Emergent Perception and Video Games that Listen: Applying Sonic Virtuality for Creative and Intelligent Virtual Agent Behaviours

Tom Garner
University of Kent, UK
t.garner@kent.ac.uk

Anna Jordanous
University of Kent, UK
a.k.jordanous@kent.ac.uk

Abstract

Computer-controlled virtual agents, typically referred to in the video games industry as 'non-player characters (NPCs)' can present well-crafted behaviours and evoke engaging and immersive player-experiences. One particular class of behaviours relates to those that portray creative problem solving; the generation of ideas, emerging from the situation and their 'character', to overcome problems and advance the narrative. However, such behaviour in current NPCs is arguably illusory and is only achievable within a controlled and linear video game context. NPCs struggle greatly to portray flexible or creative behaviours within an adaptive or procedurally generated environment and this is even more apparent in their relationship with sound. Contemporary psychology and neuroscience research increasingly advocates embodied cognition perspectives and this paper posits that such theoretical developments offer significant opportunity to advance NPC-AI. This paper proposes that an intelligence framework based upon the recently published theory of Sonic Virtuality, and integrated within an NPC, would offer distinct advantages over current systems. With regards to computational creativity, the particular progression for NPCs relates to creative problem solving. To illustrate this vision, a roadmap for future work is laid out, using Sonic Virtuality as the foundation for a 'synthetic listener', an NPC capable of responding to procedurally generated and external (player-domain) audio. The closing section of this paper further posits that the principles of Sonic Virtuality could be applied equally well to other perception modalities such as visual feedback, presenting a unique avenue of AI research that, ultimately, could dramatically improve NPC's 'humanness' and enable them to evoke a player's sense of immersion and presence that is equivalent to existing scripted AI but in much bigger, more complex virtual worlds.

Introduction

Whether as a means of increasing player engagement and excitement in a video game, learning potential in serious simulations or therapeutic efficacy in a virtual reality therapy; immersion and presence is a valuable commodity. Concurrently, virtual environments (inclusive of virtual/augmented/mixed reality, video games and serious games/simulations) that are increasingly responsive and dynamic (seemingly living environments full of content that reacts to the presence and actions of the human user) are in great demand due to their enhanced capacity to evoke immersion (Shaker & Yannakakis 2012). Immersion can be best thought of as a bi-directional process with two alternative forms. Firstly, 'outward immersion' describes pulling the virtual world closer to the physical world by way of ever-increasingly realistic and engaging environments, an example of which includes using head-tracking and virtual reality head-mounted displays (HMD-VR). In contrast, 'inward immersion' refers to bringing the physical closer to the virtual. This can be exemplified by methods of inputting real-world information into responsive virtual environments such as biometrics, voice commands and gesture detection. Regarding non-player character artificial intelligence (NPC-AI), particularly that which emulates human appearance and behaviours, outward immersion is achieved via displays of relevant and timely actions (from speech and facial expressions to gross motor movements, etc.) that closely reflect the situation as a whole ([virtual] environmental factors, narrative, one's own and others' motivations, affective tone, etc.). An important aspect of these actions is that they reflect a creative approach to problem solving by way of ideas generation. For example, when close to a 'Witch' in *Left 4 Dead* (Valve 2008), friendly NPCs will often call for silence and instruct you to switch off your flashlight as a means of evading the adversary. Of course here the environment and situation are controlled and predictable, allowing this creative strategy to be entirely pre-written by the programmer.

Leading on from the above, Procedural Content Generation (PCG: Algorithmically created content, determined by complex interactions in the virtual environment) is a contemporary approach to deepening immersion via facilitating larger, more complex (outward immersion) and more player-reactive (inward immersion) virtual environments. This raises the question of how NPC-AI approaches that present us with crafted behaviours within a restricted (and often linear) virtual environment can produce the same results in a complex and unpredictable PCG context.

In a related concern, progress in both PCG and HMD-VR largely deprioritises sound with the former focussing upon automated terrain and graphical object generation, whilst the latter primarily examines head-set responsive graphics and integration of user-kinaesthetics (tracking, position, motion, gesture, etc.). To an even greater extent, research and development in this area is overlooking our interactive relationship with sound, with neither the human-controlled character nor the NPC-AI engaging with sound creatively as part of the gameplay or narrative – ultimately reducing sound to window dressing. The concern here is that sound (particularly as an interactive entity) is an extremely powerful means of deepening immersion and, consequently, neglecting sound will significantly limit both PCG and HMD-VR's immersive potential.

The research interest that is laid out within this paper integrates various psychological perspectives by way of a proposed practical innovation; realising contemporary theory on the nature of sound and demonstrating its application in a virtual environment context. The expectation is that this avenue of research will contribute to both theoretical understanding and practical application for AI and virtual environment research, presenting opportunities for greater functionality and more realistic NPC behaviour that goes beyond the limitations of scripted AI to accommodate adaptive and procedurally generated game environments. Video games is the most obvious beneficiary here, specifically relating to player-experience improvements such as heightened immersion and greater engagement. However, it is anticipated that successes in this context could be extrapolated across human-computer interaction (HCI) in general, generating potential application for serious games, simulations/training environments and digital avatars that front HCI systems such as interactive healthcare and e-commerce.

