Audio Engineering Society
Convention Paper
Presented at the 117th Convention
2004 October 28–31 San Francisco, CA, USA

Correction of Wow and Flutter Effects in Analog Tape Transfers

Jamie Howarth[1] and Patrick J. Wolfe[2]

[1] *Plangent Processes, 58 Miacomet Avenue, Nantucket, MA 02554-4354, USA*

[2] *Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK*

Correspondence should be addressed to Jamie Howarth (`info@plangentprocesses.com`)

ABSTRACT

Here we describe a system whereby analog hardware is combined with the theory of nonuniform sampling in order to correct for wow and flutter effects in analog tape transfers. We show how in certain instances the medium itself can provide an accurate measurement of a recording's timing irregularities, in which case digital signal processing techniques permit a playback-rate correction of what is essentially an irregularly sampled audio waveform. Results using both real and synthetic data demonstrate the effectiveness of the method, both in cases of severe degradation as well as high-quality analog transfers heretofore considered normal.

1. INTRODUCTION

Implicit in the nature of most analog-to-digital transfers is the assumption that the actual recording and playback mechanism is operating at a constant rate; however, due to a variety of mechanical considerations (as well as potentially other causes), this assumption is only approximately valid in practice. While work in the past has focused on the correction of either severe speed variations or those exhibiting a regular structure (such a gramophone disc with an off-center spindle hole; see, e.g., [1, Chapter 8]), we consider here the potentially more subtle and yet arguably more widespread problem of wow and flutter in analog tape media. Using both artificially degraded and commercial-quality recordings as examples, we show how artifacts present in tape media may be utilized to correct for such defects during the process of digitization.

2. PRINCIPLES OF OPERATION

We first provide a brief overview of the underlying principles involved in a restoration system for affected media. A principal step in any such procedure is to extract accurate information about the timing variations involved, using either the recorded signal itself or—if possible—artifacts of the recording process. Indeed, such information may manifest itself via ultrasonic signals such as bias or logic noise in the case of magnetic tape, or (for example) the "chip whistle" phenomenon found in certain 78-rpm wax/shellac recordings, wherein the cutter collects scrap material, causing a friction-based chatter to be impressed into the groove wall. In cylindrical recordings various motor rumbles, spindle eccentricities, and bearing noises present periodic information observable within the program material.

A key observation is that a sinusoidal signal (i.e., a pure tone), when subjected to timing variations upon recording or playback, will act as an ideal frequency-modulation (FM) carrier—with these same timing variations comprising the modulating signal. If such a carrier can be recovered intact as an artifact of the recording/playback process, and in a manner which preserves its time alignment with the audio program material, then there exists (at least in theory) the possibility of restoring the audio program completely.

While one may attempt to estimate global timing variations from the program material itself (see [1, Chapter 8] and references therein), we consider here the use of the bias signal accompanying the audio program material on an analog tape. In theory, such a signal serves as an ideal FM carrier, providing the necessary timing information when demodulated.

3. BASIC ASPECTS OF ALGORITHMIC AND HARDWARE DESIGN

In practice, however, the following difficulties may be encountered when attempting to correct for wow and flutter in analog tape transfers:

- The true bias frequency is unknown *a priori*, and may typically range from 40–435 kHz depending on the recording era and machine employed;
- Instability in the bias oscillator may not be easily distinguishable from the mechanical effects of wow and flutter;
- Edits and drop-outs affect the quality and continuity of the bias signal;
- Its amplitude may be extremely weak in comparison to the tape's background noise level; both the signal and noise levels may vary significantly over time.

To overcome the first of these difficulties, a heterodyning circuit may be employed to mix the bias signal down to an intermediate frequency suitable for analog-to-digital conversion (ADC). Such an approach minimizes storage requirements, and enables any subsequent signal processing—for instance, that needed to address the latter difficulties, as well as to correct for the timing variations—to be done in the digital domain. Fig. 1 illustrates this via a block diagram of the overall restoration system, from tape machine playback to final digital output.

