Postprint of: Czyzewski A., Maziewski P., Kupryjanow A., Reduction of parasitic pitch variations in archival musical recordings, Signal Processing, Vol. 90, Iss. 4 (2010), pp. 981-990, DOI: 10.1016/j.sigpro.2009.09.015

Title: Reduction of Parasitic Pitch Variations in Archival Musical Recordings

Authors: Andrzej Czyzewski, Przemyslaw Maziewski, Adam Kupryjanow

Affiliation: Gdansk University of Technology, Multimedia Systems Department, ul. Narutowicza 11/12, 80-226 Gdansk, POLAND; {andcz, przemas, adamq}@sound.eti.pg.gda.pl, tel. +48 58 3471301, fax +48 58 3471114.

Abstract: A new method for reducing parasitic pitch variations in archival audio recordings is presented. The method is intended for analyzing movie soundtracks recorded in optical films. It utilizes image processing for calculating and reducing effects of tape shrinkage being one of the main reasons for parasitic pitch variations in audio accompanying moving images. As long as the film tape characteristics are known the new method can be easily tuned to analyze archival recordings. The new method is also compared to some previous approaches to pitch variation correction.

List of unusual symbols used in the article:

f_m – parasitic modulation frequency

PVC(t) – pitch variation curve

P_{distnom} – standardized distance between two perforation holes

$P_{\mathrm{dist}n}$ – real distance between the n-th and the n+1perforation hole

$H_{1\mathrm{sec}}$ – length of a movie tape recording of one second time duration

t_n – sampling time of the n-th PVC value

$f_{\max}$ – maximal modulation frequency properly captured in the determined PVC

F_s – sampling frequency of a digitized audio recording

M – processing frame length

MSE – mean square error

$\mathrm{PVC}_{\mathrm{sim}}(t)$ – simulated PVC

$\mathrm{PVC}_{\mathrm{det}}(t)$ – determined PVC

N – total number of PVC values

Number of:

pages - 22, tables - 9, figures - 9.

Keywords: audio restoration, parasitic pitch variations, wow and flutter.

1. Introduction

The purpose of this work is presenting a new method for reducing parasitic pitch variations in archival audio recordings and comparing it with the already known approaches. Parasitic pitch variations, mainly wow and flutter, can be defined as unwanted frequency modulations added to the original signal during recording, producing, monitoring or duplicating stages. Wow and flutter are defined based on the range of the parasitic modulation frequency (f_m): in case of wow $f_m \in \langle 0.5\ \text{Hz}, 6\ \text{Hz})$, in case of flutter $f_m \in \langle 6\ \text{Hz}, 100\ \text{Hz}\rangle$ [1]. The distortions are introduced to the signal by the irregular velocity of its analogue carrier. The irregularities can originate from various mechanisms, depending on the carrier type, its condition and damages, its production technique, and other factors [2], [3], [4]. The irregularities are commonly found in the ethnic music recordings made by an unskilled staff using nonprofessional recording systems. Especially in the eastern Europe low quality of the audio carriers, poor alignment and maintenance of the recording and reproducing equipment enhanced the risk of wow and flutter.

An example of a badly produced movie tape splice, taken from a Polish archive film, is given in Fig. 1.

Fig. 1 Example of a badly produced movie tape splice

When such a tape makes a contact with the projector's sprocket roller (or telecine's) it moves slightly back or towards its original movement direction, depending on the difference between the interval of successive perforation holes and the distance between the cogs of the transpiration system (see Fig. 2). Both cases result in a momentary pitch variation of the soundtrack.

The other reason for wow and flutter occurring in the movie soundtracks is the tape shrinkage. It is caused by loss of water, solvent and plasticizer in movie tapes. Both nitro-based and acetate-based films are exposed to this process. Movie tape shrinkage causes similar problems as the ones presented with the badly produced splice. As the shrunk perforation do not match the distances between the cogs of the sprocket roller, it causes movie tape displacements each time the transportation cogs enter the perforation hole. This situation is illustrated in Fig. 2.

Fig. 2 Tape shrinkage effects on the movie tape transportation system [5]

Different reasons for wow and flutter can be found in magnetic and mechanical recordings [6]. Again most of them are caused by the poor equipment or lack of appropriate production standards, thus they are highly perceivable in the ethnic recordings. The authors of this papers conceived and introduced several DSP algorithms for correcting wow and flutter. The algorithms utilize some ancillary information retrievable from the audio documents, i.e. the initial frequencies of the power line hum, of the high frequency bias or of the 15.734 kHz pilot tone, which can be found in the NTSC television audio recordings. Short overview of those will be presented in the next section, whereas the details can be found in the literature [6], [7]. New approach was recently studied by the authors [8]. It is dedicated to movie soundtracks and it utilizes image processing of the movie tape where the ancillary historical information, like the tape format, its initial speed or the recording procedure, play the crucial role. After providing those information, which can be very unique in case of some obsolete recordings, the reduction of the pitch variations of the movie soundtrack can be greatly improved and even done automatically. This kind of a new approach will be presented in details in the paper and then compared with the previously elaborated algorithms.

2. State of the Art

The idea of DSP algorithms for wow and flutter determination and reduction was introduced by Gerzon [9]. He suggested using the magnetic leftover of the high-frequency bias, which can be found in magnetic tape recordings, in the form of a pilot tone depicting the distortions' course. Additionally, he proposed the non-uniform resampling technique for reducing above distortions. However, it seems that Gerzon did not verified his ideas, practically. First DSP algorithms for wow course determination and reduction were presented by Godsill [10]. Then, further developed, they were reported by him together with several co-researchers [11], [12], [13], [14]. The algorithms aimed at automatic correction of smooth pitch variations over long time scales. They were build using a three-step processing scheme. First the preprocessing, i.e. the standard sinusoidal modelling analysis, provided information on the frequency partials. Secondly the partials were analyzed in order to determine the wow course. Subsequently, the non-uniform resampling was used for the distortion reduction. The main emphasis was put on the second step, i.e. the determination of the wow course. Different *a priori* models of the distortion course, ranging from simple sinusoidal to more complex regressive processes, were examined in this context, employing the Bayes' estimation procedure. It was through those models that the ancillary information about the recording medium, gramophone records in this case, could be used.

