Intelligent Algorithms for Optical Track Audio Restoration

Andrzej Czyzewski, Marek Dziubinski,
Lukasz Litwic, Przemyslaw Maziewski

Multimedia Systems Department
Gdansk University of Technology
ul. Narutowicza 11/12, 80-952 Gdansk, Poland
ac@pg.gda.pl

Abstract. The Unpredictability Measure computation algorithm applied to psychoacoustic model-based broadband noise attenuation is discussed. A learning decision algorithm based on a neural network is employed for determining audio signal useful components acting as maskers of the spectral components classified as noise. An iterative algorithm for calculating the sound masking pattern is presented. The routines for precise extraction of sinusoidal components from sound spectrum were examined, such as estimation of pitch variations in the optical track audio affected by parasitic frequency modulation. The results obtained employing proposed intelligent signal processing algorithms will be presented and discussed in the paper.

1 Introduction

An approach of spectral subtraction, employing perceptual filtering driven by an intelligent algorithm, used for signal enhancement is presented in this paper. A neural network is a main part of the decision system employed to classify noisy patterns (see [11] for more details). The idea of applying psychoacoustic fundamentals for signal enhancements in terms of perceptual filtering was also demonstrated by others [8]. Number of methods related to noise reduction problem have been proposed, among which Wiener and Kalman adaptive filtration, or spectral subtraction belong to the most frequently applied[4]. However, these methods do not take into account some subjective proprieties of human auditory system[5] successfully exploited in some audio coding standards[2][3]. Additionally, adaptive noise reduction (Wiener and Kalman) suffers from several drawbacks such as: computational complexity, problems in estimation of filter parameters and slow convergence of parameters, which in case of dynamically varying signals (music) result in significant distortions[24][28][25][26]. Spectral subtraction techniques, which are computationally efficient and do not face slow convergence problems, are far more popular in acoustic noise reduction [27]. However, these methods may suffer from artifacts produced by non-ideal estimation of signal and noise estimation in the frequency domain. Both adaptive and spectral subtraction do not employ proprieties of sinusoidal components

versus components with chaotic phase in the signal's and noise parameters estimation.
As it was reported in earlier work [7], employing human auditory system models for parasite noise reduction may be very effective. Application of precise sound perception modeling appears to be necessary for this task and as discussed in [10], it requires implementation of a complex psychoacoustic model [6]rather than the simplified one exploited in the MPEG standard [3].
Another problem related to archive audio recorded in optical tracks is parasitic frequency modulation originated from motor speed fluctuations, tape damages and inappropriate editing techniques. This kind of distortion is usually defined as wow or flutter, or modulation noise (depending on the frequency range of the parasitic modulation frequency). As particularly wow leads to undesirable changes of all of the sound frequency components, sinusoidal sound analysis originally proposed by McAulay and Quatieri [12] was found to be very useful in the defects evaluation. In such approach tracks depicting tonal components changes are processed to obtain precise wow characteristics [13–17]. Notwithstanding all of the cited proposals there is still a need for further algorithmic approach to the wow restoration as it can be very complex sharing periodic or accidental nature. Therefore this paper, similarly as the previous one [16] addresses the problem of wow extraction, however in this case also employing soft computing.

2 Noise removal algorithm

The masking phenomena are fundamental for contemporary audio coding standards [2] [3], although it can be also exploited in noise reduction [7] [10]. More detailed information on psychoacoustics principles of signal processing can be found in abundant literature [5] [6] [9] also including our papers [7] [10] [11]. Significant role in the psychoacoustic modeling play tonality descriptors of spectral components. The tonality may be represented by the *Unpredictability Measure* parameter [1] used for calculation of the masking offset. Masking offset for the excitation of b_x Barks at frequency of b_x Barks is given by the formula:

$$O_{k,x} = \alpha_k^t \cdot (14.5 + bark(x)) + (1 - \alpha_k^t) \cdot 5.5 \tag{1}$$

