ARCHIVES OF ACOUSTICS
33, 1, 117–123 (2008)

MTS PILOT TONE TRACKING FOR "WOW" DISTORTION DETERMINATION

Piotr PASTUSZAK

Gdańsk University of Technology
Multimedia Systems Department
Narutowicza 11/12, 80-952 Gdańsk, Poland
e-mail: ugm@go2.pl

(*received June 15, 2007; accepted November 30, 2007*)

New algorithm for the wow distortion characteristic determination using the Multichannel Television Sound (MTS) pilot tone tracking is introduced. By observing variations of the pilot frequency, the depth of the parasitic modulation can be estimated. MTS is an audio coding system proposed by the National Television Standards Committee (NTSC), thus the algorithm is applicable only to NTSC TV standard-based recordings. It is based on the high-frequency bias tracking algorithm and further modified to minimize the influence of acoustic-band audio signals which may interfere with the 15.734 kHz MTS pilot tone.

Keywords: wow distortion, Multichannel Television Sound, High Frequency Bias.

1. Introduction

"Wow" distortion (further referred to as wow) is a pitch variation of a relatively low frequency (up to 6 Hz [1]) commonly found in gramophone or tape recordings. In most cases, it is caused by the irregular velocity of a recording medium, which can be triggered by the unstable motor or unsymmetrical geometry of the carrier. The depth and character of such modulation is in most cases unknown thus hard to evaluate or model accurately.

The pitch-variation curve (further referred to as the PVC) is the characteristic which depicts the depth of the parasitic modulation triggered by wow [2–4]. Such depth is measured as the nominal to actual frequency ratio, so the PVC value of 1 indicates the signal without distortion, the PVC value of 0.5 is the actual frequency two times less than the correct one and so on. One of the methods of the PVC determination is to track the modulation of some tone with a theoretically stable frequency that is assumed to be known. The high-frequency bias in magnetic recordings (further referred to as the HFB), the power-line hum or the MTS pilot tone in the NTSC stereo recordings are examples of such tones. The difference between them is the carrier frequency which implies specific approaches and techniques of tracking. One of such examples is the algorithm for tracking the HFB constructed at the Multimedia Systems Department of the Gdansk University of Technology [4–7].

2. The MTS pilot and the HFB tracking algorithm

The aim of this work is to present the algorithm for tracking the MTS pilot tone in the NTSC signals which is based on the mentioned algorithm for the HFB tracking [4–7].

The main differences between these two signals are:

- the MTS pilot tone is at 15.734 kHz, which is inside the acoustic band, and can be masked by useful signals (in contradiction to the HFB which is high above 20 kHz);
- when copying the NTSC recordings, a new pilot tone is added (in contradiction to magnetic tapes where the existing bias is overwritten) – such phenomenon is called "the background pilot tone" later in this work;
- the pilot tone has a relatively low and unstable level (complete decays may occur as well).

The techniques used to determine the PVC in the HFB tracking algorithm (Fig. 1) are described below. First, a part of the signal is taken for analysis, Hann windowing is applied to it, and then the FFT is performed to operate on a frequency domain. Next, the entire acoustic band is removed, preemphasis and spectral expansion algorithms are applied, and finally the maximum of the spectrum energy is looked for. Parabolic interpolation helps to find the fractional index of the bias frequency bin. Precision can

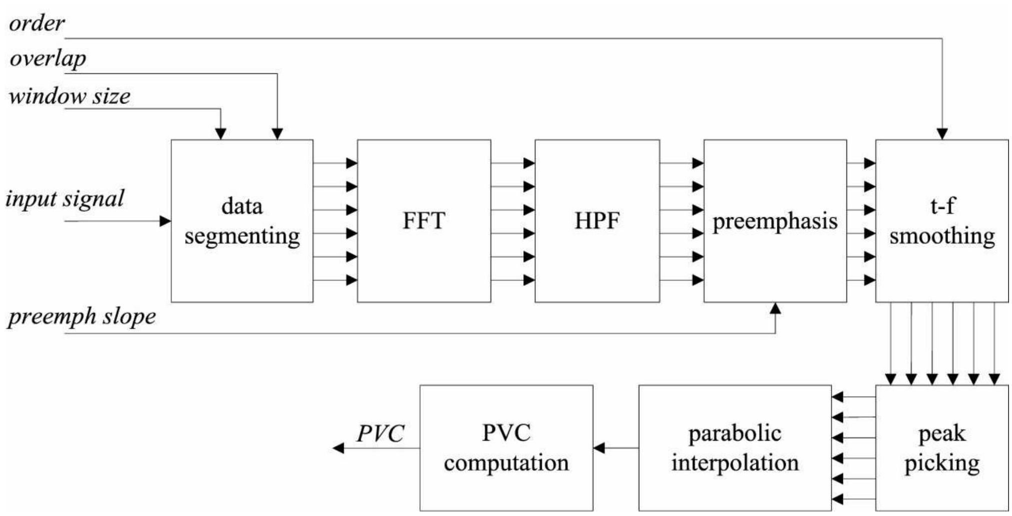

Fig. 1. Block diagram of the HFB tracking algorithm [4].

