BINAURAL RECORDING SYSTEM AND SOUND MAP OF MALAGA

PACS:

Rosas Pérez, Carmen; Luna Ramírez, Salvador
Universidad de Málaga
Campus de Teatinos, 29071
Málaga, España
Tel:+34 952 137 186
Fax: +34 952 132 027
E-Mail: crosas@ic.uma.es; sluna@ic.uma.es

Palabras clave: Binaural recording, binaural microphones, spatial sound, HRTF, dummy head, KEMAR, MIT, sound map, soundscapes, sonic heritage, Malaga,

ABTRACT

The aim of this project is the construction and the characterisation of a pair of in-ear binaural microphones with high-quality capsules. A set of HRTF measurements was obtained and applied to different audio signals for the realisation of a psychoacoustic experiment to assess the spatiality provided by the system. For the system assessment, another set of audio samples was generated from the MIT's HRTFs, and both results have been compared. Additionally, different soundscapes have been recorded with the binaural system, and a binaural sound map of Malaga has been developed, which aims to create an archive to collect and conserve the most distinctive sounds of the city using an immersive technology.

INTRODUCTION

Immersive sound can be generated using multiple recordings and mixing techniques. Binaural microphones have several disadvantages to making a binaural mix from a set of recorded sounds, e.g. the impossibility of head tracking or moving sound sources using processing software [1]. However, binaural microphones implement an immediate and practical method that allows to record in almost any place and circumstances, and produce acceptable immersive recordings with a low budget.

MICROPHONES CHARACTERISATION

The binaural microphones created for this project have been built using two Primo EM172-Z1 capsules. These condenser capsules present low noise, highly flat response and wide dynamic range [2]. The capsules have been introduced into a pair of silicone in-ear monitors and soldered to a stereo audio cable with a 3.5mm jack connector.

Figure 1. Binaural microphones used on the project.

HRTF Measurements

The measurement of the Head-Related Transfer Function, using the binaural microphones, was performed in the anechoic chamber of the Superior Technical School of Telecommunication Engineering of Malaga. The impulse responses (HRIR) were obtained with the software *Smaart* using pink noise as stimulus and a neutral response loudspeaker. The subject wearing the microphones was situated in a series of different positions in elevation and azimuth with respect to the loudspeaker.

EXPERIMENT DESCRIPTION

Stimuli

Each pair of HRIR (right and left) was applied to two different mono audio files (a low voice and a maraca recording) using *Matlab* to obtain the binaural audio signals, which were reproduced using Sennheiser HD 449 headphones.

To compare the results obtained for the binaural microphones, another series of binaural audio samples was generated using the HRTFs from the MIT [3] and the same source mono files.

Group of Subjects

A total of 13 subjects, with ages between 24 and 47 years old and without any known hearing problem, participated in the localisation test.

Procedure

Each volunteer was asked to indicate on circumference diagrams the approximate location of the reproduced sound. The circumferences to indicate the azimuth were divided in 8 equal sectors: front, front-right, right, rear-right, rear, rear-left, left and front-left [4]. The semi circumferences for the elevation were formed by 4 sectors and 2 extreme positions: total elevation, strong elevation, soft-elevation, soft negative elevation, strong negative elevation, total negative elevation. A total of twelve sounds were virtually situated on the space around the subjects for both sets of source files (microphones and MIT). The first six sounds had variations only in azimuth, and other six sounds had also variations in elevation. The number of sounds to be located was chosen with the aim of not creating an excessive fatigue on the subjects.

RESULTS

The perception of sound sources presents the common problematic positions as other previous psychoacoustic experiments [5], being more difficult to detect the difference between sounds located within the cones of confusion (ambiguity for sources between 45º and 135º or 255º and 315º) and at 0º or 180º. In these cases, the percentages of correct answers are smaller than those from lateral positions, for both types of sounds (i.e. voice and maraca) and HRIR origin (binaural microphones and MIT's responses). The average of correct answers in the horizontal plane is very similar in both cases, as shown in Fig. 2 and Fig. 3.

The high-frequency content in the maraca sound leads to a more detectable variation of the spectral profile when the sound is located at rear positions, and a higher percentage of correct answers compared to results with percentages with voice sound (69% against 46%, respectively).

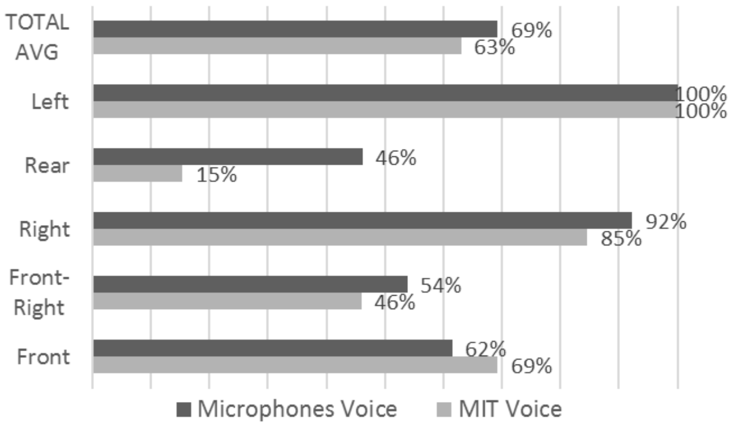

Figure 2. Percentage of correct answers in the horizontal plane (voice sound).

