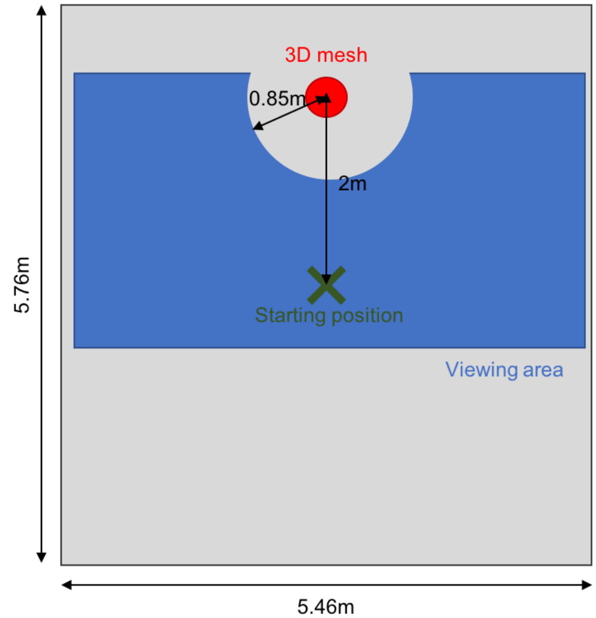

Figure 2: Diagram of the test room.

Figure 3: Photo of the test scene.

Evaluation Methodology

The visual quality of the 3D stimuli was assessed using the ACR-HR methodology. It consists in rating stimuli independently on a five-grade category scale, from 5 (Excellent) to 1 (Bad). The hidden reference condition means that the test procedure includes a reference (R) version (undistorted) of each source content shown as test stimuli.

The observers were asked to freely explore the displayed 3D models (moving around and closer to the objects within the limits of the viewing area depicted in Fig. 2) and rate the perceived quality. Each object was displayed for 15 seconds, and then a grey cube was displayed showing the number corresponding to the observed object, so the users were able to rate its quality writing a mark in the corresponding box in a paper questionnaire. After that, they clicked in the box to continue with the next object using the HoloLens remote control. Also, the observers were asked to go back to the starting position to rate the objects, so they start exploring all the objects from the origin.

The whole test was split in three sessions to evaluate the 28 objects with the three different lighting conditions. Thus, the total length of each session was around 10 minutes. It is worth noting that the order of the lighting conditions was randomized for each observer, and a different random order of the sequence of test stimuli was used in each session (with the condition of not showing the same source content twice consecutively). After each session, the observers took a small break and filled the Simulator Sickness Questionnaire (SSQ) [14] to evaluate their discomfort, visual fatigue, and sickness, and to check if the length of the test was acceptable in this sense.

Furthermore, a training session was done before the test to explain to the participants the procedure, to show them some examples of 3D objects (using two models different from those used in the test) with different levels of the degradations considered in the test, and to make them familiar with the headset, the exploration of the objects within the interaction area, and the test and rating methodology.

Observers

A total of 24 participants took part in the test (11 women, 13 men), with ages ranging between 21 and 55 (average of 29). Vision screening was carried out before the test, to assure that they have normal or corrected-to-normal vision. A post-screening of the results provided by the participants was performed that led to discard one observer. Also, the participants were asked to fill a questionnaire about their experience in using VR/AR headsets, which showed that 25% of the participant were using it the first time, 50% of them had used it less than 5 times, 17% had used it between 5 and 20 times, and 8% had used it more than 20 times.

Results

Subjective Test Results

The Mean Opinion Scores (MOS) from the quality ratings provided by the observers in the questionnaires are shown in Fig. 4, together with the 95% confidence intervals. Given the space limitations, each sub-figure represents the results obtained for each 3D object and all the distortions (both geometry and texture) evaluated in the tests.

As shown in Fig. 4, on one side, for geometry distortions, the used methodology make possible the perception of the quantization effects for all test contents, and the scores obtained for the different quality levels are well distributed along the scale used in the test. This behavior is in line with the results reported by the authors of the dataset in their subjective tests [1], even though the differences between both studies make difficult a direct comparison of the results (e.g., they used a TV screen to show rendered videos of the 3D objects and they used a double-stimulus methodology). On the other side, for the texture degradations, our results show a lesser discriminability among the different quality levels, covering a limited range of the five-grade MOS scale between around 2.5 and 4.5. These degradation levels were selected by the authors of the dataset to cover a wide range of perceptual quality [1] and their subjective results show a better distribution than in our case. Therefore, although a double-stimulus methodology would definitely help in identifying differences among the different quality levels, another important factor to consider is the HMD technology (e.g., low resolution of HoloLens systems, properties of the glasses, etc.), which can make necessary the adjustment of the quality levels for VR/MR scenarios.