be further increased by applying other methods such as the time-frequency reassignment [8]. More information on the described algorithm can be found in several publications [4–7].

It must be mentioned that the presented algorithm differs form the standardized wow measurement techniques [1] were only one value of the weighted peak flutter index is calculated per item (recorder or reproducer). In the described algorithm wow is measured continuously for the period of time represented by the recording. The algorithm captures wow variations with a high precision (per sample). Also the pilot tone frequency is different (in the Audio Engineering Society standard it equals 3150 Hz [1]). More information on the described HFB tracking algorithm can be found in several publications [4–7].

3. The MTS Tracking Algorithm

In the HFB tracking algorithm, the signal was filtered to remove spectral components of up to 25 kHz (HPF block in Fig. 1). Then the preemphasis filter was applied and the entire spectrum was expanded (raised to the 4th power). The first operation was to counterbalance the decrease of energy towards high frequencies and the latter was to bring forth the HFB tone against the background of a spectral noise. Also the time- and frequency-domain smoothing (both 3rd order) were applied. This had the effect of blurring the spectrogram which was helpful to reduce noise and improve frequency estimation.

The MTS tracking algorithm is presented in Fig. 2. It operates in two phases (loops). The first one is similar to the HFB tracking algorithm, while the latter is more complex.

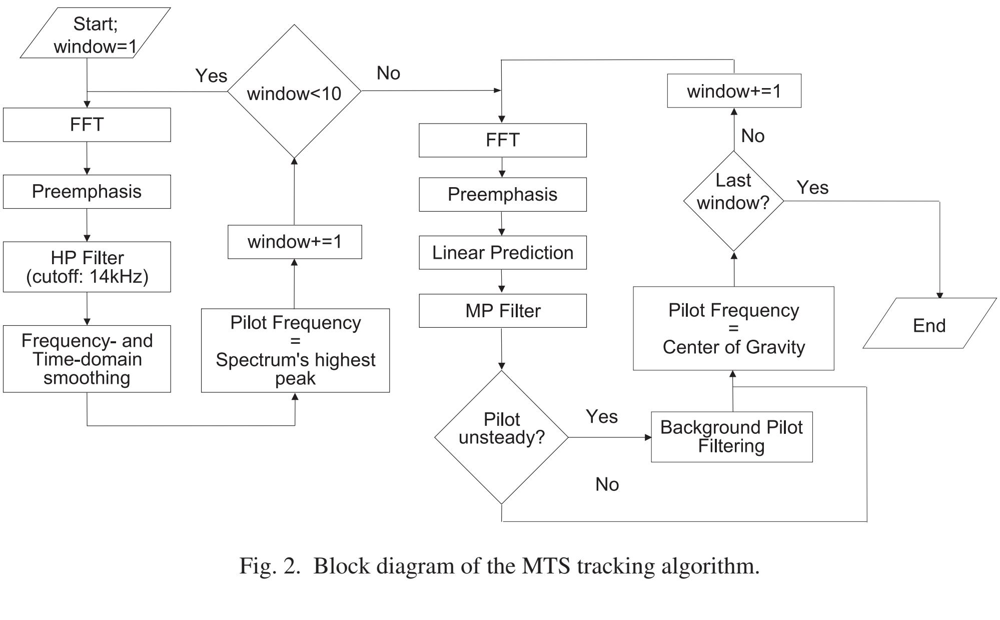

Fig. 2. Block diagram of the MTS tracking algorithm.

In the first phase, the only modification made to the HFB tracking functions was to lower the cutoff frequency to 14 kHz because of the importance of the frequencies around 15.7 kHz. It is assumed that the beginning of the recording under analysis is free of wow distortion, so the only peak existing above the cutoff frequency is the pilot tone. Parabolic interpolation is used to estimate the pilot tone frequency more precisely. It is assumed that acoustic signals above 16 kHz that could have been interfered, can be disregarded due to their low energy.

