SOME NEW POSSIBILITIES IN AUDIO RESTORATION

John S. Allen

INTRODUCTION

Forty years ago, audio recording entered a new era with the transition from disc mastering to magnetic tape mastering, comparable to the transition in written language from clay tablets to papyrus scrolls several thousand years earlier.

Only ten years ago, computer word processing supplanted scissors and paste editing of written language, and now, delayed by its greater requirements for computer power, digital editing of audio signals is supplanting razor editing of recorded tape. We have reached a second historical divide.

Any technology is regarded first as a more convenient or successful way to do what has been done before. But as a technology becomes widely adopted, it points to new possibilities; for example, today's word processing has onscreen spelling checking and typeface selection, features unimaginable to a writer whose technology went only as far as the clay tablet, the papyrus scroll or even the typewriter.

In the first decade of digital audio, the main effort has been toward duplication of pre-existing audio techniques such as recording, mixing, editing and filtering. I would like to look ahead to some additional possibilities which will be realizable within a few years, and which will have a very significant impact on the way that archives, collectors and audio restorers go about their work.

These possibilities result from the ability precisely to control the time relationship between digital audio signals. Since digital signals consist of a train of samples taken at precisely defined moments in time, they can be aligned and realigned precisely to each other in time. One digital track can easily be stretched or shrunk in time to align it to another.[1]

Time alignment is possible in a limited way with analogue signals. Tape editing is a crude form of time alignment. Dual-system film sound (one film for the picture and another for the audio) is synchronized by the clapstick at the beginning of each take, and from there onwards by sprocket holes kept in step by synchronized motors. The tracks of a stereo tape or disc are synchronized by being physically side by side on the recording medium so the playback head or stylus reads them at the same time.

There are severe limitations on synchronization of analogue signals, however. If an analogue signal has a noticeable time-base error ("wow" or "flutter"), this is difficult or impossible to eliminate. If two audio signals--or an audio signal and a moving video or film image--were not recorded along with a synchronizing reference, aligning them precisely is as hard as putting Humpty Dumpty together again. It must be done by trial and error, and sometimes it cannot be done with sufficient accuracy. Nonetheless, it is often necessary: for example, to synchronize film or video with a "wild" soundtrack from an unsynchronized audio recorder.

SUMMARY OF POSSIBILITIES OF DIGITAL TIME ALIGNMENT

Time alignment with digital techniques removes most of the difficulty and presents the following possibilities:

1) Synchronous playback of multiple copies of analogue recordings to achieve an improved signal-to-noise ratio;
2) Automated level-matching of copies of the same audio recording which have had different gain-control processing, allowing switching between them without level changes;
3) Removal of print-through and groove echoes without degradation of the desired audio signal;
4) Much easier synchronization of free-running (unclocked, "wild") audio recordings with moving film or video images;
5) Automated, exhaustively detailed reconstruction of the editing history of recordings, useful in restoration work and for forensic purposes;
6) Easy generation of true stereo recordings by synchronizing separate monaural recordings made at the same time and location;
7) Automated or semi-automated correction of time-base errors of analogue recordings such as those caused by disc warp and mechanical speed variations.

DETAILED DESCRIPTION OF POSSIBILITIES

Possibility 1, noise reduction, is easily understood by comparing playback of a full-track analogue tape by a full-track tape head against playback by a quarter-track head. The audio signal is the same across the full width of the tape, but the noise, which results from the particulate structure of the tape oxide layer, varies both lengthwise and crosswise on the tape. Therefore, the signal level builds up proportionally to increasing track width read by the playback head, but the noise does not build up as fast. In our example of a full-track vs. quarter-track head, the signal quadruples, while the noise only doubles: a 6 dB improvement in signal-to-noise ratio.

The same improvement can be realized by playing multiple copies of the same recording, if they can be synchronized accurately. For example, if one has eight copies of Caruso singing "Vesti la Giubba" from 1907, all with an equal amount of surface noise, playing them synchronously will reduce the noise by 9 dB--about as much as Dolby B processing of cassette tapes.[2] Where one copy has a transient tick or pop, eliminating that copy from the mix will result in an even greater improvement. Note that the gain in signal-to-noise does not involve any filtering or other tampering with the recorded signal. I predict that software to realize the synchronous playback technique will be available on moderate-priced desktop computers within a few years.

Possibility 2, automated level matching, is achieved by synchronizing copies of a signal, comparing their levels and adjusting the level of one to match that of the other. One example is to synchronize a master tape with dropouts to a disc pressing which has been subjected to level compression. The disc pressing can be used to fill in the dropouts while maintaining the uncompressed dynamic range and--except where the dropouts are replaced--the better audio quality of the master tape.[3]

Possibility 3, removal of print-through and groove echoes, follows from automated level-matching. The signal and its echo are synchronized, level-matched and subtracted to cancel the echo. Again, the improvement is realized with no filtering, level-shifting or other degradation of the recorded signal.

