A QoE Evaluation on the Influence of Spatial Audio on Visual Attention in 360° Videos

Amit Hirway

Dr. Niall Murray, Dr. Yuansong Qiao

Motivation

Research on neurological analysis of the human brain shows evidence of **multimodal information processing**. However, research in the area of human visual attention do not take into consideration the impact of the audio modality. In practice, visual signals often come along with audio. Therefore, it is natural to investigate the influence of audio on visual attention.

Previous research in the area of "Audio-Visual Attention" [4][5] addresses audio-visual focus of attention in traditional videos. Many experiments suggest that acoustic modality and the audio-visual cross-modal interaction play important roles in attention and quality perception.

Public datasets (with head & eye tracking data) [1][2] and viewing behaviors [3] of users watching 360° spherical videos have been made available for research. However, none of these have considered audio as a stimulus.

Research Aim and Objectives

This research aims to understand audio-visual cross-modal interaction for 360° spherical content with spatial audio and its effect on user's QoE .

The objectives of this research are:

1) To collect a set of 4K resolution 360° videos with non-spatial and spatial audio - duration of the videos to be between 3 and 5 minutes to cover as many different viewing behaviors as possible.
2) To classify the videos according to a set of categories that attempt to describe the expected viewing behavior of a user watching the video
3) To perform eye-tracking experiments in AV (video with non-spatial and spatial audio) conditions for each video
4) To compare eye-tracking data for each of the two conditions in order to verify the salient objects
5) To perform a QoE evaluation of the influence of spatial audio on visual attention

