A High-performance, Low-cost Wax Cylinder Transcription System

James Nichols
Ifixaudio.com
Paris, France

We describe a high-performance, low-cost wax cylinder transcription system. Using readily available parts and software, the BURP-ONE (Basic Utility Record Player) system gives archivists a pragmatic means for preserving the heritage of antique cylinder recordings. A practical methodology for quality restoration of wax cylinders that can be followed by archivists of limited technical means is described. The resulting transcriptions represent unprecedented performance at minimal cost for serious audio archival work.

INTRODUCTION

A tremendous audio legacy rests in wax cylinder archives throughout the world. These cylinders, fragile and subject to organic decay, are often in an advanced state of deterioration. Attempts to play these cylinders on restored period phonographs often result in serious wear or outright destruction of the original media, and rarely yield high quality transcriptions. BURP-ONE was developed to fill the need for an inexpensive high-performance cylinder transcription system. The purpose of this paper is to facilitate cylinder playback by the general public and to encourage cylinder restoration efforts.

Several elegant cylinder reproduction machines have been constructed [1]. These designs are intended to minimize mechanically-related defects in the transcription chain. Unfortunately, these custom-made, hand-tooled devices are inherently expensive [2]. At the other end of the spectrum, some antique phonograph makers rely upon vintage hardware, and place less reliance on modern technology in the anticipation of capturing an “authentic” sound. Unfortunately, most wax cylinders transcriptions are performed using period machines, many a century old. These old phonographs mechanically couple the reproducer stage to a drive screw that pulls the reproducer across the face of the cylinder. If the cylinder thread pitch differs from that of the player, the cylinder will be seriously damaged.

BURP-ONE is designed to make quality wax cylinder transcriptions at minimal cost, through simplified hardware requirements. We make the assumption that signal processing will be used to correct defects found in both the cylinder and the transcription chain itself.

The paper is organized as follows. First the transcription hardware is described. Choice of cartridge, stylus, and transcription system tradeoffs are considered. Next, typical defects encountered in wax cylinder work are reviewed, along with potential treatments. A signal processing protocol is described that can yield high-quality transcriptions.

TRANSCRIPTION HARDWARE

The BURP-ONE system consists of the following components:

- Elongated tone-arm
- Belt-driven turntable platter
- Mandrel and base plate
- Variable-speed D.C. motor
- Magnetic cartridge
- Glass or conical diamond 3-mil stylus
- Signal processing workstation

Most of the components are salvaged from a conventional modern belt-driven phonograph. The tone-arm pivot bearing and platter bearing should be removed from the phonograph while keeping the motor and platter support subassemblies intact for subsequent mounting on an aluminum chassis.

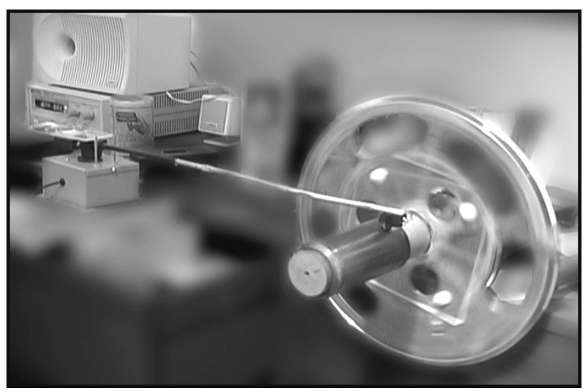


Figure 1: BURP-ONE transcription system in action.

Tone-arm

A linear tracking mechanism is preferred to minimize tracking error. Although some successful work has been done in adapting commercial record players to this application [3], they are complex and the raw materials are relatively scarce. A much simpler solution is to use a pivoting tone-arm, with an extension made from thin (≈ 3mm) aluminum tubing. The choice of tone-arm length is a tradeoff between angular tracking error and tracking force, which should be limited to three grams or less. The overall length of the BURP-ONE tone-arm from pivot to cartridge is 450 mm. *Very* thin wire (28-32 gauge) should be used in the tone-arm wiring to minimize tone-arm mass.

The additional mass of the extended tone-arm requires a corresponding counterbalance. To make a simple yet effective counterbalance, attach a small piece of acrylic plastic sheet to the top of the tone-arm's existing counterbalance using two-sided adhesive foam tape. Place a second strip of foam tape along the top of the counterbalance. The adhesive foam is a good surface for holding counterweights (coins work well) in place.

The tone-arm assembly is mounted to an aluminum box. The delicate wires coming from the tone-arm (again, use *very* thin wire to minimize tone-arm mass) are soldered to an internal terminal block, which serves as the mounting point for the shielded low-level audio output cable.

The longer tone-arm may introduce hum into the audio chain. The hum can be reduced by loosely looping a few turns of thin (24-26 gauge) wire around the base of the tone-arm (near the pivot) and attaching it to the pre-amplifier ground.

Platter and base plate

Construct the mandrel base plate from a 150 mm square piece of acrylic plastic. Drill a 12 mm hole in the center to accommodate the turntable center spindle. Drill a 12 mm hole at points 20 mm in from each corner. Roughen the center 50 mm of the base plate to ensure a good adhesive contact surface for the mandrel.

