35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Atlas Image Labeling of Subcortical Structures and Vascular Territories in Brain CT Images

Kaifang Du[1], Li Zhang[2], Tony Nguyen[2], Vincent Ordy[2],
Heinz Fichte[3], Hendrik Ditt[3], Christophe Chefd'hotel[2]

***Abstract*— We propose a multi-atlas labeling method for subcortical structures and cerebral vascular territories in brain CT images. Each atlas image is registered to the query image by a non-rigid registration and the deformation is then applied to the labeling of the atlas image to obtain the labeling of the query image. Four label fusion strategies (single atlas, most similar atlas, major voting, and STAPLE) were compared. Image similarity values in non-rigid registration were calculated and used to select and rank atlases. Major voting fusion strategy gave the best accuracy, with DSC (Dice similarity coefficient) around 0.85 $\pm$ 0.03 for caudate, putamen, and thalamus. The experimental results also show that fusing more atlases does not necessarily yield higher accuracy and we should be able to improve accuracy and decrease computation cost at the same time by selecting a preferred set with the minimum number of atlases.**

I. Introduction

Brain CT (computed tomography) imaging plays an essential role in clinical diseases diagnosis with the advantages of quantitative measurements, short imaging time, ease of access, high resolution of bony detail, and detection of hemorrhage. Therefore CT is the method of choice especially in stroke perfusion imaging. Detection of the morphological signatures in brain CT images provides useful diagnostic information for brain diseases. For example, interruption in the blood supply to a certain area of the brain will cause the ischaemia, infarction and eventual necrosis of tissue. These changes can be interpreted and localized from CT image for early diagnosis. The quantitative analysis of brain CT images usually benefits from labeling of subcortical structures and vascular territories. Reliable and accurate CT labeling of vascular territories can help localize brain lesion and detect cerebral infarction. For instance, the Alberta Stroke Program Early CT Score (ASPECT) was proposed as a 10-point quantitative topographic CT scan score used in patients with middle cerebral artery (MCA) stoke, therefore the segmental assessment of MCA territory will help the diagnosis [1].

Unlike brain MRI labeling which has rich literature, there are only a few studies on labeling of subcortical structures and vascular territories for brain CT data. Maldjian et al [2] presented an atlas-based automated method for identifying potential areas (lentiform nucleus, internal capsule, and insula) of acute ischemia on CT scans. Their approach can be extended to other anatomical regions and vascular territories using image registration and multi-atlas labeling approach proposed in this paper. Rutczyńska et al [3] presented structures segmentation from brain CT by employing adaptive filtering, Gaussian mixture modeling and context-based enhancement. Their approach was claimed better than region growing segmentation results.

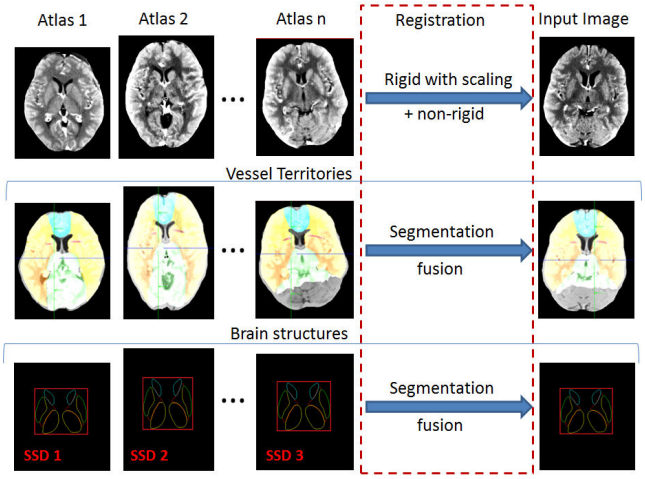

Fig. 1. Processing Pipeline of the Atlas-based Image Labeling.

