Chapter 5
Improving Audio Spatialization Using Customizable Pinna Based Anthropometric Model of Head–Related Transfer Functions

Navarun Gupta
University of Bridgeport, USA

Armando Barreto
Florida International University, USA

ABSTRACT

The role of binaural and immersive sound is becoming crucial in virtual reality and HCI related systems. This chapter proposes a structural model for the pinna, to be used as a block within structural models for the synthesis of Head-Related Transfer Functions, needed for digital audio spatialization. An anthropometrically plausible pinna model is presented, justified and verified by comparison with measured Head-Related Impulse Responses (HRIRs). Similarity levels better than 90% are found in this comparison. Further, the relationships between key anthropometric features of the listener and the parameters of the model are established, as sets of predictive equations. Modeled HRIRs are obtained substituting anthropometric features measured from 10 volunteers into the predictive equations to find the model parameters. These modeled HRIRs are used in listening tests by the subjects to assess the elevation of spatialized sound sources. The modeled HRIRs yielded a smaller average elevation error (29.9°) than “generic” HRIRs (31.4°), but higher than the individually measured HRIRs for the subjects (23.7°).

INTRODUCTION

Immersive sound systems are an essential part of any modern virtual reality system. Computer games, navigation systems, and teleconferencing – all require a life like reproduction of sound that surrounds the user and makes human computer interaction effortless and authentic.

Audio spatialization, the ability to impart a virtual originating location to digital sounds, is gaining increasing relevance due to its continuously expanding applications in the entertainment
industry, computer gaming, and also in the broader fields of Virtual Reality (VR) and Human-Computer Interaction (HCI). Current advances in this field are ultimately predicated on an inextricable interplay between physics, signal processing, audiology, cognitive neuroscience and anthropometry.

It is known that we are able to discern the location of the source of a sound, around us, by exploiting localization clues that are part of the sounds that reach our eardrums. If a well-defined sound, such as the buzzer of an alarm clock is originated at a given location relative to the position of a static listener who faces North (e.g., two meters away from the listener, at ear level, from the NE direction), the deflections of the listener’s right eardrum will not be the same as those of the buzzer’s diaphragm. The acoustic signal originated in the buzzer will be transformed as it is propagated to the listener’s right ear, due to multi-path reflections involving the torso, head and (right) outer ear of the listener. Furthermore, the acoustic signal reaching the listener’s left ear will be affected by the same effects and, in addition, it may be affected by diffraction, as the listener’s head is interposed between the buzzer and the his/her left eardrum. From a dynamic system’s point of view, the transformation of sound from source (buzzer) to the destination (e.g., right eardrum) can be modeled as a transfer function, which describes how each frequency component of the buzzer sound (and any other sound originated at the same source location, for that matter), is modified in magnitude and phase, as it travels from source to destination. Evidently, the HRTF pair changes as we consider different sound source locations around the listener. Consider, for example, that the buzzer is now placed two meters away from the static listener, but in a position west from his/her location. Clearly, now the right HRTF will indicate a more severe attenuation of high frequencies than the left HRTF. Perceptually, our brains are capable of discerning the location of a sound source by comparing the transformation suffered by the sound from its origin to each one of our eardrums (“binaural localization clues”), or by learned recognition of characteristic features present in even the sounds reaching a single one of our eardrums (“monoaural localization clues”).

Clearly, if the HRTF pair corresponding to a given source location around a listener can be specified and, later implemented by engineering means, we could process a single-channel sound signal (e.g., a monoaural recording of the buzzer sound collected right next to the buzzer), and process it by both HRTFs in the pair, resulting in a binaural sound that, when delivered through headphones to a listener, should produce a similar sound localization perception to the one experienced when the buzzer was physically placed there. If HRTF pairs are defined for many locations around the listener, one could change the HRTF pair used for the sound spatialization process described and, therefore re-assign the buzzer sound to a different “virtual source location”.

Even the highly simplified overview of the foundations of most contemporary sound spatialization approaches, summarized above, highlights the critical role played by the “correct” specification of the HRTF pairs for locations around the listener and its proper implementation, toward the generation of convincing spatialized binaural sounds. The description given also implies that the required HRTF pairs should be defined individually for the prospective listener, since the physical elements cited as their defining factors (torso, head and outer ears) are clearly individual and must match the localization clues that each
Related Content

Counting People Using Blobs and Contours
[www.igi-global.com/article/counting-people-using-blobs-and-contours/87247?camid=4v1a](www.igi-global.com/article/counting-people-using-blobs-and-contours/87247?camid=4v1a)

Analysis and Comparison of Clustering Techniques for Chronic Kidney Disease With Genetic Algorithm

Image Processing for Localization and Parameterization of the Glandular Ducts of Colon in Inflammatory Bowel Diseases
[www.igi-global.com/chapter/image-processing-localization-parameterization-glandular/77570?camid=4v1a](www.igi-global.com/chapter/image-processing-localization-parameterization-glandular/77570?camid=4v1a)

[www.igi-global.com/chapter/machine-vision-application-on-science-and-industry/197036?camid=4v1a](www.igi-global.com/chapter/machine-vision-application-on-science-and-industry/197036?camid=4v1a)