Registration of *in-vitro* Micro CT Images For Construction of Statistical Shape Models And Atlas-based Segmentation of *in-vivo* CT Images of Cochlea

Kasper Korsholm Marstal, s112392

***Abstract*—Cochlear implants significantly improves quality of life for hearing impaired patients. Today, standard CT images of the head are used during surgical planning but the low-resolution and large amounts of noise limit the amount of available anatomical information. It is hypothesized that a model-based statistical understanding of the shape variability of the middle and inner ear among patients will be of great benefit during image analysis. In this report we register microCT images and obtain correspondences that are well suited for construction of statistical shape models of inner-ear anatomy. We also synthesize standard CT images by inducing a significant amount of noise in microCT images and obtain promising registration results. The results also provide a baseline for comparison with future model-based registration approaches that may further improve registration accuracy by incorporating prior knowledge from statistical shape models in the registration algorithm.**

***Index Terms*—cochlea implant, elastix, registration, statistical shape models**

I. Introduction

COCHLEAR implantation is a surgical procedure that aims to overcome hearing loss by direct stimulation of the spiral ganglion cells in the cochlea of the inner ear. Technological progress in this area has led to the development of inner ear implantable devices, which have proved to be of great benefit to patients suffering from moderate to severe hearing loss. The surgical scenario is very complex and requires high clinical expertise in order to efficiently assess the surgical site, the cochlea, localize nearby critical structures and optimize the position of the implantable device inside the cochlea. Because of high anatomical variability amongst patients, it is critical that individuals' cochlea shape is considered during the surgical planning process in order to reduce intra-cochlear damage and misplaced electrodes. Today, CT images are used during presurgical planning but the low-resolution of *in-vivo* CT images compared to the size of the cochlea limits the amount of available anatomical information. It is hypothesized that a model-based statistical understanding of the shape variability of the middle and inner ear among patients will be of great benefit during image analysis [1]. Such prior knowledge may guide atlas-based segmentation algorithms in difficult regions of the image where registration would otherwise fail.

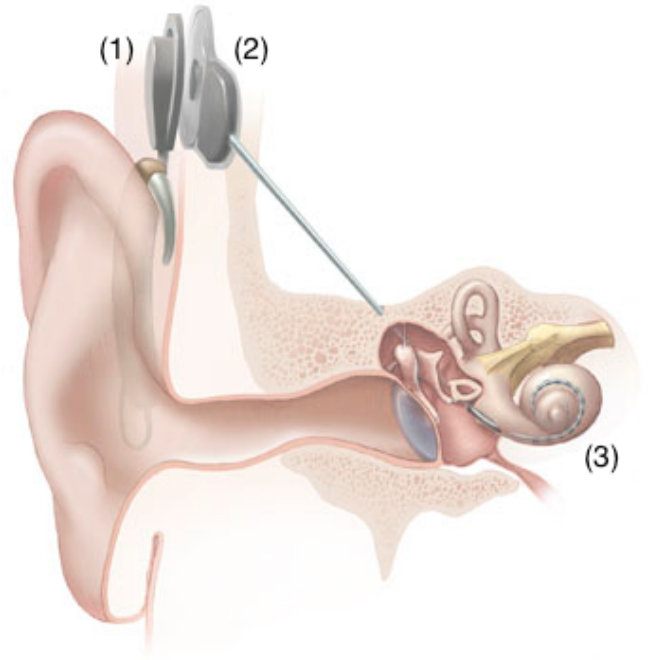


Fig. 1. Schematic depiction of a cochlear implant. (1) Microphone and transmitter, (2) receiver and stimulator inside skull and (3) intra-cochlea electrode array. Retrieved from \cite{fig:cochlearImplantSchematic}.

Obtaining accurate correspondences across a set

of images is a key algorithmic component in building such models. However, significant levels of noise and the complex structure of the inner ear canals pose challenges even for modern non-rigid multi-resolution registration methods. A potential solution is therefore to build the statistical models of inner ear anatomy on higher-resolution *in-vitro* micro CT images.

