Thesis subject 2017

Laboratory: Institut ean Le Rond d’Alembert
University: UPMC
Title of the thesis: The Spatial Voice
Thesis supervisor: Christophe d’Alessandro
Email contact: cda@dalembert.upmc.fr, brian.katz@upmc.fr
Co supervisor: Brian FG Katz
Collaborations within the thesis: Markus Noisternig, IRCAM/STMS
Program affiliation:
Cotutelle: no
University: no
This subject can be published on the doctoral school’s web site: Yes

Thesis’s summary (abstract):
Human interactions through voice are always situated within a given space or environment and at a given distance between speaker/signer and listener. The aim of this thesis is to investigate acoustic and perceptual issues related to space in vocal communication (radiation directivity patterns, distance, and acoustic response of the surrounding environment). In a first part, phoneme dependent radiation patterns are measured and analyzed for a number of subjects, under conversational speech, theater speech, and singing conditions. Correlations between mouth geometry and directivity patterns are analyzed and acoustic models are investigated. Data reduction of directivity patterns is performed for spatial audio rendering (e.g. via clustering and spherical harmonic decomposition). The effect of distance in voice communication is also investigated, both with regards to perception and production. Perceived speaker or singer distance and changes in voice production according to the supposed distance to the listener are studied. Finally, interaction between room acoustics (e.g. in opera house, theater, concert hall, or virtual conference room) and spatial perception and production for singing and vocal performances are investigated. The results of this thesis will be a better understanding of ecological voice interactions, including adaptation to communication distance and relative head positions. It will improve rendering in virtual audio applications: position of the speaker or singer avatars in the scene, orientation of the avatar’s head, and interaction with the acoustics of the virtual room as well as rendering environment.
Subject

Context

Human interactions through voice are always situated in a given space or environment and at a given distance between speaker/singer and listener. Spatial audio is a growing research field with impressive results and applications. However, few studies have been devoted to the spatial aspects of the voice, although it is by far the most important sound source in communication, entertainment, art, etc. This project will merge two important fields: spatial audio and voice studies.

Aims and Objectives

The aim of this thesis is to investigate acoustic and perceptual issues related to space in vocal communications. Three main aspects will be investigated: phoneme dependent radiation directivity patterns, effect of distance on voice perception and production, and voice transformation and adaptation as a function of the acoustic response of the surrounding environment.

Methodology

The project is structured in three main parts: voice directivity, effect of distance, and effect of the environment. In a first part, phoneme dependent radiation patterns are measured and analyzed for a number of subjects, under conversational speech, theater speech, and various singing conditions. Correlations between mouth geometry and directivity patterns are analyzed and acoustic models are investigated through numerical methods previously used for hearing directivity research efforts (e.g. Boundary Element Method).

Data reduction of directivity patterns is performed for improved integration of vocal directivity in spatial audio renderings (e.g. via clustering and spherical harmonic decomposition). These methods will allow for adjustments of vocal directivity in real-time systems, without the need for lengthy individual specific simulation pre-calculation steps.

In a second part, the effect of distance in voice communication is investigated. Perceived speaker or singer distance and changes in voice production according to the supposed distance to the listener or audience are studied. It is obvious that distance has an effect on perception with the perception of distance being one of the most important ecological features in human audio processing. Another important and often neglected aspect is the adaptation of voice production to the listener or audience distance. This very important effect will be modeled in term of voice quality changes according to the audience distances.

In a third part, interactions between room acoustics (e.g. in opera house, theater, concert hall, or virtual conference room) and spatial perception and production for singing and vocal performances are investigated. This includes the interactions of the vocal radiation directivity patterns, singer/speaker head orientations, dynamic movements, and the architectural space. In the context of virtual conference rooms or collaborative virtual environments, this includes interactions between the acoustic and vocal directivity of the real environment and the remote environment, in order to create an improved sense of presence and co-locatedness between participants.

Expected results

The results of this thesis will be a better understanding of ecological voice interaction, including adaptation to communication distance and relative head positions. It will improve rendering in virtual audio applications: position of the speaker or singer avatars in the scene, orientation of the avatar’s head, and interaction with the acoustics of the virtual room as well as rendering environment.
Ecole doctorale SMAER  
*Sciences Mécaniques, Acoustique, Electronique, Robotique*

**References**

7. M. Noisternig, B. FG Katz, C. d’Alessandro, « Spatial rendering of audiovisual synthetic speech used for immersive environment » Abstract in Acoustics'08