



Technology Corner: Randy Hoffner

ATSC Recommends AV Sync Rules

Technology Corner: Randy Hoffner

01.07.04

This column has frequently addressed audio-video synchronization or lip sync issues. The lip sync problems in television today result from the proliferation of digital processing that audio and video signals undergo in the contemporary television plant. These problems appeared with the dawn of digital processing of video and, later, audio signals; we can rest assured that Uncle Milt's show did not have any of these problems. The fact is that both audio and video signals remain in the digital domain for much, if not all, of their travels through the production, post-production, distribution and transmission pathways. There is much digital processing applied along the way, and each digital processing step makes its own contribution to the cumulative audio-video synchronization problem. It is a fact that video signals are almost always subjected to more delay than audio signals, and this results, at the end of the line, in audio leading its associated video.

DIGITAL DELAYS

This is the worst possible outcome, of course, because in nature, light travels many times faster than sound, so the associated sound always lags behind the visual event for a remote observer. We therefore can accept a certain degree of audio-lagging video, because it is what we experience in the natural world. The visual component of an event never lags behind its associated sound in nature, and for this reason, our sensory apparatus cannot inherently make sense of such an observation.

The reasons for audio-video synchronization errors then are grounded in digital processing. This led the Advanced Television System Committee's (ATSC) Implementation Subcommittee to charge its Systems Evaluation Working Group to address the systemwide audio-video synchronization problem in DTV. Its investigation ultimately led to an Implementation Subcommittee Finding, IS/191, ATSC Implementation Subcommittee Finding: Relative Timing of Sound and Vision for Broadcast Operations, which is available on the ATSC Web site, http://www.atsc.org/standards/is_191.pdf.

PROCESSING PROBLEMS

The crux of the audio-video synchronization problem in DTV systems is well-stated in the first paragraph of IS/191:

"The end-to-end DTV audio-video production, distribution and broadcast system is a complex array of digital processing, compression, decompression and storage devices. Each component in the system imposes a latency on the audio and/or video signals flowing through it. System design goals often call for the relative audio-video latency through each component to be in the sub-millisecond range. Operationally, unequal delays can be imposed on the audio and video signals respectively, and these delays compromise audio-video synchronization."

The Finding divides the end-to-end DTV system into four segments: acquisition and production/post production; release facility and distribution system; local broadcast station; and home receiver. Synchronization problems may be introduced in all these segments, and the finding states that, "...steps must be taken to ensure that the audio and video signals delivered at the output stage of each of the four segments are synchronized within a tight tolerance."

The Systems Evaluation Working Group, which authored IS/191, conceived that for the purpose of assigning synchronization tolerances, the end-to-end DTV system could be conveniently divided into two parts with the first part ending at the DTV broadcast encoder inputs. The SEWG determined that overall, end-to-end audio-video synchronization should be held within a tolerance of +30 milliseconds (audio leading video), to -90 milliseconds (audio lagging video). These tolerances correspond reasonably well with the traditional lip sync rules of thumb for NTSC television, which set the tolerance at +1, -2 frames, or +33, -66 milliseconds.

The finding states that, "IS finds that under all operational situations, at the input to the DTV [broadcast] encoding devices... (T)he sound program should never lead the video program by more than 15 milliseconds, and should never lag the video program by more than 45 milliseconds." It was conceived that the second half of the end-to-end chain, that from the encoder input terminals to the receiver's output, would have a similar +15, -45 millisecond tolerance. The detailed description of the second half of the system's tolerance was to be the subject of a future finding.

These numbers are significantly tighter than those specified in ITU-R Recommendation BT.1359-1 (1998), which has been discussed here previously. For this reason, the Implementation Subcommittee noted that Rec. 1359 was considered and found inadequate for purposes of audio-video synchronization for DTV broadcasting. IS has further made the appropriate ITU-R body aware of IS/191, and suggested that Rec. 1359 be reassessed in light of this finding.

Audio-video synchronization has been one of the major casualties of the DTV era. The industry has become well-aware of this problem, and the Implementation Subcommittee of the ATSC has stepped up with a recommendation that, if observed in implementation, will make a material contribution toward rectifying this problem.

Randy Hoffner is manager of technology and strategic planning at ABC, New York, N.Y. The views expressed in his column are his own, and not necessarily those of ABC. [Write to him c/o TV Technology.](#)