

SEEING USING SOUND - DESIGN OVERVIEW*

Richard Hall
Jared Flatow

This work is produced by OpenStax-CNX and licensed under the
Creative Commons Attribution License 2.0[†]

1 Input Filtering

The first step in our process is to filter the input image. This process helps solve the "busy" sound problem from the vOICE[‡]. We decided to first smooth the image with a low pass filter, leaving only the most prominent features of the image behind. We then wanted to filter the result with an edge detector, essentially a high pass filter of some sort. We chose to use a Canny filter for the edge detection. The advantage of using an edge detector lies in simplifying the image while at the same time highlighting the most structurally significant components of an image. This is especially applicable to using the system for the blind, as the structural features of the image are the most important to find your way around a room.

2 The Mapping Process

Simply put, the mapping process is the actual transformation between visual information and sound. This block takes the data from the filtered input, and produces a sequence of notes representing the image. The process of mapping images to sound is a matter of interpretation, there is no known "optimal" solution to the mapping for the human brain. Thus, we simply chose an interpretation that made sense to us.

First of all, it seemed clear to us that the most intuitive use of frequency would be to correlate it to the relative vertical position of an edge in the picture. That is, higher frequencies should correspond to edges that are higher in the image than lower frequencies. The only other idea that we wanted to stick to was making the center the focus of the attention. For a complete description of this component, see the mapping process.

*Version 1.1: Dec 15, 2005 12:03 am -0600

[†]<http://creativecommons.org/licenses/by/2.0/>

[‡]<http://www.visualprosthesis.com/javoice.htm>