In fact, capturing the necessary signals via tape playback is in itself a non-trivial task, as both audio program material and ultrasonic artifacts are read simultaneously from the same head gap in order to ensure synchronization. Basic theory on head geometry and magnetics being both well known and documented elsewhere (see, e.g., [2]), we do not detail it here; however, we note that the application at hand requires a high degree of isolation of the low-level artifacts from the much higher-level audio content. In turn, the filtering necessary to achieve the requisite out-of-band rejection requires an amplifier design incorporating a very large dynamic range.

By extracting information about the speed variations from the recording medium in this manner and then converting it to digital format, we are able to apply efficient digital signal processing techniques such as cubic Hermite, spline, or other polynomial interpolation methods to effect a playback-rate correction of what is in essence a nonuniformly sampled audio waveform [3]. We have found such techniques to be preferable to the windowed-sinc methods for sample-rate conversion currently described in the literature.

The interpolation procedure itself comprises the following steps (shown in Fig. 2):

1. Read in the input waveform and FM carrier waveform.
2. Demodulate the FM carrier waveform to obtain a speed variation function.

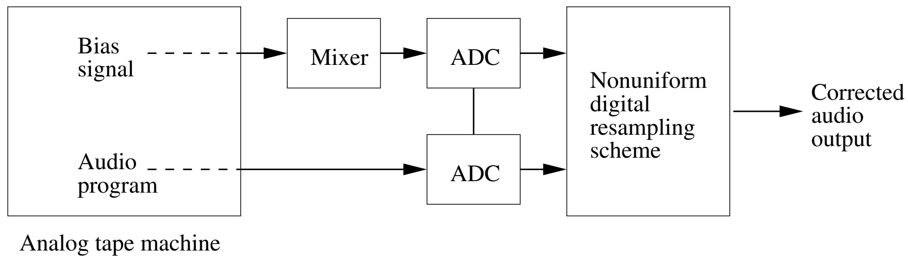

Fig. 1: A block diagram of the restoration procedure for wow and flutter correction

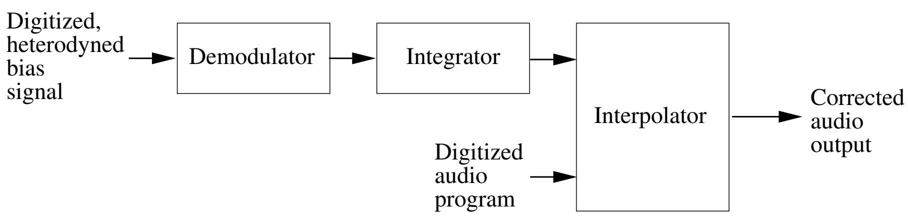

Fig. 2: An expanded block diagram detailing the nonuniform digital resampling scheme shown in Fig. 1

3. Integrate the speed variation function to obtain the time delay corresponding to a given sample point.

4. Interpolate between irregular samples of the input waveform, given at the times determined in step 3, in order to determine a set of output samples at a regular interval corresponding to the desired sampling rate.

4. EXPERIMENTAL RESULTS

We now consider the results of experiments performed both with artificially degraded recordings and existing commercial archive material, from which the efficacy of the proposed restoration method may be seen. In all cases, tape playback was conducted using an Ampex ATR 102 transport, rebuilt by Mike Spitz of ATR Services (York, Pennsylvania). Custom-built wide-bandwidth heads and electronics were employed to retrieve the audio and ultrasonic data, and analog-to-digital conversion was performed using state-of-the-art professional equipment, after careful level calibration.

In addition to the results presented here, audio examples and accompanying figures are available at the first author's home page, http://www.plangentprocesses.com.

4.1. Artificially Degraded Recordings

We first consider artificially degraded recordings, to which severe speed fluctuations were applied under controlled conditions via manipulation of the tape machine roller guides with a rubber-tipped pencil eraser. This manipulation was applied in real time during playback, at which time the signals were captured and subjected to the restoration scheme described in Section 3.