Additionally, the center of gravity (CoG) within the neighborhood of the nominal pilot tone frequency (15734 $\pm$ 250 Hz) is sought. It is used later to eliminate the background pilot tone signal and to correct any potential inexactness of the signal's pitch.

After 10 steps, the algorithm starts to work differently. It enters the second phase. The pilot tone frequency is estimated basing on its 3 previous frequencies and using the second order linear prediction. Then the analyzed band is limited to 7.5% of the estimated frequency at the bottom and 10% at the top. The margins are not statically assigned, so if the pilot tone goes towards low frequencies (due to a deeper modulation), they are still narrow enough to reduce the influence of "useful" yet unwanted spectral components. These components always interfere with the lower frequencies and have more energy (the spectrum energy seldom increases towards higher frequencies) – that is the reason for making the bottom margin narrower than the top one. Next, a simple weighting function is created, where the element of the predicted frequency has the weight of 1, border elements have the weight of 0 and weights of elements between the center and the borders are linearly approximated and raised to the second power.

When the predicted frequency is up to 1 kHz from the nominal pilot tone, a very narrow (200 Hz bandwidth) mid-pass filter (MP) is applied at the actual pilot tone frequency to eliminate the background pilot tone. It is done to avoid following the wrong pilot and altering the tracking of the right one. This filtering is applied only if the pilot tone is unstable for some time, so that the algorithm does not filter the correct pilot tone which is stable and at the right frequency. Stable state is defined when the frequency of the pilot tone does not differ more than 250 Hz from the nominal pilot frequency in the 10 previous steps.

Finally, the pilot tone frequency for the current step is calculated as the CoG of such narrowed spectrum. Choosing the CoG instead of just picking up a maximum of spectrum reduces accidental jumps to nearby strong interfering signals that are making pilot tone harder to find in the next steps (due to incorrect prediction caused by the false sample). Such accidental jumps are also limited by checking if the difference between the last and current frequency is not greater than 800 Hz. If it is so the last frequency is taken.

4. Experiments and results

The MTS tracking algorithm was tested using the wow-distorted audio sample of the length of 143 sec. As seen in its spectrogram (Fig. 3, upper part), there are three fragments with oscillating FM distortions (around 20–30, 45–65 and the third fragment,

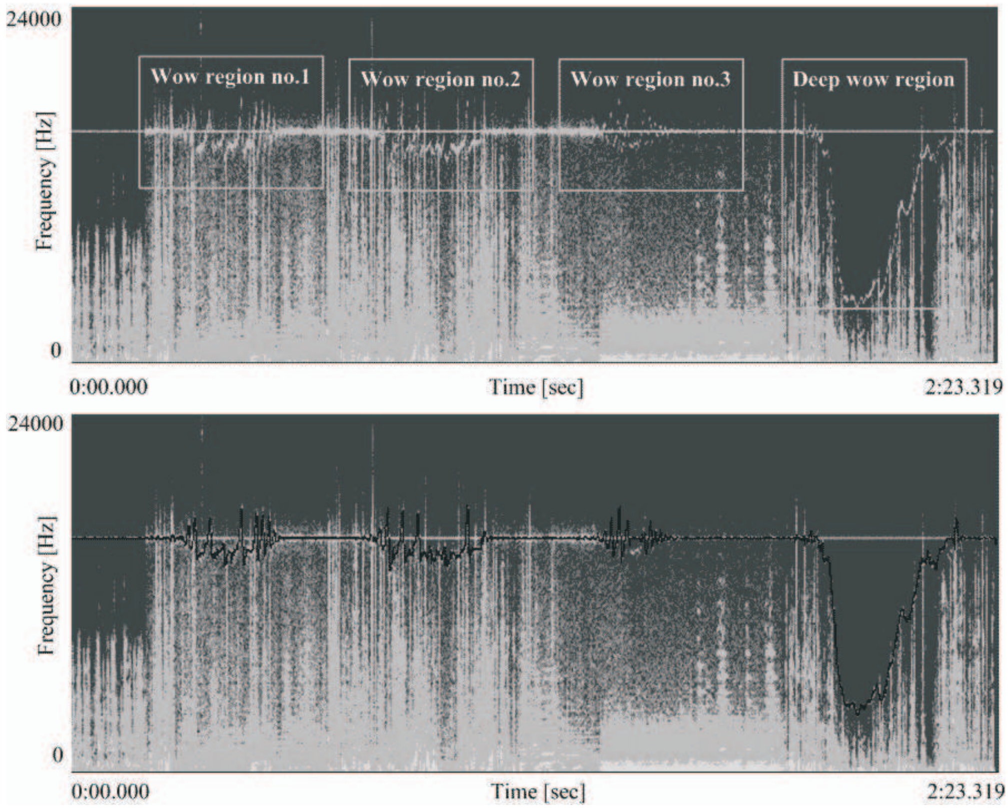

Fig. 3. Wow-distorted audio sample spectrograms with wow regions (upper part) and tracked MTS pilot tone (black line in the lower part).

