

# **THE ELECTRONIC MUSICIAN**

## **Into New Worlds**

### **Virtual Reality and the Electronic Musician**

**Scott Fisher**

You're ready to enter an alternate reality. You slip on a pair of headphones, a helmet with a tiny video screen for each eye, and a special glove. Three-dimensional computer generated images are projected onto your eye screens. Video "hands" match the movements of your own limbs. Sound seems to be coming from all around, not just inside your head, as you'd normally expect with phones.

There's a mixing console floating before you. The music is coming from an apparent distance of a couple yards, but when a channel is "soloed", that sound moves to a point inches from your ear, while the rest of the mix stays put. By grabbing the edge of the board and pulling, you get as many input channels as you need stretching away to infinity.

This isn't as far-fetched as you might think. In fact this technology exists today at NASA, Ames Research Center and other labs. The concept, called "virtual reality" (VR for short), gives us a new way to think about the use of electronic and computer systems.

Virtual reality takes you places you've never been and lets you interact with your surroundings in ways not possible in the real world. In these new spaces, physical laws can be modified or ignored. The computer/user interface, until now bound by keyboards, mice, and video display terminals, makes the jump from our desktops to inside our heads, while our bodies begin to enter our machines.

#### **The Challenges**

Virtual reality poses a real challenge to our imagination. Cast into unfamiliar territory, there is a danger that we'll restrict ourselves to old ideas and needlessly limit ourselves. "The sooner we discard our old concepts and treat virtual reality as a new medium, the further we'll go with the idea," says Mark Bolas, president of Fake Space Labs, a consultant to the Virtual Environment Workstation (VIEW) project at NASA Ames. "VR can free us from old concepts," continues Bolas. "For instance, the reason we use knobs in the physical world is not because they're the best way for people to interact with equipment. It's the physical requirements of their function that dictates their form. With VR, the link between form and function can be severed."

No sudden breakthrough has made virtual reality possible. In fact, most of the components have been available for some time. So, before we get into applications, let's look at the parts that make up a typical VR system.

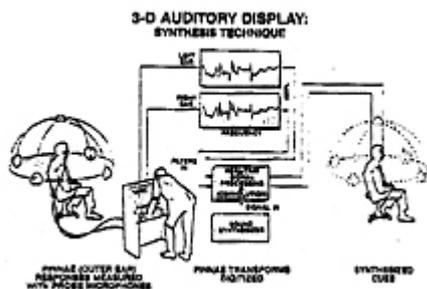
Sight is the sense most often stimulated by VR systems. To help the user feel like a part of the virtual environment, a helmet fitted with LCD video displays (one per eye) is worn. Three-dimensional video is created by showing a slightly different image in each eye. These head-mounted displays generate images that nearly encompass your entire field of view. The system incorporates a means of sensing and responding to the user's head position, so the stationary objects behave as they would in the real world. Objects in your view pan to the right when you turn your head to the left and vice versa.

If sight is viewed as input to the user, then human gesture could be considered the primary input to the VR system. A common gestural input device is the VPL DataGlove. VPL builds VR systems that incorporate the DataGlove and the EyePhone, a head-mounted video display. The glove is wired with fiber-optic cable that refracts light differently depending on whether

each finger is straight or bent. A magnetic sensor, called the Polhemus 3-Space tracking device, determines the location of the glove in space up to fifteen times per second. Technically speaking, six axes of movement can be determined by the DataGlove: X, Y, and Z position, as well as roll, tilt, and pan.

Gesture and the sense of touch are important, since music performance always involves gesture. The tactile feedback we get from real instruments is an important part of controlling these gestures but virtual instruments don't provide a real object to touch. Some prototype VR systems can simulate the sense of touch, a concept called "force feedback". Research is being done in this area, and tactile feedback is becoming a realistic goal for VR systems.

## Sound in Virtual Reality



The Convolvotron allows you to determine the three-dimensional location of a given sound within a virtual reality.

In the NASA-Ames virtual system, a device called the "Convolvotron" creates three-dimensional sound within a pair of normal stereo headphones. Up to four discrete audio channels can be individually placed and/or moved in an imaginary sphere surrounding the listener (see Figure). As with VR video displays, the perceived location of the sound remains constant regardless of head position. The Convolvotron is a two-board set that works with IBM PC's.

## Synthesizing audio in 3-D

A three-dimensional visual display, an ongoing project of the National Aeronautics and Space Administration (NASA), has hatched an audio counterpart — a signal processor that synthesizes three-dimensional sound over headphones. The device is under development at the NASA Ames Research Center in Moffett Field, Calif.

Hearing 3-D sound can be useful when visual information is absent or limited. For instance, pilots flying in bad weather could receive a 3-D signal from the control tower indicating whether a nearby airplane was above, below, or in any three-dimensional direction from them.

Begun about three years ago by NASA Ames researcher Elizabeth M. Wenzel, the project uses a signal processor to process sound, such as the analog output of a record player, so that a listener wearing headphones can locate the source, even one in a different room.

Human ability to locate sound is largely due to the difference in the time it takes for an acoustic wave to reach each ear. But the folds of each outer ear also affect the way the inner ear receives a signal by attenuating some frequencies and boosting others, or filtering it. The device synthesizes sound by taking into account these two effects. The technique assumes that if ear canal waveforms identical to those produced by a free-field, or natural, source can be duplicated, the free-field sound can be heard with head phones.

Psychophysical tests by two researchers, Frederic L. Wightman and Doris J. Klatier, at the University of Wisconsin's Waisman Center in Madison, found actual and synthesized listening experiences comparable for static sound sources varying in horizontal distance but not height. The researchers plan to extend the comparison to moving sources.

