Implementation and Integration of Sound Spatialization

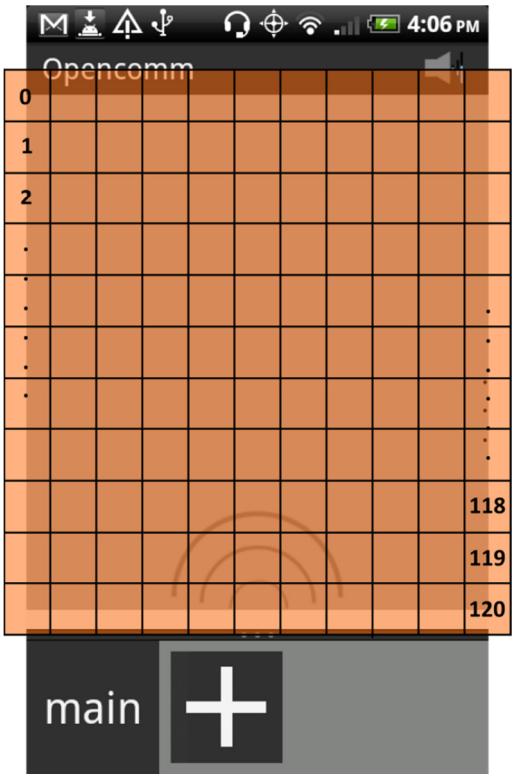
Risa Naka (rn96) modified March 26, 2012

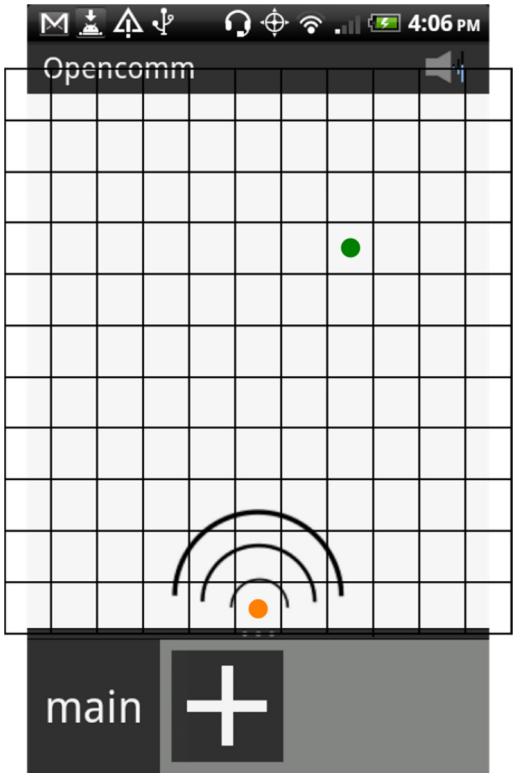
Regions

The conference space is divided up into 121 different regions, identified as Region 0 – 120. The region covers 110% of the available conference space.

The center point of each user icon is used to determine an icon's region.

conference space dimension: width: 480px, height: 537px
region dimension: width: 528, height: 590.7px





Point of Reference for Calculations

The interaural time delay (ITD) and volume difference for each region is calculated based on the region's center point (green for Region 80) with the primary user's location (orange) as the bottom center edge of the conference space.

The primary user's location is calculated as the origin (0, 0). In the case of a user located in Region 80, the point used to calculate the spatialization variables is (96, 376).

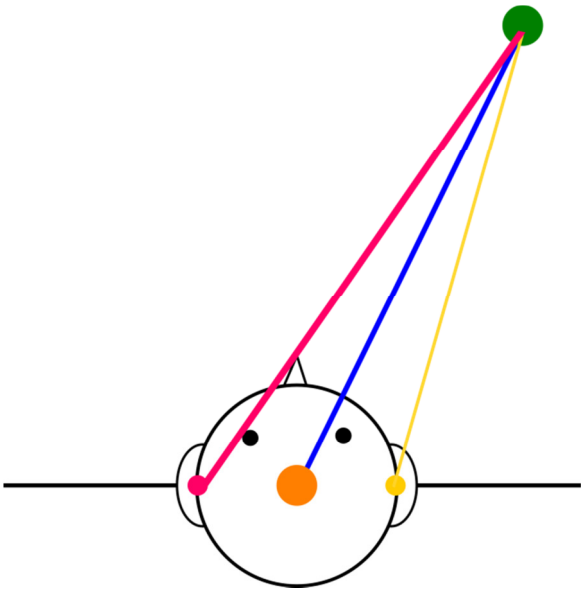
Implementation and Integration of Sound Spatialization

