Archives of Acoustics Vol. **47**, No. 4, pp. 581–586 (2022), doi: 10.24425/aoa.2022.142904

Chronicle

19th Symposium on New Trends in Audio and Video Technology NTAV2022 October 13 – 15, 2022, Wroclaw, Poland

19th Symposium on New Trends in Audio and Video Technology NTAV2022 was held in Wroclaw on October 13–15, 2022. It was another edition of the regular event organized by the Polish Section of the Audio Engineering Society in cooperation with the Polish Acoustical Society and the Department of Acoustics, Multimedia and Signal Processing of Wroclaw University of Science and Technology.

The conference was attended over 60 participants representing all academic centers from Poland working in the field of audio and video and signal processing. The Symposium was also attended by representatives of companies dealing with multimedia technology and IT in the wide sense. There were 5 plenary lectures and more than 20 contributed papers. One of the elements of the Symposium was a Competition for Young Authors, in which 5 prizes were awarded, sponsored by the Foundation for the Support of the Development of Radio and Multimedia Techniques and the Polish Acoustical Society. There was also a competition for audio and video recordings, with the 3 winners receiving prizes sponsored by Dolby Poland, the Polish Section of AES and the Polish Acoustical Society.

During the Symposium, there were presentations of spatial audio, personalized HRTF, virtual reality systems, as well as workshops given by practical specialists in multimedia techniques, virtual reality and data processing. Two technical tours were also organized: to Dolby Poland and the National Forum of Music. The conference topics covered a wide range of issues related to audio engineering, quality assessment and virtual reality creation and data management.

Abstracts

Educational computer game supporting the skills with in timbre solfege

Paulina Bielesz

University of Silesia in Katowice
Katowice, Poland

An educational computer game based on the issues inherent in timbre solfege. An innovative tool is created, a computer environment based on the Unity game engine, working with the FMOD sound engine, which allows shaping the tone of the sound. The game operating with a coherent, composed world of sounds supposed to be also a didactic tool. Supporting the developing of skills in the field of timbre solfege, having all features the computer game should have. The aim of the study is making a prototype this type of a game. A Game Design Document was created, i.e. a document describing the game's plot, level, description, mechanics, it presents graphics, sound layers of the game.

⋆ ⋆ ⋆

Comparing performance of deep learning methods for fundamental frequency extraction

Filip Bończyk, Ewa Łukasik

Poznan University of Technology
Poznan, Poland

Automatic fundamental frequency (F0) extraction is one of the central topics in research areas like Music Information Retrieval (MIR) or prosody of speech analysis. Over the years, many algorithms have been proposed for its successful tracking. Recently, deep learning methods have been used in many audio applications, including fundamental frequency extraction. It is interesting to know if the deep neural networks outperform classical state-of-the-art classical algorithms and if they are appropriate for music transcription. Thus, the paper aims to compare the results of certain deep learning algorithms performance with the classical ones. The main conclusion is that the CREPE algorithm is the deep learning model surpassing other handcrafted methods. It can successfully replace baseline algorithms, e.g. pYIN.

⋆ ⋆ ⋆

Evaluation of an automated method for inter-speaker synchronization measurement

Jędrzej Borowski

Wroclaw University of Science and Technology
Wroclaw, Poland

Interspeaker synchronization measurement is an important aspect of evaluating quality of multichannel wireless speaker systems. Synchronization of the speakers may be affected by issues such as decoding and processing latency, or instability of wireless network. Maintaining correct synchronization is necessary to prevent phase issues and maintain proper stereo or multichannel imaging. The author presents an automated method of interspeaker synchronization measurement of such systems, done by calculation of cross-correlation between output signals captured by microphones. Effects of using different types of measurement signals are evaluated. The method is tested with a synthetic test case, using microphones in a controlled environment with wired speakers, and with a real system consisting of a pair of synchronized wireless speakers.

⋆ ⋆ ⋆

Subjective quality assessment of the video signal coded in H.264 and H.265 standards

Stefan Brachmański, Michał Łuczyński

Wroclaw University of Science and Technology
Wroclaw, Poland

The paper presents the results of subjective tests for assessing the quality of video signals subjected to H.264/AVC (Audio Video Coding) and H.265/HEVC (High-Efficiency Video Coding) encoding. The evaluation was done at home, which is when young people can potentially watch movies on their laptops, as well as in a laboratory setting. In both experiments, the video signal was evaluated by young viewers who were not experts in the field of quality assessment. The tests were carried out with two image resolutions (1280×720 and 1920×1080) and different bit rates (from 300 kbps to 6000 kbps). The experiments performed showed that the video quality rating under laboratory conditions is only slightly higher than the home rating. It was also found that for the standard. The results showed that for the H.265 standard the best quality was obtained for the minimum bit rate of 3000 kbps, while the value of 2000 kbps already ensures acceptable image quality. In turn, in the H.264 standard, the bit rates are higher and amount to 4000 and 2500 kbps, respectively.