To further analyze the statistical significance of the results, given that they were not all normally distributed, Friedman tests (i.e., non-parametric version of the repeated-measures ANOVA) were performed, together with post-hoc Wilcoxon Signed-Rank tests. Thus, to compare each pair of test conditions, the sets of raw scores provided by the observers were used and, a significance level of 95% have been considered for all the analyses in the paper. To consider the effect of multiple comparisons, the Bonferroni correction was applied to the threshold "p-value", so the common 0.05 value was divided by the number of conditions compared in each case [15].

Firstly, the impact of the lighting for each HRC was analyzed, showing statistical significance in few pairs, both for geometry quantization and texture compression, reported in Table 2. For this analysis, as three levels of light were compared for each HRC, the threshold "p-value" for statistical significance was $p = 0.017$. Although Fig. 4 reflects very similar results for each test stimulus with the three levels of lighting conditions considered in our MR test, some significant differences were found both for texture and geometry degradations for "HeadStatue" and "Dwarf" (see Table 2), and a particular behavior for the medium level of quantization (i.e., Q2) for those contents, where the MOS values increase with the brightness of the light. These facts suggest an influence of background lighting conditions and the contents' nature on the perceived quality, given the properties of the noisy meshes "HeadStatue" and "Dwarf" (obtained from reconstruction and scanning) [1], which should be explored with more objects in future tests.

Then, the impact of the levels of distortions considered in the tests for each lighting conditions was similarly analyzed. On one side, for each content and lighting condition, each pair of degradations coming from geometry quantization was compared using the Wilcoxon Signed-Rank test. The results are summarized in Table 3, where all the non-significantly different pairs are shown (the threshold "p-value" for statistical significance was $p = 0.0125$, given that four quality levels were considered). On the other side, a similar analysis was performed considering the degradations resulting from JPEG texture compression. In this case, given that there is no clear predominance of significantly or non-signignificalt different pairs, the "p-values" resulting from the Wilcoxon Signed-Rank test are shown for all the compared conditions in Table 4. As before, the threshold "p-value" for statistical significance was $p = 0.0125$, so the significant results are marked in bold. The results show a potential impact on the discriminability among different quality levels for a given object, and thus, on the design of subjective tests in MR scenarios. For example, almost all the non-significant differences between two quantization levels appear with medium and bright lights (see Table 3), which suggest that dark lighting conditions make easier to perceive quantization effects, especially on the high-quality range. Similarly, environment lighting affects the noticeability of texture degradations, since from the significantly different pairs reported in Table 4 (especially for "HeadStatue" and "Dwarf"), in the high-quality range it is easier to perceive degradations with dark light,

