Complexity measurement and characterization of 360-degree content

Francesca De Simone[1], Jesús Gutiérrez[2], Patrick Le Callet[2]

[1] *Distributed & Interactive Systems Group, Centrum Wiskunde & Informatica (CWI) , Amsterdam, The Netherlands*

[2] *Image, Interaction, Perception Group, Laboratoire des Sciences du Numérique de Nantes (LS2N), Université de Nantes, France*

Abstract

The appropriate characterization of the test material, used for subjective evaluation tests and for benchmarking image and video processing algorithms and quality metrics, can be crucial in order to perform comparative studies that provide useful insights. This paper focuses on the characterisation of 360-degree images. We discuss why it is important to take into account the geometry of the signal and the interactive nature of 360-degree content navigation, for a perceptual characterization of these signals. Particularly, we show that the computation of classical indicators of spatial complexity, commonly used for 2D images, might lead to different conclusions depending on the geometrical domain used to represent the 360-degree signal. Finally, new complexity measures based on the analysis of visual attention and content exploration patterns are proposed.

Introduction

The selection of visual content, (i.e., test material), in order to validate and benchmark the performance of image and video processing algorithms, can be crucial for the successful development of new technologies. Meaningful image and video datasets are required to evaluate the performance of compression and streaming algorithms and objective quality metrics, as well as to design subjective quality assessment experiments [1]. Despite its importance, content selection has been often done based on convenience or content availability [2]. This is also due to the lack of clear and comprehensive guidelines for appropriate content characterization and selection.

A simple classification of the content can be based on how the content is generated (e.g., synthetic vs. natural), its technical properties (e.g., duration, resolution, frame rate, etc.), and its semantic (e.g., sports, news, movies, etc.). In addition, specific measures have been proposed in the literature, for example to quantify the spatial and temporal complexity of a 2D video sequence [3], the colorfulness of the visual information [4], depth features [5] and disparity [6] for 3D content, or to reflect perceptual issues related to the visualization for high dynamic range [7] and wide color gamut content [8].

These measures take into account intrinsic features of the given signal. Therefore, their usage might be not appropriate when new kinds of signals are considered, so the design of new measures might be needed. For example, recently, new measures have been proposed to characterize light-field content, taking into account the full-parallax and refocusing possibilities offered by this type of signals [9], as well as the navigation trajectories of the users exploring it [10].

This paper focuses on the characterisation of an emerging kind of visual signals: *360-degree*, i.e., *omnidirectional*, images [11][1]. These images are becoming widely spread nowadays for Virtual Reality (VR) and immersive applications. They are captured by cameras with a field of view of 360-degree, which can be modelled as central cameras projecting a point in the 3D space to a point on a spherical imaging surface [13]. Thus, an omnidirectional image can be considered as a signal lying on a sphere. In practice, the signal is stored as a rectangular array of samples, referred to as *planar representation*, obtained by projecting the sphere to a plane. Such representation allows to re-use existing processing algorithms designed for planar images and videos [14]. Accordingly, existing measures, commonly used to characterise 2D and 3D images, can be computed on the planar representation. Nevertheless, depending on the projection used, these might provide inaccurate indications upon the perceptual characteristics of the 360-degree visual content. In this paper, we analyse such inaccuracy, addressing the question: *does the geometrical domain of the 360-degree image matter when characterizing the signal using existing measures?* We limit the analysis to the spatial characterization of 360-degree images.

Regardless of the planar representation used for storage, processing and distribution, the planar signal is mapped back to its native spherical domain when presented to the end user. Particularly, at each instant, the portion of the sphere in the field of view of the user is rendered by projecting it to a plane, i.e., the *viewport*, which corresponds to the display of the rendering device. When consumed via a Head Mounted Display (HMD), the viewports (one per eye) are seamlessly updated following the users head movements to provide a sense of presence. The interactive and immersive dimension, characteristic of 360-degree content rendering, represents a significant novelty with respect to existing imaging technologies. In fact, it has practical implications that have been exploited for video compression and streaming [15, 16] and might also influence the characterisation of the content. For example, 360-degree content could be characterized depending on whether the attention of the observers is guided by the storytelling [17] or the content is purely exploratory [18]. So, in this paper, we also address the question: *can 360-degree content be characterized based on users' navigation and visual attention patterns?* We show preliminary results on new indexes that could be used to this scope.