Here the foundations are set for future work. It is not the intention of this paper to provide any computational work, but rather to contribute by way of a roadmap for practical work that, based upon the Sonic Virtuality (and associated) theory, has significant potential to help advance NPC-AI humanness. The remainder of this Introduction gives contextual background; the subsequent section presents a theoretical foundation of a 'synthetic listener' for games, based on the Virtual Sound theory of auditory perception. Then we explore potential applications of this theoretical framework in games and go on to outline how the roadmap presented in this paper will be taken forward in future work. The final section summarises the paper and considers how these ideas can be expanded beyond the auditory domain

Theoretical Context: Sound & Perception

The most widely recognised definition of sound is physicalist (materialist) and treats sound as an acoustic wave emitted from a physical source and transmitted by perturbations along a physical medium (Asutay et al., 2012). Examination of alternative perspectives reveals distinct contradictions with regards to both the nature (*what is sound?*) and the location (*where is sound?*) of sound. Competing perspectives include sound as a property of an object/source (Pasnau, 1999), sound as an event (O'Callaghan, 2009), and sound as both object and event (Scruton, 2009). According to further conflicting perspectives, sound may be located at the object (distal theory, O'Callaghan, 2007), the listener (proximal theory, Nudds, 2009) or somewhere between (medial theory, O'Shaughnessy, 2009). It may even be located nowhere (aspatial theory, Casati and Dokic, 2009). The problem shared by these positions is that they are inherently contradictory and do not account for numerous factors that impact upon auditory perception, in particular those closely related to the listener (their physiology, body chemistry, affective state, cognitive state, etc.). They also lack explanatory power, particularly with regards to non-cochlear sound such as tinnitus (Riddoch, 2012). With regards to practical NPC-AI limitations, understanding sound as a soundwave encourages sound-related information (source, function, semantics, etc.) to be positioned at the sound object (as is commonplace in contemporary game design). If an NPC is affected by the sound object, its sound-related information will be imposed upon the NPC rather than the NPC creating its own information in response to the sound. Therefore, one NPC will always respond to repetitions of the same sound object in the exact same way, causing the NPC to very quickly expose its 'machine-ness' to the player and thereby breaking diegetic immersion.

Application Context: NPC-AI in Video Games

According to Forbes (Gaudiosi, 2012) the computer video games industry is currently worth over £42 billion and is forecast to exceed £52 billion by 2017. For many years now there has been an observable demand for NPCs to be increasingly realistic and engaging (Buede et al., 2013; Drennan, Viller and Wyeth, 2004; Verhagen, Johansson and Eladhari, 2012). This need also extends to serious games applications such as education and training (Verhagen, Johansson and Eladhari, 2011; Zeilke et al., 2009). In recent years, research has called upon the computer games industry to move away from heavily graphics-centric development towards more sophisticated artificial intelligence and there has been an increasing demand for NPCs to exhibit more 'humanness' (Nareyek, 2001; Sweetser and Wiles, 2002). Games developers have responded to this in titles such as *Slender: The Eight pages* (Parsec Productions, 2012) that incorporates randomisation variables into the programmed movements of the antagonist, utilising unpredictability to create the perception (or some might say, illusion) of more organic NPC behaviour. With a focus upon the auditory modality, games such as *Far Cry* 3 (Ubisoft, 2012) and *Alien Isolation* (The Creative Assembly, 2014) endow their NPCs with the capacity to respond, with some sophistication, to auditory information relative to actions made by the human player. Consequently, gameplay is beginning to demand that players maintain a more heightened and detailed awareness of game sound; inclusive of both received auditory cues and the sounds (in-game) that they themselves produce.

Artificial intelligence has an established history of increasing a computer user's senses of presence and immersion with regards to human-computer interactions, by way of natural language processing, that makes the machine itself appear more human (McMahan, 2003). Well-designed AI has been attributed to enhanced player experience and enjoyment, primarily by way of its capacity to craft gameplay challenge (Yannakakis and Hallam, 2007). With regards to immersion, AI within games has been described as "perhaps one of the most influential ingredients for enabling a game player to suspend disbelief long enough to become properly immersed" (Charles, 2003: 9). Emotionally intelligent non-player characters endowed with systems that enable responsiveness to affective game information have been posited as a significant opportunity to increase players' sense of immersion (Nogueria et al., 2013). 'Narrative intelligence' progresses AI-lead immersion a step further, conceptualising that NPC-AI has the potential to procedurally generate ideas and actions that contribute to a game's storyline as the NPCs interact directly with the situation. This draws players into the game by way of attributing greater meaning to their actions by requiring them to consider their choices and behaviours more carefully in the face of greater uncertainty and more numerous and differing consequences (Riedl and Bulitko, 2012). The consistent underlying theme of the above positions is that games AI (in particular NPC-AI) immerses the player by way of greater perceived realism that works in two distinct ways. Firstly, it lowers the degree of difference between worlds as the virtual world and its components appear and behave with increasing similarity to the real world, thereby increasing suspension of disbelief. Secondly, the AIs demand of the player a greater cognitive engagement that pulls them deeper into the virtual world as they ascribe greater meaning and value to their actions.

