# *Maps and Legends*: Designing FPS-Based Interfaces for Multi-user Composition, Improvisation and Immersive Performance

Robert Hamilton

Center for Computer Research in Music and Acoustics (CCRMA),
Stanford University,
660 Lomita Drive,
Stanford CA, 94305, USA
rob@ccrma.stanford.edu
http://ccrma.stanford.edu/~rob

**Abstract.** This paper describes an interactive multi-channel multi-user networked system for real-time composition and immersive performance built using a modified version of the Quake III gaming engine. By tracking users' positional and action data within a virtual space, and by streaming that data over a network using OSC messages formatted as UDP packets to a multi-channel Pure Data patch, actions in virtual space are correlated to sonic output in a physical space. Virtual environments designed as abstract compositional maps or representative models of the users' actual physical space are investigated as means to guide and shape compositional and performance choices. This paper analyzes both the technological concerns for building and realizing the system as well as the compositional and perceptual issues inherent in the project itself.

## 1  Introduction

In the context of highly realistic 3-dimensional video games, sound is commonly utilized as a critical element in the communication of virtual spatial cues as well as in the enhancement and definition of actions by gamer-controlled avatars and other game entities alike. By presenting a user-centric sound field to the gamer - where the intended listening audience is one player alone within a front-focused stereo, 5.1 or other commercially standard sound field - designers create insular sound-worlds, reinforcing the immersive experience through their use of realistic 3D models and environments with auditory cues and behaviors based in the reality of the virtual world.

As platforms for creating, interacting with and acting within virtual environments, "First-Person Shooter" or FPS-style video game engines such as the open-source Quake III engine offer artists a powerful new paradigm to explore novel methodologies for the control of sound and music with a low cost to entry and a flexible and extensible development environment. Used in conjunction with a data-transmission protocol like Open Sound Control (OSC) over UDP,

in-game parameters controlled by multiple game users can be passed to any number of music and sound generation software environments, creating a seamless network-based transmission of data from the virtual realm to physical auditory space. In this manner, composers and performers alike can explore new relationships with composed and improvised musical materials by subverting traditional models of performer/composer/audience relationships and by distributing them in new contexts across networked physical and virtual space.

Towards this goal, from the standpoint of the composer, it becomes necessary to explore novel forms of compositional structure based in the virtual environment which are designed to fully exploit the dynamic nature of both control and sound-generation systems alike. At the same time, for composers wishing to retain a level of control or influence over the musical work as a whole, it is necessary to design a system capable of imparting a level of intention and logical structure upon what could easily become a fully improvisatory musical form. One solution - referred to henceforth as the "compositional map" - draws inspiration and structure from the topography of the virtual environment itself and aims to build upon this visual and musical structure by leveraging the inherently flexible and indeterminate system of player motions and actions. Within this composer-defined visual and musical space, in-game performers reacting with their virtual environment as well as with one another via their avatars, collectively move the musical work forwards from inception to completion in a quasi-improvisatory fashion.

## 2   System Overview

*Maps and Legends* is a dynamic software system making use of the immersive environment of the Quake III engine as the user-interface for a flexible composition and performance system. Pre-composed computer-generated musical cells are assigned to each in-game performer and are triggered, processed and controlled through performers' interactions with and paths through the environment. By tracking multiple performers' coordinate locations within virtual space and by subsequently spatializing those locations across a multi-channel performance space, an auditory correlation can be formed between the physical and virtual environments, engaging the attentions of both performers and audience members alike within the musical soundscape.

*Maps and Legends* was designed as a fully-rendered 3-dimensional virtual compositional map built using the GtkRadiant game-level editing software [5] (see Figure 1). While clearly-visible pathways, directional arrows, and active directional jump-pads are built into the map to encourage or force performer motion in certain predefined or composed directions, each performer retains a high-level of independence and improvisatory flexibility, allowing for spontaneous new interpretations of the pre-composed materials.