In this report, we seek to obtain the best possible correspondences between *in-vitro* high-resolution microCT images for subsequent construction of statistical shape models (the actual construction and evaluation thereof is beyond the scope of this report). We also register micro CT images to synthesized *in-vivo* CT images *without* the use of prior knowledge to provide a baseline for comparison with future model-based registration approaches, e.g. Active Appearance Models (AAMs), Active Shape Models (ASMs) or registration regularized by shape models.

A. Previous Work

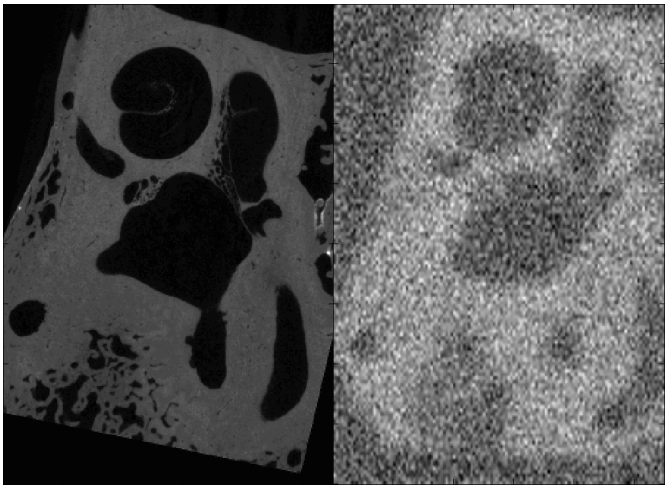
Image registration is a common but difficult problem in medical image analysis. The typical approach has been to manually segment groups of images. While manual labeling is a time-consuming but feasible solution for 2D models with a limited number of landmarks, it is highly impractical in the 3D domain. Not only is the required number of landmarks higher than in the 2D case, but it also becomes increasingly difficult to identify and pinpoint corresponding points, even for experts [2]. Manual labeling inevitably introduces bias in landmark selection and resulting correspondences, which may lead to weak models [3]. In the past decade a significant amount of research has focused on methods for automatically aligning images to reduce time-consuming and error prone human interaction [4].

Early registration methods assumed that between image acquisitions, the anatomical and pathological structures of interest do not deform or distort, e.g. [5]. This rigid body assumption simplifies the registration procedure but these techniques have limited applicability to this work due to the non-linear anatomical variability of cochlea.

Non-rigid registration methods are capable of aligning images where correspondence cannot be achieved without localized deformations. They are most popular in the literature, generally fully automatic and yield good results in controlled environments [6, 7]. Pioneered by Rueckert et al [8], these methods commonly consist of multi-resolution optimization with free-form deformations (FFDs) based on multi-level B-splines. These methods are able to register most well-posed datasets. Landmark-based methods try to mitigate problems in difficult cases by naturally incorporating expert knowledge in the registration scheme. A drawback of these methods is that landmark extraction is prone to error. In this work, it was investigated whether a landmark-based approach could aid our registration algorithm but the approach was abandoned due to a significant amount of manual intervention and the preference for a fully automatic method.

Groupwise registration methods try to mitigate uncertainties associated with any one image by simultaneously registering all images in a population. A common approach is to calculate the group mean image and then register all images towards this image, see e.g. [8-16]. In this work, it was investigated whether a groupwise approach was able to register the micro CT images but performance was undermined presumably by high anatomical variability and, consequently, a fuzzy mean image in the beginning of the registration procedure. Optimization was further challenged by the high-dimensional parameter search space associated with groupwise registration.

In this work, we therefore settled on a state-of-the-art pair-wise multi-resolution non-rigid registration approach. The goal is to quantify the level of error introduced when registering (without the use of prior knowledge) micro CT images and inspect if obtained correspondences are good enough construction of statistical shape models A second goal is to register microCT images to synthesized standard CT images and probe whether our method is capable of atlas-based segmentation of standard



(a) (b)

Fig. 2. High-resolution microCT images of inner-ear anatomy. (a) original subsampled image and (b) synthesized microCT image.

