HYPERTHERMIA CRITICAL TISSUES AUTOMATIC SEGMENTATION OF HEAD AND NECK CT IMAGES USING ATLAS REGISTRATION AND GRAPH CUTS

V. Fortunati[1], R.F. Verhaart[2] , F. van der Lijn[1], W.J. Niessen[1,3], J.F. Veenland[1], M.M. Paulides[2] and T. van Walsum[1]

[1]Biomedical Imaging Group of Rotterdam, Department of Medical Informatics and Radiology, Erasmus MC, Rotterdam, The Netherlands
[2]Hyperthermia Unit, Department of Radiation Oncology, Daniel Den Hoed Cancer Center Erasmus MC, Rotterdam, The Netherlands
[3]Faculty of Applied Sciences, Delft University of Technology, Delft, The Netherlands

ABSTRACT

Outcome optimization of hyperthermia tumor treatment in the head and neck requires accurate hyperthermia treatment planning. Hyperthermia treatment planning is based on tissue segmentation for 3D patient model generation. We present here an automatic atlas-based segmentation algorithm for the organs at risk from CT images of the head and neck. To overcome the large anatomical variability, atlas registration and intensity-based classification were combined. A cost function composed of an intensity energy term, a spatial prior energy term based on the atlas registration and a regularization term is globally minimized using graph cut. The method was evaluated by measuring Dice similarity coefficient, mean and Hausdorff surface distances with respect to manual delineation. Overall a high correspondence was found with Dice similarity coefficient higher than 0.86 and a mean distance lower than the voxel resolution.

Index Terms— Atlas-based method, Graph-cuts, Image segmentation, Hyperthermia, Computed Tomography,

1. INTRODUCTION

During hyperthermia treatments (HT), tumor temperature is elevated to temperatures in the range of 40º-44º. The effectiveness of radiotherapy (RT) is significantly enhanced when combined with HT [1-2]. To further improve the proven effectiveness of HT in the head and neck (H&N) region a site-specific applicator was developed for precise HT in the entire (H&N) region using electromagnetic fields. This device, called the HYPERcollar, permits the application of deep heating (>4cm under the skin) to the H&N cancers [3-4]. Simultaneously, electromagnetic based hyperthermia treatment planning (HTP) has been developed to exploit the specific heat-focusing capabilities of the novel applicator. HTP requires the generation of patient-specific anatomical models. Their accuracy directly affects the quality of the optimized settings that HTP provides and, hence, treatment outcome. The generation of these models is particularly challenging due to 1) the complexity of the underlying anatomy, 2) the large number of anatomical structures in a relatively small area and 3) the low contrast between relevant tissues in the CT images. Moreover the presence of the tumor increases the anatomical variability between different subjects, adding to the complexity.

In this paper we focus on the accurate segmentation of the thermo-sensitive organs at risk (OAR), for which the treatment temperature level has to be restricted, i.e. brain, spinal cord [5] and eyes [6]. Exposure of the OAR must be kept below the threshold levels to avoid toxicity of the treatment. Using HTP, the ratio between tumor dose and thermal dose deposited in the OAR is optimized. Hence, the maximum thermal dose in OARs determines directly the possibility to achieve high temperatures in the tumor. Accurate segmentation of OARs is therefore a critical issue in HTP. This procedure is currently done manually in clinical practice. Since manual delineation is a tedious and labor intensive procedure, automation is needed.

This work proposes an automatic method for OARs segmentation. Multi-atlas registration-based segmentation is refined by a graph-cut optimization step [7] which takes into account label probability distribution and tissue intensities.

2. PREVIOUS WORK

Atlas based techniques have been used by several groups to segment H&N CT images [8-13]. The success of these techniques can be explained by the easy integration of context knowledge to discriminate between neighboring anatomical structures with unclearly resolved boundaries. Below we discuss registration approaches and atlas fusion strategies commonly applied in H&N atlas-based segmentation.

2.1. Registration methods

Han et al. [8-9] have used object boundary points to compute the deformation field in a three step registration algorithm: an initial linear step, a block matching to align boundary points and a dense deformable registration for refinement. Teng et al. [10] used a knowledge based unsupervised segmentation algorithm to initialize a cubic B-Spline non-rigid registration and to select the most similar atlas. Gorthi at al. [11] applied an active contour based atlas

registration. The registration is driven by region based forces that brings the atlas contour to move towards target regions with similar mean intensity value.