Users running the game-client software and connected over a standard high-speed network control their avatars and through them the musical work using standard Quake III game control methodologies - typically a combination of computer-keyboard controls for motion and a mouse for view-angle. Game-clients
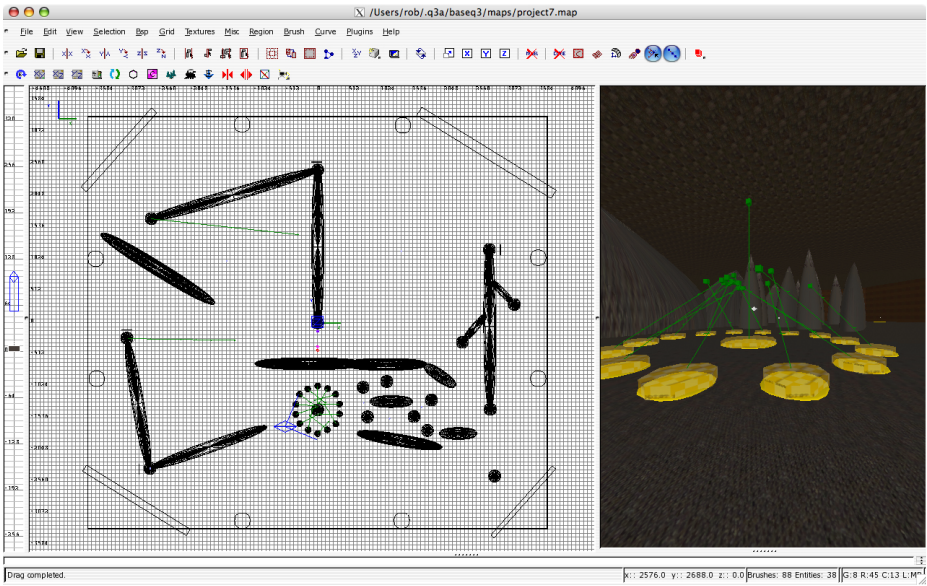
**Fig. 1.** *Maps and Legends* compositional map shown from both a top-down structural view (left) and a rendered "ground-level" view (right) in the GtkRadiant editor

connect to a host game-server which in-turn streams OSC-formatted data reflecting their players' actions and coordinates to a sound-server in the performance venue running Pure Data (PD). Sound generation and processing for each independent performer are handled by and spatialized within an 8-channel PD patch, circumventing the sound-system of the game itself and substituting the composer's musical environment for Quake III's stock in-game sounds and music.

At the heart of *Maps and Legends* is a set of software modules, currently Linux-only, created by multimedia artists Julian Oliver and Stephen Pickles entitled *q3apd* [6] which modify the open-source network code of the Quake III game engine to stream a variety of ingame data - including global XYZ player positioning coordinates, directional velocity and view-angle - formatted as OSC messages over a network as UDP packets. As users connect to the host game server, data from their characters' movements and actions are sent to a PD patch which parses incoming OSC messages and extracts multiple player-specific data points. In this manner, the global position of individual game players within the virtual game-space, and certain subsequent actions performed by each user, are mapped to a number of sound-generation and spatialization control parameters creating a rich interactive-system for musical control.

## 3    Prior Work

The use of networked/multi-user video game paradigms for music and sound generation has become increasingly common as generations of musicians who have
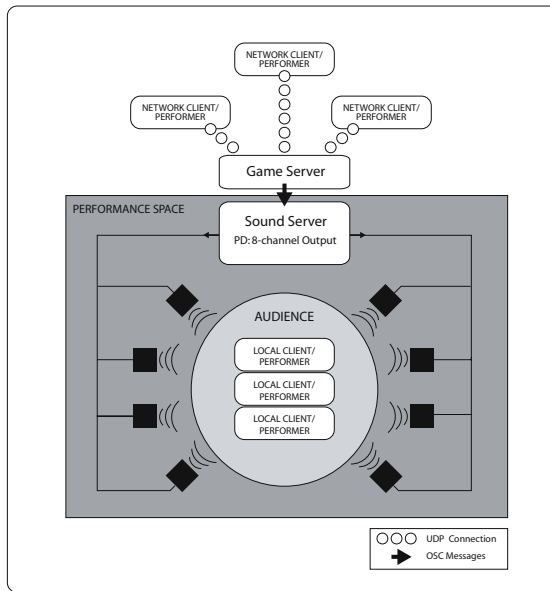
**Fig. 2.** System diagram of Client/Server interactions

grown up with readily accessible home video game systems, internet access and personal computers seek to bring together visually immersive graphical game-worlds, wide-spanning networks and interactive control systems with musical systems. Though its graphical display is rendered in 2-dimensions, *Small_Fish* by Kiyoshi Furukawa, Masaki Fujihata and Wolfgang Muench [4] is a game-like musical interface which allows performers/players to create rich musical tapestries using a variety of control methods. *Auracle* [3], by Max Neuhaus, Phil Burk and their team from Akademie Schloss Solitude allows networked users to collaborate and improvise using vocal gesture. Oliver and Pickles' own works, including *q3apd* and *Fijuu2* [7], a fully-rendered three-dimensional audio/visual installation controlled with a game-pad, tightly marry the videogame and musical worlds through the use of immersive graphics and familiar game control systems. And work on the *Co-Audicle* by Ge Wang, Perry Cook and the Princeton Soundlabs team is actively seeking to build a user-community of collaborative performers through networked extension of the ChucK language and its *Audicle* front-end [11].

