# Assessing Visual Defect Saliency on 3D Meshes Through Gaze-Based Metrics

1st Sonain Jamil
*LIUM*
*Le Mans Université*
F-72000 Le Mans, France
sonainjamil@ieee.org

2nd Julien Pettré
*Univ Rennes,*
*Inria, CNRS, IRISA*
Rennes, France
julien.pettre@inria.fr

3rd Sébastien Iksal
*LIUM*
*Le Mans Université*
F-72000 Le Mans, France
sebastien.iksal@univ-lemans.fr

4th Erwan David
*LIUM*
*Le Mans Université*
F-72000 Le Mans, France
erwan.david@univ-lemans.fr

***Abstract*—Visual artefacts in 3D meshes can act as local distractors and perturb visual attention, which may affect how users inspect and interpret interactive 3D content in applications such as technology-enhanced learning and training. An important open question is how localized geometric distortions translate into measurable perceptual disruption during interactive viewing, as understanding this relationship is necessary for developing perceptually grounded evaluation methods for 3D assets. In this work, we investigate whether gaze can serve as an implicit, spatially localized indicator of such effects. To this end, we introduce a surface-based gaze analysis approach that maps eye-tracking samples onto the visible 3D mesh geometry under user-driven interaction, enabling view-consistent analysis of localized attention. This enables defect-level, interaction-aware attention analysis that is not supported by conventional image-plane or fixed-viewpoint methods. In a study with 22 participants viewing 27 textureless meshes with four defect types, we observe that self-intersection defects produce significantly higher fixation counts, while time-to-first fixation (TFF) does not show significant differences compared to corresponding regions in unmodified meshes. In contrast, smoothing defects shows weaker trends, while low-polygon and semantic defects do not produce reliable localized effects under the present conditions. These results suggest that visual attention reflects the perceptual severity of localized mesh distortions, positioning gaze as a complementary signal for perceptually grounded quality-of-experience assessment in 3D content. The findings are conditioned by the controlled viewing setup and defect placement, and should be interpreted accordingly.**

***Index Terms*—3D mesh quality, visual defect saliency, gaze analysis, perceptual quality assessment, eye tracking, quality of experience**

## I. Introduction

Interactive 3D content is increasingly used in technology-enhanced learning, training, and simulation to support visualization and exploration [1], [2]. In such settings, unintended visual artefacts may act as local distractors, drawing attention away from relevant content and potentially increasing cognitive load [3], [4]. Understanding how such distortions affect visual attention is therefore important for the perceptual evaluation of 3D assets.

Quality assessment of 3D meshes is still largely based on explicit measures such as subjective ratings or global similarity metrics [5]–[7]. While these approaches provide useful overall quality estimates, they do not capture how localized defects influence perceptual processing during interaction. A key open question is how localized geometric distortions translate into perceptual disruption under interactive viewing conditions.

Gaze offers a complementary, implicit signal, as visual attention is known to be sensitive to structural irregularities and perceptual inconsistencies [8]–[10]. Unlike explicit ratings, gaze reflects moment-to-moment attentional responses without requiring user reporting, making it particularly suitable for capturing localized perceptual effects. Prior work has shown that mesh distortions and saliency cues can influence gaze [11]–[14], but localized defect-level analysis under interactive viewing remains limited.

A key challenge is that, during interactive viewing, users freely rotate and zoom the object, making it insufficient to analyze gaze in the 2D image plane. Traditional approaches that rely on 2D projections or fixed viewpoints cannot reliably capture how attention relates to the underlying 3D geometry under interaction [15], [16]. Moreover, existing experimental studies have not systematically analyzed how localized geometric defects influence visual attention under such conditions.

More recent work has focused on methodological frameworks for processing gaze in 3D environments [17], [18]. These approaches enable surface-based gaze analysis by mapping eye-tracking data onto mesh geometry. For example, [18] provides a toolbox that projects gaze onto surface samples and addresses mesh-dependent biases through adaptive resampling strategies. Related representations, such as mesh-color techniques [19], associate data directly with surface elements, enabling a one-to-one correspondence between geometry and encoded information.

Building on these surface-based approaches, we adopt a simplified method that maps gaze samples directly onto the original mesh and assigns them to surface-based areas of interest (AOIs) using vertex-level encoding. This choice reflects a design trade-off for the present study, while more advanced surface representations such as those proposed in [18] could be incorporated in future work.

We apply this approach in an eye-tracking study with 22 participants viewing 27 meshes containing four defect types: self-intersection, smoothing, low-polygon simplification, and semantic part removal. We compare fixation behaviour within defect AOIs against corresponding regions in matched unmodified meshes. Our results show that self-intersection defects produce significantly stronger localized attentional capture than the other defect types, while smoothing shows weaker trends and low-polygon and semantic defects do not produce reliable effects under the present conditions.

