GTAC ANNUAL RESEARCH REPORT 2018-2019

HEAD OF THE GROUP RESEARCH REPORT

The Audio and Communications Signal Processing Group (known by their acronym GTAC from its Spanish name Grupo de Tratamiento de señal en Audio y Comunicaciones) has developed its research during the scholar year 2018-19 mainly on active noise control, spatial audio perception and rendering, and sound quality improvement for multi-channel audio systems. GTAC has carried out several research projects and has published their most relevant results in several scientific journals and conference proceedings. In particular, the European project "Distributed Network of Active Noise Equalizers for Multi-User Sound Control" has finished on October 2018 with great success, achieving their objectives by creating a laboratory prototype of a distributed active noise equalizer system for two independent car seats. Moreover, the whole system has been filed as a patent in the United States. On the other hand, the national project "Smart Sound Processing for the Digital Living" has also finished on December 2019 achieving several applications as: (1) Design of personal sound zones with spatial smoothness (2) Measurement and monitoring of noise conditions in industrial environments to detect high sound pressure levels (SPL), and (3) recording, monitoring and classification of sound events.

Regarding the GTAC facilities, a new laboratory for perceptual spatial sound (see Fig. 1) has been built within the iTEAM place. The purpose of the new audio lab is to measure the Head-Related Transfer Function (HRTF) of many people with a very high precision, in order to improve the way spatial sound is rendered to a particular person. The HRTF is in somehow a personal acoustic fingerprint that changes from one person to another. This new set-up is formed by a 4-meter-diameter circular array of 72 loudspeakers placed in the same horizontal plane, plus two sets of 8 loudspeakers, one placed in the ceiling and one on the floor.

As said before, GTAC current research includes the design optimization of personal audio zones. Personal audio systems aim to create listening (or bright) and quiet (or dark) zones in a room using an array of loudspeakers (see Fig. 2). It allows rendering a target soundfield in the bright zone while having control over the mean acoustic energy in the quiet zone. Moreover, it can create different listening sounds in different audio zones in the same room. For this purpose, a new array of 24 loudspeakers in blocks of 8 speakers each has been built to control mid-range frequencies. The new array is formed by small tweeters that allow for a separation of 4.5 cm between the center points. Fig. 3 shows in the first plane the array of microphones that has been used to control and validate the acoustic zones, and in second place two blocks of the new array (16 tweeters inserted in two wooden boxes) can be seen. This research line has been granted with the iTEAM Science Pills Video Award 2018 (link to the video: https://youtu.be/tSfrVNdAvEI).

Fig. 1. Perceptual spatial acoustic laboratory including a render system of 72 + 16 loudspeakers.

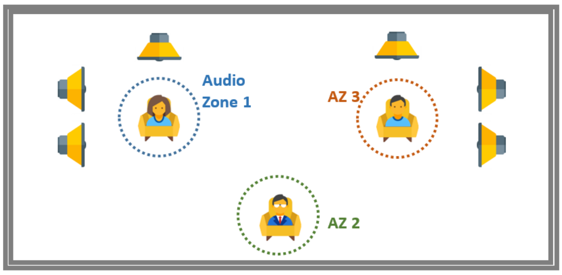

Fig. 2. Personal audio zones using loudspeakers.

Fig. 3. Rendering and measurement of personal sound zones.

Moreover, three new projects have started during 2019. They are summarized in the following.

1.- ONGOING PROJECTS

1.1 *DYNAMIC ACOUSTIC NETWORKS FOR CHANGING ENVIRONMENTS (DANCE)*

Webpage: *www.dance.upv.es*

Summary

The ``Dynamic Acoustic Networks for Changing Environments'' (DANCE) is a coordinated project that intends to demonstrate the usefulness and applicability of Dynamic Acoustic Sensor Networks by the development of distributed algorithms and systems to different audio applications, such as, self-localization of nodes, dynamic room impulse responses (RIR) and inverse RIR estimation, dynamic characterization and control of acoustic wave fields, adaptation of communication and processing to changes of network topology and/or objective. Furthermore, go forward to real-time implementation by using emerging computing tools.

DANCE proposal will carry out research in the field of acoustic sensor networks, when dynamism, in the sensor network or in the environment, is considered. Dynamism is here referred to movement of the nodes in the sensor network, changes in the network configuration, movement of the sound sources or obstacles in the considered environment, or time-variant characteristics of any of them.

The project considers different complementary applications: control and generation (user selected) of time-variant sound scenes (in the audible band), monitoring of acoustic scenes for detecting and tracking acoustic threats in security applications (such as unmanned aerial vehicles-UAV in critical infrastructure protection or impulsive sound sources like electric sparks that could be the origin of fires), indoor human behaviour characterization using acoustic sensors, etc.

These applications will be based on the usage and development of innovative signal processing methods, such as space-time adaptive signal processing techniques (STAP), time-frequency signal models, algorithms for self-determination of sensor network topology, dynamic array signal processing, tracking algorithms based on Kalman and particles filters, detection algorithms based on machine learning tools, emerging pattern recognition tools such as Deep Learning, etc.