Partials' processing was also used by Nichols in his approach to the wow determination and reduction [15], [16]. The first one among his algorithms was similar to the Gosill's approach with only minor improvements added, such as the spectrum warping or the iterative procedure for improving the spectrum's peaks frequency estimation. His second algorithm was build on a novel concept of graphical processing of the spectrogram. In this algorithm partials were build using image processing routines allowing of pixels selection from the spectrogram which represented the tonal components. In both of Nichol's algorithms, partials were identically processed in order to determine the wow course. As the algorithms designed by Nichols were dedicated to phonograph recordings restoration, the simple sinusoidal model was used to describe wow. The non-uniform resampling was then applied to reduce the distortions.

Also Howart and Wolfe reported algorithms for wow and flutter determination and removal [17]. They designed a system being a combination of a hardware part for capturing the high frequency bias and a software part for wow and flutter reduction. Additionally, those authors presented the study of the non-uniform resampling technique used for wow and flutter reduction [18].

Wow and flutter determination and reduction was studied also by the authors of this paper. Different algorithms were proposed for wow course determination such as [6]: the power line hum tracker, the high frequency bias tracker, which was also successfully used for the multichannel television sound pilot tone tracking [7], as well as the spectrum centre of gravity analyser. The algorithms were tested employing real-life archive recordings [6]. Simulations were also made in order to determine their limitations in terms of the maximal modulation frequency [19]. It was found that only the high frequency bias tracker can determine both wow and flutter characteristics, whereas the other algorithms can process flutter only to some extent. We have also studied the non-uniform resampling influence on the quality of restored recordings [20].

All of the algorithms for wow and flutter correction presented above were based on a similar idea. The audio documents' content, either the audio itself or embedded signals like the power line hum or the high frequency bias, were analyzed in order to determine the distortions' course. Meanwhile, a different approach can be applied: the analysis of the audio carrier. Some interesting work related to his subject already exists. Namely, Fadeyev and Haber proposed image processing of the mechanically made recordings (modulated groves) [21]. In their approach the grove shapes are being scanned, their images corrected, and then translated into sound. Image corrections allow virtually repairing even severe damages of the carrier. This kind of a processing allows also for wow and flutter reduction but only with regard to the ones triggered by the ellipticity or eccentricity of the gramophone records. Sounds from other carriers can be similarly restored. Consequently, it was reported that image processing can be used for movie soundtracks restoration [22], [23].

Image processing of optical tapes can be used for wow and flutter determination and reduction in movie soundtracks. The algorithms presented in the next section allow to determine characteristics of wow and to some extent the course of flutter in order to reduce them.

3. New Approach

Typically in the DSP algorithms for wow and flutter determination these distortions are characterized using the pitch variation curve (PVC) [6]. This curve depicts the distortions' course showing their changes in time. If there are no distortions the PVC equals one. The PVC deviations from unity, illustrating pitch variations, indicate a depth of wow and flutter. In most real-life recordings the PVC is close to a constant unity function with occasionally varying fragments indicating wow and flutter [6]. The PVC can be estimated using the tape shrinkage analysis (shrinkage results in pitch changes of the soundtrack). To estimate the local tape shrinkage, physical dimensions - heights of selected elements of the movie tape must be compared to their original values. Therefore, the original values must be known *a priori*. In most cases they can be learnt from standards (16 mm or 35 mm movie tapes). In rare ethnic recordings, made with some obsolete formats, the knowledge about the movie tape format must be obtained from the historical sources or presumptions must at first be made and then corrected after obtaining the initial results.

It is expected for the presented method that the movie tape will be digitized properly with a high spatial resolution and at least 48-bit colour depth. A proper scan is obtained when the image of the movie frame is parallel to the edge of the scanned image content and the sharpness of the picture is correct. In Fig. 3 an image of a properly scanned, typical movie frame is shown.

Fig. 3 Example of typical 35 mm movie frame

The PVC can be estimated via analysis of the standardized elements of the movie tape (i.e. distances between frames or perforation holes). When analysing the perforation of the film, the first value of the PVC is determined in accordance with expression (1):

$$\mathrm{PVC}(t_1) = \frac{P_{\mathrm{dist1}}}{P_{\mathrm{distnom}}} \tag{1}$$

where P_{distnom} represents the known (standardized) distance between two perforation holes, P_{dist1} is the distance between the first and the second perforation hole in the analyzed movie (see Fig. 3) and t_1 is the sampling time of the calculated PVC value that can be expressed in seconds as:

$$t_1 = \frac{P_{\mathrm{dist1}}}{2H_{\mathrm{1sec}}} \tag{2}$$

H_{1sec} represents length of a movie tape recording of one second time duration in Eq. 2.

The successive PVC values can be calculated as:

$$\mathrm{PVC}(t_n) = \frac{P_{\mathrm{dist}n}}{P_{\mathrm{distnom}}} \tag{3}$$

where: t_n represents the sampling time of the n-th PVC value and it can be calculated in seconds as:

$$t_n = \frac{\sum_{m=1}^{n-1} P_{\mathrm{dist}m} + \frac{P_{\mathrm{dist}n}}{2}}{H_{1\mathrm{sec}}} \tag{4}$$

In the next section some algorithms for calculating the distances between successive perforation holes as well as between the frames' beginnings will be presented.

3.1 Algorithm for Calculating Distances Between Successive Perforation Holes

As the perforation holes have rectangular shape (or close to it) the $P_{\mathrm{dist}n}$ can be calculated using locations of two parallel lines representing the successive perforation holes' beginnings. In the presented algorithm the edge detection is applied to determine the perforation beginnings. The image colour is analysed for capturing the transition between the darkest colour – representing the perforation hole's border – and the lighter colours representing either the tape or the hole. The scheme of the algorithm is presented in Fig. 4.

Fig. 4 Algorithm for calculating the distances between successive perforation holes in one movie frame

In the algorithm only one frame perforation image is being analyzed. The processing is repeated for all of the successive frames. The first processing block represents the cropping procedure where the image is cut accordingly to the perforation width. To facilitate the determination of the perforation distances, the decimal-to-binary conversion is made, emphasising perforation borders. This results in a black and white image. Further the processed image is analyzed as a two dimensional matrix. First the mean value in each column of the matrix is calculated and then the mean values' derivate is computed. The obtained vector corresponds to the intensity changes in the perforation's picture. Its maximal values represent the beginnings of the successive perforation holes (see Fig. 5). The $P_{\mathrm{dist}n}$ parameter is calculated as a difference of two successive perforation holes' beginnings.

