Manoj Dinakaran, Peter Grosche, Fabian Brinkmann, Stefan Weinzierl

Extraction of Anthropometric Measures from 3D-Meshes for the Individualization of Head-Related Transfer Functions

Extraction of anthropometric measures from 3D-meshes for the individualization of head-related transfer functions

Manoj Dinakaran[1,2], Peter Grosche[1], Fabian Brinkmann[2], and Stefan Weinzierl[2]

[1]*Huawei Technologies, European Research Centre, Riesstraße 25, 80992 Munich, Germany.*
[2]*Technical University Berlin, Audio Communication Group, Einsteinufer 17c, 10587 Berlin, Germany.*

Correspondence should be addressed to Manoj Dinakaran (manoj.dinakaran@huawei.com)

ABSTRACT

Anthropometric measures can be used for the individualizing of head-related transfer functions (HRTFs), e.g. by selecting best match HRTFs from a large library or by manipulating the HRTF with respect to anthropometric features. For this purpose, an accurate extraction of anthropometric measures is crucial, since small inaccuracies may already influence the quality of the individualization. Anthropometric features can be measured in different ways, e.g. from scale pictures or anthropometers. However, these approaches are time-consuming and prone to bias of the investigator. In the current project, we evaluated the precision of anthropometric features automatically extracted from individual 3D head-and-shoulder meshes generated with a low-cost 3D scanning device (the Kinect sensor), by identifying and measuring distances between characteristic points on the outline of each mesh. A comprehensive set of anthropometric features was automatically extracted for 61 subjects. By cross-validation with the manually extracted values, the method was found to yield accurate and reliable estimates of corresponding features.

1 Introduction

The individualization of head-related transfer functions (HRTFs) provides an approach to improve the quality of binaural synthesis, e.g. by maintaining an accuracy of localization comparable to the corresponding real sound fields [1]. The most precise approach to obtain individual HRTFs is a direct acoustic measurement. However, this approach requires elaborate experimental equipment [2, 3]. It is thus of special interest to individualize HRTFs by directly relating the listener's anatomy to acoustic features of HRTFs, which seems reasonable as the salient auditory features in HRTFs originate from the impact of the listener's torso, head and outer ears (pinnae) on the incident sound field. This correspondence was statistically confirmed by Jin *et al.* [4] who showed that a functional model relating morphological measurements to HRTF magnitude spectra by multivariable linear regression could explain about 90% of the variance in the HRTF spectra. Other possibilities for matching morphologic and acoustic features include numeric solutions, machine learning algorithms, geometric head models, or sparse representation methods [5, 6, 7, 8]. Moreover, in localization experiments conducted by Zotkin *et al.* [9], six out of eight subjects

showed improved localization of elevation when using best-match HRTFs compared to those of a head and torso simulator (KEMAR). HRTFs were chosen by the best-matching anthropometric measures in the CIPIC HRTF database [10].

All methods above require a reliable tool for measuring anthropometric parameters which can be done in several ways. The simplest solution to this is a manual measurement either using anthropometers or scale pictures as done in [9, 10], which, however, is time consuming, and might introduce bias from the investigator. In case the pictures are taken from a small distance, additional error might be introduced if the object to be measured (e.g. the pinna) and the scale have different distances to the camera.
Torres-Gallegos *et al.* [11] suggested an automatic extraction from scale pictures using trained active shape models. The pictures were taken in a blue screen photo studio with good lighting conditions and known camera-to-subject distance. It was shown that the mean differences between manually and automatically extracted measures were in the range of 0.7 mm to 1.6 mm for the ear parameters, 4 mm to 6.5 mm for head parameters and 2.8 mm to 5.6 mm for neck parameters of 11 subjects respectively. The influence of the measurement error was evaluated by calculating spectral differences between best-match HRTFs that were selected with and without adding error to one anthropometric measure. Spectral differences were small in this case, but it was not investigated to what extent the results would change, if the error were introduced to all measures at the same time, which would be a more realistic test scenario.
In this paper, we have used the Kinect sensor to generate 3D surface meshes of different listeners' torso, head and pinna. From these representations, anthropometric measures were automatically extracted, aiming at a consumer solution which is robust against the scanning conditions and easy to handle. In section 2, we introduce the scanning method, and the mesh post-processing. In section 3, the extraction of anthropometric measures is outlined, followed by an evaluation in section 4.

