

# Überatlas: Robust Speed-Up of Feature-Based Registration and Multi-Atlas Segmentation

Jennifer Alvé, Alexander Norlén, Olof Enqvist, and Fredrik Kahl

Department of Signals and Systems  
Chalmers University of Technology

**Abstract.** Registration is a key component in multi-atlas approaches to medical image segmentation. Current state of the art uses intensity-based registration methods, but such methods tend to be slow, which sets practical limitations on the size of the atlas set. In this paper, a novel feature-based registration method for affine registration is presented. The algorithm constructs an abstract representation of the entire atlas set, an überatlas, through clustering of features that are similar and detected consistently through the atlas set. This is done offline. At runtime only the feature clusters are matched to the target image, simultaneously yielding robust correspondences to all atlases in the atlas set from which the affine transformations can be estimated efficiently. The method is evaluated on 20 CT images of the heart and 30 MR images of the brain with corresponding gold standards. Our approach succeeds in producing better and more robust segmentation results compared to two baseline methods, one intensity-based and one feature-based, and significantly reduces the running times.

**Keywords:** feature-based registration, pericardium segmentation, brain segmentation, multi-atlas segmentation

## 1 Introduction

Segmentation is one of the most fundamental problems in medical image analysis and may be used in order to locate tumors, measure tissue volumes, for the study of anatomical structures, surgery planning, virtual surgery simulation, intra-surgery navigation etc. [22]. The value of automatic segmentation is huge, since manual delineation is time-consuming and sensitive to the skill of the expert. It is important that the segmentation algorithm is robust and fast in order to be useful in clinical care. Moreover, the segmentation algorithm should produce results comparable to those by a skilled expert.

One segmentation method that has become popular in recent years is multi-atlas segmentation [15, 12, 21, 5, 26] as it produces state-of-the-art results. Multi-atlas segmentation relies on a set of atlases (images with corresponding manual delineations), which are separately registered to an unlabeled target image. Labels are transferred from the atlas images to the target image and fused by a voting scheme. In contrast to segmentation based on shape models, for example,

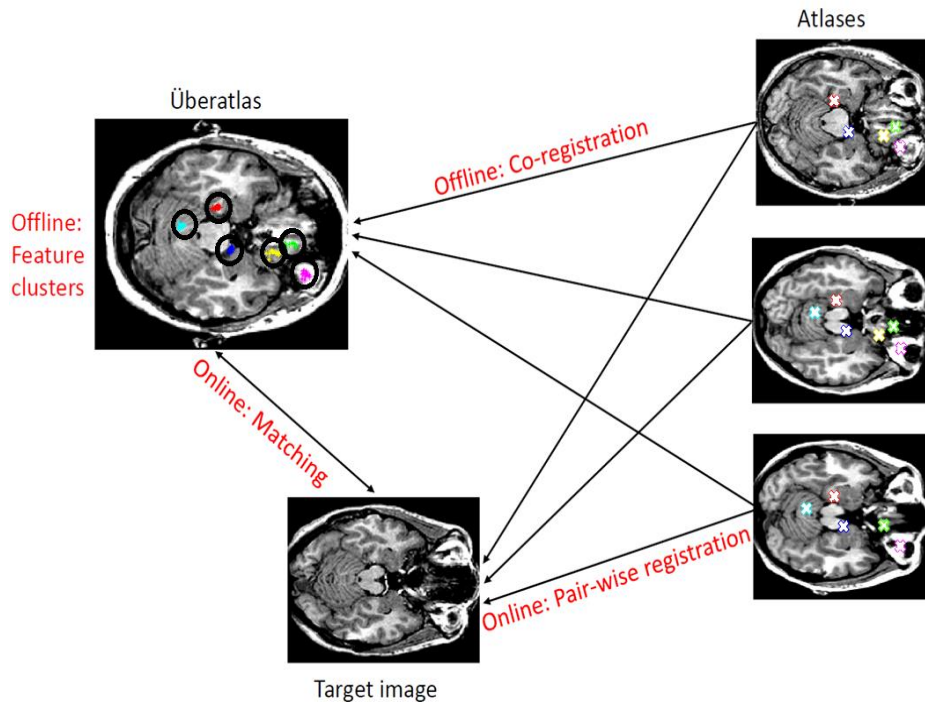


Fig. 1: A schematic drawing of the proposed framework. Offline: The atlases are co-registered to a reference atlas where the features are clustered both according to descriptor distance and spatial distance to make sure that the features in the cluster describe the same anatomical structure and to exclude outliers. Online: Robust matching between the target image and the überatlas is performed, directly gives feature correspondences between the target and all of the atlases. Using these correspondences all the atlases are robustly and efficiently registered to the target image.