Sound is also recognised (but often overlooked) as a key contributor to enhancing player experience and immersion within games, often at a subconscious level (Collins, 2013). Lopes and colleagues (2015) have recently drawn attention to how computational creativity (CC) can be harnessed for the sonification of game levels, reporting ongoing work on their *Sonancia* system. With CC generation of sound comes an additional requirement: the need for IVAs to be able to interact with and create appropriate reactions to computationally generated sound outside of an expected, limited range of pre-scripted sounds

A review of game AI by Yannakakis (2012) confirms that artificial intelligence for NPCs within computer video games has improved significantly and notes that increasing industry attention is now being paid to embodied NPC architectures (where the overall AI system incorporates the NPC body and the game environment) but that, at present, there is a distinct requirement for further research to develop this area.

Expanding beyond video games/NPCs, we can also consider the potential benefit to intelligent agent system that facilitate human-computer interaction for other purposes. Applications are both numerous and diverse, with intelligent agents providing solutions as museum tour guides (Kopp et al., 2005), teachers/mentors (Rickel, 2001; Verhagen et al., 2011), storytellers (Theune et al., 2003) and healthcare assistants (Rizzo et al., 2011) to name a few. Research within this area has argued that such systems must, as a primary aim, be able to communicate with human users in a manner that is natural (Gratch et al., 2007; Rizzo et al., 2001). Maatman and colleagues (2005) assert that intelligent agents must be able to provide real-time, multi-dimensional feedback to a user that is appropriate and organic, much as one would expect in an interaction between two humans. Such feedback requires a complex cognitive architecture that can enable the agent to reason and act effectively, the main benefit of which is improved communication that leads to the agent being more able to fulfil its function (Kenny et al., 2008).

Outlined later within this paper, it is intended that the close associations between the theoretical and practical dimensions of the study will enable both aspects to benefit within a single track of research. To elucidate, the expectation is that the research outputs (auditory perception models, AI frameworks and NPC designs) will have immediate practical/technical value as prototype systems and will also provide a proof of concept for the underlying Sonic Virtuality thesis. Successful realisation of this ambition has the potential to dramatically alter the way in which we approach sound from both theoretical and technical standpoints, with prospective applications including sound processing techniques, designs for audio hardware and software technologies, and the implementation of sound within multimedia outputs across the various creative industries. It is also anticipated that this research presents an opportunity to further increase the profile of sound (in particular non-musical, non-speech sound) within multimedia outputs such as film, television and computer video games that often deprioritise sound (Alves and Roque, 2011) despite its significant capacity to immerse, engage and excite (Grimshaw et al., 2008; Wells and Hakanen, 1991).

Theoretical Foundation of Sonic Virtuality

The central proposition of Sonic Virtuality (Grimshaw and Garner, 2015, see figure 1) states that: "Sound is an emergent perception arising primarily in the auditory cortex and that is formed through spatio-temporal processes in an embodied system". Within this concept, sound exists within the mind whilst the acoustic soundwave is only one of many components that together form the sonic aggregate and contribute to the final actualization of sound as experienced by the listener. Some of these components are material (the exosonus) and relate to physical phenomena (the environment, the body and the brain) whilst others are immaterial (the endosonus) and exist within the psychological domain (emotions, beliefs, memories).

The components of the endosonus and exosonus together form the sonic aggregate and the final actualised sound emerges from the precise position and nature of this aggregate, much like the

Kanizsa triangle (see figure 1) emerges from the precise position and shape of the 'pac-men' figures.

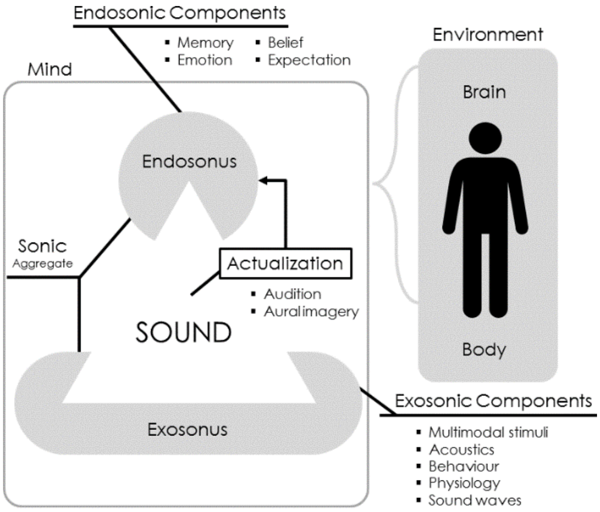

Figure 1: Visualisation of the virtual (emergent perception) perspective of sound