The tonality index α_k^t of the excitation of b_x Barks is assumed to be directly related to the *Unpredictability Measure* parameter ($\alpha_k^t = c_k^t$), where c_k^t is calculated in the following way:

$$c_k^t = \frac{\sqrt{(r_k^t \cdot \cos \Phi_k^t - \hat{r}_k^t \cdot \cos \hat{\Phi}_k^t)^2 + (r_k^t \cdot \sin \Phi_k^t - \hat{r}_k^t \cdot \sin \hat{\Phi}_k^t)^2}}{r_k^t + |\hat{r}_k^t|} \tag{2}$$

for r_k^t denoting spectral magnitude and Φ_k^t denoting phase, both at time t , while $\hat{r}_k^t$ and $\hat{\Phi}_k^t$ represent the predicted values of Φ_k^t , and are referred to the past information (calculated for two previous signal sample frames):

$$\begin{cases} \hat{r}_k^t = r_k^{t-1} + (r_k^{t-1} - r_k^{t-2}) \\ \hat{\Phi}_k^t = \Phi_k^{t-1} + (\Phi_k^{t-1} - \Phi_k^{t-2}) \end{cases} \Rightarrow \begin{cases} \hat{r}_k^t = 2r_k^{t-1} - r_k^{t-2} \\ \hat{\Phi}_k^t = 2\Phi_k^{t-1} - \Phi_k^{t-2} \end{cases} \tag{3}$$

Thus, based on the literature [6], the masking threshold of the Basilar membrane T , stimulated by the single excitation of b_x Barks and of magnitude equal to S_x is calculated with regard to:

$$\begin{cases} T_{i,x} = S_i \cdot 10^{-s_1 \cdot (b_x - b_i)/10 - O_{i,x}}, & b_x \leq b_i \\ T_{j,x} = S_j \cdot 10^{-s_2 \cdot (b_j - b_x)/10 - O_{j,x}}, & b_x > b_j \end{cases} \tag{4}$$

where S_i,S_j are magnitudes related to excitations b_i,b_j and global masking threshold is obtained by summing up all of individual excitations.

2.1 Perceptual noise reduction system

In the perceptual noise reduction system (Fig. 1), published first at the KES'2004 conference [11], it is assumed that noise is of additive type. Spectral representation of the disturbance is calculated with regard to spectral subtraction techniques [4].

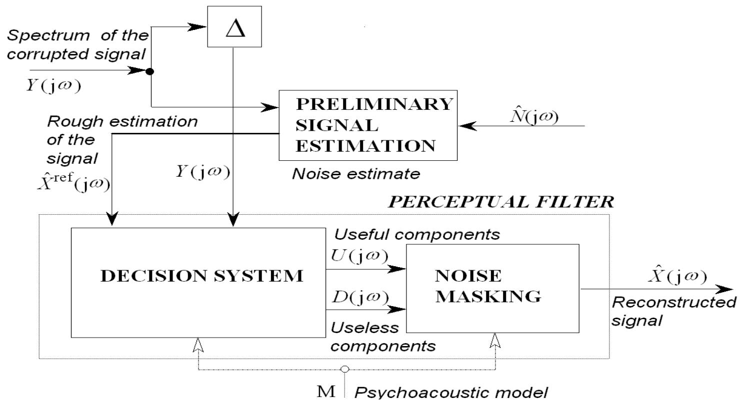

Fig. 1. General lay-out of noise reduction system

Because noise suppression in this approach is based on masking some spectral components of the disturbing noise, it is necessary to determine which components should be masked and which should act as maskers. For this reason, so called rough estimate $\hat{X}^{ref}(j\omega)$ of the clean signal's spectrum is obtained with accordance to spectral subtraction method [4] based on the iterative algorithm represented by the Noise Masking block in Fig. 1.

