Data Sonification - Report

Hanley Weng
309248434

Goal

To provide a new perspective of our physical surroundings via the conversion of visual data into sound. Presented as an interactive system.

Description

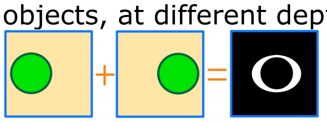
Conceptual Explanation

Overview

The developed sonification relies on data mined from two mini cameras (or a single camera oscillating between left and right), attached to an efficiently portable system (like a pair of sunglasses, a system similar to vOICe[1]). Binaural sound is produced in real time with sound outputted through earphones attached to the portable system. The system also relies lightly on the user's ability to learn and derive the meanings of sets of sounds as they continue to use the system whilst interacting with the environment.

Data Mining

With the two camera setup, a single video feed is produced, one that alternates between the two cameras frame by frame. This (or single camera feed) is compared with itself for any pixel changes every two frames. Differences allow for different objects, at different depths and also any movement within the camera's field of vision to be distinguished. 'Comparison images'generated from this are blended together over time to compensate for slow computer processors and to smooth out erratic movements.



[Dual-Camera: Camera #1 + Camera #2 = Compared Image]

[Moving-Camera: Camera 1ms + Camera 2ms = Compared Image]

Data Sonification

Each pixel in 'Comparison Image' generates their own sound as they are traversed. The Cartesian locations (x,y) of the pixel dictates the localization of the sound. Both the y-co-ordinate and color of the pixel (from the pre-compared image), categorizes sounds, uniting and diversifying them via timbre and pitch. The significance of a pixel is represented by its brightness in the compared image, with groups of sounds with higher significance becoming louder and rhythmically faster than dimmer pixels.

Patch Explanation

In researching methods of mining appropriate data, Andrew Benson's Jitter Recipes [2] and Jitter and Max Tutorials within MaxMSP5 were explored. Methods regarding the sonifcation of data utilized methods found in Max and MSP Tutorials, provided by MaxMSP5 [3].

[Significant actions are listed on the next page in order of appearance]:

Data Mining is handled via Jitter.

1. Input (main) video; combination of the two mini-cameras or single moving camera.
2. "jit.rgb2luma" - gray-scales the video.
3. "jit.op" - finds and contrasts differences in pixels per frame.
4. "jit.slide" - smooths out movements and lingers the image for slow processors.
   Data mining produces a 'comparison video' for the sonification to use.

Data Traversal is handled mainly via Max.

5. "metro, counter" - used to traverse pixels in 'comparison images'.
6. "jit.peek~" [jitter] - outputs value of a pixel at a given location (x,y)
7. "value __" - stores/access' values (pixel (compared & real) brightness, x, y, etc)

Data Sonification is handled mainly via MSP.

8. 'Pixel values' are fed into a subpatch, 'mixer'.
9. 'Mixer' routes the pixel sound into an un-used sound slot via a counter.
10. These multiple soundslots are subpatches, called 'pixelSound'.
11. 'PixelSound' outputs a sound to 'mixer' based off 'pixel values'.
    - (sound based off 'Data Sonification' in 'Conceptual Explanation' above.)
12. 'Mixer' multiplies all the 'pixelSound's together, outputting it to main patch.

Results

Interest in this perception of the environment was witnessed. The responsiveness of the system is dependent on the power of the processor it is run off. Moving and static objects, their relative depths and positions are determinable to an approximate degree, especially at far distances. Natural association and user responsiveness to the interactive system was achieved, with users interacting with the system without the aid of instructions. Attempting to navigate with the system returned adequate results within a familiarized environment.

Evaluation

The sonification satisfies its goal. It provides an interesting audio perception of a physical space which we are accustomed to visualizing as opposed to sonifying, even allowing for simple navigation after familiarization with the system and environment.

Improvements

Like all systems, nothing is perfect, and can be continually improved. As computer chips become faster and more portable with time (Moore's law [4]), the software that is part of this sonification can become much more efficient and accurate. With more processing power, more pixels can be analyzed, accuracy improved, better localization of sound can be obtained in 3 dimensions (x,y,z) and implementation of further equations for distinguishing objects, such as pixel-clusters, can be executed. With higher processing power, more sounds can also be effectively mixed together in real-time.

In conclusion, though the system can be continuously improved, the state of the current sonification satisfactorily accomplishes its goal, providing an interestingly new perspective of our physical environment.

References

[1] vOICe, by Peter B.L. Meijer, at seeingWithSound.com

[2] A. Benson, Jitter Recipes: Books 1,2,3, cycling74.com/section/tutorials

[3] Jitter, Max, MSP Tutorials, from the program MaxMSP, Version 5.0.6, Cycling '74

[4] Moore, Gordan E. "Cramming more components onto integrated circuits", Electronic Magazine 1965