This work contributes (1) a surface-based framework for linking gaze behaviour to 3D geometry under interaction, enabling view-consistent analysis of localized perceptual effects, and (2) empirical evidence that attentional responses vary with the perceptual severity of mesh defects. Rather than proposing a complete QoE metric, this work isolates visual attention as a measurable component of perceptual response that can complement explicit and computational quality models.

More broadly, this work takes a step toward bridging visual attention analysis and quality-of-experience assessment in interactive 3D content. However, results should be interpreted in the context of the controlled, textureless, and screen-based experimental setup used here.

## II. Method

We conducted a controlled laboratory eye-tracking study to investigate how localized geometric defects translate into perceptual disruption during interactive 3D viewing. The method combines sample-wise gaze-to-surface projection with AOI-based fixation analysis, enabling defect-level assessment of visual attention directly on mesh geometry.

### A. Participants and Ethics

Twenty-two participants took part in the study. All reported normal or corrected-to-normal vision. Informed consent was obtained prior to participation, and participants were compensated 7 EUR for their time. The study was conducted in accordance with institutional ethical guidelines. Although eye-tracking data may be considered personally identifiable under current regulations, all data were anonymized and handled in compliance with applicable data protection standards.

### B. Stimuli

We used 27 textureless 3D meshes from [11]. For each mesh, four localized defect variants (self-intersection, smoothing, low-polygon, semantic) were generated, yielding 135 stimuli (27 unmodified, 108 distorted).

Defects were manually introduced in Blender to ensure visibility while preserving object identity. AOIs were defined on corresponding regions of matched unmodified meshes for paired comparison; for topology-changing defects, AOIs reflect intended regions rather than exact correspondence. Because defect placement was not systematically controlled, differences in regional salience may influence attention, and results should be interpreted as relative comparisons within matched mesh pairs.

### C. Apparatus

Gaze was recorded with a Tobii Pro Fusion eye tracker. Stimuli were displayed on a 2460×1440 monitor at 60 Hz. Participants sat at approximately 60–65 cm from the screen and used a chin rest to reduce head motion and improve tracking stability.

### D. Experimental Design and Procedure

Each session began with a five-point calibration, and each trial started with a central fixation cross to standardize initial gaze position. Meshes were presented in Unity for 10 s, allowing participants to freely rotate and zoom the objects, with a mid-session break and recalibration included to reduce fatigue and drift. Each participant viewed 27 meshes, each presented once in either an unmodified or distorted condition across the four defect categories. Trial order was randomized, and all meshes used predefined initial orientations to ensure initial defect visibility. Because interaction changed defect visibility, AOI hits were counted only when the gaze ray intersected the defect region on the camera-visible mesh surface.

### E. Surface-based Gaze Projection

At each sample, gaze was represented in normalized display coordinates $g = (g_x, g_y) \in [0, 1]^2$. When both eyes were valid, the gaze position was defined as their average; otherwise the valid eye was used.

To associate gaze with the mesh surface, we reconstructed the camera pose and object transform for each sample and cast a ray $\mathbf{r}(g)$ from the camera through the gaze position $g$ into the scene. The first intersection between this ray and the mesh surface $\mathcal{M}$ defines the surface hit point $\mathbf{x}(g)$:

$$\mathbf{x}(g) = \mathcal{M} \cap \mathbf{r}(g), \tag{1}$$

where $\mathbf{r}(g)$ denotes the gaze ray, $\mathcal{M}$ the mesh surface, and $\mathbf{x}(g)$ the resulting intersection point.

This sample-wise projection keeps gaze assignment consistent under user-driven rotation and zoom, while restricting analysis to surface regions visible to the camera.

### F. AOI Definition

Defect regions were encoded on the mesh using the vertex color attribute, where vertices belonging to a defect AOI were labeled by setting their red channel to 1, while non-defect vertices remained at 0.

For each raycast hit, the intersected triangle was retrieved, and the gaze sample was assigned to the AOI if at least one vertex $\mathbf{v}$ of the triangle satisfied the condition:

$$\mathrm{AOI}(\mathbf{v}) = \begin{cases} 1, & \text{if } R(\mathbf{v}) > \tau_R, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

where $\tau_R = 0$ in our implementation.

This definition enables view-consistent labeling of gaze samples directly on the 3D surface, avoiding reliance on 2D screen-space regions or fixed viewpoints, and ensuring that AOIs remain anchored to the underlying geometry during interaction.

Because AOIs are defined at the vertex level, their spatial precision depends on mesh resolution: regions with higher vertex density allow more precise AOI boundaries, whereas coarser regions may lead to approximate localization. As discussed in the introduction, prior work [18] addresses this limitation through surface resampling strategies that ensure more uniform spatial coverage. While the 3D objects in our current dataset do not exhibit significant variations in vertex distribution, we intend to incorporate Plume's resampling technique in future experiments to further enhance robustness.

