COMPARISON OF DIFFERENT METODS FOR ATLAS CONSTRUCTION

Hrvoje Kalinić[1], *Sven Lončarić*[1], *Maja Čikeš*[2], *Davor Miličić*[2], *Bart Bijnens*[3]

[1] Department of Electronic Systems and Information Processing, University of Zagreb
Unska 3, 10000 Zagreb, Croatia
phone: + (385) 1 6129940, fax: + (385) 6129652
email: hrvoje.kalinic@fer.hr
web: ipg.zesoi.fer.hr

[2] Department for Cardiovascular Diseases, University Hospital Centre Zagreb, 10000 Zagreb, Croatia

[3] Catalan Institution for Research and Advanced Studies (ICREA) and Universitat Pompeu Fabra (CISTIB)
E08003 Barcelona, Spain

ABSTRACT

When comparing experimental data obtained from different subjects, a standard approach is to display results on an atlas, as a common anatomical substrate. Through time, different methodologies for atlas constructions have been developed. A lot of works on this topic discuss the selection of the "best" atlas, using different criteria. In this article we discuss several methods for atlas construction and propose a novel method for least-biased atlas construction. The four different methods proposed in this article are compared based to their segmentation results. The results show no significant error reduction if the same methodology but with different atlas is implemented. Finally, we propose a methodological improvement that is not based solely on "best" atlas selection but rather on "best" segmentation selection as well. The image set used in this work consist out of 140 ultrasound images of cardiac outflow velocity profiles from both healthy volunteers and patients.

1. INTRODUCTION

When comparing experimental data obtained from different subjects, a standard approach is to display results on an atlas, as a common anatomical substrate. This is done to bring useful prior information to segmentation and registration tasks, so that variation within population can be described with fewer (transformation) parameters. Atlases have broad application in medical image segmentation and registration and are often used in computer aided diagnosis to measure the shape of an object or to detect the morphological differences between patient groups. Various techniques for atlas construction are developed for different human organs, like the heart [1, 2, 3] and especially the brain [4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. Some of these authors do not discuss the selection of a "best" atlas for a specific task, while various others do. We will focus our attention on these works.

To our knowledge the only work which compares different methods for atlas construction is the work of Rohlfing et al., where few of the techniques for atlas construction are described (in [14]) and in more details explored (in [15]). In these papers, four different strategies for atlas selection are investigated. First, one individual from the set was selected, second, the average shape atlas was constructed, third, the most similar instance from the set was selected as atlas and fourth, several individual images were used as atlases and multi-classifier approach was introduced before final segmentation. However, although this work incorporate basic principle of atlas construction, it does not present all the possible options for atlas selection, therefore, we will list few more approaches for atlas selection and construction.

While it is the most trivial approach for atlas selection is just to pick a random image form a set, by doing this, we have a great chance to pick an instance far from the population mean and therefore biased toward part of the samples in the set. To produce better result, an atlas from several images can be constructed. A common way to do this is to transform other images onto an image from a set, to have the same spatial frame for further processing, and to calculate an average image [16, 17, 18]. Intensity averaged atlases based on such image have the same disadvantage, which is especially visible if the reference is picked far from the population mean, and has been noted by several authors (see [19, 20]).

To overcome this problem different methods were used, which generally reduces to two approaches: ether to pick the sample (whereas is the best to pick a sample closest to the mean), or to try to estimate (or converge to) the true mean of the population. This has led to a problem on how to construct the least biased atlas, which several authors have discussed.

In [21], Guimond et al. developed an iterative averaging algorithm to reduce the bias. Masland et al. in [22] proposed a method for least biased selection of target image, using iterative algorithm that minimizes the distance and maximizes the mutual information. Park et al. in [17] proposed an alternative algorithm that estimate the target image based solely on distance and argues that the least biased atlas should be done in this way because it is more robust since it is less affected by inherent noise in the images, and, since it uses estimation technique instead of iterative algorithm, is faster. In [13] the iterative technique whereby the atlas converges to the unknown population mean is also described by Toga and Thompson. It is important to notice that they used an algorithm that independently averages shape and intensity. Bathia et al. in [6] proposed the similar approach where one arbitrary image is used just as an intensity reference, after which the similarity between images is maximized using non-rigid transformation. To assure that the image calculated in this fashion is actually the mean (with respect to the transformation) they put the constrain that the sum of all transformation is equal to zero. In [5], Joshi et al. proposed the method which is invariant to target image selection since after the construction of the atlas in the space frame of the target image, the target image is transformed to the space frame of the mean transformation. The mayor improvement of this work is that this was done for large deformation which was not the case in [6]. As the dissimilarity measure the squared error distance was used and it was shown that the optimal

atlas (for the selected dissimilarity measure) is an average intensity atlas. A similar work using Kullback-Liebler divergence is described in [7] by Lorenzen et al..