Compositional precedents for modular composed forms allowing for performer control over a work's structure can be found in the polyvalent form of Karlheinz Stockhausen's *Zyklus* [9] for one percussionist, as well as in the variable form of his *Klavierstück XI* [10] for solo piano. In *Zyklus*, Stockhausen's strictly composed sectional materials are designed to be interpreted in a performer-chosen direction, reading through the score either forwards or backwards, starting performance on any given page. Similarly, in *Klavierstück XI*, nineteen composed and precisely notated musical fragments are ordered by the performer. These flexible structural concepts are also prevalent in John Cage's body of chance-based musical works

including his *Music of Changes* [1] or *String Quartet in Four Parts* [2], where pre-composed musical cells were selected through chance operations and formed into a definitive compositional structure.

## 4    Mappings and Sound-Generation

By mapping OSC data-streams from the q3apd mod to various sound processing and spatialization controls in PD, a virtually-enactive control system is created allowing for musically expressive and flexible gestural control through the virtual physicality of performers' avatar motion. The linking of virtual gesture to sound and spatialized auditory motion sets the stage for an exploration of musical control through a manner of virtual interactive choreography. Towards this end, sound-generating subsystems making use of sample-playback, active filters, delay and reverb parameters are all linked to various possible performer motions or actions.
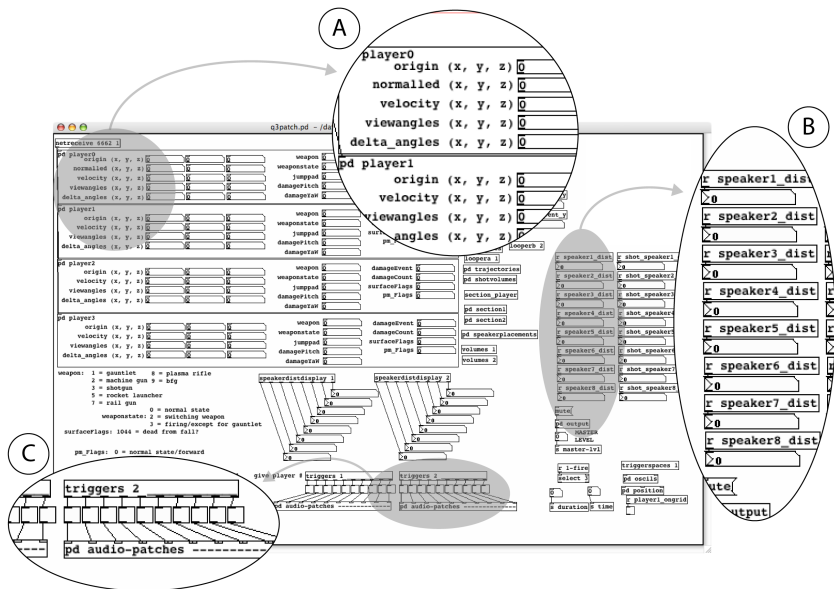
### 4.1    *q3apd* Data Streams

Making use of the Quake III engine's native ability to run user-created software library modifications or "mods," q3apd's customized code streams a number of game-state and positioning parameters for each connected user from the game server to a specified IP and Port address as OSC messages. The q3apd libraries export player-specific data, including XYZ positioning and view-angle, directional velocity, selected weapon, and player states such as jumping, crouching or falling. q3apd also formats each message with a prepended user-id tag ensuring that multiple user data-streams can easily be separated in PD and tracked independently.

### 4.2    PD Mappings

Basic control values supplied by q3apd such as player-motion and XYZ position, (Figure 3: section A) are used to calculate constantly changing horizontal-plane distance vectors from a performer's current position to pre-defined virtual speaker locations within the compositional map (Figure 3: section B). Similarly, distance between multiple performers is calculated and mapped to sound events - complementary or disruptive depending on composer-defined performance states - in an effort to support or discourage performers from moving too close or too far from one another.

At key points in the compositional map, circular trigger regions lie in the suggested path of motion. Precomposed musical materials are triggered when performers move over these coordinate spaces, viewed in-game as bright yellow circles on the virtual ground. In this manner, separate sets of mono sound files, all complementary parts of the composition, are assigned to individual performers and triggered in PD based on each performers' position (Figure 3: section C).

