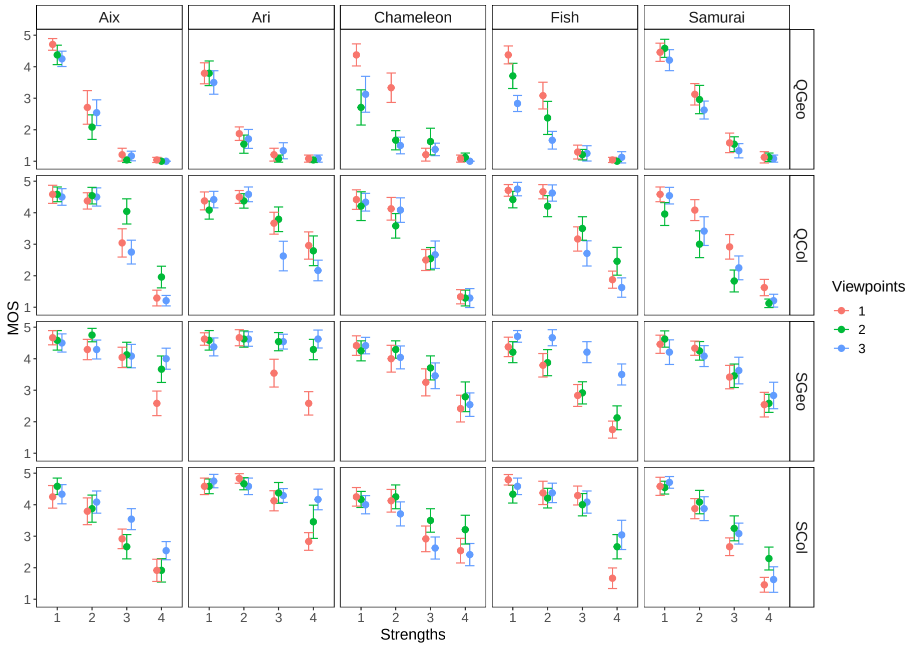
— Supplementary Material —
Visual Quality of 3D Meshes with Diffuse Colors in Virtual Reality: Subjective and Objective Evaluation

Yana Nehmé   Florent Dupont   Jean-Philippe Farrugia
Patrick Le Callet   Guillaume Lavoué

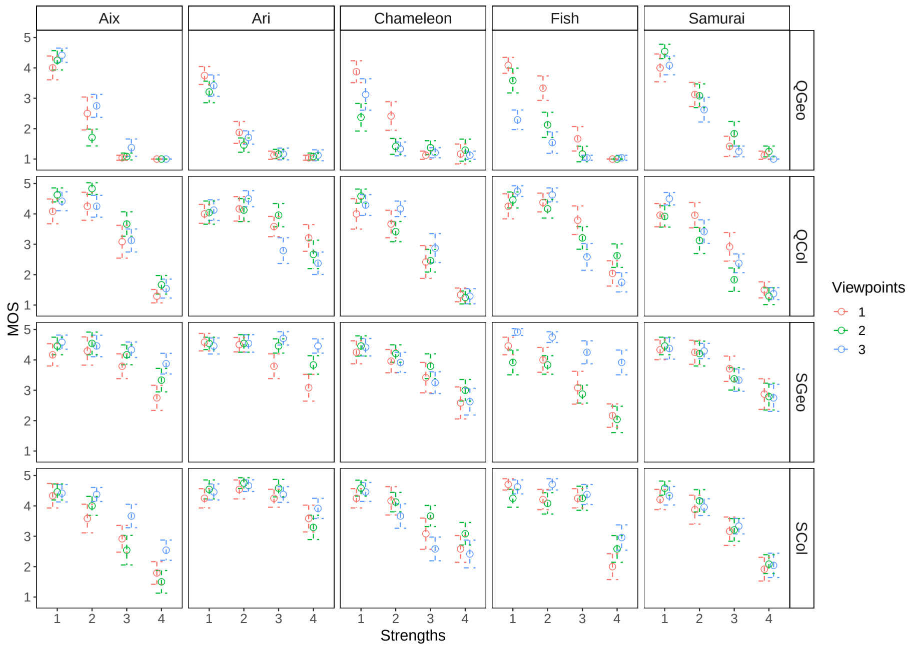
This supplementary material is organized as follows. Section 1 shows the mean opinion scores (MOSs) of all the stimuli of our database, associated with their confidence intervals (CIs), for the 2 animations separately. Section 2 presents the content ambiguity analysis obtained by the MLE model. In section 3, we justify the choice of the scales used in our objective metric *CMDM*. Section 4 evaluates the prediction performance of each feature implemented in CMDM after removing certain distortions. Section 5 provides the parameters of the tested images quality metrics, as well as snapshots of the camera positions. We illustrate, in section 6, the subjective scores with respect to objective metric values for the textured 3D meshes dataset. Finally, section 7 assesses the performance of the the tested metrics with and without integrating the viewpoint (the visible parts) of 3D models.