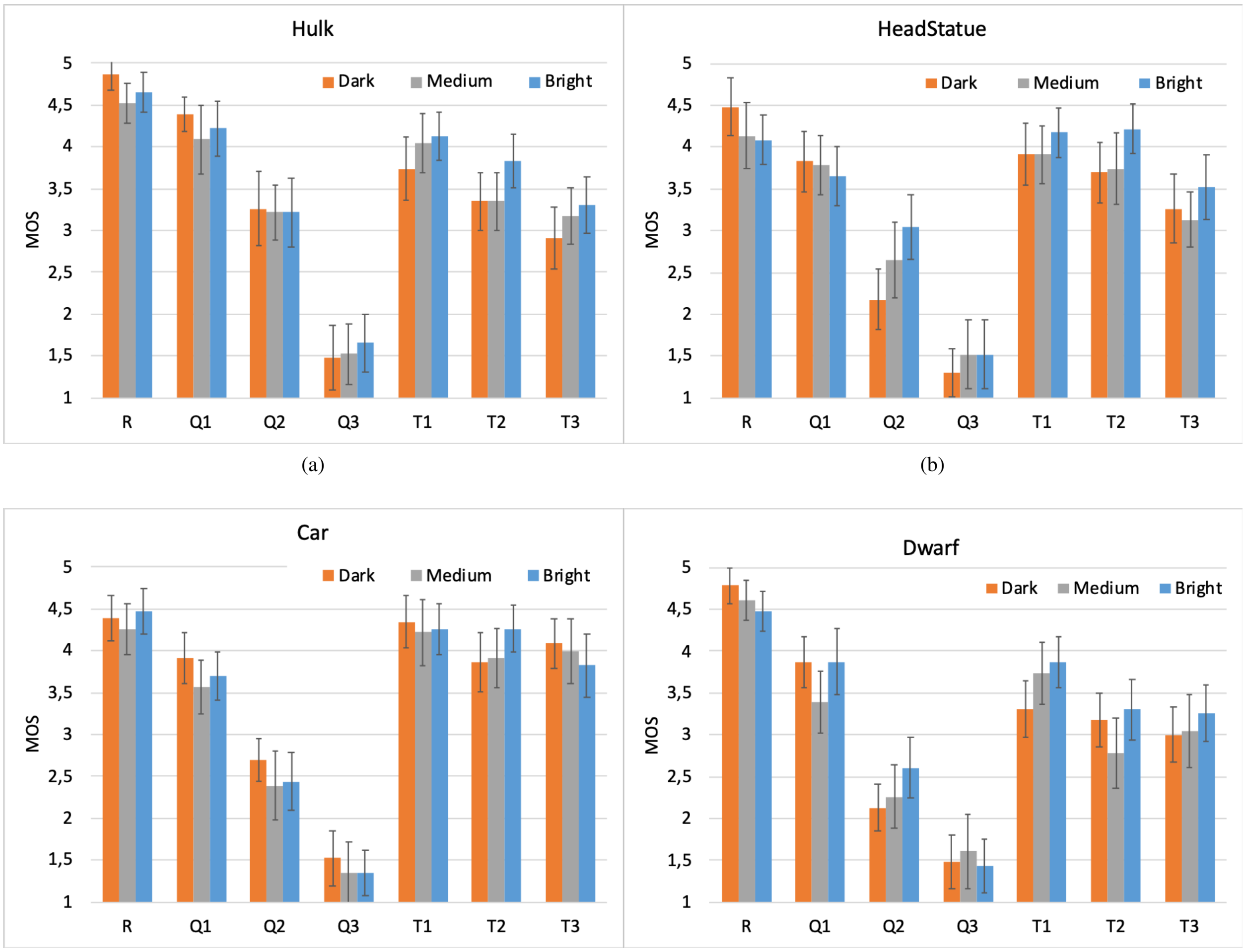

Figure 4: MOS scores for the test conditions evaluated for each 3D object.

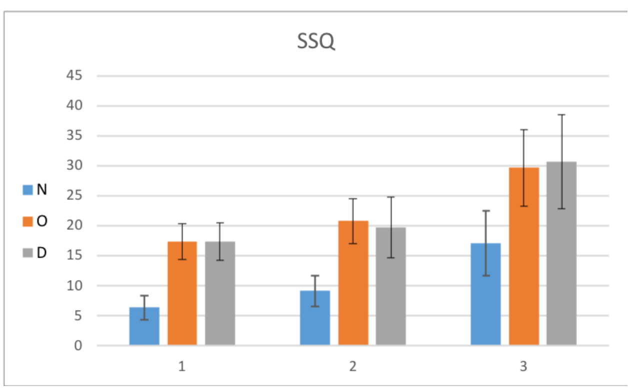

Figure 5: SSQ results.

while in the low-quality range bright light helps on noticing them.

Finally, the results from the SSQs are depicted in Fig. 5, which show an augmentation of symptoms on Nausea (N), Oculomotor (O) and Disorientation (D) scores from session 1 to session 3 [14]. However, only one observer scored symptoms as severe during the last session. Also, from these results, we can assume that the duration of the test was acceptable, while they also imply the necessity to consider fatigue and sickness in subjective testing in MR in relation with the duration of experiments.

