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

Link back to DTU Orbit

Citation (APA):
Numerical simulations offer a feasible alternative to the 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 this paper, we systematically analyze the variations between different approaches and evaluate the influence of the accuracy of 3D scans on the resulting simulated HRTF. To assess this effect, HRTFs were numerically simulated based on 3D scans of the head and pinna of the FABIAN dummy head and 6 different methods. These HRTFs were analyzed in terms of interaural time difference, interaural level difference, energetic error in auditory filters and by their modeled localization performance. From the results, it is found that a geometric precision of about 1 mm is needed to maintain accurate localization cues, while a precision of about 0.4 mm is sufficient to maintain the overall spectral shape.

Motivation

➢ In recent years, several approaches have been proposed with a focus on increasing the accuracy of the simulated HRTF by acquiring high quality 3D scans of head and pinna. Different techniques for the acquisition of 3D surface scans exist such as MRTF scans, structured light scanners, laser scanners, infrared scanners, stationary scanners, hand held scanners, or by using mobile camera phones [1].

➢ Each of them provides a different resolution and its accuracy directly affects the quality of the numerically simulated HRTFs which are subject to research.

➢ Here, we systematically analyze the accuracy of 3D surface scans obtained by different approaches and study their influence on the resulting HRTFs by means of interaural time difference (ITD), interaural level difference (ILD), energetic error in equivalent rectangular bandwidth (ERB) auditory filters, and their simulated localization performance.

➢ To isolate the influence of the scanning method on the HRTF, the different scanning methods were evaluated against a high resolution structured light scan (ground truth) which showed a very good agreement to its acoustically measured counterpart in an anechoic studio [2].

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).

1. gOM ATO5-I (GOM-Ref): Stationary, structured light scanner (0.01 mm point resolution).
2. Artic Space (SPY): Hand-held structured light scanner, scanning at a working distance of 0.2 m to 0.3 m (0.05 mm point resolution).
3. Canfield Vector M3 (CAN): Stationary, stereo photogrammetry technology scanning at a working distance of 1 m (0.1 mm point resolution).
4. Microsight KIN (KIN): Low cost IR scanner with a working distance of 0.5 m to 1.5 m (0.2 mm point resolution).
5. Autodesk 123D catch (123D): Mobile application which allows the user to scan a 3D model from at least 5 to 6 overlapping photos.
6. The Python Photogrammetry Toolbox (PPT): An open source tool which has a pipeline to construct a 3D model from a set of photos.

\[ \text{Fig.1: FABIAN 3D Surface Scans using (a) GOM ATO5-I Scanner (GOM-Ref), (b) Artic Space Scanner (SPY), (c) Canfield Vector M3 scanner (CAN), (d) Kinect scanner (KIN), (e) Autodesk 123D (123D), and (f) PPT (PPT).} \]

Acoustic comparison of the different scanning systems

We compared the 3D meshes (GOM-Ref, SPY, CAN, KIN, 123D, and PPT) using the following metrics:

- **Scan Quality Metrics**
  - **Resolution** (SPY, CAN, KIN, 123D, PPT): Mesh resolution in mm.
  - **Point Cloud Error (PCE)** (SPY, CAN, KIN, 123D, PPT): Error in mm.
  - **Feature Consistency (FC)** (SPY, CAN, KIN, 123D, PPT): Consistency in mm.
  - **Surface Consistency (SC)** (SPY, CAN, KIN, 123D, PPT): Consistency in mm.
  - **Surface Consistency (SC)** (SPY, CAN, KIN, 123D, PPT): Consistency in mm.

- **Geometric Accuracy** (SPY, CAN, KIN, 123D, PPT): Accuracy in mm.
  - **Average spectral difference (ASD)** (SPY, CAN, KIN, 123D, PPT): ASD in dB.

\[ \text{Table 2: PCE, GE, FC, and SC with respect to GOM-Ref (in mm).} \]

Conclusion

- A high precision of about 1 mm is needed when capturing the pinna 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

Reference