UNIVERSITY of LIMERICK

OLLSCOIL LUIMNIGH

Exploration of the impact of high-definition (multi-channel) spatial audio on the manipulation of non-linear audio components in an interactive environment

SUMMER BURSURY RESEARCH SCHEME (2017)

WALTER NIED walternied@gmail.com

SUPERVISOR: DR. KERRY HAGAN kerry.hagan@ul.ie

Literature Review

Introduction

The concept of hypermedia has been solidified as an important aspect of our lives in today's world. First hypothesised by Vannevar Bush's groundbreaking concept of the 'Memex' (Bush 1945) from which the hypertext protocol and later, the world wide web, took inspiration from. Hyperaudio is an area of hypermedia in which "an arrangement of auditorily presented material represented within locally coherent hyperlinked nodes." (Zumbach et al 2014). This idea still rests upon the former idea of hypermedia as the information can be traversed non-linearly.

Spatial sound is an area of audio reproduction in which sound gains the addition of space to its dimensions. This can be done in several different loudspeaker set-ups or through headphones. If these systems can be combined successfully they could be used for blind accessibility to the internet, information immersion and information exploration. In this literature review, I will be discussing the topic of hyperaudio with more scope than the aspect of spatial sound as hyperaudio can be seen more as a 'problem', which I will discuss.

Hyperaudio

Hyperaudio's history is a very complicated one but origins seem to lie in the 90's where the first notions of the idea where in the article *'Hyperspeech: Navigation in a speech only environment'* (Arons 1991). Even though the article does not mention hyperaudio at any point, the basis of the system prototyped employs the exact layout and thinking behind the concept of hyperaudio. Many projects followed in the years to come such as the '*AudioStreamer*' (Schmandt et al 1995), '*Dynamic Soundscape: mapping time to space for audio browsing.*' (Kobayashi et al 1997) and '*A 3D sound hypermedial system for the blind*' (Lumbreras et al 1995). Of these examples, all employed spatialisation to realize their prototypes. One problem discerned early on by Arons was that:

> While speech is a powerful communications medium, it exists only temporally—the ear cannot browse around a set of recordings the way the eye can scan a screen of text and images.
>
> (Arons 1991)

The implications of this is a loss of orientation when navigating auditory information. Lumbreras et al discovered this problem and tackled it in their prototype:

> In our approach by using a spatial metaphor, the clickable [assistant] speaker exists all the time regardless of whether it is speaking or silent. This concept can be of great value when applying the proposed grab-and-drop technique.
>
> (Lumbreras et al 1995)

The employed 'grab-and-drop' technique is an effective navigational interface between user and system allowing the user to interact with the audio and provided an intuitive way to utilise the assisting functionality. One of these was the 'assistant speaker' who was ever present next to the user. This assistant would provide constant feedback on system status and actions executed by the user. The article concluded with an indecisive view but clear was that further work was needed with more advanced technology to realise a truly usable system. On top of that, it was clear that 3D audio was a fundamental element that helped this project come to fruition. This was also clear from ESPACE 2 (Sawhney et al 1996) where they concluded "Users agreed that 3D spatialisation of the sound sources in the environment would improve navigation and representation of simultaneous audio". As it is clear from all the above prototypes is that the main objective of the implied systems is to enable the visually disabled to access information. A huge drawback of all concepts is the lack of integration to the world wide web as it was difficult to achieve with technology present at that time, but when considering a system today, it should be a compulsory goal.

Before going on any further, an important auditory phenomenon that needs introduction is the cocktail party effect. Explained perfectly in the article *'A perspective on the limited potential for simultaneity in auditory display'* (Gossman 2012) it shows the thinking behind many research attempting to exploit this effect:

> A speaking voice, figuring prominently in the famous auditory scene example of the cocktail-party, is characterized by a persistence that allows the party guest to navigate the auditory scene with their attentional focus. But the interior, the content of the stream is characterized by variability: What is being talked about, how it is being said, the specific sounds of vowels, consonants, phonemes, how the physiological performance of the speaker contextualize the individual voice, etcetera: The modal stream can be interpreted as an interface that allows the discovery of previously unknown aspects and properties of the environment. Upon closer inspection,

streams can in turn disintegrate into a manifold of independently observable features: Streams within streams, accessible within one another through progressive attentional disclosure as it was described for example in Merleau-Ponty's phenomenological analysis of perception.