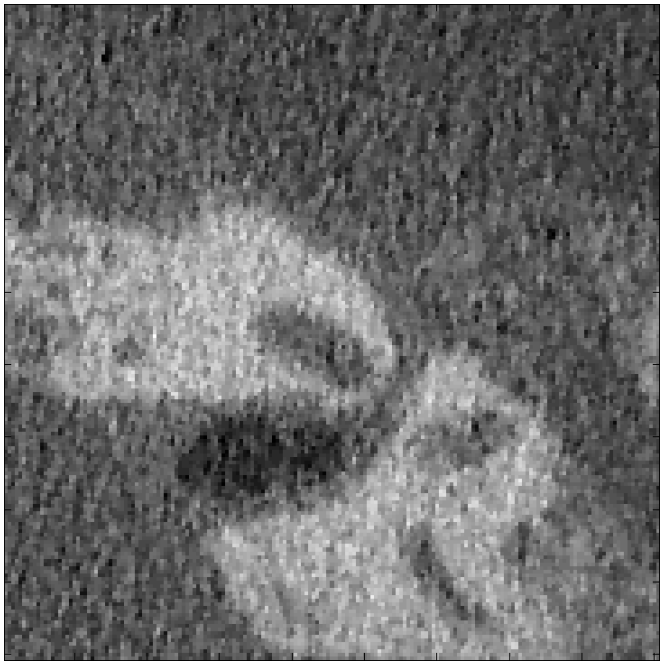


Fig. 3. Standard CT image of the head zoomed in on inner-ear anatomy.

CT images.

The remainder of the report is laid out as follows. In the following section the method, data and results are presented. In section III the results are discussed and the report is concluded in section IV.

II. Methods and Material

A. Image registration

Registration is the act of deforming a moving image $I_M(x)$ to fit a fixed image $I_F(x)$. This corresponds to finding a coordinate transformation $T(x)$ that makes $I_M(T(x))$ spatially aligned with $I_F(x)$. Mathematically, the registration problem is formulated as an optimization problem in which the cost function $C(\cdot)$ is minimized with respect to $T(x)$. A penalty term $P(T(x))$ is usually incorporated to regularize the search procedure. The optimizer can interface with the moving image by introducing a parameterization of the transformation. The optimization problem reads,

$$\hat{\mu} = \arg\min C\left(T_\mu; I_F; I_M; P_{T_\mu}\right) \quad (1)$$

where μ denotes that the transform has been parameterized. This vector contains the transform parameters.

In this work, the registration method is based on a 3D free-form B-spline deformation model, normalized correlation coefficient (NCC) similarity metric and stochastic Robbins-Monro gradient descent. The B-spline transform is widely used in non-rigid registration algorithms. The choice of NCC is motivated by its invariance with respect to linear transformation of intensity scales since the intensity scales in units of Hounsfield of any two CT images is related by a linear transformation. NCC was chosen over the more general Normalized Mutual Information (NMI) since NCC is more problem specific. The gradient descent algorithm is chosen due to the availability of a robust implementation with automatic hyper-parameter tuning. The method is embedded in a multi-resolution approach with a regularization term that promotes smooth deformation by penalizing transform bending energy (modeled as the bending energy of a thin metal sheet). The method is available in the public `elastix` registration library. The reader is referred to the accompanying paper [17] for details.

B. Data

14 microCT images are provided by the HEAR-EU project [1]. The temporal bone was dissected and imaged using a Scanco μCT100 system (Scanco Medical AG, Switzerland). Images have been subsampled to 400 x 275 x 200 voxels with an isotropic voxel spacing of 0.049 x 0.049 x 0.049

mm to reduce the computational burden. DTU Compute provides segmentations of the inner ear that facilitate evaluation of registration performance using the Dice Similarity Coefficient (DCE) and surface reconstructions that facilitates evaluation of registration performance in Hausdorff distance (mm). Figure 2 shows an example of a microCT image.

In this work, segmentations and surface reconstruction of inner-ear anatomy are not available for normal CT images. We have therefore opted to "synthesize" standard CT images by subsampling microCT images and induce Gaussian noise with a standard deviation of $\sigma = 1$ mm. The level of noise was chosen ad hoc to roughly match that of standard the CT image (see figure 3). This serves two purposes. Ground truth is readily available for performance assessment and initialization is a non-issue. A future application will have to sufficiently initialize images. This is a big problem on its own and beyond the scope of this work. The standard CT images have a voxel spacing of 0.15 x 0.15 x 0.15 mm so the microCT images are subsampled 1:3.

C. Registration

We perform two experiments.