active shapes [6], the solution is not constrained to be in the subspace spanned by the shape model. While being robust to non-satisfactory registrations multi-atlas segmentation has the disadvantage of being dependent on multiple image registrations which in general take a lot of time to compute and consequently limits the practical size of the atlas set.

In principle, there are two different approaches to image registration, feature-based and intensity-based registration, see the surveys [14, 25]. Intensity-based methods are capable of producing very accurate registrations, but are often slow and sensitive to initialization. On the other hand, feature-based methods are fast, but risk failing due to the difficulty in establishing correct point-to-point correspondences between the images.

In this paper, a novel feature-based registration method is presented that combines the information of the entire atlas set and efficiently finds robust transformations between a target image and *all the images* in the atlas set. The general idea is to cluster the feature points in the atlas set, both according to descriptor distances and according to spatial distances after a careful co-registration of the atlases, into what will be referred to as an *überatlas*. The construction of the *überatlas* is done offline. At runtime, one only needs to register the *überatlas* to the target image, and correspondences to all images in the atlas set are automatically obtained. As the *überatlas* only contains the best fitted feature points for matching from the co-registration phase, the approach has the potential not only to improve speed, but also to reduce the amount of classification errors and improve the segmentation. Figure 1 shows a schematic drawing of the proposed solution.

## 2 Related Work

The work that is most closely related to ours is the standard multi-atlas approach, where registrations are computed between all the atlases and the target image independently, see [15, 12]. Typically, the images are first registered using an affine transformation. If needed, this is followed by a refined nonrigid registration. In this paper, we focus on the first step of estimating an affine transformation. We experimentally compare to two such standard baselines, one using intensity-based registration and one using features. The intensity-based method performs image registration with the popular software package `Niftyreg`, meaning that an affine transformation is estimated via a block-matching strategy, proposed and implemented by Ourselin et al. [19, 20]. The feature-based method is based on standard RANSAC optimization [9] which is further described in the sections below.

Dey et al. [7] proposed to first co-register a set of atlases of the heart using a nonrigid intensity based method and then, at runtime, to only register one of the atlases to a target image and letting the other atlases vote indirectly. To some extent this is contrary to the idea of multi-atlas segmentation as it relies on obtaining a single very accurate registration at run time. Gill et al. [10] proposed to create a mean atlas including feature points of a set of atlases of the lung, and use this mean atlas to initialize an active shape model.

Hence, the concept of using a mean atlas is not new, but we are not aware of any work using an intermediate representation for efficiently estimating all the transformations between atlases and target images. Our main contributions are the development of such a representation, which we refer to as an *überatlas*, and to experimentally demonstrate that we obtain comparable accuracy with respect to the two baselines while significantly reducing runtimes.

### 3 Proposed Solution

In order to perform a multi-atlas segmentation, one needs to register the atlases to a target image as accurately as possible. As previously mentioned, a common way of doing this is to estimate an affine transformation followed by a nonrigid transformation. In this paper, we focus on the affine estimation step.

The underlying idea of our approach is to construct an intermediate representation of the atlas images which can be used to perform the most time-consuming part of the process, namely the robust matching, that is, obtaining correct feature correspondences between the atlases and the target images. The überatlas is constructed in such a way that if a feature correspondence between the überatlas and the target image has been identified, one can immediately derive in which atlas images this particular feature correspondence was originally present, see Figure 1. Once good correspondences are obtained to all the atlases, one can quickly compute an affine transformation for each atlas individually.

First, we will describe how to construct the überatlas from the atlas set (Section 3.1). Note that this is an offline process done only once, so speed is less important than for the online steps. Then, we will describe how to register the überatlas to a new, unlabeled target image (Section 3.2), which is a process done online.