Within the Sonic Virtuality model, exosonic entities (such as soundwaves) do not carry any semantic content, only physical data (frequency and intensity). Meaning can only be attributed as part of the emergent perception as the sound is actualised within the mind. Akin to this proposition, the wider implication asserted by Grimshaw and Garner is that the human mind, within which perceptions and experiences emerge, is not centralised within the brain, but instead is the culmination of the brain, the body and the environment. Within artificial intelligence, this proposition suggests that models of AI should be integrated and embodied such that embodiment variables (e.g. emotional state, memory, multi-modal effects, etc.) would have the potential to significantly impact upon perception and response outputs within AI architecture.

The proposition of sound as an emergent perception is both controversial and, currently, theoretical; drawing largely upon, but not limited to, theories of virtuality (Deleuze, 2002; Massumi, 2014), embodied cognition (Shapiro, 2011; Wilson, 2002), acoustic ecology (Wrightson, 2000) and Heidegger's geworfenheit (thrownness, 1927). The proposition can be elucidated with the assertion that sound (and indeed all of human experience across its multiple forms and modalities) exists not as a physical or material entity, but as a deeply unique and exclusively personal experience of the listener. What we experience as sound is not the oscillation of an acoustic soundwave, but a significantly more complex terminus formed from great numbers of contextual variables. These include the acoustic wave but also accommodate the embodiment factors of body (from the shape of the outer ear to the precise neural network structure of the auditory cortex), the ecological factors of the surrounding environment (from competing soundwaves to the dominant colour of the visual landscape) and the psychological factors of the brain (from current emotional state, to the evoking of associated memories). Prior studies assert that emotional state is of particular significance in terms of its impact upon auditory perception (see Juslin and Laukka, 2004) and the some of our own prior research (Garner, 2013a; Garner 2013b) further argues that this extends to non-musical, non-speech sounds. Consequently, we argue that the role of emotion should be afforded increased attention when constructing a neural network and, particularly, models of auditory perception.

Exploring Potential Applications

This section deliberates further on the anticipated outputs of this research roadmap. Regarding the video games application, the next logical step change in NPC development will involve facilitating a form of sensory processing, utilising synthetic cognition and emotional architectures connected to the NPC's physiology and (virtual) environment. It should be noted that the expectation here is not to create a system of complexity and generalisable application equivalent to an actual human but to attain some form of middle ground between that and the current scripted approach to NPC-AI. With regards to auditory feedback, such processing could enable NPCs to assess sound cues - not only from within the game world (i.e. generated by the game engine) but also from audio input detected from the physical (player's) domain - and generate artificial auditory perceptions that are more complex and contextually grounded. These 'perceptions' could then facilitate creative idea generation for strategic problem solving. As the NPCs in *Left 4 Dead* present the illusion of creative strategy in a controlled environment, potentially so too could our Sonic Virtuality-inspired NPC – only the latter would be able to do so in an more unpredictable PCG environment.