1. Leave-one-out registration of microCT images to a reference microCT image. Using this approach all images are registered to all other images so the results will not be biased towards a chosen reference image. The experiment serves as a baseline for comparing of registration performance in DSC. A subset of images comes with subsampled surface models with point correspondence for which we report Hausdorff.
2. Registration of microCT images to synthesized CT images. This experiment measures the registration performance when registering microCT images to standard CT images. Again, a subset of the images comes with subsampled surface models with point correspondence for which we report Hausdorff distance.

TABLE I
RESULTS

EXPERIMENT	DSC	HD
Before registration	0.565±0.130	1.468±0.355 mm
μCT to μCT	0.931±0.024	1.194±0.169 mm
μCT to CT	0.894±0.027	1.219±0.211 mm

Results are reported in Dice Similarity Coefficient (DSC), Hausdorff Distance (HD). Note that one voxel is 0.049 mm^3 so a Hausdorff distance of 1.468 mm corresponds to a maximum distance of 29 voxels between two surfaces.

The results are shown in Table 1.

In the following section results are discussed.

III. DISCUSSION

Figure 4 shows a successful registration of a microCT image to another microCT image and the result is visually appealing. The semi-circular canals are well aligned and the method also captures the spiral structure of cochlea. The apexes of the cochleae are misaligned by a few voxels but this is probably due to the regularization term that impedes the transformation to a certain degree (however small it might be). The Hausdorff distances show are relative large gap (1.19 mm ~ 24 voxels) between some of the surfaces. This is probably due to the complex structure of cochlea. For all practical purposes, the significant increase in DSC from 0.565±0.130 to 0.931±0.024 for $(14-1)^2 = 169$ registrations validates this approach for constructing statistical shape models from microCT images. However, poor registrations that would lead to weak models should be discarded.

Figure 5 shows registration of the same two images but here the reference image has been synthesized to look like a regular CT image. The alignment is a little rough around the edges of the semi-circular canals and tips of cochleae. Considering the large amount of noise present in the reference CT image results are overall positive.

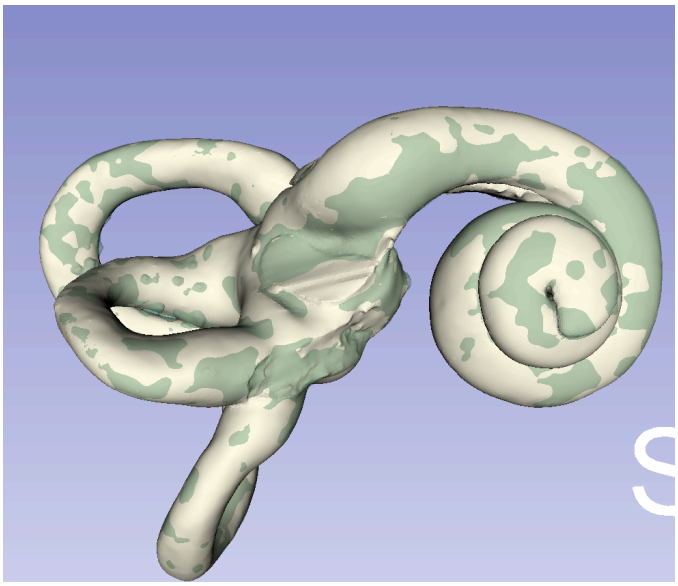


Fig. 4. Example of a successful registration of a microCT (transparent green) to microCT image (white).

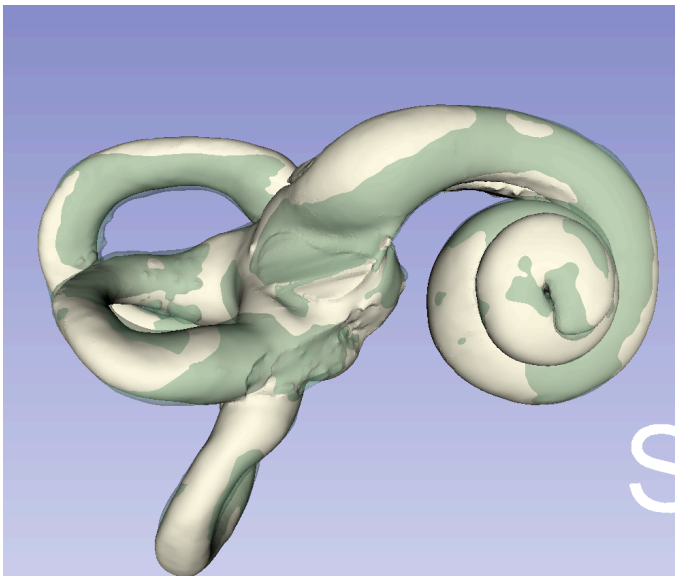


Fig. 5. Example of a successful registration of a microCT (transparent green) to synthesized CT image (white).

