Responses to Reviewers' Comments for Manuscript TOG-2026-0142

Perceptual Error Metrics for Real-Time Volumetric Rendering

Addressed Comments for Publication to

ACM Transactions on Graphics

by

S. Adeyemi and L. Marchetti

Dear Prof. Amara Osei,

Please find enclosed the revised version of our manuscript entitled “Perceptual Error Metrics for Real-Time Volumetric Rendering” with manuscript number TOG-2026-0142. We thank you and the two reviewers for their careful reading and constructive comments, which have substantially improved the paper. Below we give a summary of the main changes, followed by a point-by-point response to each reviewer’s comments, in the order they appeared in the decision letter.

Sincerely,

S. Adeyemi and L. Marchetti

Note: To improve legibility, each reviewer comment is set in a boxed block, our response follows immediately below it, and any new or rephrased manuscript text we quote is set in a shaded block. The corresponding passages have also been highlighted in the revised manuscript.

Response to the Meta Review

Summary Comment 1

Both reviewers see merit in the proposed error metric but raise concerns about the scale of the user study and the clarity of the comparison against existing perceptual metrics. Please address these points in a revision.

Response S1: We thank the editor for the constructive summary.

We have expanded the user study from 12 to 34 participants and added a direct quantitative comparison against SSIM and FLIP in Section 5. A summary of all changes follows in the reviewer responses below.

Response to Reviewer 1

Comment 1.1

The user study has only 12 participants, which is too small to support the claim that the metric correlates with human judgement across content types.

Response 1.1: Thank you for the comment.

We agree the original sample size was too small to support a general claim. We recruited an additional 22 participants (34 total) spanning three content categories (animation, ray-traced stills, and volumetric fog), and re-ran the correlation analysis. The revised Section 5.2 now reads:

> Across all 34 participants and three content categories, our metric achieves a Pearson correlation of $r = 0.87$ with mean opinion scores, compared to $r = 0.71$ for SSIM and $r = 0.79$ for FLIP on the same stimuli.

Comment 1.2

Figure 4 does not clearly show how the proposed metric differs from FLIP on the same inputs.

Response 1.2: Thank you for the comment.

We have replaced Figure 4 with a side-by-side comparison showing our metric's error map alongside FLIP's on all four test scenes, with a shared colour scale so the difference in sensitivity to temporal flicker is directly visible.

Response to Reviewer 2

Comment 2.1

The paper does not report the computational cost of the metric relative to SSIM and FLIP, which matters for the real-time use case the authors motivate in the introduction.

Response 2.1: Thank you for the comment.

This is a fair omission on our part. We added a new Table 3 reporting per-frame evaluation time at 1920 × 1080 on the hardware described in Section 4.1:

> Our metric evaluates in 1.8 ms per frame, compared to 0.6 ms for SSIM and 2.3 ms for FLIP, comfortably within budget for a 60 Hz real-time pipeline.

Comment 2.2

Minor: the related-work section conflates perceptual metrics for still images with those designed for video; this distinction matters for the paper's contribution.

Response 2.2: Thank you for the comment.

We have split Section 2 into two subsections, "Still-Image Perceptual Metrics" and "Temporal and Video Metrics", and moved the relevant citations accordingly.