TRAIL TRUSTED AI LABS BY DIGITALWALLONIA / SPW-RECHERCHE

TRAIL Summer Workshop' 25
Project Proposal

Full Name of Team Leader	Ludovic Pirard l.pirard@imperial.ac.uk
Project Title	From Sparse Photogrammetry Data to High-Resolution 3D Head Mesh for Individual HRTF Synthesis
Profile of the Team Leader(s)	Ludovic Pirard joined Imperial College London in 2024 as a PhD student taking part in the Marie Sklodowska-Curie Doctoral Network 'Cochlear Implants and Spatial Hearing: Enabling access to the next dimension of hearing'. Ludovic worked at Multitel Belgium, a research center specialised in signal processing, artificial intelligence, embedded systems and optical fiber, where he contributed to high-resolution beamforming developments. Prior to this, he completed a MSc in Artificial Intelligence (AI) and Smart Communication at the University of Mons (UMons) in Belgium where his master's thesis focused on personalised HRTF using a Convolutional Neural Network. Ludovic is currently working on enhancing photogrammetry-reconstructed head meshes for individual HRTF synthesis.
Abstract	This project explores a machine learning pipeline that reconstructs high-resolution 3D head meshes from a small number of photogrammetry images. The goal is to develop a Neural Network (NN) capable of generating anatomically accurate 3D head meshes suitable for individual Head-Related Transfer Function (HRTF) synthesis. The input will be a minimal set of images (e.g., 3, 12, 36, or 72) per subject taken on the horizontal plane, and the output a high-resolution 3D head mesh. This task bridges accessibility and personalisation in spatial audio rendering, particularly where traditional 3D scanner solutions are impractical for end users. The project is aligned with the broader WAL4XR initiative, aiming to enable fast, secure, and ethical creation of multisensory immersive environments.
Project Objectives	During the workshop, the aim is to design and train a deep learning model capable of reconstructing high-resolution 3D head meshes from limited photogrammetry inputs. The dataset will consist of image batches of varying sizes (3, 12, 36, and 72 views) from the SONICOM HRTF dataset, along with corresponding high-resolution 3D scans as ground truth for supervised learning. The project has been explored by generating 3D head meshes utilising the available 72 images per subject from a subset of 250 subjects evaluating diverse commercial solutions. The obtained meshes and the synthetic HRTF computed are currently under evaluation. The primary objective is to reduce the number of images utilised for facilitating public access.

	By the conclusion of the workshop, we anticipate achieving the following outcomes: - (1) define the best-performing neural network architecture for mesh prediction from sparse photogrammetry input, - (2) evaluate mesh quality against ground truth data using geometric similarity metrics (e.g., Chamfer, Hausdorff distances), - (3) evaluate model robustness and computational efficiency across varying numbers of input views, - (4) demonstrate feasibility for wide public access to individual HRTF numerical synthesis. These outcomes will support the preparation of a journal publication documenting the technical approach, experimental results, and broader implications for the field.
Project Dataset	The **SONICOM Head-Related Transfer Function (HRTF) Dataset** [1,2] is a large-scale, high-resolution dataset developed as part of the SONICOM project, which is funded by the European Union's Horizon 2020 programme. Created by the Audio Experience Design group at Imperial College London, the dataset is intended to advance research in spatial audio, immersive sound technologies, and personalised binaural rendering. It includes acoustic and anatomical data from 300 human subjects, with plans for further expansion. The dataset contains several components. Firstly, it provides Head-Related Transfer Functions (HRTFs) sampled at 96 kHz and 24-bit resolution, with additional versions available at 44.1 kHz and 48 kHz. These HRTFs are provided in the widely used SOFA (Spatially Oriented Format for Acoustics) format, as well as in .3dti format for compatibility with the 3DTune-In Toolkit. Anatomical data are also provided, including high-resolution 3D scans of the participants' heads, ears, and torsos. These are supplied as .stl files, both with and without watertight post-processing. The dataset further includes high-resolution RGB depth images captured every 5 degrees of horizontal rotation around each subject. These images are in .tif and .HEIC format, accompanied by gravity data in .txt files. The dataset is available under the MIT License, allowing for open and flexible use in research and development with appropriate attribution. It can be accessed and downloaded through the following link: **https://transfer.ic.ac.uk:9090/#/2022_SONICOM-HRTF-DATASET/** For a detailed technical description and background, consult the following publication: Engel, I., Daugintis, R., Vicente, T., Hogg, A. O. T., Pauwels, J., Tournier, A. J., & Picinali, L. (2023). *The SONICOM HRTF Dataset*. *Journal of the Audio Engineering Society*, 71(5), 241–253. AES E-Library Link