Content	Degradation	Compared Lights	p-value
HeadStatue	Q2	Dark vs. Bright	0.003
HeadStatue	T2	Dark vs. Bright	0.012
Dwarf	Q1	Dark vs. Medium	0.016
Dwarf	T1	Dark vs. Bright	0.005

Table 2: Statistically significant pairs from the Wilcoxon test for the impact of light conditions

Objective Results

The performance results of state-of-the-art metrics on estimating the quality perceived by the observers in the considered MR scenario are shown in Table 5. On one side, two metrics for geometry distortions (the commonly used Hausdorff distance[1] and the good-performing perceptual metric MSDM [16]) were

[1]Hausdorff distance: Mean result with respect to the bounding box obtained with Meshlab: http://www.meshlab.net/. MSDM: Computed with the Mesh Processing Plateform (MEPP): https://projet.liris.cnrs.fr/mepp/index.html

Content	Light	Compared Conditions	p-value
Hulk	Medium	R vs. Q1	0.057
Hulk	Bright	R vs. Q1	0.027
HeadStatue	Medium	R vs. Q1	0.201
HeadStatue	Bright	R vs. Q1	0.125
HeadStatue	Bright	Q1 vs. Q2	0.029
Car	Dark	R vs. Q1	0.031
Dwarf	Bright	R vs. Q1	0.032

Table 3: All non-significant pairs from Wilcoxon test for the impact of geometry quantization

Content	Light	Compared conditions					
		R-T1	R-T2	R-T3	T1-T2	T1-T3	T2-T3
Hulk	Dark	**0.000**	**0.000**	**0.000**	0.0456	**0.004**	0.099
	Med.	0.033	**0.001**	**0.000**	**0.011**	**0.001**	0.381
	Bright	**0.011**	**0.001**	**0.000**	0.092	**0.001**	0.022
Head Statue	Dark	0.017	**0.001**	**0.002**	0.431	**0.009**	0.082
	Med.	0.305	0.154	**0.004**	0.526	**0.001**	0.016
	Bright	0.745	0.513	0.018	0.734	**0.012**	**0.005**
Car	Dark	1.000	0.039	0.118	0.034	0.285	0.270
	Med.	1.000	0.097	0.460	0.232	0.180	0.694
	Bright	0.392	0.340	0.020	1.000	0.065	0.063
Dwarf	Dark	**0.000**	**0.000**	**0.000**	0.479	0.198	0.344
	Med.	**0.001**	**0.000**	**0.000**	**0.000**	0.030	0.502
	Bright	0.026	**0.001**	**0.000**	**0.006**	**0.010**	0.960

Table 4: Results from the Wilcoxon test for the impact of JPEG texture compression

computed over the test stimuli resulting from geometry quantization. On the other side, two image quality metrics (PSNR and SSIM [17]) were computed over the texture images resulting from JPEG compression. Also, to analyze the performance of a combined metric, a linear combination of the best performing metrics (i.e., MSDM and SSIM) for geometry and texture quality was also applied over the whole set of test contents [1]. The performance of the metric is reported in terms of Pearson Linear Correlation Coefficient (PLCC), Spearman Rank Order Correlation Coefficient (SROCC) and Root-Mean-Sqaure Error (RMSE) between the outputs of the metrics and the MOS obtained form the subjective tests (per light condition and aggregated), as recommended in ITU-T P.1401 [18]. A cubic polynomial function was used on the objective scores to compare them with the subjective ones, since it provided the best performance for all metrics[2].

The correlations of the objective scores and the subjective MOS show that: 1) the impact of geometry quantization are acceptably estimated by dedicated objective metrics, especially perceptual metrics as MSDM [16]; 2) image quality metrics on texture images do not provide a good performance, given that the effects related to the mapping of the texture on the geometry are not considered and, as shown by the subjective results, the impact of texture compression is more reduced in our MR scenario; 3) a simple combination of geometry and texture metrics can provide an acceptable prediction, but, given the greater influence of geometry distortions in our test set, this should be further explored, especially with mixed distortions. Regarding the influence of light conditions on the performance of the objective metrics, it can be

[2]The coefficients of the mapping function and the linear combination of MSDM and SSIM were obtained with the Matlab function *fminsearch*. The outlier ratio was also computed but no outliers were obtained.