In this paper we propose four different approaches to atlas construction which incorporate some of the basic principles of the works mentioned above. The atlases are compared with respect to their segmentation performance, which we believe should be the criteria which atlas is the "best" atlas. The atlases are constructed from the aortic outflow velocity profile images. The images represent the blood-flow velocities from the left ventricle to aorta through time. Images are acquired using continuous wave Doppler ultrasound method by a echocardiographic scanner (Vivid 7, GE Healthcare). An instance of an aortic outflow velocity profile of a patient and a healthy volunteer is presented in the Figure 1.

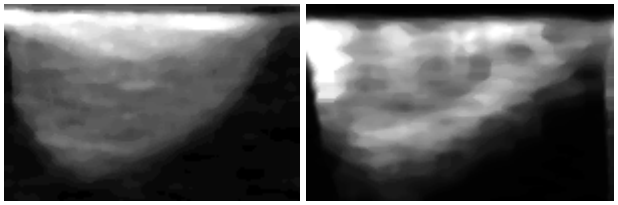

Figure 1: An example of the patient (left) and the healthy volunteer (right) aortic otuflow velocity profile.

2. METHOD

In the further subsections an image registration method used for atlas evaluation is introduced, and next, four methods for atlas construction are described, along with the framework for atlas evaluation. In all methods for atlas construction, described below, it is assumed that all the images are already rigidly aligned and resized to the same resolution. The images are denoted with I_i, and the set of images can be denoted as $S = \{I_i; i = 1..N\}$, where N stands for number of images. In this work we uses 140 images out of which 26 belongs to healthy volunteers, while the rest 114 belong to patients.

2.1 Image registration

To register two images transformation function, similarity measure and optimization algorithm have to be defined. For registration of cardiac outflow profiles we have used the transformation function that non-linearly transforms image only along y-axis since this should be enough to describe all the the possible physical changes. This can be expressed with the following formulae:

$$T(x,y) = t(x) \cdot y \tag{1}$$

where $t(x)$ denotes warping of image space frame along y-axis. This function was only estimated from N stripes selected from the image and calculating only the samples $t[k]$, after which the function $t(x)$ was reconstructed using linear interpolation. To measure the quality of the alignment between image, the similarity measure was defined in the form of normalized mutual information [23, 24, 25]. As an optimization algorithm a version of the gradient ascent algorithm [26] with multiresolution implementation was used.

2.2 Average intensity atlas

The average intensity atlas is probably the easiest way of construction an statistical atlas. The construction of an atlas can be explained with formula:

$$A_{av}(x,y) = \frac{1}{N}\sum_{i=1}^{N} I_i(x,y) \tag{2}$$

2.3 Median intensity atlas

With median intensity atlas we try to produce the atlas image only from the set of the pixels already existent in the image set S. The value of an each pixel at the position (x,y) in the mean intensity atlas was calculated as median of all pixels at the position (x,y), selected from the set of images I_i:

$$A_m(x,y) = median(I_i(x,y)); i = 1..N \tag{3}$$

2.4 Construction of least biased atlas with respect to transformation function

Let's assume that all the images are already registered onto each other using the approach described in Section 2.1. Now we can define the distance measure to calculate how far is any image (let's say I_i) of the set from the rest of the images in the set (i.e. $I_j \forall i \neq j$). Let's define the distance measure as:

$$d_{ij} = \sum_{j=1, j\neq i}^{N} \sum_{k=1}^{n} |log(t_{ij}[k])| \tag{4}$$

Where $t_i j$ denote transfer vector that transforms the image I_i onto the image I_j, and k stands for k-th element of the transformation vector. The logarithm was used in distance measure since the scaling of the image is done by multiplication. In this way we assure the symmetry of the distance measure, i.e. that two images one stretched by factor α and other squeezed by the same factor have the same distance from theirs originals.

Defined in this way, the distance measure from the Equation 4 will lead to selection of the image on which we can root the atlas. Now, when the root image is selected, the rest of the images are transformed on this image and the average intensity image is calculated (similar to the formula from the Equation 2). Finally, this image is then transformed so that it has the same distance from all the other images, where distance is again calculated as defined in Equation 4, with difference that it is now done for fixed j (the index of the root image).