2 Mesh acquisition

High precision surface meshes can be obtained using magnet resonance imaging [12, 13], but such scanners can only be operated by medical personnel and are extremely costly. In this paper, a Kinect 3D scanner, and the Kinect fusion with the developer toolkit browser v1.8.0 [14] was used to generate 3D surface meshes from which the anthropometric measures can be extracted automatically.

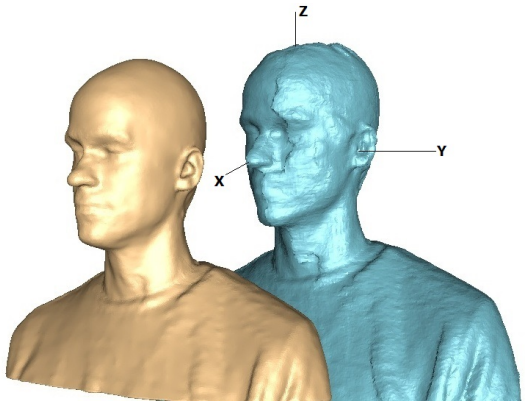

Fig. 1: Example for a 3D surface mesh before (right) and after post processing (left).

The Kinect was set up at ear level at a distance of about one meter from the subject who was sitting in a swivel chair, and was wearing a swim cap to reduce the influence of hair on the scans and the extracted anthropometric measures. The scans were taken in a two-step procedure with the Kinect resolution set to the maximum of 768 voxels per meter. First, a complete mesh was generated by slowly rotating the subject by 360 degrees. In this case, the rotation caused a slight spatial smoothing in the mesh. To obtain non-smoothed meshes, separate scans from the subject's ears, and left, and right side of the face were taken, while carefully rotating it back and forth by a couple of degrees until the result was satisfactory.
Post-processing involved four manual steps: (a) removing unwanted parts in the mesh (e.g. surroundings, unwanted parts of the torso, and irregular parts close to holes), (b) filling holes, (c) merging, and (d) aligning. Removing unwanted parts and filling holes, which commonly occurred at the top of the head and below the jaw, was done using Meshlab [15]. In the merging stage, which was done using Geomagic's [16] point based glue tool, the non-smoothed meshes were used to replace corresponding parts of the complete, but smoothed mesh. In the last step, the interaural center, which is defined as the midpoint of the axis connecting the entrances to the ear canals, was aligned to the origin of coordinates. An example of the resulting mesh, before and after post-processing is shown in Fig. 1.

The whole procedure took approximately 3.5 hours per subject including 15 minutes for scanning.

3 Automatic extraction of anthropometric measures

For individualizing HRTF using anthropometric features, a set of 27 parameters as defined by the CIPIC HRTF database [10] is widely used, describing the basic dimension of head, neck, shoulder, torso, and pinna. In this study, we present an approach for automatically estimating a subset of 14 anthropometric measurements of head, neck and pinna from 3D surface. (cf. Table.1). With these automatic measurements we avoid the difficult manual selection of the correct points in the mesh which would lead to inaccurate measurements.

3.1 Extraction of the mesh outline

As the first step to measure anthropometrics automatically, we extracted 2D outlines from the 3D surface meshes. Different outlines need to be extracted for the above mentioned 14 anthropometric parameters: one from the front view, and one from the side view and four of each pinna (cf. Fig. 3-7). From the front view outline, parameters x_1 and x_6 are extracted, from the side view x_2, x_3, x_4, x_5, x_7 and x_8 are extracted, and parameters d_1 - d_6 were taken from the pinna outlines. Outlines were extracted with a resolution of $\Delta z = 0.75$ mm by finding the points with minimum, and maximum x-values for the front, and back part of the side view outline, and by finding the points with the minimum, and maximum y-values for the left, and right part of the front view, and pinna outlines. For a robust extraction of the outlines at the current z-value, points were grouped by $[z-\Delta z/2,\ z+\Delta z/2]$ before finding the minimum/maximum x/y-values of all points in this search range.

3.2 Head measures

Head anthropometric parameters are head width (x_1), height (x_2), and depth (x_3). The head width (x_1) was calculated by averaging the distances in y-direction between the points directly above and below the ears (cf. Fig. 3). For this purpose, differences in y-coordinates $\Delta y_{i,i+1}$ were calculated for successive points on the left and right part of the outline between $z = [-Z/2, Z/2]$, where Z is the maximum z-value of the entire mesh corresponding to a peak point on the top of the head. The points above and below the ear were then extracted by finding the maximum, and minimum of $\Delta y_{i,i+1}$.

