

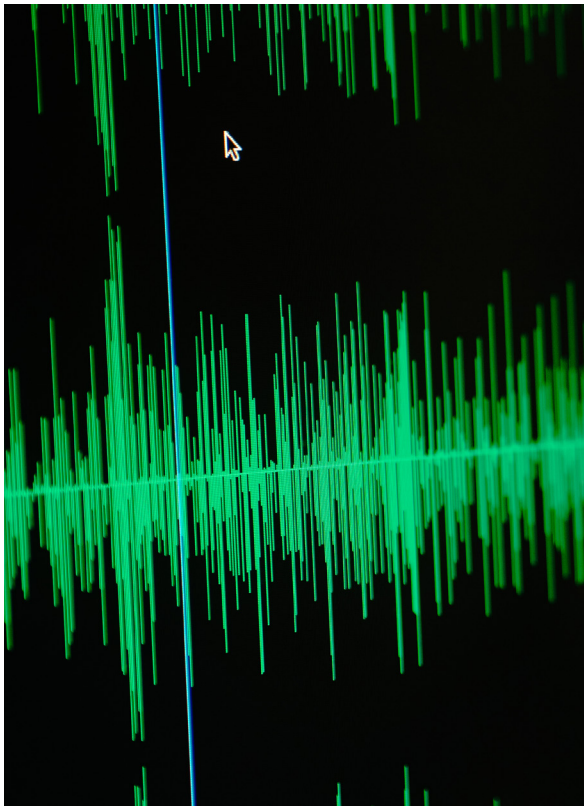

Progress Report on Audio

By Brian Vessa

The world of audio has progressed rapidly over the last few years, ushering in new tools, infrastructure, and workflows that significantly upgrade quality and efficiency. As technology has advanced, so have computer power and Digital Signal Processing (DSP) capabilities. The processing power of a single computer chip, such as Apple Silicon, is astonishing. Digital Audio Workstations (DAW) can now handle enormous numbers of tracks and plug-ins. Combined with the introduction of tailored AI into the audio toolset, this enables capabilities that were unimaginable a few years ago.

Networking and Storage

Let's start with audio networking and storage, which have advanced dramatically. Many channels of discrete audio can now travel long distances over a single network cable with minimal latency, aided by CAT 8 (the fastest copper twisted-pair Ethernet cable standard available) and fiber-optic cabling. MADI (Multichannel Audio Digital Interface) remains a reliable workhorse, while AES67, which provides high-performance audio-over-IP (AoP) and audio-over-Ethernet (AoE) interoperability, has enabled fast, accessible multichannel networking implementations such as Dante. Ten-gigabit Ethernet is now native in personal computers, with even faster connections available in enterprise environments. Networked storage has therefore become incredibly fast, with many terabytes of data available at the drop of a hat. Storage itself has also taken a huge leap. Solid-state storage (SSD) has increased in speed and capacity. Hard disk drives (HDDs) continue to grow and can now handle 32TB on a single drive. Both can be arrayed for far greater storage. Internet connectivity speeds have also ramped up considerably, enabling quick data transfers and real-time collaboration. Together, these advances let us store, process, and create at speeds never thought possible.

Powerful New Audio Tools

Tools for music, audio production, and post-production, in the form of plug-ins for DAWs and stand-alone applications, have made incredible strides thanks to increased DSP capability, AI neural networks, and machine learning.

"Stem Separation" tools allow separating composite audio into its components so each can be processed and mixed separately. In the post-production world, dialog, music, and effects can be separated from a mixed track, and further divided into individual actors, sound effects, and musical instruments. Older movies with a monaural or stereo mix can take on a new life and be remixed into immersive sound. Similarly, in the music world, the ability to separate individual instruments from a mixed track allows for greater clarity and upmixing. AI voice recognition is also being used to split overlapping dialog, singers and speakers. The AI can recognize the individual voices and create separate tracks for each. Popular tools include the RX suite and SpectraLayers.

