

# Surround Sound for Video

An ambisonic visual postcard.

By David Moore

After several years of intense behind-the-scenes activity, ambisonic surround-sound recording techniques have passed from the experimental stages into real-world applications and popularization. Despite the regrettable confusion with earlier attempts at 4-channel sound, engineers and producers are beginning to recognize the operational and psycho-acoustical advantages inherent in working with B-Format and UHJ-encoded ambisonic material.

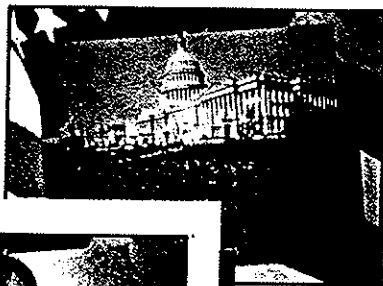
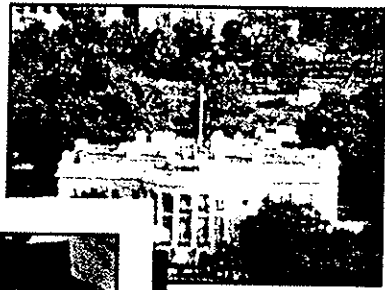
Recent advances in consumer technology, including Compact Disc and stereo television, have come a long way in eliminating the distribution problems presented by earlier multichannel attempts at *surround sound*. Many users consider that ambisonically encoded material may find its way into mainstream records and television far faster than one might think possible.

Stereo television has already put a new emphasis on the quality of sound that accompanies the visuals, and UHJ-encoded ambisonics could very well open up a new dimension for home entertainment.

Recently, a new genre of home video was released by a Washington, DC-based production company, Nicholas Communications. The 30-minute video was designed to take advantage of the realism offered by ambisonic production techniques. Producer Stephen J. Nicholas conceived the idea of a video and audio postcard—not just a simple travel log, heavy with narration and stock-look film footage, but rather a carefully assembled series of images and music working to-

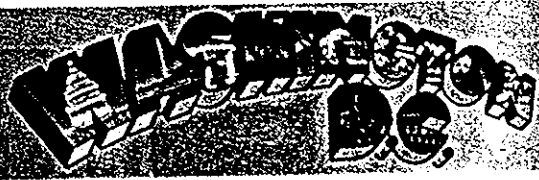
*Scenes from the visual postcard showing some of the more memorable sights of Washington, D.C.*

David Moore is a regular RE/P contributor, and director of product training at Sony Corporation's Fort Lauderdale, FL, facility.



  
**nicholas communications**  
presents

"WASHINGTON, D.C."—A thrilling state of the art visualized music adventure featuring the sights and sounds of Washington, D.C.



together to evoke memories of experiences.

The goal was to project a sense of immediacy, a sense of being there in the imagery. According to Nicholas, the project seemed a natural marriage between visual images and realism of ambisonic techniques.

"Since my background is in the visual arts as well as in audio production," the producer says, "it seemed the natural thing to do: synchronizing musical imagery—colors, textures, hues and moods—and visual images to produce the sights and sounds of some of the great cities on our planet. We're presenting this form of information/entertainment in a long-form music video that takes the viewer on a thrilling journey."

His thinking was as follows: A venue would be selected for broad appeal, perhaps one that was already known as a tourist attraction, and one full of compelling, easily identifiable images. The sights and sounds of this location would then be recorded in the appropriate formats, always paying careful attention to the producer's requirements for high quality and containing as much ambisonically encoded information as possible.

Finally, these visual and aural images would be mixed over an original music score, which would be paced and timed for the shifting images.

#### Audio-video synergy

The first location chosen for the postcard treatment was Washington, DC; the eventual production spanned just about every form of visual and sound recording, as well as having to bring together elements from a variety of diverse media for final editing, mixing, duplication and sale.

Having a strong recording and musical background, Nicholas considered that the best way to break out with his new product would be to have a home video release that was strongly differentiated by an audibly superior soundtrack, matched with the highest quality visuals.