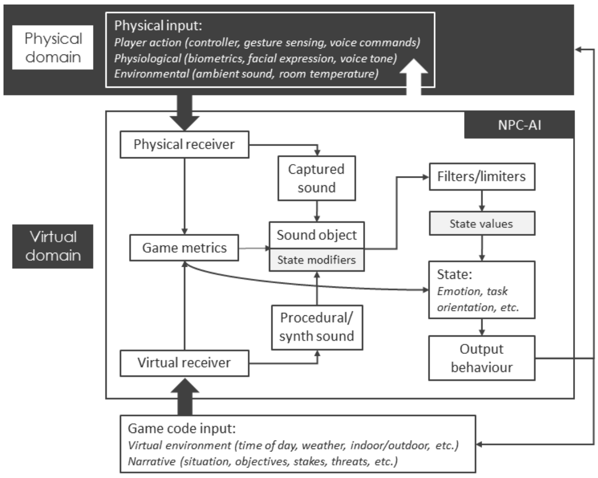

Figure 2: NPC-AI outline

Figure 2 (above) provides the initial concept of a mechanism for such a system. Here, input from the physical and virtual domains is brought together, giving the NPC-AI access to a much greater wealth of game-relevant information. This input can be used to influence the state of the NPC and also facilitate a basic idea-generation framework, the output of which would feedback into the NPC state and also influence output behaviour. To elucidate this proposition (and figure 2), we can ruminate upon our previous *Left 4 Dead* example: Our human player advances through the darkened corridors of Mercy Hospital, dispatching hordes of zombies with characteristic enthusiasm and vigour. Accompanied by her NPC compatriots, the faint sound of a woman crying creeps along the corridor. The NPC's immediate response is to consider this the sound of a survivor in need of help but is then halted when they notice the human player, who has stopped moving and is expressing a great deal of fear. This prompts the NPC to recollect that (based on a rather harrowing prior experience), when alone, Witch-zombies also cry. The NPC also has access to other relevant pieces of information, such as that Witch-zombies are enraged by bright lights and human voices, meaning that calling out or shining torchlight into the darkness to identify the source is too dangerous. They are aware that one of their party is injured, that they are low on ammunition and that their position would make running from an enraged Witch very difficult. These values result in their fear getting the better of them. They begin to turn back when they see that the human player is progressing anyway. Inspired by this they follow. As they edge closer, the NPC considers the problem of identifying the source of the crying without risking an attack. From the available information, they form a creative potential solution; that a gentle, non-verbal communication ('shave and a haircut' tapping on a nearby pipe) could encourage a human to stop crying and call out to them whilst, at the same time, not risking enraging the Witch. Where this system reveals its distinctiveness is that neither the initial interpretation of the sound or the process of generating a creative solution to the source-identification problem, is predetermined by the programmer. Instead the emotional and cognitive states of the NPC (themselves determined by numerous virtual domain and physical domain factors) determine the outcome.

Of course, using present computing power and theoretical understanding, the above example could not be achieved in an entirely uncontrolled environment and the NPC itself could not be endowed with limitless decision-making and response behaviour options. The challenge that this research roadmap would seek to address is in finding exactly where this 'middle-ground' lies and to what extend we can generate responsive NPC behaviours within the confines of contemporary technology

The above attempts to exemplify the practical application of Sonic Virtuality as a perspective that reveals underlying variables that impact upon how sound is perceived and responded to; variables that are commonly hidden from view in more traditional forms of sound analysis (see Grimshaw and Garner, 2015). For the player, such additional complexities within AI would mean that they must attempt to consider such factors if they are to make an interaction that leads to success. In 1999, *Half-life* (Valve) was universally praised for its cooperative opponent behaviour. Enemies would work together and carry out sophisticated manoeuvres during gun battles, requiring players to carefully strategise their actions. For the players, the AI generated significant degrees of immersion and presence, as enemies appeared more real and required players to engage with the game in a more complex way (Charles, 2003). The principle is the same for the Sonic Virtuality model of auditory perception, which expects to progress game engagement a step further by way of advancing AI realism and demanding increased cerebral engagement to draw in the player for a more immersive experience.

Sonic Virtuality in Wider HCI Applications

As stated earlier, realistic feedback and communication is the primary functionality of AI agents within applications such as healthcare, education and tourism. It is envisaged that the Sonic Virtuality research would connect to this function by way of facilitating an underlying intelligent system capable of considering a much wider range of factors when processing input from a user, and also producing more nuanced feedback behaviours that align more closely to the situated environment and the precise context of the communication between user and machine.

To again use a hypothetical example to best explain: consider an exchange between a human user and a virtual agent specifically designed to provide ongoing emotional welfare support to older adults. In this scenario the user has been awoken in the night and believes that she has heard the noise of a kitchen cupboard door creaking downstairs. She asks the agent to confirm if there was indeed a noise emanating from below just now, but does not describe or identify it. Theoretically, with Sonic Virtuality architecture the agent would be able to provide an appropriate response to a request that, for a human, may appear exceedingly simple but that contains several highly sophisticated sub-tasks. Presuming it is equipped with the required sensing hardware, the agent would need to detect the relevant soundwave and have it stored for processing. It would need to correctly identify the source and/or event that generated the soundwave then correctly assume that this particular sound is indeed the one to which the user has referred. Following this feat, the agent would finally need to communicate the appropriate information regarding that sound to alleviate the user's concern, requiring the agent to correctly infer their requirements.

Within this scenario, an intelligent virtual agent would (in theory) have access to the necessary cognitive and affective frameworks that would enable such a complex task to be performed appropriately. Firstly, the agent would build a detailed awareness of the contextual factors surrounding the request. The agent would need to be situated both temporally (to correctly infer the sound to which the user refers based upon the time the request was made) and geographically (to infer the relative position of the sound via localisation of both the user and itself). Awareness of the emotional (anxious) and physical (suddenly awoken) state of the user would also be utilised to help infer the correct sound. Short and long-term memory architecture (itself connected to an emotion/cognition framework) would filter insignificant soundwaves leaving the agent to store only soundwaves that meet certain affective (e.g. irregular sounds evoke anxiety) and cognitive (e.g. soundwave that has existing connections to a potential threat) criteria for later processing. Once the most likely candidate soundwave has been identified and the agent has built a complete perception of 'sound' via the sensory data and auto-generated semantic/contextual associations it can then communicate this wealth of information concisely by telling the user that it 'heard' a kitchen cupboard door closing. Further information/inferences also enable the agent to inform the user that they have a guest staying and this is the likely 'sound object'. The user calls out for her son. He replies. He was making a sandwich. All is well.