In the first such experiment, a relatively high-level, pure 12.5-kHz tone was first recorded onto magnetic tape running at a rate of 15 inches per second; upon playback the distorted output was captured and converted to a 96-kHz, 24-bit digital signal prior to processing. Figure 3 provides a before-and-after spectral comparison of a 10-second portion of the result, from which the successful correction of both small- and large-scale pitch deviations may be seen.[1] It is also apparent from Fig. 3, however, that the tone level is significantly above the natural noise floor of

[1] In this and all subsequent cases, spectrograms were calculated using Hanning windows with a length of 30 ms and an overlap of 50%.

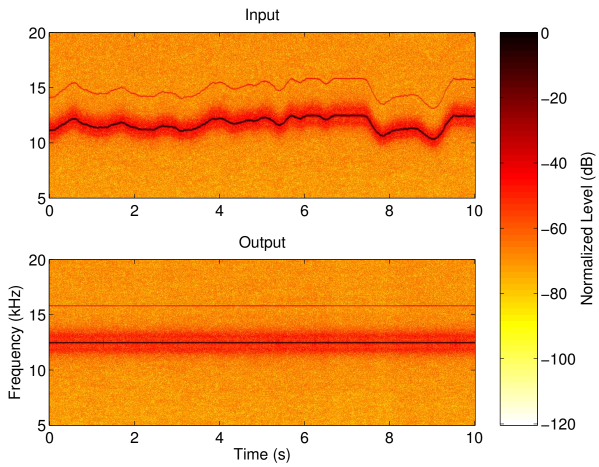

Fig. 3: Spectrograms of a 12.5-kHz tone subjected to extreme mechanical wow and flutter, shown both before and after processing (with slight bleed-through from the adjacent channel visible near 16 kHz)

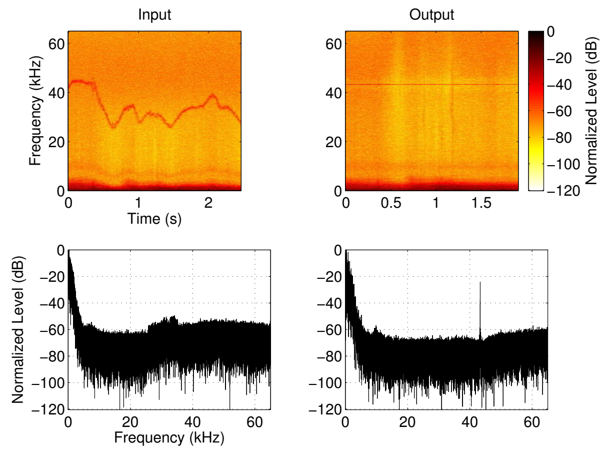

Fig. 4: Time-varying and average spectra of a tape segment with low-frequency audio content and bias tone at approximately 43.3 kHz, shown before and after processing

the tape medium—a situation which is not necessarily likely to be encountered in practice.

As a more realistic example, we next consider a tape containing both audio program content and an actual bias signal—but synthetically degraded according to the method described earlier. Input/output comparisons of a 2.5-s signal segment, via spectrograms as well as the discrete Fourier transform (DFT), are shown in Fig. 4. Successful recovery of the signal, even in the case of a low-level carrier tone, is evident in this figure.

4.2. Commercial-Quality Recordings

While damaged and otherwise severely degraded tapes are obvious candidates for the treatment described above, an often overlooked audible distortion in analog tape recording and playback is that of frequency modulation caused by periodic speed variations occurring at a higher frequency than what is typically considered to be flutter.

As is well known (see, e.g., [4]), flutter occurring at a rate above several hertz will cease to be perceptible as such, but instead will take on a character variously described as "roughness" or "fuzziness". While we do not undertake to provide a full-scale analysis of flutter profiles here, we do note that the specifications adopted by the Audio Engineering Society (AES) for the measurement of flutter range only to 200 Hz, while others have argued that high-frequency "scrape flutter" is more properly characterized in the region of 3–5 kHz [5]. Indeed, it has been the authors' experience that tapes suffering from fast flutter and scrape flutter share the sonic characteristics of classic FM distortion.