Using the base plate as a template, center it on the platter, mark the four corner locations and drill a 10 mm hole at each point. Using epoxy resin, attach the head of a 12mm nylon screw at each point from the bottom of the platter, with the threads oriented towards the top of the platter (to receive the mandrel base plate).

Remove the platter and its bearing assembly from the donor phonograph and install it vertically on an aluminum chassis. Ensure that the platter and motor pulley are in alignment for the drive belt, which should be placed around the platter subassembly before fastening the subassembly to the chassis. Note that the chassis may need some counterbalancing to offset the weight of the vertical platter.

Mandrel

The nominal dimensions of a typical wax cylinder are 105 mm in length with an outer diameter of 56 mm. The inner bore is a roughly linear taper with the inside (start of recording) diameter of 46 mm and an outside (end of recorded area) diameter of 43 mm. Depending on variables such as cylinder manufacture, age, and condition, these values will vary somewhat. To accommodate a wide range of cylinders, the mandrel should be machined to a length of 140 mm with a linear taper from 42 mm to 47 mm. A lightweight metal such as aluminum works well, although a fairly dense wood such as mahogany is easier to work and yields good results.

An inexpensive mandrel can be easily constructed using a 140 mm length of 40 mm (or 1.5") diameter acrylic tubing. Ensure that the tube is cut with at least one end as perpendicular as possible; this side will be cemented to the center of the base plate using epoxy resin. Apply adhesive foam strips at 90 degree intervals along the length of the tube. Keep the protective wrapping tape on the exposed strips. Wrap additional foam strips over the longitudinal strips at both ends end. This provides a good approximation to the cylinder taper, with the advantage that the soft foam mandrel reduces the possibility of shattering a cylinder when sliding it on the mandrel. The plastic mandrel and base plate is shown in Figure 2.

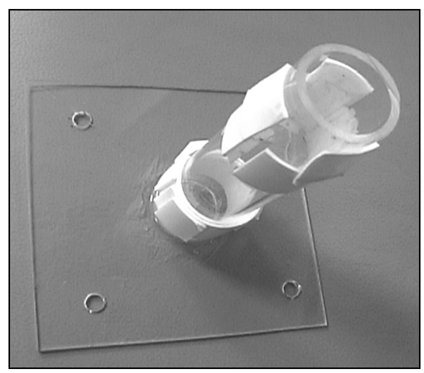


Figure 2: Inexpensive plastic base plate and mandrel.

Motor

Commercial cylinder recordings were made at a nominal speed of 160 rpm. Many recordings, especially those made in the field or at home, were recorded at considerably lower speeds (perhaps 140 rpm). Thus a variable-speed motor is desirable for first-order pitch correction.

Many modern record players use a variable-speed D.C. motor that drives the platter via an elastic belt. The speed can be easily adjusted using an variable (0-15V D.C.) power supply. Mount the motor/platter assembly vertically onto an aluminum chassis using the motor's elastic shock-mounting material. Transcriptions are very susceptible to mechanical noise and vibration since cylinder recordings are amplitude modulated.

Transcription work at 160 rpm can be difficult, particularly with a cylinder that is warped, cracked, or has other gross defects. In these cases the half-speed rate (80 rpm) is a good compromise between improved trackability and extended transcription time. Particularly damaged cylinders may require slower playback, at quarter-speed (40 rpm) or an intermediate rate such as

$$\frac{16{,}000\,sps}{22{,}050\,sps}*\frac{160\,rpm}{2}=58\,rpm$$ [1]

Pitch is restored by sampling at 16,000 and processing at 44,100 sps. For greater processing efficiency the audio stream is down-sampled to 11,025 sps before processing.

An effective way to improve transcription quality is to perform cylinder playback *backward* (counter rotation). This approach has several benefits:

- The vertically-encoded modulation will often be less worn on the "back" side of the modulation than the "front" side that the stylus first encounters on normal playback. Closer tracking of the relatively undamaged modulation leads to better transcriptions.
- The cylinder "lead-in" grooves are often the most damaged part of the recording, and are subsequently difficult to track. Transcribing these grooves backward from the end to the beginning will often yield usable results where conventional playback would fail.
- There is less physical damage to the cylinder in the region of the lead-in, which often contains valuable material such as self-announcements and performance details.

Cartridge

Cylinders were recorded with vertical ("hill-and-dale") modulation, actually the mechanical imprint of the recorded acoustic wave. A conventional modern stereophonic magnetic cartridge with a suitable stylus can yield excellent results, but it must be properly wired for monophonic playback of the vertically-encoded signal. This is accomplished by connecting the stereo left- and right-channel outputs together, with the monophonic vertical signal now taken from left-channel ground (+) and right-channel ground (-). Figure 3 shows monophonic wiring for vertical and, for reference, lateral cut recordings. This wiring sums the output due to vertical stylus motion, while suppressing the horizontal component, providing true monophonic playback. The MONO switch found on a stereo amplifier, which usually connects the two channels in parallel, will produce a noisier transfer.