The head height (x_2) was identified based on the front part of the side view outline (i.e. points with positive x-coordinates) by calculating the distance in z-direction between Z, and the point with minimum $\Delta z_{i,i+1}$ within $z = [-Z,\ -Z/2]$ (green point in Fig. 4). Here, $\Delta z_{i,i+1}$ gives the differences in z-coordinates between successive points on the outline.

The head depth (x_3) was estimated based on the side view outline evaluated within $z = [Z/3,\ Z]$, by calculating the distance between the yellow point at the front of head and the blue point at the back of head shown in Fig. 4. The points on the forehead, and backhead were selected by finding the minimum $\Delta x_{i,i+1}$ for points

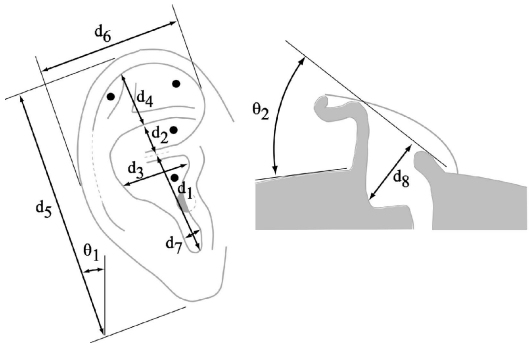

Fig. 2: Definition of pinna measures, taken from [10].

ID	Anthropometric measure	ID	Anthropometric measure
x_1	head width	x_{15}	seated height
x_2	head height	x_{16}	head circumference
x_3	head depth	x_{17}	shoulder circumference
x_4	pinna offset down	d_1	cavum conchae height
x_5	pinna offset back	d_2	cymba concha height
x_6	neck width	d_3	cavum concha width
x_7	neck height	d_4	fossa height
x_8	neck depth	d_5	pinna height
x_9	torso top width	d_6	pinna width
x_{10}	torso top height	d_7	intertragal incisure width
x_{11}	torso top depth	d_8	cavum concha depth
x_{12}	shoulder width	Θ_1	pinna rotation angle
x_{13}	head offset forward	Θ_2	pinna flare angle
x_{14}	height		

Table 1: Anthropometric measures from the CIPIC database [10]. Measures with gray font were not acquired in this study.

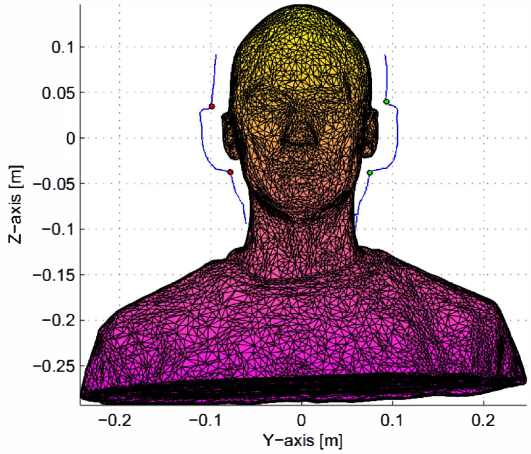

Fig. 3: Front view of 3D surface mesh with the extracted outlines for measuring head width.

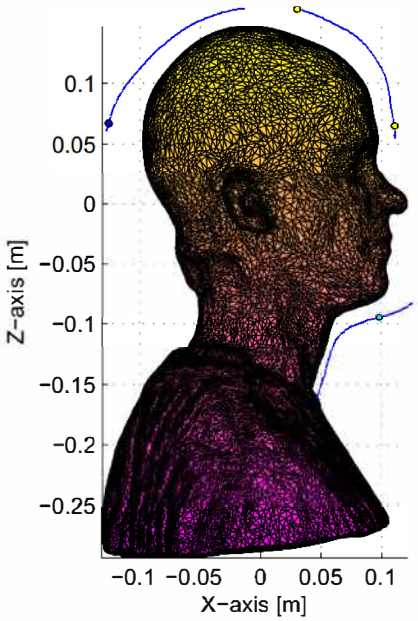

Fig. 4: Side view of 3D surface mesh with the extracted outlines for measuring head depth and head height.

with positive, and negative x-coordinates, respectively, where $\Delta x_{i,i+1}$ gives the differences in x-coordinates between successive points on the outline.