In fact, the AES standard specifies a frequency-weighted curve with a maximum at 4 Hz, which serves to discount flutter phenomena at higher (and lower) frequencies. As it became possible to better model the effects of high-frequency flutter in the 1960s, efforts were made to develop transports (such as, e.g., the 3M Isoloop) that minimized these. While the results of such improvements may have been audible in terms of noticeably reduced roughness, they were not necessarily reflected in the corresponding weighted measurement. However, the advent of new methodologies for modeling—and even mediating *ex post facto*—the effects of wow and flutter via digital signal processing suggests that the sonic attributes and measurable characteristics of high-frequency flutter be revisited.

Hence, as the next phase of enquiry we consider the application of the restoration methodology described above to typical "real-world" tapes heretofore con-

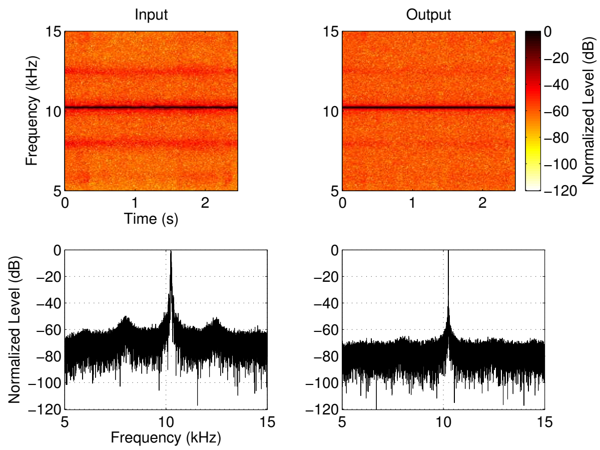

Fig. 5: Spectral comparison of The Association's *Insight Out* calibration tone before and after processing, showing a reduction in flutter-induced sideband amplitude

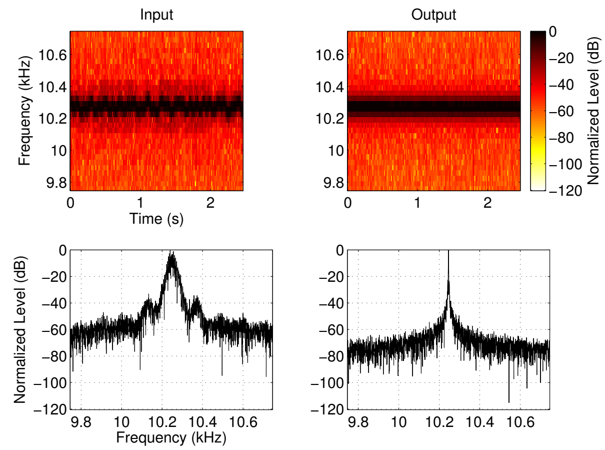

Fig. 6: A close-up of The Association's *Insight Out* calibration tone before and after processing

sidered normal. In order to demonstrate the measurable effects visually, we focus here on calibration tones found on existing recordings rather than audio program content. In each case the control signal used for restoration was derived from the tape's actual bias signal.

To this end, Fig. 5 represents a short segment of a (nominally) 10-kHz calibration head tone taken from side two of The Association's *Insight Out* album master (courtesy Warner Music Group), the bias frequency in this case being approximately 99 kHz. Here flutter may be seen in both the spectrogram and the DFT representations of the signal; in particular, the broad peaks at approximately 8 and 12 kHz appear to be scrape-related flutter rather than lower-frequency phenomena. Figure 6 shows a detailed view of the same segment.