The proposed algorithm [7] [10] has been recently improved and extended with a learning decision algorithm. The new *Decision System* module [11] containing a neural network is responsible for determining which components are going to be treated as maskers U (*useful* components), and which represent distortions and are going to be masked D (*useless* components). The basic classification (without neural network application described in Sect. 4) can be carried out on the basis of the following expressions:

$$U = \{\hat{X}_i^{ref}; \quad |\hat{X}_i^{ref}| > T_i^{ref} \wedge |Y_i| > T_i^Y, 1 \leq i \leq N/2\} \tag{5}$$

$$D = \{Y_i; \quad |\hat{X}_i^{ref}| \leq T_i^{ref} \vee |Y_i| \leq T_i^Y, 1 \leq i \leq N/2\} \tag{6}$$

where i denotes spectrum component indexes, U and D are sets containing useful and useless information. T^{ref} is the masking threshold caused by the presence of $\hat{X}^{ref}(j\omega)$, and T^{Y} is the masking threshold of the input signal: $Y^{ref}(j\omega)$. Lowering of the masking threshold preserves more noise of the input signal, so the influence of the reconstruction filter is significantly smaller than it is in case of the uplifting method, giving less distorted output signal. Modified global masking threshold T_x^{β} at barks can be calculated with regard to formula:

$$T_x^{\beta} = \sum_{j \in U_L(x)} T_{j,x} + \sum_{j \in D_L(x)} T_{j,x}^{\beta} + \sum_{i \in U_H(x)} T_{i,x} + \sum_{i \in D_H(x)} T_{i,x}^{\beta} \tag{7}$$

where $T_{i,x}^{\beta}$ and $T_{j,x}^{\beta}$ represent new masking thresholds, caused by reduced single excitations and β is vector containing reduction factor values for the noisy components. $U_L(x)$ and $U_H(x)$ (similarly $D_L(x)$ and $D_H(x)$) denote subset of U(or subset of D) containing elements with frequencies lower or equal (L) to b_x barks, and frequencies higher than b_x barks (H).
Since values of β may differ for the elements of D, and changing each value affects T_x^{β} , thus it is impractical to calculate all reducing factor values directly. For this reason sub-optimal iterative algorithm was implemented [11].

2.2 Unpredictability Measure application

Calculation of the masking offset, described by (1) plays a significant role in the masking threshold calculation. In noisy signals, tonal components that are occurring just above the noise floor, may be not very well represented by the *Unpredictability Measure* (UM) parameter due to the strong influence of the noisy content. A practical solution to this problem is extending time domain resolution, by increasing overlap of the frames used only for unpredictability calculation. Standard *Unpredictability Measure* (2-3) refers to the fragment of the signal represented by 3 consecutive frames, i.e. beginning of this fragment (T_{start}) is at the beginning of the frame with $t-2$ index and the end of the fragment (T_{start}) is at the end of frame with t index, with accordance to (3). Consequently, the same fragment is divided into N equally spaced frames, so that the improved UM can be expressed as:

$$\bar{c}_k^t = \frac{1}{N-2} \sum_{n=1}^{N-2} c_k^{t^n} \tag{8}$$

$$\text{where} \qquad c_k^{t^n} = \frac{dist\Big((\hat{r}_k^{t^n}, \hat{\Phi}_k^{t^n}), (r_k^{t^n}, \Phi_k^{t^n})\Big)}{r_k^{t^n} + |\hat{r}_k^{t^n}|} \tag{9}$$

$$\text{and} \begin{cases} \hat{r}_k^{t^n} = r_k^{t^n-1} + (r_k^{t^n-1} - r_k^{t^n-2}) \\ \hat{\Phi}_k^{t^n} = \Phi_k^{t^n-1} + (\Phi_k^{t^n-1} - \Phi_k^{t^n t-2}) \end{cases} \Rightarrow \begin{cases} \hat{r}_k^{t^n} = 2r_k^{t^n-1} - r_k^{t^n-2} \\ \hat{\Phi}_k^{t^n} = 2\Phi_k^{t^n-1} - \Phi_k^{t^n-2} \end{cases} \tag{10}$$

