AUTOMATED EARDRUM REGISTRATION FROM LIGHT-FIELD DATA

Sofia Karygianni[2], *Manuel Martinello*[1], *Leonidas Spinoulas*[1], *Pascal Frossard*[2], *Ivana Tošić*[1]

[1]Ricoh Innovations Corp. (USA), [2]EPFL (Switzerland)

ABSTRACT

The performance of automated classification algorithms for medical images needs to be very high and especially robust in order to be adopted into healthcare. Most of the time the main challenge is *unregistered* data, since it is usually captured: 1) from different patients, 2) with different devices, and 3) at different time. Registration and normalization of the captured data is a necessary condition for success.

In this paper we present for the first time an automated method to register eardrums from light-field data. This procedure uses the shape information captured by a light-field otoscope and compensates for the natural tilt of the eardrum, its size, and the camera viewpoint. Results on clinical data show that the proposed algorithm is robust and works well for different types of ear conditions.

Index Terms— medical imaging, registration, 3D data.

1. INTRODUCTION

Automated classification algorithms using medical images typically require a registration step or an object detection step before passing the information to the classifier. In this step, a given object is localized in a given image, namely its position is expressed in pixel coordinates. This allows for the description of the object to be constructed relative to its position and makes the classifier invariant to the object translation. Moreover, an object in the image can have different rotations from one image to another.

In the case of middle ear images, the eardrum (or tympanic membrane, TM) can be positioned and rotated differently, depending on how the otoscope was held during the image acquisition. Examples of images of different TMs are shown in Figure 1. These examples are of different ear conditions (both healthy and unhealthy): notice that the position and the orientation of the TM are different in each image, as well as the tilt of the eardrum within the ear canal.

In the past, a work on automated diagnosis of otitis media [1] has tried to register this type of images based on appearance (colors and specularity). However, with the recent introduction of the Light Field Otoscope (LFO) - a device that enables non-invasive three-dimensional (3D) imaging of the middle ear [2, 3] - we have now access to the 3D shape information of eardrums, which can make the registration more robust and automated; this could be a key element to improve the performance of automated diagnosis of otitis media [4].

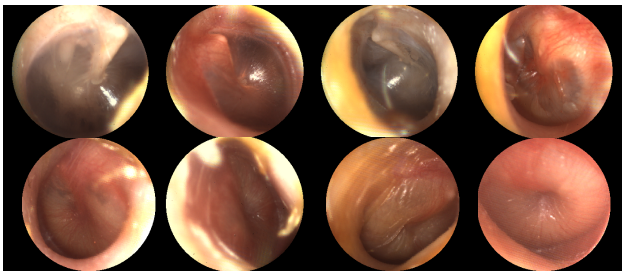


Fig. 1: Examples of 2D eardrum images prior to any registration.

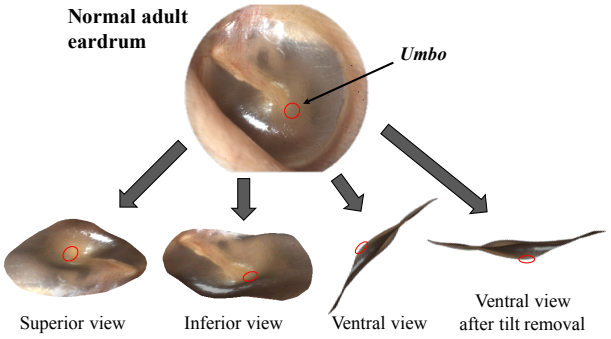


Fig. 2: Example of a normal eardrum, along with its 3D reconstruction. The ventral view shows that the TM is tilted within the ear canal; the *umbo* (red circle), which we designate as the center of the eardrum, indicates the lowest point on the TM (after the tilt removal).

In this work we present for the first time an automated algorithm for registration of 3D eardrums obtained from light-field data. The registration is designed to provide the center of the TM, together with its rotation and scaling information.

