

# EVALUATION OF THE MAXIMAL MODULATION FREQUENCY FOR WOW AND FLUTTER DETERMINATION

P. MAZIEWSKI

Multimedia Systems Department,  
Gdansk University of Technology  
ul. Narutowicza 11/12, 80-952 Gdansk, Poland

Author of this paper participated in the development of several algorithms for wow and flutter characteristic determination. The algorithms were found useful in some real-life restoration procedures, however, there was no study of their capabilities in terms of the maximal determined modulation frequency. This paper presents such a study for three algorithms, i.e. the power-line hum tracker, high frequency bias tracker, and the centre of gravity tracker. In the paper a short description of the algorithms is followed by a theoretical study of the maximal modulation frequency determined by the algorithms. Next, the experimental part is given with simulations which confirm the findings from the theoretical part. The paper ends with conclusions.

**Wow and flutter determination, maximal modulation frequency.**

## 1. Introduction

Wow and flutter are audio distortions perceived as an undesired frequency modulation (FM) in the range of: a) wow from 0.5 Hz to 6 Hz, b) flutter from 6 Hz to 100 Hz [1]. The distortions are introduced to a signal by an irregular velocity of the analogue medium. As the irregularities can originate from various mechanisms, the resulting parasitic FMs can range from periodic to accidental, having different instantaneous values.

The problem of reducing wow and flutter remains unsolved in many respects in the analogue domain, however, digital signal processing (DSP) era enabled new approaches. First ideas on how to reduce wow and flutter in digitized analogue recordings were introduced by Gerzon [2]. Next, they were employed in a more practical implementation [3]. Some research on wow reduction was also reported by other authors [4-5]. Despite the available research, the wow reduction in the digital domain remains an open topic in many aspects.

The author of this paper took part in researching and developing new algorithms for wow and flutter reduction [6-8]. Different approaches were studied, some inspired by available literature, which resulted in several algorithms. The most successful algorithms are briefly presented in the next section. A more in-depth description is available in author's previous papers [8]. Third section gives a theoretical study of the maximal modulation frequency ( $FM_{\max}$ ) determined by the developed algorithms, whereas the fifth section presents simulations' results confirming the theoretical findings. The paper ends with conclusions.

## 2. Algorithms overview

The wow distortion can be characterized by the pitch variation curve (PVC) [4, 8]. This function describes the parasitic FM caused by the irregular velocity  $V(t)$  of the recording medium:

$$PVC(t) = V(t)/V_{nom} \quad (1)$$

where  $V_{nom}$  represents the nominal, constant speed value. If the pitch (speed) is constant, i.e. there is no wow, the PVC equals one. PVC deviations from unity illustrating pitch variations, indicate the wow depth.

### 2.1 Genuine Audio Content Analysis

Wow introduces FM to all spectral components of the distorted audio signal. Thus the spectrum of the useful audio, which is also affected by the wow modulations, can provide information necessary to determine the distortion characteristic.

Analyzing separated tonal audio components allows a PVC determination. Tonal components analysis can be performed for several components simultaneously, by using algorithms adapted from the sinusoidal modelling [4-5, 7]. However, such an analysis is more reliable provided that one tonal component depicting the wow most noticeably, is used. Additionally, as in most cases wow is mostly accidental and short in time the presented approach involves an human operator who selects the time-frequency region with the varying tonal component. Based on this selection, the spectral centre-of-gravity (COG) is calculated. The COG values from successive time frames are used to compute the final PVC [8].

### 2.2 Pilot tone analysis

Additional tones, besides the genuine audio, can be found in archival recordings. Some of the magnetic recordings contain a high-frequency bias (bias). Another pilot tone (15.734 kHz) can be found in the NTSC stereo soundtracks. Moreover, older recordings likely contain a low-frequency hum (hum). Tracking the additional tones' frequency variations was found useful in the PVC determination.