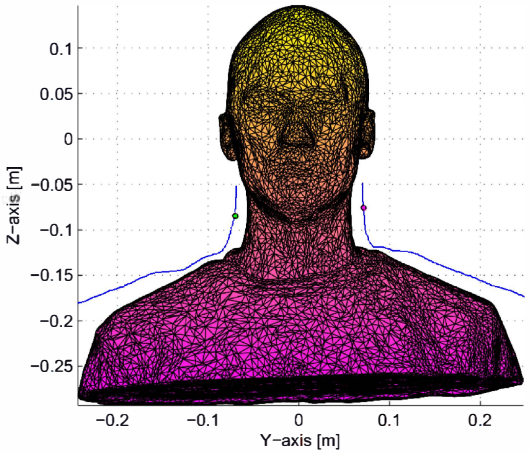

Fig. 5: Front view of 3D surface mesh with the extracted outlines for measuring neck width.

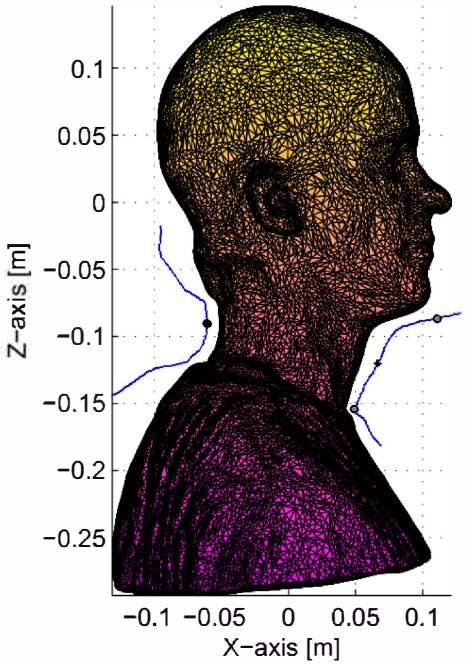

Fig. 6: Side view of 3D surface mesh with the extracted outlines for measuring neck height and neck depth.

3.3 Neck measures

Neck anthropometric paramaters are neck width (x_6), height (x_7), and depth (x_8). In the first step, the front, and side view outline was extracted in the range $z =$

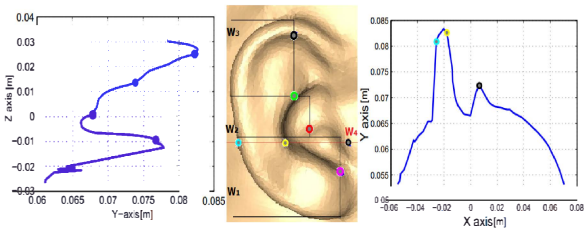

Fig. 7: Vertical (left), horizontal pinna outline (right) and pinna mesh of subject-26 (middle). Points marked in the pinna mesh correspond to the points of the same color marked in the pinna outlines.

$[-Z-Z/8,\ -Z/2]$. Then, the neck width (x_6) was estimated by calculating the distance between the red point in the right neck outline and the green point in the left neck outline shown in Fig. 5. The points were selected by finding the minimum, and maximum $\Delta y_{i,i+1}$ for points with positive, and negative y-coordinates, respectively

The neck height (x_7) was calculated as the distance in z-direction between the gray points on the frontal part of the side view outline (cf Fig. 6). The lower anchor is defined by the point with the smallest x-value, the upper anchor was identical to the point used for calculating the head height (see above).

The neck depth (x_8) was calculated as the distance in x-direction between the point with the largest x-value on the back part of the side view outline, and the point half way between the two points that were used to measure the neck height (cf. Fig. 6).

3.4 Pinna offset measures

The Pinna offset is defined as the displacement of the left and right ear channel entrance from the center of the head in x-, and z-direction [10]. Since the meshes were manually aligned so that the ear channel's x-, and z-coordinates were zero, the pinna offset is given by the displacement of the head center. The latter was simply estimated by the arithmetic mean of the points that were used to calculate the head depth and height (see above).

3.5 Pinna measures

Pinna parameters measured here are cavum concha height (d_1), cymba concha height (d_2), cavum concha width (d_3), fossa height (d_4), pinna height (d_5) and pinna width (d_6) as shown in Fig. 2. In this case, the range $z = [z_{\min}, z_{\max}]$ for extracting the outline was taken from the four points used to calculate the head width (cf. Fig. 4). The same points were used to estimate the left and right pinna height (d_5). This was done by calculating the distance between the two red points on the left and two green points on the right part of the front-view outline.

