Localization with artificial head recordings

Sandvad, Jesper; Christensen, Flemming; Olesen, Søren Krarup; Møller, Henrik

Published in:
Journal of the Acoustical Society of America

Publication date:
1997

Session 2pSP

Signal Processing in Acoustics, Psychological and Physiological Acoustics and Speech Communication: Acoustics in Multimedia II: Head Related Transfer Functions

Isaac Graf, Chair
Speech and Hearing Sciences Program, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Invited Papers

1:30

2pSP1. Spatial perception, the acoustics of the external ear, and interactions with earphones. Edgar A. G. Shaw (Inst. for Microstruct. Sci., Natl. Res. Council, Ottawa, ON K1A 0R6, Canada)

In free-field listening, the perception of source elevation and discrimination between forward and rearward source positions appear to be largely dependent on the availability of high-frequency spectral cues associated with the mode structure of the concha and diffraction by the pinna, respectively. The role of the concha is particularly clear between 5 and 10 kHz, while pinna-diffraction effects are strongly developed between 2.5 and 6 kHz. When sounds are presented through earphones, the free-field characteristics of the ear are replaced by very different characteristics that are dependent on complex and unreliable interactions between the individual earphone and the individual ear. Since these interactions are highly dependent on earphone design, it may be appropriate to focus attention on the special characteristics that may be desirable in earphones intended for use in the creation of virtual acoustical environments. These may include coupling that (i) allows the response at the eardrum to follow the primary resonance curve of the individual ear and (ii) minimizes the inter- and intrasubject variations in response to other frequencies, including frequencies above 5 kHz where the wave characteristics of the concha come into play.

2:00

2pSP2. Application of head-related transfer functions for binaural synthesis. Dorte Hammershoei, Jesper Sandvad, and Henrik Moeller (Acoust. Lab., Aalborg Univ., Fredrik Bajers Vej 7 B4, DK-9220 Aalborg, Denmark, acoustics@kom.auc.dk)

Binaural synthesis offers a method for creating a truly immersive three-dimensional auditory environment. The sound output is produced by convolving a sound input with sets of head-related transfer functions (HRTFs). The result is a set of two signals typically reproduced by means of headphones, and the technique is therefore well suited for systems with limited hardware for sound reproduction, such as most multimedia systems. In the present paper, HRTFs are presented in the time and frequency domains, and are commented upon in relation to their application for binaural synthesis using digital signal processing. Results obtained from experiments conducted at the Acoustics Laboratory at Aalborg University in Denmark are reviewed, giving hints for the selection of filter lengths, update rate, and spatial resolution for dynamic environments.

Contributed Papers

2:30

2pSP3. Localization with individual and nonindividual binaural recordings. Henrik Moeller, Dorte Hammershoei, Clemen B. Jensen, and Michael F. Soerensen (Acoust. Lab., Aalborg Univ., Fredrik Bajers Vej 7 B4, DK-9220 Aalborg, Denmark, acoustics@kom.auc.dk)

The localization performance was studied when subjects listened (1) to a real sound field and (2) to binaural recordings of the same sound field, made (a) in their own ears, and (b) in the ears of other subjects. The sounds to be localized were loudspeaker reproductions of female speech at natural level, from 19 different positions in a standard listening room. The binaural recordings were reproduced by carefully equalized headphones. With individual recordings the performance was preserved compared to real life, whereas nonindividual recordings resulted in significantly more errors for sound sources in the median plane. Errors were seen in terms of confusion not only between nearby directions, but also between directions further away, such as between sound sources in front and behind the subject. However, nonindividual recordings made in the ears of a carefully selected "typical" subject resulted in a performance much closer to the real-life performance, although still inferior.

2:50

2pSP4. Localization with artificial head recordings. Jesper Sandvad, Flemming Christensen, Soeren K. Olesen, and Henrik Moeller (Acoust. Lab., Aalborg Univ., Fredrik Bajers Vej 7 B4, DK-9220 Aalborg, Denmark, acoustics@kom.auc.dk)

Twelve artificial heads, of which one was designed at the authors' laboratory, were evaluated in a localization test. The localization performance was compared in two situations: First, the subjects localized sound sources in a real sound field, then the localization test was repeated with artificial head recordings of the same sound field. The sounds to be localized were loudspeaker reproductions of female speech at natural level, from 19 different positions in a standard listening room. The artificial head recordings were reproduced by carefully equalized headphones. Results from eight of the heads showed an increased number of localization errors compared to real-life performance. The directions in the median plane were most often confused, not only with nearby directions, but also with directions further away. The number of errors was significantly higher than what can be obtained with recordings from a carefully selected human head. Results from the last four heads, including the one from the authors' laboratory, were not available at the time of abstract submission.