2.2. Atlas selection and fusion

Atlas selection and fusion strategies have been adopted to improve the segmentation results. Han et al. [12] have used STAPLE [14] to weigh the different atlas results in an optimal segmentation. In [10], the best atlas is selected based on the minimum Hausdorff distance between landmark surface meshes. Gorthi et al. [14] have chosen atlas fusion by majority voting for their registration framework [11].

2.3. Graph-cuts

Adding an intensity model and regularization to atlas-based segmentation has given good results in brain structures segmentation from MR images [8][15-16]. However this approach has not been applied to H&N CT image segmentation yet. To efficiently combine this information in an atlas-based segmentation framework graph-cut method is adopted. The advantages of graph-cut approaches [17] are well known: global optima, practical efficiency, numerical robustness and ability to fuse a wide range of visual feature and constraints.

Our combined atlas registration – graph cut approach permits us to combine strengths of both approaches: the atlas-based registration gives a suitable spatial prior for a more accurate gradient- and intensity-based graph-cut segmentation approach. In this way, we prevent inaccurate atlas-based segmentations that are caused by a deformation field that is constrained by the degrees of freedom of the transformation model, and by an atlas-fusion process that generally does not use intensity information.

3. METHOD

In our method, we solve for each critical tissue type a binary labeling problem: each voxel is assigned a label L_p which is 1 if p is a critical tissue voxel and 0 if p is a background voxel. The segmentation problem can be written as an energy minimization problem [7], where we are looking for labeling $\hat{\mathbf{L}}$ that minimizes the negative logarithm of the probability $p(\mathbf{I}|\mathbf{L})$ which, according to Bayes and ignoring the constant term $p(\mathbf{I})$ can be written as

$$\begin{aligned}\hat{\mathbf{L}} &= \arg\min_{\mathbf{L}}(-\ln p(\mathbf{I}|\mathbf{L}) - \ln p(\mathbf{L})) \\ &= \arg\min_{\mathbf{L}}\left(E_{intensity}(\mathbf{L}) + E_{prior}(\mathbf{L})\right), \quad (1)\end{aligned}$$

where $\mathbf{I}$ and $\mathbf{L}$ are the set of observed intensities and labels respectively. The prior energy is a function of a particular label distribution $\mathbf{L}$. A Markov random field, with cliques consisting of one and two voxels [18], is used to approximate $p(\mathbf{L})$ resulting in a prior energy:

$$\begin{aligned}E_{prior}(\mathbf{L}) &= \sum_{p\epsilon P}\beta_p(\mathrm{L}_p) + \sum_{p\epsilon P}\sum_{n\epsilon N_p}\beta_{p,n}(\mathrm{L}_p, \mathrm{L}_n) \\ &= E_{sprior}(\mathbf{L}) + E_{rprior}(\mathbf{L})\end{aligned}$$

$E_{prior}(\mathbf{L})$ is composed by a spatial prior information term $E_{sprior}(\mathbf{L})$, computed from the label probability distribution, and a discontinuity preserving function $E_{rprior}(\mathbf{L})$.

3.1. Intensity model

The term $E_{intensity}(\mathbf{L})$ depends on the intensity distribution within the foreground and background regions. The intensity distribution is estimated by sampling foreground and background voxels from the atlases. Here, each voxel is considered an independent and identically distributed random variable so:

$$p(\mathbf{I}|\mathbf{L}) = \prod_{p\epsilon P} p(I_p|L_p) \Rightarrow E_{intensity}(\mathbf{L}) = -\sum_{p\epsilon P}\ln p(I_p|L_p)$$

The intensity distribution is modeled by a Parzen window estimator for both the foreground and background voxels:

$$p(I_p|L_p) = \frac{1}{\mathrm{M}_0}\sum_{m=1}^{\mathrm{M}_0}\frac{1}{h}\phi\left(\frac{\mathrm{I_p} - \mathrm{I}_{\boldsymbol{m}}}{h}\right)$$

Where $\phi(\mathrm{I})$ is a gaussian kernel function and h is its width. m indexes the M_0 samples drawn from the training image. For our experiment we have used 250 random samples from background and foreground voxels of each training image and a window width h of 50 HU was chosen empirically.