These requirements led him to digitally master the project on Mitsubishi X-800 multitrack and X-80 stereo machines at south Florida's Criteria Recording Studios and, more importantly, the use of ambisonic technology whenever and wherever possible.

Mac Emmerman, Criteria owner and past president of SPARS, was brought in early as the project's sound designer. His enthusiasm for the project, allied with the producer's quest for quality, quickly

involved other high-tech aspects

A decision was made to offer the product only on Beta hi-fi and VHS hi-fi videocassettes, not only with an all-digitially mastered music soundtrack, but also with 2-channel ambisonic UHJ-Format surround sound. To this end, a Calrec Soundfield microphone was used to record as much of the audio material as possible, with special attention being

given to its use in capturing ambient sounds on location.

Although the ambisonic system is capable of capturing a 3-dimensional picture of sound for later decoding and playback over a domestic surround-sound system, Nicholas thought that the technique's outstanding mono compatibility was also of value. Not only would a 2-channel UHJ soundtrack enable the critical

## Ambisonic Surround-Sound Technology

*Ambisonics is a total systems technology, aimed at capturing, recording and replaying surround-sound information to meet a wide range of professional and consumer needs.*

*In the studio, ambisonic information is recorded onto tape in the form of four B-Format signals, known as W, X, Y and Z. The W signal represents the omnidirectional soundfield arriving at the microphone position, while X, Y and Z signals are equivalent to bidirectional (figure-of-eight) microphones, pointing respectively forward to back, left to right, and up and down.*

*For horizontal surround soundfields, the Z signal is set to zero (or omitted), and only the W, X and Y signals used. (It should be noted that these four B-Format signals bear no relationship to the signals routed to a 4-speaker monitoring array and only serve as recording channels.)*

*The Calrec Soundfield microphone houses a tetrahedral array of four cardioid capsules. A companion control unit allows sounds to be panned and steered in horizontal and vertical planes either in real-time or in post-production from off-tape B-Format signals. Azimuth and elevation controls rotate and tilt the microphone's apparent spatial orientation, while dominance and zoom controls focus and emphasize sound on front and back planes.*

*To process ambisonic sound, three pieces of equipment are required:*

- A transcoder unit, which enables both encode (B-Format to 2-channel UHJ) and transcode (conventional 4-channel mixes to 2-channel UHJ) capabilities. With this unit, control of front and back stereo width is provided.
- A multitrack pan/rotate unit, which features sine/cosine panpots and switchable controls that allow a 360° soundfield rotation. The panpots can be switched before or after the rotate control, and the unit can be cascaded

*via external B-Format inputs for recording or post-production control, to produce a suitable UHJ (2-, 3-, or 4-channel) master.*

- A converter unit, allowing conventional console panpots to localize a soundfield.

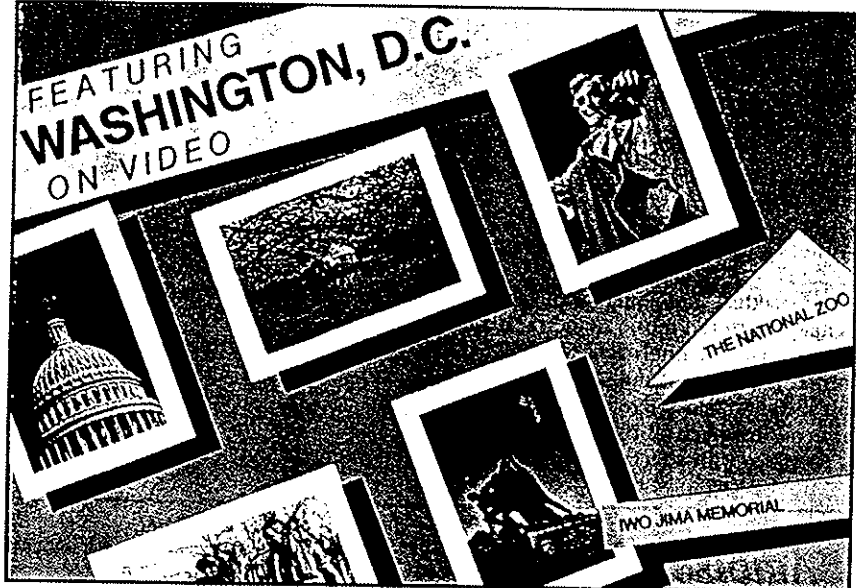
*In its simplest form, UHJ uses the same two recording or transmission channels as conventional stereo, using both amplitude and phase information to convey a horizontal surround-soundfield to the listener. Via a suitable UHJ decoding system, an ambisonics playback system can process information from multiple devices, including AM/FM transmissions, albums, cassettes and Compact Discs, allowing the listener to move around the listening area and face any direction without losing the sound image.*