To analyze these two questions, we consider as a case

---

[1] At the best of our knowledge, the only study addressing the problem of 360-degree content characterization is that by Afzal *et al.* [12], where 360-degree videos available on an online platform are classified from a content provider perspective, based on statistics on video duration, resolution, bitrate, and motion vectors.
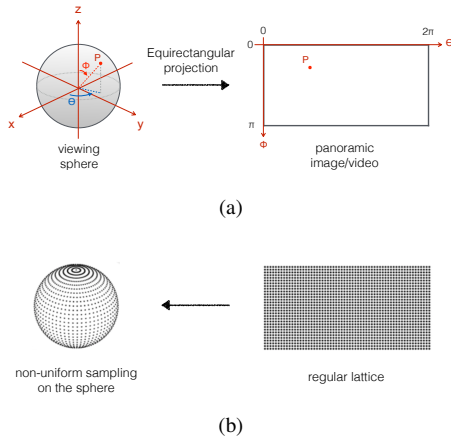
(a)

(b)

Figure 1: Equirectangular planar representation of a 360-degree image (a) and corresponding sampling density on the sphere (b).

of study the "Salient360!" dataset [19], which contains eighty-five omnidirectional images with the corresponding tracking data (saliency maps and scanpaths) from head and eye movements of sixty-three observers.

The rest of the paper is organized as follows. First, a brief overview of the most used planar representations for 360-degree content is presented. Then, the two research questions are addressed, providing insights on how to characterize 360-degree images. Finally, conclusions and future-work directions are discussed.

Planar representations

The sphere is a non-developable surface: projecting visual content from the sphere surface to a plane inevitably introduces distortions [20]. Consequently, any planar representation of a 360-degree image is not a typical natural 2D signal.

Most of the existing commercial 360-degree cameras store the captured images in the so-called *equirectangular* (ER) planar representation. This corresponds to a projection that samples the spherical signal on an equi-angular grid and uses the longitude and latitude of each sample on the sphere as coordinates of the sample projected on the plane (Fig. 1(a)). Such representation presents strong warping distortions around the polar areas (Fig. 2(a)) and corresponds to a highly non-uniform sample distribution on the spherical surface (Fig. 1(b)).

Other planar representations, characterized by less strong warping distortions, have been proposed for 360-degree content compression and streaming [14]. They are based on the projection of the visual information from the spherical surface to the faces of a polyhedron enclosing the sphere. The more faces the polyhedron has, the more subtle the warping distortions of the visual content projected on each face are [20]. Nevertheless, the faces are usually arranged on a plane [14]. Therefore, increasing the number of faces, increases the number of discontinuities that appear in the planar representation, at borders between faces or due to zero padding.

Nowadays, the *cube map* (CM) planar representation is commonly used for content delivery. It is the result of the gnomonic projection [20] of the visual samples from the spherical surface to the faces of a cube enclosing the sphere (Fig. 3 (a)). The vi-

(a)   (b)

Figure 2: A 360-degree image in equirectangular (a) and cube map (b) representation.
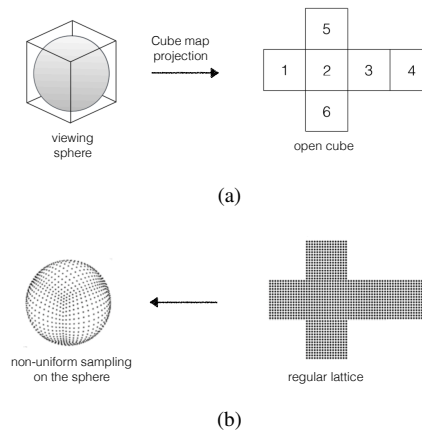