Possibility 4, synchronizing wild audio to film or video, is of use in reconstructing a sound film or videotape from an audio recording and a silent film or video image, or one with a poor or dubbed soundtrack. Film and audio of many historic events are preserved in bits and pieces from different sources, and synchronizing them could lead to much more complete and higher-quality sound film or video records of these events. Also, a theatrical film whose soundtrack has deteriorated can be restored by substituting a tape or disc soundtrack copy in good condition.

Possibility 5, reconstruction of editing history, is of use to collectors and archives seeking to determine the sources and authenticity of a recording. If two recordings are from the same source, synchronizing them will reveal consistent waveforms in both. Differences will reveal themselves through loss of synchronization. The process of comparing recordings could easily be automated, making it easy to trace versions of a recording through successive generations of editing.

Possibility 6, generation of true stereo recordings, has already been realized using analogue equipment--arduous work, though ARSC member Brad Kay has produced a number of stunningly successful stereo restorations from 1930s disc pressings.[4] Analogue restoration of stereo is possible, however, only with unedited sources that have very low levels of wow or flutter.

As Brad Kay has pointed out, recording companies running a second lathe for a safety backup at a recording session would often feed it with a separate microphone, in order further to reduce the chance of losing a take due to technical problems.

Discs marked, for example, take 1 and take 1A (typically at 9 o'clock to the label on American Victors) are different recordings of the same take and may be the two halves of a stereo recording. As an example, among my two copies and ARSC member Clark Johnsen's one of Victor set M/DM 85, Koussevitzky and the Boston Symphony Orchestra playing Tchaikovsky's 6th Symphony, five of the ten sides are unintentional stereo due to substitutions for worn masters. More stereo sides--maybe all of the rest--are gathering dust in the RCA vault!

It is well-substantiated that two orchestra microphones were used at the Toscanini/NBC Symphony concerts from the late 1930s to the late 1940s: one for the feed with an English-speaking announcer and one for the feed with a Spanish-speaking announcer. It is not known whether anyone accidentally tuned in two radios and heard the broadcasts live in stereo, but the transcriptions await stereo restoration.[5]

Audio recordings of many historic events also are available from separate sources, made with separate microphones. As an example, I have collected five different recordings of President Franklin D. Roosevelt's December 8, 1941 declaration of war speech, all clearly made with separate microphones. Most news items still are miked in mono, so events right up to the present are candidates for restoration. If one can capture them live from separate radios and/or TV sets onto a multitrack recorder, one is spared the trouble of synchronizing them--though it may be necessary to time-delay one channel to match another, especially with satellite links in the transmission path.

Stereo restoration may even be useful with recordings already available in stereo. For example, where different microphones or different mixes were used for mono and stereo masters of a recording, they may be combined into multichannel surround-sound recording.

The technique of synchronizing unintentional stereo differs from that for noise reduction. The two stereo channels do not have the same signal, so they must be synchronized by ear, even when using digital techniques. However, digital time alignment allows the recordings to be aligned automatically once a number of points of good synchronization have been adjusted by ear and marked.

Possibility 7, correction of time-base error, involves automated as well as operator-controlled techniques.

"Wow," which is time-base error due to eccentricity and warpage of a phonograph record, is a well-defined mathematical function of the shape of the record. Playing the record with a tonearm which is instrumented to register its position as well as the audio signal allows time realignment according to this mathematical function in order to eliminate the "wow."

Tape flutter, or any time-base error due to changes in speed of the recorded medium, is more difficult to correct. Sometimes a steady signal such as AC power-line hum, vibrato-free musical note or ultrasonic bias tone can serve as a timing reference[6]; or there may be another, less well-preserved copy of the same recording, but without the flutter and which can be used as a timing reference. If there is no timing reference, one must be established by ear. Digital interpolation then can smoothly adjust the speed between the reference points. This approach will greatly ease the task of maintaining constant pitch when transferring old multi-side recordings made on lathes with poor speed control.

Digital interpolation--actually in this case, extrapolation--can also extend the correction of a cyclic speed error such as a tape flutter into parts of the signal where there is no good reference. In cases where it is important to restore steady pitch to a valuable, historic recording, the repeated attempts needed to reach a precise correction may be justified.

IMPLICATIONS FOR COLLECTORS AND ARCHIVISTS

All of the possibilities discussed here bear directly on the choices which collectors and archivists should be making today about the material they preserve.