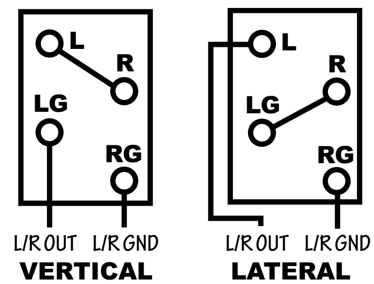


Figure 3: Cartridge wiring for true monophonic playback.

Connect the cartridge output to a suitable A/D converter via shielded cable. 16-bit sampling is adequate. A conventional phono pre-amp works well, but the signal will require subsequent equalization. For convenience, the preamplifier analog output can be fed directly to the analog input of a quality PC soundcard. For higher-quality work, it is better to record directly into a DAT or other standalone recorder, transferring the signal to the workstation via optical link.[2] As a practical matter, many personal computer sound cards use the left input for monophonic recording.

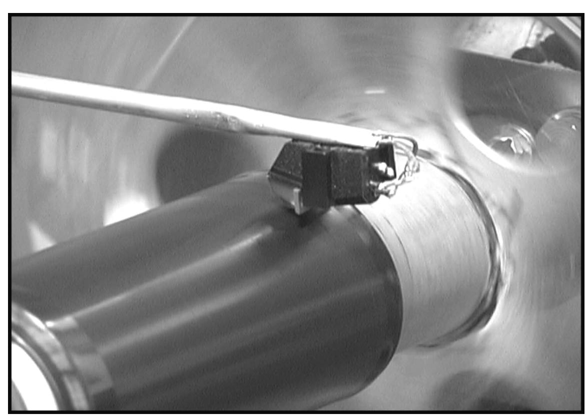


Figure 4: Reverse-mounted cartridge on tone-arm. Note cartridge position relative to cylinder.

[1] Vertically-cut Edison Diamond Discs can be played on many recent turntables that are equipped with speed-adjustment controls. Set the turntable speed to 45 rpm. If possible, set the pitch adjustment potentiometer (usually located on the bottom of the chassis) to a nominal speed of 58 rpm. Sample at 16,000 sps, then process the PCM stream at 22,050 sps to restore original pitch.

[2] Set the DAT to 32,000 sps and work at 58 rpm.

Attach the cartridge to the end of the tone-arm using a *small* amount of cyanoacrylate adhesive, so that the cartridge can be snapped off later if necessary. The backward cartridge orientation is shown in Figure 4. Note that the stylus touches the cylinder roughly 10 mm ahead of the customary vertical tracking point. This provides better tracking on difficult cylinders.

Stylus

Later-era mass-produced cylinders such as the Edison Gold-Molded and Amberol records are fairly hard and reproduce well using a commercial 3-mil diameter, truncated conical diamond stylus. A larger diameter (4-mil) stylus is preferred for use on very early (19^{th} century) cylinders, which are softer and have wider grooves.

Brown wax cylinders, in particular those formulated for home recording, are extremely soft and easily scratch. In no case should these cylinders be played on period equipment. A custom glass-bead stylus [4], which is attached to the shank of the original stylus using a cyanoacrylate adhesive[3], yields better results with less wear to the recording media.

Making a glass stylus requires some trial and error, but the materials are inexpensive. Start with a supply of crushed glass. Select two sharply-pointed pieces and, holding each with a heat-proof tool, touch the two ends together and place them over a hot flame; a gas stove will do. When the tips of the two pieces start to melt, allow the tips to fuse together. Then, simultaneously pull the two pieces away from the flame and quickly draw the two pieces apart, to a distance of about one-half meter. The cooling glass will produce a “glass wire.” Select a very thin portion of the glass wire and cut a 3 mm portion. Holding the wire at one end high above the flame, slowly lower it towards the heat. When the tip starts to glow, a molten ball will form. Immediately remove it from the heat. Place the finished glass stylus in a high-contrast, clean container; a finished stylus is very hard to see. One can quickly make several glass styli in this manner.

Now take a conventional diamond-tipped magnetic stylus and cut the diamond tip from the shank. Attach the glass stylus to the shank using cyanoacrylate adhesive. To facilitate this step, it is helpful to mount the cartridge on an adjustable vise. Place the stylus *tip* in a piece of silicon putty, then move the stylus shank to touch the top-half of the glass stylus. Apply adhesive using a needle. When the adhesive has hardened, the stylus can be safely lifted from its putty base. The residual putty will be removed after a few minutes of use.

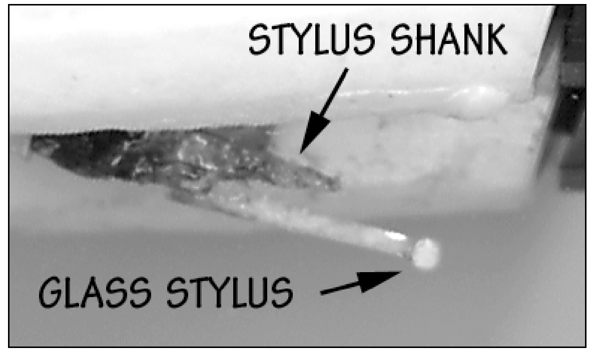


Figure 5: Mounted glass-bead stylus.