3.2. Atlas based spatial prior

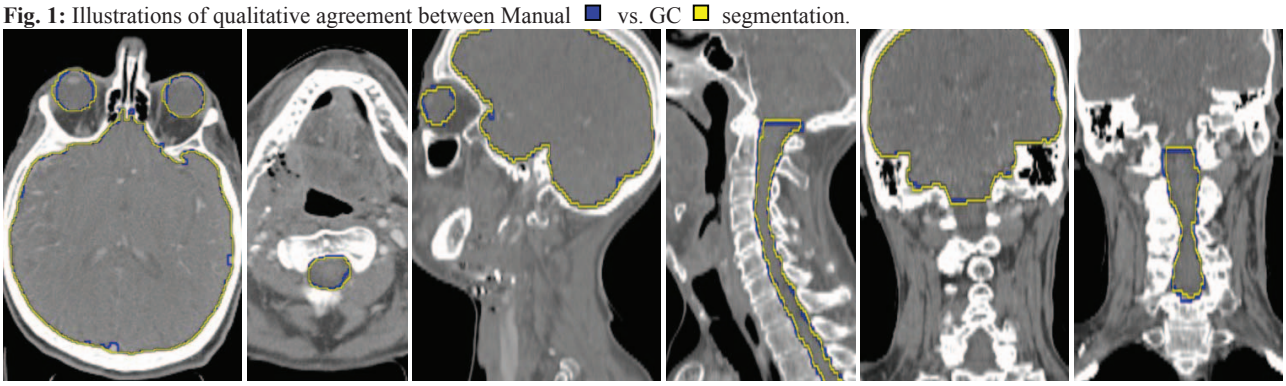
The spatial prior term E_{sprior} is constructed from the N training images. First, every atlas image is registered to the target and its label is propagated to the target image space. Then the N propagated labels are used to generate a probability map $p(\mathrm{L}_p)$ for foreground and background labels. The map is taken as single one-site clique potential:

$$\beta_p(\mathrm{L}_p) = -\ln p(\mathrm{L}_p)$$

For registration a two-step approach is followed. First a rigid registration step is applied: a sum of squared distances rigid registration with 6 degrees of freedom (3 rotations and 3 translations parameters) is employed; next a cubic B-Spline, multi resolution, mutual information registration [19] is used to obtain a fine non-rigid alignment. Some registration results in totally wrong alignment. To exclude these cases a safety criterion is adopted: if normalized cross correlation coefficient (NCC) between registered and the target image is $\mathrm{NCC} < 0.96$ the registration result is considered as erroneous and that image is not used as an atlas for that specific target image.

3.3 Regularization

Fig. 1: Illustrations of qualitative agreement between Manual ■ vs. GC ■ segmentation.



The regularization term E_{rprior} is given by [17]:

$$\beta_{p,n}(\mathrm{L}_p, \mathrm{L}_n) = \frac{1}{2} exp\left(-\frac{(\mathrm{I}_p - \mathrm{I}_n)^2}{2\xi^2}\right) \cdot \frac{1}{\mathrm{dist}(p,n)}, \quad \text{if } \mathrm{L}_p \neq \mathrm{L}_n$$

This gives a penalty to the energy if two neighboring voxels have a different label and intensity values within the difference range ξ. This value is chosen as the standard deviation of the foreground region (learned from the training set).

3.4. Graph cut

With the terms introduced above the total energy functional reads:

$$E(\mathbf{L}) = \sum_{p \epsilon P} \lambda_1 E_{intensity}(\mathrm{L}_p) + \lambda_2 E_{sprior}(\mathrm{L}_p) + \sum_{p \epsilon P} \sum_{n \epsilon N_p} E_{rprior}(\mathrm{L}_p, \mathrm{L}_p) \quad (2)$$

This function is converted into a directional graph to compute the global minimum, finding the minimum s-t cut. λ_1 and λ_2 are parameters that weight the single term contributions on the functional value. A standard graph-cut algorithm [20] is used to find the minimum energy cut. For each tissue a graph-cut is computed.

4. EXPERIMENTS AND RESULTS

4.1. Experiments

The method was trained and tested on two disjoint sets of 11 CT images for a total number of 22 sets of images corresponding to 22 patients. All the images were manually delineated by an expert for hyperthermia treatment planning. In total 6 different experts delineated the images (about 4 delineations for expert). The in-plane resolution is 1.0 mm for all images while the slice thickness ranges from 1.5 to 3 mm. We compared the graph-cut (GC) segmentation, with a multi-atlas segmentation (MA). MA results were obtained minimizing (2) without considering intensity and regularization priors. This is equivalent to fusing the single atlas segmentations for the same target, with majority voting.