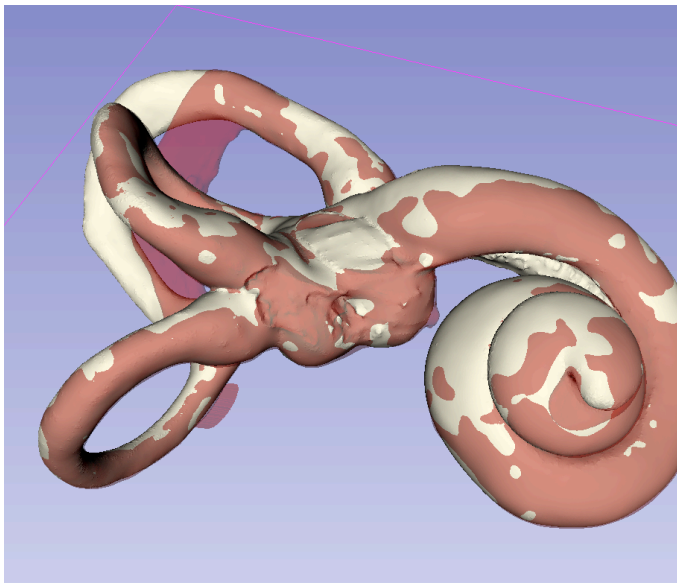


Fig. 6. Example of a bad registration of a microCT (transparent red) to microCT image (white).

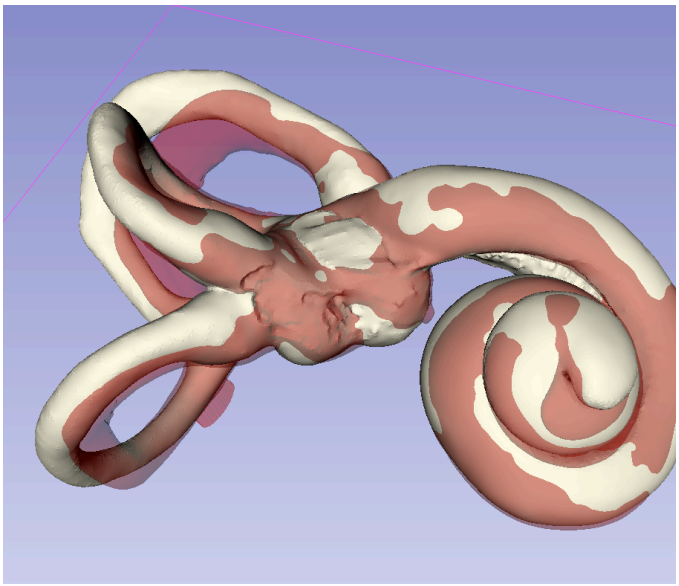


Fig. 7. Example of a bad registration of a microCT (transparent red) to a symthesized CT image (white).

With that in mind, figure 6 and 7 shows that the method is somewhat sensitive to initialization. Here, images were ill aligned prior to registration. This is reflected in a gross misalignment of the tips of cochleae and superior canals. The example shows that initialization is especially important when registering microCT images to standard CT images in all likelihood due to the complex structure of inner-ear anatomy and the abundance of noise in synthesized CT images confusing the search space. A model-based approach may potentially obtain a better result because the prior knowledge embedded in the search procedure may guide the optimizer past local minima in the search space.

Results show that registrations of microCT images are applicable for construction of statistical shape models. However, due to poor alignment in some cases we suggest future approaches should involve a bootstrap procedure where a model is built from good registrations and subsequently used to aid the registration of challenging images.