Outlining an Initial Research Direction

The primary components of this system around which all anticipated research will focus are: the NPC-AI framework and associated avatar model; a game engine that utilises procedural content generation as part of the virtual environment in which the NPC-AI is embedded; biometric and body-sensing hardware to facilitate human-player emotion detection; and finally, physical environment detection and interpretation.

Presenting a practical application, built upon the principles of Sonic Virtuality, that can demonstrate improved functionality and effectiveness when compared against existing contemporary systems, is arguably a powerful means of testing the validity of the Sonic Virtuality concept. We posit that this method of testing is particularly appropriate when the application in question is a digital representation of a human, as demonstrating that a particular model of perception increases a NPCs perceived humanness could imply that such a model is closer in design to actual perception and human experience than contemporary alternatives. Working towards proving Sonic Virtuality, it is anticipated that a bespoke virtual environment and NPC will need to be built, with an underlying AI architecture that accurately reflects the Sonic Virtuality concept. The following are research questions that are deemed essential milestone matters to be addressed en route to the primary goal:

- What impact do embodiment and ecological factors (primarily emotional state) have upon auditory perception of non-musical/non-speech sounds?
- What is best practice with regards to emotion detection via biometrics and can such data support an embodied model of AI auditory perception?
- What is a 'human response' to sound and which features of this should be emulated by AI to make an NPC appear more human?
- What embodiment/ecological factors are most relevant to generating a 'synthetic listener'?

The basic structure of this research track would need to commence with initial investigations (a series of surveys, interviews, secondary research, etc.) to address the above questions. In order to develop an NPC-AI realisation of the Sonic Virtuality concept, it is essential that we first obtain a more robust understanding of how some of the components within this model interact with one another. Exactly what effects emotion, memory, beliefs, behaviour and acoustics (amongst others) have upon our perception and reaction to sound (and inversely, what effects sound has upon such things) is not fully understood and therefore it is asserted that initial research should be directed towards relevant investigation.

To acquire such understanding will arguably also require a series of experimental trials to generate further, more highly relevant data. Much contemporary research has turned to psychophysiology as a means of making psychological observations by way of objective, quantitative data (see Cacioppo, Tassinary and Bernston, 2007; Mulholland, 1973). Electro-dermal activity (see Kylliäinen and Hietanen, 2006) and electro-encephalography (see Murugappan, Ramachandran and Salazi, 2009) are prevalent measures in emotion-analysis studies. The strengths of psychophysiological study support its implementation as part of the research and it would facilitate an in depth analysis of the relationship between physiology, psychological phenomena and sound cues. It is acknowledged that much literature already exists within this field as it pertains to music and speech. The specific interest of this research track however would be exclusively non-music and non-speech sound. Implementation of biometrics is of particular good fit here as physiological measurement is a pivotal input as part of the NPC-AI framework outlined earlier in figure 2.

Connecting the psychophysiological research to the NPC development; the associations between emotion and sound, collected by way of the experimental trials, could then be integrated into the NPC-AI development. The principle being that the design would be informed by directly relevant human activity. The process of translation between experimental findings and the informing of the NPC design involves taking the observed physiological changes in response to various sounds, contextualising the data by way of the qualitative emotion type (as identified by the participant) and then associating both the data and the affect class with behavioural and cognitive factors. For example, certain sounds could be found to reliably invoke a rise in skin conductance response and an emotional experience commonly described by the participant as 'anxious'. This could then be incorporated into the NPC, where both their increased perspiration and state of anxiety could act as variables with the potential to impact upon various output behaviours and also affect processing of future information input.

Finally, in order that the practical outputs above might support the validity of Sonic Virtuality, a series of evaluative trials comparing the NPC designs to a control system is also an essential component of this research track. The authors' suggestion is that this should be conducted on two fronts. The first: a subjective assessment measuring the qualitative differences in participant engagement, enjoyment and immersion. Secondly, a direct comparison of functionality between systems to evaluate the quantitative features of the Sonic Virtuality model (e.g. number of observable variations in NPC output behaviour).

Summary & Expansion Beyond Sound

This research aims to substantiate a highly contemporary and controversial thesis on sound. Sonic Virtuality is controversial in that it challenges long-held presumptions regarding the fundamental nature of sound. It is anticipated that successes in this research could have a significant impact upon video games development, both in general and with regards to sound design. Using embodied AI within a game sound context opens up the opportunity to examine the potential of a new approach to game NPC design which, if yielding positive results, would support a step progression in human-like characters with wide-ranging impact beyond gaming into any industry of technology that implements virtual agents (healthcare, customer service and education to name a few), as increased 'humanness' opens the doors for many exciting new developments.

