PAN ARCHIVES OF ACOUSTICS
Vol. 37, No. 3, pp. 395–398 (2012)

DOI: 10.2478/v10168-012-0048-x

Chronicle

Joint NTAV/SPA 2012 Conference
Lodz University of Technology
Poland, September 27 – 29, 2012

The NTAV/SPA 2012 conference was held on 27–29th September 2012 and was organized by the Institute of Electronics, Lodz University of Technology (www.eletel.p.lodz.pl) with the support of the IEEE Polish Section Region 8, Polish Section of the Audio Engineering Society, Department of Acoustics, Wroclaw University of Technology and the Division of Signal Processing and Electronic Systems, Poznan University of Technology.

This scientific event has joined two successful series of conferences: the New Trends in Audio and Video (NTAV) conference – organized biannually since 1994 by major technical universities in Poland and the IEEE Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA) conference – first held in 1993 and then organized yearly since 1999 by the Poznan University of Technology.

The scope of the NTAV/SPA 2012 joint conference has covered a broad range of topics related to audio, image and video processing, digital signal processing theory, and numerous practical applications including human-computer interaction systems and medicine (see the conference webpage at www.ntavspa.pl for a conference programme).

Fourteen of the accepted submissions were devoted to audio processing, analysis or acoustics. They can be grouped into the following thematic fields: processing and coding techniques of audio signals (including cepstral analysis and time-scale decomposition), music modeling and matching, sound and melody recognition, multimedia, acoustics and psychoacousitcs. We hope that the provided list of abstracts will encourage the readers of the Archives of Acoustics to probe further into the presented new trends of audio processing and analysis.

On behalf of the conference co-organizers the Polish Section of the Audio Engineering Society and the Polish Association of Theoretical and Applied Electrotechnics the NTAVSPA 2012 Organizing Committee acknowledges the financial support of the Ministry of Science and Higher Education.

Pawel Strumillo
Chairman of the Organizing Committee

Abstracts

Hybrid sinusoidal modeling of music with near transparent audio quality

Bartkowiak Maciej, mbartkow@multimedia.edu.pl
Januszkiewicz Łukasz

Poznań University of Technology
Chair of Multimedia Telecommunications
and Microelectronics
Polanka 3, 60-965 Poznań, Poland

It is often believed that sinusoidal as well as sinusoidal plus noise modeling is not capable of delivering high audio quality for complex signals such as wideband music. We identify the key sources of modeling artifacts in sinusoidal modeling systems and demonstrate a hybrid system that offers near transparent quality of reconstructed audio, thanks to application of a dedicated transient model, an accurate parameter estimation method, an advanced tracking algorithm and a warped frequency spectral model of noise.

Keywords: audio analysis/synthesis; sinusoidal model; estimation; transient modeling; noise modelling.

⋆ ⋆ ⋆

Detection of tampering in lossy compressed digital audio recordings

Korycki Rafał, r.korycki@ire.pw.edu.pl

Warsaw University of Technology
Institute of Radioelectronics
Nowowiejska 15/19, 00-665 Warszawa, Poland

This paper addresses the problem of tampering detection and discusses one of the methods used for authenticity analysis of digital audio recordings. Presented approach is based on checking frame offsets in audio files compressed by using perceptual audio coding. This method can be further improved by applying histogram analysis of modified discrete cosine transform spectrum and detection of maximum number of nonzero spectral coefficients. The influence of compression algorithms' parameters on detection of forgeries are presented by applying AAC and Ogg Vorbis encoders as examples. The effectiveness of tampering detection algorithms proposed in this paper is tested on a predefined music database and compared graphically using ROC-like curves.

Keywords: component: tampering detection, digital audio authenticity, lossy compression, frame offsets, MDCT, spectral coefficients.