### G. Fixation Extraction and Measures

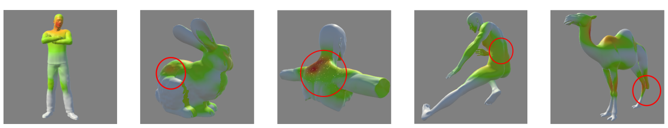
Fixations were extracted from the gaze stream using the Dispersion-Threshold Identification (IDT) algorithm [20]. A fixation was defined as a sequence of gaze samples whose spatial dispersion remained below $1.27^{\circ}$ of visual angle and whose duration exceeded $T_{\min} = 0.2\,\mathrm{s}$, consistent with commonly used thresholds in eye-tracking studies [10].

From the detected fixations, we computed two complementary AOI-based measures for each trial: (1) *AOI fixation count*, i.e., the number of fixations intersecting the AOI, and (2) *AOI time-to-first fixation (TFF)*, i.e., the latency between stimulus onset and the first fixation within the AOI. These measures capture distinct aspects of visual attention: fixation count reflects attentional allocation, whereas TFF reflects how quickly attention is directed toward a region.

For each defect condition, AOI-based metrics were compared with the corresponding AOI region defined on the matched unmodified version of the same base mesh, enabling paired statistical analysis of localized attentional effects.

## III. Results and Discussion

Surface-based gaze projection as shown in Fig. 1 enables spatially localized attention analysis directly on the mesh geometry.



(a) N (b) SI (c) SM (d) LP (e) SE

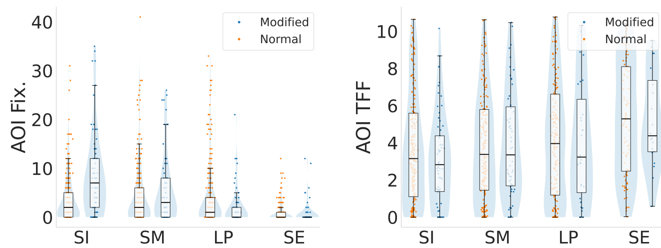
Fig. 1: Example of localized defects: (a) normal (N), (b) self-intersection (SI), (c) smoothing (SM), (d) low-polygon (LP), and (e) semantic (SE).

To quantify localized perceptual disruption, we compared AOI fixation count and AOI time-to-first fixation (TFF) between distorted meshes and corresponding AOI regions in unmodified meshes. For each defect type, analysis was performed within a defect-specific AOI and compared to the same spatial region on the matched unmodified mesh. Because defect locations differ across conditions, the corresponding AOIs on unmodified meshes also differ, resulting in distinct "normal" distributions for each defect type. Because distributions were non-normal, all comparisons used Wilcoxon signed-rank tests. Fig. 2 shows the distributions, and Table I summarizes medians and statistics.

The strongest and most consistent effects were observed for self-intersection defects. As shown in Fig. 2 and Table I, self-intersections yielded substantially higher fixation counts than the corresponding regions in unmodified meshes ($p < .001$). TFF values show no significant differences ($p = 0.250$), indicating that while these defects strongly attract attention, they are not necessarily fixated earlier than other regions. These results indicate that strong geometric inconsistencies act as salient local distractors, capturing visual attention.

TABLE I: Median AOI fixation count (Fix.) and AOI TFF for modified (Mod.) and normal (Norm.) meshes.

| Defect type | Fix. | | | TFF | | |
|---|---|---|---|---|---|---|
| | Mod. | Norm. | $p$ | Mod. | Norm. | $p$ |
| Self-intersection | 7 | 2.0 | $< .001$ | 2.83 | 3.15 | 0.250 |
| Smoothing | 3 | 2.0 | 0.110 | 3.35 | 3.38 | 0.724 |
| Low polygon | 0 | 1.0 | 0.006 | 3.23 | 3.96 | 0.978 |
| Semantic | 0 | 0.0 | 0.320 | 4.38 | 5.29 | 0.987 |




(a) Fix. (b) TFF

Fig. 2: AOI fixation count and AOI TFF for modified (Mod.) and normal (Norm.) meshes across defect types (SI: self-intersection, SM: smoothing, LP: low-polygon, SE: semantic).

Smoothing artefacts showed weaker, non-significant trends ($p = 0.11$), indicating limited attentional impact. Low-polygon distortions produced reduced attention ($p = 0.006$), while semantic defects showed no reliable effects ($p \geq 0.32$). As summarized in Table I, these conditions exhibit minimal or inverse differences. This pattern suggests that attentional responses are sensitive to geometric inconsistency rather than semantic disruption under the present conditions.