(a)

(b)

Figure 3: Cube map planar representation of a 360-degree image (a) and corresponding sampling density on the sphere (b).

sual content on each cube face can be considered approximately as a perspective image, with subtle warping distortions. The cube faces are then arranged in a rectangular shape, without need for zero padding, for example as depicted in Fig. 2(b). This representation is more compact than the ER: the uncompressed image can represent the same equatorial sampling density as an ER one with 25% less pixels. It also corresponds to a more uniform density of samples in the native spherical domain (Fig. 3 (b)). Nevertheless, the arrangement of cube faces in a rectangular lattice is inevitably characterised by visual discontinuities, due to the mosaicking of non-neighbouring faces.

Does the geometrical domain of the 360-degree image matter for its charecterization?

We consider the spatial characterization of 360-degree images, as one example to analyze whether the geometrical domain of the signal matters. A measure that is commonly computed to characterize perspective images is the *spatial information* (SI) [21]. This is an indicator of edge energy, thus spatial complexity of the content. It is usually computed on the luminance channel of the signal, as the root mean square of the edge magnitude over the image: $SI = \sqrt{\sum s_r^2}$ with $s_r = s_h^2 + s_v^2$, where s_h and s_v denote the gray-scale images, filtered with horizontal and vertical Sobel kernels, respectively, and s_r represents the edge magnitude at every pixel [2].

The SI computed in ER domain might lead to a misleading perceptual charecterization of the test material, due to the strong warping of the visual content around the poles. Similarly, when

---

[2] To avoid the scale/resolution-dependence of SI and compare images with different resolutions, a normalization factor is sometimes considered [1].

Domain	ER	CM-faces	CM-mosaic
ER	-	0.97534	0.96476
CM-faces	0.97534	-	0.97586
CM-mosaic	0.96476	0.97586	-

Table 1: SROCC coefficients between the SI values computed in ER domain, in CM domain as mean across cube faces (CM-faces), and in CM domain on the mosaicked image (CM-mosaic), for a set of 360-degree images [19].

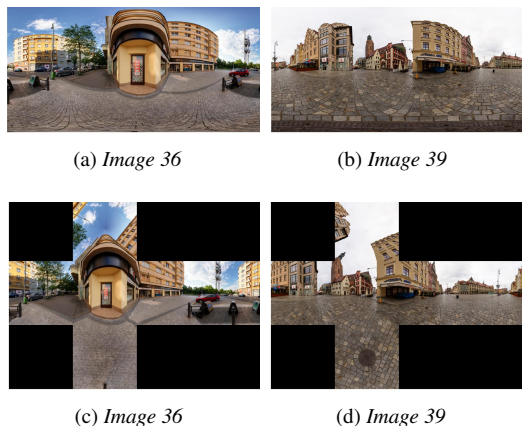(a) *Image 36*

(b) *Image 39*

(c) *Image 36*

(d) *Image 39*

Figure 4: Example of inconsistent ranking depending on planar domain in which the SI is computed: *Image 36* has an higher SI than *Image 39* in ER domain, while the opposite is true in CM-faces domain.
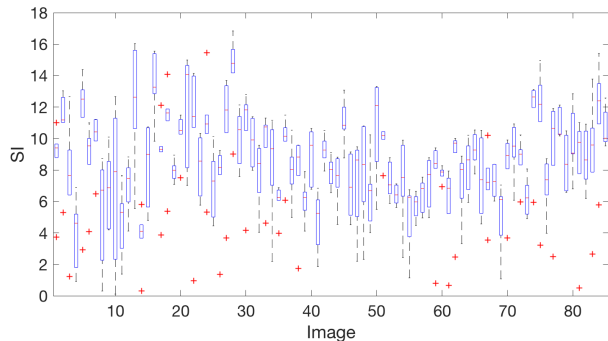