while $T_{start} \le t^n - 2 < t^n - 1 < t^n \le T_{stop}$ and $c_k^t = \bar{c}_k^t$. Additionally, classification of the spectrum components in non-linear spectral subtraction, can be extended by some psychoacoustic parameters, i.e. the tonality description values. By analyzing time-frequency domain behavior of the UM vectors calculated for each frame, it is easy to spot tracks representing harmonic content of the signal. Basing on this observation, artificial neural network was deployed as the decision system for classifying, c_k^{tn} patterns. A set of training data was obtained from the noise fragment and from the noisy signal - c_k^{tn} vectors of the noise represented useless components, while those obtained from the noisy input signal, classified as useful components with standard spectral subtraction algorithm, represented patterns of the useful signal. A three-layer neural network of the feed-forward type was used in the experiments. Its structure was defined as follows:

- Number of neurons in the initial layer is equal to the number of elements in the feature vector
- Number of neurons in the hidden layer is equal to the number of neurons in the initial layer
- Output layer contained one neuron in the output neurons in the initial and the output layers have log-sigmoid transfer functions, while neurons in the hidden layer have tan-sigmoid transfer functions.The weights and biases, were updated during the training process, according to Levenberg-Marquardt optimization method. A method of controlling the generalization process was also used. Such an approach is very effective for recovering sinusoidal components, however it does not significantly improve recovery of non-tonal components. Therefore it should be considered as an extension to the spectral subtraction decision process. The algorithm was verified experimentally (the results are discussed in Sect. 4).

3 Parasitic modulation compensating algorithm

The proposed algorithm (presented in Fig.2) for wow defect evaluation is to estimate the pitch variation function $p_w(n)$ from the contaminated (wow) input signal $x_w(n)$[23]. The first stage of the algorithm, depicted as STFT in Fig.2,

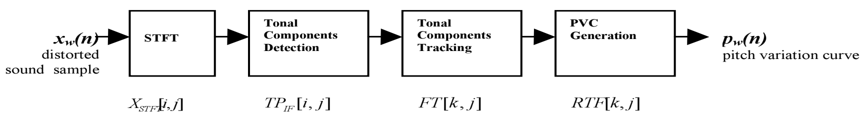

Fig. 2. Block diagram of wow estimation process

results in a time-frequency representation of the input signal. The distorted input signal, initially divided into time-frames (with an appropriate overlapping) is windowed with the Hamming window for better side-lobe suppression. To gain frequency resolution the windowed signal is zero-padded and then it is packed into the buffer for a zero phase spectrum [18]. Finally, Discrete Fourier Transform (DFT) is evaluated for every time-frame buffer to obtain the time-frequency representation .

The next stage of the algorithm (presented in Fig.3) is to detect tonal components of the signal. In the first step candidates for tonal components are detected

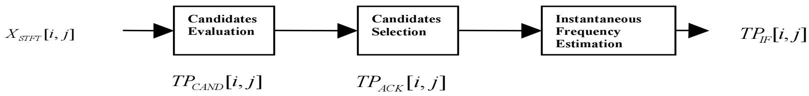

Fig. 3. Structure of the Tonal Components Detection Block

as local maxima (peaks) of a magnitude spectrum stored in X_{STFT}. Since these peaks result either form the main-lobes or the side-lobes of the spectral components it is necessary to exclude the latter ones. It also essential to reject the peaks resulting from localized noise [23]. This task has to be performed according to an appropriate criterion. The most intuitive, an amplitude threshold, recognizes component as a tonal peak when its amplitude is above the certain threshold. As this criterion suffers from many drawbacks other criteria for peak validation have been proposed [19, 20, 23]. Nevertheless, none is enough sufficient by oneself for distorted and noisy musical signals which are of interest of this paper.
We propose an intelligent algorithm, depicted as Candidates Selection block in Fig.2, containing three independent criteria. Those are as follow:

- Sinusoidal Likeness Measure (SLM) - the criterion assumes the tonal peaks in analysis spectrum provide the result of multiplication of the tonal components by the window function. Thus the maxima of cross-correlation function of the main-lobe's spectrum of the candidate and the analysis window would indicate the presence of a sinusoidal components [21]:

$$\Gamma(\omega) = \Big| \sum_{\substack{k \\ |\omega-\omega_k|<B}} X(\omega_k)W(\omega-\omega_k)\Big| \tag{11}$$

  where X and W are spectra of analyzed signal and window respectively, B is the low-pass bandwidth within the cross-correlation evaluated.
  The main drawback of this preliminary criterion is that it may acknowledge a side-lobe component as a tonal part.
- Phase Measure - this criterion assumes that the phase of the tonal component's main-lobe varies significantly less than those resulted from noise. This criterion validates the result of the SLM criterion selection.
- Relative Amplitude Threshold [21] - this criterion can discard some side-lobes as well as components of a minor significance:

$$h(k_p) = |X(k_p)| - 0,5 \cdot |X(k_{v^+})| + |X(k_{v^-})| \tag{12}$$

  where, X represents the DFT array, k - the frequency bin index, subscript p represents the peak, and subscripts v^+,v^- represent the adjacent local minima.

The next step is to estimate their true (instantaneous) frequency values as they are relevant due to the time-frequency resolution trade-off in the STFT representation. As mentioned earlier, the zero padding was applied to gain frequency resolution but still this resolution is constant over the signal bandwidth. Moreover, for signals that are non-stationary within the analysis frame, inter-frame modulations, which effect in peaks smearing may occur [24].

To overcome above mentioned drawbacks, the spectral reassignment method, depicted as Instantaneous Frequency Estimation block (Fig. 3), is employed. This method [22] assigns the tonal component's value to the STFT frequency bin's center of gravity:

$$\hat{\omega}(x;t,\omega) = \omega + \Im\left\{\frac{STFT_{dw}(x;t,\omega)}{STFT_{w}(x;t,\omega)}\right\} \tag{13}$$

where, $STFT_w$ is STFT using analysis window w, $STFT_{dw}$ is STFT employing the first derivative of a window function.
After the tonal components detection the tracking stage (presented in Fig.4), in which the peaks are linked to create trajectories is launched. Since the pitch

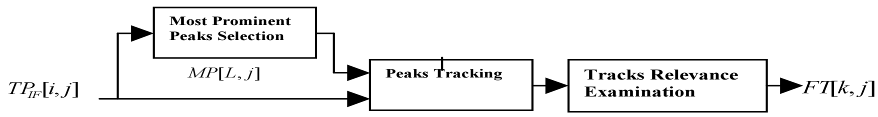

Fig. 4. Structure of the Tonal Components Tracking Block

variation curve $p_w(n)$ is evaluated from the trajectories it is proposed to take into account only the relevant tracks, that can depict the wow defect well. Therefore it is reasonable to form only these trajectories which are based on the most prominent tonal peaks from the magnitude spectrum. The idea for joining tonal components together applies the following frequency criterion [12]:
The $K-th$ tonal component $TP_{IF}[K,j]$ is joined together with the $P-th$ track $FT[P,j]$ when:

$$|TP_{IF}[K,j] - FT[P,j-1]| = min\big(|TP_{IF}[:,j] - FT[P,j-1]|\big) \tag{14}$$