1 Resulting MOSs

In this section, we present the MOSs and CIs acquired for our ground truth database of 480 animated 3D graphics. Figure 1.a shows the results of the stimuli in rotation, while Figure 1.b shows the results of those in zoom.



(a) Stimuli animated with a slow rotation (R).



(b) Stimuli animated with a slow Zoom (Z).

Figure 1: Overview of mean opinion scores of the stimuli, associated with their confidence intervals. For a given distortion strength, the dots are horizontally spaced apart to avoid overlapping.

2 Content ambiguity

We analyzed the ambiguity of our source contents (a_c), obtained by the MLE model, for each viewpoint and animation (i.e. for each of our 6 HRTs: combinations of 3 viewpoints and 2 animations). Figure 2 shows that the source models animated with a zoom movement and displayed in viewpoint *1*, are associated with the highest content ambiguity. We recall that viewpoint *1* is the viewpoint that covers most of the shape and carries the most information on color and geometry.

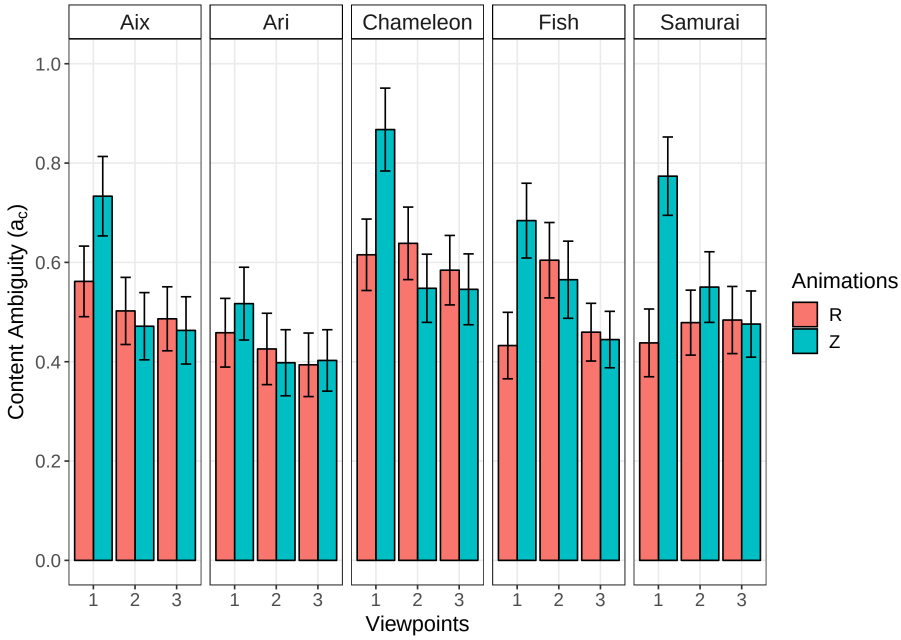
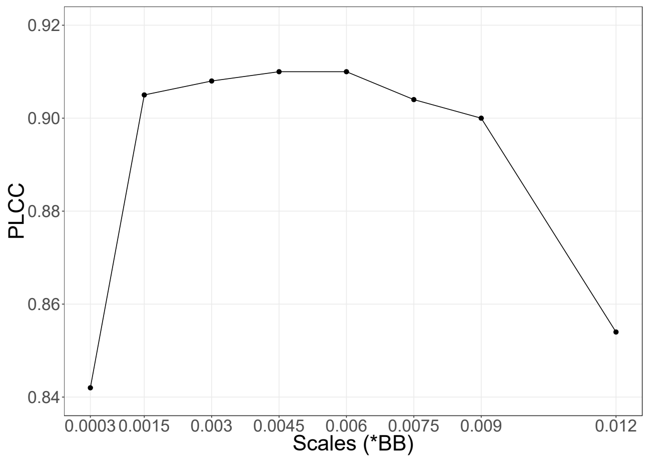