hardly seen at 80–90 sec. due to the lowered pilot tone level) and one with deep asymmetrical modulation (starting around 115 with a peak at around 120 sec, and lasting almost until the end of the sample). The background pilot tone is noticeable in the entire clip. The mentioned varying level of the pilot tone is also easy to see (it goes up after 11 and down again after 82 sec in region no.3). Unstable level causes difficulties in finding the correct pilot tone (the background one has actually a higher level than the distorted one as seen in upper spectrogram in Fig. 3). The second and third region was difficult to track properly due to some rapid changes in the modulation level within the neighborhood of the background pilot tone. The problem with the last region was to avoid interfering with the audio spectral part of the signal. An impulse-type, high-energy useful audio content introduced a danger of losing the track. This was minimized by the varying band of the MP filter, narrower at the bottom.

The results of the pilot tracking are shown in the lower part of Fig. 3. It presents the spectrogram with the marked MTS pilot track (black line). Optimal selection of various parameters led to satisfactory results, where the pilot tone was found and properly followed.

5. Disadvantages and possible modifications

The weakest point of the presented algorithm is its iterative character with the backward dependence. It means that there are no means to verify the confidence of the signal being followed. A simple confidence value could be added in future – e.g. based on the spectral level of the estimated pilot tone frequency. If the confidence value were too low in a certain number of steps, the algorithm could try to find the pilot tone by searching for a high peak starting from upper frequencies and moving downwards. Distinguishing the background pilot tone from the actual one would be problematic, though.

Another simple concept of the confidence value is to track the MTS pilot tone from both sides of the distorted sample (forwards and backwards). Significant differences in the estimated pilot tone track would mark regions of uncertainty.

The fact that the algorithm was tested on a small variety of NTSC signals due to lack of sufficient examples of distorted audio is another weak point. Artificial NTSC samples could not be forged because the MTS pilot tone itself was problematic to model (rudimentary documentation).

Further improvements to the precision of the frequency determination could be made. Basing on the computed track of the pilot tone, an additional algorithm could try to find the real frequency in close neighborhood. It was not introduced directly because of the required stability of the algorithm (following the pilot tone track). Any future correction of the pilot frequency will not result in "error propagation" because the track will be fixed.

Acknowledgment

The research was 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*, Standard No. 6–1982, 2003.

[2] GODSILL S. J., RAYNER P., *The restoration of pitch variation defects in gramophone recordings*, Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 148–151, Mohonk, NY State, USA, 17–20 October 1993.

[3] HOWARTH J., WOLFE P., *Correction of wow and flutter effects in analog tape transfers*, Proceedings of the 117th Audio Engineering Society Convention, paper no. 6213, San Francisco, USA, 28–31 October 2004.

[4] CZYŻEWSKI A., CIARKOWSKI A., KACZMAREK A., KOTUS J., KULESZA M., MAZIEWSKI P., *DSP techniques for determining "Wow" distortion*, Journal of the Audio Engineering Society, **55**, 4, 266–284 (2007).

[5] CZYŻEWSKI A., DZIUBIŃSKI M., CIARKOWSKI A., KULESZA M., MAZIEWSKI P., KOTUS J., *New algorithms for wow and flutter detection and compensation in audio*, Proceedings of the 118th Audio Engineering Society Convention, paper no. 6353, Barcelona, Spain, 28–31 May 2005.

[6] CZYŻEWSKI A., MAZIEWSKI P., DZIUBIŃSKI M., KACZMAREK A., KULESZA M., CIARKOWSKI A., *Methods for detection and removal of parasitic frequency modulation in audio recordings*, Proceedings of the Audio Engineering Society 26th International Conference, paper no. 3–3, Denver, USA, 7-9 July 2005.

[7] CZYŻEWSKI A., MAZIEWSKI P., *Some techniques for wow effect reduction*, Presented at the IEEE International Conference Image Processing 2007, San Antonio, Texas, USA, 16–19 September 2007.

[8] AUGER F., FLANDRIN P., *Improving the readability of time-frequency and time-scale representations by the reassignment method*, IEEE Transactions on Signal Processing, **43**, 1068–1089 (1995).