Figure 5: Boxplot of the SI values across all cube faces, for each of the eighty-five images in the set: the SI varies significantly across the cube faces in some images, while it is quite constant in others.
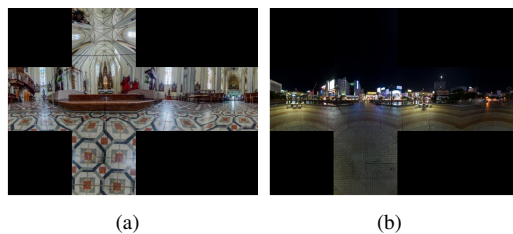

(a)

(b)

Figure 6: Images with the lowest (a) and highest (b) SI variability across the cube faces, in the set of images considered in Fig.5.

the SI is computed on the mosaicked CM planar image, vertical and horizontal edges corresponding to discontinuities at the borders between cube faces are taken into account, even if they are not features of the 360-degree image. To avoid the problem of handling the warping distortions and the discontinuities, one could consider the CM representation and compute the SI on each cube face separately, considering the mean value across all faces as a measure of the spatial complexity of the entire 360-degree image. Since each cube face can be considered as a perspective image, the SI computed on each cube face can be considered as a meaningful indicator of its spatial complexity.

Table 1 reports the Spearman's rank-order correlation (SROCC) coefficients between the SI values computed on eighty-five 360-degree images, belonging to the dataset described at [19], both in ER and in CM format, when the overall SI value is the mean across all cube faces (CM-faces) or it is computed on the mosaicked images (CM-mosaic). There are cases in which the rank of the same content differs depending on the domain used. Figure 4 shows, as an example, a pair of images that obtain different ranks in different domains.

As illustrated in these examples, *the geometrical domain of the 360-degree image does matter for its characterization*. In order to remove such dependency, a suggested approach could be *to compute the same index on the entire set of images in multiple domains and select a subset of images that have the same rank order independently from the domain*. If the set of images is too small to perform a meaningful subsampling, we suggest that the domain that implies less warping distortions is chosen, i.e., CM-faces, to characterize a visual signal that is close to what the user will visualize. Nevertheless, a single spatial complexity value might not

be very informative to characterize the entire 360-image. Figure 5 shows, for each of the eighty-five images in the set, the boxplot of the SI value across all cube faces. It can be noticed that, in some images, there is a significant variability in terms of spatial complexity across cube faces, i.e., depending on the viewing direction, the portion of 360-degree image attended by the user can have very different spatial complexity. Figure 6 shows the images in the set that have the lowest and highest variance of SI across cube faces. To account for this variability, which is an intrinsic feature of the 360-degree content, we suggest *to take into account the variance of the measure over the cube faces* when characterizing the content. To select images having different characteristics, a suggested approach could be *to select images that have different variability across the cube faces*.

Can 360-degree content be characterized based on users navigation and visual attention patterns?

As reflected by the previous examples, the characterization of 360-degree content is influenced by the region of the omnidirectional image that is considered. Furthermore, the freedom of the observers to explore 360-degree content entails that some regions of the images may not even be seen by the observers. On the contrary, other regions may be particularly important and attract users' attention. Thus, *proxies for content complexity could be inferred by analyzing users' visual attention patterns*.