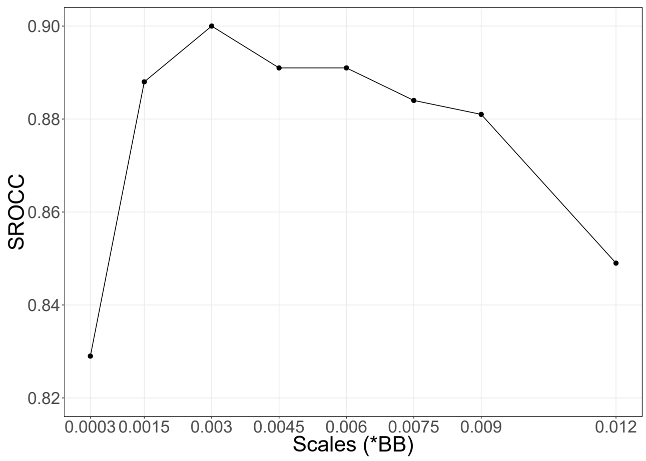
Figure 2: Content ambiguity a_c of each source model associated with HRTs, and its confidence interval.

3 Scales

As stated in the paper, we developed a full-reference multiscale metric ($CMDM$) for predicting the quality of colored meshes. We relied on Figures 3 to select the most appropriate /relevant scales. We recall that the scale (h_i) defines the radius of the spherical neighborhood around each vertex v of the distorted mesh. For each scale h_i, we compute geometry and color based features over the local corresponding neighborhood of v.



(a)



(b)

Figure 3: Variation of Pearson (a) and Spearman (b) correlations between $CMDM$ and subjective scores, with respect to several scale values. BB is the maximum bounding box length of the stimulus.

As can be seen, CMDM provides the best performances, in terms of PLCC and SROCC correlations, for the following 3 scales: $h_i \in \{0.003BB, 0.0045BB, 0.006BB\}$, where BB is the maximum bounding box length of the stimulus.

4 Single feature prediction performance

This section evaluates the prediction performance of each feature implemented in our multiscale metric. In this analysis, we did not consider stimuli distorted with color quantization (QCol) to assess the performance of geometry-based features (f_1, f_2, and f_3), since this type of distortion is applied only on the vertex colors and does not affect the model geometry at all. The correlations of these individual features with the recovered MOSs, as well as their classification abilities are reported in Table 1.
Similarly, for the color-based features (f_4, f_5, f_6, f_7, and f_8), stimuli geometrically quantized (QGeo) were not taken into account. Indeed, this distortion superimposes the vertices of the stimulus, meaning that we cannot know or control exactly which vertex color is taken into account in Unity's import and render pipelines. Results are reported in Table 2.

Table 1: Performance of geometry-based features.

Feature	**Id**	$\boldsymbol{PLCC}$	$\boldsymbol{SROCC}$	$\boldsymbol{AUC_{DS}}$	$\boldsymbol{AUC_{BW}}$
Curvature comparison	f_1	0.78	0.752	0.647	0.902
Curvature contrast	f_2	0.723	0.736	0.616	0.882
Curvature structure	f_3	0.502	0.558	0.53	0.789

Table 2: Performance of color-based features.

Feature	**Id**	$\boldsymbol{PLCC}$	$\boldsymbol{SROCC}$	$\boldsymbol{AUC_{DS}}$	$\boldsymbol{AUC_{BW}}$
Lightness comparison	f_4	0.644	0.785	0.713	0.878
Lightness contrast	f_5	0.743	0.796	0.729	0.911
Lightness structure	f_6	0.632	0.761	0.7	0.888
Chroma comparison	f_7	0.494	0.707	0.678	0.842
Hue comparison	f_8	0.457	0.533	0.627	0.761

Removing these distortions improves feature performances, especially the geometry-based features.

5 Settings for image quality metrics

We compared our metric with 3 state-of-the-art full-reference image quality metrics (IQMs): *SSIM*, *HDR-VDP2*, *iCID*. For *SSIM*, we considered a local window of size 11 × 11 pixels. For the resolution used for *HDR-VDP2*, we considered 33.7 pixels per degree, which corresponds to the following experimental setting: stimuli presented on a calibrated 23" LCD display (resolution 1920 × 1080 pixel) at a constant viewing distance of 0.5m. The peak sensitivity parameter of *HDR-VDP2* was set to 2.4 and the selected output from this metric was the quality prediction Q. For the *iCID* metric, we considered the default parameters:equal weight of lightness, chroma, and hue, and use of chroma contrast and chroma structure.

To apply these IQMs, we generate for each 3D object in our database, a set of 18 snapshots taken from different viewpoints. To do so, the camera was placed at regularly sampled positions around the stimulus, as shown in Figure 4.

Figure 4: Camera positions regularly sampled around the 3D object.