Fig. 5 Example of determined distances between successive perforation holes for 35 mm film (single frame)

Analysing the perforation in 35 mm films allows determining 4 PVC values for each movie frame. It is different, however, in case of 16 mm films, where only one perforation hole comes with each movie frame. In such cases, or when the perforation is not available, since it was not scanned, the movie frames' analysis can be applied for the $P_{\mathrm{dist}n}$ calculation.

3.2 Algorithm for Calculating Distances of Successive Movie Frames

Movie frames' heights can be calculated using the similar processing scheme as the one presented in Fig. 4. It is because they share the same characteristics as the perforation, i.e. they are rectangular with well defined borders. In case of movie frames in the first processing step the perforation must be cut off from the analyzed image. Then the same processing is applied as it was used for the perforation. The maxima of the obtained intensity changes vector can now be applied to determine the movie frames beginnings (see Fig. 6).

Fig. 6 Example of determined distances between successive movie frames for 35 mm film

Knowing the PVC allows one's to reduce the soundtrack's wow and flutter caused by the tape shrinkage. The reduction can be done either via the soundtrack's picture processing – through inverting the shrinkage - or it can be done for the digital sound obtained from the soundtrack. The first approach involves a two-dimensional and the second one a one-dimensional non-uniform resampling [20]. In both cases the algorithm for translating the image of the soundtrack into digital sound can be applicable and very useful.

3.3. Translating the Image of the Soundtrack Into Digital Sound

The translation algorithm was designed to process movie soundtracks frame-by-frame. In order to concatenate the contents of sound buffers from successive frames a redundancy must be provided to the processed images. The image of one movie frame must contain a part of the previous and of the next frame (like it was shown in Fig. 3). The scheme of the translation algorithm is presented in Fig. 7.

Fig. 7 Scheme of algorithm for translating image of optical soundtrack to digital sound

At the preprocessing stage the image of the soundtrack is edited out of the entire frame. Then the algorithm follows up the way of working of a standard cinema film projector. The amplitude of the sound signal is determined based on intensity of light in the projector that has passed through the soundtrack changes, . Thus, in order to calculate the intensity of light the mean values in columns of the soundtrack's image are calculated. The obtained vector of mean values represents the sound signal's amplitudes. As in most cases the obtained signal has a considerable DC component, this component is removed by zeroing the signal's mean value.

If more than one image frame has to be translated, sound buffers from each successive frame have to be connected. The concatenation is made by calculating the cross-correlation function between the sound signal taken from the end of the image of the previous frame and the sound signal taken from the beginning of the image of the currently processed frame. The maximum value of the cross-correlation function determines the similarity region of the sound signals of the previous and of the current buffer, allowing connecting and mixing them both.

When the digital sound is finally obtained, the wow and flutter reduction can be applied. Here the known non-uniform resampling techniques can be used [20].

4. Experiments

The experiments aimed at comparing the new method for wow and flutter determination and reduction with the known approaches. Particularly, two factors were considered. The first one is the maximal modulation frequency (f_{max}) which can be properly captured in the determined PVC. The second one represents the precision of the determined PVC, i.e. the difference between the true and the determined characteristics of wow and flutter calculated for frequencies below f_{max}.

The preliminary study of different factors limiting f_{max} that can be determined using the new method, was done by the authors of this paper [8]. It was shown that f_{max} can have different values depending on the different film format. Additionally, the selection of movie tape elements which were utilized for calculating the PVC, i.e. frames or perforations, also influences the maximal modulation frequency value. The determined values of f_{max} are given in Tab. 1 and in Tab. 2. In the tables different film formats are characterized using the movie tape velocity parameter, which is expressed either in terms of frame per second (fps) or in terms of perforation per second (pps), depending on the movie tape elements which were used to calculate the PVC.

Tab. 1 f_{max} values calculated for different film formats utilizing movie frame analysis [8]

Tab. 2 f_{max} values calculated for different film formats utilizing tape perforation analysis [8]

As 35 mm film is one of the most popular formats in film archives [3], thus this format was chosen for experiments. Additionally, as for this film format the f_{max} values are higher for the perforation analyses (see Tab.

2), the algorithm for calculating the distances between successive perforation holes (see Section 3.1) was used in performed simulations.

Among the reported audio-based algorithms for the PVC determination (see Section 2) only 3 have been studied with regards to $f_{\max}$ [19]. Those are: the power line hum tracker, the centre of gravity (COG) tracker and the high frequency bias tracker. The results of this study are given in Tab. 3. In the table the F_s stands for the sampling frequency of the digitized audio recordings, whereas M represents the processing frame length.

Tab. 3 $f_{\max}$ values calculated for different, audio-based, wow and flutter tracking algorithms [19]

Among those 3 algorithms the power line hum tracker is the most appropriate one to be used for the comparison with the new method. That is because power line hum can be easily found in many of the archive films' soundtracks, whereas the high frequency bias was never used to produce optical recordings. Moreover, the power line hum tracker can be used almost automatically, i.e. the human operator has to tune up the initial processing parameters only with no further interaction needed,. Such an approach helps avoiding problems with the wow versus vibrato discrimination, which are common when using the COG tracker, while the operator has to decide which of the analyzed audio parts represent distortions and which correspond to natural pitch changes (i.e. vibrato).

The experiments were performed as follows. First wow and flutter simulations were prepared. For the power line hum tracker 4 s long sound files with 50 Hz and 60 Hz tones were prepared, simulating the European- and USA-originated recordings respectively with regards to power line hum inclusion. For the new method 100 movie frames (digital images) in the 35 mm format were prepared, corresponding to 4 s long audio and video recording, assuming 25 fps tape velocity. As the typical tape resolution lies between 60 to 80 lines per mm [5], which represents the range from 1524 DPI to 2032 DPI, the 3600 DPI image resolution was chosen (i.e. the lowest DPI resolution available in typical scanners, without any DSP interpolation, satisfying the higher tape resolution).