1.1. Prior Work

The 3D shape of the eardrum is initially given as an unstructured point cloud, where rigid registration approaches cannot be applied [5]. Some methods, based on the Iterative Closest Points (ICP) algorithm, rely on an overlap (or reliable common features) between point clouds [6, 7, 8]. Others (e.g., in the field of 3D object recognition) create local descriptors on points of interest to be then matched to find correspondences between different object instances [9, 10, 11]. In our case, features are sparse and they appear very different in eardrums with different medical conditions: this lack of common features may severely deteriorate the accuracy of the registration [12]. In addition, we have some partial stretching or shrinking of the object.

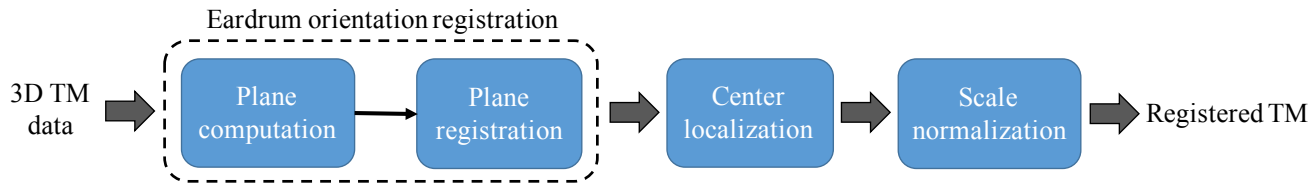


Fig. 3: Block diagram of the registration algorithm.

In healthcare, non-rigid registration algorithms have been used to accurately aligned a soft-body model with a set of 3D measurements (e.g., cancer-tissue detections, hole detection, and artefact recognition) [13, 14].

2. DATA REPRESENTATION

The proposed eardrum registration method takes a 3D mesh $M = \{V, O\}$ of the eardrum as input, where V is the point cloud and O is the mesh connectivity, obtained using Delaunay triangulation. In this paper, all the meshes have been obtained using the procedure described in [15] on eardrum images captured by the LFO [2]. Alternatively, one could also create the mesh by computing the Lisad I [15] and/or the Lisad Hessian [16] keypoints from the light field data.

3. OVERVIEW OF THE ALGORITHM

We present for the first time an algorithm for eardrum registration, using information from its shape; this algorithm is composed of four stages (Fig. 3):

1. Plane computation (or out-of-plane rotation), which automatically estimates and removes the tilt of the TM;
2. Plane registration (or in-plane rotation), which finds the zero-orientation to correct for the camera viewpoints;
3. Center localization, which estimates the center of the eardrum;
4. Scale normalization, which corrects for any difference in size.

3.1. Plane Computation

As illustrated in Fig. 2, the eardrum is tilted towards the inside of the ear canal, where the top of the TM lies closest to the tip of the otoscope tip (see ventral view). In order to compare different eardrums, we need to estimate and remove the slope from the captured data. First, we fit a plane on the 3D depth data by minimizing the sum of the square distances of the mesh points to the proposed plane. The recovered plane passes from the mean of the data and its normal vector equals the least important principal component of the mesh points, i.e., the one corresponding to the smallest eigenvalue of their covariance matrix. To make sure that the plane has the correct orientation, we always select the normal vector that points towards positive depth values, i.e., towards the otoscope sensor. An example of tilt removal from the 3D TM data is shown in Fig. 4. These TM depth maps have been segmented to include only the TM and not the ear canal. Initially the 3D mesh representation is viewed from the otoscope's point of view (V_{LFO}); after tilt removal, the representation is relative to the estimated tilt plane (viewed from V_{plane}).

However, because we cannot control the camera viewpoint, different eardrums can still have different orientations.

3.2. Plane Registration

To define the zero orientation - on the tilt planes - for different eardrums, we have to find a common landmark in the TM anatomy and designate it as a zero orientation landmark. We choose this landmark to be the top of the TM in the ear canal (the closest part to the tip of the otoscope).