⋆ ⋆ ⋆

Analysis of methodologies used in sound system tuning algorithms

Łukasz Burek, Bartłomiej Kruk

Wroclaw University of Science and Technology
Wroclaw, Poland

The key aspect of research was to compare various methods of sound system tuning with use of subjective and objective approach. Three methods have been chosen: manual tuning with use of external DSP system, automatic with use of professional GLM and consumer Audyssey algorithm. In order to achieve the goal, it was required to design advance electroacoustic system with ability of simultaneous switching between different tunings. Listening tests have been carried out with custom designed method, which have been based on ITU-T BS 1116-3 and EBU Tech. 3286. The main obstacle in use of existing recommendations was presence of reference sample in it. Ultimately the best algorithm of sound system tuning is hard to be define. Everything depends of listener preferences.

⋆ ⋆ ⋆

Directivity characteristics of distributed mode loudspeakers. Selected aspects

Karol Czesak, Piotr Kleczkowski

AGH University of Science and Technology
Krakow, Poland

Distributed Mode Loudspeakers (DML) are characterized by properties, significantly differing from those related to pistonic loudspeakers. The differences occur due to design assumptions of DML, that are totally differing from design principles of pistonic loudspeakers. The oscillations of DML consist of bending waves, propagating across the loudspeaker surface, which is rectangular. That is the cause, why transducers of this type present frequency characteristics with ripples occurring at various frequencies, depending on the angle between the loud-speaker's surface and a microphone. Also, the directivity characteristics of DML are differing from those related to pistonic loudspeakers. The beam is not tightening with increasing frequency, but starts splitting into several side beams, what occurs for frequencies above 500 Hz. The strongest radiation does not occur for the axis of transducer, what determines that the Sound Pressure Level measured on the axis of the loudspeaker ought be not conceived as the reference level when normalizing measurement results and calculating directivity parameters. The attempts to analyze the directivity characteristics of DML raise a question, whether the common and well-known means of characterizing acoustics transducers are applicable to DML.

⋆ ⋆ ⋆

Flexible room acoustics simulation platform for impulse response estimation

Aleksander Kaczmarek, Piotr Cenda,
Jakub Wasilewski, Tomasz Woźniak,
Maria Peńsko, Łukasz Januszkiewicz

SoftServe Poland
Wroclaw, Poland

Measuring room impulse responses (RIR) of the complex acoustic environments is a crucial starting point of many audio engineering scenarios. Usually performing a series of necessary experiments in all required acoustic arrangements is demanding in terms of time and effort. Using synthetic or simulated RIRs is one of the data augmentation techniques used for developing spatial audio applications based on machine learning algorithms. In this paper we present a customized simulation platform that allows for flexible estimating impulse responses of various acoustic scenarios. Simulations are based on k-Wave – an open source acoustics toolbox for MATLAB and C++ to compute a time-domain acoustic wave propagation based on the k-space pseudo-spectral method. We demonstrate the capabilities of the platform by simulating selected physical phenomena such as inverse square law, sound diffraction and early reflections. We also show the results of applying our platform to a specific use case: spherical microphone array impulse response estimation. We compare our results with the analytical calculations of impulse responses based on Bessel-family functions and Legendre functions.

⋆ ⋆ ⋆

Analysis of the influence of selected coding techniques on the quality assessment of various music genres

Maurycy Kin, Stefan Brachmański

Wroclaw University of Science and Technology
Wroclaw, Poland

The main aim of this study is to examine how the type of coding and bit rate affect subjective quality assessment of different music genres. The study was conducted for AAC, HE-AAC, and MP3 encoding methods, at bit rates of 64 and 128 kbps. A listening experiment was conducted in which listeners evaluated overall sound quality and two attributes of the listening experience: timbre and spatiality. The study was performed using the Absolute Category Rating method (ACR), and 21 subjects participated in the study. The results of the experiment confirm the influence of coding techniques on the evaluation of the sound quality of a piece of music for each genre studied. In addition, it can be deduced from the study that AAC and HE-AAC techniques make it possible to obtain satisfactory sound sample quality at lower bit rates than MP3 coding. The use of higher bit rates for some music genres does not always result in better timbre, spatiality or overall sound quality.