#### 3.1 Überatlas Construction

The purpose of the überatlas is to gather information about corresponding feature points in all atlases, as well as an approximate way of describing all the feature points. To compute sparse features we use the code by Svärm et al. [24]. This code uses the method of SIFT for feature detection [17] and a descriptor similar to that of SURF proposed by Bay et al. [2].

Constructing an überatlas is done in two steps: co-registration and feature clustering.

*Co-registration of atlases.* In this step, a nonrigid transformation  $\hat{T}_{i,\hat{i}}$  is estimated between each atlas,  $i$ , and the *reference atlas*  $\hat{i}$ , that has been chosen at random. An initial estimate of  $\hat{T}_{i,\hat{i}}$ , was given by affine feature-based registration. We used truncated  $l_2$  as a loss function with a truncation threshold of 20 mm and optimized it using RANSAC [9] with one million iterations.

For the nonrigid registration, each atlas image,  $\mathcal{I}_i$ , is concatenated with its corresponding labeling,  $\mathcal{L}_i$  into a two-channel 3D image. The labeling is weighted with a factor  $\alpha_i$  to determine its impact on the solution. For our experiments  $\alpha$  was set to the intensity span of the image. Finally, the actual registration is performed with `NiftyReg` using normalized mutual information, due to its capability of dealing with multi-channel images. For all other parameters we used the default values.

*Feature clustering.* A feature point  $\mathbf{f}$  is uniquely determined by the index of the atlas from which it is extracted,  $i$ , the coordinates of the feature point in the coordinate frame of the atlas,  $\mathbf{x}$ , and the corresponding descriptor,  $\mathbf{d}$ , i.e.,  $\mathbf{f} = (i, \mathbf{x}, \mathbf{d})$ . Using the transformations,  $\hat{\mathbf{T}}_{i,\hat{i}}$ , obtained by co-registration, all feature points are transformed into a common coordinate system. The transformed coordinates are in other words the coordinates of a feature detected in any atlas, transformed accurately into the coordinate system of the reference atlas.

In order to construct the feature clusters,  $\mathcal{F} = \{\mathbf{f}_k : k \in \mathcal{K}\}$  where  $\mathcal{K}$  is the set of the feature indices in a cluster, agglomerative clustering is used [27]. A distance matrix based on the distances between all descriptors is constructed and sorted. When clustering  $M$  atlases, the  $3 \times M$  smallest distances to each of the feature descriptors are taken into account and the remaining distances are considered infinitely large. The set of clusters is initialized with one cluster for every feature point (a cluster with size 1) and the clustering proceeds by merging pairs of clusters.

For merging clusters, a variant of single linkage is used, i.e., the shortest distance between two clusters is considered. However, these clusters are only merged if no features in the two clusters come from the same image and if all features in the two clusters are spatially consistent. More precisely feature  $k$  from image  $i$  is spatially consistent with feature  $l$  from image  $j$  if their coordinates satisfy,

$$\left| \hat{\mathbf{T}}_{i,\hat{i}} \circ \mathbf{x}_k - \hat{\mathbf{T}}_{j,\hat{i}} \circ \mathbf{x}_l \right| < \epsilon_s. \quad (1)$$

In practice, the spatial threshold,  $\epsilon_s$ , was set to 10. Finally, we only keep clusters that contain at least 3 features.

These criteria make sure that the features in the clusters with high certainty describe the same anatomical feature. Further, it tells us that this anatomical feature is often detected by the feature detector meaning that we can expect to find the feature in a new image as well. Finally, it cleans the überatlas of features that are not found consistently between the atlases. These are features that most likely are outliers, noise, misplaced features or features describing uncertain anatomical regions.

An *überatlas* is a set of non-overlapping feature clusters with a corresponding descriptor, i.e.,

$$\{(\mathcal{F}_u, \hat{\mathbf{d}}_u) : u \in \mathcal{U}\} \quad (2)$$

where  $\mathcal{U}$  is the set of feature cluster indices and  $\hat{\mathbf{d}}_u$  is calculated as the mean of the feature descriptors included in the cluster  $\mathcal{F}_u$ .

