Binaural-Visual Avatars: live streaming binaural audio and video from a mobile point of audition/view

Dr. Jon Aveyard, University of Central Lancashire

‘Balconi’ was a 2021 performance that took place at the Harris Museum, Preston amidst restrictions caused by the Corona-virus pandemic. The project was led by Les Gillon and event managed by Tony Rigg. At this event, the music ensemble Cold Bath Street, led by Simon Partridge, performed a largely improvised piece entitled Clouds. The musicians were distributed across the four floors of the museum’s central atrium whilst three ‘avatars’ moved amongst them live streaming video and binaural audio. No audience was permitted in the performance space; instead they were given a weblink which would provide access to three Youtube pages, one for each avatar. The audience could stick with one avatar or switch between the three as and when they chose to. Binaural audio was used so as to create an immersive experience when heard over headphones. Youtube was chosen because it was considered to be the most accessible and widely understood platform available.

My brief was to find two people to join me as avatars for the pilot project and the public event, to choose and acquire the necessary technology to equip the avatars, and to liaise with Rob Conroy who was managing the webpage and the Youtube channels. The technology I was gathering was expected to provide the binaural audio and moving image to be live streamed over Youtube in synchronicity or near synchronicity by three avatars. There was also the wish from the project leader to minimise the number of steps between the audio-visual data being captured and it being received by the audience. The use of binaural audio, the use of Youtube, and the minimising of steps were all imposed with good reason but each brought problems.

A small amount of funding was available but not so much that equipment could be obtained speculatively. What was purchased or hired had to be fit for purpose. There

were also the significant constraints of trying to carry out this research whilst in lockdown meaning that none of the equipment I might ordinarily have been able to borrow from the university or from colleagues was available for trying out ideas.

The first technological solutions considered were smartphones and tablets. These are lightweight and highly manoeuvrable devices. However, most models don't accept stereo audio inputs and those that do aren't necessary able to take stereo input from any given source. Additionally, live streaming to Youtube from a mobile device requires the Youtube channel to have over 1000 subscribers, something that wasn't available to us.

GoPro cameras have the same strengths and weaknesses as smartphones – they are lightweight and manoeuvrable but unable to take stereo audio inputs and require over 1000 subscribers for Youtube.

Combining smartphones or GoPros with wireless transmitters for the binaural microphones, and sending the video and audio signals to be mixed and streamed from a laptop was also considered but rejected as having too many stages that could potentially go wrong, particularly given that the lockdown meant there were very few opportunities to test this as a team and the number of personnel who would be able to be present on the day to manage the technology would be greatly restricted.

Laptops were considered: they can accept a stereo audio input through the use of a USB audio interface and don't require +1000 subscribers to live stream to Youtube. However, they are bulky to carry and rarely have a rear camera meaning that the avatars would have to use the front camera and would be streaming without having sight of their own image. The only way for them to see their image would be to carry a second

device on which they could watch their own stream over Youtube – a complex and impractical solution.

The selected approach was to use Microsoft Surface Pros. These have the benefits of the laptops in that they can receive stereo audio input via a USB audio interface and they do not require +1000 subscribers to live stream to Youtube, but go someway towards matching the smartphones for convenience in that they have a rear camera, are comparatively lightweight for a full-size device, and have detachable keyboards so they can be used as tablets. Additionally, it was possible to obtain one of these devices on which tests could be carried out to confirm that this approach would work before additional devices were obtained.

For the pilot and the public event, three matching Surface Pro 5s were hired, each connected to a Scarlett audio interface and then to a pair of OKM Soundman Studio binaural microphones. Each device was installed with OBSStudio outside broadcasting software. The binaural audio and laptop camera images were received by OBSStudio and then streamed to one of the three Youtube channels.

The three avatars were myself, Jonathan Robinson who had previously been a musician as part of my Points of Audition binaural audio project and so was familiar with some the techniques required, and my wife Anna Debbage with whom I could spend whatever time was required to brief her on the techniques despite the coronavirus restrictions.

The guidance that was given to the avatars was as follows:

We can stand still whenever we want and move whenever we want. We are aiming for steady head movements to present coherent binaural audio. Your head movements should match the direction of the camera as much as possible. Movement relative to the drum kit creates particularly convincing binaural effects.

We are aiming for steady movement of the video from the rear camera of the tablet. The video is a point-of-view shot so avoid walking backwards, avoid impossible points-of-view that a person couldn't manage, and avoid lingering shots of things that a visitor to the gallery would not choose to linger on. Generally, you should give musicians attention when the music is going through a period of change and be drawn to individual musicians as they start playing or are in a period of change. Occasional changes of elevation of screen (e.g. looking over the balcony, looking up to the floor above or to the ceiling) work well.

Don't spend too long observing the musicians from afar. It works well to vary the distances to the musicians.

It works well to linger on the artworks. Hold the shot and make sure the artwork is close enough and central. On the larger canvases, it is interesting to have the painting entirely fill the frame straight on. Closeups of the statues also work well.

Use the musicians' gaze to help inform your shots, e.g. if a musician is clearly watching someone on a different floor, maybe move from a shot of that musician to a shot of whoever the musician is looking at.

If you ever feel that you have made a poor choice in your direction of movement, don't be too obvious about correcting your path.