⋆ ⋆ ⋆

Microphone versus laser displacement sensor in measuring the diaphragm movement of dynamic loudspeakers

Michał Kmiecik

AGH University of Science and Technology
Krakow, Poland

The distortions generated by the loudspeaker are largely due to the processing of electrical signal into mechanical movement. However, displacement of the loudspeaker coil does not exactly match the changes in acoustic pressure and velocity. Therefore, recording movement of the loudspeaker coil based on the generated acoustic signal is subject to significant errors, particularly in the low-frequency range. Obtaining information about constant displacement related to, e.g. reluctance force is impossible. Contactless measurement of speaker diaphragm movement is problematic and requires an optical method. A high-speed laser displacement sensor is needed to provide a wide band of the measurement, which is then limited, however, by the sensor's amplitude range and accuracy.

The paper presents a comparison of the displacement of the speaker diaphragm measured with the laser sensor and the displacement calculated on the basis of the acoustic signal recorded with a microphone. It discusses the problem of the measurement range and the error related to nonlinearity (sensitivity) of the sensor as a function of frequency.

⋆ ⋆ ⋆

Lofotr longhouse acoustic parameters analysis

Piotr Książek, Adam Pilch

AGH University of Science and Technology
Krakow, Poland

Acoustics is a multi-disciplinary science which can greatly aid other fields of study when applied. Such a field is archaeology, where material degradation can irrecoverably erase information, which must then be inferred from other sources. Room acoustics modelling may enable archaeologists to better understand historical places and in turn help support or oppose certain theories concerning both interior and exterior spaces. The following work is an exploration of acoustic modelling in the context of aiding archaeology. An acoustic model of the largest excavated Viking-age longhouse located in Lofotr was created. The model was used to calculate the distribution of the STI parameter with differing source locations and types. The potential impact of the high seat location on STI distribution of chieftain's speech was discussed as well as potential artistic performance locations which would provide balanced sound level distribution to all feasting guests with the chieftain's seat being the place of focus in all discussed hall arrangements.

⋆ ⋆ ⋆

Equalization of sound system parameters based on atmospheric conditions

Marcin Lewandowski, Tomasz Duda

Warsaw University of Technology
Warsaw, Poland

Meteorological conditions significantly impact the propagation of a sound wave outdoors. Change in temperature, humidity, and atmospheric pressure means that the sound wave can travel at different speeds with varying attenuation. This work aims to make a device/signal processor that will automatically correct the parameters of the sound system and adapt them to the actual weather conditions. The processor's task is to update sound delays and compensation for frequency response changes. Two methods were used to calculate losses resulting from changes in atmospheric conditions. The ANSI S1.26-1995 standard calculates the energy absorption of a sound wave in the air. In contrast, the method developed by Owen Cramer is used to precisely determine the speed of sound and delays for a sound system.

⋆ ⋆ ⋆

Simplified method of measuring spatial impulse responses and their application in the objective assessment of acoustic quality of rooms and surround reverberation

Witold Mickiewicz,
Kaja Kosmenda,
Grzegorz Pawełkiewicz

West Pomeranian University of Technology in Szczecin
Szczecin, Poland

The primary purpose of this article is to present issues related to the study of some acoustic properties of spaces using a portable integrated measurement system capable of measuring spatial impulse responses. For a basic introduction to the topic, parameters characterizing the sound field and acoustic properties of enclosed spaces will be described. The next section will explain the methodology for measuring the sound intensity and spatial impulse response of a room using a broadband excitation signal and a single-microphone 2D probe. The conclusion will be devoted to a description of the measurement system itself and methods for processing and interpreting the measurements obtained with it to describe the acoustic properties of confined spaces and the application of spatial impulse responses to auralization (spatial reverberation) algorithms.

⋆ ⋆ ⋆

Waveguide matrix concept for shaping the acoustic wave front

Tomasz Nowak,
Andrzej Dobrucki

Wroclaw University of Science and Technology
Wroclaw, Poland

The author proposes the use of a lens in the form of a matrix of waveguides, which divide the wave front into a finite number of fragments and introduce a controlled delay of each of them. The result is a discretizing structure that allows to adjust the curvature of the generated wave in two planes. It is achieved by an output matrix of a size and position other than the input one. Scaling the size and position of the output matrix in both directions allows the wavefront to be shaped as flat, cylindrical, ellipsoidal, and other spatial curvatures.