The stylus attached to the cartridge stylus shank is shown in Figure 5. The resulting cartridge/stylus is quite rugged and will give several hours of service.

Signal processing workstation

A fast computer (clock > 1 GHz) is preferred to run modern restoration algorithms with reasonable performance. Several audio processing packages with credible denoising modules are available for various operating systems. BURP-ONE uses additional experimental denoising modules that were written in MATLAB. Perl scripts were used to interconnect MATLAB algorithms with other denoising processes.

CYLINDER TRANSCRIPTION PREPARATION

Once assembled, the transcription system should be calibrated and tested. Mandrel speed for the desired rate(s) can be set by noting the D.C. voltage on the power supply corresponding to the desired rpm. We recommend 40, 58, and 80 rpm, where the slower rates are used for cylinders in poor condition. If the platter spins in the wrong direction, simply reverse the power supply polarity.

Accurately mounting the mandrel requires some patience. To avoid annoying “wow”, the mandrel must be perfectly concentric with the platter, and perpendicular to the base. Place the chassis so that the turntable platter is horizontal. Using a felt-tipped marker, draw a 45 mm diameter circle on the platter centered on the spindle. Slip the mandrel/base plate assembly over the nylon screw threads. Center the mandrel by aligning the mandrel base to the circle just drawn on the platter. Finger-tighten the nuts and apply power to the motor. Spin the platter at about 60 rpm and, holding a finger stationary, gently prod the mandrel end until it is concentric. Stop the motor and tighten the nuts. Verify the installation, then draw an outline of the mandrel base-plate on the platter. This will facilitate re-mounting the mandrel in the future. If the

[3] A gel-type formulation for joining rough surfaces is ideal.

mandrel is not normal to the base (the end will wobble), insert thin paper shims under the base plate until the problem is corrected.

Tracking force depends on cylinder material and stylus diameter. Three grams is a good starting point. Eccentric cylinders may require heavier tracking, but if more than four grams is required, it is probably better to choose a slower playback speed.

Since cylinder programs are only two to four minutes long, it is convenient to capture a series of cylinders in one session, then perform initial post-processing (time reversal and rate normalization) in one operation.

Cylinders are very fragile. It is important to remember that it is surprisingly easy to shatter a cylinder, particularly when sliding it onto the mandrel. If excessive pressure is used, a lateral hairline crack will suddenly run up the length of the cylinder.[4] Hold a cylinder touching only its non-recorded surfaces. Moving a cylinder by placing spread fingers inside the cylinder and lifting may shatter it.

The organic wax cylinder, housed in its paper storage tube with felt lining, fosters the growth of mold, which eats the wax. In minor cases the mold appears as small (1-10 mm) circular patches, which impart an annoying periodic scratch to the recording. In more severe cases the cylinder may be completely enveloped in mold and irrevocably damaged.

Cylinders should be carefully cleaned before playback. A commercial detergent designed to remove organic deposits from medical or lab equipment is ideal for removing mold. Mix a batch of sudsy detergent preferably using distilled water at 30° C and soak the cylinders, labeled-end up, in the solution for fifteen minutes. Gently wipe the cylinders with a soft *clean* sponge, rinse, and allow them to thoroughly air dry.

Sampling rate tradeoffs

Cylinder audio fidelity is limited by the resonance of the recording horn and recording head inertia. The usable frequency response is roughly 200 to 4,000 Hz. A computationally-efficient sampling rate common to PC hardware is 11,025 sps. Good results are obtained using this sample-rate, but with *half-speed* (80 rpm nominal) playback. The quantized audio will be processed at 22,050 sps to restore original pitch. Regardless of the sampling rate used for audio capture, we prefer to sample-rate convert the audio stream to 11,025 sps, 32-bit floating-point before performing restoration work. The lower sample rate preserves the audio content, while allowing greater efficiency in the signal processing that follows. One should always make a backup of the raw PCM audio data, as restoration is a subjective process, subject to errors, revisions, and improvements in the art. Every restoration is in essence a "work in progress", so it is wise to preserve the original data.

The final sampling rate used for the restored audio will depend on the target application. Good quality MP3 encoding at low sample rates such as 11,025 sps requires a fairly clean transcription. The noisy background common to cylinders is a challenge to compress, resulting in unpleasant compression artifacts. In theory one could store over 5,000 two-minute cylinders on a single audio CD, but the audio quality would be compromised.

CYLINDER TRANSCRIPTION PROTOCOL

Restoration of heavily degraded cylinders requires some or all of the following processing steps, to be performed in the following order:

- Treating thumps and mold noise
- Removing impulsive noises such as clicks and pops
- Suppressing continuous noise (hiss)
- Equalization and dynamic range expansion
- Pitch correction

Denoising

Cylinders are corrupted by a number of noise processes ranging from hairline cracks or even repaired breakages to periodic disturbances such as thumping or scratchiness caused by mold. Wide-band noise is particularly troublesome, especially on worn cylinders. An excellent overview of the denoising process is given in [5]. Here we focus on practical aspects of denoising wax cylinders.

