Palliating Visual Artifacts Through Audio Rendering

Hui Ding and Christian Jacquemin

LIMSI–CNRS - University of Paris-Sud, France
{hui.ding,christian.jacquemin}@limsi.fr

Abstract. In this paper, we present a pipeline for combining graphical rendering through an impostor-based level of detail (LOD) technique with audio rendering of an environment sound at different LODs. Two experiments were designed to investigate how parameters used to control the impostors and an additional audio modality can impact the visual detection of artifacts produced by the impostor-based LOD rendering technique. Results show that in general, simple stereo sound hardly impact the perception of image artifacts such as graphical discontinuities.

1 Introduction

Cross-modal dependency deals with the role of sensory perceptions (e.g., vision, audio, etc.) and their combinations for generating and improving the multimodal perception of virtual environments. Major works on user perception for interactive applications have been done on spatialized sound rendering [2][5], auditory facilitation on visual modality [4][3], and audio-visual rendering driven by crossmodal perception[1]. These works have confirmed that a multi-modal virtual environment able to synchronize efficiently the audio and visual renderings can provide users with an enhanced visual perception of the virtual environment.

In this paper, we are also interested to evaluate to what extent the audio modality can compensate imperfections in the visual modality. Our purpose is to consider these issues from another point of view, i.e., to investigate whether the audio modality can impact the perception of visual artifacts due to the perspective distortion of a 2D image onto which are projected the 3D geometry.

2 Implementation of Audio-Graphical Rendering

In our graphical implementation, we design a real time impostor-based method at five LODs for representing a tree in the clipping volume (frustum). LOD1 is detailed geometric rendering. LOD2 to LOD5 are impostor-based rendering. The selection of LOD is controlled by distance as the selection factor. Once the tree reaches an area corresponding to a different value of LOD selection factor, we use a detailed rendering of the tree and we take snapshots of different sub-parts of rendered tree according to the selected LOD. The snapshots (impostors) will then be used to construct the whole tree during the following steps/frames as

long as the tree remains in the same value of LOD selection factor. For LOD2, the branches of terminal level of tree are rendered through impostors instead of detailed geometry; and for LOD3, the branches of the last two levels of tree are rendered through impostors (see Fig.1), and so on. As was expected, the impostors produce visual artifacts, mostly graphical discontinuities in our case. The perception of this kind of visual artifacts depends on human vision capacities.

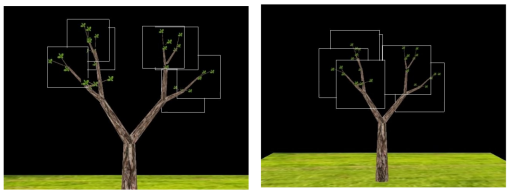

Fig. 1. Left snapshot is the LOD2; right snapshot is the LOD3.

In our case, only a recorded sound and its impact on the perceived image quality are considered as audio rendering to be combined with graphical rendering. So, sound is rendered in a non-realistic manner through stereo and level defined by the distance between 3D graphical object and the viewer. The 3D static graphical tree is accompanied by a continuous stereo sound simulation of a tree in the wind. The stereo sound volume changes according to the user's motion in the scene.

3 Experiment Design

The discontinuity artifacts obviously depend on the parameters that control the impostor rendering such as distance and and angle of view. We have designed two experiments: the first experiment is to illustrate how the parameters impact the perception of visual artifacts, and the second experiment is to test if an additional audio modality will impact the perception of visual artifacts.

Since the purpose is the perception of artifacts in the rendered image, the experiments are designed as a test of perception of image quality. Thus, the subjects are required to observe the rendered images instead of blindly navigating in the scene. Besides, since we purposely only present static scenes, the experiment of additional audio modality is also performed by the observation of the same rendered images but combined with stereo sound.

Twenty subjects participated in the experiment and each was presented with two series of snapshots captured from the renderings of the same 3D audio-visual tree scene at different LODs varying from LOD2 to LOD5. The snapshots were captured at ten different angles of view for every LOD. Each LOD snapshot has a reference twin tree snapshot with full geometry, i.e., this tree has the same view but is rendered in full detail. The pairs of snapshots with and without LOD are selected randomly in the first series of snapshots, we test if and when the subjects distinguish the impostor and full detail rendering by perceiving visual artifacts. The second series of snapshots are the same as the first but are accompanied by

stereo sound simulation with different volume based on the distance to viewer. After observing each snapshot, the subjects had to score the level of perception of visual artifacts ranging from 1 to 5, in which '1' and '5' refer to the highest (i.e., obvious visual artifacts) and lowest (i.e., no perception of visual artifacts) levels, respectively.

4 Statistical Analysis of Experiment

We applied statistical method on the score data for analyzing the two similarities firstly between impostor-based rendering and full detail rendering and secondly between impostor-based rendering with sound and without sound.