Both the sound files and the movie frames were frequency-modulated. The frequency modulations were driven by the PVCs (hereafter called the simulated PVCs) build on frequencies no higher than $f_{\max}$. Each simulated PVC was generated as a white Gauss noise, low-pass filtered with the cut-off frequency equal to the particular $f_{\max}$ value. Five $f_{\max}$ values were used: 0.5 Hz, 6 Hz, 12 Hz, 19 Hz and 38 Hz. First two of them correspond to the typical wow boundary frequency limits (f_m from 0.5 Hz to 6 Hz). The third one is the $f_{\max}$ value which can be properly captured using the power line hum tracker (see Tab. 3). The fourth one is the $f_{\max}$ value which can be established using the new method (see Tab. 2), whereas the last value doubles it. The last value was used to study the performance of both algorithms when tracking distortions located outside the allowable range of frequency modulations. To simulate different depths of wow & flutter two amplitude ranges of the simulated PVCs were used in the experiments: <0.9, 1.1> and <0.99, 1.01>. The first range was used for simulating deep wow and flutter, whereas the second was employed to simulate the smaller amplitudes of the distortions. Frequency modulations were performed using the non-uniform resampling. Mono sound files – represented by one dimensional vectors - were resampled using one-dimensional, spline-based resampler [20]. Two-dimensional resampling was used for the movie frames processing (as pictures represent two-dimensional objects). The simulations were repeated 100 times for each algorithm, for all $f_{\max}$ values, and for both amplitude ranges of the simulated PVCs. The mean square error (MSE) among the simulated ($\mathrm{PVC}_{\mathrm{sim}}$) and the determined ($\mathrm{PVC}_{\mathrm{det}}$) PVCs were calculated for each simulation:

$$\mathrm{MSE} = \frac{1}{N}\sum_{n=1}^{N}(\mathrm{PVC}_{\mathrm{sim}}(t_n) - \mathrm{PVC}_{\mathrm{det}}(t_n))^2 \quad (5)$$

where N represents the total number of PVCs samples. Then the results were summed up using the mean and standard deviation values of the calculated MSEs.

Results obtained using the power line hum tracker are given in Tabs 4 to 7. First, the results for the simulations involving 50 Hz tone are given. They are followed by the results for the sound files with the 60 Hz tone. In each case the results from the simulations with the wider range of PVC amplitudes are presented first.

Tab. 4 MSE of the PVCs determined using the power line hum tracker, 50 Hz tone, PVC amplitude range <0.9, 1.1>

Tab. 5 MSE of the PVCs determined using the power line hum tracker, 50 Hz tone, PVC amplitude range <0.99, 1.01>

Tab. 6 MSE of the PVCs determined using the power line hum tracker, 60 Hz tone, PVC amplitude range <0.9, 1.1>

Tab. 7 MSE of the PVCs determined using the power line hum tracker, 60 Hz tone, PVC amplitude range <0.99, 1.01>

The results obtained for the new method are given in Tab. 8 and Tab. 9.

Tab. 8 MSE of the PVCs determined using local tape shrinkage analysis, PVC amplitude range <0.9, 1.1>

Tab. 9 MSE of the PVCs determined using local tape shrinkage analysis, PVC amplitude range <0.99, 1.01>

To better visualize the obtained results additional charts are given. Fig. 8 and 9 present the box-plots of the mean MSEs obtained for the simulation with deep and shallow wow and flutter respectively.

Fig. 8 Mean MSEs obtained for the simulations with the PVCs amplitude range <0.9, 1.1>

Fig. 9 Mean MSEs obtained for the simulations with the PVCs amplitude range <0.99, 1.01>

For the simulations involving samples with the wider PVC range it can be seen that the new method provides better results. The overall MSE, calculated for all of the examined f_{max} values, is smaller for this method comparing to the total MSE of the power line hum tracker. For both algorithms the MSE increases with the higher f_{max} values, however, the degree of increase is smaller for the new method.

For the simulations involving samples with the narrower PVC range the results are different. The overall MSE is higher for the new method. It can be seen that the mean MSEs are the same for all of the examined f_{max} values (here only the standard deviation values allows to differentiate results). The high and constant values of the mean MSEs are caused by the precision limitation of the proposed method. Using the 3600 DPI resolution images the distance between successive perforation holes equals 670 pixels. Knowing that the maximal allowable amplitude variations of the simulated PVCs are +/- 0.01, the maximal length difference in the modulated perforations' distances equals 7 pixels, and in most cases is less than that. Comparing this with the limited precision of the algorithm for calculating distances between successive perforation holes, which is +/- 2 pixels, explains the reported constant value of the MSEs.

5. Conclusions

The new method presented in this paper allows ones to determine course of wow in the entire range of frequency modulations The flutter course can be determined up to the f_{max} equalling around 19 Hz. Additionally, when this value is exceeded, the tracking error increases slowly comparing to errors obtained for the power line hum tracker.

For the deeper pitch variations the new method allows for a better accuracy then the power line hum tracker. However, for the smaller pitch variations precision of the new method is lower. The low precision is caused by the new method accuracy limitation due to the movie tape image resolution.

The new method can be used automatically, provided the movie tape scans are prepared properly. Additionally, the new method can help to overcome problems with comparing of natural and parasitic pitch variations (i.e. wow versus vibrato). Hence, the new method can be used to restore obsolete and rare archival recordings.

6. Summary

In this paper the mew method for reducing parasitic pitch variations in archival movie soundtracks, namely wow and flutter, was presented. The method utilises image processing for calculating the characteristics of movie tape shrinkage causing wow and flutter. Two algorithms for shrinkage evaluation were presented. The first one analyses distances between successive perforation holes and the second one is dedicated to movie frames analysis. Additionally, the algorithm for translating soundtrack images into digital sound was also presented. The new method was compared with the previous approaches to wow and flutter estimation and reduction. It was found on this basis that it outperforms them when analysing distortions with higher amplitudes, whereas the small amplitudes of wow and flutter are better estimated using the previously known algorithms. Moreover, the new method can be used to analyse the entire range of wow modulation frequesncy, however, recordings affected by flutter can be restored to some extent, only.

Acknowledgements

Research subsidized by the Polish Ministry of Science and Higher Education within the Projects No. PBZ-MNiSW-02/II/2007 and No. N N517 4381 33.