To find the zero orientation from the tilt plane we use the plane elevation vector. We assume that the plane elevation, defined as the projection of the z-axis on the tilt plane, is indicative of the top of the eardrum. Fig 5 reports some examples of depth maps of different TMs, before tilt removal; on each depth map we overlaid the estimated plane elevation vector. Notice how these vectors point to the area of the highest depth (warmest color), i.e., to the top of the TM.

In our coordinate system, the plane elevation vector defines the direction of the y-axis on the plane. Then, given the directions of the y-axis and the normal vector of the plane, we define the x-axis based on the right-hand rule for Cartesian systems.

Since the TMs of the right and left eardrums are mirror-symmetric, we choose to flip left-to-right the input images of the right eardrum such that they are registered correctly with the left eardrum images (alternatively, one can flip the left eardrum images to fit the right eardrum, without any loss). The eardrum depth representation, after the tilt removal and rotation to the zero orientation, is an updated mesh $M_R = \{V_R, O_R\}$.

3.3. Center Localization

Now that the orientation of the 3D TM has been registered, we need to define - and successively estimate - its center. We designate as TM center the location of the *umbo*, which is the most depressed part of the tympanic membrane (see Figure 2). The overall procedure to localize the TM center is presented in the block diagram of Figure 6 and described in details in the next paragraphs.

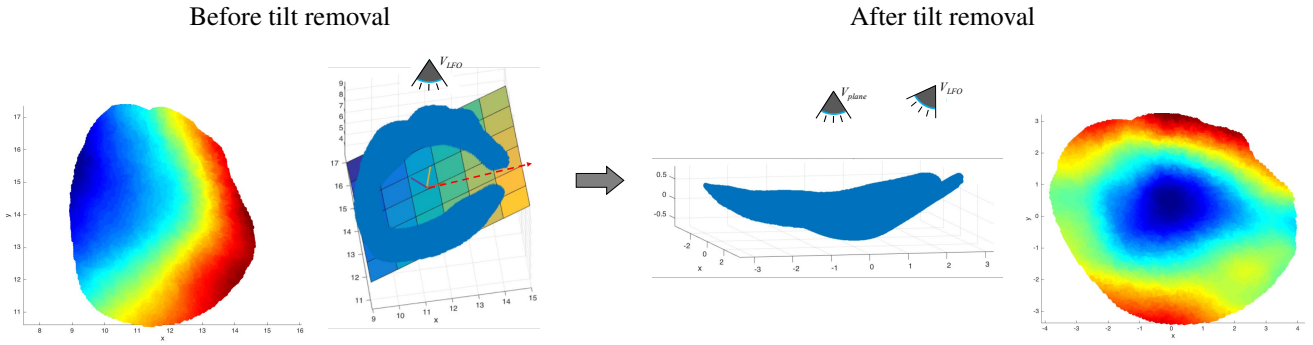


Fig. 4: Examples of eardrum tilt removal: (a) coded depth information showed in the 2D image domain (warmer colors indicate areas closer to the otoscope) relative to view point of the LFO (V_{LFO}); (b) estimated tilted plane together with the original 3D points of the eardrum (the chosen local basis as well as the plane normal vectors are shown with the red, yellow and purple vectors); (c) 3D mesh of the TM in the new coordinates system after tilt removal; (d) coded depth information relative to the tilt plane (viewed from V_{plane}).

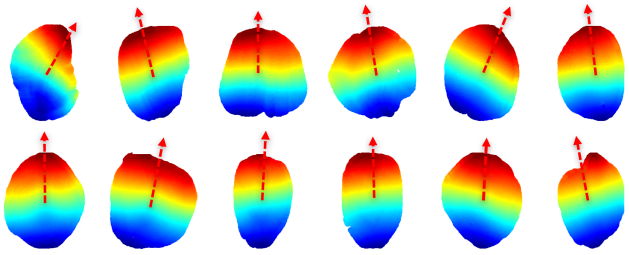


Fig. 5: Examples of plane elevation vectors overlaid on depth maps.