(Gossman 2012)

The cocktail party effect is seemingly the ideal phenomenon to exploit in terms of spatialisation in a hyperaudio system. It would allow a user to use their attention to browse through auditory information just like the eye scans a large body of information. Kobayashi et al, Lumbreras et al, Sawhney et al, and Schmandt et al; all decided to exploit this phenomenon. Later in Gossman's article, he goes on to explain how streams that are played simultaneously will either have the effect of a *'congruent merge'* or a *'destructive merge'* depending on the nature of the audio present. He sums up the section with the ramifications of the cocktail party effect:

For example, it seems evident that we only have the potential to fully engage and understand a single stream of type speech. Multiple simultaneous language streams will lead to a discrimination of the streams into attended and peripherally attended speech—or, if that is not possible, confusion and unintelligibility are the consequence.

(Gossman 2012)

What is implied is that while it may seem plausible that a system that exploits such a phenomenon would enable the user to ingest more information, full concentration for extended periods of time can only comfortably be given to one dominant stream.

More work has been done since on the topic of hyperaudio but not enough to say that there is a clear concise way of conceptualising a system. Further research from Zumbach et al (2014) article *"Hyperaudio learning for non-linear auditory knowledge acquisition"* delved into the more psychological side and investigated whether non-linearity increased cognitive load (amount of stress a person's thinking is under to understand the information presented). Evidence from Zumbach et al show that hyperaudio did not prove to be superior to traditional text based or linear auditory information. In its abstract the authors state:

Interaction effects indicate that non-linearity increases cognitive load assessed with a self-report measure in auditory instruction compared to linear information presentation while cognitive load in processing written text is not affected by linearity. Further, effects reveal that the text type (ex-pository vs. linear text type) interacts with presentation format showing that expository text leads to comparable learning outcomes in linear and non-linear formats, while presenting linear text type as hypertext or Hyperaudio is here rather unbeneficial.

(Zumbach et al 2014)

Even though the evidence is not final as the difference is not monumental and in the conclusion of the article several hypotheses are brought forward to account for the discrepancy, cognitive load must be considered when presenting information auditorily. It can be implied that because of this, recent attempts in hyperaudio are scarce with more research being done in the areas of sonification and auditory displays.

Sodnik et al (2012) conducted an experiment to explore the feasibility of a 3D auditory interface for blind and visually impaired and how it would compete against a braille reader. 3D referred to the use of binaural audio and tested was completion speed, information gained, text decoration, text alignment and table structure. Even though not a hyperaudio system, it shares much of the same genetics and ideas. Results showed that the reading speed cut in half with the system implemented while information accuracy was upheld.

The auditory interface with spatialised speech proved to be more than 160% faster than the tactile interface. The majority of participants reported on their subjective perception of the difference in speed between the two interfaces. It is important to note that although the auditory interface was faster, it did not cause any degradation in accuracy (Sodnik et al 2012).

Impressive was also the subjective feedback collected from all participants as each where asked to use both methods. The overall conclusion was that with more work, it would prove to be a viable extension to current screen reading technologies. With more advanced technology, further research and extension into HTML, this could be an important aspect for blind people's accessibility to the internet.

On the topic of HTML extension is a conceptual prototype thought of by Goose et al (1999). This prototype focused on the aspect of how an application could be made to read and interpret a HTML document and spatialise the information intuitively. Even though no experiments with subjects were conducted, the project brought forward some interesting ideas. One of these was devising earcons to inform the user of location. A specified set of earcons was implemented to inform when a user clicked on an intra-link (from one place to another inside the same document) or an inter-link (one document to another). These were over embellished but ensured the listener knew what was happening.