These results show that gaze-based AOI measures differentiate the *perceptual strength* of defect types, with strong geometric inconsistencies producing greater attentional capture. Despite limitations in defect visibility, placement, and semantic salience, self-intersections emerge as the most perceptually disruptive defect type.

## IV. Conclusion

This paper introduced a surface-based gaze analysis method that maps eye-tracking data onto visible 3D mesh geometry during interactive viewing. By assigning gaze samples to geometry-anchored AOIs, the method enables view-consistent, localized analysis beyond conventional image-plane approaches. It supports defect-level assessment of perceptual impact by quantifying how specific regions attract attention. Results show that different defect types produce distinct attentional responses. Rather than a complete QoE metric, this work presents gaze-based surface analysis as a complementary tool for perceptually grounded 3D content evaluation.

## REFERENCES


[1] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Computers & education*, vol. 147, p. 103778, 2020.

[2] M. Coban, Y. I. Bolat, and I. Goksu, "The potential of immersive virtual reality to enhance learning: A meta-analysis," *Educational Research Review*, vol. 36, p. 100452, 2022.

[3] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognitive science*, vol. 12, no. 2, pp. 257–285, 1988.

[4] G. Makransky and G. B. Petersen, "The cognitive affective model of immersive learning (camil): A theoretical research-based model of learning in immersive virtual reality," *Educational psychology review*, vol. 33, no. 3, pp. 937–958, 2021.

[5] G. Lavoué and R. Mantiuk, "Quality assessment in computer graphics," in *Visual Signal Quality Assessment: Quality of Experience (QoE)*. Springer, 2014, pp. 243–286.

[6] 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 (TAP)*, vol. 14, no. 2, pp. 1–20, 2016.

[7] Y. Nehmé, F. Dupont, J.-P. Farrugia, P. Le Callet, and G. Lavoué, "Visual quality of 3d meshes with diffuse colors in virtual reality: Subjective and objective evaluation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 3, pp. 2202–2219, 2020.

[8] A. L. Yarbus, *Eye movements and vision*. Springer, 2013.

[9] J. Wagemans, J. Feldman, S. Gepshtein, R. Kimchi, J. R. Pomerantz, P. A. Van der Helm, and C. Van Leeuwen, "A century of gestalt psychology in visual perception: Ii. conceptual and theoretical foundations." *Psychological bulletin*, vol. 138, no. 6, p. 1218, 2012.

[10] A. T. Duchowski, *Eye tracking methodology: Theory and practice*. Springer, 2017.

[11] G. Lavoué, F. Cordier, H. Seo, and M.-C. Larabi, "Visual attention for rendered 3d shapes," in *Computer Graphics Forum*, vol. 37, no. 2. Wiley Online Library, 2018, pp. 191–203.

[12] J. Guo, V. Vidal, A. Baskurt, and G. Lavoué, "Evaluating the local visibility of geometric artifacts," in *Proceedings of the acm siggraph symposium on applied perception*, 2015, pp. 91–98.

[13] Y. Wang, Q. Dai, M. Bâce, K. Klein, and A. Bulling, "Saliency3d: a 3d saliency dataset collected on screen," in *Proceedings of the 2024 Symposium on Eye Tracking Research and Applications*, 2024, pp. 1–6.

[14] K. Zhang, D. Zhu, X. Min, and G. Zhai, "Textured mesh saliency: Bridging geometry and texture for human perception in 3d graphics," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 9, 2025, pp. 9977–9984.

[15] M. Abid, M. P. Da Silva, and P. Le Callet, "Towards visual saliency computation on 3d graphical contents for interactive visualization," in *2020 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2020, pp. 3448–3452.

[16] M. Maurus, J. H. Hammer, and J. Beyerer, "Realistic heatmap visualization for interactive analysis of 3d gaze data," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, 2014, pp. 295–298.

[17] E. David, J. Gutiérrez, M. L.-H. Võ, A. Coutrot, M. P. Da Silva, and P. Le Callet, "The salient360! toolbox: Handling gaze data in 3d made easy," *Computers & Graphics*, vol. 119, p. 103890, 2024.

[18] C. Javerliat, S. Villenave, P. Raimbaud, and G. Lavoué, "Plume: Record, replay, analyze and share user behavior in 6dof xr experiences," *IEEE Transactions on Visualization and Computer Graphics*, vol. 30, no. 5, pp. 2087–2097, 2024.

[19] C. Yuksel, J. Keyser, and D. H. House, "Mesh colors," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 2, pp. 1–11, 2010.

[20] D. D. Salvucci and J. H. Goldberg, "Identifying fixations and saccades in eye-tracking protocols," in *Proceedings of the 2000 symposium on Eye tracking research & applications*, 2000, pp. 71–78.