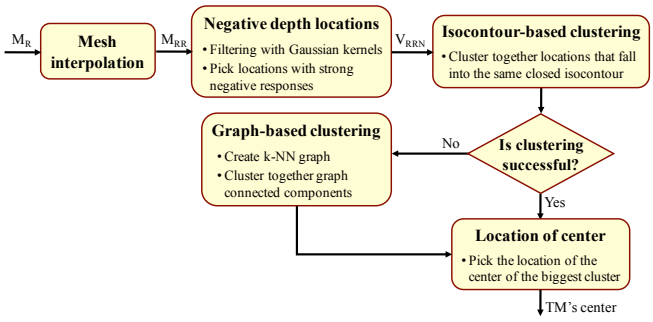


Fig. 6: Block diagram for center registration.

3.3.1. *Negative depth locations.*

To find the umbo, we consider the mesh M_R and locates the area with the lowest depth. In order to make the required computations more efficient, and without loss of generality, we first re-sample and interpolate the mesh M_R to get a new mesh $M_{RR} = \{V_{RR}, O_{RR}\}$ which has vertices on a regular grid in the first two spatial directions. Then, we locate areas of negative depth (below the estimated plane) on M_{RR} by filtering with Gaussian kernels and selecting the locations of the vertices with sufficiently negative responses, denoted as $V_{RRN} \subset V_{RR}$. For our experiments, we have used filters with $\sigma = 2$ and a threshold value of -0.1.

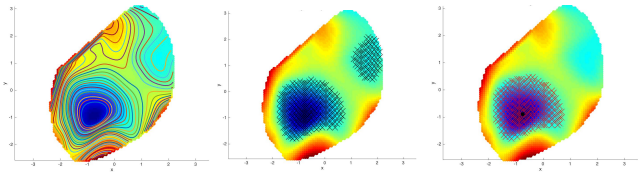
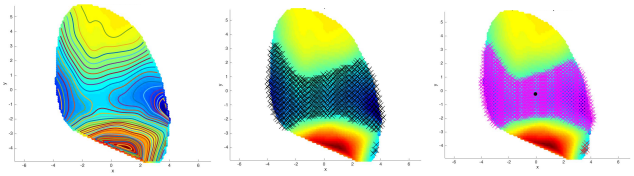


Fig. 7: Examples of center localization using only the isocontour-based clustering. (a) isocontours superimposed on a depth mesh. (b) Locations of negative depth as detected by our algorithm with Gaussian filtering; notice the clear overlap between the location of closed isocontours and the locations of negative depth. (c) The isocontour-based clustering is successful and reveals the central region of negative depth while the border region is discarded; the center of the biggest cluster is picked as the TM's center (black dot).

3.3.2. *Isocontour-based clustering.*

Using only the negative depth locations, we usually end up with many neighboring vertices that belong to the same area (Fig. 7(a-b)). To alleviate this issue, we have designed an additional clustering step for the vertices in V_{RRN}.

During the clustering, we take into account the fact that the umbo is usually located in an inner area of the eardrum, i.e., areas that are physically separated from the border of the eardrum; such areas often exhibit closed depth isocontours (contours of equal depth in the mesh after tilt removal), which are calculated using [17]. Therefore, we use the isocontours of the surface represented by the mesh M_{RR} in our clustering procedure. To be more specific, our scheme examines one by one all closed isocontours of the surface, starting from the ones with most negative depth, and creates a new cluster for all the vertices in V_{RRN} that lie inside each isocontour. In this way, all the vertices that belong to the same inner eardrum area will be clustered together based on the largest closed isocontour in this area, while the vertices belonging to outer areas of negative depth will remain unclustered. Such an example is shown in Fig. 7.



