Research Masters or Engineering Internship (Stage de Recherche MASTER)

Using Audio-Visual Crossmodal Integration to Guide 3D Interaction

INRIA Sophia-Antipolis, France
REVES (http://www-sop.inria.fr/reves)
Supervisor : George Drettakis George.Drettakis@inria.fr

General Presentation of the Research Domain

The focus of the REVES research group is on image and sound synthesis for virtual environments. Our research is on the development of new algorithms to treat complex scenes in real time, both for image rendering (for example the capture and rendering of trees using an image-based technique [1]) or for sound (for example using perceptual masking and clustering to render complex sound scenes [2]). We are coordinating the new EU project IST/FET CROSSMOD (http://www.crossmod.org), which started on December 1st 2005, on the perceptual interaction between the audio and visual channel and the effects of this interaction on rendering and user attention for both sound and images.

Objectives

Interaction in 3D environments is a very hard problem. Recent studies [3] have shown that the combination of auditory and visual information can help in navigation tasks in 3D virtual environments. In the context of CROSSMOD, we have extensively studied various cross-modal integration phenomena, notably relating to spatio-temporal integration (for example [4][5]), which indicate that it should be possible to enhance the capabilities of the user in terms of perceiving spatial position or temporal occurrence compared to a unimodal (sound or graphics only) environment, by combining audiovisual signals.

In this internship we will be choosing a small set of tasks in an audio-visual environment, such as localisation of sources in a “pursuit”, “assembly” or other game-like scenario, and we will try to determine whether the performance of the user is improved using audio-visual signals. Such experiments typically use very simple stimuli (“beeps and flashes”). We will define a task so that we can first perform a simple test, and then add a more realistic “ecological” setting, with meaningful objects and sounds. We will examine whether the more realistic stimuli enhance performance, and which factors of the difference signals seem to produce best results.

We hope to determine that certain audio signal characteristics significantly improve interaction performance, using the appropriate signal descriptors, such as audio saliency [6] or some variant developed in the context of CROSSMOD. If this is the case, we will construct a short VR application which clearly demonstrates the improvement obtained and its utility for interactive applications.

Methodology

The successful candidate will use C++, OpenGL or other graphics API's and the GPU programming language CG to develop the software required for this research. The interface system will be implemented in our in-house VR extension of the Ogre3D [7] engine, which is supported in our group. Knowledge of psychophysical experimental setups and statistical analysis methods will be a plus for the analysis of the user experiments.

References

[1] Reche, A., Martin, I., & Drettakis, G. (2004, August). *Volumetric reconstruction and interactive rendering of trees from photographs*. ACM Transactions on Graphics (SIGGRAPH Conf. Proceedings), 23(3).

[2] Tsingos, N., Gallo, E., & Drettakis, G. (2004, August). *Perceptual audio rendering of complex virtual environments*. ACM Transactions on Graphics (SIGGRAPH Conf. Proceedings), 23(3)

[3] Matti Gröhn, Tapio Lokki, Tapio Takala, "Comparison of auditory, visual, and audiovisual navigation in a 3D space"ACM Transactions on Applied Perception (TAP) Volume 2 , Issue 4 (October 2005) 564 - 570

[4] Wallace MT, Roberson GE, Hairston WD, Stein BE, Vaughan JW, Schirillo JA. Unifying multisensory signals across time and space. *Exp Brain Res.* 2004;158(2):252-8

[5] Lewald J, Ehrenstein WH, Guski R. Spatio-temporal constraints for auditory--visual integration. *Behav Brain Res.* 2001;121(1-2):69-79.

[6] Kayser C, Petkov CI, Lippert M, Logothetis NK. Mechanisms for allocating auditory attention: an auditory saliency map. Curr Biol. 2005;15(21):1943-7.

[7] http://www.ogre3d.org