In a next step, the pinna height was used to subdivide the pinna into three vertical segments (w_1, w_2, and w_3) whose lengths in z-direction are defined by the ratio 2:1:2. To access the fine structure of the pinna, corresponding outlines were extracted at different x-coordinates: $x = 0 \pm \Delta x$ mm for w_1, $x = 7 \pm \Delta x$ mm for w_2, and $x = 12 \pm \Delta x$ mm for w_3 (Fig. 7, middle). Δx was introduced for a more robust detection of the outline and was set to 1 mm.

After extraction of the segments, the cavum concha height (d_1) was estimated as the distance between the magenta point and the red point from w_1 and w_2, the cymba concha height (d_2) was estimated as the distance between the red point and the green point from w_2 and w_3, and the fossa height (d_4) was estimated as the distance between the green point and the black point from w_2 shown in Fig. 7 (right). The corresponding points on the segments were selected by finding the minima and maxima of $\Delta y_{i,i+1}$, i.e. the difference in y-coordinates between successive points of the segments.

In the next step, the horizontal outline of the pinna w_4 was extracted by grouping points between $z = \pm 1$ mm in the range of x = $[-X/2,\ X/2]$, where X is the maximum absolute x-value of all points in the mesh. The cavum concha width (d_3) was then estimated as the distance between the black point and the yellow point, and the pinna width (d_6) was estimated as the distance between the black point and the cyan point as shown in Fig. 7 (left). The black point was selected by finding the point with the maximum y-coordinate in positive x-direction. The cyan, and yellow points were selected by finding the minimum $\Delta x_{i,i+1.0}$ for points whose x-value are larger, and smaller than the x-value of the point with the maximum y-value, respectively.

4 Evaluation

In the following, we present an evaluation of the proposed algorithm regarding four aspects: The influence of (a) the mesh quality, and (b) the head geometry on the automatic parameter extraction on one hand, and

ID	St. light [cm]	Kinect [cm]	RE. [%]
x_1	14.81	14.76	0.3
x_2	21.55	21.57	0.1
x_3	18.73	18.65	0.4
x_6	11.85	11.91	0.5
x_7	8.37	8.38	0.1
x_8	11.80	11.82	0.2
d_1	1.55	1.59	2.6
d_2	1.08	1.08	0
d_3	1.79	1.79	0
d_4	2.18	2.18	0
d_5	5.99	6.06	1.2
d_6	3.24	3.26	0.6

Table 2: Anthropometric measures extracted from scans acquired with a structured light (St. light) and Kinect scanner, and Relative error (RE) between the two.

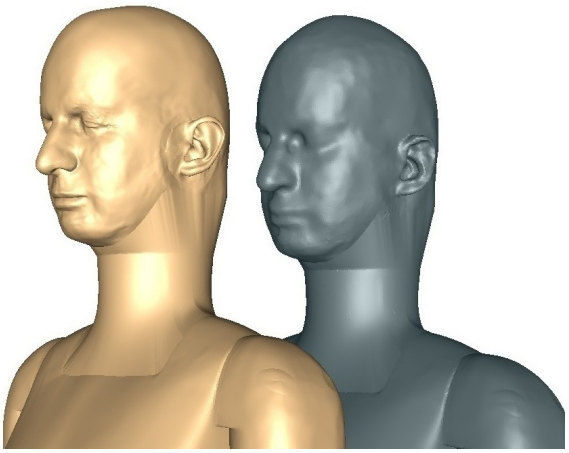

Fig. 8: 3D surface meshes of FABIAN obtained with the structured light scanner (left) and the Kinect scanner (right).

the suitability of the automatically extracted parameters for (c) ITD individualization, and (d) best match HRTF selection on the other.

4.1 Influence of mesh quality

For assessing the influence of the mesh quality, anthropometric features were manually extracted for the FABIAN head and torso simulator using a mesh acquired with a Kinect, and a high resolution mesh obtained from a structured light scan, with a precision of approx. $1/100$ mm [17] [18] (cf. Fig. 8).
The absolute mean difference for the head, neck and pinna anthropometric measurements from the high resolution scan to the Kinect scan is 0.3 mm (Table. 2), whereas the relative error is below 1% in most of the cases, and is always smaller then differences reported by Torres-Gallegos for the analysis of scale pictures *et al.* [11]. It is thus assumed, that the mesh quality that can be achieved with the Kinect scanner is sufficient for the automatic measurement of anthropometrics and should not bias an HRTF individualization process.