*To fully use ambisonic technology, six or more monitor loudspeakers are necessary for full-sphere or periphonic reproduction. Otherwise, horizontal surround-sound monitoring (without Z-channel or height information) require only a decoder and four speakers, arranged around the listener.*

*Consumer systems are normally based on 2-channel ambisonic UHJ Format. This 2-channel signal, when properly decoded, provides four or more loudspeaker channels for surround-sound reproduction. And, in contrast to B-Format signals, the 2-channel UHJ signals can be monitored in either mono or stereo via conventional replay system.*

*Unlike conventional 4-channel quadrafonic sound—which uses variants of a 4-2-4 matrix—ambisonic technology allows the use of virtually any number of loudspeakers to recreate localized phantom images, even between side loudspeakers.*

Editor's note: The information in this sidebar was extracted from "Ambisonics Surround-Sound Technology for Recording and Broadcast," published in the December 1983 issue of RE/P.



listener to recover location and music ambience contained within the original mix via a surround-sound decoder, the less well-equipped audio enthusiast could replay a perfectly acceptable monophonic sound. (Which, after all, represents the bulk of the current VCR market.)

With the selection of recording methods and venue decided, Emmerman and Nicholas linked up with Miami-based arranger/composer Mike Lewis, who was commissioned to write the original score. The producer's intention was to first record the score, shoot the visuals to fit, add location sound effects and ambience, and then edit, conform and synchronize all of the elements in a final mix.

#### Music scoring sessions

More than 30 of the Southeast's best session players were used at Criteria Studio E during the music tracking date. With only a rough idea of how the images were to go together, Lewis provided far more music than was required to accompany the finished product. The concept was to record a group of separate pieces, each related to a certain part of the nation's capital, and timed to the footage that would be shot on-location.

A total of 12 individual pieces appear in the final cut: "City Scapes," the opener; "Cherry Blossom;" "Tribute;" "Celebration;" "Zoo;" "Peace;" "Hirshorn Sculpture Garden" (shot at the Smithsonian Institution Garden); "Capitol;" "WDC Breakdance Capital USA;" "Pulse;" "Vietnam Memorial;" and "Solitude."

Some of these titles are self-explanatory; others make sense only to those who have visited these sites around the Capitol. (But, then again, you don't generally send a postcard from a place you haven't visited.) Together, all of these subjects form one producer's idea of what would be remembered about a

visit to this historic city.

Two of the pieces presented technical problems for the producers of this ambitious project: "Vietnam Memorial" and "Breakdancing." Both segments required a great deal of preplanning—even then, as is often the case, there were massive saves made in the mix/final post process.

#### Location shoots

Both sequences were shot on location, using a Sony Betacam system. The memorial piece required an ambisonic sound-effects recording of helicopters in-flight. By special arrangement of the Air Force, Nicholas and Tom Gandy of Audio+Design/Calrec were granted access to Andrews Air Force Base. Their plan was to record the live sound of the president's Mission, a 6-chopper flight of helicopters designed to provide presidential transport in case of an emergency.

These sounds, properly mixed, equalized and brought in over the music, are intended to provide a haunting introduction. Using a Soundfield microphone and UHJ encoder, Nicholas and Gandy recorded the 2-channel sound onto a portable VHS hi-fi videocassette recorder, so that the tracks could be synchronized using the video sync reference.