6 Validation on a dataset of textured 3D meshes

To evaluate the robustness of our metric and to verify that it did not just learn the distortions that are specific to our dataset, we tested $CMDM$ on the LIRIS Textured Mesh Database [1]. We included results of the IQMs presented previously, as well as the results obtained by Guo et al. [1] for different metrics, applied either on rendered videos of the stimuli (DCT, $MS\!-\!SSIM$ and $PSNR$) or directly on textured meshes (FQM, CM_1 and CM_2). Figure 5 illustrates the subjective scores with respect to the values of these metrics. Note that, Figures 5.e, 5.f, 5.g, and 5.h are reprinted from [1].

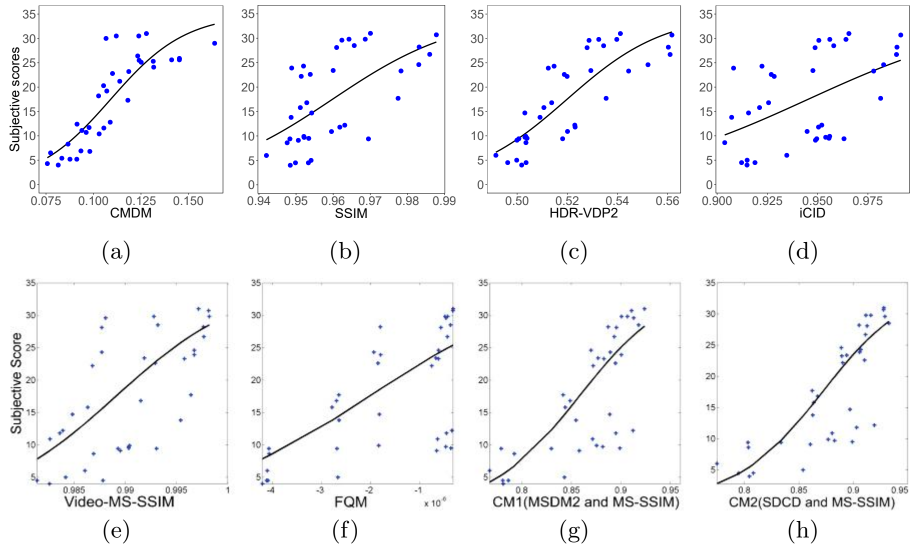


Figure 5: Scatter plot of subjective scores versus objective metric values for the LIRIS Textured Mesh Database. Each point represents one stimulus. The fitted logistic function is displayed in black.

[1] J. Guo, V. Vidal, I. Cheng, A. Basu, A. Baskurt, and G. Lavoue, “Subjective and objective visual quality assessment of textured 3D meshes,” ACM Transactions on Applied Perception, vol. 14, no. 2, 2016.

7 Integration of the viewpoint

In section 7 of the paper, we studied the relevance of incorporating the viewpoint (the visible parts) of the 3D model into objective metrics. Thus, we tested the performance of our metric and that of IQMs according to 2 scenarios, using a subset of 240 stimuli from our database:

- (1) Without integrating the visibility: we computed the IQMs on multiple snapshots (18) taken from different viewpoints of the object. *CMDM* was computed over all the vertices of the stimuli. The results are reported in Table 3.
- (2) With integrating the visibility: IQMs were directly computed on the snapshot taken from the real viewpoint displayed to the observer (IQM_{vis}) and *CMDM* was computed only over the visible vertices ($CMDM_{vis}$). The results are reported in Table 4.

Table 3: Performance comparison of different metrics Without integrating the viewpoint.

	$\boldsymbol{PLCC}$	$\boldsymbol{SROCC}$	$\boldsymbol{AUC_{DS}}$	$\boldsymbol{AUC_{BW}}$
CMDM	0.886	0.871	0.756	0.967
SSIM	0.773	0.768	0.697	0.915
HDR-VDP2	0.827	0.808	0.714	0.942
iCID	0.8	0.8	0.727	0.927

Table 4: Performance comparison of different metrics when integrating the viewpoint.

	$\boldsymbol{PLCC}$	$\boldsymbol{SROCC}$	$\boldsymbol{AUC_{DS}}$	$\boldsymbol{AUC_{BW}}$
$\mathrm{CMDM_{vis}}$	0.886	0.866	0.755	0.967
$\mathrm{SSIM_{vis}}$	0.791	0.798	0.722	0.927
$\mathrm{HDR\text{-}VDP2_{vis}}$	0.805	0.826	0.661	0.943
$\mathrm{iCID_{vis}}$	0.857	0.871	0.776	0.957