This paper discusses the use of visualization to enhance audience and performer understanding of a distributed networked musical performance by addressing issues of anticipation of upcoming sonic events.

Network Performance, Anticipation, Netrooms, Latency, Visualization

1. INTRODUCTION
This paper examines the use of visualization as an augmentation to networked musical performance to allow performers to anticipate the actions of other remote participants, and enhance audiences experience and engagement with network topology in a performance context. Specifically, we address considerations made when developing a visualization made for the participative network performance work Netrooms (Rebelo, 2008).

2. BACKGROUND
Research in networked music performance has focused on high quality audio exchange between peers with a view to facilitate musical interaction (Carrót, 2006, Chafe 2004). Although the need for short latencies is critical for many types of music performance over the network, the medium also provides topologies, which take latency as a key characteristic of a distributed environment. Numerous works have made use of latency to explore notions of distance or indeed to sonify network connections according to temporal delays involved in sending messages between two nodes such as in Ping by Chris Chafe (Chafe, 2001). The range of latencies and jitter inherent in a network connection is intrinsically linked to the type of music that is deemed performable and affects the nature of musical interaction (Chafe, 2004), The Apart Study (Schroeder, 2007) artificial latencies of up to 125 milliseconds are introduced between three studio locations while asking musicians to perform both metrically determined music and free improvisation. Audio-visual documentation of the study shows that even within the same performance, the perception and affect of latency can vary tremendously depending on the intended durational relationships between sound events produced in different location. The study also introduced a number of visual platforms for facilitating or expanding interaction between musicians. It became clear that immediate, micro event interaction occurs mainly through sound. However the introduction of gesture-informed graphic avatars was seen to facilitate a sense of large-scale form and structure. This becomes particularly relevant when a graphic environment entails its own temporality that is related but not the same as what is occurring in sound, potentially allowing for anticipation. Visuals can therefore engage performers and audiences in experiencing music by addressing different temporal frameworks and hence promote readability of structure, form and topology.

Figure 1: The Netrooms Visualization
3. NETROOMS

“Netrooms: The Long Feedback” (Rebelo, 2008) is a participative network piece which invites the public to contribute to an extended feedback loop and delay line across the internet. The work explores the juxtaposition of multiple spaces as the acoustic, the social and the personal environment becomes permanently networked. The performance consists of live manipulation of multiple real-time streams from different locations that receive a common sound source, itself made out of contributions from each stream. Netrooms celebrates the private acoustic environment as defined by the space between one audio input (microphone) and output (loudspeaker). The performance of the piece consists of live mixing a feedback loop with the signals from each stream. Aspects of distributed dramaturgy (Rebelo 2009) that are possibly articulated by this work centre around the ambiguity between participant and performer, contributor and spectator. Numerous performances in 2008/10 reveal how the dynamics of the network can be modulated by the intention and expectation of individuals, causing the work to range from a relatively anonymous acoustic process to an environment clearly articulated by the intervention of performers. Perhaps most significantly, in a single performance, each participant takes both a passive and an active role depending on the constant reconfiguration of the audible network.

By virtue of relatively long network latency inherent in a server style configuration (up to 12 seconds roundtrip), the role of anticipation becomes crucial for how participants engage with the temporality of the piece. As each participant hears the piece from a different perspective, based on the nature of their contribution and the topology of the network, expectation becomes a key factor in sustaining a performative condition. This type of distribution of events across the network means that all participants are intervening and listening to different temporal realities. As a sound event produced in one site takes 6 seconds to “arrive” at the site of performance/mixing, there is not only the need for anticipation (from the point of view of the performer mixing the streams but also the audience) but also a temporal window during which visualization can articulate how events unfold.

Differentiation between each participant’s stream is articulated through spatialised audio using multi-speaker arrays (one per participant). Each stream’s unique timbral characteristics allows for localization as well as for a “in house” mix created by the combination of the streams. Visually, it is important to reflect more directly the number of participants and the relationship between their temporal framework and that of the performer and audience.

Interaction between participants is further articulated through comments and remarks displayed in a live blogging context. This allows not only for immediate communication, but for a sense of shared language to develop. This live blog serves as a useful display for audiences as well as for documentation.

By allowing the performer to anticipate upcoming audio events, they can make proactive decisions regarding the direction to take the performance rather than only rely on reactive intervention.

In past Netrooms performances audience members had no visual cues indicating which sounds were coming from which sites, or how many sites were participating. Towards these ends, the visualization should be in some way representative of the network topology, be extensible to arbitrary numbers of participants (within reason), and provide visual cues as to aspects of sounds being stream from each participant.

Finally, since the visualization is intended for public performance and serves as the sole visual focus during the performance of the piece, the visualization had to be aesthetically interesting.
5. VISUALIZATION MECHANICS

5.1 Data Flow

The visualization for Netrooms was designed to take advantage of the exaggerated (but predictable) latency introduced to audio by the Icecast streaming server. Barring network latency (which is usually less than 100ms) the Icecast server introduces a six second delay between audio transmission and reception due to encoding and decoding. For the purposes of the visualization we introduced a second data channel between remote participants and a second computer used exclusively for projection of the visualization in the main performance site. This secondary data channel is a direct peer-to-peer UDP connection with no mediating server.