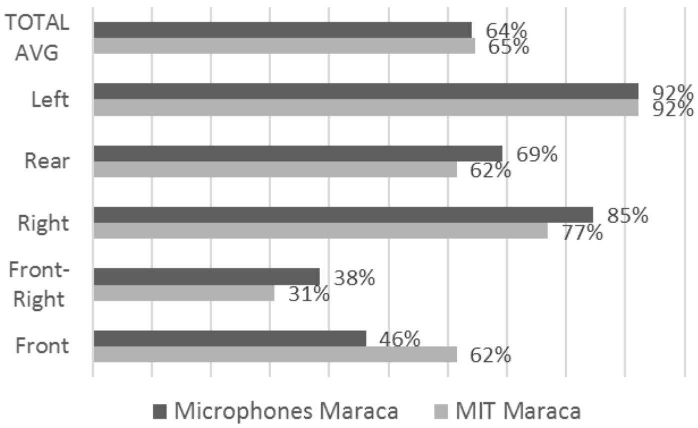

Figure 3. Percentage of correct answers in the horizontal plane (maraca sound).

Fig. 4 presents the percentage of correct answers for the median plane. Results are worse than those obtained in the horizontal plane. This was expected due to different reasons, such as a) the perception of the elevation is highly related to the shape and folds of the pinna [6], b) the differences between listener's and KEMAR's ears, and c) the impossibility of placing the microphones at the same position of the eardrum and other anthropometric factors. Additionally, and in this case, correct answers using the MIT signals are higher than answers from the binaural microphones.

Fig. 5 shows results when some elevation is introduced when locating the voice sound. In this case, localisation with microphones' responses performs better than MIT responses. Compared to Fig. 2 (i.e. horizontal plane without elevation) correct answers decrease.

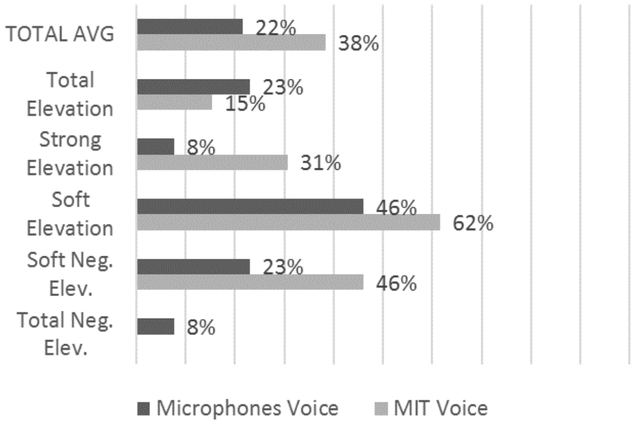

Figure 4. Percentage of correct answers in the median plane (voice sound).

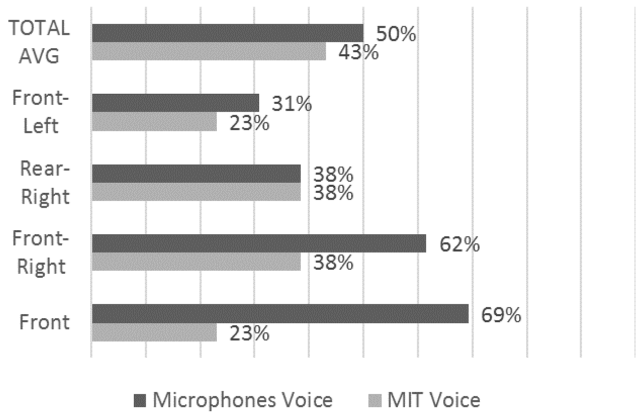

Figure 5. Percentage of correct answers in the horizontal plane adding elevation (voice sound).

Finally, Fig. 6 and Fig. 7 detail angles answered by the listeners with voice and maraca sounds, respectively. Summarising, listener estimations have resulted as expected considering drawbacks as the non-personalisation of the HRIRs and the equalisation of the headphones.