Beyond the video game sound application there is potential for wider impact across multiple regions of the sound design industry, should the research convince designers to incorporate embodiment and ecological factors into their method. In an age where a substantial quantity of enthusiastic research energy is being funnelled towards the recent generation of head-mounted virtual reality systems, this is arguably an excellent time to ensure that both sound design and NPC-AI are properly represented alongside graphics.

Whilst the focus of this paper has been upon the perception of sound, the wider elements of the Sonic Virtuality thesis need not be confined to the auditory modality. Sonic Virtuality is essentially the auditory modality of a hybridisation of emergent perception and philosophical (see Deleuze, 2002) Virtuality that positions the mind, and therefore all perceptual systems, as an integrated system of the brain, body and environment. Ultimately it is anticipated that the application of Sonic Virtuality across all modalities will be an essential step for development as, within a comprehensive integrated system, we must be able to account for cross/multi-modal effects. Consequently, as this research would expand to accommodate all five senses we have the potential to develop not only a synthetic listener, but also a synthetic seer, taster, and so on.

References

Alves, V. and Roque, L. (2011) Guidelines for Sound Design in Computer Games, in: *Game Sound Technology and Player Interaction: Concepts and Developments* (Ed. Mark Grimshaw) IGI Global, pp.362-383

Asutay, E., D. Västfjäll, A. Tajadura-Jimenez, A. Genell, P. Bergman, and M. Kleiner. (2012). Emoacoustics: A Study of the Psychoacoustical and Psychological Dimensions of Emotional Sound Design. *Journal of the Audio Engineering Society* 60 (1/2): 21–28.

Buede, D., DeBlois, B., Maxwell, D., and McCarter, B. (2013, January). Filling the need for intelligent, adaptive non-player characters. In *The Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)* (Vol. 2013, No. 1). National Training Systems Association.

Cacioppo, J. T., Tassinary, L. G., and Berntson, G. G. (Eds.). (2007). *Handbook of psychophysiology* (Vol. 2). New York: Cambridge University Press.

Casati, R., and J. Dokic. (2005/2010). Sounds. *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/ entries/sounds. (Accessed 27.04.2015)

Charles, D. (2003). Enhancing gameplay: Challenges for artificial intelligence in digital games. In *Proceedings of the 1st World Conference on Digital Games*(pp. 4-15).

Collins, K. (2013). *Playing with sound: a theory of interacting with sound and music in video games.* Cambridge: MIT Press.

Deleuze, G. (2002). The Actual and the Virtual. *Dialogues II* (revised edition), translated by E. R. Albert, 148–152. New York and Chichester: Columbia University Press.

Drennan, P., Viller, S., and Wyeth, P. (2004). Engaging game characters: Informing design with player perspectives. In *Entertainment Computing–ICEC 2004* (pp. 355-358). Springer Berlin Heidelberg.

Garner, T. (2013a) *Game Sound from Behind the Sofa*, PhD thesis, University of Aalborg

Garner, T. (2013b) Identifying Habitual Statistical Features of EEG in Response to Fear-related Stimuli in an Audio-only Virtual Environment, *Audio Mostly* 2013, Piteå, Sweden.

Gaudiosi, J. (2012) *New Reports Forecast Global Video Game Industry Will Reach $82 Billion by 2017*, Forbes Online, http://www.forbes.com/sites/johng audiosi/2012/07/18/new-reports-forecasts-global-video-game-industry-will-reach-82-billion-by-2017/ [accessed 30.04.2015]

Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007, January). Creating rapport with virtual agents. In *Intelligent Virtual Agents* (pp. 125-138). Springer Berlin Heidelberg.

Grimshaw, M. and Garner, T. (2015) Sonic Virtuality, Oxford University Press (In press)

Grimshaw, M., Lindley, C. A., and Nacke, L. (2008, October). Sound and immersion in the first-person shooter: mixed measurement of the player's sonic experience. In *Proceedings of Audio Mostly Conference* (pp. 1-7).

Heidegger, M. (1962) *Being and Time*, translated by John Macquarrie and Edward Robinson. San Francisco: Harper & Row.

Juslin, P. N., and Laukka, P. (2004). Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening.*Journal of New Music Research*, *33*(3), 217-238

Kenny, P., Parsons, T., Gratch, J., & Rizzo, A. (2008). Virtual humans for assisted health care. In *Proceedings of the 1st international conference on PErvasive Technologies Related to Assistive Environments* (p. 6). ACM.

Kopp, S., Gesellensetter, L., Krämer, N. C., & Wachsmuth, I. (2005). A conversational agent as museum guide–design and evaluation of a real-world application. In *Intelligent Virtual Agents* (pp. 329-343). Springer Berlin Heidelberg.