Atlas-based labeling is a commonly used technique to segment the query image by propagating labels from the atlas with image registration. An atlas is defined as the pairing of an original CT scan and a corresponding "ground truth" labeling. When a moving image (the atlas) is registered non-rigidly to a fixed image (the query image), the calculated transformation is used to propagate the labels from the atlas to the space of the query image. The labeling accuracy can be improved considerably by multi-atlas labeling with different label fusion methods [4]. Usually multi-atlas method has better accuracy than single atlas.However, the major drawback of multi-atlas segmentation is its expensive computation cost.

The objective of this study is to develop a multi-atlas labeling approach for brain subcortical structures and vascular territories for clinical use. Several label fusion strategies are compared. The labeling accuracy is reported as Dice similarity coefficient and Hausdorff distance. Image similarity values in non-rigid registration are calculated and the information helps select the preferred atlas. The change in labeling

[1] K. Du was with Siemens Corporate Technology, Imaging and Computer Vision, Princeton, NJ, USA. He is now with the Department of Biomedical Engineering, University of Iowa, Iowa City, IA, USA (e-mail: kaifang-du@uiowa.edu)

[2] L. Zhang (corresponding author), T. Nguyen, V. Ordy, and Chefd'hotel are with Siemens Corporate Technology, Imaging and Computer Vision, Princeton, NJ, USA (e-mail: lizhang@siemens.com)

[3] H. Fichte and H. Ditt are with Siemens AG, Healthcare Sector, Imaging & Therapy Division, CT, Forchheim, Germany

accuracy versus atlas number is analyzed. The subcortical structures labeling shows good accuracy and demonstrates the potential use of multi-atlas labeling for brain vascular territories.

II. METHODS

A. Atlas-based Labeling for Brain CT Images

Figure 1 shows the processing pipeline in this study. With the deformation field from registrations between atlases and query image, subcortical structures and vascular territories labels are warped and fused to finally label the query image.

Image registration performs the task of finding a spatial transformation from the atlas to the query image. Before registration, all images were preprocessed to remove the skull bone using thresholding and voxel connection. Rigid registration is firstly applied to the entire brain area to place the images in global spatial alignment. Scaling is also optimized with the rigid registration parameters to compensate inter-subject size variation. The classification of a query image using the atlas can be done by finding the transformation such that $L_{sub} = \Phi_i * I_i$. In this study, we acquire this transformation by a diffeomorphic non-rigid registration using a variational approach [5].

B. Multi-atlas Label Fusion Strategies

Labeling from multiple atlases can be fused to provide a consensus labeling estimate for the query image. The multi-atlas approach reduces the effect of errors associated with an individual atlas, such as registration error and anatomy variability. The fusion of labels is at the voxel level and can be achieved in different strategies. Table I summarizes the four label fusion strategies that are analyzed in this study - SINGLE, SIMILAR, MV, and STAPLE [4].

Label Fusion	Description
SINGLE	One of the atlases with ground truth labels is chosen to serve as an individual atlas
SIMILAR	One of the atlases with minimum SSD between atlas and query image is chosen to serve as the atlas
MV	All atlases except the query image from the database are used by "major voting" label fusion strategy
STAPLE	All atlases except the query image from the database are used by STAPLE label fusion strategy [4]

TABLE I

SUMMARY OF FOUR LABEL FUSION STRATEGIES.

In SIMILAR strategy image similarity between the atlas and query image is calculated to select the preferred atlas. As shown in figure 1, to help select the most similar atlas, the sum of squared difference (SSD) is computed in a fixed ROI for each pair of registration after registration as an estimate of image similarity. The ROI is defined as the minimum bounding box in the query image (target image in registration) that contains all subcortical structures analyzed in this paper - caudate, putamen, thalamus, and internal capsule. Then the most similar atlas with the smallest SSD criterion is selected.

