A framework for geometry acquisition, 3-D printing, simulation, and measurement of head-related transfer functions with a focus on hearing-assistive devices

Stine Harder [a,*], Rasmus R. Paulsen [a], Martin Larsen [b], Søren Laugesen [c], Michael Mihocic [d], Piotr Majdak [d]

[a] Technical University of Denmark, DTU Compute, DK-2800 Lyngby, Denmark
[b] Oticon A/S, DK-2765 Smørum, Denmark
[c] Eriksholm Research Centre, DK-3070 Snekkersten, Denmark
[d] Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria

ARTICLE INFO

ABSTRACT

Individual head-related transfer functions (HRTFs) are essential in applications like fitting hearing-assistive devices (HADs) for providing accurate sound localization performance. Individual HRTFs are usually obtained through intricate acoustic measurements. This paper investigates the use of a three-dimensional (3D) head model for acquisition of individual HRTFs. Two aspects were investigated; whether a 3D-printed model can replace measurements on a human listener and whether numerical simulations can replace acoustic measurements. For this purpose, HRTFs were acoustically measured for four human listeners and for a 3D printed head model of one of these listeners. Further, HRTFs were simulated by applying the finite element method to the 3D head model. The monaural spectral features and spectral distortions were very similar between re-measurements and between human and printed measurements, however larger deviations were observed between measurement and simulation. The binaural cues were in agreement among all HRTFs of the same listener, indicating that the 3D model is able to provide localization cues potentially accessible to HAD users. Hence, the pipeline of geometry acquisition, printing, and acoustic measurements or simulations, seems to be a promising step forward towards in-silico design of HADs.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Human listeners are able to localize sounds in space in terms of assigning direction and distance to the perceived auditory image [1]. This ability is an essential function of spatial hearing, which involves further perceptual effects like the estimation of the apparent source width, spatial unmasking of speech (cocktail party effect), and externalization (out-of-head perception) of sounds. Generally, spatial hearing relies on physical directional acoustic features which are the consequence of acoustic filtering of the sound by the pinna, head, and torso. The filtering can be described by the head-related transfer functions (HRTFs), which represent the transfer functions from sound sources to a sound receiver, usually placed at the entrance of the ear canal. For far-field sources, HRTFs depend on the direction of sound incidence [2]. For near-field sources, HRTFs additionally depend on the distance between the source and the ear [3].

An HRTF encodes monaural spectral cues, which are used by human listeners to estimate the sound-source position along sagittal planes (top, down, front, back). For the localization of the source along horizontal planes (left, right), interaural cues are used, namely, interaural time and level differences (ITDs, ILDs). In particular, ITD cues in the frequency range below 1.5 kHz are important for the sound localization in the horizontal planes. The interaural cues are encoded in a binaural pair of HRTFs, thus, a binaural set of HRTFs can be used to describe all directional cues required for spatial hearing.

HRTFs can be used for many purposes, e.g., for models of spatial hearing [4–6], for fitting of hearing aids [7,8], and for presenting virtual binaural audio signals via headphones in so-called virtual auditory displays [9,10]. As HRTFs depend on the

individual geometry of the listener's head and ear, HRTFs are listener-specific [11]. Listener-specific HRTFs are often acoustically measured by placing small microphones at the entrance of the listener's ear canals. During a measurement session, often consisting of measurements for many spatial positions, the listener must sit still for tens of minutes depending on the measuring facilities [12–14]. Thus, it is not surprising that in some applications, generic HRTFs, i.e., HRTFs of a manikin representing an average of the human population [15–17], are used. However, in applications like fitting hearing aids to children [18] or providing accurate sound localization performance via headphones [19], it might be important to consider listener-specific HRTFs. As an alternative to demanding acoustic measurements, HRTF can also be numerically calculated from a three-dimensional (3D) representation of a human geometry [20–23] and established methods have shown good congruence between measured and simulated HRTFs for frequencies below 7 kHz [23,24].

