DOI: 10.61782/fa.2023.1082

forum acusticum 2023

COMBINATION OF PHOTOGRAMMETRY AND NON-RIGID PINNA REGISTRATION FOR THE CALCULATION OF PERSONALISED HEAD-RELATED TRANSFER FUNCTIONS

Katharina Pollack[1*] **Piotr Majdak**[1] **Hugo Furtado**[2]

[1] Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria
[2] Dreamwaves GmbH, Vienna, Austria

ABSTRACT

For the calculation of personalised head-related transfer functions (HRTFs), the pinna is the most relevant geometry of a listener, determining the sound-localisation performance in the sagittal planes. One popular approach to acquire the pinna geometry is photogrammetric reconstruction (PR), in which photos from various directions of the listener's geometry are processed yielding a 3D representation, i.e., a point cloud. However, because of the pinna structure and texture, the PR often yields a noisy and non-uniformly sampled point cloud corrupted by outliers, which is disturbed especially in the concave pinna regions that are relevant for plausible HRTFs. Non-rigid registration (NRR) algorithms can be applied to register a clean but generalised pinna point cloud to the noisy but personalised pinna point cloud. However, NRR algorithms can fail in concave pinna geometry regions by means of estimated sound-localisation errors in calculated HRTFs using auditory models. In our work, we propose to integrate confidence levels obtained from the PR in the NRR process in order to improve the quality of the acquired pinna geometry, and subsequently, the quality of personalised HRTFs.

Keywords: *photogrammetry, head-related transfer function, personalised spatial audio*

**Corresponding author: katharina.pollack@oeaw.ac.at.*

1. INTRODUCTION

Head-related transfer functions (HRTFs) play an important role when it comes to personalised binaural audio [1]. They can be acquired with numerical calculation given an accurate 3D representation of the listener-specific pinna [2]. One popular approach to capture the pinna geometry is photogrammetry [3, 4]. Previous research has shown that non-rigid registration (NRR) algorithms can be used to improve the quality of PR-based pinna geometry acquisition [5], but there are still problems for pinna regions that are hard to capture albeit important for plausible personalised HRTFs [6–8]. In this investigation, we proposed to utilise *additional information* from the PR software in order to introduce thresholds to filter the PR point clouds with. With this approach, we want to improve the performance of existing NRR algorithms for the application of pinna point clouds registration.

2. METHODS

2.1 Photogrammetric reconstruction

Photogrammetric reconstruction (PR) describes the process of obtaining a three-dimensional (3D) geometry from 2D information, such as photos or videos. PR previously has been used to acquire the pinna geometry [3], however, the spatial accuracy of regions especially important for plausible HRTFs is not satisfying [5, 7]. We considered three PRs of the same listener in this investigation, to account for some variability in the data.

In the used software (Metashape version 1.6.5, Agisoft LLC), "confidence levels" describe how many depth maps of image pairs were used in the reconstruction of the

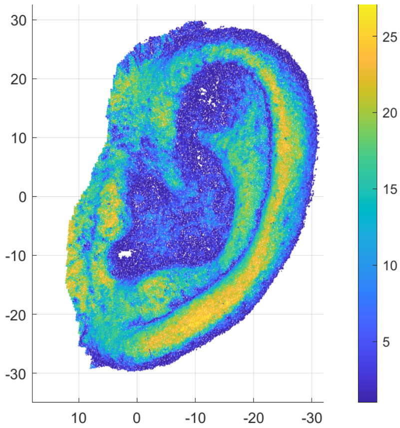


Figure 1: Photogrammetric reconstruction of a pinna, showing confidence levels for each point. The colourbar denotes the confidence level, i.e., the number of depth maps of image pairs which where used to reconstruct the respective point in 3D.

corresponding point in the point cloud. We utilize this additional value for each point to introduce a threshold to the registration performance. Figure 1 show the spatial distribution of these confidence values for one PR, and the colourbar denotes the number of depth maps being used for each reconstructed point. Note that the confidence levels in this figure are displayed as absolute values, and for the remainder of the investigation, we use normalised confidence levels (between 0 and 100%). For the PRs in this investigation, 54, 77, and 125 photos were taken around a vertical hemisphere of the left ear of NH5, details of which are described in [4], yielding confidence levels of up to 28, 57, and 74, respectively. These confidence levels have been normalised within each PR set, i.e., the maximum confidence of 100% corresponds to the respective maximum number of depth maps used.