A study conducted by Kobayashi et al (1997) showed the exploitation of spatial memory where a sound source defined as the ‘speaker’ was mapped to orbit the listeners head in an attempt to utilise human’s ability to remember where something was uttered in relevance to their location. Kobayashi et al conclude the article with:

In contrast to the initial implementation where users could not remember the location of audio events, most users reported that they could use their spatial memory for audio navigation with the refined system. When the Speakers were moving at an adequate speed to form memory of the topics, the space seemed to help users to memorise the topics. By observing subjects, we are led to believe that the association between the topics and spatial locations helps to transfer the memory of topics to the long term memory.

(Kobayashi et al 1997)

As 3D binaural audio was implemented as spatialisation of the audio stream, an area that would be of interest to explore would be to see if these results improve or diminish within an eight channel loudspeaker array.

Spatial Audio

The history of spatial sound is more definitive. With early attempts in cinema by Disney Studio’s 1940 movie *‘Fantasia’,* the real use of audio-centric spatial environments where in the 1950’s by composers such as Karlheinz Stockhausen, John Cage and Pierre Schaeffer. These early endeavors where extremely cost expensive as everything had to be done in the analog domain therefore keeping them very scarce. Yet these attempts inspired many researchers and artists to branch out and develop new technologies to enable people to explore the field further. Below I will discuss some of these developments that would be relevant in the terms of devising a widely applicable hyperaudio system.

Binaural audio

This is a method of sound reproduction which provide a convincing auditory experience over two channels by presenting stereophonic audio cues:

Binaural audio is based on a simple assumption: if the signals that would be received at the ears of a listener as a result of an acoustic event are provided to the listener with sufficient accuracy, the person will perceive an auditory event corresponding to the original acoustic event.

(Hacıhabibogˇlu et al 2017)

Since inter-aural time differences and inter-aural level differences are exploited, left and right channels must be kept separate implying the use of headphones. But this is also where its strength lies as it allows a cost-effective method where spatialisation can be of lower resolution and elevation is not an important aspect. Many hyperaudio prototypes have employed this method such as Kobayashi et al (1997), Lumbreras et al (1995), Zumbach et al (2014), and Schmandt et al (1996); which allowed them to easily implement their ideas. However, Binaural audio does come with the draw backs of being reproducible only on headphones and requiring individualised Head Related Transfer Functions (HRTF) to maximise localisation as found in the study by Rothwell et al (2016). Binaural audio is starting to become of interest again especially on the sectors of virtual reality, augmented reality and gaming as it is a cost-effective method of immersion.

Ambisonics

In 1969 Michael Gerzon developed a new method called Ambisonics, which was developed to compete with Dolby Surround as a more flexible and efficient way of producing surround sound (Gerzon 1973). First order Ambisonics is based upon the b-format encoding which allows decoding to any number channel loudspeaker arrays making it a versatile system that is capable of transcription.

This would make Ambisonic’s a great spatialisation technique for a hyperaudio system because of the b-format utilised. It would allow a program to be devised that would decode its multichannel audio stream to an end-user’s needs, even binaural. Higher Order Ambisonics (HOA) are used for a more accurate soundfield reproduction but also comes with greater computational and technological costs.

Vector-Based Amplitude Panning

Devised in 1997 by Ville Pulkki to accommodate for more irregular loudspeaker arrays. He says in his opening abstract:

Using the method, vector base amplitude panning (VBAP), it is possible to create two- or threedimensional sound fields where any number of loudspeakers can be placed arbitrarily. The method produces virtual sound sources that are as sharp as is possible with current loudspeaker configuration and amplitude panning methods.

(Pulkki 1997)

This is achieved by a method of vector formulisation in the horizontal plane which utilises tangent panning rather than the generic sine panning found in stereo (Hacıhabiboğlu et al 2017). Initially thought of for the use in geodesic domes where a loudspeaker would be present at each vertices of the dome skeleton and therefore, with added mathematical calculations, capable of being extended into the vertical plane. This could be implied in a hyperaudio system installed in a permanent geodesic dome which would allow for a more immersive environment to be produced in conjunction with visual stimuli.