In this study, a method for obtaining listener-specific HRTFs with the focus on virtual product design is evaluated. An application could be optimization of the directional microphone using the Directivity Index as an evaluation criteria, as is performed in.[1] Virtual product design is an emerging discipline that has the potential of reducing production costs and creating more comfortable and better functioning wearables as for example clothes, helmets, and in our case, hearing-assistive devices (HAD), which are commonly used to treat hearing impairment. An example of virtual product design is the size-China project that aims at creating a population statistics on human heads for product design [25]. In [26], a parametric model of the entire human body was computed based on 250 full body scans. This model was then used to synthesize plausible body shapes as input to product design. In another example, foot shape was investigated in several studies driven by the large footwear industry and recently, 50 surface scanned feet were used as input to a statistical analysis of shape with the goal to produce optimized shoe lasts [27]. Similarly, in the design of HADs, the position of the microphones is essential, not only for the design of the casings, but also for capturing spatial acoustic cues. Thus, different types of microphones at various places have been proposed, for a recent review, see [28]. The measurement of HRTFs with HADs is even more demanding than the measurement done in the ear canal because of the much larger degree of freedom: a simple positioning of the HAD might yield different HRTFs, and many microphone positions have to be considered in the individualization of the HAD. On the other hand, since most of the HADs focus on transmitting frequencies below 8 kHz [29], the evaluation of HRTFs can be limited to the upper frequency 8 kHz.

This paper presents yet another step towards full in-silico design of HADs. The goal is to evaluate whether simulated HRTFs can replace expensive acoustic measurements in the future. We simulate listener-specific HRTFs based on a reconstruction of the listeners' geometry from a set of 3D surface scans with missing data. An evaluation of methods for optical 3D scanning of human pinnas is available in [30]. HRTF measurements are expensive with respect to measurement equipment and measuring time for the listener, in particular when measuring HRTFs for different microphone positions in an HAD. In a simulation re-positioning of the microphones does not require additional participation of the listener. HRTF simulations can have a further impact on the future product design: by using statistical shape modeling (as in for example [26,27]), HRTFs of future HADs could be obtained for

hundreds or even thousands of plausible virtual heads, in contrast to the current practice where one or a few generic manikin heads are used because of the time-consuming measurement procedure.

In the following, the three stages of investigations are described. First, a high-quality 3D geometrical model of a human listener was used to numerically simulate corresponding HRTFs. Second, computer-aided-design modeling and 3D printing were used to obtain a listener-specific 3D printed head model, from which acoustically measured HRTFs were obtained. Third, acoustical measurements were performed to obtain HRTFs of the corresponding actual human listener. This three-step pipeline allowed for a thorough evaluation of the differences among acoustically measured and simulated HRTFs of the model and the listener. The acoustically measured HRTFs for the printed head model made it possible to distinguish between HRTF deviations caused by differences in shape (human versus obtained 3D model) and differences in method (measurements versus simulations).

2. Material and methods

2.1. The surface model

A complete human head model was generated using a framework which combines a number of 3D surface scans captured from different angles surrounding the individual. Listener NH167 from the ARI database (http://sofaconventions.org) was scanned. The method was presented in [31] and it consists of both scanning, stitching, and meshing. Here, the method is briefly reviewed.

Sixteen 3D surface scans were acquired using a 3D facial scanner (Canfield scientific Vectra M3) and subsequently aligned with a rigid-body transformation based on a few manually annotated landmark points. The result was a roughly aligned point cloud where many parts of the head were well covered by samples, but several regions lacked well-defined sample points. The regions which lack sample points were in particular hairy areas that were not covered by a cap, such as eyebrows or sideburns.

Surface reconstruction was applied on the point cloud. It was based on the Markov Random Field reconstruction in [32], which is well-suited for human head scans because it (1) is robust to noise, (2) interpolates in areas with missing data in an anatomically plausible way, and (3) accurately aligns the partially overlapping scans. An implementation of the higher-order mathematical model are found at (http://www2.imm.dtu.dk/image/MRFSurface/) along with examples. The scanning process lasted for approximately 15 min, while the stitching required half a day's work. The stitching process has however not been optimized or completely automated yet, which could decrease the required time dramatically.

By visual inspection it can be seen that the anatomy is correctly described by the final surface model. All anatomical structures are present and the surface contains no holes, see Fig. 1(a). Further information on the surface reconstruction accuracy can be found in [32].

2.2. Adapting 3D surface models to simulation and printing

By nature, the 3D surface model acquired from range scanning is a two-dimensional (2D) surface embedded in 3D space. Technically it consists of vertices and triangles, which also applies to the results of the surface merging and stitching. Several numerical simulation packages exist that can use triangulated surfaces as input but most programs, such as Ansys, Comsol Multiphysics, or Abaqus, work better with surfaces represented as non-uniform rational B-splines (NURBS) [33]. Thus, the surface model was converted to NURBS (Geomagic Studio, www.geomagic.com), see Fig. 1(b).

The head model represented in NURBS was further processed using a computer-aided design software package (Creo Parametric

---

[1] Optimizing hearing-aid directionality from measurements and simulations, S. Harder, R.R. Paulsen, M. Larsen, M.S. Pedersen, S. Laugesen, M. Mihocic, and P. Majdak, manuscript in preparation.