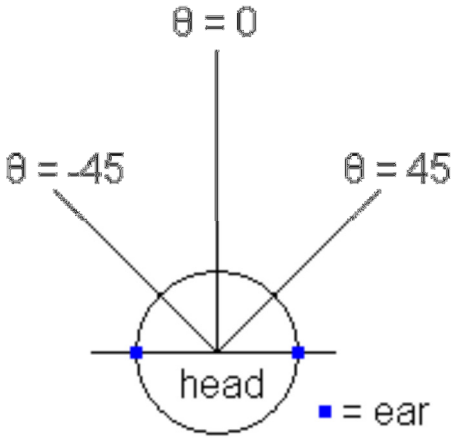
Risa Naka (rn96) modified March 26, 2012

Concept of ITD and Volume Difference

Interaural time difference is the difference in time for the sound to reach the left ear as opposed to the ear.

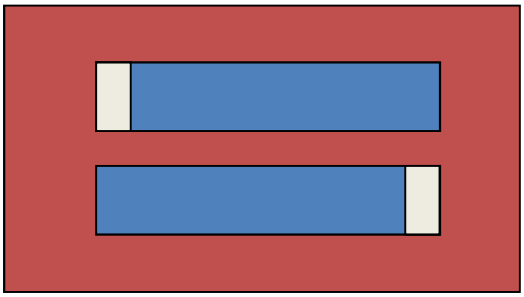
Volume difference is the difference in volume between the left and the right ear.





Implementation of ITD

1. Read in short array of sound
2. Make copies for left and right stereo channel
3. Add ITD (short) either front or back of the sound source, depending on whether the sound source is coming from the left or the right
4. Combine them to create a short array formatted for stereo output



Calculation of ITD and Volume Difference: Example for Region 80

$coordinate_{primaryUser} = (0, 0)$

$coordinate_{userReg80} = (96, 376)$

$radius_{head} = 0.085m$

$speed_{sound} = 343.42m$

Implementation and Integration of Sound Spatialization

Risa Naka (rn96) modified March 26, 2012

Input sound: 8kHz 16-bit linear PCM, mono channel (8 shorts/ms)

Output sound: 8kHz 16-bit linear PCM, stereo channel (8 shorts/ms)

ITD in bytes

$$\theta(radian) = \frac{\pi}{2} - \cos^{-1}\left(\frac{x_{userReg80}}{\sqrt{\left(x_{userReg80}\right)^2 + \left(y_{userReg80}\right)^2}}\right) = 0.249979$$

$$ITD(s) = \frac{radius_{head}}{speed_{sound}} \times (\theta + \sin\theta) = 0.000123102\ seconds$$

$$ITD(short) = \left\lfloor \frac{radius_{head}}{speed_{sound}} \times (\theta + \sin\theta) \times \frac{1000ms}{1s} \times \frac{8\ shorts}{1ms} \right\rfloor = 0\ shorts$$

Stereo Volume (0 ≤ vol ≤ 1)

$$\Delta_{distance} = ITD(s) \times speed_{sound} = 0.042275759m$$

$$distance_{left} = \left(\sqrt{\left(x_{userReg80}\right)^2 + \left(y_{userReg80}\right)^2}\right) \times \frac{0.005m}{1px} + \frac{\Delta_{distance}}{2} = 1.961447m$$

$$distance_{right} = \left(\sqrt{\left(x_{userReg80}\right)^2 + \left(y_{userReg80}\right)^2}\right) \times \frac{0.005m}{1px} - \frac{\Delta_{distance}}{2} = 1.919171m$$

$$vol_{left} = 1 - \frac{distance_{left}}{3m^*} = 0.3462$$

$$vol_{right} = 1 - \frac{distance_{right}}{3m^*} = 0.3603$$

** maximum distance: <3mt, assuming 1px = 0.005m*