Funded by: Spanish Ministry of Science, Innovation and University.

1.2. *INTELLIGENT SPATIAL AUDIO: SYNTHESIS AND CUSTOMIZATION (ISLA-THESON)*

Summary

The sound industry has been experiencing profound changes in recent years under the perspective of 3 complementary approaches: the individual, the group and the contents.

On the one hand, the irruption of mobile devices has modified the way people listen to music, play and interact with machines and other people. The use of headphones has spread enormously, and the need to reproduce highly realistic spatial sound through them is a great opportunity for the industry. For a very immersive experience, the sound must be customized for each individual based on their anatomy, in particular the head and pinna shape, which define their particular Head-Related Transfer Function (HRTF). Measuring a subject's HRTF is still a costly process that requires specialized facilities and its indirect estimation remains an unresolved problem. By employing Deep Learning techniques and photographs of the ear/head, we aim to achieve a personalization of HRTF of better quality than other proposed methods. This will in turn allow mobile devices to incorporate personalized responses for their direct application in 3D sound, virtual and augmented reality, video games, etc.

On the other hand, the sound and entertainment industry has been redirected during the recent years to big live shows, where the spatialization of sound is still a challenge and an opportunity for using sound field synthesis algorithms and others for recreating virtual spaces. Array processing techniques should be developed aimed at allowing the control of sound in different listening areas while synthesizing the different live sound objects (musicians, actors, presenters, effects, etc.), adapting the synthesis of each object to its own movement (which must be captured and monitored), and achieving greater realism over the audience. In addition, there are other scenarios, such as museums, exhibitions, restaurants or smart homes, which would also benefit from the creation of independent audio zones or with different needs, using similar techniques employing loudspeaker and sensor arrays.

Finally, from the contents point of view, the viewer demands more and more a complete and interactive audiovisual experience. Within this line, this subproject will work on creating new methods for the analysis of audio and music based on Machine Learning, with application to synchronized audiovisual effects and live enriched events. The aim is to develop Machine Learning algorithms that extract features from music that allow the synchronization of audiovisual material such as 3D animations, lights or lasers. In addition to the improvement in rhythm extraction algorithms with low latency, it is necessary to work on the segmentation of songs into their different parts to link different animations. Thus, Machine Learning techniques trained from labeled songs are profiled as a solution for real-time song segmentation in a causal way, allowing for their application in live scenarios.

Funded by: Spanish Ministry of Science, Innovation and University.

1.3. SMART SOCIAL COMPUTING AND COMMUNICATION (IN SPANISH: COMUNICACIÓN Y COMPUTACIÓN INTELIGENTES Y SOCIALES - COMTACTS)

Summary

The advances made in the field of distributed computing and the hardware-software available right now make possible to develop increasingly powerful information processing and exchange systems, which can interact with the environment through numerous sets of transducers. These transducers, in turn, provide an ever-increasing volume of signals and data, making possible a more precise knowledge of the social environment and the physical environment in which living beings, particularly humans, work and live.

At the same time, we can observe the boom in applications arising from computing and communication devices for personal use, and their massive use with the advance of communications; we can highlight some applications such as: human-machine interaction, control systems, location and tracking systems, telepresence, automatic classification, high-speed communications, diagnostic assistance systems, etc. Within this framework, intelligent and social computing and communication is defined as the hybrid mix of the two disciplines in order to face challenges of high socio-economic interest. Science is used for the purpose of communications and computing, but taking into account ubiquity, versatility, scalability, efficiency. and cooperative processing of heterogeneous computing and data acquisition device networks.

Moreover, an insightful consideration is given in this project regarding the physical aspects of computing, signal processing, energy consumption, technology, communication, etc., particularly in distributed, collaborative scenarios and provided with massive and heterogeneous data. In this way, the research group of the present proposal addresses the design, development and implementation of products, systems, programs and algorithms for signal processing and communications, which make use of state-of-the-art architectures, advanced computing and efficient communications within the framework of intelligent computing and communication aimed at tackling social challenges.

Funded by: Regional Government – Generalitat Valenciana.

2.- RESEARCH RESULTS

The most important results of the GTAC publications over the past year are summarized in the following. For a more detailed description, visit our webpage: www.gtac.upv.es where a complete list of projects and papers can be found.

2.1.- FEATURED JOURNAL PUBLICATIONS

- ***Combined precoding for multiuser Multiple-Input Multiple-Output satellite communications.*** *M.A. Simarro-Haro, Beatriz Puig, Francisco José Martínez Zaldívar, Alberto Gonzalez, Computers & Electrical Engineering, vol. 71, pp. 704-713, 2018.*

DOI: 10.1016/j.compeleceng.2018.08.006.