Thumps

Thumps arise from cartridge non-linearity and when large scratches are encountered. An eccentric cylinder will cause the stylus assembly to periodically accelerate as the cylinder revolves. The resulting "thump" is a low-frequency perturbation from the violent motion imparted to the stylus. When the stylus encounters a large scratch, the resulting "thump" approximates the system's impulse response. A template of this response for a particular tone-arm/cartridge/stylus combination is made from a col-

[4] Fortunately, these long hairline cracks are often easy to repair using a small amount of cyanoacrylate adhesive, de-thumping software and a bit of patience.

lection of thump samples. The template is then aligned and subtracted from the audio. This effectively suppresses the low-frequency thump component. Two major challenge with this approach are the creation of a template library and applying the template without leaving residual clicks.

In mild cases the thumping sound is effectively removed during the frequency equalization step, which suppresses low-frequency signals. More severe cases may benefit from transcription at a lower speed. Wavelets can also be used to create “thump markers”, with subsequent interpolation in the spectral-frequency domain. This is demonstrated in Figure 6.

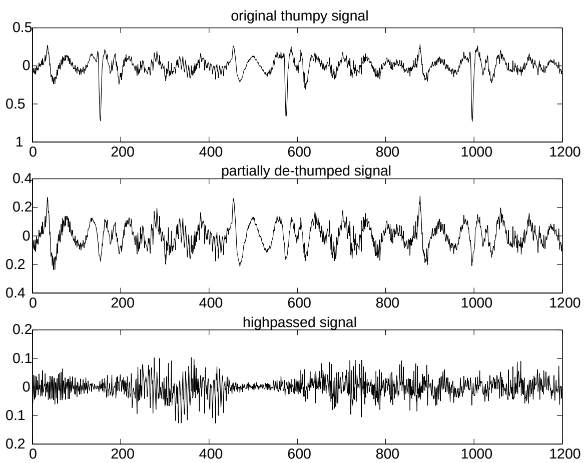


Figure 6: Thumps, partially removed thumps, and subsequently high-passed audio. Note scale change.

Mold noise

Mold is difficult to treat; not only have the grooves been destroyed, the mold patches may span tens or even hundreds of milliseconds. Paradoxically, mold damage is fairly difficult to identify due to its broadband “scratchy” nature, which is difficult to separate from genuine content spectra.

One approach that shows promise is to use a higher-order wavelet from the symlet family to generate “mold markers” from squared detail coefficients which are subsequently thresholded. The mold regions thus identified are then interpolated using a wide-area interpolation technique, such as spectral-time processing using Fourier analysis [6].

Figure 7 shows a portion of a mold-damaged cornet solo (top), the corresponding mold-marker obtained from the squared symlet order 8 detail coefficients at level 6 (middle), and the solo with mold suppressed (bottom). An iterative algorithm determines the thresholds to be used. The markers are qualified using a peak detection algorithm.

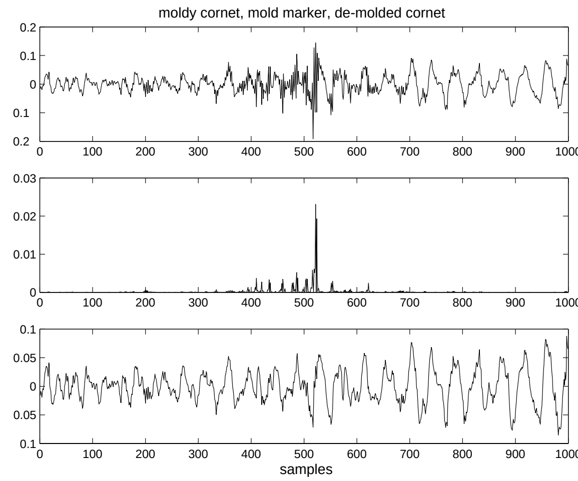


Figure 7: Mold-corrupted cornet solo, mold marker, and resulting mold-suppressed signal.

Clicks and pops

Removing clicks, pops, and other impulsive noises presents two challenges. The first is to accurately identify the location of the impulse events. The second is to eliminate them through a suitable interpolative process.

Wavelet analysis at fine resolution is excellent at locating singularities such as clicks [7]. Wavelets, however, are computationally demanding. Autoregressive (AR) models [8], in conjunction with Bayesian estimation [9], have proven to be particularly successful [10][11]. These techniques have been successfully used in commercial software, often at modest cost.

After the gross impulsive defects have been removed, a more difficult problem remains: eliminating the residual small clicks or “crackle” that remains after a first-pass de-clicking process. It is hard to separate the low-level clicks from genuine program content. Aggressive (lower detection thresholds) de-clicking using algorithms effective for more obvious defects has a tendency to “scrub” the audio, leaving a hollow-sounding remnant of the original audio. On the other hand, conservative de-clicking leaves annoying noise unchecked.