Kylliäinen, A., and Hietanen, J. K. (2006). Skin conductance responses to another person's gaze in children with autism. *Journal of autism and developmental disorders*, 36(4):517-525

Lopez, P., Liapis, A., & Yannakakis, G. N., (2015) Sonancia: Sonification of procedurally generated game levels. In 1st *Computational Creativity and Games workshop.*

Maatman, R. M., Gratch, J., & Marsella, S. (2005). Natural behavior of a listening agent. In *Intelligent Virtual Agents* (pp. 25-36). Springer Berlin Heidelberg.

Massumi, B. (2014). Envisioning the Virtual. In *The Oxford Handbook of Virtuality*, edited by M. Grimshaw, 55–70. New York: Oxford University Press.

McMahan, A. (2003). Immersion, engagement and presence. *The video game theory reader*, 67-86.

Mulholland, T. (1973). Objective EEG methods for studying covert shifts of visual attention. In F. J. McGuigan and R. A. Schoonover (Eds.), *The psychophysiology of thinking*. New York: Academic Press.

Murugappan, M., Ramachandran, N., Sazali, Y. (2009) Classification of human emotion from EEG using discrete wavelet transform, *Journal of Biomedical Science and Engineering*, Vol.3 no. 4, pp.390-396

Nareyek, A. (2001). Review: Intelligent agents for computer games. In*Computers and Games* (pp. 414-422). Springer Berlin Heidelberg.

Nogueira, P. A., Rodrigues, R., Oliveira, E. C., & Nacke, L. E. (2013, November). Guided Emotional State Regulation: Understanding and Shaping Players' Affective Experiences in Digital Games. In *AIIDE*.

O'Callaghan, C. (2007). *Sounds*. Oxford: Oxford University Press.

O'Shaughnessy, C. (2009). The Location of a Perceived Sound. In *Sounds & Perception*, edited by M. Nudds and C. O'Callaghan, 111–125. Oxford: Oxford University Press.

Pasnau, R. (1999). What is Sound? *The Philosophical Quarterly* 49 (196): 309–324.

Rickel, J. (2001, January). Intelligent virtual agents for education and training: Opportunities and challenges. In *Intelligent Virtual Agents* (pp. 15-22). Springer Berlin Heidelberg.

Riddoch, M. (2012). On the Non-cochlearity of the Sounds Themselves. Paper presented to *The International Computer Music Conference*, September 9–14, Ljubljana, Slovenia.

Riedl, M. O., & Bulitko, V. (2012). Interactive narrative: An intelligent systems approach. *AI Magazine*, *34*(1), 67.

Rizzo, A. A., Lange, B., Buckwalter, J. G., Forbell, E., Kim, J., Sagae, K., ... & Kenny, P. (2011). *An intelligent virtual human system for providing healthcare information and support*. Madigan Army Medical Center Tacoma WA.

Shaker, N., Yannakakis, G. N., & Togelius, J. (2012). Towards player-driven procedural content generation. In *Proceedings of the 9th conference on Computing Frontiers* (pp. 237-240). ACM.

Shapiro, L. A. (2011). *Embodied cognition*. London: Routledge.

Sweetser, P., and Wiles, J. (2002). Current AI in Games: A review. *Australian Journal of Intelligent Information Processing Systems*, *8*(1), 24-42.

Theune, M., Faas, S., Heylen, D. K. J., & Nijholt, A. (2003). The virtual storyteller: Story creation by intelligent agents.

Verhagen, H., Johansson, M., and Eladhari, M. P. (2012). The Need for Socially Believable NPCs: Game Designers' View. In *ECREA's Pre-Conference: Experiencing Digital Games: Use, Effects and Culture of Gaming*.

Verhagen, H., Johansson, M., and Eladhari, M. P. (2011). Model of Social Believable NPCs for Teacher Training. In *Proceedingsof the 5th European Conference on Games Based Learning* (pp. 771-774). The National and Kapodistrian University of Athens.

Wells, A., & Hakanen, E. A. (1991). The emotional use of popular music by adolescents. *Journalism & Mass Communication Quarterly*, *68*(3), 445-454.

Wilson, M. (2002). Six Views of Embodied Cognition. *Psychonomic Bulletin & Review* 9 (4): 625–636.

Wrightson, K. (2000). An Introduction to Acoustic Ecology. *Soundscape: the Journal of Acoustic Ecology* 1 (1): 10–13.

Yannakakis, G. N. (2012, May). Game AI revisited. In *Proceedings of the 9th conference on Computing Frontiers* (pp. 285-292). ACM.

Yannakakis, G. N., & Hallam, J. (2007). Capturing player enjoyment in computer games. In *Advanced Intelligent Paradigms in Computer Games* (pp. 175-201). Springer Berlin Heidelberg.

Zielke, M. A., Evans, M. J., Dufour, F., Christopher, T. V., Donahue, J. K., Johnson, P., ... and Flores, R. (2009). Serious games for immersive cultural training: creating a living world. *IEEE computer graphics and applications*, *29*(2), 49-60.