$$\text{and} \qquad |TP_{IF}[K,j] - FT[P,j-1]| < f_{Dev} \tag{15}$$

where, $min(\cdot)$ denotes minimal value and f_{Dev} is the maximum frequency deviation.
The most prominent tonal components, stored in MP matrix, correspond to the peaks of the greatest magnitude hence are considered to be the most perceivable ones. A new track is "born" only from the components stored in MP matrix that were not fitted to any existing tracks. A track is "dead" when there is no continuation according to the frequency criterion.
The last stage of the presented algorithm, according to the diagram in Fig.2, is the PVC generation stage in which the pitch variation function $p_w(n)$ is computed. First, the relative frequencies are calculated for all of the tracks stored in the FT matrix. Next, RFT (Relative Frequency Tracks) matrix is obtained by dividing each track frequency values by the preceding point values. Secondly, median is calculated in RTF columns (i.e. discrete time moments)[16]. Finally, $p_w(n)$ is obtained as a cumulative product of the mean values computed for each discrete time moment.
Since the optimal values of Sinusoidal Likeness Measure (SLM), Phase measure, and Relative Amplitude Threshold are unknown, we applied standard fuzzy logic

reasoning to determining the tonal components $TP_{ACK}[i,j]$ true (instantaneous) frequency values. Due to space limitation for this paper more details on that will be discussed during the paper presenting at the RSFDGrC 2005 in Regina, Canada.

4 Experiments and results

4.1 Experiments concerning noise reduction

It is important to notice, that for the comparison purposes in the informal subjective tests the same spectral subtraction algorithm was used to calculate the *rough estimate* $\hat{X}^{ref}$ as for perceptual reconstruction. Figure 5 presents time-domain changes of the masked noise for a saxophone melody recorded with 44100 Hz sampling rate. The second part of the experiments was devoted to analyze performance of the intelligent unpredictability measure pattern classification employed in spectral subtraction. Below spectrograms (Fig. 6) present signal recovered with standard linear spectral subtraction method, and with spectral subtraction improved by UM vector classification system (as described in Sect. 2.2).

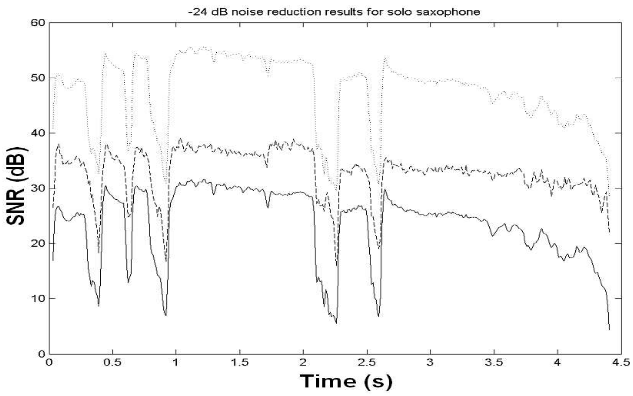

Fig. 5. Time varying SNR for 24 dB noise attenuation, calculated for each processing frame, for input signal (solid line), for perceptually reconstructed signal (dashed line) and for signal restored with spectral subtraction (dotted line), which was used as the *rough estimate* of the restored signal

4.2 Experiments concerning wow compensation

The proposed wow compensating algorithm was tested on several archival sound samples recorded in the Polish National Film Library and the Documentary and Feature Film Studio. Presented example epitomize obtained results. Figure 7 depicts the spectrogram of the sample simultaneously with the detected tracks and the evaluated pitch variation curve (PVC). As can be noticed

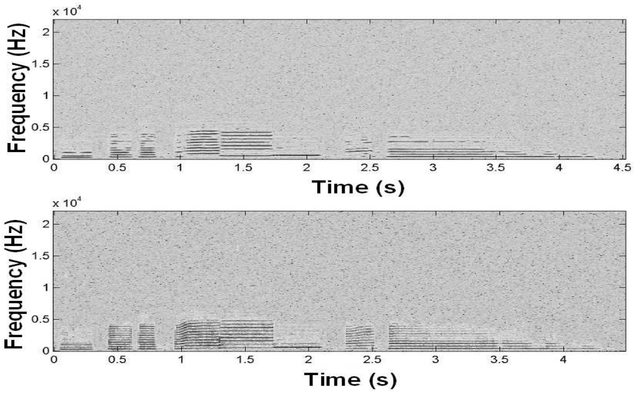

Fig. 6. Spectrograms of signal restored with spectral subtraction (upper plot), and with spectral subtraction enhanced by intelligent pattern recognition system (lower plot)

from the spectrogram only the most relevant tracks are taken into account for the PVC extraction.