The results were evaluated by comparison with the manual delineations. Dice similarity coefficients (DSC), Hausdorff and mean surface distances between automatic segmentation and reference standard were used as similarity measures. The evaluation is performed per tissue.

4.2. Parameter settings

A leave-one out experiment on the training set was performed to determine the optimal settings of the algorithm parameters. The DSC was used for the optimization of B-Spline grid-spacing [19] in the registration and of the weights λ_1 and λ_2 in the graph cut method. We used a 5-resolution B-Spline control points grid [19], and found optimal results with grid-spacing going from 112 mm to 7 mm, coarse-to-fine. The lowest resolution permits to cope with large non-rigid deformations. The graph-cut segmentation parameters were determined for each tissue separately: a combination of weights $\lambda_1 = 0.5$ and $\lambda_2 = 0.5$ is used for brain and eyes while $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$ for spinal cord. The same setting was used for all the images.

4.3. Results

The method was evaluated both on the training and test set using the optimal parameters. For the training set a leave one out approach was adopted. For the evaluation of the test images, we segmented each test image using the training set images as atlases. On average 9 atlases were used for the registration. The average results for training and test set are shown in Table 1. In Figure 2 the segmentation of one patient is shown: manual labels are compared with graph cut segmentation. DSC and mean distance values show that the overall segmentation correspondence is high for all the tissues: average DSC is higher than 0.86 for all the tissues and the average mean distance indicates accuracy within the voxel resolution. The Hausdorff distance values magnitude states the presence of outliers; these errors are caused by registration errors which the GC refinement could not recover, in fact there is hardly any improvement in

Hausdorff distance between MA and GC. The results on the training set are similar to the results on the test set. The improvements between MA and GC results were found to be statistically significant at the 0.05 level using paired t-test for spinal cord and eyes DSC.

Table 1: Training (Train) and Test set evaluation, MA vs GC. Mean DSC, mean ± standard deviation, Hausdorff and mean surface distances.

	DSC		Hausd. dist. (mm)		Mean dist. (mm)	
Train	**MA**	**GC**	**MA**	**GC**	**MA**	**GC**
Brain	0.97±0.01	0.98±0.01	14.3±5.4	14.7±5.1	0.7±0.2	0.5±0.1
S.Cord	0.80±0.03	0.86±0.04	10.2±5.6	9.4±4.9	1.2±0.6	0.9±0.4
Eyes	0.85±0.03	0.88±0.03	4.1±0.9	3.8±0.8	0.8 ±0.2	0.7±0.2
Test	**MA**	**GC**	**MA**	**GC**	**MA**	**GC**
Brain	0.97±0.01	0.98±0.01	11.5±3.1	11.2±3.2	0.7±0.3	0.6±0.2
S.Cord	0.81±0.03	0.87±0.04	8.9±3.4	8.6±4.4	1.2±0.5	0.9±0.5
Eyes	0.85±0.03	0.87±0.03	4.8±2.2	5.1±2.1	0.8 ±0.3	0.8±0.3

5. DISCUSSION AND CONCLUSION

In this study, we investigated a fully automatic method to segment thermal sensitive tissues using a combination of atlas registration and intensity modeling in a graph-cuts segmentation framework.

The results of the method over the training and test set are similar, this indicates that parameter optimization was good and that the training set is representative for the test set. When comparing the results of the original MA with the GC approach, the latter approach consistently performs better. For all tissue types, both average Dice and mean distance values are better for the GC approach. Compared to the original application of the method, hippocampus segmentation in MR images, the improvement of GC over MA is larger [8], suggesting that the introduction of the intensity model and regularization is more effective in our application. It may be caused by the fact that CT intensities have an absolute value, whereas intensities in MR are relative and may suffer from in-homogeneities.

We are planning to investigate the inter- and intra-observer variability in method accuracy using different atlases for segmentation. Further the influences of the variation of accuracy on the SAR and temperature distribution prediction will be analyzed.

We conclude that we implemented and evaluated an atlas-based graph-cut approach that has potential to improve the efficiency of temperature sensitive tissue delineation in HTP.

6. ACKNOWLEDGEMENT

This research was funded by technology foundation STW.