The pair recorded about a half dozen passes of the helicopters—just enough to *layer up* in the studio and provide the effects required by the production.

The crew then headed downtown to capture the traffic noise that later accompanies "City Scapes," the piece used to introduce Nicholas' audio/video postcard. The use of a portable VHS hi-fi VCR to record the UHJ-encoded ambisonic ambience sounds was typical of the kind of decision a producer must make everyday. With a limited budget and the need to transport a lot of equipment to location, sometimes expediency must win out over perfectionism.

Because all of the material recorded on location needed some sort of reference for synchronization during post-production, using a VCR or other sync recorder for audio was mandatory. FM-encoded VHS hi-fi technology was in the right place at the right time, for the right price. (Later, we'll see that even a control-track reference may not help in post-production, unless some careful planning is done up-front.)

After making UHJ 2-track recordings of traffic sounds, the crew moved on to videotape the "WDC Breakdance Capitol USA" segment. Working with previously casted local talent, the high-energy piece captures the feeling of this urban art form, with a costumed Uncle Sam breaking away in a routine called "Style and Profile."

Sequences performed by the "Mighty Poppalots," a local dance group, were shot to a rough mix of the material recorded at Criteria, which in the case of the breakdance music contained no live sounds. Everything was built up of synthesizer and sequencer parts taken direct. In much the same way that rough tracks are used while producing a music video, the audio material was necessary to establish the duration of the piece, and to cue the performer's moves.

The video was shot without the customary practice of *dumping* the scratch audio onto the VTR's audio tracks. Instead, a time code generator resident in



Seen here during the mixdown for the "Washington DC" visual postcard project: (clockwise from top left) Stephen J. Nicholas (producer/director), Tom Gandy (Audio+Design/Calrec field engineer), Mike Lewis (composer/arranger) and Mack Emmerman (Criteria Studios president/sound designer).

## Ambisonic Radio Drama

Late last year, KWMU-FM, the public radio station of the University of Missouri-St. Louis, broadcast Stanley Elkin's radio drama, *The Coffee Room*, which was produced in the 2-channel UHJ ambisonic format. The play was the first ambisonic radio drama produced in the United States, and featured Elkin with 13 other St. Louis actors. It was directed by John Grassili and produced by KWMU's Lorin Cuoco and Sean Collins.

In order to take full advantage of the opportunities offered by ambisonics, the play was recorded in a room thought to have the sound of a typical university coffee room, with the actors working 360° around a Calrec Soundfield microphone. The voice tracks produced from these sessions were edited

and put onto an Otari 8-track along with sound effects recorded ambisonically on the Washington University campus.

An Audio+Design pan/rotate unit was employed to place effects and additional speech in the ambisonic soundfield. The final Dolby-A encoded UHJ ambisonic broadcast mix was done through a Harrison PRO-7 console by Collins onto a Studer A-810 2-track. Monitoring during mixdown was accomplished using four KEF 101s, two stereo amplifiers and a Minim AD10 ambisonic decoder.

KWMU premiered U.S. ambisonic broadcasting in May 1984, and has continually sought to integrate ambisonics into its program offerings.

the Betacam unit was used to provide a stripe for later use in editing.

### Video post-production

The basic elements were now complete, but much more shooting and location work were needed to fill out the rest of the producer's concept. Following these two segments further, production moved to Editel/Chicago for video post.

By this time some of the other segments had been shot on other formats, including film. Other pieces took the form of film footage picked up from sources such as award-winning cinematographer Louis Schwartzberg, who provided the time-lapse and aerial sequences used so effectively in this postcard. Editel provided the equipment necessary to transfer all of this material to a common 1-inch C-Format for editing.

Off-line workprints were created, and the production team went to work editing the material to a rough mix of the audio. Editel's Montage editing system

provided the ability to preview a large quantity of edits, and quickly assemble an off-line edit decision list.