### 3.2 Überatlas Registration

Registration using the überatlas is done in two steps. First the features of the target image are matched to the feature clusters in the überatlas in a robust manner. Then each atlas is registered efficiently to the target image using iteratively reweighted least squares (IRLS).

*Matching.* Given a new target image, feature points are detected and descriptors are computed as explained in Section 3.1. For each cluster the cluster descriptor is matched to its nearest neighbor of the feature descriptors in the target image. The distance between the descriptors is measured using the Euclidean norm.

A match between a feature  $\mathbf{f}_t$  in the target image and a cluster  $\mathcal{F}_u$  in the überatlas simultaneously establishes correspondences to all features present in the cluster and therefore also to the atlases from which these features were extracted.

In order to eliminate outliers, a restriction on the distance between the descriptors of the correspondences is introduced. A correspondence between  $\mathbf{f}_t$  and a feature in an atlas  $\mathbf{f}_s$  found through matching to the cluster  $\mathcal{F}_u$  is considered an inlier if the following criterion is fulfilled

$$\frac{\|\mathbf{d}_t - \mathbf{d}_s\|}{D_u} < \epsilon_d, \quad (3)$$

where  $\epsilon_d$  is a parameter and  $D_u$  is the maximum distance between the descriptors within the cluster  $\mathcal{F}_u$ , i.e.,

$$D_u = \max\{\|\mathbf{d}_i - \mathbf{d}_j\| : \mathbf{f}_i, \mathbf{f}_j \in \mathcal{F}_u\}. \quad (4)$$

This can be viewed as a generalization of Lowe’s ratio criterion for feature matching (see [17]), to the case of cluster matching. In practice, we used  $\epsilon_d = 1$ .

Finally, RANSAC is used in order to remove matches that are not consistent with an affine transformation between the überatlas and the target image. The coordinates of the clusters are set to the mean of the transformed feature coordinates. Also, the affine transformation is used as an initialization to the estimation of the affine transformation between the atlases and the target image, described below.

*Affine transformation.* As described in Section 1, the problem of using feature-based registration is the great amount of outliers produced. Even though the amount of outliers is reduced thanks to the construction of the überatlas, the problem is not entirely eliminated. Normally this is dealt with using RANSAC with a high number of iterations, but überatlas-based matching produces fewer outliers, so we can use a faster method for estimating the affine transformation, namely iteratively reweighted least squares (IRLS), see Chartrand and Yin [4].

Often truncated  $l_2$ -norm is viewed as an appropriate loss function in the presence of outliers, see Blake et al. [3]. However, we have found that using the truncated  $l_1$ -norm instead, reduces the sensitivity to the choice of outlier threshold and enables larger thresholds without losing robustness. The reason is that the  $l_2$ -norm is suitable mainly when the noise is Gaussian. This is a good assumption for the measurement errors of inliers but works poorly for the outliers. When increasing the value of the outlier threshold, we include more outliers making the the assumption of normally distributed residuals unsuitable. Truncated  $l_1$ -norm is less sensitive.

Given a set of  $N$  local correspondences, where the coordinates of the correspondences are given by  $\{(\mathbf{x}_n, \mathbf{y}_n) : n = 1, \dots, N\}$ , the affine transformation is estimated via IRLS. The following minimization problem is solved iteratively with the  $l_2$ -norm as a loss function,

$$(\hat{\mathbf{A}}_i, \hat{\mathbf{t}}_i) = \operatorname{argmin}_{\mathbf{A}, \mathbf{t}} \sum_{n=1}^N \omega_{i,n} |\mathbf{A}\mathbf{y}_n + \mathbf{t} - \mathbf{x}_n|^2, \quad (5)$$

where  $i$  is the iteration number and the weight  $\omega_{i,n}$  is a function of the errors in the previous iteration. In order to avoid dividing with zero, a regularization,  $|\delta| \ll 1$ , is introduced. For details on IRLS, see, e.g., Aftab and Hartley [1]. In the end, the weights are given by

$$\omega_{i,n} = \begin{cases} 1/\max\{|r_{i-1,n}|, |\delta|\}, & |r_{i-1,n}| < \epsilon, \\ 0, & |r_{i-1,n}| \geq \epsilon, \end{cases} \quad (6)$$

where the residuals are given by  $r_{i,n} = \hat{\mathbf{A}}_i \mathbf{y}_n + \hat{\mathbf{t}}_i - \mathbf{x}_n$ .