An interesting alternative approach is to perform aggressive de-clicking, but instead of processing the audio directly, we process the residual noise signal. The noise signal is obtained by a Wiener filtering operation in the spectral domain, as described under hiss suppression. The impulsive clicks, containing wide-spectrum energy, appear in the noise residual. More aggressive de-clicking can be performed on the noise residual, yielding better crackle suppression with less compromise to the program audio. The de-crackling procedure is as follows:

- Obtain a fairly accurate estimate of the audio noise floor, using the guidelines discussed. *The quality of this estimate determines the effectiveness of this technique.*

- Separate the noise from the audio using a Wiener filter in the spectral domain. Save the resulting residual noise vector **x**.

- Process the noise residual **x** using an aggressive setting of the de-clicking tool, to yield the de-clicked noise residual vector **y**.

- Sample-wise subtract[5] the original noise residual **x** from the de-clicked noise residual **y**. The resulting click mask **z** represents an inverted estimate of the residual clicks not detected in the original de-click phase.

- Superimpose (sample-wise add) the click mask **z** to the original signal. This step effectively removes low-level crackle with minimal impact to the underlying audio.

Figure 8 shows a typical click-mask resulting from these operations. Note that the amplitude of the click-mask is fairly low (≈ -60dB relative to the de-clicked audio). However, removing these residual clicks results in a noticeable improvement to the sound.

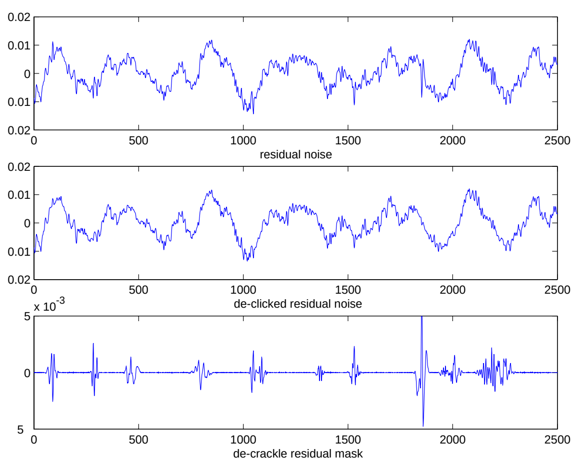


Figure 8: Residual noise, de-clicked residual noise, and resulting difference click-mask.

In some cases performing a second de-clicking pass on the *time-reversed* audio, using the same parameters as the first pass, may provide additional benefit.

[5] or alternatively, add the inverted vector **x** to **y**

Wide-band (hiss) denoising

Perhaps the most popular technique for reducing wide-band noise is based upon Wiener filtering in the spectral domain [12]. The sample stream is processed in overlapping, windowed blocks. Each block is multiplied in the frequency domain by the noise reduction function $f(\Psi)$. The Wiener filter that minimizes the mean-squared error of the reconstructed signal in the time domain is given by

$$H(\omega) = \frac{S(\omega)}{S(\omega) + N(\omega)}$$

where $S(\omega)$ is the signal power spectrum and $N(\omega)$ is the noise power spectrum. The success of this filtering operation is obviously highly dependent upon an accurate estimation of $N(\omega)$. The noise spectrum is derived from sampling a section of "quiet" audio which is assumed to have no program content. In restoration work the mean of $N(\omega)$ is seldom stationary, and its variance is not, so multiple noise estimates may need to be taken and subsequently applied in the denoising process across the recording.

Care must be taken when selecting the noise sample estimate. Lead-in and lead-out grooves often have impulsive noise that leads to an inaccurate noise estimate, causing a so-called "musical noise" with its characteristic "burbly" or "spacey" overtones. Attempts have been made to minimize this effect [13][14]. In practice, a reliable noise estimate should be taken, if possible, from a known truly "silent" passage within the program material. The selected passage should be de-clicked before analysis to improve the estimate's reliability. At 11,025 sps, a DFT size of 4,096 or 8,192 gives good results.

The quality of the noise estimate can be subjectively determined by auditioning a segment of the denoised audio. Be careful to avoid the common mistake of focusing on the eliminating noise component thus aggressively denoising for the sake of obtaining a "clean" transcription. To verify a noise estimate, an effective trick is to apply the denoising process, but retain and audition a copy of the *noise* component. With a good noise estimate, very little if any of the program content will be heard in the noise component. Conversely, with a good noise estimate, the denoised audio will retain its original musical quality.

Once a good noise estimate is obtained, save it as a template. A library of templates, one for each combination of cylinder type, cartridge, and stylus will prove useful later when transcribing a troublesome cylinder where no accurate noise sample can be found.

Wavelet Denoising

Wavelets have been studied extensively for their value in denoising in audio and image processing applications. Berger et. al. [15] demonstrated that wavelets can provide significant denoising performance, in this case treating an extremely noisy cylinder recording of Brahms. Wavelets are particularly effective for denoising recordings of the spoken word.

