Real Time Binaural Audio Spatialization Using MAX/MSP

BSc Thesis Presentation

By Nikoloz Kapanadze

Motivation

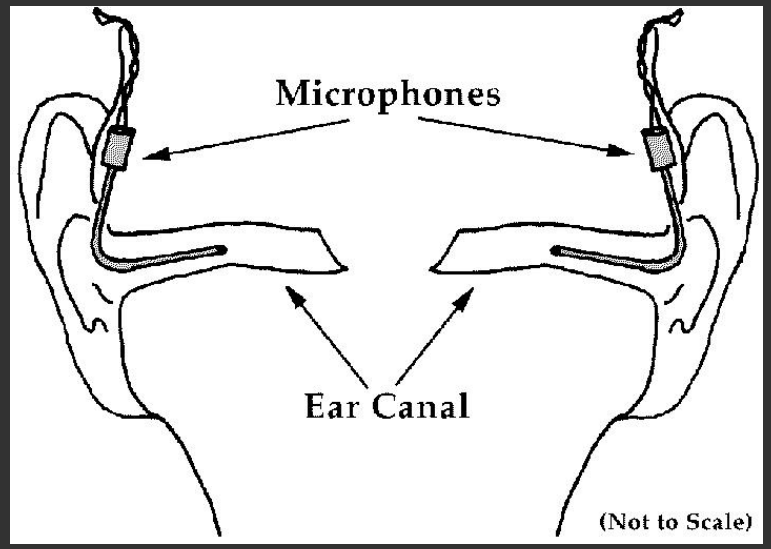
- Stereo Hearing allows us to pinpoint the sound sources
- Magnitude based panning methods do not preserve spatial cues
- Binaural Spatialization usually used as a post processing technique
- Can be used to increase immersion levels in applications such as VR, Telepresence
- Audio based orientation for the visually impaired

State of the art

- Modern computers have the capacity to perform complex DSP operations
- Several DAWs have static spatial panners built in
- High end audio-cards use HRTFs to optimize audio for 3D applications
- No good solution for adapting the audio environment to changes in listeners orientation in real time

A bit of theory

- Binaural recording can capture a particular auditory scene
- Binaural processing can map sounds to arbitrary locations in space
- HRTF – Head Related Transfer Function



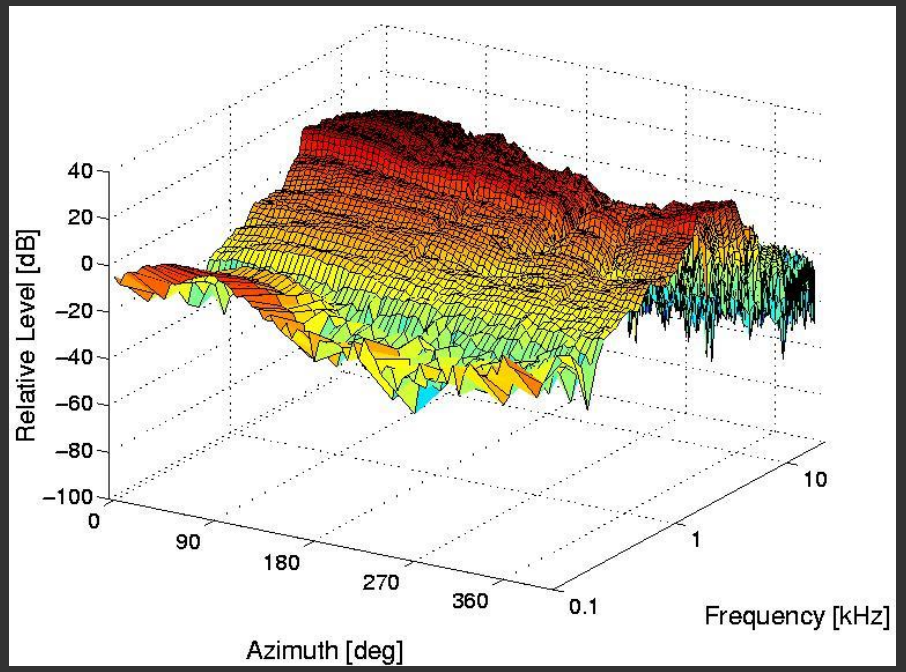
Binaural recording/HRTF measurement setup

Binaural Processing

- Uses HRTFs (Head Related Transfer Functions) to model the listeners had.
- Filters audio based on the model
- Accounts for the low pass filtering a human head introduces between left and right ears
- Accounts for IDT(Interaural Delay Time) - Difference in ToA od signals

CIPIC HRTF Database

- Measurements performed by the Department of Electrical and Computer Engineering at the UCDavis
- Public domain database
- Defined for 25 azimuth points, 50 elevation points and 200 samples
- IDT values for the 25*50 points



HRTFs plotted against azimuth elevation and time

The CIPIC HRTF Database

Tech&Hardware

Max/MSP 7 + Max SDK 6.0.4 - Modular graphical programming language.