Assisted by Larry Sexton, Nicholas finished the EDL and then moved into one of Editel's 1-inch, on-line suites for the final cut and addition of optical and special effects. Visual pacing of the production was assisted by extensive use of the facility's Ampex ADO system and various other digital visual effects, manipulated by editor Jerry Doskoczynsky under Nicholas' direction.

At this point, there was only rough conformance and synchronization of audio to video; the producer thought that the audio tracks could be massaged into place during the final mixdown process. A ¾-inch U-Matic workprint was dubbed from the 1-inch Edit Master, and the print moved back to Criteria for the final conformance and synchronization.

### Music and effects assembly

Parts of the soundtrack needed to be flown-in to the 32-track digital master on

a Mitsubishi X-800 supplied by Digital Associates, Nashville, TN. Although control tracks have been provided on all of the audio material, it was discovered that the digital multitrack could not be slaved to the VHS-format videocassette recorder.

As many readers may already be aware, a digital recorder has its own control track, and generally prefers to be the boss in synchronization situations. (Try to servo-control the capstan of most digital recorders beyond a tight tolerance window and, with few exceptions, the audio mutes.)

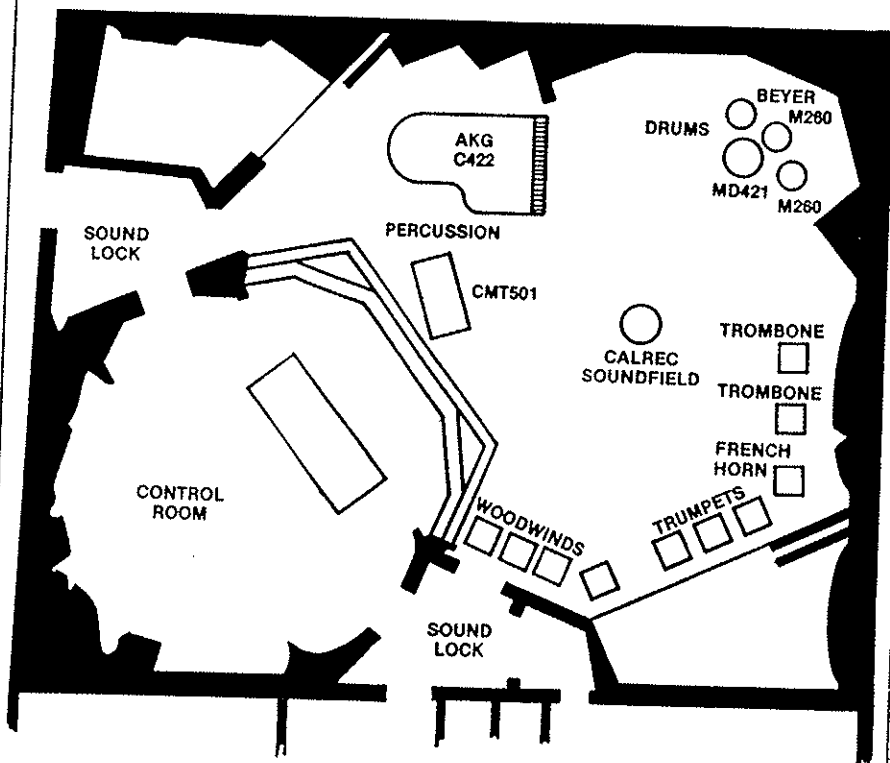
During the original music-tracking dates, both ambisonic and spot microphones were used. Although this meant that the multitrack was pretty loaded up, by doing some bounce downs (no problem with digital material) a few tracks could be freed. Spare tracks were necessary because to provide creative control over the final ambisonic surround-sound mix, the UHJ-format material recorded in the field had to be decoded and re-recorded as four B-Format signals repre-

senting the horizontal W, X and Y plus vertical Z information.