Multiple examples of wave fronts were calculated and presented in this paper, including parametric ones, where there is a change in shape as a function of one or both dimensions of the output matrix. A prototype lens was made for the AMT transducer, which significantly increased its angular radiation range in the vertical plane. The operation of the device has been confirmed by measurements.

⋆ ⋆ ⋆

Perceptual evaluation of first and third order ambisonic recordings

Amaia Sagasti[1],
Agnieszka Paula Pietrzak[1],
Ricardo San Martín[2],
Rubén Eguinoa[2]

[1] Warsaw University of Technology
Warsaw, Poland
[2] Public University of Navarre
Navarra, Spain

First-order or higher-order ambisonic microphones are used for ambisonic recordings. The higher the order of ambisonics, the more accurately the acoustic field can be recorded, giving more information about the location of the sound sources. However, in the case of subsequent binaural reproduction of the ambisonic sound, errors, and distortions may occur, which make the listener's ability to locate the sound source in the recordings made with a higher-order microphone not significantly higher than in the case of a first-order microphone. Higher resolution of spatial audio information can also be achieved by using upmixing techniques. In this work, auditory tests were carried out on the ability to locate the sound source in recordings made using first and third order ambisonic microphones. Perceptual evaluation of the quality of recordings upmixed from first to third order ambisonic recordings was also carried out.

⋆ ⋆ ⋆

The use of e-learning methods in musical ear training

Sara Saja, Bartłomiej Kruk

Wroclaw University of Science and Technology
Wroclaw, Poland

Paper applies to influence of ear training in e-learning system due to their growth during the pandemic. Furthermore, dependencies and clues were made so they can be used in future training. Research aimed at ability in timbre solfege recognition and music memory among course participants. Exercises were made empirically, based on previous tests results which were made on an ongoing basis. Different samples of signals as well as music were used to ensure full differentiation. The analysis was performed both in specific groups and for individual cases. Various test layouts were tested for the best outcome and future guidance. Ultimately the analysis for typical in audio industry nomenclature related to timbre and was made.

⋆ ⋆ ⋆

Design and development of management systems in unconventional surround sound installations

Jan Skorupa, Maciej Głowiak

Institute of Bioorganic Chemistry of the Polish Academy of Sciences
Poznan, Poland
Poznan Supercomputing and Networking Center
Poznan, Poland

Surround sound systems are a practice that has been explored for years by engineers and artists alike. The activities of such centers as Radiodiffusion-Télévision Française, GRM and composers Ilanis Xenakis and Karlheinz Stockhausen have exemplified the possibilities inherent in combining science and art. The article, on the example of two author's realizations, will present contemporary practices of creating unconventional surround sound systems. Contemporary technological possibilities in the design of multi-speaker installations will be considered, as well as the programming of computer systems that control audio signals.

⋆ ⋆ ⋆

Restoration of the 1950s speech recording studio in the Polish Radio in Warsaw

Radosław Smoliński

Polish Radio S.A.
Warsaw, Poland

The Polish Radio in Warsaw has restored the 1950s speech recording studio. In the studio historic elements of the interior are restored and supplemented with modern recording equipment and new technological furniture. The studio has received corrections in the field of interior acoustics, taking into account modern technical requirements. Historical acoustic elements are composed with modern sound-absorbing and sound-diffusing solutions. The studio has an original, unique, historic design, at the same time is adopted to modern technological, acoustic, and image realities.

⋆ ⋆ ⋆

Employing the MUSHRA test in the study of the benefits of using hearing prostheses

Piotr Szymański[1], Tomasz Poremski[1], Bożena Kostek[2]

[1]Sonova Audiological Care Sp. z o.o.
Lodz, Poland
[2]Gdansk University of Technology
Gdansk, Poland

Evaluating the quality of hearing aid fit in terms of the benefits a prosthesis can provide is a complex issue. It is easy to determine objective parameters of hearing aids such as gain, harmonic distortion, frequency response, etc. However, these parameters do not always have a direct and decisive influence on the patient's subjective evaluation of the quality of the hearing prosthesis fit. Today's hearing aids have a number of advanced features that make speech understanding, in particular, easier and more manageable in various difficult listening situations. Still, it is not entirely possible to compare or measure them. Measurements of hearing aid effectiveness can address many aspects, including hearing loss compensation, acceptance, gain, or satisfaction with the prosthesis.