4.2 Influence of head geometry

To evaluate the robustness of our method with respect to different head and ear shapes, the Kinect was used to generate 3D surface meshes of 61 human subjects, and the automatically extracted anthropometric measures were compared to the corresponding manually extracted counterparts. Manual extraction was done on the basis of the same meshes, but using Meshlab's distance tool on manually selected points. In 16 cases, the hair covered parts of the neck and the pinna lead to unreliable results not only in automatic but also in the manual approach. In another 5 cases for the female subjects who got hair of very high density covering the entire pinna and neck, both the side view outline and front view outline could not be extracted and the automatic measurements failed for those subjects because no characteristic points can be extracted from the 3D mesh.
Differences between manually and automatically extracted anthropometric measures for the 40 *good* subjects are given in Table. 3. The second column indicates the mean difference (μ) and standard deviation (σ) across all 40 subjects. The third column indicates the absolute mean difference ($|\mu|$) and standard deviation (σ), the fourth column the absolute relative deviation in percent (RE).
In general, difference decrease by only 1.7% if considering only the *good* subjects compared to considering *all* subjects (not shown here). This suggests that most parameters can still be reliably extracted in the presence of smaller mesh irregularities. Mean differences μ show systematic offsets in some cases (x_2, x_6, x_8, d_1, and d_5), and appear to be randomly distributed in the other cases ($\mu \leq 0.5$ mm). Absolute mean (and

ID	μ (σ) [cm]	$\lvert\mu\rvert$ (σ) [cm]	RE. [%]
x_1	-0.05 (0.31)	0.22 (0.22)	1.50
x_2	0.28 (0.41)	0.36 (0.33)	1.69
x_3	0.05 (0.28)	0.18 (0.22)	0.95
x_6	-0.56 (1.20)	0.60 (1.18)	5.4
x_7	0.01 (1.38)	0.87 (1.07)	11.98
x_8	0.24 (0.91)	0.64 (0.68)	5.39
d_1	-0.13 (0.35)	0.23 (0.28)	15.01
d_2	-0.04 (0.22)	0.17 (0.15)	16.17
d_3	-0.01 (0.37)	0.25 (0.27)	14.89
d_4	-0.04 (0.23)	0.17 (0.15)	9.15
d_5	0.20 (0.26)	0.24 (0.21)	4.10
d_6	0.01 (0.30)	0.22 (0.20)	7.55

Table 3: Differences between manually and automatically extracted anthropometric measures of 40 subjects: μ is mean difference, σ is standard deviation, $\lvert\mu\rvert$ is absolute mean difference and RE is relative error using the absolute measurement difference.

relative) differences are in the range of 0.18 cm to 0.87 cm (0.95% to 16.17%) which is comparable to Torres-Gallegos *et al.* [11]. However, standard deviations are larger in our case, suggesting at least some erroneous detections for the automatic feature extraction. The pinna displacement measurements x_4 and x_5 were not compared here, because their manual measurement is prone to errors.

4.3 Suitability for ITD individualization

In the third step, the suitability of the proposed approach for individualizing the ITD based on an ellipsoidal head model using anthropometric measures of head, and pinnae displacement [19] was investigated. The difference between the ITD calculated based on automatic (ITD_A) and manual (ITD_M) measurements is given by

$$\overline{ITD}_D(\theta) = \frac{1}{K}\sum_{k=1}^{K} \left| ITD_M(\theta,k) - ITD_A(\theta,k) \right| , \quad (1)$$

where θ is the azimuth angle, and K is the number of subjects. $\overline{ITD}_D$ is found to be less than 12 μs along the azimuth plane for 56 subjects and less than 5 μs along the azimuth plane for 40 subjects which is in the range of the most critical JND (just noticeable difference) values reported in literature [20, Tab. 2.3] (cf. Fig. 9).

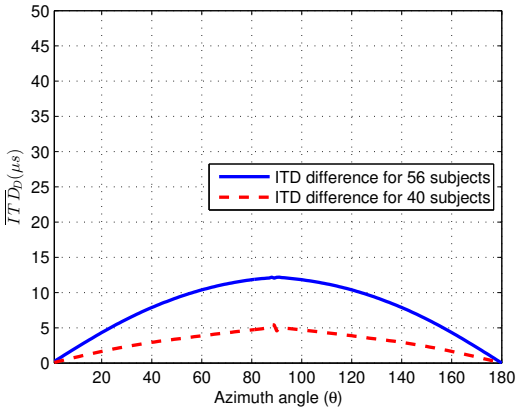

Fig. 9: Absolute mean ITD difference between the automatic and manual anthropometric measurements.