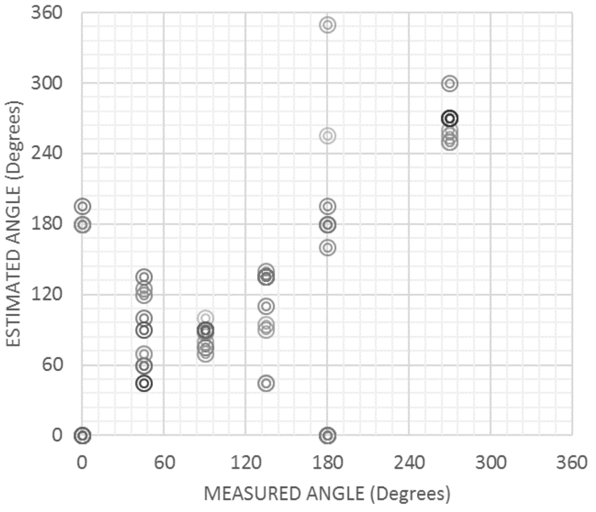

Figure 6. Azimuth answers for the voice synthesised with the HRIR of the microphones.

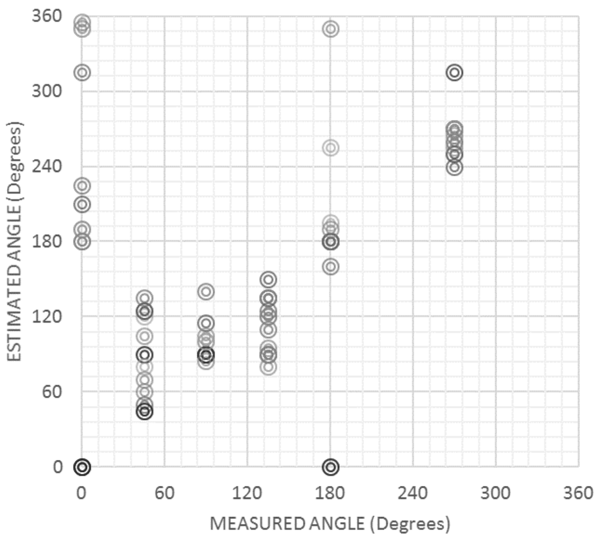

Figure 7. Azimuth answers for the maraca synthesised with the HRIR of the microphones.

SOUND MAP

Constructed binaural microphones were used to record a set of soundscapes in and around Malaga. Audio recordings were implemented using a Sony PCM-M10 recorder at 48 kHz, 16 bit resolution. The recordings, a total of 45 at this moment, have been included in an online sound map developed for this project: http://mapasonoromalaga.com. The audio files do not have any other processing than slight fades and a conversion to Free Lossless Audio Codec (FLAC) files to be uploaded to the hosting platform. The online player embedded in the sound map offers only MP3 quality, but the sounds can be downloaded in high quality. The map has been generated using *uMap*, a free tool based on *OpenStreetMap*.

The recording locations were determined by their sound richness and diversity offered by the city of Malaga, although the project is continuously growing. The website aims to provide a practical tool that collects all the most distinctive soundscapes of Malaga so that they can be listened to by people from all over the world as an immersive experience, becoming part of the city's cultural heritage. The sound map also serves as an archive and catalogue of the conserved soundscapes. In contrast with noise maps, a sound map allows the characterisation of the territory from a different perspective, in which the identity of the depicted area at a certain time is defined by all its sounds.

CONCLUSIONS

This e-brief describes the construction and the characterisation of a pair of binaural microphones with high-quality capsules. A set of spatialised sounds were constructed from HRIRs obtained from the binaural system and, with comparison purposes, from a publicly available set of HRIRs. A set of listeners were asked to locate the spatialised sounds from both HRIRs and for different sounds and locations.

Results suggest that the spatiality achieved with the binaural microphones is not exceptional, but similar to that obtained from sounds generated from HRIRs obtained with a professional KEMAR dummy head. In-ear binaural microphones present some advantages over dummy head equipment, like portability and discretion, so the interference with the natural ambience of the recorded soundscapes is minimal.

REFERENCES

[1] M. C. Rosas Pérez, “Sistema de grabación binaural y mapa sonoro”, Master Thesis, Dir. S. Luna-Ramírez, Universidad de Málaga, Spain (2016).

[2] *EM172 Electret Omni-Direction Product Details*, Primo Microphones, Inc. http://www.primomic.com/php/get_products_details.php?model=EM172

[3] B. Gardner, K. Martin, “HRTF measurements of a KEMAR dummy-head microphone”, MIT Media Lab, Tec. Report 280 (1994) http://sound.media.mit.edu/resources/KEMAR.html

[4] B. Payri, R. Rodríguez-Mariño, “Perceptual comparison of localization with Soundman binaural microphones vs HRTF post-processing”, Audio Engineering Society 140th Convention, Paris (2016).

[5] C. J. Plack, *The Sense of Hearing*, Routledge (2005).

[6] J. Blauert, *Spatial Hearing. The psychophysics of human sound localization,* The MIT Press (1997).