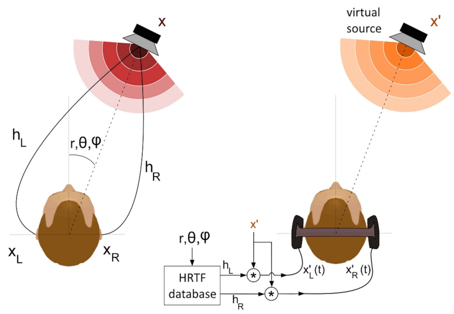
Gian
Danuser

Michael
Eugster

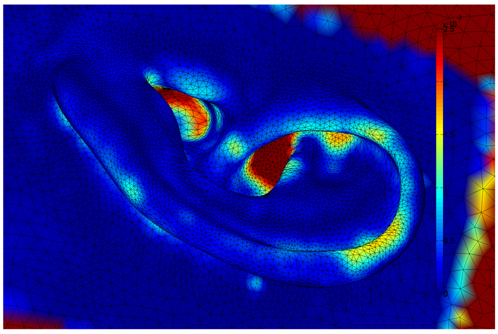

| | |
|---|---|
| Graduate Candidates | Gian Danuser, Michael Eugster |
| Examiner | Prof. Dr. Guido Schuster |
| Co-Examiner | ... ... |
| Subject Area | Sensor, Actuator and Communication Systems |
| Project Partner | Phonak |


# 3D Ear Reconstruction from a Series of Photographs

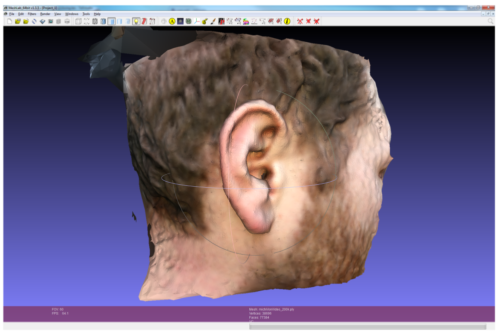
## Term Thesis 2




Natural binaural 3D hearing compared with the virtual 3D hearing with HRTF and headphones (source: PhD thesis Pierre Guillon)



Quality verification between the reconstructed and the reference ear (dark blue: 0mm distance to reference, dark red: >2.5mm distance to reference)



3D model reconstructed out of a movie sequence taken by a smartphone camera

Introduction: Humans can localise a sound source in the 3D space with binaural hearing. Therefore the brain uses the time difference between the ears and the transfer function of the head, pinna and torso. This transfer function is called head-related transfer function (HRTF).The HRTF can be used to generate a virtual auditory space via headphones and could be applied in virtual reality applications, video games, hearing research, etc.The measurement of the individual HRTF is expensive and time consuming. Another approach is to calculate the HRTF, based on a 3D model. For this an accurate model of the ear pinna is needed. The main aim of this project is to generate such a 3D model of the listener's ear out of a movie sequence or a few pictures taken by a smartphone camera without using structured light, landmarks, etc.

Approach/Technologies: In a first step, it was evaluated how the images need to be taken and prepared for a good further processing. After the image acquisition, so called scale-invariant feature transformation (SIFT) is applied to extract key points out of every image. With a nearest neighbour search key points are matched between the series of photographs. These matches are used in the structure-from-motion (SFM) method to generate a sparse 3D point cloud. The next step is to find a dense and smooth surface. This is done with a multi-view reconstruction software for weakly supported surfaces, which is based on visual hull.To verify the quality of the used pipeline, the whole reconstruction process is carried out with a sequence of photographs from a reference ear. The benefit of this approach is that the reconstructed 3D model can be compared to the original one.

Result: The used pipeline produces acceptable 3D models out of short movie sequences. With the reference ear, with almost ideal conditions, the basic shape of the created 3D models are virtually identical. Some parts of the pinna are error prone, due to missing SIFT points. Even in terms of real conditions the algorithm creates promising reconstructions. For a good reconstruction images with good quality, enough sharpness on the ear and high resolution are necessary. Additionally a uniform illumination is essential. In order to achieve the required quality, it's important to make proper adjustments of the degrees of freedom. A crucial factor is the maximal number of allowed vertices, otherwise the reconstruction algorithm may fail.