AI voice recognition is also being used to split overlapping dialog, singers, and speakers. The AI can recognize the individual voices and create separate tracks for each. AI has also improved searches of sound libraries. A person can use a keyword or even mimic the desired sound, and AI will present a palette from the library that resembles it. This has greatly accelerated searches through large sound databases. A popular example is Orion-SFX.

Noise reduction tools have really taken on a life of their own, and many new products have hit the market that can reduce noise in recorded tracks. In the post-production world, they are widely used to clean up production dialog and production sound effects (PFX). Many of these work on a three-layer approach, dividing a dialog recording into dialog, noise, and reverb, and can further divide these into frequency bands. The audio operator can then balance each for the most natural, clean sound. Tools for clicks, pops, and crackle are numerous, faster, and more accurate after several generations of refinement. For movie and music remastering, noise reduction tools have become indispensable for lowering noise and increasing dynamic range. Popular tools include Absentia, Clean, Digital Dialog Extract, DX Revive, DX Split, Hush, SpectraLayers, and the RX suite.

Equalization (EQ) matching tools allow any sound to be frequency-matched to a reference. AI figures out the frequency characteristics and suggests EQ curves. Using these tools, re-recording mixers can quickly match automated dialogue replacement (ADR) to production, newer sound effects can be matched to older ones in restoration, and music tonality can be matched in mastering. Popular tools include FabFilter and Ozone.

Audio alignment and phase-matching tools have been around for a while, but have greatly improved in accuracy and quality. These allow precise timing matches between audio events. In post-production, ADR can be tightly matched to production, Foley to PFX, and multiple microphone recordings can be phase-aligned. In music recording, an instrument recorded directly and with a microphone or through an amplifier can be phase-matched and mixed. Popular tools include Auto Align Post, IBP Phase Align, and VocAlign.

One of the latest plug-in tool sets allows the acoustics of a recording to be analyzed, sampled, and automatically matched. In post-production, this can be used to apply the same acoustics from a production recording to ADR, or match an actor's lavalier mic recording to the boom recording. Similarly, for music, an overdub can be matched to the live room track. Definitely a game-changer! A popular example is Chameleon Reverb, which received a technical Emmy in September 2026.

As these plug-ins are automatable, the parameters can be memorized and changed instantly on the workstation timeline, allowing a tremendous amount of precision.

As with all tools, there is no "one-size-fits-all" application. Each has strengths and weaknesses, and it is up to the talent of audio creators to find the best tool, or set of tools, for each job. The old adage "less is more" definitely applies here—overprocessing is all too easy to do. The skilled operator needs to have good ears and deep familiarity with the power in these tools for them to be the most effective. AI can aid the cause, but humans still call the shots!

Custom Software

Custom software to streamline workflows is a chat away, as AI chatbots can now code software! A person describes the desired function and provides example data if available, and voila! A code script is generated. This allows all kinds of individually customized software that would not have been possible even a year ago. These can range from scripts for repetitive tasks to analytical tools such as acoustic analysis, and much more.

Immersive Audio Advances

The world of immersive audio has expanded! The increased power of DAWs now makes it easy to create 9.1.6 or 7.1.4 mixes. Most high-end mix stages and recording studios are equipped to create it, and most major movies and TV shows are mixed natively in immersive audio. Music has now made the leap to immersive, with many recordings created or remixed in formats such as Apple Spatial Audio, Dolby Atmos, and Sony Reality Audio. Many premium cinemas have immersive sound systems installed, and they are increasingly available for the home. Consumers can now buy a complete loudspeaker playback system at a reasonable price. There are a number of soundbars that will do a fair job, especially if you add a couple of rear loudspeakers, which can easily be connected wirelessly or via Bluetooth. On the higher end, increased demand for home theaters has fostered a burgeoning business, with very knowledgeable acousticians designing and installing high-end equipment.

Headphone Technology

But the biggest change for consumers has come via headphone technology and encoding algorithms. Headphones have seen a big bump in quality thanks to advances in

1545-0279/26©2026SMPTE

psychoacoustics, binaural technology, and transducers. Earbuds have also seen a significant quality improvement. Individually customized Head-Related Transfer Function (HRTF) curves can now be created by testing the user's Head-Related Impulse Response and applying convolution. This optimizes the headphone experience for each listener and allows for very realistic sound localization. Because every consumer cannot be tested, an HRTF can instead be selected from a database using head and ear measurements or photographs. Once the HRTF is optimized, the listener can truly hear immersive audio and experience a full 360-degree sound field on standard headphones. Music, movie, and gaming aficionados are loving this!