Tracking the bias allows determining the PVC. Since the bias is of high frequency, the Short Time Fourier Transform (STFT) is a suitable tool for detecting its time-frequency variations. Thus, in the researched algorithm the input signal is divided into STFT frames. The Hann window weights each frame and then the Fourier spectra are calculated. As a result of the spectrum calculations, the spectrogram matrix, representing the time-frequency properties of the signal, is obtained. Low-frequency spectral components are set to zero in order to remove the high-energy genuine audio content which may obscure the bias. Additionally, each amplitude spectrum (each column of the spectrogram matrix) is weighted by an appropriate preemphasis curve, allowing the bias enhancement. Then all columns are searched for maximal peaks, which, after correcting their amplitude and frequency estimation accuracy, are processed to obtain the PVC [8]. It was proven that the same algorithm can be used for tracking the NTSC-related tone.

Another approach was used for tracking the hum. In all European countries the power-line frequency equals to 50 Hz. In Americas it is typically 60 Hz, and always one of these two values in all other countries. Thus, the hum is placed in a frequency region where achieving a desired high STFT resolution is impossible (due to the “uncertainty principle”). Therefore, the proposed algorithm utilizes an auto-regressive (AR) modelling to track the hum frequency.

In the presented approach the input signal is downsampled prior to being split into frames. The downsampling allows eliminating most of non-hum related tonal components from the signal’s spectrum. To reduce the presence of noise further, and other non-hum related components, the signal is bandpass-filtered. The next step is tracking properly the hum which employs the AR-modeling. The downsampled and filtered signal is split into frames for this purpose. The coefficients of the corresponding AR filter are calculated using the modified-autocovariance method for a given frame. The coordinates of transmittance poles are obtained as a result. Subsequently, one pole is selected and the frequency corresponding to its angle, normalized with the base hum frequency, is chosen as the PVC value [8].

### 3. Maximal modulation frequency

The algorithms presented in the previous section aims at the PVC determination, however, the maximal modulation frequency determined by those algorithms is unknown. In order to establish the  $FM_{\max}$  two factors must be analyzed: a) the Nyquist sampling theorem for the discrete PVC, b) the quasi-stationarity criteria for the infra-frame modulations.

The discrete version of the wow and flutter characteristic, i.e.  $PVC(nT_{sPVC})$ , is defined by the PVC sampling period  $T_{sPVC}$  (or the PVC sampling frequency  $F_{sPVC}$ ), and represents its continuous counterpart until the maximal frequency  $FM_{\max} < 0.5F_{sPVC}$ . The PVC sampling period can be expressed in a number of distorted signal’s samples as:

$$R = T_{sPVC} / T_s = F_s / F_{sPVC} \quad (3)$$

where  $T_s$  ( $F_s$ ) represents the sampling period (frequency) of the distorted signal.  $R$  can be interpreted as the hop size of the time-windowing used in the algorithms presented in the previous section. Knowing that in the presented algorithms the hop size is always a quarter of the frame size ( $M$ ), the  $FM_{\max}$  limitation, being a consequence of the sampling theorem, can be expressed as:

$$FM_{\max} < F_s / 0.5M \quad (4)$$

The second factor which limits the  $FM_{\max}$ , comes from the quasi-stationarity criteria. The criteria eliminates time aliasing between the adjacent frames, and can be express as *a priori* assumption that the spectral components must be constant within the analysis frame. For a specific frame size, it can be formulated as:

$$M \ll T_{FM_{\max}} / T_s = F_s / FM_{\max} \quad (5)$$

Minimal  $M$  value which meets this criteria must be at least two times smaller then the FM period, assuming small frequency deviations. Thus the  $FM_{\max}$  limitation coming from the quasi-stationarity criteria can be expressed as:

$$FM_{\max} < F_s / 2M \quad (6)$$

Assuming that  $R$  is always smaller then  $M$ , which is true for the algorithms presented in this paper, Eq. 6 fulfils also the Nyquist theorem for the discrete PVC given by Eq. 4. Thus the final  $FM_{\max}$  limitation is given by the quasi-stationarity criteria.