Arduino UNO R3 – Atmega 328p based AVR microprocessor dev board

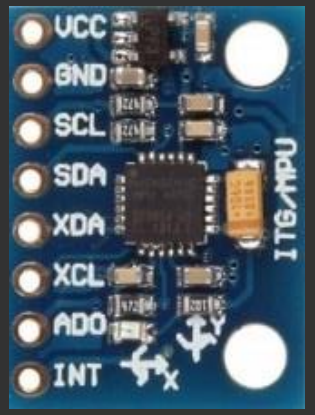
Invensense MPU 6050 – 6DoF Accelerometer/Gyro breakout board

Max / MSP

- PureData(PD) based GUI programming language
- Modular design: Easily extensible through a thoroughly documented SDK for C programming language. Compiles into shared libraries
- Good wrapper for audio I/O and driver communications
- Allows changing the patcher while audio is running
- Multiple DSP chains can run in parallel, takes care of concurency

Invensense MPU 6050

- Combination of an accelerometer and a gyroscope
- Uses the common I2C serial protocol
- Can be extended with a magnetometer over slave I2C pins
- Easy to interface
- 6DoF for low drift over time



Arduino

- Good alternative for dedicated I2C adapter
- Easy to program
- Good support for a wide range of peripheral
- Good libraries for working with I2C and MPU line of motion sensors in particular
- Available through Embedded Systems lab

Implementation

Can be split into two sections:

- DSP - Audio Processing
- Head-tracking – Generates control signals for the DSP objects.

Convolver

- MSP family of Max Externals
- The main method of the object is called within the callback of an audio driver
- Picks appropriate HRIR based on control input
- Does fixed size convolution
- Overlap add to merge consecutive buffers

IDT

- Shares the common structure with Convolver object
- Simple two channel delay line
- Picks delay time based on control input

Serial communication

Arduino(with custom firmware) + MPU 6050 for orientation sensing

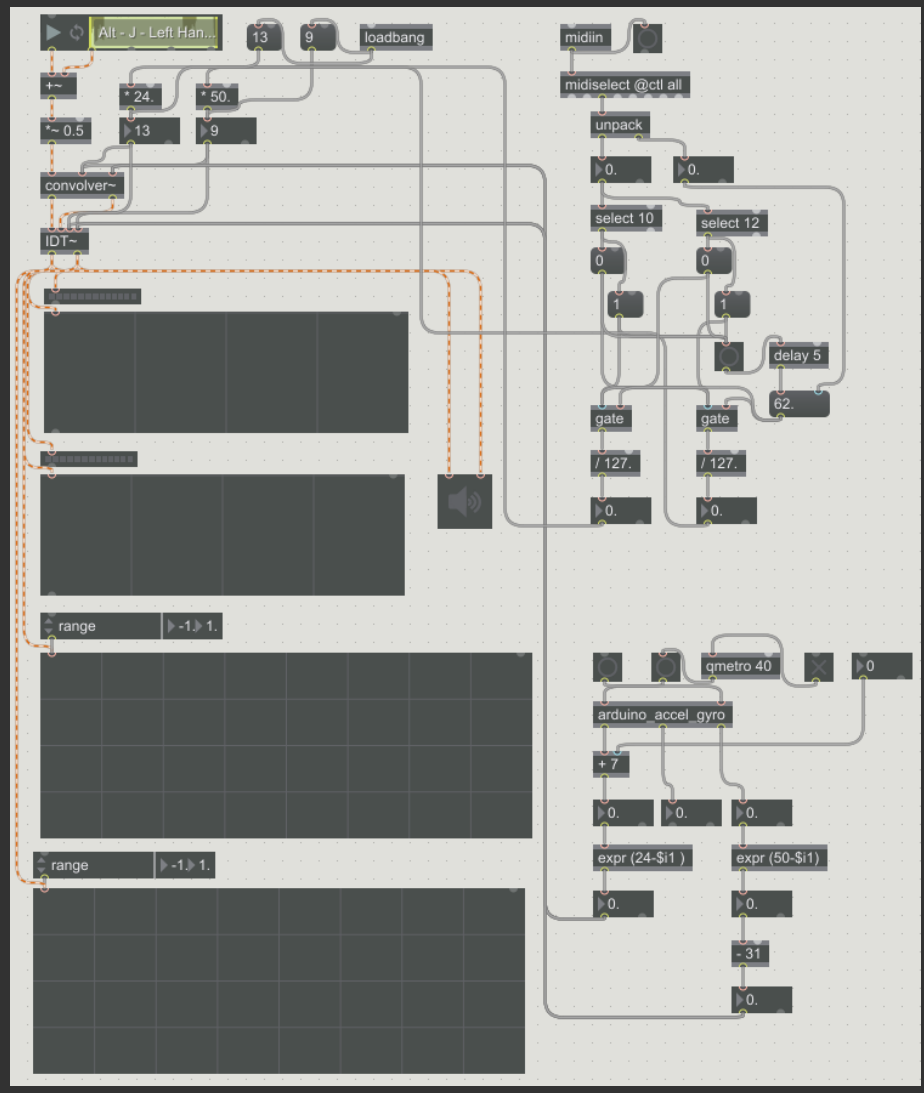
Talk to PC over virtual COM port

Generates output in the form of quaternion

arduino_accel_gyro Max external polls Arduino for updated vales at regular specifiable intervals

Patcher

- The sections discussed above come together in the patcher
- Spectroscope and Osciloscope for audio monitoring
- Takes input from an audio file
- Can be reconfigured to accept audio from elementary signal generators



Testing and Evaluation

Precision of head-tracking:

- Average Deviation - 3.7°
- Maximum Deviation - 10°

Drift Accumulation:

- 5° Over 15 minutes of “intended” use

Precision of localization:

- Average Deviation - 4.6°
- Maximum Deviation - 10°

Q&A

Questions?