Wherever restoration of a recording may involve one of the possibilities described above, all copies of the recording which may be useful in the restoration should be preserved in the earliest-generation version possible.

A disc pressing or a photographically copied sprocketed film soundtrack preserves geometric timing cues and uncorrupted power-line hum from the master, but electronic copying destroys these cues. Only a first-generation tape has uncorrupted power-line hum and also the original ultrasonic bias tone which can serve as timing references. Because uncorrupted versions may have dropouts, transient noises or other defects, they are invaluable if synchronization or time-base error correction is to be attempted.

An attempt should be made to preserve a number of copies rather than just one "best" copy of a recording, so that the copies may be synchronized for noise reduction. The copies preserved should differ back to the earliest possible generation, to assure that noise will be different on each of them: for example, several copies of a disc pressing rather than copies of a tape dubbed from the same disc. Note, however, that transfers of the same disc using styli of different diameters will have different surface noise, and will provide some noise reduction when played back synchronously.

When unintentional stereo is a possibility, different takes or different networks' recordings of an event should be preserved, rather than just the "best" take or network version.

When a transfer must be made to preserve or acquire a recording, attempts should be made to preserve timing cues. Film and video soundtracks should be transferred along with a time code, or transferred synchronously to film or videotape. Equipment to do this is widely available.

Providing a time code for disc recordings requires equipment which is not yet available. One easily-implemented measure is to install an optical gate or other once-per-revolution index sensor on a turntable and feed its output to an audio track of the tape to which the recording is transferred. Building a position-sensing tonearm is a more complicated technical chore. Once it is built, though, there should be no trouble in storing its output in coded form on an audio track. Until a position-sensing tonearm is available, you can get by with tape copies except for one archival disc, preserved as a geometric reference. This should be the least warped (and in the case of wax-mastered discs, the least oval) of the lot, rather than the cleanest-sounding one.

Power-line hum on a disc or tape can be preserved uncorrupted by rerecording at a different speed, with a different power line frequency, or with equipment which has been carefully selected and tweaked to eliminate hum. A tape's ultrasonic bias tone should ideally be preserved by frequency-dividing it and storing it on an audio track; lacking equipment to do that, a special copy of the tape could be made at 1/8 speed to preserve the bias tone. This copy would later be synchronized to a copy made at normal speed which would preserve the audio information better.

Good archiving practices should be maintained for all transfers: the full audible frequency range should be retained, without filtering--both because noise-reduction through synchronous playback could extend the usable frequency range and because timing cues may be present in frequency ranges which are noisy or distorted. As is also normal archiving practice, transfers should be integral, with no editing--and especially, no excision of transient noises, which would disturb the time continuity of the recording and make synchronizing much more difficult. Removal of transient noises by oxide scraping or by an analogue or digital interpolation process which does not disturb time continuity is, however, acceptable except where these noises are copied from an earlier generation and may be useful as timing cues.

The watchword to remember in connection with digital synchronization techniques is that the total is greater than the sum of the parts. Once we are able to put the parts together, the potential of audio restoration will increase very significantly--as long as we preserve the source material to allow the application of the new techniques.

I invite anyone who is interested in the techniques I have described to contact me. I have written a technical paper describing them in detail, and I am actively seeking collaborators in an effort to build hardware and write software to realize these techniques.

NOTES

[1] A readable description of the technique for stretching or shrinking digital signals, called "sample interpolation," "resampling" or "sample rate conversion" is found in Hal Chamberlin, *Musical Applications of Microprocessors*, 2nd ed. Rochelle Park, NJ: Hayden Books, 1985, p. 518 ff.

[2] Synchronous playback reduces noise which differs between copies. This accounts for most of the noise on shellac pressings. Noise from the master or from earlier-generation copy common to all those used in restoration--such as a metal mother or stamper--does not differ. When this is the preponderant remaining noise, there is no advantage in adding more copies to the mix unless they differ at an earlier generation. The point of diminishing returns will vary depending on the copying history of the recording.

[3] Lexicon, Inc. (100 Beaver St., Waltham, Massachusetts 02154) already uses automatic time alignment and level matching for tape head azimuth error and level differences, preventing L+R "dialog" channel signals from spilling over to the L-R rear channel of a Dolby stereo matrix.

[4] Brad Kay, 722 Superba Avenue, Venice, CA 90291.

[5] Brad Kay discovered unintentional stereo in 1985 but points out that it had been discovered earlier in the 1970s by Clyde Key of the Toscanini Society, as documented in the Toscanini Society newsletter.

[6] The straight tone of the piano, organ and other vibratoless instruments, which is most revealing of wow and flutter, fortunately also provides the best musical reference to remove it.