Light	Metric	PLCC	SROCC	RMSE
Dark	Haussdorf (G)	0.817	0.844	0.768
	MSDM (G)	0.906	0.862	0.563
	PSNR (T)	0.612	0.616	0.517
	SSIM (T)	0.737	0.476	0.359
	MSDM & SSIM (G & T)	0.888	0.759	0.493
Medium	Haussdorf (G)	0.835	0.851	0.632
	MSDM (G)	0.887	0.795	0.531
	PSNR (T)	0.527	0.584	0.532
	SSIM (T)	0.665	0.807	0.410
	MSDM & SSIM (G & T)	0.857	0.719	0.517
Bright	Haussdorf (G)	0.855	0.839	0.633
	MSDM (G)	0.899	0.895	0.533
	PSNR (T)	0.440	0.411	0.467
	SSIM (T)	0.506	0.487	0.398
	MSDM & SSIM (G & T)	0.887	0.779	0.489
All	Haussdorf (G)	0.844	0.825	0.657
	MSDM (G)	0.905	0.860	0.519
	PSNR (T)	0.527	0.562	0.501
	SSIM (T)	0.657	0.680	0.371
	MSDM & SSIM (G & T)	0.881	0.742	0.484

Table 5: Performance of the objective metrics for each type of degradation (G: Geometry, T: Texture), and considering the different light conditions and aggregating them.

observed that the performance of texture metrics decrease as light conditions are brighter, obtaining the best performance for dark conditions. On the other hand, the performance of the metrics for geometry distortions do not show a general trend depending on light conditions, in particular MSDM (perceptual metric) seems to perform better in dark and bright conditions and Haussdorf with bright light, but all differences are small. These results emphasize the need of metrics more sensitive to light effects when visualizing the objects in MR scenarios.

Conclusion

A study of QoE evaluation of 3D objects with geometry and texture degradations in MR was performed, based on a subjective test using HoloLens and a benchmark of state-of-the-art metrics. The results showed: 1) the appropriateness of the ACR-HR methodology for similar tests, 2) the effects of ambient lighting conditions on the perceptibility of degradations, and the influence on these effects of the content's nature and the type of degradations, and 3) the acceptable performance of geometry metrics and the need of metrics for both geometry and texture artifacts. Future work will be oriented to: 1) evaluate the quality of dynamic 3D objects and mixed degradations, 2) make available an extensive dataset, 3) compare with other methodologies, and 4) develop and benchmark objective quality metrics.

Acknowledgments

The work of J. Gutiérrez was supported by the People Programme (Marie Curie Actions) of the European Union's 7th Framework Programme (FP7/2007-2013) under REA grant agreement n. PCOFUND-GA-2013-609102, through the PRESTIGE Programme coordinated by Campus France.

References

[1] J. Guo, V. Vidal, I. Cheng, A. Basu, A. Baskurt, and G. Lavoue. Subjective and Objective Visual Qual-

ity Assessment of Textured 3D Meshes. *ACM Transactions on Applied Perception*, 14(2):1–20, Oct. 2016. https://perso.liris.cnrs.fr/guillaume.lavoue/data/datasets.

[2] A. Javaheri, C. Brites, F. Pereira, and J. Ascenso. Subjective and Objective Quality Evaluation of Compressed Point Clouds. In *IEEE Int. Workshop on Multimedia Signal Processing*, Luton, UK, Oct. 2017.

[3] E. Alexiou, M. V. Bernardo, L. A. Da Silva Cruz, L. G. Dmitrovic, C. Duarte, E. Dumic, T. Ebrahimi, D. Matkovics, M. Pereira, A. Pinheiro, and A. Skodras. Point cloud subjective evaluation methodology based on 2D rendering. In *International Conference on Quality of Multimedia Experience*, Cagliari, Italy, May 2018.

[4] J. Zhang, W. Huang, X. Zhu, and J. N. Hwang. A subjective quality evaluation for 3D point cloud models. In *International Conference on Audio, Language and Image Processing*, pages 827–831, Jul. 2014.

[5] E. Zerman, P. Gao, C. Ozcinar, and A. Smolic. Subjective and Objective Quality Assessment for Volumetric Video Compression. In *IS&T Electronic Imaging, Image Quality and System Performance XVI*, Burlingame, USA, Jan. 2019.

[6] R. Mekuria, K. Blom, and P. Cesar. Design, Implementation, and Evaluation of a Point Cloud Codec for Tele-Immersive Video. *IEEE Transactions on Circuits and Systems for Video Technology*, 27(4):828–842, Mar. 2017.