The authors present a modification of the commonly used hearing aid benefit assessment questionnaire Abbreviated Profile of Hearing Aid Benefit (APHAB) by combining it with the MUltiple Stimuli with Hidden Reference and Anchor (MUSHRA) test. The MUSHRA test is used to assess the sound quality of hearing aids.

The APHAB is a closed-ended questionnaire completed independently by the hearing aid user. Specific sound situations are evaluated. They are rated on a 7-point scale. Each degree of the scale from A to G includes a description and an associated percentage value (minimum value 1%, maximum 99%). The combination of letter, percentage, and descriptive scales can be challenging to interpret for the subjects, most of whom are elderly. Therefore, the concept of modifying the questionnaire was developed, i.e. mapping the 7-point APHAB scale to a 100-point scale consistent with the MUSHRA test scale using fuzzy logic.

⋆ ⋆ ⋆

Using the MATLAB for the synthesis of spatial sound based on personalized measurements of HRTF

Zbigniew Świętach, Przemysław Plaskota

Wroclaw University of Science and Technology
Wroclaw, Poland

This paper deals with synthesis of spatial sound based on personalized measurements of Head Related Transfer Functions (HRTF). The main work has been done using MATLAB computing environment that is well adapted to solve manifold engineering problems.

Authors used the personalized base of HRTF measurements. Synthesis of the virtually moving sound source as the still sound source has been performed, however an attempt that was made to adaptively adjust the length of adjacent time frames between successive positions of the sound source did not achieve satisfactory results yet. It will be extensively studied in the future.

⋆ ⋆ ⋆

Adaptive near-field beamforming system based on machine learning

Agnieszka Wielgus

Wroclaw University of Science and Technology
Wroclaw, Poland

This work deals with the problem of near-field broadband beamforming. Since the issue of the number and position of individual microphones is of great importance in the design process of broadband beamforming, we try to propose an adaptive beamforming system in which the number and position of microphones depend on the current speaker's position (we assume that the speaker can move in a restricted, previously defined area). For this purpose, we consider a big rectangular microphone matrix, in which any number of microphones can be active depending on the speaker's position. The applied method is based on reinforcement learning (RL) – a methodology of machine learning in which agent learns how to maximize returns or achieve the given goal by system of rewards and punishments. RL has close connections to both adaptive control and optimization. We try to learn our system to adapt to changing speaker's positions by changing the set of active microphones and its filters coefficients so that the output of the system is as close as possible to optimal output in the sense of the l2 norm.

⋆ ⋆ ⋆

Automatic classification of pathological speech

Martyna Włoszczyńska, Bożena Kostek

Gdansk University of Technology
Gdansk, Poland

The purpose of this paper is to present an application that is used to automatically detect pathological speech based on an annotated database containing recorded words and sentences. The background of the study belonging to human-computer communication, a critical review of the literature, the algorithms used, and the features of the speech signal that discern between undisturbed and pathological speech are presented first. The assumptions underlying the experiments conducted are also given, along with the selection of the pathological speech dataset. In the next step, the neural network architecture and its parameters for speech type classification are proposed, and the preprocessing of the speech signal is presented. The obtained results of the classification of undisturbed and pathological speech are compared with literature sources. The trained neural networks are then used in an application to perform the classification of the speech signal. The summary of the experiments also includes conclusions and suggestions for the development of this research study.

⋆ ⋆ ⋆

Sound spatialization algorithm based on phase shifting

Kamil Zimny, Teresa Makuch

AGH University of Science and Technology
Krakow, Poland

Most of the currently used signal spatialization algorithms are based on adding reverberation, signal repetitions, delays. Usually they offer a number of parameters corresponding to the parameters of the modelled process. However, these parameters poorly relate to the subjective impression of the signal spatiality. Scientific research indicates that much more complex mechanisms are responsible for perception of the spatiality of signals. According to them, the auditory system is sensitive to phase dependencies between the harmonics of the signal, which translates to a greater extent into the sense of spatiality.

For this reason, the subject of this paper is description of the spatialization algorithm based on phase modifications. They are performed on the phase spectrum of the windowed signal obtained using the Fourier transform. Then the phase modifications are applied to certain harmonics in a given frequency band. The proposed algorithm describes several methods of phase modification: random with the use of several distributions, and with the use of mathematical functions, including periodic functions. Then the modified signal is transferred back to the time domain by inverse Fourier transform and overlap-add reconstruction.

⋆ ⋆ ⋆