References

[1] AES Std. 6-1982, Method for Measurement of Weighted Peak Flutter of Sound Recording and Reproducing Equipment, Audio Engineering Society, New York, 2003.

[2] E. D. Daniel, C. D. Mee, M. H. Clark, Magnetic Recording: The First 100 Years, IEEE Press, New York, 1999.

[3] P. Read and M. P. Meyer, Restoration of Motion Picture Film, Butterworth Heinemann, Oxford, UK, 2000.

[4] E. S. Busby, Analog Tape Recording, in: J. Whitaker, B. K. Benson (Eds.), Standard Handbook of Audio and Radio Engineering, McGraw-Hill, New York, 2003, pp. 6–21-6–42.

[5] I. Malecki, Technika Nagrywania i Odtwarzania Dzwiekow, PWT, Warszawa 1953. (in Polish)

[6] A. Czyzewski, A. Ciarkowski, A. Kaczmarek, J. Kotus, M. Kulesza, P. Maziewski, DSP Techniques for Determining "Wow" Distortion, Journal of the Audio Engineering Society 55 (4) (2007) 266-284.

[7] P. Pastuszak, MTS Pilote Tone Tracking for "Wow" Determination, Archives of Acoustics 33 (1) (2008) 117-123.

[8] P. Maziewski, A. Kupryjanow, A. Czyzewski, Drift, Wow and Flutter Measurement and Reduction in Shrunken Movie Soundtracks, presented at the 124th AES Convention, Amsterdam, The Netherlands, 17-20 May 2008, preprint no. 7392.

[9] M. Gerzon, Don't Destroy the Archives!, unpublished (1992); available at ww.waves.com/german/htmls/service/faq/dont.html (accessed 2006 Oct.).

[10] S. J. Godsill, The Restoration of Degraded Audio Signals, Ph.D. dissertation, Cambridge University, Cambridge, UK, 1993.

[11] 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, 17-20 October 1993, pp. 148-151.

[12] S. J. Godsill, Recursive Restoration of Pitch Variation Defects in Musical Recordings, Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing 2, Adelaide, Australia, 19-22 April 1994, pp. 233-236.

[13] S. J. Godsill, P. Rayner, Digital Audio Restoration - A Statistical Model-Based Approach, Springer, London, 1998, Chapter 8, pp. 171–190.

[14] S. J. Godsill, G. Reid, and C. Hicks, "Restoration of Smooth Pitch Variations over Long Timescales," presented at the Joint Technical Symposium, Toronto, Canada, 24–26 June 2004; available at http://www.jts2004.org/english/proceedings/Reid.htm (accessed October 2008).

[15] J. Nichols, An Interactive Pitch Defect Correction System for Archival Audio, Proc. AES 20th Int. Conf., Budapest, Hungary, 5-7 October 2001, paper no. 1941.

[16] J. Nichols, High Performance, Low-Cost Wax Cylinder Transcription System, Proc. AES 20th Int. Conf., Budapest, Hungary, 5-7 October 2001, paper no. 1937.

[17] J. Howarth, P. Wolfe, Correction of Wow and Flutter Effects in Analog Tape Transfers, presented at the 117th AES Convention, San Francisco, USA, 28-31 October 2004, preprint no.6213.

[18] P. Wolfe, J. Howarth, Nonuniform Sampling Theory in Audio Signal Processing, presented at the 116th AES Convention, Berlin, Germany, 8-11 May 2004, preprint no. 6123.

[19] P. Maziewski, Modulation Frequency Constrains on Wow and Flutter Determination, Archives of Acoustics 33 (1) (2008) 125-131.

[20] P. Maziewski, Wow defect reduction based on interpolation techniques, Bulletin of the Polish Academy of Science: Technical Sciences 54 (4) (2006) 469-477; available at http://bulletin.pan.pl/(54-4)469.pdf (accessed October 2008).

[21] V. Fadeyev, C. Haber, A novel application of high energy physics technology to the problem of audio preservation, Nuclear Instruments and Methods in Physics Research A 518 (2004) 456-462.

[22] E. Brun, A. Hassaine, B. Besserer, E. Decenciere1, Restoration of Variable Area Soundtracks, IEEE International Conference on Image Processing, ICIP 2007, San Antonio, Texas, USA, 16-19 September 2007.

[23] D. Richter, D. Poetsch, A, Kuiper, Localization of Faults in Multiple Double Sided Variable Area Code Sound Tracks on Motion Picture Films Using Digital Image Processing, 13th International Czech-Slovak Scientific Conference Radioelektronika, 2003.

Vitae

Andrzej Czyzewski was born in Gdansk, Poland, in 1956. He received an M.Sc. degree in sound engineering from the Gdansk University of Technology in 1982, a Ph.D. degree in 1987, and a D.Sc. degree in 1992. He was granted the title of professor in 1999 and he has held the position of full professor at Gdansk University since 2002. His main interest is in research and education in the domain of multimedia and studio recording as well as in modern information and communication technologies.

Przemyslaw Maziewski was born in Suwalki, Poland, in 1978. He received an M.Sc. degree from the Faculty of Electronics, Telecommunications and Informatics of the Gdansk University of Technology, Poland. His thesis was related to the problem of sound virtualization techniques implemented in stereo headphones. His fields of interests are related to audio recording, digital signal processing, and sound archiving. Currently he continues his studies at the Gdansk University of Technology as a Ph.D. candidate.

Adam Kupryjanow was born in Bialystok, Poland in 1983. He received his M.Sc. degree in 2007 from the Faculty of Electronics, Telecommunications and Informatics of the Gdansk University of Technology, Poland. His thesis was related to optical soundtrack restoration. Currently he is a student of Ph.D. at the Gdansk University of Technology

Figures

Fig. 1 Example of a badly produced movie tape splice

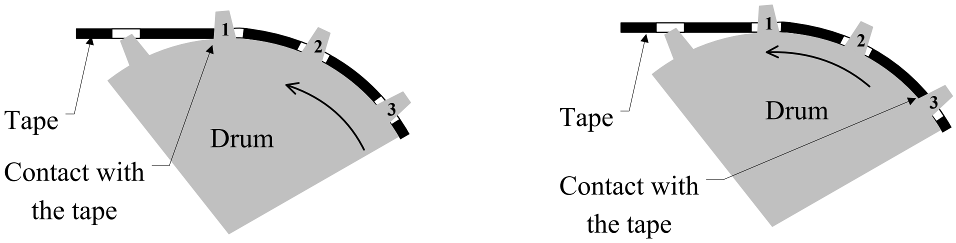

Fig. 2 Tape shrinkage effects on the movie tape transportation system [5]