Some efforts have been already made to characterize 2D content in terms of visual attention complexity. For instance, Le Meur *et al.* proposed the Inter-Observer Congruency (IOC), as a measure of the observers' agreement when exploring 2D images to be applied for image ranking [22]. Results showed that the IOC
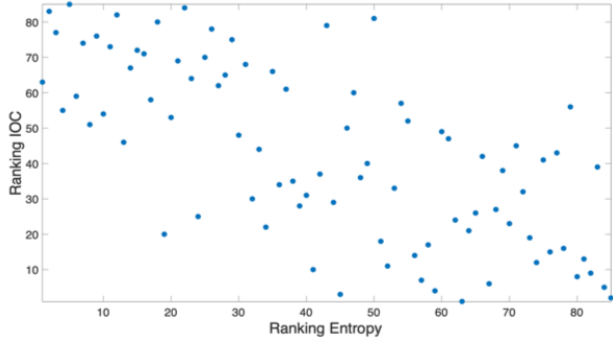
Figure 7: Ranking of the images obtained from entropy (horizontal) vs. ranking from IOC (vertical). Each point represents one image. Both measures were computed on the "head+eye" data.
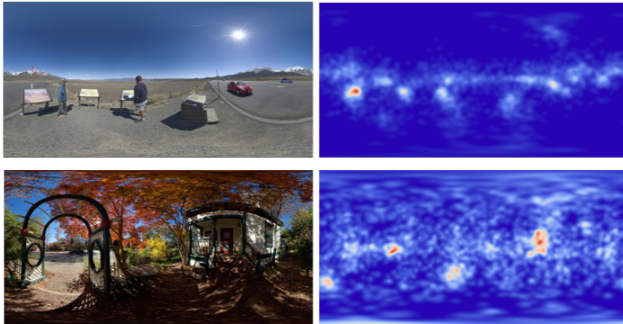

Figure 8: Examples of a "focused image" (top) and "exploratory image" (bottom), according to both IOC and entropy of the saliency maps, together with their corresponding saliency maps (from "head+eye" data).

is small when there is nothing in the scene that clearly catches the observers' attention ("exploratory images"), while IOC increases when some objects stand out ("focused images"). To compute this metric, the fixation data of each observer (i.e., ground truth) is required. Since this is not always available, other metrics have been proposed, such as using the entropy of the saliency maps (combined information from all observers) as a measure of saliency dispersion [23]. In this case, when the entropy is low, the saliency is concentrated in certain parts of the content ("focused images"). On the contrary, when the entropy is high, saliency is diffused representing "exploratory images". Although these two measures normally present an acceptable correlation, they can show some limitations and inconsistencies in certain cases [24].

To check the validity of these measures with 360-degree content, we computed them on the images of the considered dataset [19], using the ground-truth "head+eye" tracking data for the IOC, and the "head+eye" saliency maps for the entropy[3]. We applied a uniform spherical sampling on the ER maps [25] to avoid the drawbacks of the ER representation mentioned in Section 2. The scatter plot of the rankings of the images according to both metrics are shown in Figure 7. As aforementioned, their performance on representing the complexity in terms of visual attention is correlated ($SROCC = -0.7$, between the IOC and en-

---

[3]As described in [19], the dataset contains saliency maps and scanpaths obtained from eye fixations (referred as "head+eye", given the joint movement of head and eyes), and also saliency maps and scanpaths obtained from only head movements, by sampling head-tracking data (referred as "head-only").

tropy of the saliency maps for the images in the dataset, which is equivalent to the Pearson Linear Correlation Coefficient of the rankings showed in Figure 7), although some inconsistencies can be also observed. As an example, in Figure 8, we can observe a "focused image" and its saliency map in the upper part, which was ranked the 2nd and 3rd by entropy and IOC, respectively; while in the lower part, the most "exploratory image" according to entropy and 2nd according to IOC is shown.

Although some studies are considering the use of eye-tracking devices integrated in the HMDs (as in the considered dataset [19]), generally only head-tracking information is obtained from exploration of 360-degree contents directly from the HMDs, also allowing the analysis of viewers' behaviors [18][26]. Thus, we have done a similar analysis with IOC and entropy of the saliency maps, but considering only the head-tracking data. The resulting image ranking with IOC is highly correlated with the one obtained with IOC from "head+eye" tracking data ($SROCC = 0.91$), which reflects that using IOC over head-tracking data may be a good proxy for visual attention complexity. Also, entropy of the head-saliency maps can be an acceptable measure of exploration complexity, given the correlation ($SROCC = -0.68$) with IOC from 'head+eye' data (which can be considered the baseline measure for attention complexity). However, the results show that entropy of the saliency maps is less stable than IOC when only head-tracking data is used.