Pilot Study

Research Question

Looking at the area of Hyperaudio showed some very distinct downfalls that audio-only information systems encounter. The greatest problem to overcome is the fact that all audio is temporal, meaning that once the auditory information has been played it ceases to exist and can only be recalled in the person's own memory. Some researchers such as Lumbreras et al (1995), used an 'assistant' method to overcome this problem. Kobayashi et al (1997) used a different approach by exploiting spatial memory to help users remember when and where certain topics were talked about. This approach will be investigated further in this experiment but using loudspeaker spatialisation instead of headphones.

The developments of spatial audio are substantial when looking at the field alone but in context with hyperaudio, the developments would only prove to be useful once a fully-fledged system has been developed and a need for greater immersion and realism is required. The greatest asset that spatial audio brings to the table is allowing for the exploitation of spatial memory. The space available is far greater and more defined over loudspeaker spatialisation than headphones.

The question that was raised for me was: Can spatial memory be applied further in the field of hyperaudio and can it be used in conjunction with loudspeaker spatialisation to help people navigate auditory information?

SpADE

The Spatial. Audio. Display. Environment or SpADE, is a 28-channel surround sound system set up in the University of Limerick. This system can be used for many things involving spatial sound such as psychoacoustic testing, prototype design and scientific research making it perfect for performing an informal user test in form of a pilot study. Even though a pantophonic fourteen loudspeaker configuration was available, only the eight-channel octagonal configuration was used.

Figure 1: SpADE Lab

Outside of using SpADE, Max/MSP was used to design the test. Max is an object-oriented coding environment with powerful capabilities in spatial sound. Max was chosen as I am comfortable using it.

Concept behind the experiment

The idea of browsing through hyperlinks is core to finding information on the web and seeing this is where a hyperaudio system would be implemented, this would be an important aspect to concentrate on. When looking at visual systems and how they execute such a function, it does not hint at a viable method of transcription to the audio realm. But when looking at spatial memory laid out by Kobayashi et al (1997) and the standard octagonal configuration of a spatial loudspeaker setup, it becomes apparent to how one might devise such a tool.

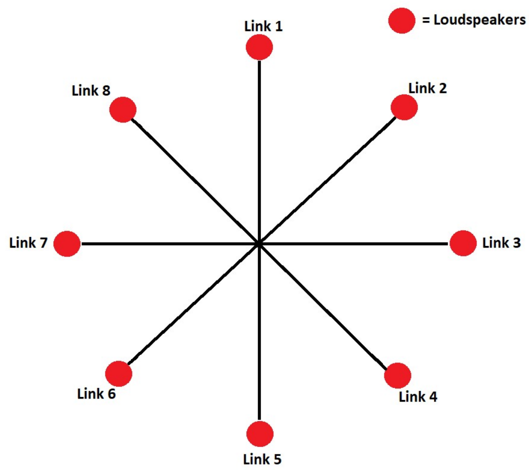

Figure 2: Hyperlink to loudspeaker mapping

As shown above, links will be spoken consecutively from one loudspeaker to the next of the octagonal configuration. This setup is in line with Kobayashi et al's (1997) findings that once exceeding quadrants of twelve in a circle, users find it hard to differ between direction so spatial memory starts to become less effective. Once a link is selected the loudspeakers will be repopulated with eight new links associated with the selected link.

Each of these links represents an article which was done to represent selecting a hyperlinked word on Wikipedia which would contain more hyperlinks to more articles. In future systems, selecting a link would contain the entire article but this was beyond the scope of this experiment. For now, the system will just comprise of singular words and phrases used as representations.

Experiment Conditions

For the input method, the numerical keypad on a computer keyboard was used. Each key around the '5' key was mapped to a loudspeaker of the octagonal configuration with '8' being the loudspeaker in front of the user. Clicking a button once would audit the link at that position and clicking it again would select it sounding a confirmation earcon for user feedback. The '5' key tells the user which article they are in at the that time. Participants were statically seated in the center of the loudspeakers facing the front speaker.