2.5 Average shape and intensity atlas

The atlas construction procedure described in this section is inspired with the idea of shape-based interpolation of multidimensional object described in the work of Raya and Udupa [27]. They presented the method that was used for shape interpolation between slices acquired from medical imaging scanner. The basic idea behind this approach is to convert the binary image (which represents the segmented object) to gray image where the gray value of the point represent the shortest distance (from the border of the binary image). The distance is defined as positive for the points within an object and negative for the points outside of an object. For an image this distance function can be observed in 3D space. For a

circular object the distance transformation will have a shape of a cone, as shown in Figure 2.

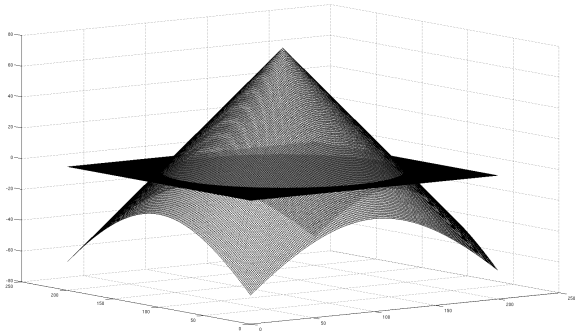

Figure 2: The distance transformation (cone) of an circular object on the plane.

Since the distance transformation of a binary image can again be represented as an (gray-level) image the process of averaging is by no means different. When the gray images from multiple objects are calculated and the images are averaged their mean shape can be extracted as a set of pixels with value zero in the average gray image. In other words, the border of the mean shape object is an isoline (contour line) with pixel value zero. We may think of the gray images from the Doppler ultrasound scanner as a set of 3-D object, if we represent the intensities as elevation (z-axis coordinate). This can help us to extend the idea of Raya and Udupa to a 3D space. Now, the shortest distance from the 3-D object is represented by isoplanes. The distance transformation will now give a function that exists in 4-D space and in its discrete form describes a set of isoplanes. After averaging this set we need to find the isoplane with value zero (i.e. the distance is zero). This isoplane is shape average of the 3-D object. Since the 3-D object contains the informations from both the intensity and shape of an object form an 2-D Doppler image we can say that this object is shape and intensity average of the images used for its construction. We convert this object back to gray (2-D) image and use it as an atlas.

2.6 A framework for atlas evaluation

In this section we will propose the method for evaluation four atlases type described in Sections 2.2 – 2.5. The proposed method for atlas evaluation is based on the segmentation accuracy of each atlas and the steps of this method are depicted in Figure 3.

The problem of segmentation evaluation lies in the fact that segmentation accuracy may wary based on error from manual segmentation of an atlas (let's denote it with e_m), registration error (e_r), error form suboptimal choice of an atlas (e_a) and golden standard error (e_g). As a golden standard a manual segmentation of an image was used, and since the same image and same segmentation is used in all experiments this error is constant across experiments. In this way, we expect that only the segmentation accuracy is affected. Similarly, one can use $\sigma(e_m + e_r)$ to denote the uncertainty of a method, since it depends on both the error form manual segmentation of an atlas and the registration error. With this, only the precision of segmentation is affected, but since the same method is used in each experiment, the additional

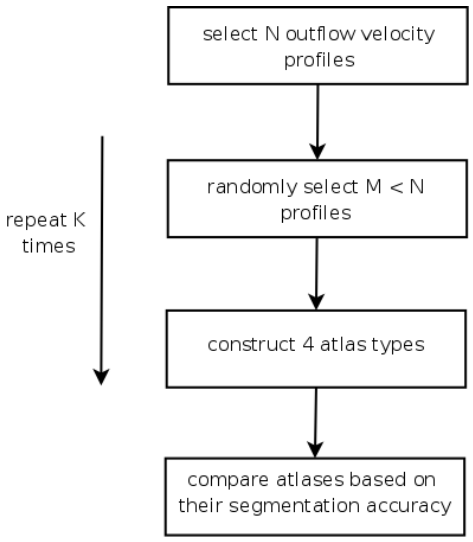

Figure 3: A flow chart of the proposed method for atlas evaluation.

variation across experiments (due to e_m and e_r) should be approximately the same.

When each instance of an atlas is constructed as shown in Figure 3 all the images that were not used for atlas construction, were transformed onto the atlas. The registration procedure for this is already described in Section 2.1. The transformation parameters were memorized and after the atlas is segmented, this segmentation was backward transformed onto the images where the segmentation evaluation is done. The segmentation accuracy and precision is used to evaluate the atlases performance. In these experiments, $M = 50$ images were used to construct an atlas, and $K = 22$ times different images from the set were selected, to compare the variation across different atlas types, and across different image selection.

3. RESULTS

The preliminary results shows that segmentation error has a small variation across different atlas types. This is shown in Table 1.