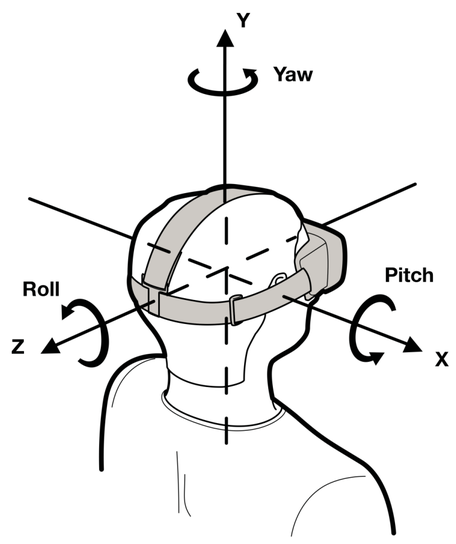

Fig 1. 3-Degrees of Freedom in Virtual Reality

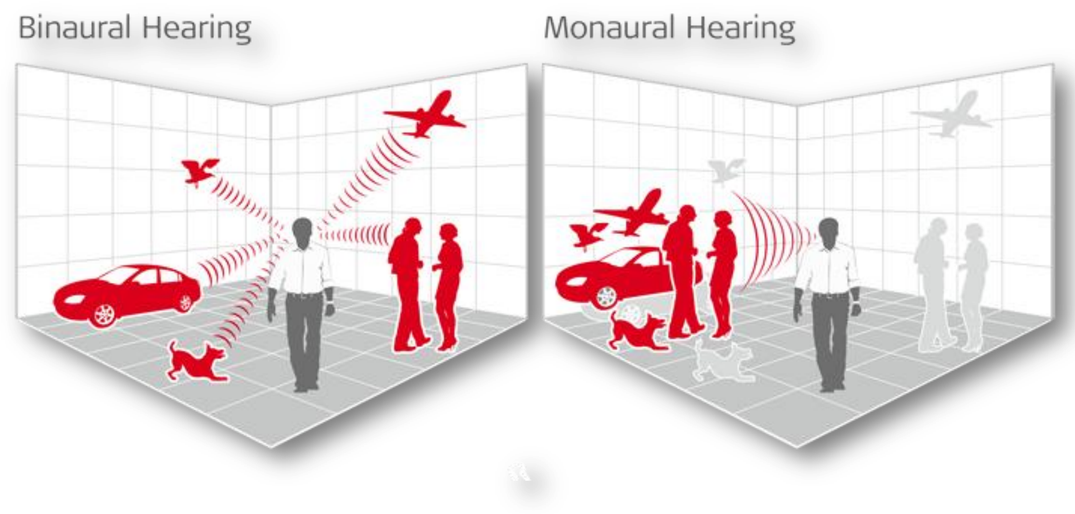

Fig 2. Binaural vs. Monoaural Hearing

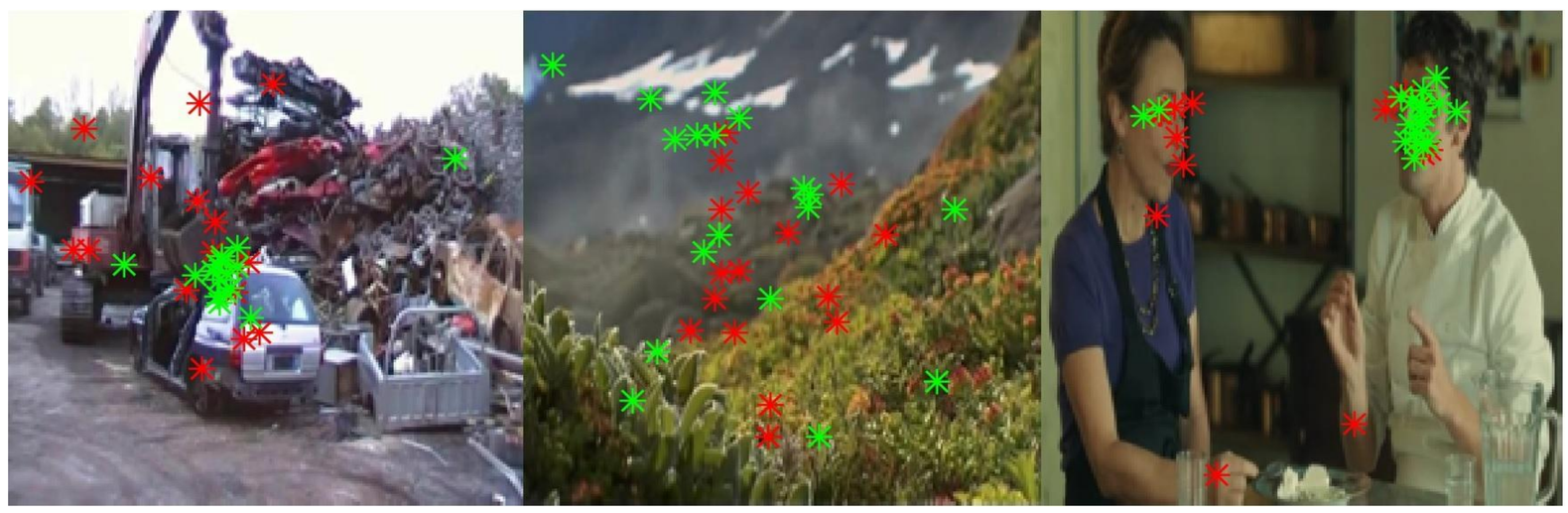

Fig 3. Examples of eye tracking- red and green points represent gaze in two different conditions

What is 360 ° video & spatial audio?

360-degree video, also known as an immersive video or spherical video, is a video recording where a view in every direction is recorded at the same time. It is captured using an omni-directional camera or a collection of cameras.[7]

Spatial audio is audio captured identically to the way we hear the world. When you listen back to binaural audio recording on any 2-channel system (so any pair of speakers, any pair of headphones) you will feel like you are there in the moment, hearing it identically to the way the recordist did when they were capturing it. [6]

Methodology

- Provide participants with information on the test and screen them for visual and auditory defects
- Use a series of training videos with non-spatial and spatial audio for familiarizing participants with the environment
- Run the actual tests
- Implicit (eye-tracking) and explicit (QoE) data capture

Conclusion & Future Work

With this research, we can conclude whether *spatial audio* can affect visual attention in 360° videos and evaluate its influence on users' QoE.

References

[1] Xavier Corbillon et al., "360-Degree Video Head Movement Dataset," Proceedings of the 8th ACM on Multimedia Systems Conference - MMSys'17 (June 2017)

[2] Erwan J. David et al., "A dataset of head and eye movements for 360° videos," Proceedings of the 9th ACM Multimedia Systems Conference, June 12-15, 2018, Amsterdam, Netherlands

[3] M. Almquist and V. Almquist, "Analysis of 360° Video Viewing Behaviours," Dissertation, 2018

[4] J.-S. Lee et al., "Influence of audio-visual attention on perceived quality of standard definition multimedia content," Procemfaedings of the International Workshop on Quality of Multimedia Experience, San Diego, California, USA, 2009

[5] G. Song et al., "Sound effect on visual gaze when looking at videos," in Proceedings of the European Signal Processing Conference, Barcelona, Spain, 2011

[6]https://hookeaudio.com/blog/binaural-3d-audio/binaural-audio-complete-guide/

[7]https://en.wikipedia.org/wiki/360-degree_video