Postgrad Seminar

Electronics in Performance:
Techniques for Creating 3D Sound Images

(Very) **Basic Techniques**

Amplitude (To create ILD - Interaural Level Differences)
Phase (To create ITD - Interaural Time Differences)

[There are other spatial cues created by shoulder/pinna reflections]
See Hyunkook Lee's Work.

A Review of Amplitude Based Panners

point source
- simple but it works **really** well

vbap (vector-based panning)
- miminal speaker usage (2 in 2D and 3 in 3D and hence optimum localisation)
- limited to a sphere/dome. All speakers must be equidistant (from a single point)

dbap (distance-based panning)
- allows speakers to be anywhere
- amplitude coefficients are based on distance
- there is a compromise between localisation and blending

mic modelling (ViMiC)
- models virtual mics in a space (imagine playback in an anechoic space with mics placed corresponding to speakers in actual playback)

MIAP (manifold interface amplitude panning)
- allows one set of speakers to be address in a number of ways
- allows grouping (or non-grouping of speakers)
- panning values are based on the position of the sound in a **virtual space** that does not have to correspond with real space
- uses a Barycentric triangular panning algorithm

(N.B. a system like Dolby Atmos is just amplitude based panning plus decorrelation)

Other Technologies

Delay-based Panning
- (need an algorithm to determine appropriate delays to single point)

Ambisonics
- based on pressure vector encoding (simple to do)
- decoder uses hybrid amplitude/phase (and is not simple to do)
- encoded audio format is

WFS (Wavefield Synthesis)
- recreates wavefields to allow point-source location not at a speaker location
- needs v. high density speaker array

Challenges of 3D / Spatial Audio

[Pessimism]

- The *Sweet Spot* problem
- System heterogeneity
- Lack of detailed understanding of height perception (Hyunkook Lee's current project)
- Requirement of working in expensive / time-limited spaces
- Choosing technologies to use to perform the actual panning
    - How do these plug into your workflow?
    - Technique proliferation - why does it matter what you use? / which implementation?
- Tie-in between representation and technologies
- Methods of *control* and *graphical representation* (especially 3D viewed in 2D)

Positives and Solutions in 3D / Spatial Audio

[Optimism]

- Localisation gets easier the more speakers you have
  - Reduced sweet spot issue in terms of accuracy
- A chance to reconsider panning on a more fundamental level
- Binaural rendering offers one way to work on headphones
  - Use relevant HRIRs to render each speaker position
  - **BUT** - additional rendering / CPU / technical complexity

What are Our Aims? / 3D Images

Localisation (Point Sources)

Sound Groups?

Diffuse Images

Complex Images (multiple layers)

Something Else?

Issues with (Complete) Localisation Priority

(Moving towards my interest in complex rich images)

- How does this translate to the audience (system/sweet spot etc.)?
- How do sounds have width (real sounding object are not point sources)?
- How do we realise seemingly simple concepts (e.g. sweeping through audience speakers in rows) in a localisation based model and representation?
- Does accuracy of localisation have *musical* meaning

The Premise

A residency to collaboratively develop creative sound processing tools for generating texturally and spatially rich, immersive materials for 3D sound environments.

A large body of work in the area of 3D sound exists concentrating on possibilities of spatial trajectories and accurate localisation. This residency would focus on a different approach, in which multiple components strands, derived from recorded source materials are spatialised to make a complex sonic environment, prioritising spatial diffuseness and complexity over accurate localisation. Granular and spectral DSP techniques are obvious starting points for exploration in this area, with the aim being to present users with creatively meaningful parameters for controlling rich sonic and spatial behaviours of the processing. Such control could be offered through physical models or content aware DSP that can simplify the generation of parameters for each strand.

The residency would draw on technical expertise at NOTAM, as well as my programming skills, but focus on the creative and aesthetic potential of such approaches, with a set of short study examples/ pieces forming a focus for addressing artistic questions and effectively supporting future larger compositional work.

How Does This Translate into Reality?

Basic Research

- Explore the effects of different panners

Principles of Approach

- Inputs are mono
- Complexity/richness results from sending different signals to different speakers
- The number and positioning of speakers is a rendering concern only
- The exact panning method is not a primary concern
- Representations of space should be simple - details are for computers

Specific Proposals

Decorrelation Algorithms

Content Aware Spatial Granular Synthesis

Demos and Discussion