Results also show it is indeed possible to recover inner-ear anatomy of standard CT images using atlas-based segmentation. However, the approach is

sensitive to initialization and a full-automatic method will require a well-designed initialization algorithm.

The reader should keep in mind that the results obtained in this report are based on synthesized CT images and therefore only gives an indication of registration performance that cannot replace analysis of registration of microCT images to true CT images of the head.

IV. Conclusion

We registered microCT images and obtained promising results that are well suited for construction of statistical shape models. We also induced noise into microCT images to study registration of microCT images to CT images. Based upon these experiments we believe it is possible to recover inner-ear anatomy in true CT images from microCT images using atlas-based segmentation. However, results are sensitive to initialization which future approaches will have to deal with accordingly.

References

[1] M. A. G. Ballester, Hear-EU project description, http://www.hear-eu.eu

[2] Tobias Heinmann and Hans-Peter Meinzer. Statistical shape models for 3d medical image segmentation: A review. *Medical Image Analysis*, 13:543–563, 2009.

[3] R.H. Davies, C.J. Twining, T.F. Cootes, J.C. Waterton, and C.J. Taylor. A minimum description length approach to statistical shape modeling. *Medical Imaging, IEEE Transactions on*, 21(5):525–537, 2002.

[4] Pei Zhang, Steve A. Adeshina, and Timothy F. Cootes. Automatic learning sparse correspondences for initialising groupwise registration, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, volume 6362 of *Lecture Notes in Computer Science*, pages 635–642. Springer Berlin Heidelberg, 2010.

[5] Mark Jenkinson and Stephen Smith. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2):143 – 156, 2001.

[6] B. Fischer and J. Modersitzki. Combination of automatic non-rigid and landmark based registration: the best of both worlds. *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 5032 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1037–1048, May 2003.

[7] Fahmi Khalifa, GarthM. Beache, Georgy Gimel'farb, JasjitS. Suri, and AymanS. El-Baz. State-of-the-art medical image registration methodologies: A survey. *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, pages 235–280. Springer US, 2011.

[8] 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. *Medical Imaging, IEEE Transactions on*, 18(8):712 –721, aug. 1999.

[9] Michael Miller, Alain Trouve, and Laurent Younes. On the met- rics and euler-lagrange equations of computational anatomy. *Annual Review of Biomedical Engineering*, 4(1):375–405, 2002.

[10] K.K. Bhatia, J.V. Hajnal, B.K. Puri, A.D. Edwards, and D. Rueck- ert. Consistent groupwise non-rigid registration for atlas con- struction. In *Biomedical Imaging: Nano to Macro, 2004. IEEE Inter- national Symposium on*, pages 908 – 911 Vol. 1, april 2004.

[11] Lilla Zöllei, Erik Learned-Miller, Eric Grimson, and William Wells. Efficient population registration of 3d data. In Yanxi Liu, Tianzi Jiang, and Changshui Zhang, editors, *Computer Vision for Biomedical Image Applications*, volume 3765 of *Lecture Notes in Computer Science*, pages 291–301. Springer Berlin Heidelberg, 2005.

[12] T. F. Cootes, C. J. Twining, V. Petrovic, R. Schestowitz, and C. J. Taylor. Groupwise construction of appearance models using piece-wise affine deformations. In *in Proceedings of 16 th British Machine Vision Conference*, pages 879–888, 2005.

[13] S. Balci, P. Golland, and W. Wells. Non-rigid groupwise regis- tration using b-spline deformation model. 07 2007.

[14] S.A. Adeshina and T.F. Cootes. Constructing part-based models for groupwise registration. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pages 1073 –1076, april 2010.

[15] C.T.Metz,S.Klein,M.Schaap,T.vanWalsum,andW.J.Niessen. Nonrigid registration of dynamic medical imaging data using nd+t b-splines and a groupwise optimization approach. *Medical Image Analysis*, 15(2):238 – 249, 2011.

[16] K. Sidorov, D. Marshall, and S. Richmond. An efficient stochas- tic approach to groupwise non-rigid image registration. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 2208–2213, 2009.

[17] S. Klein, M. Staring, K. Murphy, M.A. Viergever, and J. Pluim. elastix: A toolbox for intensity-based medical image registration. *Medical Imaging, IEEE Transactions on*, 29(1):196 –205, jan. 2010.