Several commercial wavelet packages with denoising functionality are available. The MATLAB wavelet toolbox is well-known. Donoho and his colleagues at Stanford University have developed a wavelet toolbox [16] for MATLAB that is currently available on the web at no cost. Taswell has developed the WavBox wavelet suite [17], also for MATLAB.

Wavelet denoising is computationally more expensive than spectral subtraction, but can yield better results in many cases. However, the resulting audio may sound "stripped" or artificial unless one chooses the coefficient thresholds with care. Using a higher-order wavelet ($n \geq 8$) tends to reduce these audible artifacts.

While there are automated wavelet denoising methods using various noise and thresholding models, better results are often possible by manually setting the thresholds at each level of decomposition. In practice, a relatively small sample of representative audio is auditioned and the thresholds are manually set. The threshold settings are then applied to the entire audio file as a series of blocks.

Pitch variation corrections

The cylinder's small diameter makes it susceptible to pitch variances due to physical distortion of the media. Rotating at 160 rpm, an annoying pitch fluctuation with a period of roughly 375 ms results.

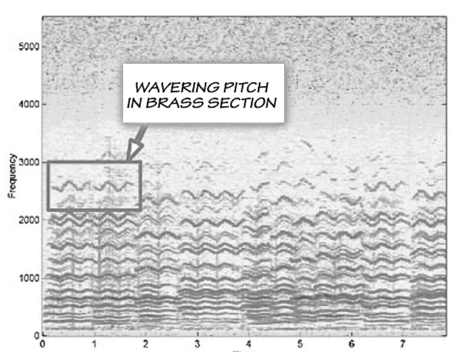


Figure 9: Pitch variations in spectrogram.

Pitch correction requires estimation of the pitch frequency θ, phase ϕ, and amplitude A. Once estimated, the signal can be resampled using a warping function dw(t) based on the pitch defect estimate (θ, ϕ, A). We use an algorithm that broadly follows that developed by Godsill and Rayner [18], and inspired by an application of non-uniform sampling described in [19]. However, we make a simplifying assumption that the gross pitch defects are primarily due to physical distortion of the cylinder. In this case the pitch analyzer examines the harmonic structure of the audio, and extracts pitch information only where it exceeds a confidence threshold (see box in Figure 9). These reliable pitch estimates $(\theta, \phi, A)_n$, taken at frame n, form the basis for the de-warping function dw(t). Intermediate values of (θ, ϕ, A) where a reliable estimate cannot be found are obtained through linear interpolation. Finally, the PCM audio is resampled using bicubic interpolation. An illustrative example of a de-warping function (at exaggerated scale) is shown in Figure 10.

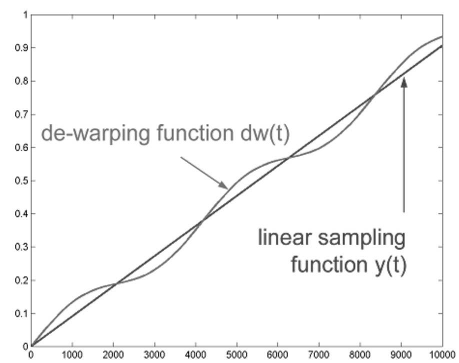


Figure 10: De-warping function.

Using the interpolated de-warping function, the signal is resampled using bicubic interpolation. A spectrogram of the pitch-corrected brass section is shown in Figure 11. The pitch defect has been effectively suppressed.

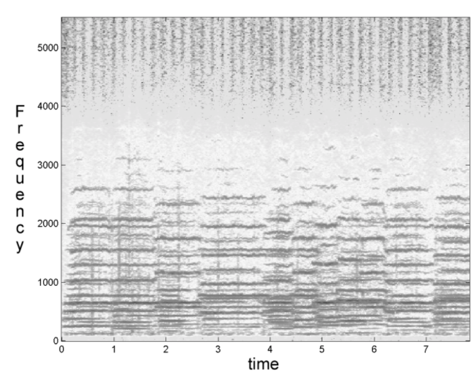


Figure 11: Pitch-corrected brass spectrogram.

The periodic vertical lines at the top of the spectrogram represent a residual pitch variation that is visible as a frequency modulation of the high frequency noise. This artifact resulted from an imperfect estimate of the pitch variation defect. The pitch-corrected audio sounds more "focused."

Frequency response and dynamics restoration

A commercial phonograph preamplifier incorporates a pre-equalization curve that is not ideal for cylinder transcription work. The curve shown in Figure 12 compensates for preamplifier pre-equalization. This equalization suppresses low frequency rumble and high frequency noise, yielding a more natural sound.

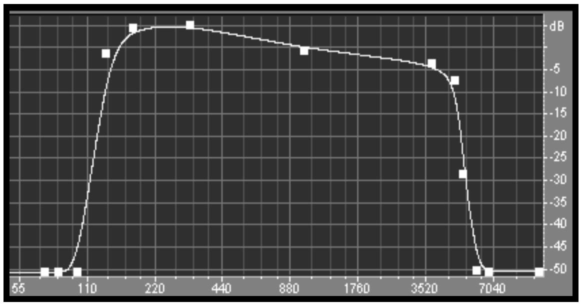


Figure 12: Cylinder frequency equalization curve.