In majority voting (MV) strategy, the final label assigned to a voxel is decided by "majority vote" from all propagated labels. The STAPLE approach presented by Warfield et al [4] uses expectation maximization to iterate between the estimation of true consensus segmentation and the estimation of reliability parameters for each rater. The reliability parameters are based on the sensitivity and specificity of each rater and are then used to weigh their contributions when generating the final consensus estimate.

C. Validation

The labeling is tested by measuring the agreement between ground truth and the segmentation produced by atlases. Dice similarity coefficient (DSC) is a commonly used validation method that measures the overlapping intersection divided by the mean volume of the two regions [6]. Another validation metric used in this paper is Hausdorff distance (HD), which calculates the maximum distance between two point sets in 3D surfaces [7].

III. EXPERIMENTAL RESULTS

The algorithm was tested by 18 brain perfusion dynamic CT scans from 18 clinical patients (age: 43 - 84 years old). Each dynamic CT scan consists of a volume series acquired over time using Siemens Adaptive 4D Spiral Scan (A4DS) technique. The number of slices for each CT volume ranges from 13 to 95, with slice spacing ranging from 1 mm to 5 mm. The dimension of each CT images is 512 $\times$ 512, with pixel spacing ranging from 0.33 mm to 1 mm. The field of view (FOV) on head-foot direction ranges from 65 mm to 90 mm, with different brain coverage. As different institutions used different scanning protocols to acquire the data, in addition to the brain coverage, the head position, especially the rotation about left-right axis, also varies.

The input data for vessel territory labeling was created from the dynamic CT data by averaging intensity over time for each voxel. The skulls in the average images were then removed before registration to avoid influence of topological skull change from craniotomy. The intensity averaging and skull removal were performed by a software application (syngo.Volume Perfusion CT Neuro) installed in Siemen CT scanners. The averaged perfusion images were manually labeled for subcortical structures and vascular territories and then used as the atlas labeling for our method. Leave-one-out cross-validation approach was used for validation. In this approach the manual labeling for the query subject was treated as the ground truth, and the remaining subjects were used as the atlases.

If single atlas was used, the average computation time is 34 seconds with Intel(R) Xeon (R) CPU 2.27 GHz (2 processors) Quad Core, 12.0 GB RAM. The computation speed and good labeling accuracy demonstrate potential use for a practical solution.

Figure 2 shows the labels for one subject in 3D view. Subcortical structures are shown on the left. Vascular territories

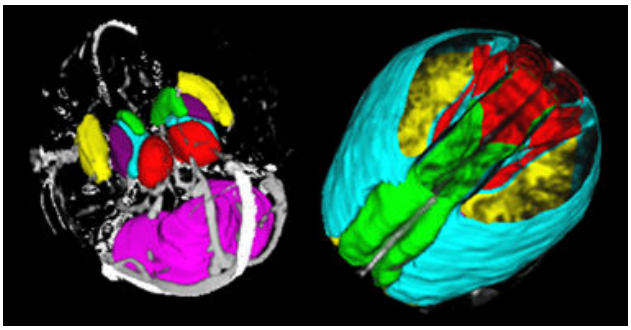

Fig. 2. 3D view of the labeling for subcortical structures (left) and vascular territory (right).

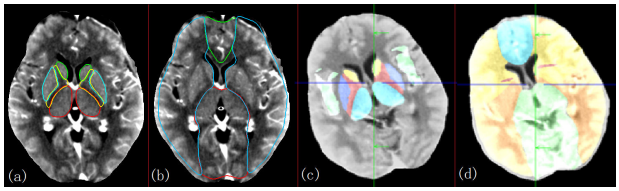

Fig. 3. (a) Manually labeled subcortical structures in atlas image. (b) Manually labeled vascular territories in atlas image. (c) Atlas-based labeling of subcortical structures in query image. (d) Atlas-based labeling of vascular territories in query image.