Knowing the typical  $F_s$  values of the distorted signals and the typical frame lengths of the presented algorithms, the  $FM_{\max}$  can be calculated – see Tab. 1.

*Table. 1.  $FM_{\max}$  values for the presented wow & flutter tracking algorithms*

Algorithms and their settings	$FM_{\max}$
Power line hum tracker; $M = 8$ Sa; $F_s = 200$ Hz;	$FM_{\max} < 12,5$ Hz
COG tracker; $M = 1024$ Sa; $F_s = 44100$ Hz;	$FM_{\max} < 21,5$ Hz
High-frequency bias tracker; $M = 1024$ Sa; $F_s = 192000$ Hz	$FM_{\max} < 93,75$ Hz

The simulations presented in the following section aim at confirming the given  $FM_{\max}$  values.

#### 4. Experiments

The experiments were performed as follows. For each algorithm three test signals were prepared. Each signal – a pure sinusoidal tone with a constant carrier frequency – was modulated with a simple FM pattern equalling a half, whole, and a double  $FM_{\max}$  (see Tab. 1.). The carrier frequencies were set to: a) 50 Hz for the hum tracker, b) 3150 Hz for the COG tracer [1], c) 30 kHz for the bias tracker. Next the modulated signals were processed by the tracing algorithms and the resulting PVCs were compared against the known FM patterns. The results are as follows.

The results of COG tracking, are given in Fig. 1. Left chart shows, that the algorithm performs well for the slowly modulated signal. The 20 Hz modulation given in the centre chart is also well determined, however, there is a minor underestimation of the frequency deviations due to the time aliasing. Right chart shows even a greater underestimation caused by the unfulfilled quasi-stationarity criteria.

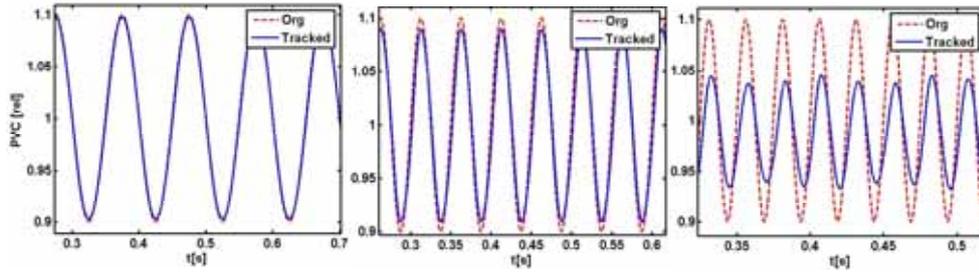


Fig. 1. Simulation results of the COG tracking. The FM values from the left side are: 10, 20 and 40 Hz.

The results of the hum tracking and the bias tracking, given in Figs. 2-3, present similar situations. The simulated tones are well tracked for the FM values below the  $FM_{max}$ .

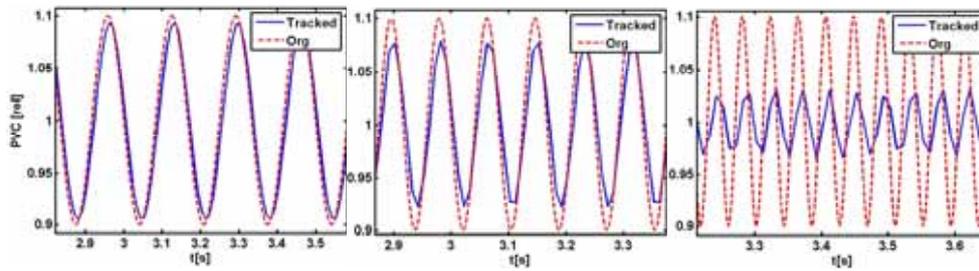


Fig. 2. Simulation results of a artificial hum tracking. The FM values from the left side are: 6, 12 and 24 Hz.