⋆ ⋆ ⋆

Analysis of damping materials in a transmission line loudspeaker system

Lusztak Krzysztof, krzysiek@lusztak.pl
Bujacz Michał, bujaczm@p.lodz.pl

Lodz University of Technology
Wólczańska 211/215, 90-924 Łódź, Poland

The presented paper contains an analysis of the influence of various types of damping materials in a transmission line loudspeaker enclosure on the acoustic emission spectrum. Damping of the tunnel is crucial in the design of this type of enclosures, as the intended effect is to maximally utilize the tunnel's "self-resonance" (when the tunnel output and loudspeaker output are in phase) at low frequencies, but dampen out and smooth the anti-resonances at higher frequencies. Six types of materials traditionally used in home speaker construction were studied: technical felt, upholstery wool, two thicknesses of plain foam and two types of profiled foam. The tests were conducted in an anechoic chamber using a custom designed loudspeaker enclosure. All types of dampening materials successfully reduced the emission of the tunnel and slowed the sound wave progression, virtually enlarging the enclosure. The profiling of the material did not influence the output as much as the material thickness, that is why the thick foams were best at smoothing the frequency characteristics.

Keywords: loudspeaker, vented enclosure, bass-reflex, sealed enclosure, transmission line, waveguide, damping materials, felt, wool, foam.

⋆ ⋆ ⋆

Follow That Tune – Dynamic Time Warping Refinement for Query by Humming

Stasiak Bartłomiej, basta@ics.p.lodz.pl

Lodz University of Technology
Institute of Information Technology
Wólczańska 211/215, 90-924 Łódź, Poland

Dynamic Time Warping is a standard algorithm used for matching time series irrespective of local tempo variations. This type of variability is inherent to audio input data obtained directly from users and, as such, it occurs in the context of Query-by-Humming interface to multimedia databases. Apart from the time-alignment problem, most of the known melody matching approaches are also affected by a second issue of aligning the pitch between the query submitted by a user and the template. The query is usually in a different key and it may be simply sung out of tune, which needs some additional, sometimes computationally expensive processing and may not guarantee the success e.g. in the presence of pitch trend or accidental key changes. The method of tune following, proposed in this paper, enables to solve the pitch alignment problem in an adaptive way inspired by the human ability of ignoring typical errors occurring in sung melodies. The experimental validation on the database containing 4431 queries and over 5000 templates confirmed the enhancement introduced by the proposed algorithm in terms of the global recognition rate.

Keywords: melody matching, query-by-humming.

⋆ ⋆ ⋆

MusicEar – a System for Real Time Analysis and Archivization of Violin Sound

Łukasik Ewa, Ewa.Lukasik@cs.put.poznan.pl
Makowski Marcel, Malchrowicz Leszek
Nawracała Tomasz, Robak Adam

Poznan University of Technology
Institute of Computing Science
Piotrowo 2, 60-965 Poznań, Poland

The paper presents a prototype of a portable PC based system, called MusicEar, allowing for real time analysis archiving and visualization of violin sound and thus supporting real time violin sound quality assessment. MusicEar visualizes the frequency spectrum in different scales including the Duennwald frequency regions designated specially for timbre characteristics of violins. It also calculates the instantaneous parameters based on Long Term Averaged Spectrum (LTAS). It enables comparison of characteristics of instruments sound already saved in the database.

Keywords: component, violin, musical instruments, sound analysis, sound quality assessment, real time system, audio database.

⋆ ⋆ ⋆

Melody Recognition System

Adamska Katarzyna
Pełczyński Paweł, pawel.pelczynski@p.lodz.pl

Lodz University of Technology
Wólczańska 211/215, 90-924 Łódź, Poland

The purpose of the presented research was the development of a melody search system that would allow to find a song in a music database, based on humming the tune of a known song fragment. The melody recognition in the developed system is based on comparing vectors of voice pitch values. The best match of the humming pitch vector against all the recordings in the database is searched. An original frequency and time scaling approach was implemented to improve recognition accuracy. The system performance was tested with the use of live humming recordings of four volunteers with a small database.

Keywords: melody search, note frequency, query by humming, melodic contour.

⋆ ⋆ ⋆

Automatic Analysis System of TV Commercial Emission Level

Spaleniak Paweł, papol@sound.eti.pg.gda.pl
Kostek Bożena, bozenka@sound.eti.pg.gda.pl

Gdańsk University of Technology
Multimedia Systems Department
Narutowicza 11/12, 80-233 Gdańsk, Poland

The aim of this paper is to present the engineered application that enables to analyze the level of the emitted TV commercials. First, a review of some chosen existing methods of the loudness measurement contained in ITU and EBU recommendations is shortly presented. Then, a prototype of a system implemented in Embarcadero C++ Builder 2010 using C++ programming language is described. A database containing recordings of television programs is prepared and employed in the experiments. Results of automatic loudness measurement of TV commercials are then shown.