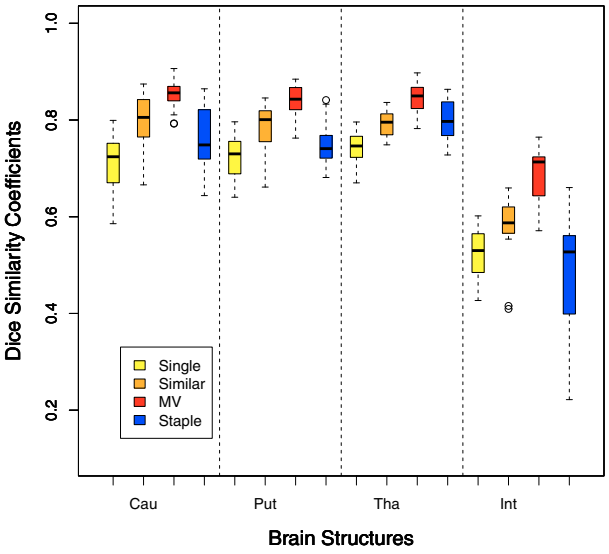

Fig. 4. Dice coefficients of subcortical structures labeling using SINGLE, SIMILAR, MV, and STAPLE label fusion strategies.

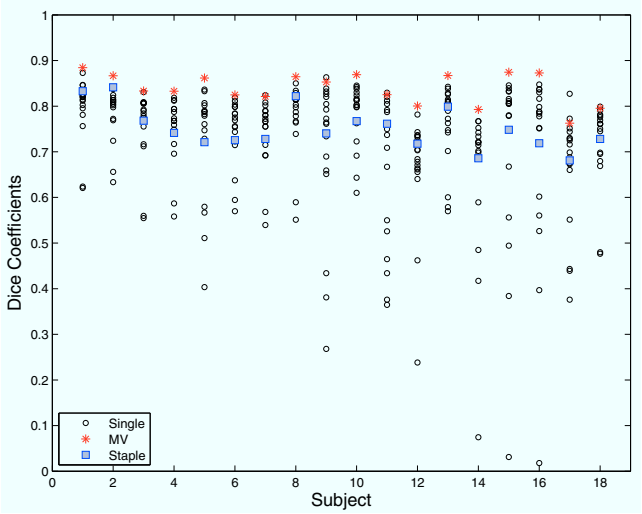

Fig. 5. DSC values of the subcortical structure putamen for each query subject, with label fusion strategy of SINGLE, MV, and STAPLE.

are shown on the right where green is anterior cerebral artery (ACA), blue and yellow are superior and inferior middle cerebral artery (MCA), and red is posterior cerebral artery (PCA). Figure 3 shows subcortical structures and vascular territories labels in a sample atlas and warped labels in a sample query subject using SINGLE atlas fusion approach.

Figure 4 shows the box plots for DSC between the ground truth and the fused labels for different subcortical structures with four different label fusion strategies. Each query subjects has 17 warped labels from SINGLE atlas strategy. Among the four subcortical structures, the DSC for internal capsule is the worst. The ground truth labeling of internal capsule was obtained by using the other subcortical structures and therefore all the manual labeling errors were accumulated here. However, this can be improved using a better atlas. For the four atlas fusion strategies, MV achieved the best labeling accuracy. SIMILAR label fusion strategy is better than SINGLE, while surprisingly the STAPLE approach is also worse than MV and sometimes even worse than SIMILAR.

Table III summarizes the labeling accuracy (mean $\pm$ standard deviation) for different label fusion strategies for all subcortical structures. Each label is tested through two metrics - DSC and Hausdorff Distance (HD). MV achieves best performance as around 0.85 DSC and less than 6 mm HD.

Figure 5 shows the labeling accuracy of subcortical structure putamen for each query subject. In the leave-one-out cross-validation, black circles dots represent DSC values from SINGLE atlas, red stars represents MV fusion with all 17 atlas, and blue square represents the STAPLE strategy. We can see the MV fusion method performs better than SINGLE and STAPLE for each query subject.

