Intelligent Algorithms for Movie Sound Tracks Restoration

Andrzej Czyżewski, Marek Dziubiński, Łukasz Litwic,
and Przemysław Maziewski

Multimedia Systems Department
Gdansk University of Technology
ul. Narutowicza 11/12, 80-952 Gdańsk, Poland
{andcz, mdziubin, llitwic, przemas}@sound.eti.pg.gda.pl

Abstract. Two algorithms for movie sound tracks restoration are discussed in the paper. The first algorithm is the unpredictability measure computation applied to the psychoacoustic model-based broadband noise attenuation. A learning decision algorithm, based on a neural network, is employed for determining useful audio signal components acting as maskers of the noisy spectral parts. An application of the rough set decision system to this task is also considered. An iterative method for calculating the sound masking pattern is presented. The second of presented algorithms is the routine for precise evaluation of parasite frequency modulations (wow) utilizing sinusoidal components extracted from the sound spectrum. The results obtained employing proposed intelligent signal processing algorithms, as well as the relationship between both routines, will be presented and discussed in the paper.

Keywords: Audio restoration, noise reduction, wow evaluation.

1 Introduction

Noise is a common disturbance in archival recordings and a suitable solution is presented in this paper. Acoustic noise reduction is a subject of extensive research, carried out in the last decades. Several approaches during this time were studied, such as adaptive filtering [1,2,3], autocorrelation [4,5] and statistical methods [6,7,8], parametric models for spectrum estimation [9,10], and some techniques based on intelligent algorithms (including rough set - based audio signal enhancement approach) have been investigated in recent years [11,12]. In addition, multi-channel representation of signals was considered, with regard to microphone matrices [13,14,15]. The main stream approaches, which are based on assumption that undistorted signal is not correlated with parasite noise and the noise is stationary, of additive type are: Kalman [10][16][17] and Wiener filtration [18,19], stochastic modeling of the signals [20,21] and spectral subtraction [22,23]. These methods however, do not utilize concepts of perceptual filtration, thus they do not take into account some subjective properties of the human auditory system [24].

J.F. Peters and A. Skowron (Eds.): Transactions on Rough Sets V, LNCS 4100, pp. 123–145, 2006.

The engineered algorithm, described in this paper, utilizes phenomena related to sound perception and is based on perceptual filtering. In addition, the restoration process is controlled by an intelligent algorithm. The intelligent reasoning, based on a neural network, is the core decision unit responsible for classifying noisy patterns. The approach has been presented in authors earlier paper [25]. Application of perceptual filtering has been exploited by several researchers [26], for various audio processing applications, such as audio coding standards [27,28]. As it was demonstrated in our earlier work [29], utilizing concepts of perceptual filtering for noise removal may be very effective. However such an approach requires employing a precise sound perception modeling [30], rather than implementation of the simplified one exploited in the MPEG coding standard [27].

Another problem related to archive audio recorded in movie sound tracks is parasitic frequency modulation (FM) 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. Wow defect is typically defined as frequency modulation in the range up to 6Hz, flutter is the frequency modulation between 6-15Hz and modulation noise (or scrape flutter)describes 96Hz frequency modulation. In this paper we will focus on the wow defect.

As particularly wow leads to undesirable changes of all of the sound frequency components, sinusoidal sound analysis originally proposed by McAulay and Quatieri [31] was found to be very useful in the defects evaluation. In such an approach tracks depicting tonal components changes (MQ tracks) are determined to obtain a precise wow characteristic [32][33]. The statistical methods for post-processing of MQ tracks for monophonic audio signals were introduced by Godsill and Rayner [34][35][36]. Their approach is build on three processing steps. Firstly, a DFT magnitude-based peak tracking algorithm is used for tonal component estimation. It provides a set of harmonic components trajectories. This set constitutes data for the FM estimation which leads to the second processing step. Each trajectory is denoted by its center frequency being unknown and considered as varying one. The variations are attributed both to the investigated FM (tracking misleading) and unknown noisy-like components (e.g. genuine musical pitch deviations). The noise-like components are assumed to be independent, identically distributed Gaussian processes. Thereby, the likelihood function of the unknown center frequencies and the parasite FM are obtained. Further the maximum likelihood solution (ML) is used for wow characteristic evaluation. Details of this method can be found in the papers [34][35][36]. Additionally a priori information on wow distortion can be introduced through Bayesian probability framework. The maximum a posteriori estimator (MAP) was proposed by Godsill [33]. In the last processing step, the estimated wow characteristic enables signal restoration. The sinc based non-uniform resampling was used for this purpose. For details on the non-uniform resampling in wow restoration the literature can be inquired [38].

A method for statistical processing of MQ trajectories for polyphonic audio signals was introduced by Walmsley, Godsill and Rayner. The method is an