The goal for the user was to complete all tasks in as little clicks as possible. Clicks are when the user selects a link using the keyboard. No time limit was placed on the user. The tasks were as follows:

- Starting on the article "Limerick City", find the article on "Polynesia".
- From there, find the article on the "Mid Atlantic Ridge"
- From there, find your way back to the starting point "Limerick City".

Only certain links were selectable as there were never more than two ways to reach your goal with only one shortest way. Not all populated articles would have the full eight links present e.g. some only had three links at the front

Participation criteria was to be between the age of eighteen and sixty-five and have no aural issues. No knowledge of the field was required. No audio or video recording was made of the experiments. All participants had no strong visual impairments.

Audio used was recordings of my own voice speaking each word. Compression was used to keep them all at the same dB level. Reverb was added in Max/MSP to the entire system for a more natural spacious ambience.

Results

Sixteen people took part in the pilot study over three days. The results were documented on separate sheets with number of clicks and nine subjective follow-up questions which will be discussed later.

Two people achieved the minimum number of clicks required which was 9. The maximum clicks used was 18 which one person needed. Average score was 13 with a standard deviation of 2.94. This abnormally high standard deviation suggests errors in the conducting of the experiment. The histogram below shows all scores.

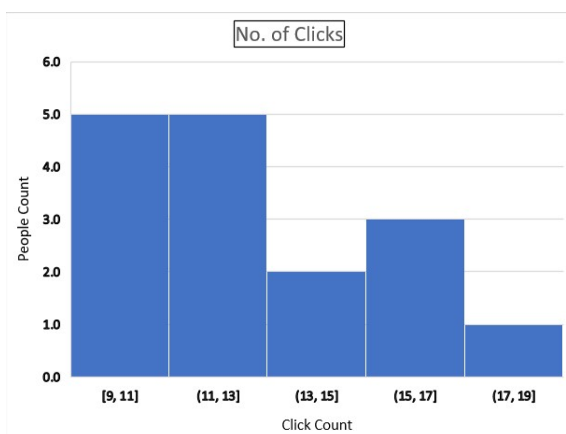

Figure 3: Histogram for number of clicks used by participants.

Following up were these nine questions intended to clarify what the user experienced:

1. Did you find all tasks manageable?
2. Were you lost at any point or did you find the audio only factor disorientating?
3. Did you find the spatialisation helped or hindered with navigating the information?
4. Did you understand the metaphor between links and directions?
5. How did you find the input method and mapping of the keyboard?
6. Did you find the directions were adequately spatialised and mapped?
7. Did you find it strenuous to remember all the words and directions?
8. Would you find this system useful as a browsing or research tool?
9. Any further thoughts/notes/questions?

The answers were documented for the participant in short form in a box under the question. Due to this, irregularities appear in the answers and becomes difficult to quantify. The table below shows these answers as condensed and accurate as possible with interesting answers discussed in the discussion section below. This data cannot be taken as scientific data but is of interest nonetheless.

Q.1	Yes = 14	No = 0		Ambiguous = 2
Q.2	Yes = 1	No = 10	Ambiguous = 2	Slightly = 3
Q.3	Helped = 13	Hindered = 1	Ambiguous = 2	
Q.4	Yes = 3	No = 13		
Q.5	Positive feedback = 13	Ambiguous = 3		
Q.6	Yes = 10	No = 4	Yes but behind listening issues = 2	
Q.7	Yes = 4	No = 6	Didn't try = 6	

Table 1: Answers given by all participants

From the data that was gathered, four people reported no complications and had an average score of 11 clicks with minimum of 9 and maximum 15. The other twelve had an average score of 14 with minimum of 9 and maximum 18. From the problems reported these twelve can be further split up with some overlap of people reporting both. Among others, the two main problems reported were issues with spatialisation and lack of prior clarification of task and system.

