Multimedia Output Devices and Techniques

COLIN WARE

University of New Brunswick, Fredericton, New Brunswick, Canada ⟨*cware@UNB.ca*⟩

This brief review of the state of the art in output devices is written from a perspective of two basic questions. How closely do current devices come to reaching the limits of the human capacity to receive information? What are the prospects for improvement? This overview is divided into sections dealing with visual, auditory and touch displays, followed by VR displays, which integrate all three types of system. For a much more detailed review, the best single reference is Barfield et al. [1995].

VISUAL DISPLAYS: SPATIAL AND TEMPORAL

Humans get perhaps 70% of all sensory input from vision. The human eye has an enormous dynamic range. The amount of light reflected from surfaces on a bright day at the beach is about five orders of magnitude greater than the amount available under dim lamp lights. By way of contrast, most computer displays have a dynamic range of a little more than two orders of magnitude, but it does not matter because the eye is quite insensitive to the overall level of the illumination. This is an important point: displays do not synthesize reality, they fool the brain. We are approaching the ideal of having screens with pixels as fine as the human eye can resolve (at least with proper anti-aliasing). The receptors in the human eye have a visual angle of about 0.8 minute of arc. High-resolution screens are commonly available that provide 1200 pixels in a 30-cm screen—about 40 pixels/cm; at normal viewing distances these pixels are about twice the size of receptors when imaged on the retina. However, there are aliasing effects that mean that image errors of less than receptor size can be seen.

Temporal resolution is as important as spatial resolution in designing a display. A light that flickers at more than 50 Hz is generally perceived as steady. Higher update rates are needed to cope with images that are rapidly changing.

In general, the prospects for better spatial and temporal resolution are good both because of improvements in the display and because of better spatial and temporal anti-aliasing techniques. We can expect to see an increase in the number of pixels available on the screen up to and beyond the 1080 × 1920 in the draft high-definition TV standard. At present screens that refresh at more than 70 Hz are rare, but this will change as the display of moving images becomes more common.

STEREOSCOPIC DISPLAYS

In a stereoscopic display slightly different images are presented to the two eyes to create a greater illusion of 3D space. The most widely used method for achieving this is the frame-sequential method in which shutters are used to ensure that the two eyes receive alternative frames of the video image. In one widely used technique, shutters are incorporated in glasses that are synchronized with the monitor via an infrared link. In another, the polarization of a screen covering the monitor can be changed for alternate frames; in this case the user wears Polaroid glasses, differently polarized for each eye. Circular polarization is currently used since this is not affected by head orientation. Either method works well if the graphics system can provide at least one hundred updates per second, fifty to each eye.

The most challenging problem relating to stereo displays is the difficulty of simulating depth of focus. This may make stereo displays both fatiguing to look at and difficult to fuse because the brain expects objects at different depths to require refocusing of the eye. There is no short-term solution to this problem.

FORCE AND HAPTIC DISPLAYS

A force display is one in which touch information is synthesized by producing forces on the skin of the operator. However, the touch sensation is extraordinarily complex, involving sensitivity to small shear forces in the skin as well as pressure sensors in the skin and in the joints. The technical problems in generating even simple touch sensations are formidable, especially on the large scale demanded by VR. Devices must be stiff in order to be able to create the sensation of solid contact and light so that they have little inertia, and there must be a tight loop between input (position) and output (force); 5 Khz may be necessary for optimal fine motor control. Nevertheless, it has been shown that even very limited force feedback improves performance in certain telerobotic applications when, for example, inserting a peg into a hole. There is little doubt that the use of force output has clear potential for applications such as arthroscopic surgery and telerobotics. There are good prospects that small, low-cost force output devices will become available, useful for example in CAD systems so that users can feel contacts between components as they are assembled.

AUDITORY DISPLAYS

In terms of sheer bandwidth the problem of auditory displays have been solved; inexpensive systems exceed the temporal resolution and dynamic range of the human ear. Sound can also be generated that can be localized in space almost as well as sounds in the environment. Of course, speech in everyday life is a tool in dialog: hearing and speaking are closely coupled. But, since machines have as yet very limited intelligence, this dialog is very asymmetric in current systems and this may be why many people turn sound features off. A major use of speech output is in telephone interfaces to computers and in multimedia and training systems. The outlook for auditory displays, both speech and non-speech, is very positive.

VIRTUAL REALITY DISPLAYS

The goal of virtual reality displays is to simulate environments so that we can understand and interact with synthetic scenes in ways we have already learned from everyday reality. The typical VR display has a helmet-mounted display coupled with a head-tracking system, giving an immersive experience. Input must be coupled to output in a tight loop because the purpose of the head-tracking system is to estimate the user's eye position and to generate a perspective image of a virtual scene from that eye position. Ideally, the tracking system should be low-lag and high-precision. One or more glove input devices may be used, although the lack of touch sensation has been a major drawback to the utility of these passive input devices. The major advantages of helmet-mounted displays are said to be twofold: a sense of presence and ease of learning. The idea is that everyday life skills can be transferred directly without new learning. In practice many techniques must be learned in VR; it is generally undesirable to do things in the same way as in everyday life, so that even a perfect simulation of reality would have limited utility.

Problems of resolution, optical distortions, and position tracking can all be solved. Problems involved in creating localized sounds largely have been solved. Unfortunately, there does not appear to be a physical method for creating virtual touchable objects with any fidelity in a large working volume. Thus we can create VR worlds that can be seen and heard but not touched.

Many flavors of VR are emerging, including fishtank VR, augmented reality and highly interactive 3D video games that many people have also begun to call VR. It is clear that highly animated 3D environments will become increasingly common. It is less clear that immersion VR will be a common mode of viewing.

A final comment: there are many areas in which the distinction between output devices and input devices is becoming blurred. We are used to regarding a screen as a passive output device and a mouse as an input device with no tactile feedback. However, the trend is towards making every object on the screen into a "widget" for some kind of interaction. Hence the division of the computer interface into input and output is an anachronism.

REFERENCES

BARFIELD, W., HENDRIX, C., BJORNESETHE, O., KACZMAREK, K. A., AND LOTENS, W. 1995. Comparison of human sensory capabilities with technical specifications of virtual environment equipment. *Presence: Teleoperators and Virtual Environments 4,*4, 329–356.

BOFF, K. R., KAUFMAN, L., AND THOMAS, J. P. (EDS). 1986. *Handbook of Perception and Human Performance, Vol I, Sensory Processes and Perception*. Wiley, New York.

LOOMIS, J. M. AND LEDERMAN, S. J. 1986. Tactual perception. In *Handbook of Perception and Human Performance: Vol II, Cognitive Processes and Performance*. K. R. Boff, L. Kaufman and J. P. Thompson, Wiley, New York.

WYSZECKI, G., AND STILES, W. S. 1982. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley, New York.