Figure 6 shows the change of DSC for subcortical structures labeling using MV, with increasing numbers of fused atlases. The atlas were ranked by the SSD criterion within the bounding box. The vertical axis shows mean DSC and the horizontal axis shows number of fused atlas. For example, only one atlas that has minimum SSD criterion is selected if the "Atlas number" is 1, and more atlases are added in sequence with increasing "Atlas number". We can see more atlases would not necessarily achieve better accuracy after certain threshold. The DSC values becomes to remain stable when the atlas number is greater than five.

		Caudate	Putamen	Thalamus	Internal Capsule
SINGLE	DSC	0.71 ± 0.058	0.72 ± 0.049	0.74 ± 0.033	0.53 ± 0.050
	HD (mm)	6.7 ± 2.9	10.2 ± 4.0	9.1 ± 1.4	16.5 ± 1.3
SIMILAR	DSC	0.79 ± 0.059	0.78 ± 0.051	0.79 ± 0.026	0.58 ± 0.069
	HD (mm)	5.0 ± 3.2	7.6 ± 4.6	8.0 ± 1.4	15.5 ± 3.5
MV	DSC	0.85 ± 0.032	0.84 ± 0.035	0.85 ± 0.030	0.69 ± 0.056
	HD (mm)	3.7 ± 2.9	6.0 ± 3.9	5.9 ± 1.4	20.3 ± 5.5
STAPLE	DSC	0.76 ± 0.066	0.75 ± 0.047	0.80 ± 0.042	0.49 ± 0.123
	HD (mm)	5.2 ± 3.7	8.0 ± 4.9	6.8 ± 1.3	17.0 ± 3.3

TABLE II

SUMMARY TABLE OF LABELING ACCURACY (MEAN ± STANDARD DEVIATION) FOR DIFFERENT LABEL FUSION STRATEGIES FOR ALL SUBCORTICAL STRUCTURES.

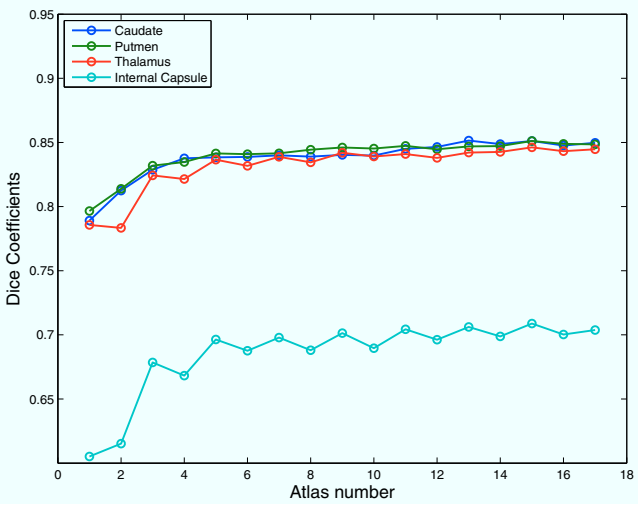

Fig. 6. Dice coefficients of subcortical structures labeling using major voting, with increasing numbers of combining atlas. The atlas were sorted by SSD criterion within the bounding box.

IV. DISCUSSION

The experimental results show DSC is around 0.85 ± 0.03 for caudate, putamen, and thalamus, which rivals the DSC results of atlas-based segmentation from brain MR images [8], [9]. This result demonstrates that using our atlas-based labeling technique, brain CT images would also achieve similar segmentation accuracy as MR images and thus provide useful tissue information for clinical diagnosis. The accurate labeling of subcortical structures also demonstrates it is feasible to use atlas-based segmentation upon brain vascular territories, which is more difficult to validate.

For all the four label fusion strategies, both DSC and HD suggest the accuracy of internal capsule is worse than the other subcortical structures. This might be explained by the curved shape of internal capsule, which may impede the label fusion from multiple atlases because label fusion is more susceptible to registration errors. The same finding was reported in atlas-based brain MRI segmentation [8].