These sets of sound files, a mixture of synthesized and sampled sound, generated using custom software systems built in Max/MSP and PD, make up the
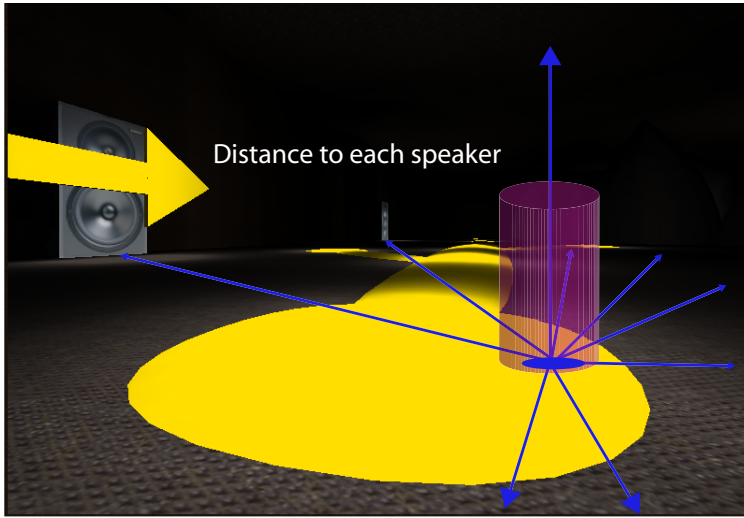
**Fig. 3.** Pure-Data patch with highlighted sections for A) player coordinates, directional velocity and viewangle tracking, B) player-to-speaker distance values, and C) player-specific map triggers

bulk of pre-composed materials for maps and legends. Other basic mappings include a light chorus and reverb processing applied to a performer's current sound when moving over a highlighted pathway - a bonus of sorts for moving in pre-composed patterns - as well as a longer reverb and longer delay applied when the user's Z coordinate indicates that they are fairly "high" in the map's vertical dimension - a state that can be triggered more often by lowering game-server-side parameters like virtual gravity.

By coordinating the speed of various "weapon" projectiles in the game, it has also been possible to create a satisfactory illusion of a "shooting-sound" which travels independently across the map, tracking the trajectory and speed of a visible fired projectile. One particularly effective mapping matches the speed of a large glowing orb moving relatively slowly across the map with the panning for a particular sound event.

### 4.3   Multi-channel Output and Spatialization

To clarify the relationship between individual performers and their respective sound sets, triggered sound files are spatialized based on a performer's distance from each of eight defined speaker locations. In short, sounds follow the performers' motions through space, spatialized between multiple speakers at any given time (see Figure 4).

**Fig. 4.** Multi-channel amplitude is calculated as a factor of virtual distance from performer to each speaker. Shown also are yellow circular trigger-locations, speaker locations, directional guide-arrows and suggested pre-composed pathways.

Speaker locations are defined in PD as XY coordinate pairs, with independent gain levels for each performer for each speaker determined by a simple distance function. In this manner, multiple speaker configurations for multi-channel output can be easily configured without any changes to the compositional map itself. While at this time additional reverb or delay-based panning cues, or more accurate multi-planar spatialization methods like vector based amplitude panning (VBAP) [8] or Ambisonic encoding are not used to either simulate the acoustics of the virtual space or to provide more realistic panning effects, such approaches are being investigated.

## 5   Discussion and Conclusions

As an immersive environment for interactive networked performance and modular composition, the system designed for *Maps and Legends* affords composers and performers alike an extremely powerful and novel musical experience. As the tools for generating virtual environments are flexible and robust, and through the use of q3apd and OSC can interface with a variety of software-based musical environments, there exist many compositional and performance methodologies which can be explored using such a system.

One important goal in building *Maps and Legends* was the development of a system capable of immersing an audience in a virtual sound world, creating a perceptual super-imposition of virtual and physical environments (see Figure 5).

And during the early stages of system development it became clear that this relationship between virtual environment and physical listening space played
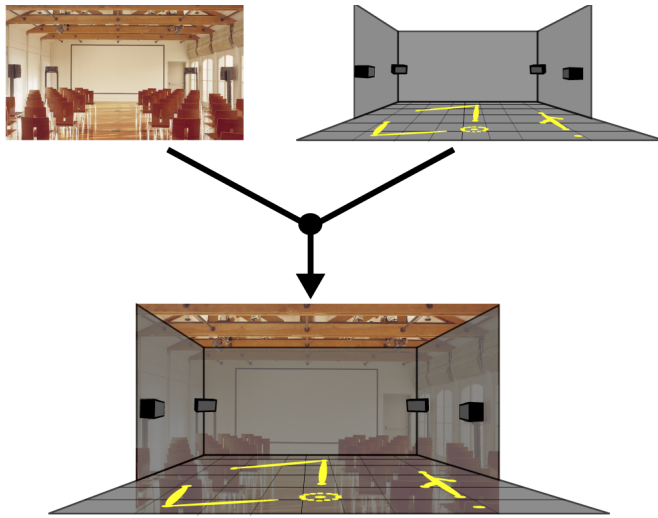
**Fig. 5.** Perceptual super-imposition of a virtual game-space on top of a physical performance space