For the experiments, we used IRLS with 10 iterations, initialized by a composition of the transformation estimated by RANSAC described in the previous section and the known affine transformation between the atlas and the überatlas that was obtained in the offline co-registration. Results for truncation levels of both 10 and 50 mm are presented in the next section.

## 4 Experiments

The experimental evaluation was carried out on two different data sets and for two different settings, the first being pairwise affine registration and the second one, multi-atlas segmentation. To make full use of the data, the experiments were performed in a leave-one-out fashion, i.e., an überatlas was constructed using all but one image and then tested on that image.

### 4.1 Data sets

The first data set consists of 20 CT images of the heart. For these images delineations of the pericardium were obtained using a gold standard approach. More precisely, the delineations were drawn for every 10th slice in all three viewing directions by an expert involved in the SCAPIS project [13]. These 2D delineations were then interpolated into a complete 3D delineation that was finally approved by the expert.

The second data set consists of 30 MR images of the brain of young adults and delineation of 83 regions of the brain, which are manually drawn according to the protocol in Hammers et al. [11]. The data set is available online on <http://www.brain-development.org>. The images of the brains were upsampled with a factor 2.

## 4.2 Pairwise Affine Registration

In this evaluation we consider the quality of the pairwise affine registrations obtained from a new target image to the individual atlas images as described in Section 3.2. Comparisons are given to two baseline methods: The first is affine intensity-based registration using `NiftyReg` with default values. The second one is standard feature-based registration, feature points were extracted by a feature detection and description program implemented by Svärm et al. [24]. The program was run with default values and without rotation invariance. The affine registration was performed with 10000-iteration RANSAC. As for the outlier threshold,  $\epsilon$ , both 10 and 50 mm were evaluated.

Comparisons are done with respect to runtime and the dice index [8], which is a similarity measure between 0 and 1. The results for the heart images and brain images are given in Tables 1 and 2. Clearly, the überatlas registration produces larger values of the dice index than the compared methods and significantly reduces running times.

Table 1: Dice index and runtimes of the pair-wise affine registrations of the hearts. For the feature-based baseline and the überatlas registration, the runtimes include feature detection, matching and affine transformation. The feature-based baseline method and the überatlas registration was tested with two different values of the outlier threshold  $\epsilon$ .

Method	$\epsilon = 10$ mm	$\epsilon = 50$ mm	Runtime
Intensity	$0.754 \pm 0.130$	$0.754 \pm 0.130$	1054 s
Features	$0.870 \pm 0.057$	$0.811 \pm 0.084$	11.7 s
Überatlas	$0.874 \pm 0.072$	$0.888 \pm 0.038$	2.7 s

Table 2: Dice index and runtimes of the pair-wise affine registrations of the brains. For the feature-based baseline and the überatlas registration, the runtimes include feature detection, matching and affine transformation. The feature-based baseline method and the überatlas registration was tested with two different values of the outlier threshold  $\epsilon$ .

Method	$\epsilon = 10$ mm	$\epsilon = 50$ mm	Runtime
Intensity	$0.637 \pm 0.024$	$0.637 \pm 0.024$	63 s
Features	$0.640 \pm 0.023$	$0.630 \pm 0.026$	7.1 s
Überatlas	$0.643 \pm 0.022$	$0.638 \pm 0.024$	1.0 s



### 4.3 Multi-Atlas Segmentation

For multi-atlas segmentation three different setups were tested:

- (II) Affine intensity-based registration + nonrigid intensity-based registration
- (UI) Affine überatlas registration + nonrigid intensity-based registration
- (UU) Affine überatlas registration + nonrigid feature-based registration

*Implementation Details.* The intensity-based nonrigid registration was based on the free-form deformation proposed by Rueckert et al. [23], and implemented by Modat et al. [18]. The affine transformation was used as an initialization for the nonrigid transformation. The program was run with default values. The feature-based nonrigid registration was based on a point-based free-form deformation implemented by D. Kroon based on a proposition of Lee et al. [16]. The correspondences marked as inliers according to the outlier threshold were used as input. The nonrigid feature-based transformation was run with default values, apart from the number of grid refinements that was set to 5.