To synthesize sound, the Wisconsin researchers placed a probe microphone near each ear drum of a listener in an anechoic chamber. The ears' frequency response was mapped for acoustic waves emanating from 144 different spherical locations around the listener's head, at intervals of 15 degrees azimuth and 18 degrees elevation. The pairs of data points were then used to construct a map of listener-specific "location fitters".

The map was fed from an IBM AT into the dual-port memory of a real-time digital signal processor designed by Scott H. Foster of Crystal River Engineering Inc., Groveland, Calif. Known as the Convolvotron, the processor filters an analog signal with coefficients determined by the coordinates of the sound's desired virtual location and the position of the listener's head. The signal consequently is perceived by the listener as located in three-dimensional space.

The Convolvotron has 128 parallel processors and can process data in real time from up to four independent, simultaneous, and mobile sound sources. For mobile sources, coefficients are derived from a linear combination of the four nearest measured directions.

Work on this device began in 1986, when Scott Fisher, project leader for the VR VIEW system at NASA-Ames, asked perceptual psychologist Elizabeth Wenzel about the feasibility of adding 3-D sound to NASA's VR system. Dr. Wenzel decided that it was possible and enlisted the aid of Professor Fred Wightman (currently at the University of Wisconsin) and Scott Foster, president of Crystal River Engineering, to develop the system. Professor Wightman was known for his highly accurate measurements of the ear canal, while Scott Foster had the necessary background to design the hardware. Besides functioning as a 3-D sound source for VR use, the Convolvotron also was designed as an aid to psychoacoustical research.

Before jumping into details on how the Convolvotron works, you need to understand some basic psychoacoustic principles. We locate sounds in space by using small differences in time phase, and amplitude of the sound that reaches each eardrum. These differences are caused by several factors: the direction we are facing in relation to the sound source, the acoustic space surrounding the listener and source, and the shape of each person's outer and inner ear. The end result is that none of us hears things in quite the same way. Although differences in each person's inner and outer ear were long suspected to be significant, they were hard to quantify. By using Fred Wightman's precise measurements, the Convolvotron can account for them. To make the measurements, the user is seated in an anechoic (echo free) chamber, and a tiny probe mic is placed inside each ear canal, next to the eardrum. Then a test tone is played from 144 different locations surrounding the subject, and the "impulse response" at each eardrum is measured. The impulse response completely characterizes the direct and reflected sound reaching the eardrum. The sum of these measurements, called a "Head Related Transfer Function" (HRTF), contains the aural cues used to determine sound location. The HRTF of a specific user can be fed into the Convolvotron and used to synthesize 3-D sound.

The four sounds going into the Convolvotron are processed through one parallel array containing 128 multiply/accumulators that are configured as tapped delay lines. Each sound is "placed" in space by a Finite Impulse Response filter whose settings are determined by the HRTF measurements. When a sound is moved, it does not "snap" between measured points. Instead, the four nearest measured points are used to interpolate the response for the unmeasured points, allowing smooth movement of sounds.

Inside a virtual reality, the Convolvotron can make sounds seem to come from within an object. Also, localized (3-D) audio cues can be used to highlight information in a crowded visual field, such as an air traffic control display. Real-world sound can be processed, as can synthesized sound generated by the MIDI capabilities of NASA's Auditory Display System (more on NASA and MIDI later).

According to Scott Foster, the Convolvotron can simulate some aspects of room acoustics more accurately than conventional digital reverbs. Instead of using recirculation (feedback) to create reverb, the Convolvotron calculates every direct and reflected path that reaches the user's ears. One program supplied with the system is called "The Reflection Kit". With it, you can move several reflective surfaces (walls) while monitoring the resulting virtual room sound in realtime. There are some limits to the size of room that can be dynamically changed, but nearly any room can be simulated statically. The Convolvotron is capable of phase vocoding, pitch shifting and other effects, as well as 3-D sound manipulation.

Though it sounds futuristic, the Convolvotron is available today. A typical system costs around US\$ 25.000., not including the host computer or headtracking equipment. Crystal River is working on a new product incorporating many of the same features, that is expected to sell for under US\$ 10.000. Tying together all the video gear, sensors, and sound processing equipment are computer hardware and software (and some pretty thick cables). Highend workstations are capable of meeting the computational and graphic rendering demands of virtual reality, but hardware capable of generating shaded solid objects at 15 frames per second (one for each eye) will cost you plenty. Simpler "wire frame" drawings can be generated at sufficient speeds on a PC.

### **Virtual Reality Applications**

The space program was an early user of virtual reality, both for training simulators and as a way to efficiently display cockpit information. The number of controls astronauts had to monitor was growing at an alarming rate. By displaying a 'virtual panel' on a video screen, only the controls needed for the current operation were displayed, in an arrangement best suited for that task. This reduced the clutter of unrelated controls and, when needs changed, the virtual panel could be instantly reconfigured.

Electronic musicians are faced with a similar problem: Many instruments have hundreds of controls hidden behind a few buttons and a small, cryptic display. A virtual panel could get us back to the days of one function, one knob, and make synth programming a more intuitive task. Patch editor programs are an existing example of virtual panels, although most are not configurable. Newer "universal" patch editors are very close in concept to VR configurable displays.

We may be seeing the beginning of a trend towards panel-less equipment. For example, DSP cards for computers cannot be physically touched when installed. In this case, the virtual panel is the only choice and can easily customize a general purpose device to look like any one of the more specific tools we're used to working with, such as samplers and reverbs.

David Trubitt