2.2 Non-rigid registration

Non-rigid registration (NRR) algorithms describe the process of registering a template to a target point cloud - in our case, pinna point clouds. Typically, the template is treated as the means of a Gaussian mixture model (GMM), and the target is treated as observations from said model. Through sophisticated constraints on the movement of the points, one point cloud can be registered to the other. As a template, we used an undistorted, high-quality point cloud from a CT scan of a listener-specific pinna cast of NH5 [9]. As a target, we used three PR point clouds of the *same* listener containing the additional information of confidence levels. We applied the NRR algorithm geodesic-based Bayesian coherent point drift (GBCPD), because it protects the relationship between the points, i.e., takes the surface of the template pinna into account [10].

3. RESULTS AND DISCUSSION

Table 1 shows the point cloud size for three PRs and the change when considering points above various confidence level thresholds, and Fig. 2 shows three thresholded point clouds. Confidence levels below 40% contain more depth information than confidence levels above 40%, which resemble almost a 2D representation.

Table 1: Point cloud size change when filtering by confidence level thresholds.

	PR1	PR2	PR3
reference (all)	1,075,135	860,035	833,082
$\geq$ 25%	354,484	326,452	245,212
$\geq$ 40%	238,134	206,405	155,172
$\geq$ 50%	193,340	143,898	100,077
$\geq$ 60%	113,334	77,863	52,992
$\geq$ 75%	28,752	21,925	6,736

Figures 3 and 4 show results of GBCPD applied to the respective confidence-level thresholded point clouds from two angles. Because of the less depth detail for confidence levels above 40%, the geometry seems to "flatten" when being registered.

We expected the algorithm to move the template towards the target with its global pinna shape intact. This is only true for no thresholding (Fig. 2). Already with a confidence thresholding of 25%, the registered template does not represent the pinna geometry sufficiently accurate for plausible HRTF calculation. Especially for regions with large spatial variations, i.e., the helix rim, the crus inferius anthelicis and the crus superius anthelicis, the pinna seems to interpolate a flat surface over regions of points below the threshold.

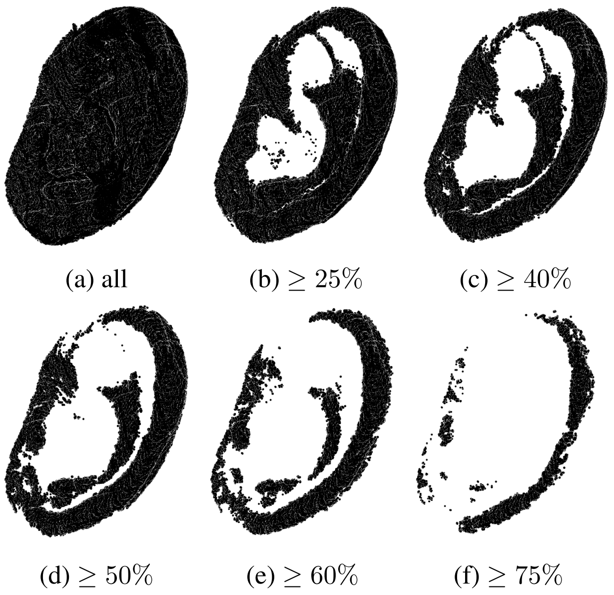


Figure 2: PR when filtering by relative confidence thresholds for above 0 (all points), 25, 40, 50, 60, and 75%, respectively.

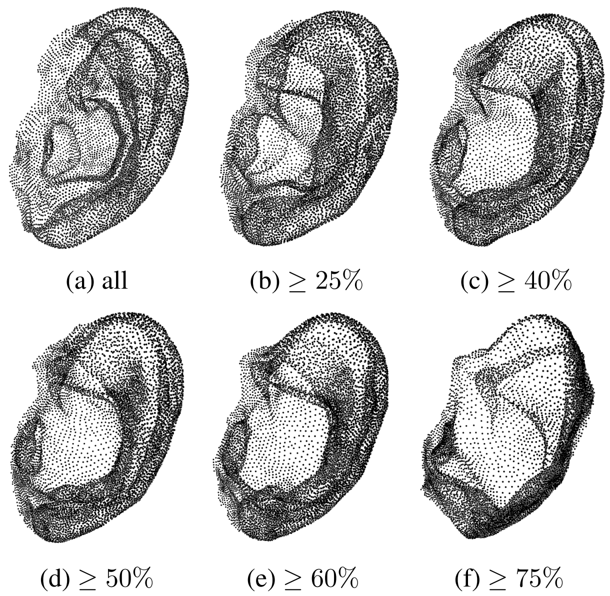


Figure 3: GBCPD-based registration of NH5 to PR1 when filtering by relative confidence level thresholds, "front" view.