	Err
A_1	5.2777%
A_2	5.2824%
A_3	4.5336%
A_4	4.7182%

Table 1: Atlas performance comparison based on segmentation accuracy. A_1 to A_4 denote atlases described in Sections 2.2 to 2.5, respectfully.

Additionally, the experiment was repeated $K = 22$ times for each method for atlas construction. The results can be depicted as shown in Figure 4, where vertical lines shows one standard deviation of the segmentation error for different atlas type. We can notice that the error distribution overlap significantly. From this we can conclude that there are cases

when carefully selected images for atlas construction outperform the carefully selected method for atlas construction.

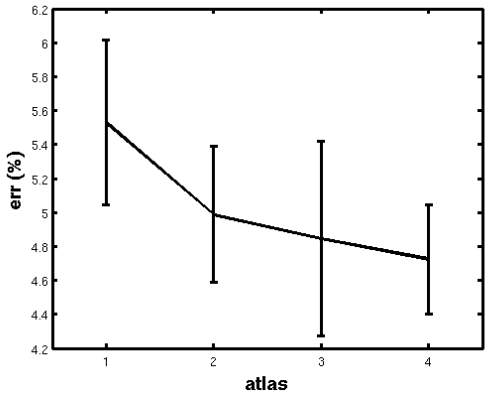

Figure 4: The average segmentation error along with denoted one standard deviation range. The 1 to 4 on y-axis denotes atlases described in Sections 2.2 to 2.5, respectfully.

Knowing that atlas segmentation also affects segmentation performance a concept of least biased segmentation selection (as compared with least biased atlas selection) is developed. All the images which contributed to atlas construction were segmented, and this segmentation was propagated along with image when atlas is constructed. The average segmentation is used as least biased segmentation of an atlas. The segmentation evaluation was conducted in the same fashion as in the previous experiment. The results are shown in Figure 5.

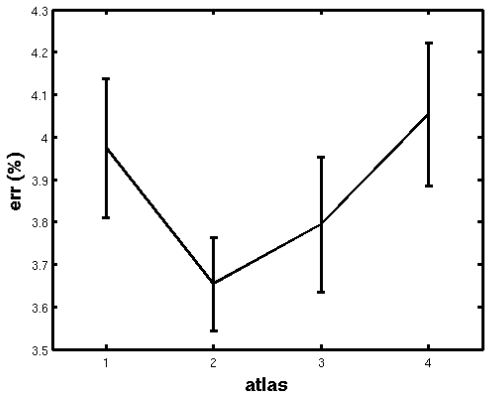

Figure 5: The average segmentation error along with denoted one standard deviation range. The 1 to 4 on y-axis denotes atlases described in Sections 2.2 to 2.5, respectfully. The results from least biased segmentation concept.

From the results we can conclude that least biased segmentation selection improves performance of any atlas since it has lower mean error as well as the standard deviation.

4. CONCLUSION

In this article we proposed a novel method for least biased atlas construction (Section 2.5), and compared the results of this method with the performance of three different atlases (Figure 4). Although our method outperforms them, this improvement is not statistically significant. Additionally, we shown that the selection of a least biased atlas did not lead to significantly better segmentation results, which may be contributed to the enough plastic transformation conducted within registration process. Finally, in this paper we have presents the results that show how carefully selected atlas segmentation may have greater impact on segmentation accuracy than the atlas selection. We also proposed a methodology for construction of a least biased (atlas) segmentation. The atlases with this property had a better segmentation performances than the atlases without this property.

REFERENCES

[1] A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 9, pp. 1151–1166, 2002.

[2] D. Perperidis, R. Mohiaddin, and D. Rueckert, "Construction of a 4D statistical atlas of the cardiac anatomy and its use in classification." *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv*, vol. 8, no. Pt 2, pp. 402–410, 2005.

[3] M. Lorenzo-Valdés, G. I. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert, "Atlas-Based Segmentation and Tracking of 3D Cardiac MR Images Using Non-rigid Registration," in *MICCAI (1)*, 2002, pp. 642–650.

[4] P. Fillard, X. Pennec, P. M. Thompson, and N. Ayache, "Evaluating Brain Anatomical Correlations via Canonical Correlation Analysis of Sulcal Lines," in *Proc. of MICCAI'07 Workshop on Statistical Registration: Pairwise and Group-wise Alignment and Atlas Formation*, Brisbane, Australia, 2007.

[5] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy." *Neuroimage*, vol. 23 Suppl 1, 2004.

[6] K. K. Bhatia, J. V. Hajnal, B. K. Puri, A. D. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *Proceedings of the IEEE Symposium on Biomedical Imaging (ISBI*, 2004, pp. 908–911.