Video Mixing

Building upon this new technology, it is now possible to create and mix immersive music and movie soundtracks through headphones that will translate extremely well to a loudspeaker system. The Sony 360 Virtual Mixing Environment (360 VME) was designed to do just that. The user is measured in the environment they want to emulate, such as a mixing stage or recording studio. Miniature microphones are placed in the ears, and sweep tones are played out of each loudspeaker. They then put on the headphones they plan to use, and the sweep tones are played through the headphones. A custom HRTF profile is then calculated. Multiple environments can be tested, and a profile is created for each. When mixing, the desired profile is loaded, the mixer wears the same headphones, and hears as though they were in that environment. Music, sound design, or mixing can be done anywhere, then brought to a recording studio or mixing stage for final tweaks. It also allows real-time collaboration between people in different locations. This workflow emerged during the pandemic, allowing creativity and production to continue while people were working from home. The technology has now been refined to a very high level.

ADM and Workflow Efficiencies

The Audio Definition Model (ADM) has been around for some time, but it was used primarily for archiving rather than audio workflows. Recently, tools to create, carry, and use ADM-BWAV (Audio Definition Model Broadcast Wave Format) have become readily available, opening a new world for delivering multitrack audio material. ADM-BWAV can represent multiple audio formats and tracks, including an immersive mix, DAW session, or any audio object, in a single interleaved .wav file containing up to 128 .wav files. Each .wav file has unique XML metadata in its header that describes the audio and how it is to be played. Since it appears as a .wav file, it can be carried over standard audio connections. DAWs can now easily export and import ADM-BWAV. One common workflow is to create an ADM-BWAV with a full multichannel or immersive mix, and rendering and conversion tools create the downstream deliverables such as 7.1, 5.1, and two-track mixes. It can also be used to carry Metadata Guided Audio (MGA), International dubbed dialog, M&E, accessibility, and audio description tracks. ADM-BWAV can now be carried in Interoperable Master Format (IMF) as part of a completely mastered A/V package.

Sound Quality—The Ultimate Goal

With all of these amazing new tools, at the end of the day, it is still about sound quality. Headphone, earbud, and soundbar quality have increased dramatically, and there have been great strides in loudspeaker design. Two new cinema loudspeaker designs introduced this past year substantially expanded what cinema sound systems can achieve. Monitor loudspeakers for home studios also took a big step up, as have loudspeakers for home theater and home stereos. While this has been aided by better hardware design, much of the change has come from greater understanding of how people hear and what they want to hear. Smarter, more sophisticated modeling programs are helping produce higher-fidelity products. And let's not forget microphones, the source for converting sound into audio. The past few years have brought new microphone designs with excellent sound at affordable prices. When it comes to high fidelity, transducers are still king!

The world of audio is filled with dedicated, passionate people who are on a constant quest for quality. We will continue to see new transducers, hardware, and software designs, along with more efficient workflows in the creative process. SMPTE currently has a project to create a Cinema B-Chain Sound System Specifications and Test Procedures Recommended Practice. See the TC-27C progress report in this issue. It is a very exciting time for audio!

Acknowledgments

The author would like to thank David Bowles, Charles Dayton, Marc Fishman, Rob Getty, Gary Gottlieb, and Avo Mardikian for their invaluable input.

About the Author



Brian Vessa is a veteran audio engineer and advocate for quality sound and standards. He chairs the SMPTE TC 27C-20 Sound Working Group. Vessa is a SMPTE Fellow and a member of the Audio Engineering Society and the Academy of Motion Picture Arts and Sciences.

A contribution received from the author for the 2026 Progress Report.
Copyright © 2026 by SMPTE.

DOI: 10.5594/JMI.2026/DGMT6561
Date of publication: 15 November 2026