A more practical scenario has been also considered checking *whether saliency predictive models can be used for characterization of 360-degree content* when the tracking data is not available (i.e., previous exploration tests are not required). So, we computed the entropy on the predicted saliency maps from the images of the dataset using the best models (with online available code) resulting from the Salient360! Benchmark [27]: 1) the model from Startsev *et al.* [28], which predicts saliency considering "head+eye" movements, and 2) the model from Lebreton *et al.* [29], which predicts "head-only" saliency. The results showed a poor ranking correlation ($SROCC < 0.5$) between entropy on predicted and ground-truth saliency maps, and also with respect to IOC, which may be due to the still improvable performance of the predictive models. Nevertheless, it is worth noting that the considered models could differentiate between clearly "focused/exploratory" images, and also that "head+eye" saliency predictions are more sensitive to entropy measures, providing a wider range of entropy (from 4.9 to 7.8) for the same dataset than "head-only" models (from 4.7 to 6.9).

Another important aspect, where both ground-truth tracking data (even only from head movements) and predictive saliency models are useful, is to represent the distribution of the salient regions on the omnidirectional content that may attract observers' attention. Actually, works have been already published analyzing this aspect with ground-truth tracking data [15][17][30][31]. Thus, we explored if *saliency models can provide insights on this, which may help on the characterization of the content without the need of gathering viewing patterns in previous exploration tests*. Thus, the model from Lebreton *et al.* [29] (which predicts head-based saliency) was used on the considered dataset. The results showed the ability of the model to predict the saliency distribution in terms of longitude and latitude, as depicted with the two examples shown in Figure 9. There, it can be seen that the model predicts the bias of the main salient objects towards the bottom of

the image, as well as it predicts that the main salient objects in the second image are not in the longitudinal center. These observations are not only useful for appropriate content characterization and selection, but also to analyze results from subjective quality tests and objective metrics according to the location of salient objects in the 360-degree content.

Conclusion

This paper provided insights on how the geometrical domain of the signal and the interactive navigation by users must be considered when characterizing 360-degree content. In fact, the same measure computed in different domains can lead to different content characterization. Moreover, given the dependency of complexity measures on viewing direction, one single value for the entire 360-degree content might be not enough. Thus, this leads to also take into account complexity measures based on visual attention and navigation patterns, and to consider the saliency distribution over the 360-degree content. Ongoing work is oriented towards extending the analysis presented in this paper work, including 360-degree videos.

Acknowledgments

References

[1] S. Winkler. Analysis of public image and video databases for quality assessment. *IEEE Journal of Selected Topics in Signal Processing*, 6(6):616–625, Oct 2012.

[2] M. H. Pinson, M. Barkowsky, and P. Le Callet. Selecting scenes for 2D and 3D subjective video quality tests. *EURASIP Journal on Image and Video Processing*, (1):50–61, August 2013.

[3] ITU-T Recommendation P.910. Subjective video quality assessment methods for multimedia applications, April 2008.

[4] D. Hasler and S. E. Susstrunk. Measuring colorfulness in natural images. In *Electronic Imaging*, June 2003.

[5] S. Winkler. Efficient measurement of stereoscopic 3D video content issues. In *Electronic Imaging*, January 2014.

[6] P. Lebreton, A. Raake, M. Barkowsky, and P. Le Callet. Measuring perceived depth in natural images and study of its relation with monocular and binocular depth cues. In *SPIE Electronic Imaging*, page 90110C, March 2014.

[7] M. Narwaria, C. Mantel, M. Perreira Da Silva, P. Le Callet, and S. Forchhammer. An objective method for High Dynamic Range source content selection. In *International Workshop on Quality of Multimedia Experience*, pages 13–18, September 2014.