Example presented in Fig.7 clearly demonstrates computed PVC and wow

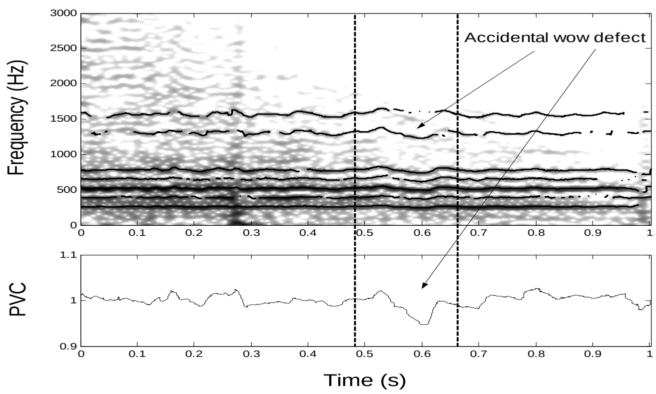

Fig. 7. Tracks detected by the algorithm plotted on the spectrogram with the simultaneously plotted pitch variation curve

defect convergence. It is also worth to mention that the obtained PVC characteristics were successfully utilized in restoration process.

5 Conclusions

The Unpredictability Measure (UM) intelligent pattern recognition system involving UM for spectrum components classification has been presented as an extension of the spectral subtraction algorithms. Applying some properties of the human auditory system to noise reduction allows one to preserve much more of the input signal's energy and consequently enables decreasing unfavourable

influence of the reconstruction filter.
As experiments resulted in satisfactory wow defect evaluation, the presented approach appears to be valid. The presented algorithm manages to detect changes of the pitch variation function $p_w(t)$. However, in case of strong variations of $p_w(t)$ (e.g. accidental wow defect), it still needs to be more robust, especially at the pitch variation tracking stage. Therefore, a further development of this algorithm is underway.

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. Standardised Practices for Audio-visual Contents Archiving in Europe". The research was also subsidized by the Foundation for Polish Science, Warsaw and by the Ministry of Science and and Information Society Technologies - Dec. No. 155/E-359/SPB/6.PR UE/DIE 207/04-05

References

1. Brandenburg, K.: Second Generation Perceptual Audio Coding: The Hybrid Coder. Proceedings of the 90th Audio Eng. Soc. Convention, Montreux (1990) Preprint No. 2937
2. MPEG-4,International Standard ISO/IEC FCD 14496-3, Subpart 4 (1998)
3. Shlien, S.: Guide to MPEG-1 Audio Standard. IEEE Trans. Broadcasting, Vol.40,(1994) 206-218
4. Vaseghi, S.: Advanced Sig. Proc. and Dig. Noise Reduct., Wiley&Teubner, N.York (1997)
5. Zwicker, E, Zwicker, T.: Audio Engineering and Psychoacoustics: Matching Signals to the Final Receiver,the Human Auditory System. J. of Audio Eng. Soc, Vol. 39, No. 3, (1991) 115-126
6. Beerends, J., Stemerdink, J.: A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation. J. of Audio Eng. Soc., Vol. 40, No. 12, (1992) 963-978
7. Czyzewski, A., Krolikowski, R.: Noise Reduction in Audio Signals Based on the Perceptual Coding Approach. Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, October, New Paltz, New York (1999) 147-150
8. Tsoukalas, D., et al.: Perceptual Filters for Audio Sig. Enh, J. of Audio. Eng. Soc., Vol.45, No.1/2, (1997) 22-36
9. Humes, L.: Models of the Additivity of Masking. J. Ac. Soc. Of Am., Vol. 85,(1989) 1285-1294
10. Krolikowski, R., Czyzewski A.: Noise Reduction in Acoustic Signals Using the Perceptual Coding. 137th Meeting, Acoust. Soc. Of Am., Berlin, Germany, (1998) CD-Preprint
11. Czyzewski, A., Dziubinski, M.: Noise Reduction in Audio Employing Spectral Unpredictability Measure and Neural Net. Knowledge-Based Intelligent Information and Engineering Systems: 8th International Conference, KES'04, LNAI [Lecture