*Experimental results.* Dice index and runtimes of the multi-atlas based segmentations are given in Tables 3 and 4 for heart and brain images, respectively. Clearly, the überatlas registration produces comparable values of the dice index for the compared methods and significantly reduces the runtimes. Considering the brain images, the multi-atlas based segmentation results are on par with the intensity-based multi-atlas approach by Heckemann et al. [12] which is current state of the art. They use the same brain data set and obtain a mean dice index equal to 0.8173. Furthermore, assuming a multi-atlas consisting of 100 images, the intensity-based multi-atlas segmentation would take 4.5 hours compared to our method that would take only 1.9 minutes. A multi-atlas segmentation with the feature-based baseline would take 13 minutes.

Table 3: Dice index and runtimes of the multi-atlas based segmentations of the hearts. For the affine überatlas registration, the runtimes include feature detection, matching and affine transformation. Abbreviations: AFF = affine registration, NR = nonrigid registration, II = intensity AFF + intensity NR, UI = überatlas AFF + intensity NR, UU = überatlas AFF + überatlas NR.

Method	Dice index	Runtime AFF	Runtime NR	Runtime total
II	$0.933 \pm 0.049$	20026 s	36955 s	56981 s
UI	$0.969 \pm 0.033$	52 s	36955 s	37007 s
UU	$0.951 \pm 0.019$	52 s	121 s	173 s

Table 4: Dice index and runtimes of the multi-atlas based segmentations of the brains. For the affine überatlas registration, the runtimes include feature detection, matching and affine transformation. Abbreviations: AFF = affine registration, NR = nonrigid registration, II = intensity AFF + intensity NR, UI = überatlas AFF + intensity NR, UU = überatlas AFF + überatlas NR.

Method	Dice index	Runtime AFF	Runtime N-R	Runtime total
II	$0.805 \pm 0.010$	1827 s	2900 s	4727 s
UI	$0.806 \pm 0.010$	30 s	2900 s	2930 s
UU	$0.769 \pm 0.013$	30 s	23 s	53 s

## 5 Conclusion

Our proposed überatlas registration framework, is a novel, feature-based method that uses co-registration of atlases and clustering of feature points as a pre-processing step in order to speed up the computations of the image registrations needed in multi-atlas segmentation, but also to reduce the amount of outliers. The robust estimation of the affine transformations are done with the truncated  $l_1$ -norm as loss function. The registration method is faster than the two methods used as a comparison. For instance, computing pair-wise affine registrations of the heart takes 2.7 seconds on average compared to 11.7 and 1054 seconds, respectively.

The experimental results show that the developed affine registration algorithm is more robust to outliers making it both less sensitive to parameter selection and allows for a higher truncation threshold. The need for tuning parameters is therefore decreased but this factor is especially valuable when an affine transformation describes the actual transformation between the images poorly.

The proposed method succeeds to produce better segmentation results with respect to the dice index and has a lower standard deviation, both compared to the intensity and the feature based baseline. It produces comparable multi-atlas segmentation results to the current state of the art. Fine-tuning of the algorithm (such as choosing reference atlas in a more structured way and optimizing parameters) could yield even better results.

## References

1. Aftab, K., Hartley, R.: Convergence of iteratively re-weighted least squares to robust M-estimators. *IEEE Winter Conference on Applications of Computer Vision* (2015)
2. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded up robust features. *Computer Vision and Image Understanding* 110(3), 346-359 (2008)
3. Blake, A., Zisserman, A.: *Visual reconstruction*. MIT Press 225 (1987)
4. Chartrand, R., Yin, W.: Iteratively reweighted algorithms for compressive sensing. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 3869-3872 (2008)