Nine participants reported spatialisation problems in the system with varying degrees of dissatisfaction. Three people answered 'no' for Q.6 with documentation showing no reason or further explanation for answer. Four users experienced problems with rear localisation which is common auditory problem. Other errors arose due to system playing back two words simultaneously creating confusion or only eight speakers being used in SpADE out of fourteen present in the lower circle, therefore mental mapping to the keyboard proved problematic for the user.

Six participants reported confusion about principles of the system, what was being tested or user objective. This was down to prior explanation offered at the start of the experiment being unsuited or ambiguous. Users reported that this confusion was cleared up after a few minutes of using the system. These problems do not reflect as heavily in the scores as they do in the scores of the spatialisation problems but this is impossible to draw conclusions from.

Discussion

Interesting suggestions for the use of this system has been brought forward by people in Q.8. One of these was the implementation as a learning environment. Two other people suggested the investigation with children with learning disabilities such as dyslexia.

Beyond that, seven people agreed with the use for browsing or research tool, two people who conveyed their desire to rather use screens and three people being unsure of the real-world application.

Three people vocalised their wish to move around rather than to stay statically seated in one place. While this is not concrete evidence, it would be of interest for further investigation.

A variety of different input controllers were suggested after Q.5 ranging from gesture controllers to X-Y pads to voice commands. Yet many people still found the simple mapping of the keyboard sufficient with one participant stating that it "kept you orientated" suggesting it as a beneficial attribution to the system.

Another participant noted for Q.7 that "yes [it was strenuous] but spatialisation helped". 81% of participants stated that spatialisation helped with navigation (Q.3) which hints at the successful use of spatial memory by these people.

Key Issues

Assessing the usability of auditory spatial memory in an objective way has been tackled incorrectly in this experiment. Instead of measuring the number of clicks used per user and discerning the usability of the system from this, a comparative approach would have been more informative. This could have been done by creating two versions of the system, one being the current eight channel spatial representation and a mono representation as a control group allowing for the comparison of results from both groups. Testing the spatial system against a visual rendering of the system would show real world application and performance for non-visually impaired individuals.

The problems that arose in the study can be surmised in three main issues:

- Poor information prior to test.
- Inadequate documentation and questionnaire.
- System design flaws.

The underlying principles and concept that was explained verbally to the users beforehand was inconsistent without a planned and concise written formulation. This created ambiguity for the participant about their goal. To remedy this, the user should have been given a trial run of different links and topic to help them familiarise themselves with the underlying concept of the experiment, spatial distribution of the links and keyboard mapping.

Debriefing of the participant was done verbally with short hand of each answer written on the experiment sheet under each question. This lead to disparity among the answers given which made quantisation of the results gathered difficult. This could have been minimised by use of a 10-point scale rating for all answers and a comment section at the end. For any verbal communication, a recording should have been made for evaluation of any interesting comments given from the participant and to clear up and doubt in the written answers.

General system design flaws such as use of my own voice, timing and lack of user feedback are all problems that could have minimised some of the discrepancies that some users encountered, skewing their results. Another mistake was the use of a specific topic like geography. This meant some participants felt lesser for not understanding the words spoken and found it harder to associate the words to their goal. A more neutral topic which does not require as much in depth knowledge is required to remove such errors from the results and would make the explanation of the participants goal simpler.

Conclusion

This pilot study aimed to construct a primitive hyperaudio system utilising loudspeaker spatialisation to allow the auditory browsing of hyperlinks. An experiment testing the usability of the system was conducted.

The results from this experiment, which have been skewed by improper conduction do not allow for any concrete evidence of the viability of such a system. Nonetheless, the overall affirmative feedback from participants coupled with the rectification of errors in this experiment would suggest further investigations as promising.

References

Arons, B. (1991). Hyperspeech: Navigating in Speech-Only Hypermedia. In: ACM Hypertext. [online] Cambridge: MIT Press. Available at: http://alumni.media.mit.edu/~barons/pdf/arons_HYPERTEXT91_hyperspeech.pdf [Accessed 10 Jun. 2017].