Background Information	Immersive technologies have come to the forefront of spatial audio research. Head-Related Transfer Functions (HRTFs) are crucial for personalised spatial audio rendering in applications such as virtual reality, augmented reality and hearing aid development. Traditional HRTFs acquisition methods rely on specialised equipment and acoustic expertise, posing accessibility challenges. Alternatively, high-resolution 3D modelling offers a pathway to numerically synthesise HRTFs using Boundary Elements Methods and others. However, the high cost and limited availability of advanced 3D scanners restrict their applicability. Photogrammetry can be seen as an alternative; it is more accessible as well as affordable, and can be performed employing consumer equipment, e.g., a smartphone or a digital camera. Photogrammetry has been explored as a solution to obtain HRTFs via mesh reconstruction and HRTFs synthesis in previous works [3,4,5]. The lack of resolution and ear details in the mesh reconstruction has limited the use of photogrammetry for personal HRTFs synthesis [6,7]. Accurate modelling of the outer ear is especially critical, as small geometric variations in the pinnae and concha influence the individual HRTF [7]. Recent developments in neural surface reconstruction provide an opportunity to enhance photogrammetry-based head models. This project builds on state-of-the-art research in geometric deep learning (e.g., MeshCNN [8], Pixel2Mesh [9], Neural Subdivision [10]) and 3D reconstruction from sparse views, to produce anatomically accurate meshes.
Bibliographic References	**[1]** Engel, I., Daugintis, R., Vicente, T., Hogg, A. O. T., Pauwels, J., Tournier, A. J., & Picinali, L. (2023). The SONICOM HRTF Dataset. Journal of the Audio Engineering Society, 71(5), 241–253. **[2]** K. C. Poole, J. Meyer, V. Martin, R. Daugintis, N. Marggraf-Turley, J.Webb, L. Pirard, N. La Magna, and L. Picinali, “The Extended SONICOM HRTF Dataset,” 2025. Forum Acusticum, Malaga, Spain. **[3]** M. Dellepiane, N. Pietroni, N. Tsingos, M. Asselot, and R. Scopigno, “Reconstructing head models from photographs for individualized 3D-audio processing,” Comput. Graph. Forum, vol. 27, pp. 1719–1727, Oct. 2008. **[4]** Meshram, A., Mehra, R., Yang, H., Dunn, E., Franm, J.-M., and Manocha, D. P-HRTF: Efficient personalized HRTF computation for high-fidelity spatial sound. In 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (Sept. 2014), pp. 53–61. **[5]** Pollack, K., Kreuzer, W., and Majdak, P. Perspective Chapter: Modern Acquisition of Personalised Head-Related Transfer Functions An Overview. In Advances in Fundamental and Applied Research on Spatial Audio, B. F. G. Katz and P. Majdak, Eds. IntechOpen, Rijeka, 2022. Section: 2.

	[6] 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. sl," Hamburg, 2023. **[7]** H. Ziegelwanger, A. Reichinger, and P. Majdak, "Calculation of listener-specific head-related transfer functions: Effect of mesh quality," Proceedings of Meetings on Acoustics, vol. 19, p. 050017, May 2013. **[8]** R. Hanocka, A. Hertz, N. Fish, R. Giryes, S. Fleishman, and D. Cohen-Or, "MeshCNN: a network with an edge," ACM Transactions on Graphics, vol. 38, pp. 1–12, July 2019. **[9]** Wang, N., Zhang, Y., Li, Z., Fu, Y., Liu, W., and Jiang, Y.-G. Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images. **[10]** H.-T. Liu, V. Kim, S. Chaudhuri, N. Aigerman, and A. Jacobson, "Neural subdivision," ACM Transactions on Graphics, vol. 39, July 2020.
Detailed Work Plan	**Step 1** – Dataset Setup - Split data into training/validation/testing sets using 300 SONICOM subjects - Prepare photogrammetry inputs in sets of 3, 12, 36, and 72 images - Pre-align and normalise meshes (scale, rotation, centering) **Step 2** – Model Architecture and Training - Design/implement architecture (e.g., MeshCNN, Graph Convolutional Network) - Choose/design an adequate loss function - Train using supervised learning on high-res meshes as ground truth - Explore fine-tuning vs. end-to-end learning from images to mesh **Step 3** – Evaluation and Iteration - Compare predicted meshes with ground truth using geometric metrics (Chamfer, Hausdorff) - Perform qualitative evaluation (e.g., ear detail preservation) - Compare model performances between different images sets sizes **Team members:** Ludovic Pirard, Alexandre Philippon and five additional researchers **Tasks Distribution:** - Data pre-processing and pipeline integration – Ludovic and one additional researcher - Neural network design and training – Alexandre and two additional researchers - Evaluation metrics and validation – Two researchers - Documentation and final presentation – All team members

	Final Deliverables - Trained model and inference script - Technical report on the mesh evaluation method and process - Slides and poster for workshop presentation - GitHub repository with reproducible pipeline - Foundations for a future paper publication
Other Remarks	This collaborative project is initiated between two universities, comprising two PhD students who will be part of the team. Ludovic Pirard, a PhD student from Imperial College London, supervised by Prof. Lorenzo Picinali, and Alexandre Philippon, a PhD student from the University of Mons, supervised by Prof. Thierry Dutoit. Alexandre Philippon's PhD funding is derived from the WAL4XR project: https://www.xrwallonia.be/

Photogrammetry-Reconstruction example for individual HRTF synthesis