(a) (b) (c)

Fig. 8: Examples of center localization using the graph-based clustering. (a) isocontours superimposed on a depth mesh; (b) Locations of negative depth as detected by our algorithm with Gaussian filtering; there is not clear overlap between the location of closed isocontours and the locations of negative depth. (c) the isocontour based clustering is unsuccessful and the graph-based clustering is invoked, resulting in two connected components; the center of the biggest cluster is picked as the TM's center (black dot).

3.3.3. Graph-based clustering.

There are cases where, due to some irregularity in the shape of the eardrum, the algorithm fails to locate appropriate closed isocontours that overlap with the locations of the vertices in V_{RRN}. Such an example is shown in Figures 8. For these cases, we employ the k-nearest neighbor (k-NN) graph of the vertices in V_{RRN} to guide the clustering [18]. More specifically, we mark the connected components of the k-NN graph as areas of negative depth.

3.3.4. Location of center.

This step identifies the largest cluster and takes the center of its most negative isocontour as the center of the eardrum.

3.4. Scale normalization.

The physical size of the eardrum might be different in different patients. To account for these variations, we perform a scale normalization of the TM: each eardrum is scaled to match the size of a *reference eardrum*; The reference eardrum can be one of the following: (1) the largest eardrum considered, (2) the smallest eardrum considered; (3) an eardrum of a given size (width and height), (4) an average eardrum of a dataset. In the results of this paper we use the latest option, considering the average size of the eardrums in our dataset.

4. EXPERIMENTAL RESULTS

We tested our automated eardrum registration method on TMs with different ear conditions: acute otitis media (AOM), otitis media with effusion (OME), mild retraction, and no effusion (NOE, healthy condition).

In Fig. 9 we show how the estimated center locations compare with the ground truth centers for some eardrums: our algorithm usually places the center sufficiently close to the ground truth location. Finally, we show in Fig. 10 examples of registered TMs. Each TM has been reprojected on the tilt plane and oriented such that the zero orientation points upward. Notice that the malleus (the bone visible behind the

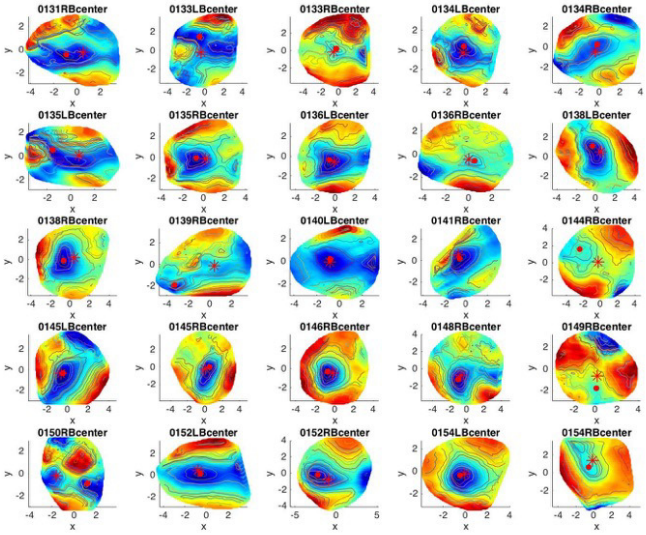


Fig. 9: Comparison between the estimated centers (red dot) and the ground truth (red star).

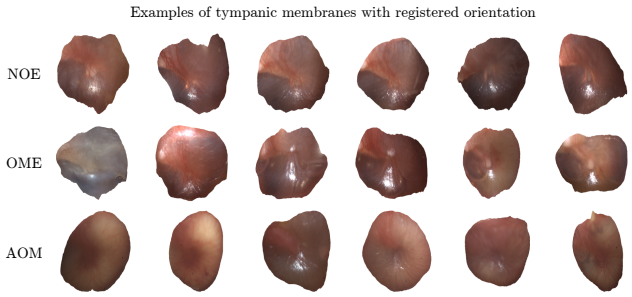


Fig. 10: Examples of registered eardrums with different conditions. All eardrums are displayed on the tilt plane and oriented such that the zero orientation points up.