4.1 Analysis for Perception of Visual Artifact in One Modality

We performed analysis of variance (ANOVA) for every view angle on similarity between impostor (with LOD/impostor) and reference (noLOD/without-impostor) as within subject factors. When *P-value* of ANOVA approaches to 1, there is a great similarity between impostor-based rendering (LOD) and full detail rendering (noLOD).

After the ANOVA, we get the table of *P-value* about every angle of view of impostor. To summarize, the results show that:

1. For LOD2, $P > 0.9$ when angle of view is smaller than approx. 7°;
2. For LOD3, $P > 0.9$ when angle of view is smaller than approx. 2°;
3. For LOD4, $P > 0.9$ when angle of view is smaller than approx. 5°;
4. For LOD5, $P = 1.0$ with all the angle of view.

We conclude that subjects cannot notice the difference between LOD and noLOD for LOD5; and for LOD2, LOD3 and LOD4, subjects observe the visual artifacts when the angle of view exceeds threshold of 7°, 2° and 5° for LOD2, LOD3 and LOD4, respectively. We deduce that under certain thresholds of angle of view and distance, the discontinuity artifacts cannot be noticed by human vision. Besides, these two parameters can be manipulated for controlling LOD selection in real time rendering.

4.2 Analysis for Perception of Visual Artifact with Audio-Graphical Effect

Here, we performed one analysis of variance on similarity between all impostors without sound and with sound as within subject factors. The result of this ANOVA is that $P\text{-}value = 0.76172$ for all data, which means that in general subjects cannot perceive the difference between the snapshot of impostors with sound and without sound. In other words, the stereo sound simulation does not obviously impact the perception of visual artifacts.

We give a supplementary figure about average scores (see Fig. 2). From this figure, we can see that the average scores slightly decrease from snapshots with sound to snapshots without sound for LOD2 (Dist1) and for LOD3 (Dist2), and slightly increase from snapshots with sound to snapshots without sound for

LOD5 (Dist4). Based on the first analysis on LOD2 and LOD3 in the previous section, we analyzed that an additional stereo sound will somehow slightly aggravate the feeling of perception of visual artifacts; conversely, users do not notice the visual artifacts caused by impostors at all for LOD5, therefore sound simulation slightly improves the quality of visual perception.

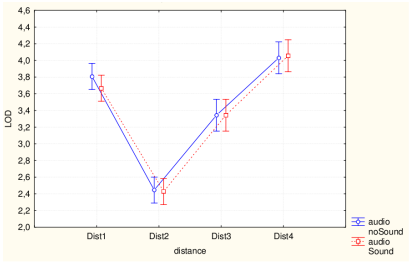

Fig. 2. The average scores for LOD2 to LOD5 with and without sound

5 Conclusion and Future Work

This paper has investigated the impact of audio modality on visual perception of artifacts. The experiments show that simple stereo sound hardly impact the perception of image discontinuity artifacts. However, there is a tendency that sound enhances the visual perception when there is no image artifact perceived and on the other hand, sound slightly aggravates the perception sense of defects when the image artifacts have been perceived. The restriction to only one type of discontinuity artifacts might be a reason that sound does not help reducing the visual perception of artifacts. In the future, we will consider different type of artifacts and integrate graphical rendering with a realistic spatialized sound.

Acknowledgements. The work presented here is partially funded by the Agence Nationale de la Recherche within the project Topophonie, ANR-09-CORD-022. We thank the project partners for the many fruitful discussions.

References

1. Bonneel, N., Suied, C., Viaud-Delmon, I., Drettakis, G.: Bimodal perception of audio-visual material properties for virtual environments. ACM Trans. Appl. Percept. 7, 1:1–1:16 (January 2010), http://doi.acm.org/10.1145/1658349.1658350
2. Funkhouser, T.A., Min, P., Carlbom, I.: Real-time acoustic modeling for distributed virtual environments. In: Proceedings of SIGGRAPH 99. pp. 365–374 (Aug 1999)
3. Moeck, T., Bonneel, N., Tsingos, N., Drettakis, G., Viaud-Delmon, I., Aloza, D.: Progressive perceptual audio rendering of complex scenes. In: ACM SIGGRAPH Symp. on Interactive 3D Graphics and Games (April 2007), http://www-sop.inria.fr/reves/Basilic/2007/MBTDVA07
4. Tsingos, N., Gallo, E., Drettakis, G.: Perceptual audio rendering of complex virtual environments. In: ACM SIGGRAPH 2004 Papers. pp. 249–258. ACM (2004)

5. Tsingos, N., Funkhouser, T., Ngan, A., Carlbom, I.: Modeling acoustics in virtual environments using the uniform theory of diffraction. In: Annual Conference on Computer Graphics and Interactive Techniques. pp. 545–552. SIGGRAPH (2001), http://doi.acm.org/10.1145/383259.383323