[7] P. Lorenzen, B. Davis, G. Gerig, E. Bullitt, and S. Joshi, "Multi-class posterior atlas formation via unbiased kullback-leibler template estimation," in *In LNCS*, 2004, pp. 95–102.

[8] P. Lorenzen, M. Prastawa, B. Davis, G. Gerig, E. Bullitt, and S. Joshi, "Multi-Modal Image Set Registration and Atlas Formation," *Medical Image Analisys*, vol. 10, no. 3, pp. 440–451, June 2006.

[9] M. Bach Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J. Villemure, and J. Thiran, "Atlas-Based Segmentation of Pathological Brain MR Images," in *Proceedings of International Conference on Image Processing 2003 , ICIP'03, Barcelona, Spain*, vol. 1. Boston/Dordrecht/London: IEEE, 2003, pp. 14–17.

[10] D. Rueckert, A. F. Frangi, and J. A. Schnabel, "Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration." *IEEE Trans*

Med Imaging, vol. 22, no. 8, pp. 1014–1025, August 2003.

[11] D. C. V. Essen and H. A. Drury, "Structural and functional analyses of human cerebral cortex using a surface-based atlas," *Journal of Neuroscience*, vol. 17, no. 18, pp. 7079–7102, September 1997.

[12] C. Studholme and V. Cardenas, "A template free approach to volumetric spatial normalization of brain anatomy," *Pattern Recognition Letters*, vol. 25, no. 10, pp. 1191–1202, July 2004.

[13] A. W. Toga and P. M. Thompson, "The role of image registration in brain mapping," *Image and Vision Computing*, vol. 19, no. 1-2, pp. 3–24, January 2001.

[14] T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer, "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains." *Neuroimage*, vol. 21, no. 4, pp. 1428–1442, April 2004.

[15] T. Rohlfing, R. Brandt, R. Menzel, D. B. Russakoff, and . C. R. Maurer, Jr., *The Handbook of Medical Image Analysis – Volume III: Registration Models*. Kluwer Academic / Plenum Publishers, 2005, ch. Quo Vadis, Atlas-Based Segmentation?, pp. 435–486.

[16] L. Zöllei, M. Shenton, W. Wells, and K. Pohl, "The Impact of Atlas Formation Methods on Atlas-Guided Brain Segmentation," *Workshop on Statistical Registration: Pair-wise and Group-wise Alignment and Atlas Formation at the 10th International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 39–46, 2007.

[17] H. Park, P. H. Bland, A. O. Hero, and C. R. Meyer, "Least Biased Target Selection in Probabilistic Atlas Construction," in *MICCAI (2)*, 2005, pp. 419–426.

[18] N. Kovacevic, J. Chen, J. G. Sled, J. Henderson, and M. Henkelman, "Deformation Based Representation of Groupwise Average and Variability," in *MICCAI (1)*, 2004, pp. 615–622.

[19] P. E. Roland and K. Zilles, "Brain Atlases - A New Research Tool," *Trends in Neurosciences*, vol. 17, no. 11, pp. 458–467, 1994.

[20] D. J. Blezek and J. V. Miller, "Atlas Stratification," in *MICCAI (1)*, 2006, pp. 712–719.

[21] A. Guimond, J. Meunier, and J.-P. Thirion, "Average brain models: a convergence study," *Computer Vision and Image Understanding*, vol. 77, no. 9, pp. 192–210, February 2000.

[22] S. R. Marsland, C. J. Twining, and C. J. Taylor, "Groupwise non-rigid registration using polyharmonic clamped-plate splines," in *in Medical Image Computing and Computer-Assisted Intervention - MICCAI*, vol. 2879 (2), June 2003, pp. 771–779.

[23] J. M. Fitzpatrick, D. L. G. Hill, and C. R. J. Maurer, *Handbook of Medical Imaging*. SPIE Press, 2000, vol. 2. Medical Image Processing and Analysis, ch. Image Registration, pp. 447–514.

[24] J. V. Hajnal, L. G. Hill, and D. J. Hawkes, Eds., *Medical Image Registration*, First ed. CRC Press, Cambridge, 2001.

[25] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information." *IEEE Transaction on Medical Imaging*, vol. 16, pp. 187–198, April 1997.

[26] R. Fletcher, Ed., *Practical Methods of Optimization*, Second ed. John Wiley and Sons, 2001.

[27] S. P. Raya and J. K. Udupa, "Shape-based interpolation of multidimensional object," *IEEE Transaction on Medical Imaging*, vol. 9, no. 1, pp. 32–42, 1990.