[7] E. M. Torlig, E. Alexiou, T. A. Fonseca, R. L. de Queiroz, and T. Ebrahimi. A novel methodology for quality assessment of voxelized point clouds. In *SPIE Applications of Digital Image Processing*, San Diego, USA, Sep. 2018.

[8] E. Alexiou, E. Upenik, and T. Ebrahimi. Towards Subjective Quality Assessment of Point Cloud Imaging in Augmented Reality. In *IEEE Int.. Workshop on Multimedia Signal Processing*, Luton, UK, Oct. 2017.

[9] Y. Yu, F. Yang, H. Liu, and W. Zhang. Perceptual Quality and Visual Experience Analysis for Polygon Mesh on Different Display Devices. *IEEE Access*, 6:42941–42949, Jul. 2018.

[10] L. Zhang, H. Dong, and A. El Saddik. Towards a QoE Model to Evaluate Holographic Augmented Reality Devices: A HoloLens Case Study. *IEEE MultiMedia*, Oct. 2018.

[11] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro. Geometric distortion metrics for point cloud compression. In *IEEE International Conference on Image Processing*, Beijing, China, Sep. 2017.

[12] A. Doumanoglou, P. Drakoulis, N. Zioulis, D. Zarpalas, and P. Daras. Benchmarking Open-Source Static 3D Mesh Codecs for Immersive Media Interactive Live Streaming. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 9(1):190–203, Mar. 2019.

[13] ITU. Subjective video quality assessment methods for multimedia applications. ITU-T Rec. P.910, Apr. 2008.

[14] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The Int. Journal of Aviation Psychology*, pages 203–220, Nov. 2009.

[15] John H. MacDonald. *Handbook of Biological Statistics*. Sparky House Publishing, Baltimore, Maryland, 2008.

[16] G. Lavoue and M. Corsini. A Comparison of Perceptually-Based Metrics for Objective Evaluation of Geometry Processing. *IEEE Transactions on Multimedia*, 12(7):636–649, Nov. 2010.

[17] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, Apr. 2004.

[18] ITU-T. Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models. Rec. P.1401, Jul. 2012.

Author Biography

Jesús Gutiérrez *received the Telecommunication Engineering degree (five-years engineering program) in 2008 from the Universidad Politécnica de Valencia, Spain. He received the Ph.D. degree in Telecommunication in 2016 from the Universidad Politécnica de Madrid, Spain. Since 2016, he is a post-doctoral researcher (Marie Curie fellow, currently under PRESTIGE Programme, previously within the PROVISION ITN) at LS2N laboratory of the Université de Nantes, France. His research interests are in the area of image and video processing, immersive media, evaluation of multimedia quality of experience, human behavior and visual perception.*

Toinon Vigier *obtained a PhD in July 2015 from the École Centrale de Nantes in the Ambiances Architectures and Urbanity lab, where she focused on virtual reality for urban studies. Since September 2016, she is an Associate Professor at Université de Nantes in the Image Perception Interaction research team of the Laboratory of Digital Sciences in Nantes (LS2N). Her research mainly focuses on the study, the analysis and the prediction of the quality of experience for immersive and interactive multimedia through subjective and objective measures. She is currently involved in various national and international interdisciplinary projects focusing on user experience in immersive VR media for various applications (health, cinema, architecture, design…). She is also active in the standardization working group IEEE 3333.1 and she served as reviewers in a lot of international conferences and journals.*

Patrick Le Callet *received the M.Sc. and Ph.D. degrees in image processing from the École Polytechnique de l'Université de Nantes. He was an Assistant Professor from 1997 to 1999 and a full time Lecturer from 1999 to 2003 with the Department of Electrical Engineering, Technical Institute of the Université de Nantes. He led the Image and Video Communication Laboratory, CNRS IRCCyN, from 2006 to 2016, and was one of the five members of the Steering Board of CNRS, from 2013 to 2016. Since 2015, he has been the Scientific Director of the cluster Ouest Industries Cratives, a five-year program gathering over ten institutions (including three universities). Since 2017, he has been one of the seven members of the Steering Board of the CNRS LS2N Laboratory (450 researchers), as a Representative of Polytech Nantes. He is mostly involved in research dealing with the application of human vision modeling in image and video processing. His current centers of interest are quality of experience assessment, visual attention modeling and applications, perceptual video coding, and immersive media processing.*