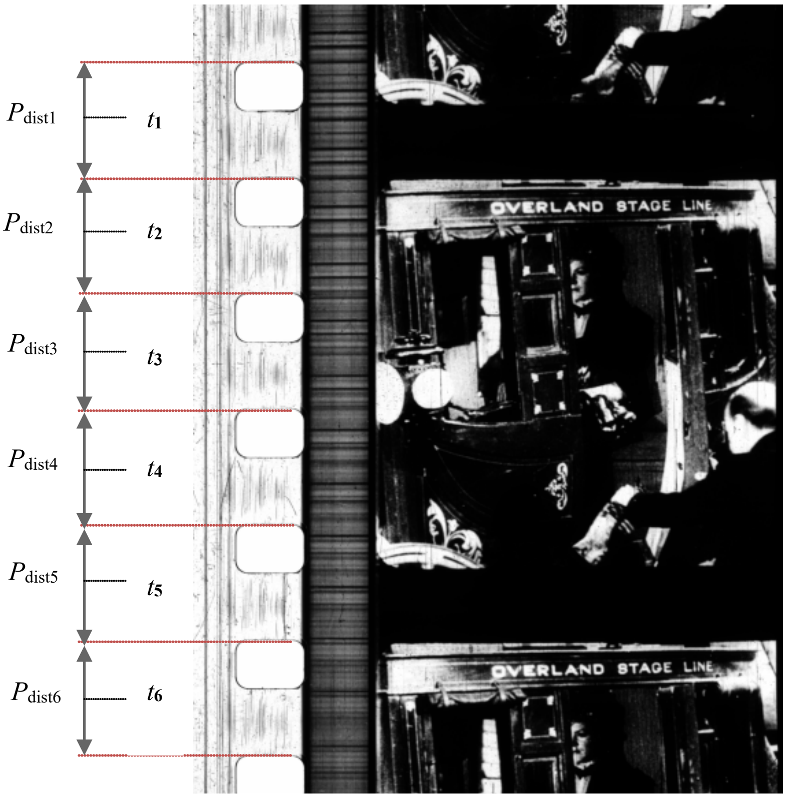

Fig. 3 Example of typical 35 mm movie frame

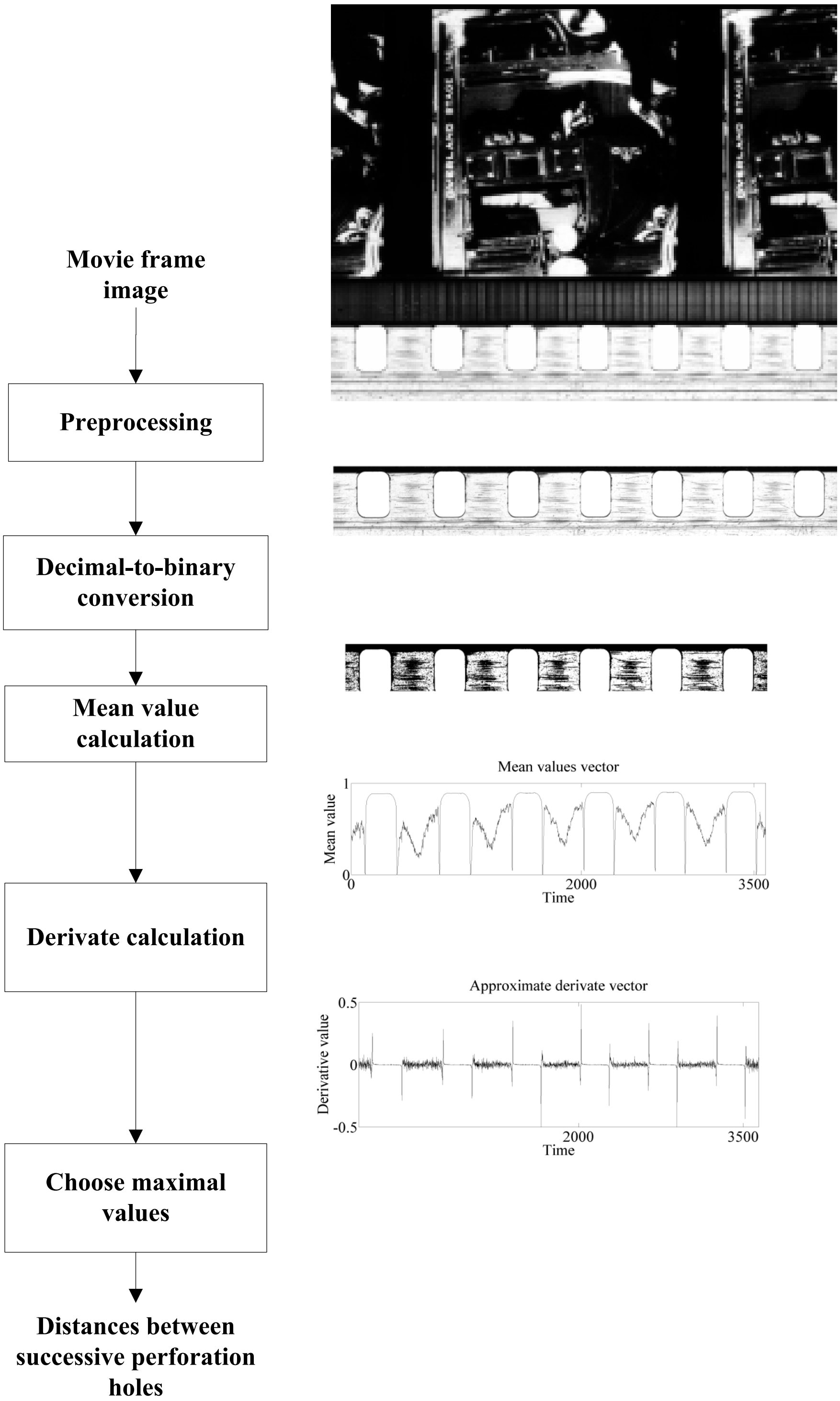

Fig. 4 Algorithm for calculating the distances between successive perforation holes in one movie frame

