

 [Print Page](#)

## Audio for DTV: The Beginning of a Journey

1/22/2003

Last time we wrapped up our detailed discussion of audio/video synchronization. Since that article was printed, I received a note from Tom Daily at Dolby mentioning that he had heard of yet another potential source of A/V sync errors-flat-panel displays! Although we are moving on, I'm sure we will revisit the subject briefly in the future once all the facts are in.

This month we will start our journey exploring audio for digital television from end to end. This will span three major areas sometimes referred to as production, distribution, and the one we have covered called emission (transmission). To understand what we were aiming for-the consumer-we had to investigate the final emission stage first. Now we can start from the true beginning: way back in post production.

### MONITOR, MONITOR, MONITOR

Post production of multichannel audio, or any audio really, can be summed up in three words: monitor, monitor, monitor. If you mix and monitor in a standard, calibrated room, you will very likely create a program that sounds just the way you intended it to when it is transmitted to consumers. An excellent resource called "5.1 Production Guidelines" can be found at [www.dolby.com/tvaudio-it](http://www.dolby.com/tvaudio-it) outlines many of the details of setting up a proper monitoring environment.

Following are some basic tips, resulting from experiences that a few of us have had over the years. First, buy an SPL meter to properly calibrate your monitoring environment. They are readily available-even Radio Shack carries them. Believe it or not, when I was working for Dolby on the film mix stages in New York, they were all calibrated with the Radio Shack model (even Todd-AO where Apollo 13 was mixed and won an Academy Award for sound). Although we calibrated the meters using a Bruel and Kjaer reference, they were usually within 1dB right out of the box. Both analog and digital models are available, and I own a few of each, but I prefer the analog version because it helps me cope with the ever-increasing amount of digital stuff we are surrounded with.

Next, you simply must have enough speakers-do not expect acceptable results when attempting to create a 5.1 channel mix while monitoring with only two speakers or a pair of headphones. I understand that it is not always practical to have a dedicated center channel, especially in tight remote production vehicles, but even a speaker that is not exactly in the perfect spot may be better than none at all. A basic setup is shown in Fig. 1.

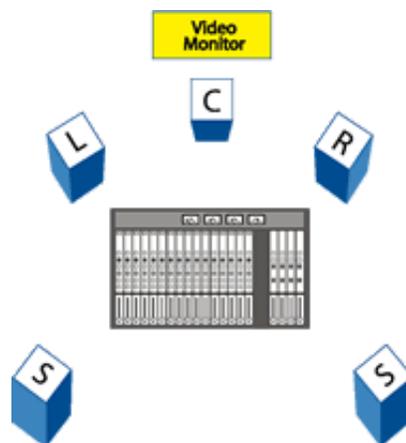
It is also good to have a collection of "reference" programming. There is so much 5.1 channel material available on DVD these days that it should be relatively easy to find a few various examples. They should not all be action-adventure, or dialogue only, or music, but a cross section of a few types.

Finally, it is also necessary to be able to emulate different speaker setups. Although the largest would be a full-range 5.1 channel system, many viewers will tune in and hear the program downmixed to stereo, or very likely downmixed to mono for the channel 3/4 RF re-modulators found in set-top boxes. There are a few monitoring systems available now from Adgil, Martinsound and others. However, you might be able to solve two requirements at once by using a product from Dolby Laboratories. The DP-570 Multichannel Audio Tool provides both accurate monitoring functions such as downmixing to all possible listening formats, as well as providing metadata authoring functionality.

Armed with your DVDs, an SPL meter, enough speakers, a monitor controller and the Dolby 5.1 Channel Production manual, you can quickly get a proper mix up-and-running.

Once you have the monitoring environment set up, you need to consider metadata. As we have discussed, the Dolby DP-570 integrates this functionality with monitoring in order for any metadata changes to be heard in real time so that appropriate decisions can be made. Many of the 5.1 channel programs created for television, such as "The Sopranos" on HBO, are using the DP-570 in production to help ensure that the program plays back properly in multiple listening environments. I can attest to the fact that it seems to be working as I have listened in my 5.1 room, the mono set in the kitchen, and in stereo in the bedroom. This was the first season broadcast in 5.1 and the mixes, although good from the beginning, seemed to my ears to get even better with each new episode.

What metadata is appropriate? That topic could easily be the subject of my column for the remainder of this new year! I will give some simple advice though. The presets stored in the DP-570 will get you 95 percent of the way home. In



(click thumbnail)  
 Fig 1. Basic monitoring setup. Note that the surrounds can be either mono or stereo, and that the subwoofer for the LFE channel can be hidden below the mixing console if necessary.

reality, most all programs will fit into the categories set up by these presets, but the Dialogue Level (dialnorm) must be measured and set for each program as it may vary substantially from program to program. Presently the ATSC's System Evaluations Working Group (we call it the sewage group), of which I chair the Audio Issues section, is creating a document that contains a table showing these categories along with descriptions of the recommended settings. Although it is not yet published, if anyone would like a copy of that section of the document, please e-mail me and I will provide it to you.

## **FITTING IT ON TAPE**

Now that you have a proper mix that has been checked for downmix compatibility and metadata that correctly represents the dialogue loudness and other properties of the program, what do you do with it all? You have up to eight channels of audio and a metadata stream. There are a few options.

One is to record the audio to an eight-track format such as DA-88 and via SMPTE timecode, synchronize this tape with the picture on a standard- or high-definition VTR. Although this scenario will work, there is no simple way to store metadata and there is always a possibility that the two tapes will lose synchronization.

Another method is to choose a VTR that can store multiple channels of audio. The Panasonic AJ-HD3700 is one such machine, capable of storing up to eight channels of 24-bit uncompressed audio when running in 24p mode. Again though, what do you do with the metadata? One solution as described in SMPTE specification 334M is to put the data into the vertical ancillary (VANC) space of the high-definition serial video signal. Currently, Norpak Corp. is manufacturing a product called the TES-7, which allows metadata and other data, such as closed captioning, to be embedded in the VANC space.

An additional solution to the problem is to use an audio format such as Dolby E. This technology was created to allow up to eight channels of audio plus professional metadata to fit into a 1.92 Mbps serial datastream. This happens to correspond directly to a 20-bit, 48-KHz AES pair just like those found on common digital videotape formats like Digital Betacam (Digibeta), HDCAM, D5, HD-D5 and even DVCAM. The use of Dolby E has become widespread, and is the source for most of the multichannel audio presently heard on television. Although reasonably simple to use, two very important points must be noted. First, the 20-bit AES channel that is expected to pass Dolby E must be bit-for-bit accurate-i.e., free of sample rate conversions (beware, they lurk in some very unexpected places such as de-embedders), level shifts, channel shifts, etc. Second, the encoding and decoding processes each impart a one video frame delay on the audio signals. This is simply and effectively compensated for during post production by advancing the audio tracks as they are being Dolby E encoded and recorded to tape. Nicely, this system accepts professional metadata at the input terminals of the encoder, and produces the metadata at the output terminals of the decoder.

Now what? How does this tape make it to the next stage? Good question. Next time we will begin the complex topic of distribution. This is where we will explore how audio and metadata make it from the front door of the network to the front door of the affiliate station. I have seen fewer roadblocks and potholes in Manhattan during the Macy's Thanksgiving parade, so be prepared for a bumpy trip.

As we begin this New Year, I would like to thank you all for your generous feedback over the past 12 months-please  keep it coming. Thanks for your time! [Print Page](#)