The secondary data channel is used to transmit an OpenSoundControl (Wright, 2005) encoded stream of data consisting of a sender identifier, and a single floating point representation of the amplitude of the audio from each participant before it is sent to the Icecast server. Additionally, the local audio master machine sends the visualization machine amplitude data for the audio stream that is being sent back to each client. Because of the direct network connection between remote performers and the visualization machine, the latter receives amplitude data approximately six seconds before the actual audio arrives at the audio master machine (Figure 3). This provides an opportunity for pre-visualization of incoming audio data.

5.2 Visualization

Early versions of the visualization consisted of a series of pairs of lines, one pair for each site. The upper line of a pair represented the amplitude of audio coming from a remote participant, and the lower line of a pair represented the amplitude of audio being returned to the participant. The left side of the visualization represented the remote sites, and the right side represented the main performance site. An audio impulse at a remote site would be seen as a peak in the line travelling from the left to the right side of the screen over the course of the six second delay. Once the peak reached the right side, the audio would be heard at the main performance site. Similarly if peaks were produced in the mix at the main performance site, they would travel from right to left along the lower line and be heard at the remote site once they reached the left side (Figure 4).

![Diagram showing audio and OSC data flow in Netrooms](image_url)
In this iteration, the centre (hub) of the formation represented the main performance site, and the ends of each spoke represented each remote site. Similar to the first iteration the upper line for each spoke represented data on route to the main performance site, and the lower line represented the data being returned to each site.

The problem with this iteration was that due to the necessity of the slow rotation of the visualization, it would be difficult to track which pairs of lines represented which site. Accordingly for the next iteration, a shift was made away from plain lines, rather opting to use differently styled brushes to represent each site. In this version the audio data modulated the size as well as vertical position of each brush stroke (Figure 6).

The latest version of the design (Figure 7) addresses most of the goals and provides a readable but not literal visualisation of how temporal events are articulated in the work, in particular in relation to how inherent network latency becomes a means for pre-monitoring (albeit in a way that only relates to envelope following).

The visualization for Netrooms was developed in the graphical programming language VVVV (http://vvvv.org). It runs on a Windows laptop separate from the main Netrooms audio computer. The stylized lines used in the current version are created by displaying 500 billboarded rectangles at regular points along a Bezier curve. Each of these rectangles is textured with a single image, and alpha blended together to create a pseudo-volumetric effect.

Although this early version of the visualization was functional it failed to be representative of the actual topology of the network. Also the number of participants visualized was limited by vertical screen real estate, and the aesthetic appeal of the visualization was limited.

To address these issues a second iteration of the design placed the line pairs in a hub and spokes configuration within a three dimensional space, using Bezier curves in the vertical direction to further differentiate the sending and returning paths for each site (figure 5). The visualization would slowly rotate around the Y-axis to ensure no site’s data was obscured for an extended period of time.

Figure 4: An audio impulse from a remote site at 1 second, approaching the main performance site at 5 seconds, and returning to the remote site at 7 seconds.

Figure 5: Second iteration of the visualization at rest (top), and with audio modulation (bottom).

Figure 6: Close up of brushed lines showing some rhythmic audio.
6. CONCLUSIONS AND FUTURE WORK

Aspects of social networking and community forming are important drivers for works like Netrooms. It is important not only for the audience but also for each participant to have a sense of how each site and its sound world relate to each other. As neither the participants, the performer in the main site, nor the audience ever hears all sound streams due to the interventions inherent in the mixing process, visualisation provides an opportunity to display not only what is heard but also what is unheard, hence contributing to render performative decision making readable.

Future work aims to distribute the visualization to each participant and to incorporate aspects of live blogging with a view to rendering verbal exchanges between participants and performers. Further work dealing with anticipation can be envisaged. The latency window allows for visualization to provide unique means for establishing the relationship between synchronous and asynchronous sound with a view to rendering the network as an audio-visual space, which operates in distributed temporalities and dynamic topologies. Performances such as Netrooms provide a rich environment for the exploration of online interaction that takes temporality as a key factor without placing itself in the realm of immediate (low latency) distributed musical performance. With increasingly low latencies (less than 90 milliseconds), network performance can begin to resemble traditional single site performance from the point of view of musical interaction. Netrooms’ open participative model aims to engage the non-performer and to celebrate the role of the active participant. Although designed as a visualization for sound events, the system described here also becomes a painterly interface from which participants and performer will take cues and exploit in the context of how they operate in their sound worlds and environments.

7. REFERENCES

Carôt, A, Krämer, U, Schuller, G. Network Music Performance (NMP) in Narrow Band Networks presented at the 120th AES convention, Paris, France, 2006 May 20-23