4.4 Suitability for best match HRTF selection

In the fourth step, we calculated distances between anthropometric feature sets for all combinations of the 40 *good* subjects as suggested by Zotkin [9], by using the features listed in Table. 3. Distances were calculated for the manually (M→M) and automatically measured anthropometric feature sets (A→A). The similarity between the distances for the cases (A→A) and (M→M) were assessed by

$$P(T) = \frac{100}{N}\sum_{n=1}^{N} \frac{X_{n,T} \cap Y_{n,T}}{T} , \quad (2)$$

where P is the precision in percent, $N = 40$ is the number of subjects, $T = \{1,2,..N\}$ is the rank, $X_{n,T}$ and $Y_{n,T}$ is the rank list of the T subjects with the smallest distance to subject n calculated based on A→A, and M→M, respectively [21].
From Fig. 10, we can see that the precision is around 80% for $T_{\max} = 20$, i.e. the top 20 rank list selected from M→M and A→A, on average contains 16 identical subjects. If we analyze only the best match ($T_{\max} = 1$), the precision between M→M to A→A is 18% which means that for 7 subjects out of 40 subjects the same best match is selected for M→M and A→A.

5 Discussion and Conclusion

In this paper, we have used the Kinect sensor as a low budget 3D scanner for consumer applications to generate 3D surface meshes of 61 human subjects and one

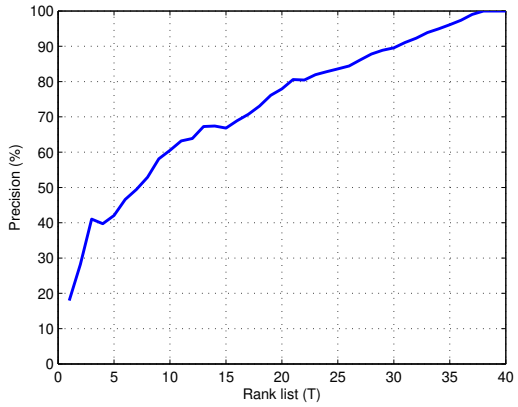

Fig. 10: Precision between the best match subjects selected from manually and automatically measured anthropometrics depending on the size of the rank list T.

head and torso simulator. From these meshes, we automatically extracted 14 anthropometric head, neck and pinna features by calculating distances between characteristic points on the mesh outlines. The suitability of the proposed method for acoustic applications was evaluated by extracting relevant anthropometric parameters using a high resolution reference mesh and a Kinect based mesh of the FABIAN head and torso simulator. For these cases, differences were well below 1% for most cases, and are assumed to be negligible.
The efficiency of the proposed method was investigated for different head geometries by comparing manually and automatically extracted features of the 61 human subjects. The automatic extraction worked for 56 out of 61 subjects and the differences between manual and automatic measurements were in the range of 2 mm for the head measures, less than 2 mm for pinna measures and in the range of 8 mm for the neck measures which is comparable to results from an analysis of scale pictures taken at well-controlled visual conditions [11]. Moreover, the extraction proved to be relatively robust against an occlusion of smaller parts of the mesh, whereas an occlusion of larger parts should be avoided during the scanning procedure.
The influence of the proposed method on HRTF individualization was assessed by calculating distances between anthropometric feature sets of all subjects based on manually and automatically extracted features and generating the rank order of best match subjects from the calculated distances. On average 6 out of the 10 best matches for each subject were identical across automatic and manual extraction, thus suggesting a general accordance between the extraction methods. This is in line with Torres-Gallegos *et al.* [11] who showed that noise in anthropometric measurements influences the best match subject. While induced spectral differences appeared to be negligible in their case, this still has to be proven for the proposed method.

6 Future works

In future works, the quality of the best match HRTFs selected based on the automatically extracted measures in terms of perceptual features such as localization and coloration could be investigated, as well as the possibility to directly generate HRTFs from the acquired 3D surface meshes using numerical approaches. For consumer applications it would be interesting to see to what extent the automatic extraction can also be performed based on 3D surface meshes generated from mobile phone pictures.