5. Chen, A., Niermann, K.J., Deeley, M.A., Dawant, B.M.: Evaluation of multiple-atlas-based strategies for segmentation of the thyroid gland in head and neck CT images for IMRT. *Physics in Medicine and Biology* 57(1), 93-111 (2012)
6. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - Their training and application. *Computer Vision and Image Understanding* 61(1), 38-59 (1995)
7. Dey, D., Ramesh, A., Slomka, P.J., Nakazato, R., Cheng, V.Y., Germano, G., Bermana, D.S.: Automated algorithm for atlas-based segmentation of the heart and pericardium from non-contrast CT. *Proceedings of SPIE* (2010)
8. Dice, L.R.: Measures of the amount of ecologic association between species. *Ecology* 26(3), 297-302 (1945)
9. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6), 381-395 (1981)
10. Gill, G., Toews, M., Beichel, R.R.: Robust initialization of active shape models for lung segmentation in CT scans: A feature-based atlas approach. *International Journal of Biomedical Imaging*, 479154 (2014)
11. Hammers, A., Allom, R., Koepp, M.J., Free, S.L., Myers, R., Lemieux, L., Mitchell, T.N., Brooks, D.J., Duncan, J.S.: Three dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. *Human Brain Mapping* 19(4), 224-247 (2003)
12. Heckemann, R. A., Keihaninejad, S., Aljabar, P., Rueckert, D., Hajnal, J. V., Hammers, A.: Improving intersubject image registration using tissue-class information benefits robustness and accuracy of multi-atlas based anatomical segmentation. *NeuroImage* 51(1), 221-227 (2010)
13. Hjärt och Lungfonden: SCAPIS - en världsunik nationell kunskapskälla. <http://www.hjart-lungfonden.se/scapis>, 2014. [Online; accessed 30- Sept-2014].
14. Khalifa, F., Beache, G.M., Gimel'farb, G., Suri, J.S., El-Baz, A.S.: State-of-the-art medical image registration methodologies: A survey. *Multi modality state-of-the-art medical image segmentation and registration methodologies*, 235-280, Springer Science+Business Media (2011)
15. Kirisli, H. A., Schaap, M., Klein, S., Neefjes, L.A., Weustink, A.C., van Walsum, T., Niessen, W.J.: Fully automatic cardiac segmentation from 3D CTA data: a multi-atlas based approach. *Proceedings of SPIE* (2010)
16. Lee, S., Wolberg, G., Shin, S.Y.: Scattered data interpolation with multilevel B-splines. *IEEE Transaction on visualization and computer graphics* 3(3), 228-244 (1997)
17. Lowe, D.G.; Object recognition from local scale-invariant features. *Proceedings of International Conference on Computer Vision* 2, 1150-1157 (1999)
18. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S.: Fast free-form deformation using graphics processing units. *Computer Methods and Programs in Biomedicine* 98(3), 278-284 (2009)
19. Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N.: Reconstructing a 3D structure from serial histological sections, *Image and Vision Computing* 19(1), 25-31 (2001)
20. Ourselin, S., Stefanescu, R., Pennec, X.: Robust registration of multi-modal images: Towards real-time clinical applications. *MICCAI 2002, LNCS 2489*, 140-147 (2002)
21. Panda S., Asman A.J., Khare S-P., Thompson L., Mawn L.A., Smith S.A., Landman B.A.: Evaluation of multi-atlas label fusion for in vivo MRI orbital segmentation. *Journal of Medical Imaging* 1(2) (2014)

22. Pham, D.L., Xu, C., Prince, J.L.: Current methods in medical image segmentation. *Annual Review of Biomedical Engineering* 2, 315-337 (2000)
23. Rueckert, D. , Sonod, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging* 18(8), 712-721 (1999)
24. Svärm, L., Enqvist, O., Kahl, F., Oskarsson, M.: Improving robustness for inter-subject medical image registration using a feature-based approach. *International Symposium on Biomedical Imaging* (2015)
25. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging* 32(7) (2013)
26. Wang, H., Suh, J.W., Das, S.R., Pluta, J., Craige, C., Yushkevich, P.A.: Multi-atlas segmentation with joint label fusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2012)
27. Xu, R.: Survey of clustering algorithms. *IEEE Transactions on Neural Networks* 16(3), 645-678 (2005)