Abstract: *Applying Multiple-Input Multiple-Output (MIMO) techniques in satellite communications can increase data rates. However, new signal processing elements have to be taken into account to fully exploit the expected advantages of MIMO communications. In this paper, we evaluate different precoding techniques over the satellite channel. A performance comparison between several precoders in terms of Bit Error Rate (BER) and complexity is given for different channel realizations. Furthermore, a novel hybrid scheme for signal precoding is proposed that optimizes the computation for a required BER. The new scheme is based on the matrix condition number of the satellite MIMO channel.*

- ***Perception of nonlinear distortion on emulation of frequency responses of headphones.*** *Pablo Gutierrez-Parera, José Javier López Monfort, The Journal of the Acoustical Society of America, vol. 143, nª 4, pp. 2085- 2088, 2018.*

DOI: 10.1121/1.5031030.

Abstract: *The equalization of headphones can force transducers to work in a non-linear condition, producing non-linear distortion. Depending on the headphone model and the reproduction level, that distortion can be audible. In this study, headphones of diverse quality and price were compelled to emulate the same target frequency response and the non-linear distortion was measured. A Diagonal Volterra model was used to simulate the different headphones with and without distortion. A perceptual test was carried out to determine the level of reproduction above which non-linear distortion is perceived for each headphone model. High correlation has been found between the level of detected distortion and retail prices of headphones.*

- ***Parallel SUMIS Soft Detector for Large MIMO Systems on Multicore and GPU.*** *Carla Ramiro Sánchez, M.A. Simarro-Haro, Alberto Gonzalez, Antonio Manuel Vidal Maciá, The Journal of Supercomputing, vol. 75, nª 3, pp. 1256-1267, 2019.*

DOI: 10.1007/s11227-018-2403-9.

Abstract: *The number of transmit and receiver antennas is an important factor that affects the performance and complexity of a MIMO system. A MIMO system with very large number of antennas is a promising candidate* *technology for next generations of wireless systems. However, the vast majority of the methods proposed for conventional MIMO system are not suitable for large dimensions. In this context, the use of high-performance computing systems, such us multicore CPUs and graphics processing units has become attractive for efficient implementation of parallel signal processing algorithms with high computational requirements. In the present work, two practical parallel approaches of the Subspace Marginalization with Interference Suppression detector for large MIMO systems have been proposed. Both approaches have been evaluated and compared in terms of performance and complexity with other detectors for different system parameters.*

- ***Fast block QR update in digital signal processing.*** *Fran J. Alventosa, Pedro Alonso-Jordá, Antonio Manuel Vidal Maciá, Gema Piñero, Enrique S. Quintana-Ortí, The Journal of Supercomputing, vol. 75, nª 3, pp. 1051-1064, 2019.*

DOI: 10.1007/s11227-018-2298-5.

Abstract: *The processing of digital sound signals often requires the computation of the QR factorization of a rectangular system matrix. However, sometimes, only a given (and probably small) part of the system matrix varies from the current sample to the next one. We exploit this fact to reuse some computations carried out to process the former sample in order to save execution time in the processing of the current sample. These savings can be critical for real-time applications running on low power consumption devices with high mobility. In addition, we propose a simple out-of-order task-parallel algorithm for the QR factorization using OpenMP that exploits the multicore capability of modern processors. Furthermore, in the presence of a Graphics Processing Unit (GPU) in the system, our algorithm is able to off-load some tasks to the GPU to accelerate the computation on these hardware devices.*

2.2.- FEATURED CONFERENCE PROCEEDINGS

- ***On Perceptual Audio Equalization for Multiple Users in Presence of Ambient Noise.*** *Juan Estreder-Campos, Gema Piñero, Fabián Aguirre-Martín, María de Diego Antón, Alberto Gonzalez, 10th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), Sheffield, UK, 2018.*
- ***Low cost algorithm for online segmentation of Electronic Dance Music.*** *Emanuel Aguilera Martí, José Javier López Monfort, Pablo Gutiérrez-Parera, Carlos Alberto Hernandez Franco, 144th International Audio Engineering Society Convention (AES 2018), Milan, Italy, 2018.*
- ***Binaural room impulse responses interpolation for multimedia real-time applications.*** *Victor García-Gómez, José Javier López Monfort, 144th International Audio Engineering Society Convention (AES 2018), Milan, Italy, 2018.*
- ***Array processing for echo cancellation in the measurement of Head-Related Transfer Functions.*** *José Javier López Monfort, Sergio Martínez-Sánchez, Pablo Gutiérrez-Parera, 11th European Congress and Exposition on Noise Control Engineering (EuroNoise 2018), Crete Island, Greece, 2017.*
- ***Optimization of line array wave guides.*** *Javier Redondo, Juan Vicente Sánchez Pérez, José Javier López Monfort, 25th International Congress on Sound and Vibration, Hiroshima, Japan, 2018.*
- ***On the performance of noise barriers based on sonic crystals.*** *Javier Redondo, Juan Vicente Sánchez Pérez, José Javier López Monfort, 25th International Congress on Sound and Vibration, Hiroshima, Japan, 2018.*
- ***Perception of noise annoyance reduction associated with acoustic screens.*** *Javier Redondo, M. Pilar Peiró-Torres, José Javier López Monfort, A..Pereira, P. Amado-Mendes, Luis Godinho, 48th International Congress and Exposition on Noise Control Engineering (Inter-noise 2019), Madrid, Spain, 2019.*