7 Acknowledgement

The authors wishes to thank all subjects for their acceptance to measure their anthropometrics, and Kainan Chen for his help with the mesh outlines.

References

[1] H. Møller, M. F. Sørensen, C. B. Jensen, D. Hammershøi, "Binaural technique: Do we need individual recordings?," *Journal of the Acoustical Society of America*, pp. 451-469, 1996.

[2] B. Masiero, M. Pollow, J. Fels, "Design of a fast broadband individual head-related transfer function measurement system," *In Forum Acusticum*, Aalborg, Denmark, 2011.

[3] A. Fuß, F. Brinkmann, F. Jurgensohn, S. Weinzierl, "Ein vollsphärisches Multikanalmesssystem zur schnellen Erfassung räumlich hochaufgeloster, individueller kopfbezogener Übertragungsfunktionen," *In Fortschritte der Akustik – DAGA*, 2015.

[4] C. Jin, P. Leong, J. Leung, A. Corderoy and S. Carlile, "Enabling individualized virtual auditory space using morphological measurements," *IEEE Pacific-Rim Conference on Multimedia*, Sydney, Australia, 2000.

[5] V. Algazi and R. O. Duda, "Estimation of a spherical-head model from anthropometry," *AES*, New York, 2001.

[6] L. Zhou, J. Zhang, H. Ma, Z. Wu, "Head Related Transfer Function Personalization Based on Multiple Regression Analysis," *Computational Intelligence and Security Conference*, Guangzhou, China, 2006.

[7] P. Bilinski, J. Ahrens, M. R. P. Thomas, I. J. Tashev and J. C. Platt, "HRTF magnitude synthesis via sparse representation of anthropometric features," *International Conference on Acoustic Speech and Signal Processing (ICASSP)*, Florence, Italy, 2014.

[8] B. F. G. Katz, "Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation," *Journal of the Acoustical Society of America*, 2001.

[9] D. N. Zotkin, J. Hwaiig, R. Duraiswami, L. S. Davis, "HRTF Personalization using anthropometric measurements," *Applications of Signal Processing to Audio and Acoustics*, New York, 2003.

[10] V. R. Algazi, C. Avendano, R. O. Duda and D. M. Thompson, "The CIPIC HRTF Database," *IEEE Workshop on Application of signal processing to Audio and Acoustics*, New York, 2001.

[11] E. A. Torres-Gallegos, F. Orduña-Bustamente, F. Arámbula-Casío , "Personalization of head-related transfer functions (HRTF) based on automatic photo-anthropometry and inference from a database," *Applied Acoustics*, pp. 84-95, 2015.

[12] A. Reichinger, P. Majdak, R. Sablatnig, S. Maierhofer, "Evaluation of Methods for Optical 3-D Scanning of Human Pinnas," *International Conference on 3D vision*, Lyon, France, 2013.

[13] C. T. Jin, P. Guillon, N. Epain, R. Zolfaghari, A. Schaik, A. I. Tew, C. Hetherington, J. Thorpe, "Creating the Sydney York Morphological and Acoustic Recordings of Ears Database," *IEEE Transactions on Multimedia*, 2014.

[14] Kinect-developer-toolkit, http://www.microsoft.com/en-us/kinectforwindows/develop/downloads-docs.aspx

[15] Meshlab, http://meshlab.sourceforge.net/

[16] Geomagic-Studio, http://www.geomagic.com/en/products-landing-pages/re-designx-wrap

[17] A. Lindau and S. Weinzierl, "FABIAN - An instrument for software-based measurement of binaural room impulse responses in multiple degrees of freedom," *VDT International Audio Convention*, Leipzig, 2006.

[18] F. Brinkmann, A. Lindau, M. Müller-Trapet, M. Vorländer, S. Weinzierl, "Cross-validation of measured and modeld head-related transfer functions," *In Fortschritte der Akustik – DAGA*, pp. 1118-1121, 2015.

[19] V. R. Algazi, C. Avendano and R. O. Duda, "An adaptable ellipsoidal head model for the interaural time difference," *IEEE international conference on acoustics speech and signal processing*, Arizona, 1999.

[20] J. Blauert, "Spatial Hearing: The psychophysics of human sound localization (Revised)," *Camebridge, Massachusetts: MIT Press*, 1997.

[21] B. Hjørland, "The foundation of the concept of relevance," *Journal of the American Society for Information Science and Technology*, pp. 217-237, 2010.