Keywords: content classification; detection of TV commercials, loudness measurements.

⋆ ⋆ ⋆

Application of Intrinsic Time-Scale Decomposition in analyzing sigma-delta modulator for audio DAC

Lewandowski Marcin, Marcin.Lewandowski@ire.pw.edu.pl

Warsaw University of Technology
Electroacoustic Division, Institute of Radioelectronics
Nowowiejska 15/19, 00-665 Warszawa, Poland

The paper introduces a new approach for analyzing and processing non-stationary audio signals obtained from underlying nonlinear systems such as digital sigma-delta ($\Sigma\Delta$) audio DACs. Their parameters and performance depend mainly on features of digital $\Sigma\Delta$ modulators. The Intrinsic Time-Scale Decomposition (ITD) method can self-adaptively decompose input signal into a monotonic trend (baseline signal) and a set of proper rotation components (PRCs) for which instantaneous parameters of signal are well defined. Finally, correlations between quantization noise and input signal (in particularly noise modulation) in digital $\Sigma\Delta$ modulators and possibility of the ITD method application for analyzing noise modulation is investigated.

Keywords: digital $\Sigma\Delta$ audio DAC, digital $\Sigma\Delta$ modulator, noise modulation, intrinsic time-scale decomposition.

⋆ ⋆ ⋆

Automatic Identification of Bird Species: a comparison between kNN and SOM classifiers

Kamińska Dorota[1], dorota.kaminska@p.lodz.pl
Gmerek Artur[2], artur.gmerek@p.lodz.pl

Lodz University of Technology
[1] Institute of Mechatronics and Information Systems
[2] Institute of Automatic Control
Stefanowskiego 18/22, 90-924 Łódź, Poland

This paper presents a system for automatic bird identification, which uses audio input. The experiments have been conducted on three group of birds, which were created basing on birdsong similarity. Starting from the training process and finishing on classification, the system is fully automated. The main problem in automatic bird recognition (ABR) is the choice of proper features and classifiers. Identification has been made using two classifiers – kNN (k Nearest Neighbor) and SOM (Self Organizing Maps). System has been tested using data extracted from natural environment.

Keywords: birds, kNN, HMM, recognition, identification, self organizing maps, SOM.

⋆ ⋆ ⋆

Estimation of Interaural Time Difference from measured Head Related Impulse Responses

Pec Michal, michal.pec@gmail.com
Strumiłło Paweł

Lodz University of Technology
Wólczańska 211/215, 90-924 Łódź, Poland

A novel approach to estimation of the Interaural Time Difference (ITD) from the measured Head Related Impulse Responses (HRIR) is proposed in the paper. Accurate estimation of the ITD is essential for a minimum phase filter implementation of the HRIRs. An innovative application of the cross-correlation function that is estimated for impulse responses corresponding to adjacent sound arrival directions makes the presented method robust and immune to inherent noise components occurring in the measured HRIRs. It was shown that the proposed method for estimation of the ITD outperforms earlier proposed methods that are based on: interaural cross correlation, threshold detection and linear phase fit.

Keywords: Head Related Transfer Functions (HRTF), Interaural Time Difference (ITD), minimum phase filter.

⋆ ⋆ ⋆

Design of DSP Supported Systems for Real-Time Voice Watermarking

Pawłowski Paweł, pawel.pawlowski@put.poznan.pl
Weychan Radosław, radoslaw.weychan@put.poznan.pl
Dąbrowski Adam, adam.dabrowski@put.poznan.pl
Drgas Szymon, szymon.drgas@put.poznan.pl
Meyer Andrzej, andrzej.meyer@put.poznan.pl

Poznan University of Technology
Department of Computing, Division of Signal Processing and Electronic Systems
Strzelecka 11, 60-965 Poznań, Poland

A problem of design of a system for real-time voice watermarking using a digital signal processor (DSP) is studied and presented in this paper. The authors prepared and compared three versions of the considered system Rusing different types of data formats (fixed-point and floating point) and various ways of programming (Matlab/Simulik compiler for Code Composer Studio and plain C/C++ programming)

Keywords: DSP, watermark, voice, DWT, real-time, design.