As a second example, Fig. 7 shows the spectrum of a 10-second segment of a (nominally) 10 kHz head tone taken from the standard 35 mm SMPTE alignment tone test reel (the bias frequency in this case being approximately 67 kHz). The bottom row shows a detailed view of the same segment, in which flutter sidebands at ± 96 Hz are clearly visible (see grid lines). A common occurrence in film soundtracks, this sprocket-cogging flutter stems from the

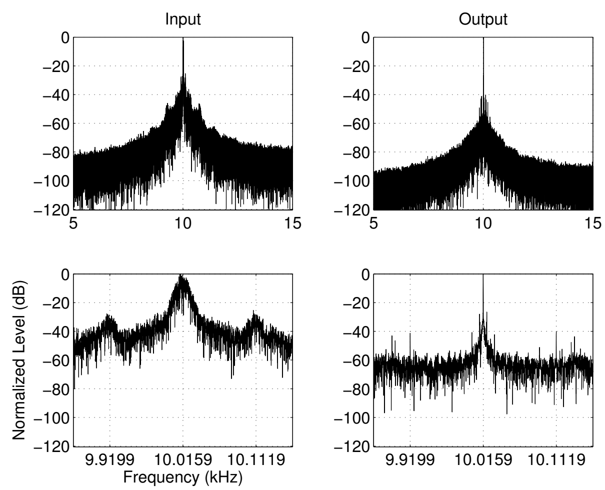

Fig. 7: Discrete Fourier spectrum of a 10-kHz SMPTE alignment tone before and after processing, shown on two different scales. Note the reduction in 96-Hz sprocket-cogging flutter, the sidebands of which may be seen to align with the vertical grid lines of the bottom row's figures.

multiple of 24 frames per second and 4 perforations per frame.

5. DISCUSSION

As it is difficult to convey the subjective effects of the proposed restoration scheme without direct reference to sonic examples, we limit our discussion here to some general observations on the subjective quality of the restorations attempted to date. To this end, it appears that the perceptual effects of the wow and flutter removal considered here correlate well with measurable improvements in signal quality. Indeed, subjectively the results align with what the measurements imply: not only a reduction in obvious pitch fluctuation, but also a significant reduction in the characteristic roughness caused by higher-frequency flutter. Overall the impression may be described perceptually as one of heightened transparency—with less interstitial "hash".

In addition, the removal of flutter modulation effects in the upper audio bandwidth appears to yield a perception of extended high frequency response, similar to that which can be observed in an earlier generation of the same material. We hypothesize that this may explain why a tape copy whose frequency response "by the tones" measures identically to the original may be perceptibly "duller" in character. Such dulling has often traditionally been assumed to be a function of the magnetic process, but we posit that such an effect may be due at least in part to the distortion and smearing of transient information caused by the transport's FM modulation of the material. While formal listening tests have yet to be conducted, listeners with many years' experience in analog recording have commented informally that when such distortion is removed, a significant portion of what has often been characterized as "generation loss" goes with it.

In summary, we have demonstrated here how analog tape recordings degraded by timing variations may be restored via the use of ultrasonic information characteristic to the medium, in a manner which enhances signal fidelity beyond the level of a standard analog-to-digital transfer. Having shown the feasibility of such a method in practice, we intend in future work to improve and extend the methodologies presented here in order to address other restoration scenarios.

ACKNOWLEDGMENTS

The authors would like to thank Dale Manquen of MANCO, Greg Orton of Flux Magnetics, Jay McKnight of Magnetic Reference Laboratory, Darryl Curley and Tim Rector of Custom Research, and David Smith of Sony Music for their invaluable advice and assistance.

6. REFERENCES

[1] S. J. Godsill and P. J. W. Rayner, *Digital Audio Restoration: A Statistical Model Based Approach*, Springer-Verlag, Berlin, 1998.

[2] M. Camras, *Magnetic Recording Handbook*, Van Nostrand Reinhold, New York, 1988.

[3] P. J. Wolfe and J. Howarth, "Nonuniform sampling theory in audio signal processing," Presented at the 116$^{\text{th}}$ Convention of the Audio Engineering Society, 2004, Preprint 6123, Audio Engineering Society, New York.

[4] "AES standard method for measurement of weighted peak flutter of sound recording and reproducing equipment," 1982, Standard AES6-1982, Audio Engineering Society, New York.

[5] D. Manquen, "A wideband tape and transport diagnostic method," Presented at the 66$^{\text{th}}$ Convention of the Audio Engineering Society, 1980, Preprint 1637, Audio Engineering Society, New York.