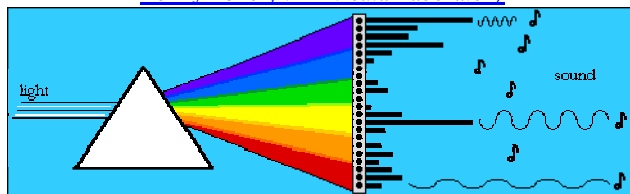


Artificial Synesthesia via Sonification: A Wearable Augmented Sensory System

Lenny Foner, MIT Media Laboratory



I am implementing a wearable artificial sensory system, which uses data sonification to compensate for normal limitations in the human visual system. The system gives insight into the complete visible-light spectra from objects being seen by the user. Long-term wear and consequent training might lead to identification of various visually-indistinguishable materials based on the sounds of their spectra.

Normal humans are trichomats (Cornsweet, 1970), meaning that almost any three wavelengths of light, if their amplitudes are adjusted correctly, can be made to look like some particular color. (For example, this is the principle behind the red, green, and blue pigments in a CRT, and is the reason why people who lack one or more of the three color-vision systems are called "colorblind".) The system described here instead acts like a spectrometer, and, unlike the human eye, cannot be fooled by such a small number of wavelengths. Instead, it images in 128 wavelengths in the visible spectrum -- and, optionally, another 128 wavelengths in the near-infrared -- then uses sonification to make the resulting histogram of wavelength vs amplitude accessible to its user. It uses sonification, rather than a visual display, to keep the user's visual field uncluttered, and to enable a sort of artificial synesthesia.

The system is wearable. It sits on the side of the user's head, and images a patch of his or her environment about two degrees wide

(the same width as the fovea), aimed in the direction that the user's head is pointed. (Full eye-tracking is too intrusive and too cumbersome to justify (Young & Sheena, 1975).) This makes it possible to use the system most of the time, and to learn the mapping between materials commonly encountered and their sounds. This in turn may allow the user to identify what things are made of, or whether objects have undergone a change. (For example, imagine looking at a car and saying, "Well, it looks like metal, but sounds like painted plastic," or looking at one's lawn and saying, "Hey, the grass sounds funny today -- is it sick?") Future extensions of the system to a wider electromagnetic bandwidth (e.g., far-infrared and near-ultraviolet) promise to improve its utility substantially. Designs that incorpor

At the moment, the sonification strategy is quite simple: each visual wavelength imaged controls a corresponding audio wavelength. The resulting collection of sine waves at different frequencies and amplitudes is then simply summed and presented to the user. It is clear from the psychoacoustics literature (Handel, 1989; Wenzel, 1994), however, that there are many better strategies for presenting the data, and some of these methods will be incorporated into future designs. For example, incorporating timbre into the individual frequencies, and slightly staggering their relative attacks, should make it easier to differentiate various spectra. In addition, because people can often remember where a given sound came from better than they can remember the individual pitches that composed it, a more ambitious design, using a more powerful digital signal processor than is present in the current system, would use stereo output channels and would spatialize th

The system is undergoing continuous improvement and redesign. Sonification is a critical part of its construction, and experiments are ongoing in determining the best way to represent the sensory information.

[More information on this system is available.](#)

References

Cornsweet, Tom. (1970). *Visual perception*. New York: Harcourt

Brace Jovanovich.

Handel, Stephen. (1989). [*Listening*](#). Cambridge: MIT Press.

Wenzel, Elizabeth. (1994). Spatial sounds and sonification. In Gregory Kramer, (ed.), [*Auditory display: sonification, audification, and auditory interfaces*](#). Reading, MA: Addison-Wesley, 1994, 127-150.

Winston, Mark. (1987). [*The biology of the honey bee*](#). Boston: Harvard University Press.

Young, Laurence, & Sheena, David. (1975). Survey of eye movement recording methods. *Behavior research methods & instrumentation*, 7(5), 397-429.



Author Information

[Lenny Foner](#)

E15-305

MIT Media Lab

20 Ames St

Cambridge, MA 02139

617/253-9601 voice

617/253-6215 fax