TM in NOE and OME cases) is at the same position on all images, showing the success of our registration.

5. CONLUSIONS AND FUTURE WORK

The main value proposition for the LFO is that it could help in diagnosing otitis media conditions in an automatic way, helping to reduce the overdiagnosis of acute otitis media and overprescription of antibiotics. However, in order for the classification methods to work well on 3D data captured from different patients, all eardrums need to be registered and normalized. A registration that can easily be automated, as presented in this paper, is a key step for ensuring translation and rotation invariance of the otitis media classifier [19], which will hopefully be able to reach performance and robustness adequate to be adopted into healthcare.

6. REFERENCES

[1] A. Kuruvilla, N. Shaikh, A. Hoberman, and J. Kovačević, "Automated diagnosis of otitis media: vocabulary and grammar," *Journal of Biomedical Imaging*, vol. 2013, pp. 27, 2013.

[2] N. Bedard, I. Tošić, L. Meng, A. Hoberman, J. Kovačević, and K. Berkner, "In vivo middle ear imaging with a light field otoscope," *Optics in the Life Sciences, OSA Technical Digest*, 2015.

[3] N. Bedard, T.R. Shope, A. Hoberman, M.A. Haralam, N. Shaikh, J. Kovačević, N. Balram, and I. Tosić, "Light field otoscope design for 3d in vivo imaging of the middle ear," *Biomedical Optics Express*, vol. 8, no. 1, pp. 260–272, 2017.

[4] N. Shaikh, A. Hoberman, H. E. Rockette, and M. Kurs-Lasky, "Development of an algorithm for the diagnosis of otitis media," *Academic Pediatrics*, vol. 12, no. 3, pp. 214–218, 2012.

[5] B. Bellekens, V. Spruyt, R. Berkvens, and M. Weyn, "A survey of rigid 3d pointcloud registration algorithms," *International Conference on Ambient Computing, Applications, Services and Technologies*, pp. 8–13, 2014.

[6] P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[7] Y. Chen and G. Medioni, "Object modeling by registration of multiple range images," *Image and Vision Computing*, vol. 10, no. 3, pp. 145–155, 1992.

[8] T. Petricek and T. Svoboda, "Point cloud registration from local feature correspondences - evaluation on challenging datasets," *Plos One*, vol. 12, no. 11, pp. 1–16, 2017.

[9] A. Zaharescu, E. Boyer, K. Varanasi, and R. Horaud, "Surface feature detection and description with applications to mesh matching," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 373–380, 2009.

[10] F. Tombari, S. Salti, and L. Di Stefano, "Unique signatures of histograms for local surface description," *European Conference on Computer Vision*, pp. 356–369, 2010.

[11] H. Chen and B. Bhanu, "3d free-form object recognition in range images using local surface patches," *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1252–1262, 2007.

[12] F. Pomerleau, F. Colas, R. Siegwart, and S. Magnenat, "Comparing icp variants on real-world data sets," *Autonomous Robots*, vol. 34, no. 3, pp. 133–148, 2013.

[13] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[14] W.R. Crum, T. Hartkens, and D.L. Hill, "Non-rigid image registration: theory and practice," *British Journal of Radiology*, vol. 77, no. 2, pp. S140–S153, 2004.

[15] I. Tošić and K. Berkner, "Light field scale-depth space transform for dense depth estimation," *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 435–442, 2014.

[16] I. Tošić and K. Berkner, "3d keypoint detection by light field scale-depth space analysis," *IEEE International Conference on Image Processing*, pp. 1927–1931, 2014.

[17] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," *SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 163–169, 1987.

[18] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

[19] M. Martinello, L. Spinoulas, S. Karygianni, P. Frossard, M. A. Haralam, T. R. Shope, N. Shaikh, A. Hoberman, and I. Tošić, "Bulged eardrum detection from 3d data," *IEEE International Conference on Image Processing*, 2018.