4. CONCLUSIONS

In this investigation, we applied an NRR algorithm to a high-quality template point cloud in order to register it to three target point clouds acquired with PR. We were able to show that, above a confidence level threshold of 40%, the quality of the resulting point clouds decreases drastically. This is reflected not only in the interaural (front) view of the registered pinna, but especially when analysing the overall shape. For future applications, however, the confidence levels can be helpful to NRR algorithms, but they have to be incorporated *into* the registration algorithm.

The decrease in performance of the NRR algorithm with highly reduced point clouds might be due to the missing 3D shape information. This suggests it may be useful to interpolate missing points and weigh them with a low confidence level, such that the shape to register still *looks* like a pinna, but the algorithm has looser boundary conditions for points with low confidence levels. In the future, confidence levels (or other potential information that the PR yields) can be utilised as weights in the registration process.

5. ACKNOWLEDGMENTS

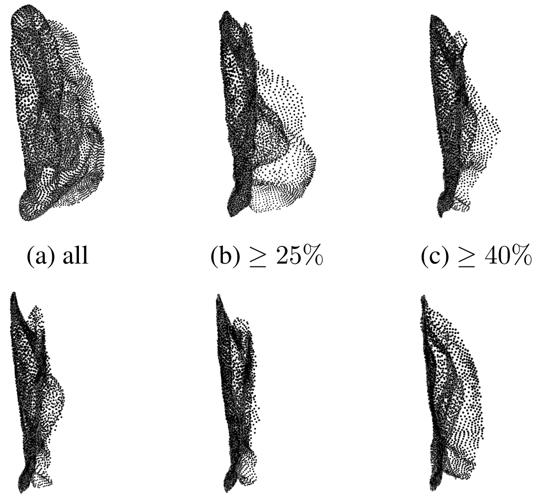
This work was supported by the Austrian Research Promotion Agency (FFG, grant no. 871263, project "soft-pinna") and the European Union's Horizon 2020 research and innovation funding programme (grant no. 101017743, project "SONICOM").

6. REFERENCES

[1] J. Blauert, *Spatial hearing. The Psychophysics of Human Sound Localization*. Cambridge, MA: The MIT Press, revised edition ed., 1997.

[2] H. Ziegelwanger, A. Reichinger, and P. Majdak, "Calculation of listener-specific head-related transfer functions: Effect of mesh quality," in *Proceedings of Meetings on Acoustics*, vol. 19, (Montreal, Canada), p. 050017, 2013.

[3] M. Dellepiane, N. Pietroni, N. Tsingos, M. Asselot, and R. Scopigno, "Reconstructing head models from photographs for individualized 3d-audio processing," in *Computer Graphics Forum*, vol. 27, pp. 1719–1727, Wiley Online Library, 2008.



(a) all (b) $\geq 25\%$ (c) $\geq 40\%$

(d) $\geq 50\%$ (e) $\geq 60\%$ (f) $\geq 75\%$

Figure 4: GBCPD-based registration of NH5 to PR1 when filtering by relative confidence level thresholds, "side" view.

[4] K. Pollack, F. Brinkmann, P. Majdak, and W. Kreuzer, "Von fotos zu personalisierter räumlicher audiowiedergabe," *e & i Elektrotechnik und Informationstechnik*, pp. 1–6, 2021.

[5] K. Pollack, P. Majdak, and H. Furtado, "Evaluation of pinna point cloud alignment by means of non-rigid registration algorithms," in *Audio Engineering Society Convention 150*, Audio Engineering Society, 2021.

[6] H. Takemoto, P. Mokhtari, H. Kato, R. Nishimura, and K. Iida, "Mechanism for generating peaks and notches of head-related transfer functions in the median plane," *The Journal of the Acoustical Society of America*, vol. 132, pp. 3832–41, Dec. 2012.

[7] K. Pollack and P. Majdak, "Importance of the acquisition accuracy of various pinna regions on the median-plane sound-localisation performance," in *Audio Engineering Society Convention 154*, pp. 1–7, Audio Engineering Society, 2023. 2.

[8] K. Pollack, P. Majdak, and H. Furtado, "Application of non-rigid registration to photogrammetrically reconstructed pinna point clouds for the calculation of personalised head-related transfer functions," in *Proceedings of 49th DAGA*, (Hamburg), pp. 1–4, 2023.

[9] A. Reichinger, P. Majdak, R. Sablatnig, and S. Maierhofer, "Evaluation of Methods for Optical 3-D Scanning of Human Pinnas," in *Proceedings of the 3D Vision Conference*, (Seattle, WA), pp. 390–397, 2013.

[10] O. Hirose, "Geodesic-based bayesian coherent point drift," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–18, 2022.