The subjective quality of a cylinder transcription can often be further enhanced by expanding the relatively limited dynamic range characteristic of cylinder recordings. In this case a mild dynamic expansion, band-limited to about 200-2,400 Hz is applied. Limiting dynamic expansion to this range minimizes amplification of spurious low frequency rumble and high frequency noise. Figure 13 shows a typical curve that gives good results. Signals at –20 dB are boosted approximately 6 dB.

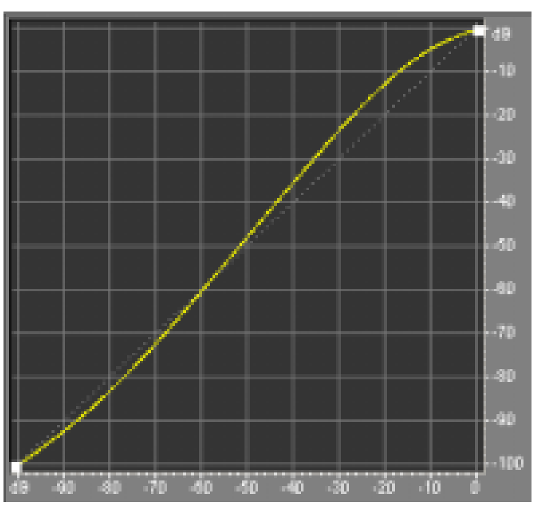
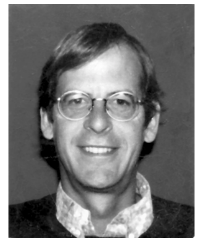


Figure 13: Dynamics restoration curve.

Low-energy signals are reduced a few dB as well. The input reference signal is roughly –13 dB average and –3 dB peak.

THE AUTHOR



James Nichols is a personal computer industry pioneer with 25 years experience in multimedia and communications signal processing with firms such as Apple Computer, Rearden Steel Technologies, Hewlett-Packard, Telebit and Advanced Micro Devices, where he is a Senior Fellow involved with software radio and computer architecture. He is active in audio restoration through ifixaudio, which specializes in historical audio preservation. He holds numerous patents in the fields of audio-visual and communications signal processing. He earned an M.S. in Computer Science at the University of California, Los Angeles. His professional interests include restoration of archival audio and photographs, classical music recording, acoustics, and communications.

REFERENCES

[1] http://members01.chello.se/christer.hamp/phono/
[2] http://members01.chello.se/christer.hamp/phono/chamoux.html
[3] http://members01.chello.se/christer.hamp/phono/hamp.html
[4] http://members01.chello.se/christer.hamp/phono/chamouxstyl.html
[5] S.J. Godsill and P.J.W. Rayner, *Digital Audio Restoration*, 1998.
[6] S. V. Vaseghi, *Advanced Digital Signal Processing and Noise Reduction (second edition)*, pp. 326-328, Wiley, 2000
[7] M. Misiti et. al., *Wavelet Toolbox for use with MATLAB*, 1996
[8] J. Makhoul, Linear prediction: A tutorial review, *Proc. IEEE*, 63(4):561-580, 1975
[9] S.J. Godsill and P.J.W. Rayner, A Bayesian approach to the restoration of degraded audio signals, *IEEE Trans. On Speech and Audio Processing*, 3(4), pp. 267-278, July 1995
[10] S. V. Vaseghi, and R. Frayling-Cork, Restoration of old gramophone recordings. *J. Audio Engineering Society*, 40(10), Oct. 1992
[11] S.V . Vaseghi and P.J.W. Rayner, Detection and suppression of impulsive noise in speech communication systems, *IEE Proceedings, Part 1*, 137(1):38-46, February 1990
[12] N. Wiener, *Extrapolation, Interpolation and*

Smoothing of Stationary Time Series with Engineering Applications, MIT Press, 1949

[13] O. Cappé, Elimination of the musical noise phenomenon with the Ephraim and Malat noise suppressor. *IEEE Trans. On Speech and Audio Processing*, 2(2):345-349, 1994

[14] O. Cappé, and J. Larouche , Evaluation of short-term spectral attenuation techniques for the restoration of musical recordings, *IEEE Trans. On Speech and Audio Processing,* 3(1): 84:93, January 1995

[15] J. Berger, R.R. Coifman and M.J. Goldberg, Removing noise from music using local trigonometric bases and wavelet packets, *J. Audio Engineering Society*, 42(10), Oct. 1994

[16] http://www-stat.stanford.edu/~wavelab/

[17] http://www.toolsmiths.com/wavbox.htm

[18] S.J Godsill, Recursive restoration of pitch defects in musical recordings, *Proc. International Conference on Acoustics, Speech and Signal Processing*, Vol. 2 pp 233-236, Adelaide, April 1994

[19] D. Koenig, Application of time–frequency analysis for optimum non–equidistant sampling of automotive signals captured at knock. *Proc. Intl Conference on Acoustics, Speech and Signal Processing*, V, p. 2746-2749, 1996