⋆ ⋆ ⋆

Influence of GSM coding on speaker recognition using short Polish sequences

Krzykowska Agnieszka
Marciniak Tomasz, tomasz.marciniak@put.poznan.pl
Weychan Radosław, Dąbrowski Adam

Poznan University of Technology
Department of Computing, Division of Signal Processing and Electronic Systems
Strzelecka 11, 60-965 Poznań, Poland

The paper presents an influence of the suitable model for effective identification of the speaker, detected on the basis of recordings of telephone conversations. Additional knowledge of the type of the encoder used during the transmission of speech allows to apply a model that takes specific characteristics of the encoder into account. This improves efficiency of the speaker recognition process. During the experimental research we used our database of short voice phrases that usually occur during emergency calls. This paper is based on our previous studies related to techniques for the GSM encoding detection and to algorithms for removing silence in the voice recordings.

Keywords: GSM, GMM, speaker identification.

⋆ ⋆ ⋆

Feature generator for speaker recognition using the fusion of cepstral and mel-cepstral parameters

Majda Ewelina, EMajda@wat.edu.pl
Dobrowolski Andrzej P., ADobrowolski@wat.edu.pl

Military University of Technology
Faculty of Electronics
S. Kaliskiego 2, 00-908 Warszawa, Poland

The paper examines issues related to the determination of features distinctive to sound generators using a fusion of mel-cepstral and cepstral information for an Automatic Speaker Recognition (ASR) system. Parameterization of the speech signal is crucial to these systems, as the chosen parameterization dictates the effectiveness of the diagnosis and the speed of the system. The authors focus on the use of speech signal processing methods that consider the phenomena connected with the speech generation process while searching for features related to speech characteristics. A well-designed system should be able to extract speech characteristics independent of the linguistic content of the speech. The research presented in this paper focuses primarily on multicriteria optimization of the generator parameters based on a set of descriptors derived from the fusion of the melcepstra and cepstra and also considers the use of additional features. Finally, the evaluation of the results was based on the analysis of the Fisher coefficients and main components of a set of descriptors.

Keywords: automatic speaker recognition, feature generator, features selection, PCA.

⋆ ⋆ ⋆

Sample Rate Conversion with Fluctuating Resampling Ratio

Blok Marek[1], mblok@eti.pg.gda.pl
Drózda Piotr[2], PDrozda@advaoptical.com

[1] Gdańsk University of Technology
Faculty of Electronics, Telecommunications and Informatics
G. Narutowicza 11/12, 80-233 Gdańsk, Poland
[2] ADVA Optical Networking Sp. z o.o.
Śląska 35/37, 81-310 Gdynia, Poland

In this paper a sample rate conversion with continuously changing resampling ratio has been presented. The proposed implementation is based on variable fractional delay filter implemented using a Farrow structure. It have been demonstrated that using the proposed approach instantaneous resampling ratio can be freely changed. This allows for simulation of audio recorded on magnetic tape with nonuniform velocity as well as removal of these distortions using the same algorithm.

Keywords: sample rate conversion, instantaneous resampling.

⋆ ⋆ ⋆

Influence of bitrate on the loudness measured in the audio broadcast stream

Plaskota Przemysław, przemyslaw.plaskota@pwr.wroc.pl

Wroclaw University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

Radio and television broadcasters increasingly encounter the problem of loudness in broadcast programmes. The problem occurs at all stages beginning with the production of a programme, ending with the transmission. In order to meet the expectations of broadcasters and receivers of radio and television content the International Telecommunication Union (ITU) has developed an algorithm for measuring the loudness of programmes simultaneously introducing a measurement unit for loudness, LU. In the paper the influence of bitrate on the loudness measured in the audio stream of TV programmes has been presented.

Keywords: loudness, lossy audio compression, broadcasting.

⋆ ⋆ ⋆