a key role in an audience's ability to become immersed in the projected audio-visual landscape. After presenting different visual displays to an audience during performance - either an in-game global vantage point or a performer's standard first-person view - early usage indicates that perhaps the perceptual illusion of super-imposed spaces succeeds or breaks down based at least in part on the perceived alignment of the virtual viewpoint with the audience's physical viewpoint. When no visual environment is projected to an audience, the resulting musical experience changes dramatically, offering less sensory confusion and a seemingly greater ability for audience members to focus on the musical work at hand. However, without visual cues to help define the virtual environment, the musical soundscape lacks the clear indications of performers' deterministic vs. improvised gestures.

While working to better understand this idea of perceptual super-imposition, a realistic slightly larger-than-scale virtual model of CCRMA's new heptagonal 16-channel Listening Room was created, complete with representative virtual speaker placements (see Figure 6). Game players sitting in the actual Listening Room could enter the model and move their avatars through the virtual room, controlling sound across an 8-channel sound field mirroring corresponding real-world physical speaker locations. For individuals controlling the system as well as individual watching on a projected video screen, effects ranging from sensory confusion to mild discomfort were reported. Though not intended as a full-blown perceptual study, the immediate reactions of those involved hint at deeper issues with users' cognitive abilities to separate virtual environments from physical environments.
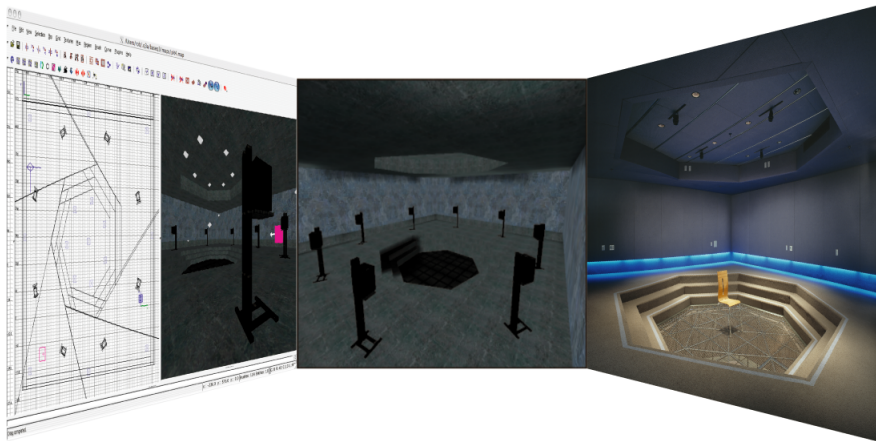
**Fig. 6.** Three representations of CCRMA's heptagonal 16-channel Listening Room (from Left to Right): GtkRadiant schematic view, in-game Quake III capture and real-world photograph

# References

1. Cage, J.: Music of Changes (Score). Henmar Press, C.F. Peters, New York (1951)
2. Cage, J.: String Quartet in Four Parts (Score). Henmar Press, C.F. Peters, New York (1950)
3. Freeman, J., et al.: The Architecture of Auracle: A Voice-Controlled, Networked Sound Instrument. In: Proceedings of the International Computer Music Conference, Barcelona, Spain (2005)
4. Furukawa, K., Fujihata, M., Muench, W.:
   `http://hosting.zkm.de/wmuench/small_fish`
5. GtkRadiant, `http://www.q3radiant.com`
6. Oliver, J., Pickles, S.: q3apd (2007) (as viewed 4/2007),
   `http://www.selectparks.et/archive/q3apd.htm`
7. Pickles, S.: fijuu2. Henmar Press, C.F. Peters, New York (1950) (as viewed 4/2007),
   `http://www.fijuu.com`
8. Pulkki, V.: Virtual sound source positioning using vector base amplitude panning. Journal of the Audio Engineering Society 45(6), 456–466 (1997)
9. Stockhausen, K.: Zyklus. Universal Edition, London (1960)
10. Stockhausen, K.: Klavierstück XI, Universal Edition, London (1957)
11. Wang, G., Misra, A., Davidson, P., Cook, P.: Co-Audicle: A Collaborative Audio Programming Space. In: Proceedings of the International Computer Music Conference, Barcelona, Spain (2005)