Moreover, the labeling accuracy is directly related with the credence of ground truth labeling in atlases. Since our study in this paper is only focused on testing the proposed labeling method, better annotation for the atlas in our future work should improve the results, including internal capsule.

We compared the four atlas fusion strategies. Figure 4, 5, and table III show MV, SIMILAR, and STAPLE all exceed SINGLE strategy, in which MV achieves best accuracy and STAPLE even yields worse DSC values than MV and SIMILAR. This result can be supported from [9] which also reported unexpected worse STAPLE accuracy. As the conjectures of reason in [9], STAPLE may yield high recognition rate but not good DSC or HD numbers. Moreover, the performance of STAPLE may vary on different image data set with different priori probabilities.

We investigated the change of labeling accuracy with increasing number of involved atlases. Figure 6 shows if the number of atlas is less than five, labeling accuracy improves with increasing number of atlases. However, the accuracy almost remains unchanged when number of atlas is more than five. Incidentally in [8] it also reported that no increase or even decrease in the accuracy was obtained when the atlas number exceeds certain limit. This result inspires us that for the time-consuming issue, the preferred selection of a limited number of atlases would always appear preferable to using the whole atlas population especially in case of arbitrarily large atlas database.

V. CONCLUSION

In this paper, we proposed a multi-atlas labeling approach for subcortical structures and vascular territories in brain CT images. The labeling accuracy of subcortical structures in CT images rivals that of MR images and demonstrates the potential use of labeling for vascular territories in brain CT images. For future work, better accuracy may be achieved by improving labeling of the atlas. Four atlas label fusion strategies (SINGLE, SIMILAR, MV, STAPLE) are compared, in which MV shows the best accuracy. This result shows that an preferred atlas selection method can be used in future work to improve accuracy and computational efficiency at the same time.

VI. ACKNOWLEDGEMENT

The authors would like to thank the following hospitals for providing data: Nakamura Memorial Hospital, Hokkaido, Japan, Klin. Radiologie UK Münster, Germany, and Neuroradiologie Uni Göttingen, Germany.

REFERENCES

[1] J. H. Pexman, P. A. Barber, M. D. Hill, R. J. Sevick, A. M. Demchuk, M. E. Hudon, W. Y. Hu, and A. M. Buchan, "Use of the Alberta Stroke Program Early CT Score (ASPECTS) for assessing CT scans in patients with acute stroke," *AJNR Am J Neuroradiol*, vol. 22, no. 8, pp. 1534–1542, Sep 2001.

[2] J. A. Maldjian, J. Chalela, S. E. Kasner, D. Liebeskind, and J. A. Detre, "Automated CT segmentation and analysis for acute middle cerebral artery stroke," *AJNR Am J Neuroradiol*, vol. 22, no. 6, pp. 1050–1055, 2001.

[3] A. Rutczyńska, A. Przelaskowski, M. Jasionowska, and G. Ostrek, "Method of brain structure extraction for CT-based stroke detection," *Information Technologies in Biomedicine*, pp. 133–144, 2010.

[4] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *IEEE Trans Med Imaging*, vol. 23, no. 7, pp. 903–921, Jul 2004.

[5] C. Chefd'hotel, G. Hermosillo, and O. Faugeras, "Flows of diffeomorphisms for multimodal image registration," in *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2002, pp. 753–756.

[6] Lee R Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[7] D.P. Huttenlocher, G.A. Klanderman, and W.J. Rucklidge, "Comparing images using the hausdorff distance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 9, pp. 850–863, 1993.

[8] J. M. Lotjonen, R. Wolz, J. R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, and D. Rueckert, "Fast and robust multi-atlas segmentation of brain magnetic resonance images," *Neuroimage*, vol. 49, no. 3, pp. 2352–2365, Feb 2010.

[9] X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de Solorzano, "Combination strategies in multi-atlas image segmentation: application to brain MR data," *IEEE Trans Med Imaging*, vol. 28, no. 8, pp. 1266–1277, Aug 2009.