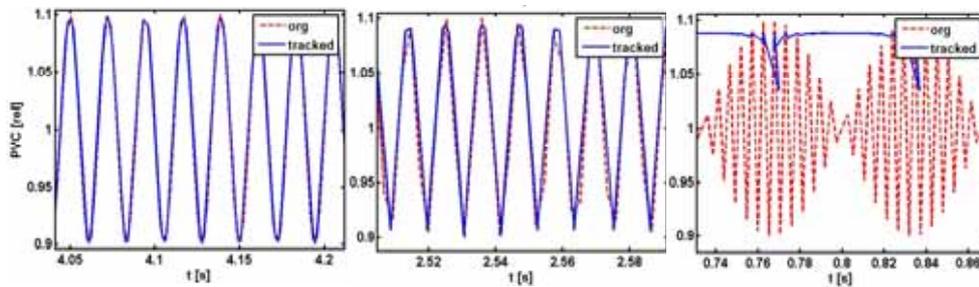


Fig. 3. Simulation results of an artificial bias tracking. The FM values from the left side are: 45, 90 and 180 Hz.

The third simulation for the bias tracker differs from the COG and hum tracking results. It is because for the first two algorithms the examples with the doubled  $FM_{max}$  still hold the Nyquist criteria (Eq. 4), thus the modulation frequency is preserved and only the deviation is underestimated. The 180 Hz modulation in the bias tracker exceeds the

$FM_{\max}$  set by the quasi-stationarity criteria as well as by the Nyquist criteria. In a result the modulation frequency in this example is not preserved.

## 5. Conclusions

Following conclusions can be derived both from the presented theoretical study and experiments. First of all the  $FM_{\max}$  is determined mostly by the quasi-stationarity criteria. Based on the criteria the  $FM_{\max}$  for each of the presented algorithms can be calculated. In a consequence it can be stated that all of the algorithms can track wow modulations. Further, the hum tracker can determined flutter modulations up to 12 Hz approximately; the COG tracker can determine flutter up to 20 Hz approximately; and the bias tracker can determine full range of flutter modulations. Additionally, the presented theoretical study allows to formulate an equation for the frame length calculation:

$$M = \lfloor F_s / 2FM_{\max} \rfloor \quad (10)$$

leading to a more intuitive algorithms' handling where the user specifies only the  $FM_{\max}$ .

## Acknowledgments

Research funded by the Commission of the European Communities, Directorate-General of the Information Society within the Integrated Project No. FP6-507336 entitled: "PRESTOSPACE – Preservation towards storage and access. Standardized Practices for Audiovisual Contents Archiving in Europe"

## References

- [1] Audio Engineering Society, Method for Measurement of Weighted Peak Flutter of Sound Recording and Reproducing Equipment, AES Standard, No. 6-1982, 2003.
- [2] M. Gerzon, "Don't Destroy the Archives!", unpublished manuscript (1992).
- [3] J. Howarth, P. Wolfe, "Correction of Wow and Flutter Effects in Analog Tape Transfers", Proc. 117 AES Convention, paper No. 6213, San Francisco, USA, 28-31 October 2004.
- [4] S. J. Godsill, P. Rayner, "The restoration of pitch variation defects in gramophone recordings", Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, Mohonk, NY State, USA, pp. 148-151, 17-20 October 1993.
- [5] J. Nichols, "An interactive pitch defect correction system for archival audio", Proc. AES 20th International Conference, paper no. 1941, Budapest, Hungary, 5-7 October 2001.
- [6] P. Maziewski, "Wow Defect Reduction Based on Interpolation Techniques", Bulletin of the Pol. Acad. of Science: Technical Sciences, vol. 54, no. 4, pp. 469-477, (2006).
- [7] P. Maziewski, L. Litwic, A. Czyzewski. "Accidental Wow Evaluation Based on Sinusoidal Modeling and Neural Nets Prediction", Proc. 120 AES Convention, paper No. 6796, Paris, France, 20-23 May 2006.
- [8] A. Czyzewski, A. Ciarkowski, A. Kaczmarek, J. Kotus, M. Kulesza, P. Maziewski, "DSP Techniques for Determining "Wow" Distortion," JAES, vol. 55, no.4, pp.266-284, (April 2007).