Goose, S. and Möller, C. (1999). A 3D audio only interactive Web browser: using spatialization to convey hypermedia document structure. In: MULTIMEDIA '99 Proceedings of the seventh ACM international conference on Multimedia. [online] New York: ACM, pp.363-371. Available at: http://dl.acm.org/citation.cfm?id=319649 [Accessed 21 Jun. 2017].

Gossman, J. (2012). A PERSPECTIVE ON THE LIMITED POTENTIAL FOR SIMULTANEITY IN AUDITORY DISPLAY. In: International Conference on Auditory Display. [online] North Ave NW, Atlanta, GA 30332, USA: Georgia Tech. Available at: https://smartech.gatech.edu/bitstream/handle/1853/44421/Gossman_ICAD2012.pdf?sequence=1&isAllowed=y [Accessed 12 Jun. 2017].

Hacihabiboglu, H., De Sena, E., Cvetkovic, Z., Johnston, J. and Smith III, J. (2017). Perceptual Spatial Audio Recording, Simulation, and Rendering: An overview of spatial-audio techniques based on psychoacoustics. IEEE Signal Processing Magazine, 34(3), pp.36-54.

Kobayashi, M. and Schmandt, C. (1997). Dynamic Soundscape: mapping time to space for audio browsing. In: CHI '97 Proceedings of the ACM SIGCHI Conference on Human factors in computing systems. [online] New York: ACM, pp.194-201. Available at: https://www.media.mit.edu/speech/projects/CHI97/CHI97.htm [Accessed 9 Jun. 2017].

Lumbreras, M., Barcia, M. and Sánchez, J. (1996). A 3D sound hypermedial system for the blind. In: 1st

European conference on disability, virtual reality and associated technologies. [online] Reading: The University of Reading. Available at: https://www.researchgate.net/profile/PM_Sharkey/publication/260854455_Proceedings_of_the_1st_European_Conference_on_Disability_Virtual_Reality_and_Associated_Technologies_ECDVRAT_1996/links/00b7d532833de149ff000000.pdf#page=199 [Accessed 10 Jun. 2017].

Pulkki, V. (1997). Virtual Sound Source Positioning Using Vector Base Amplitude Panning. Audio Engineering Society, 45(6), pp.456-466.

Rothwell, C., Romigh, G. and Simpson, B. (2016). Virtual-Audio Aided Visual Search on a Desktop Display. In: International Conference on Auditory Display. [online] Atlanta: Georgia Tech. Available at: https://smartech.gatech.edu/bitstream/handle/1853/56581/ICAD2016_paper_34.pdf?sequence=1&isAllowed=y [Accessed 15 Jun. 2017].

Sawhney, N. and Murphy, A. (1996). ESPACE 2: an experimental hyperaudio environment. In: Conference Companion on Human Factors in Computing Systems. [online] New York, NY, USA: ACM. Available at: http://www.sigchi.org/chi96/proceedings/intpost/Sawhney/ns_txt.htm [Accessed 6 Jun. 2017].

Schmandt, C. and Mullins, A. (1995). AudioStreamer: Exploiting Simultaneity for Listening. In: CHI '95 Conference Companion on Human Factors in Computing Systems. [online] New York: ACM, pp.218-219. Available at: https://www.media.mit.edu/speech/papers/1995/schmandt_CHI95_audiostreamer.pdf [Accessed 11 Jun. 2017].

Sodnik, J., Jakus, G. and Tomaži, S. (2012). The Use of Spatialized Speech in Auditory Interfaces for Computer Users Who Are Visually Impaired. Journal of Visual Impairment & Blindness, [online] 106(10), pp.634-645. Available at: http://search.proquest.com/openview/6f590d9b8febee9d77527270d60fe702/1?pq-origsite=gscholar&cbl=48011 [Accessed 21 Jun. 2017].

Zumbach, J. and Schwartz, N. (2014). Hyperaudio learning for non-linear auditory knowledge acquisition. Computers in Human Behavior, 41, pp.365-373.