Perceptually motivated analysis of numerically simulated head-related transfer functions generated by various 3D surface scanning systems

Dinakaran, Manoj; Brinkmann, Fabian; Harder, Stine; Pelzer, Robert; Grosche, Peter; Paulsen, Rasmus Reinhold; Weinzierl, Stefan

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Numerical simulations offer a feasible alternative to direct acoustic measurement of individual head-related transfer functions (HRTFs). For the acquisition of high quality 3D surface scans, as required for these simulations, several approaches exist in the literature. This paper systematically analyzes the variations between different approaches and evaluate the influence of the accuracy of 3D scans on the resulting simulated HRTFs. To assess this effect, HRTFs were numerically simulated based on 3D scans of the head and pinna of the FABIAN dummy head scanner using 6 different methods. These HRTFs were analyzed in terms of interaural difference, interaural level difference, energetic error in auditory filters and by their modeled localization performance. From the results, it is found that a geometrical precision of about 1 mm is needed to maintain accurate localization cues, while a precision of about 4 mm is sufficient to maintain the overall spectral shape.

Acquisition of Meshes using different scanning systems

We acquired 3D surface scans of the head and pinna of the FABIAN dummy head by using 6 different methods (cf. Fig.1). a) GOM ATOS 3 (GOM-Ref): Stationary, structured light scanner (0.01 mm point resolution).  
 b) Artect Spacel (SPY): Hand-held structured light scanner, scanning at a working distance of 0.6 m to 0.3 m (0.01 mm point resolution).  
 c) Canfield Vectra M3 (CAN): Stationary, stereo photogrammetry technology system, scanning at a working distance of 1 m (0.1 mm point resolution).  
 d) Microscan (KIN): Low cost IR scanner with a working distance between 0.5 m to 0.3 m (0.1 mm point resolution).  
 e) Autodesk 123D catch (123D): Mobile application which allows the user to scan a 3D model from at least 5 to 6 overlapping photos.  
 f) The Python Photogrammetry Toolbox (PPT): An open source tool which has a pipeline to construct a 3D model from a set of photos.

Alignment, Re-meshing & HRTF Simulations

1. In the first step, interaural axis and interaural center of the GOM-Ref mesh were aligned to the origin of coordinates.  
2. Then, the remaining FABIAN surface scans were then aligned with respect to GOM-Ref using the iterative closest point (ICP) algorithm from the surface manipulation and transformation toolkit (SYMATHA) [3].  
3. A priori mesh grading algorithm (resulting in non-uniform meshing) was deployed according to Ziegelwanger et al. which result in meshes with respect to the distance from the ear [4].  
4. Two different models were generated for each scanning method: One for the left pinna (with small mesh elements at the earfringe and large elements at the right), and one for the right pinna.  
5. The target lengths used were 1 mm to 10 mm, which resulted in around 20,000 elements per mesh.  
6. For numerical HRTF simulation, the MaskHRTF implementation of the 3-dimensional Butter-Miller calculation BEM was used [5].

Conclusion

A high precision of about 1 mm is needed when capturing the pinnae geometry to assure accurate localization cues. This criterion was met only by the SPY and CAN scanning methods.  
However, the overall coloration showed to be below 1 dB, even for geometric errors of up to 4 mm, which occurred for the KIN method.  
The remaining methods (123D and PPT) showed geometric deviation of up to 5 mm and slightly larger coloration of up to 1.5 dB.