[8] J. Lee, T. Vigier, P. Le Callet, and J-S. Lee. A Perception-Based Framework for Wide Color Gamut Content Selection. In *IEEE International Conference on Image Processing*, pages 709–713, October 2018.

[9] P. Paudyal, J. Gutiérrez, P. Le Callet, M. Carli, and F. Battisti. Characterization and selection of light field content for perceptual assessment. In *Int. Conference on Quality of Multimedia Experience*, Erfurt, Germany, June 2017.

[10] F. Battisti, M. Carli, and P. Le Callet. A study on the impact of visualization techniques on Light Field perception. In *European Signal Processing Conference*, pages 2155–2159, Rome, Italy, September 2018.

[11] F. De Simone, P. Frossard, C. Brown, N. Birkbeck, and B. Adsumilli. Omnidirectional video communications: new challenges for the quality assessment community. *VQEG eLetter*, 3(1):33–38, November 2018.

[12] S. Afzal, J. Chen, and K. Ramakrishnan. Characterization of 360-degree Videos. In *Workshop on Virtual Reality and Augmented Reality Network*, pages 1–6, New York, New York, USA, August 2017.

[13] B. Micusik. *Two view geometry of omnidirectional cameras.* PhD thesis, Czech Technical University in Prague, 2004.

[14] Z. Chen, Y. Li, and Y. Zhang. Recent advances in omnidirectional video coding for virtual reality: Projection and evaluation. *Signal Processing*, 146:66 – 78, 2018.

[15] X. Corbillon, F. De Simone, and G. Simon. 360-Degree Video Head Movement Dataset. In *ACM on Multimedia Systems Con.*, pages 199–204, Taipei, Taiwan, June 2017.

[16] F. Qian, B. Han, Q. Xiao, and V. Gopalakrishnan. Flare: Practical viewport-adaptive 360-degree video streaming for mobile devices. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, MobiCom '18, pages 99–114, New York, NY, USA, 2018. ACM.

[17] C. Ozcinar and A. Smolic. Visual Attention in Omnidirectional Video for Virtual Reality Applications. In *International Conference on Quality of Multimedia Experience*, pages 1–6, May 2018.

[18] S. Rossi, F. De Simone, P. Frossard, and L. Toni. Spherical clustering of users navigating 360°content. In *International Conference on Acoustics, Speech, and Signal Processing*, Brighton, United Kingdom, May 2019.

[19] Y. Rai, J. Gutiérrez, and P. Le Callet. A dataset of head and eye movements for 360 degree images. In *ACM Multimedia Systems Conference*, Taipei, Taiwan, June 2017.

[20] F. Pearson. *Map Projections: Theory and Applications.* CRC Press, 1990.

[21] ANSI T1.801.03. Digital transport of one-way video signals parameters for objective performance assessment. American National Standards Institute, New York, NY, 1996.

[22] O. Le Meur, T. Baccino, and A. Roumy. Prediction of the inter-observer visual congruency (IOVC) and application to image ranking. In *ACM international conference on Multimedia*, New York, USA, November 2011.

[23] W. Zhang, R. R. Martin, and H. Liu. A Saliency Dispersion Measure for Improving Saliency-Based Image Quality Metrics. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(6):1462–1466, 2018.

[24] A. Volokitin, M. Gygli, and X. Boix. Predicting When Saliency Maps are Accurate and Eye Fixations Consistent. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 544–552, December 2016.

[25] M. Yu, H. Lakshman, and B. Girod. A framework to evaluate omnidirectional video coding schemes. In *2015 IEEE International Symposium on Mixed and Augmented Reality*, pages 31–36, Sep. 2015.