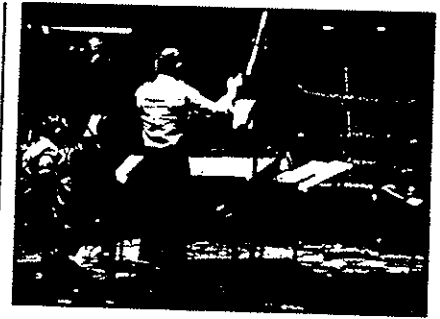
Suitable monitoring equipment was set up in the control room, and the process of flying in the audio began. Fortunately, the combination of hits and rhythms that the producer wanted to develop did not contain any lip sync (there's none in the entire postcard), and so most timing could be worked out manually. Of course, the finished visuals had to look right, which meant many long hours of tries and retries before everyone was satisfied.

By now the project was moving forward, with a time code track holding the digital multitrack mix in sync with the video workprint by means of an Audio Kinetics Q.Lock 3.10 synchronizer. Following standard practice, the next step would have been to a final automated mix of the audio material, and then transfer this mix to a digital 2-track.

If the multitrack and 2-track time-code tracks were conformed, the audio would sync up to the picture perfectly,



Room and instrument layout during the recording of the "Vietnam Memorial" segment at Eastwing (Studio E) Criteria Studios, Miami.



The orchestral music recording sessions at Criteria before the video shoot. The location visuals were shot to the musical tempos and edited during subsequent post-production.



Photo by Alan Jerriam

Video post-production editing sessions at Editel/Chicago, with Nicholas (right) spotting time code locations during preparation of the edit-decision list.

and a complete mix would be ready to go out for duplication. But a Mitsubishi X-80 (now superseded by the X-86) only knows one way to record—the time code, control track and audio all at once—or what is often referred to as the *crash record mode*.

All of which meant that it would be impossible to produce a conformed Audio Edit Master, a format that could have represented a real time-saver during the process of readying the material for videocassette duplication. Again, one of Criteria's engineers, Richard Aker, had to work with the producer to fly-in each segment of the final mix onto the 2-track Edit Master, with the remixed ambisonic surround-sound information again encoded to 2-channel UHJ.

Once the audio material was completed, the final UHJ mix was transferred to dbx model 700 CPDM digital format before being sent out for videocassette duplication. CBS/Fox was chosen as production house for the first run of dupes, in both Beta hi-fi and VHS hi-fi formats, and specified CPDM-encoded material to ensure compatibility with its system. After a number of small glitches—like forgetting to run video sync to both the X-80 reel-to-reel digital recorder and the dbx/VTR recording system—CBS/Fox

produced the dupes and readied them for market.

Since then, the project has been circulated among many in the audio community, and initial sales of the *Washington, DC* videocassette are said to be brisk. Demonstrations have proven that ambisonic information captured with the Calrec Soundfield microphone is sufficiently robust to withstand all of the different recording methods used during production, and still is capable of decoding and playback as convincing surround-sound information. Although not many people have access to a consumer UHJ decoder, Nicholas is content in the knowledge that the information is there for those who are interested in ambisonics and the equipment to match.

Final mixing of the project was done by Criteria's Mac Emmerman and Steve Johnson. According to Emmerman, although most of those involved in the project knew that the audio would largely be heard in mono, careful use of the ambisonic surround-sound information yielded the "best mono sum I've heard." When this writer heard the audio tracks for the first time, it was on a basic VHS videodeck. My initial impression was that there did indeed seem to be something very special about the sound quali-

ty, even when replayed over a 4-inch speaker.

Nicholas is already involved in pre-production of his next visual-postcard project. He has become a firm believer in the value of ambisonic recording to video and TV production. Interestingly enough, Nicholas would definitely use ambisonics technology in his next project, but does have reservations about video production.

"We all learned a lot during this first project," he says. "And I have developed a great respect for sprocket holes. If we have the budget, my next project will be done entirely on *film*, so we can avoid all of the interformat problems that can arise when a critical sound mix is married to pictures that are gathered from a lot of different sources."

Nicholas will continue his role as a proponent of ambisonic surround-sound, particularly for audio/video presentations. The technique is now beginning to be noticed, and recent business developments may make the producer a pivotal player in the technique gaining more widespread acceptance.

Ambisonics may well be in that unique position as an idea whose time has come, and bears some careful watching during the next several years.

**R/E/D**