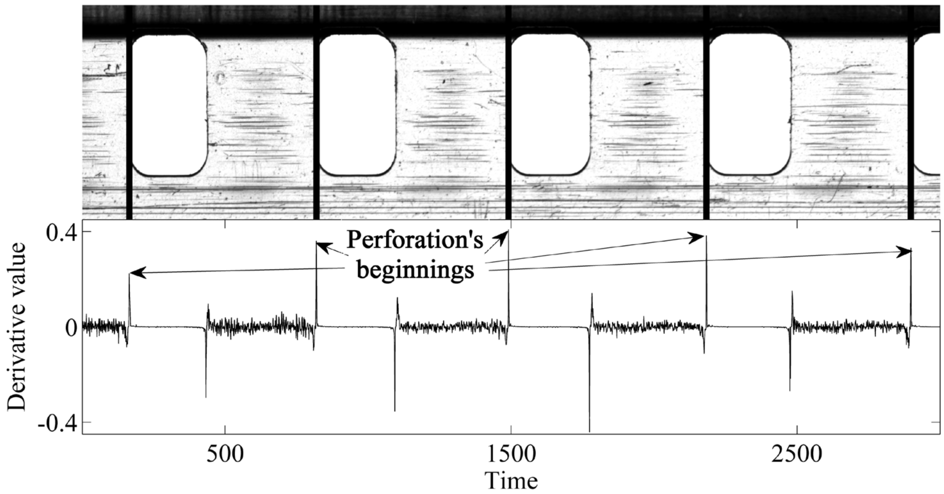

Fig. 5 Example of determined distances between successive perforation holes for 35 mm film (single frame)

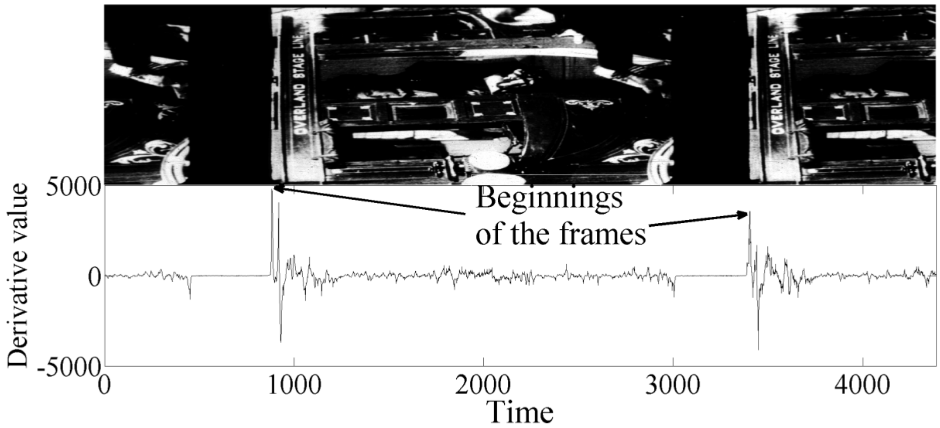

Fig. 6 Example of determined distances between successive movie frames for 35 mm film

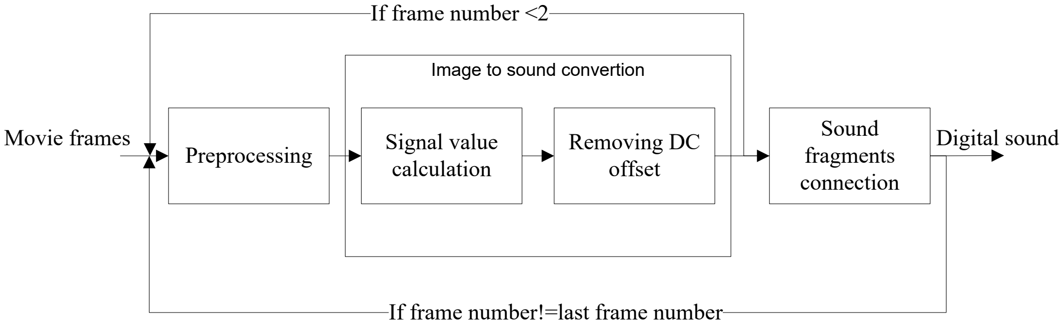

Fig. 7 Scheme of algorithm for translating image of the optical soundtrack to digital sound

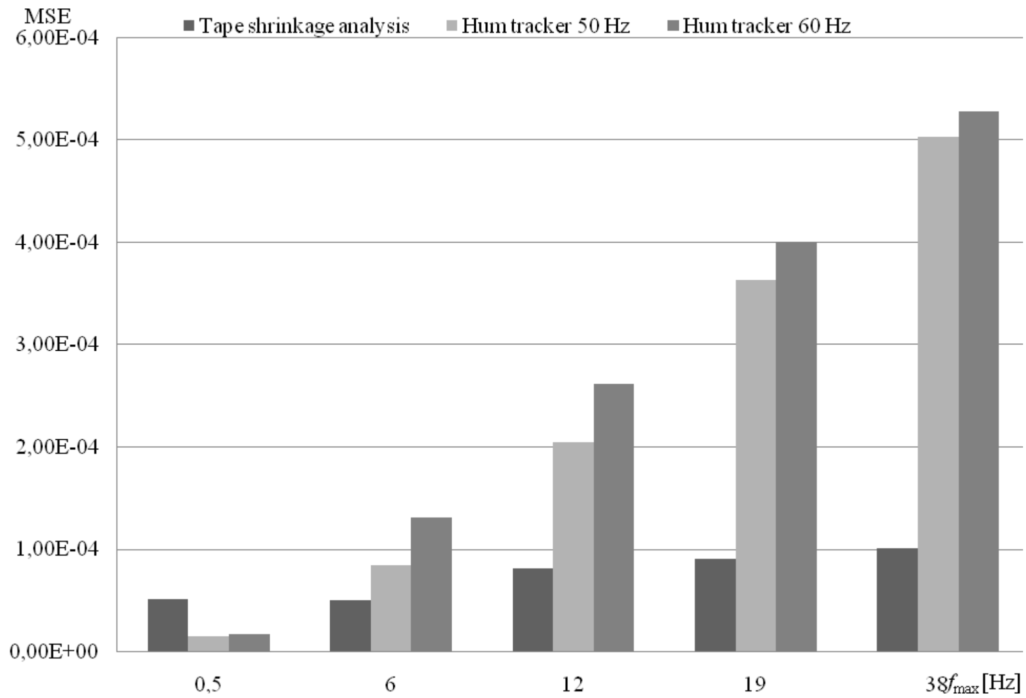

Fig. 8 Mean MSEs obtained for the simulations with the PVC amplitude range <0.9, 1.1>