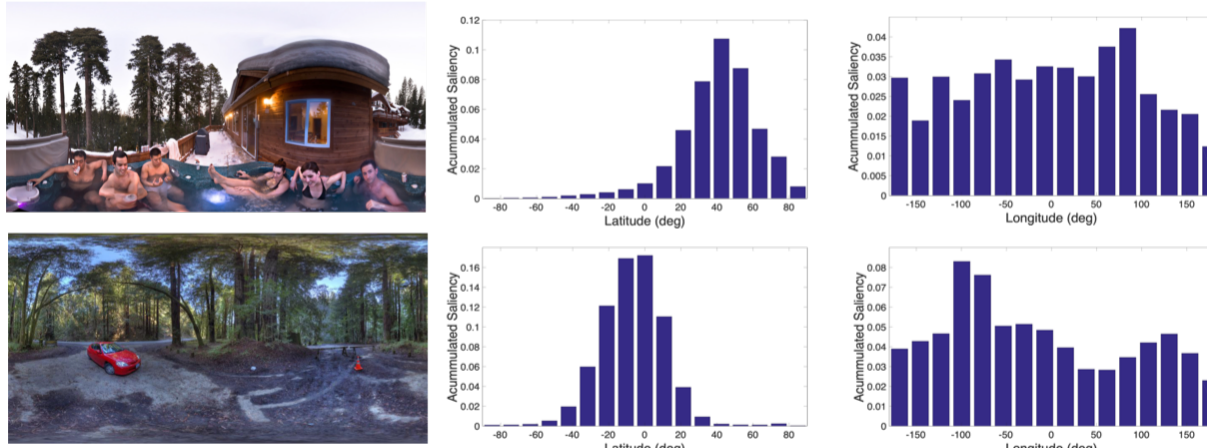[26] I. D. D. Curcio. On Streaming Services for Omnidirec-

Figure 9: Examples of saliency distributions for two images of the dataset, obtained from the saliency maps obtained with the model from Lebreton *et al.* [29]. Latitude: -90°- north pole, 0°- equator, +90°- south pole; Longitude: 0°center/front; +-180° back.

tional Video and its Subjective Assessment. *VQEG eLetter*, 3(1):33–38, November 2018.

[27] J. Gutiérrez, E. J. David, A. Coutrot, M. Perreira Da Silva, and P. Le Callet. Introducing UN Salient360! Benchmark: A platform for evaluating visual attention models for 360°contents. In *International Conference on Quality of Multimedia Experience*, Sardinia, Italy, May 2018. Website: https://salient360.ls2n.fr/.

[28] M. Startsev and M. Dorr. 360-aware saliency estimation with conventional image saliency predictors. *Signal Processing: Image Communication*, April 2018.

[29] P. Lebreton and A. Raake. GBVS360, BMS360, ProSal: Extending existing saliency prediction models from 2D to omnidirectional images. *Signal Processing: Image Communication*, March 2018.

[30] S. Fremerey, K. Meseberg, and A. Raake. AVTrack360: An open Dataset and Software recording people's Head Rotations watching 360°Videos on an HMD. In *ACM Multimedia Systems Conf.*, Amsterdam, Netherlands, June 2018.

[31] V. Sitzmann, A. Serrano, A. Pavel, M. Agrawala, D. Gutierrez, B. Masia, and G. Wetzstein. Saliency in VR: How do people explore virtual environments? *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1633 – 1642, April 2018.

Author Biography

Francesca De Simone *received the M. Sc. degree in electronics engineering from Universit degli Studi Roma Tre, Italy, in 2006, and the Ph.D. degree in computer and information science from the Swiss Federal Institute of Technology (EPFL), Switzerland in 2012. Between 2012 and 2014, she was post-doctoral fellow at the Multimedia Signal Processing Group of Institut Mines Telecom ParisTech, France. In 2015, she worked as senior engineer at the cybersecurity department of Kudelski Security, Switzerland. Between 2015 and 2017, she worked as post-doctoral fellow at the Signal Processing Laboratory (LTS4) at EPFL. Since February 2017, she is tenure track researcher at Centrum Wiskunde & Informatica, in Amsterdam, the Netherlands. Her research interests include subjective and objective multimedia quality assessment, image and video compression, and multimedia streaming strategies.*

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.*

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.*