7. REFERENCES

[1] R. Valdagni, M. Amichetti, “Report of long-term follow-up in a randomized trial comparing radiation therapy and radiation therapy plus hyperthermia to metastatic lymphnodes in stage IV head and neck patient”, Int. J. Rad. Oncology Biol. Phys., 28, 163-169, 1993

[2] Y. Hua, S. Ma, Z. Fu, Q. Hu, L. Wang, Y. Piao, “Intracavity hyperthermia in nasopharingealcancer: a phase III clinical study” Int. J. Hyperthermia, 27(2), 180–186, 2011

[3] M.M. Paulides, J.F. Bakker, E. Neufeld, J. van der Zee, P.P. Jansen, P.C. Levendag, G.C. van Rhoon, “The HYPERcollar: a novel applicator for hyperthermia in the head and neck”, Int. J. Hyperthermia, 23(7), 567-76., 2007.

[4] M.M. Paulides, J.F. Bakker, M. Linthorst, J. van der Zee, Z. Rijnen, E. Neufeld, P.M. T. Pattynama, P.P. Jansen, P.C. Levendag and G.C. van Rhoon, “The clinical feasibility of deep hyperthermia treatment in the head and neck: new challenges for positioning and temperature measurement” Phys. Med. Biol. 55, 2465–2480, 2010

[5] P. Sminia, J. van der Zee, J. Wondergem, J. Haveman, “Effect of hyperthermia on the central nervous system” Int. J. Hyperthermia, 10(1), 1-30. 1994.

[6] J.A. Elder, “Ocular Effects of Radiofrequency Energy”, Bioelectromagnetics Supplement, 6, 148-161, 2003

[7] F. van der Lijn, T. den Heijer, M.M.B. Breteler, W.J. Niessen, “Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts” Neuroimage, 43, 708-720, 2008.

[8] X. Han, M. S. Hoogeman, P. C. Levendag, L. S. Hibbard, D. N. Teguh, P. Voet, A. C.Cowen, and T. K.Wolf, “Atlas-based auto-segmentation of head and neck ct images,” In MICCAI 2008.

[9] X. Han, L.S. Hibbard, N. O'Connell, and V. Willcut, "Automatic Segmentation of Parotids in Head and Neck CT Images using Multi-atlas Fusion", In MICCAI 2010.

[10] C.C. Teng, L. G. Shapiro, and I. Kalet, “Head and neck lymph node region with image registration”, BioMedical Engineering OnLine, 9:30, 2010

[11] S. Gorthi, V. Duay, N. Houhou, M. Bach Cuadra, U. Schick, M. Becker, A.S. Allal, J.P. Thiran, “Segmentation of head and neck lymph node regions for radiotherapy planning using active contour-based atlas registration”, IEEE J. on Selected Topics in Signal Processing, 3(1), 135–147, 2009.

[12] X. Han, L. Hibbard, N. O’Connel, and V. Willcut. “Automatic segmentation of head and neck CT images by GPU- accelerated multi-atlas fusion”. The MIDAS Journal, 2009.

[13] S. Gorthi, M. Bach Cuadra, U. Schick, P.-A. Tercier, A. S. Allal, and J.P. Thiran, “Multi-Atlas based Segmentation of Head and Neck CT Images using Active Contour Framework” In MICCAI, 2010.

[14] S. Warfield, K. Zou, W. Wells, “Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation”, IEEE Trans. Med. Imag., 23(7), 903–921, 2004

[15] B. Fischl, D.H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, A.M. Dale, “Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain.” Neuron, 33(3), 341–355, 2002.

[16] K.K Leung, J. Barnes, G.R. Ridgway, J.W.Bartlett, M.J. Clarkson, K Macdonald, N. Schuff, N.C. Fox, S. Ourselin, “Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer's disease”, Neuroimage 51(4), 1345-59, 2010

[17] Y. Boykov, O. Veksler, R. Zabih, “Fast approximate energy minimization via graph cuts”, IEEE Trans. Pattern Anal. Mach. Int. 23 (11), 1222–1239, 2001.

[18] S.Z. Li, “Markov Random Field Modeling in computer vision”, Springer-Verlag, 1994.

[19] S. Klein, M. Staring, K. Murphy, M. A. Viergever and J.P.W. Pluim, “elastix: a toolbox for intensity based image registration”, IEEE Trans. Med. Imag., 29, 196-205, 2010.

[20] Y. Boykov, V. Kolgomorov, “An experimental comparison of min-cut/max-flow algorithm for energy minimization in vision”, IEEE Trans. Pattern Anal. Mach. Int., 26(9), 1124-1137, 2004.