Notes in Artificial Intelligence] 3213, Springer - Verlag, Berlin, Heidelberg, Wellington, New Zealand, September (2004), Part I: 743-750
12. McAulay, J., Quatieri, T.F.: Speech analysis/synthesis based on a sinusoidal representation.IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. 34, No. 4, August, (1986) 744-754
13. Godsill, J. S., Rayner, J. W.: The restoration of pitch variation defects in gramophone recordings. Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics,October, New Paltz, New York(1993)
14. Godsill, J. S.: Recursive restoration of pitch variation defects in musical recordings, Proc. International Conference on Acoustics, Speech, and Signal Processing, Vol. 2, April, Adelaide, (1994) 233-236
15. Walmsley, P. J., Godsill, S. J., Rayner, P. J. W.: Polyphonic pitch tracking using joint Bayesian estimation of multiple frame parameters. Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics,October, New Paltz, New York (1999)
16. Czyzewski, A., et. al.: Wow detection and compensation employing spectral processing of audio. 117 Audio Engineering Society Convention, Convention Paper 6212, October, San Francisco (2004)
17. Nichols, J.: An interactive pitch defect correction system for archival audio. AES 20th International Conference, October, Budapest (2001)
18. Serra, X.: Musical Sound Modeling with Sinusoids plus Noise.In: Pope, S., Picalli, A., De Poli, G., Roads, C. (eds.):Musical Signal Processing, Swets & Zeitlinger Publishers, (1997)
19. Rodet, X.:Musical Sound Signal Analysis/Synthesis: Sinusoidal + Residual and Elementary Waveform Models. Proc. IEEE Symp. Time-Frequency and Time-Scale Analysis, (1997)
20. Lagrange,M.:, Marchand,S., Rault, J.B.: Sinusoidal parameter extraction and component selection in a non-stationary model. Proc. of the 5th Int. Conference on Digital Audio Effects, September, Hamburg (2002)
21. Masri, P.: Computer Modeling of Sound for Transformation and Synthesis of Musical Signals, PhD thesis, University of Bristol, (1996)
22. Auger, F., Flandrin, P.: Improving the readability of time-frequency and time-scale representations by the reassignment method. IEEE Trans. on Signal Processing, Vol.43, No. 5, May (1995) 1068-1089
23. Czyzewski, A., et. al.: New Algorithms for Wow and Flutter Detection and Compensation in Audio. Paper to be presented (material submitted to pre-printing) for the 118th Audio Engineering Society Conv.,May , Barcelona (2005)
24. Ifeachor, E., Jervis, B.: Digital Signal Processing. A Practical Approach, Addison-Wesley Publishing Company, (1993)
25. Kalman, R.:A New Approach to Linear Filtering and Prediction Problems. Trans.of the ASME, J. of Basing Engineering, Vol. 82, March (1960) 34-35
26. Moghaddamjoo, A., Kirlin, L.: Robust Kalman Filtering with Unknown Inputs. IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. 37, No. 8, August (1989) 1166-1175
27. Vaseghi, S., Frayling-Cork, R.:Restoration of Old Gramophone Recordings. J. of Audio Eng. Soc., Vol. 40, No. 10, October (1997) 791-800.
28. Widrow, B., Stearns, S.:Adaptive Signal Processing, Prentice-Hall Intl. Inc., New Jersey (1985)