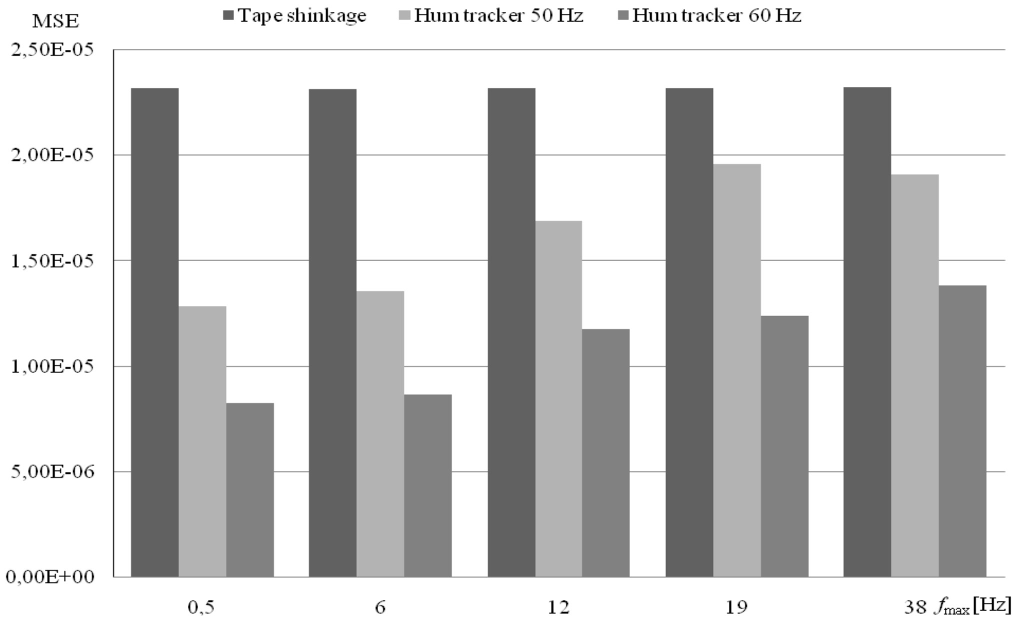

Fig. 9 Mean MSEs obtained for the simulations with the PVC amplitude range <0.99, 1.01>

Tables

Tab. 1 f_{max} values calculated for different film formats utilizing movie frame analysis [8]

Film format [mm]	8	16	35	65/70
Tape velocity [fps]	16	24	24	24
f_{max} [Hz]	3.2	4.8	4.8	4.8

Tab. 2 f_{max} values calculated for different film formats utilizing tape perforation analysis [8]

Film format [mm]	8	16	35	65/70
Tape velocity [pps]	16	24	96	120
f_{max} [Hz]	3.2	4.8	19.2	72

Tab. 3 f_{max} values calculated for different, audio-based, wow and flutter tracking algorithms [19]

Algorithms and their settings	f_{max} [Hz]
Power line hum tracker; M = 8 Sa; F_s =200 Hz;	12.5
COG tracker; M = 1024 Sa; F_s =44100 Hz;	21.5
High-frequency bias tracker; M = 1024 Sa; F_s =192000 Hz	93.75

Tab. 4 MSE of the PVCs determined using the power line hum tracker, 50 Hz tone, PVC amplitude range <0.9, 1.1>

f_{max} [Hz]	0.5	6	12	19	38
Mean value (Std. deviation)	1,48E-05 (1,10E-05)	8,38E-05 (2,78E-05)	2,04E-04 (4,16E-05)	3,63E-04 (6,99E-05)	5,03E-04 (9,40E-05)

Tab. 5 MSE of the PVCs determined using the power line hum tracker, 50 Hz tone, PVC amplitude range <0.99, 1.01>

f_{max} [Hz]	0.5	6	12	19	38
Mean value (Std. deviation)	1,29E-05 (8,78E-06)	1,36E-05 (8,05E-06)	1,69E-05 (7,06E-06)	1,96E-05 (7,10E-06)	1,91E-05 (6,70E-06)

Tab. 6 MSE of the PVCs determined using the power line hum tracker, 60 Hz tone, PVC amplitude range <0.9, 1.1>

f_{max} [Hz]	0.5	6	12	19	38
Mean value (Std. deviation)	1,69E-05 (1,19E-05)	1,31E-04 (7,18E-05)	2,61E-04 (5,14E-05)	4,00E-04 (7,27E-05)	5,27E-04 (8,20E-05)

Tab. 7 MSE of the PVCs determined using the power line hum tracker, 60 Hz tone, PVC amplitude range <0.99, 1.01>

f_{max} [Hz]	0.5	6	12	19	38
Mean value (Std. deviation)	8,27E-06 (6,23E-06)	8,69E-06 (5,64E-06)	1,18E-05 (5,91E-06)	1,24E-05 (5,49E-06)	1,38E-05 (5,11E-06)

Tab. 8 MSE of the PVCs determined using local tape shrinkage analysis, PVC amplitude range <0.9, 1.1>

f_{max} [Hz]	0.5	6	12	19	38
Mean value (Std. deviation)	5,09E-05 (4,74E-05)	5,02E-05 (5,43E-05)	8,09E-05 (0,000154)	9,05E-05 (0,000153)	1,01E-04 (0,000179)

Tab. 9 MSE of the PVCs determined using local tape shrinkage analysis, PVC amplitude range <0.99, 1.01>

f_{max} [Hz]	0.5	6	12	19	38
Mean value (Std. deviation)	2,32E-05 (2,99E-07)	2,32E-05 (2,63E-07)	2,32E-05 (2,84E-07)	2,32E-05 (2,22E-07)	2,32E-05 (2,92E-07)