Modeling of Intensity Priors for Knowledge-Based Level Set Algorithm in Calvarial Tumors Segmentation

Aleksandra Popovic[1], Ting Wu[1], Martin Engelhardt[2], and Klaus Radermacher[1]

[1] Chair of Medical Engineering, Helmholtz-Institute for Biomedical Engineering,
RWTH Aachen University, Germany
`popovic@hia.rwth-aachen.de`
[2] Neurosurgical Clinic, Ruhr-University Bochum, Germany

Abstract. In this paper, an automatic knowledge-based framework for level set segmentation of 3D calvarial tumors from Computed Tomography images is presented. Calvarial tumors can be located in both soft and bone tissue, occupying wide range of image intensities, making automatic segmentation and computational modeling a challenging task. The objective of this study is to analyze and validate different approaches in intensity priors modeling with an attention to multiclass problems. One, two, and three class Gaussian mixture models and a discrete model are evaluated considering probability density modeling accuracy and segmentation outcome. Segmentation results were validated in comparison to manually segmented golden standards, using analysis in ROC (Receiver Operating Curve) space and Dice similarity coefficient.

1 Introduction

Segmentation of calvarial tumors in Computed Tomography (CT) images is an important research topic in the framework of the CRANIO project for computer and robot assisted craniotomy [1]. Calvarial tumors can be located in both soft and bone tissue, occupying wide range of image intensities, making automatic segmentation and computational modeling a challenging task. CT is the imaging modality of choice in the CRANIO project due to superiority Magnetic Resonance Imaging (MRI) in geometrical precision [2], required for intraoperative registration and robotic resection accuracy, and in depiction of the cranial bone involvement [3].

In our previous studies, the advantage of the level set approach in comparison with other methods has been demonstrated [4] followed by the extension of the level set framework with intensity priors [5]. We have demonstrated improvements in segmentation accuracy and convergence of level set propagation as prior knowledge about the intensity distribution of calvarial tumors is integrated into the level set speed function alongside image specific gradients. Monomodal Gaussian intensity modeling has been used.

The objective of this study is to analyze methods to model intensity distribution of calvarial tumors for establishing a *belief map*, the likelihood that an image element, i.e. voxel, belongs to the tumor class. The *belief map* is used as one of the

speed terms in the level set propagation and for seed points generation. Hence, it is an essential part of the automatic segmentation process presented here.

Gaussian mixture models (GMMs) are commonly used in the tissue classification problems under the hypothesis that each physical tissue type corresponds to one Gaussian distribution [6]. Expectation-Maximization algorithm is a standardly used method to predict parameters of the GMM. Touhami et al. [7] have proposed a discrete method to model prior intensities for knowledge-based segmentation of kidneys from 2D CT images, attractive for its computational simplicity in the contrast to the complexness of parameter optimization for the mixture models.

Intensity priors have already been used within the level set framework, utilizing different appearance of brain tumors in T1 and T2 weighted MRI images [8] or with trained mixture models [9,10].

In this paper the following open issues will be tackled:

- Which method for modeling of intensity priors is the most suitable for calvarial tumors?
- If Gaussian mixture model is used, how many classes should be assumed?
- What is the influence of this modeling on accuracy and robustness of the segmentation algorithm?

2   Methods

2.1   Intensity Priors

Let $I(\boldsymbol{x})$, $I : \mathbb{R}^3 \to \mathbb{R}$, $\boldsymbol{x} = [x, y, z]^T$, be a 3D gray-level CT image, and $S(\boldsymbol{x})$, $S : \mathbb{R}^3 \to \{0, 1\}$, the result of an expert manual segmentation, then $T(\boldsymbol{x})$, $I : \mathbb{R}^3 \to \mathbb{R}$, an extracted tumor gray-level region is defined as follows:

$$T(\boldsymbol{x}) = \begin{cases} I(\boldsymbol{x}), & \forall \boldsymbol{x} \quad | \quad S(\boldsymbol{x}) = 1 \\ 0, & \forall \boldsymbol{x} \quad | \quad S(\boldsymbol{x}) = 0 \end{cases} . \tag{1}$$

The objective is to model intensities of manually extracted tumor regions $T_i(\boldsymbol{x})$, $i = 1...N$, where N is the total number of extracted tumors in the training set. The Gaussian mixture model approach assumes that the intensities take values of n Gauss distributions, with probability density function (PDF):

$$PDF_g(t) = \sum_{i=0}^{n-1} \pi_i \cdot G_{\mu_i, \sigma_i}(t), \tag{2}$$

where π_i is a proportion of the i-th class and $G_{\mu, \sigma}(t)$ is a Gaussian distribution:

$$G_{\mu, \sigma}(t) = \frac{1}{\sigma \sqrt{2\pi}} \cdot e^{-\frac{t-\mu}{2\sigma^2}} . \tag{3}$$

The optimization objective is to find three parameters for each class, i.e. mean value, standard deviation, and proportion. Expectation Maximization (EM) algorithm can optimize likelihood of the distribution, but could fail to find the global maximum [11].

Important to notice is that the EM-GMM parameter prediction is performed upon the entire training set, treating it as a single sample, rather then finding parameter sets for each tumor and averaging them, as in our previous study [5].

An alternative approach is to discretely model distribution, as in [7]:

$$PDF_d(t) = \frac{1}{N} \sum_{i=0}^{N-1} \left(\frac{1}{k} \sum_{T_i(\boldsymbol{x})=1} \delta(T_i(\boldsymbol{x}) - t) \right), \tag{4}$$

where k is the number of voxels $\boldsymbol{x}$ so that $T_i(\boldsymbol{x}) = 1$ and δ is Dirac delta function. Obvious advantage of this approach is its forward computation without initialization problems.

Following computation of prior PDF from the training set, a *belief map* is generated for each image under investigation. In each voxel, $\boldsymbol{x}$, of the image, $I(\boldsymbol{x})$, assumed *belief* that the voxel is inside of the tumor region could be defined as follows:

$$B(\boldsymbol{x}) = PDF_{g,d}(I(\boldsymbol{x})). \tag{5}$$

Since probability densities for a given intensity are typically very small, the image obtained from Eq. 5 is rescaled to the intensity range $[-1, 1]$. Negative probability is the surface evolution self correcting term allowing negative values of the speed function (see section 2.2).

2.2   Level Set Algorithm

Level sets provide a computational and tracking framework for evolution of closed surface(s) along the perpendicular direction under a given velocity field. For image segmentation problems, the movement rule is specified with a spatially varying speed function, $F(\boldsymbol{x})$, dependant on the image features. The surface evolution is defined as:

$$\frac{\partial \Psi(\boldsymbol{x}, t)}{\partial t} = -\nabla \Psi(\boldsymbol{x}, t) \cdot F(\boldsymbol{x}), \tag{6}$$

where $\Psi(\boldsymbol{x}, t)$ is a 4D implicit surface model. Here, the following framework is used for the speed function:

$$F(\boldsymbol{x}) = F_p(\boldsymbol{x}) + F_c(\boldsymbol{x}) + F_a(\boldsymbol{x}), \tag{7}$$

$$F_p(\boldsymbol{x}) = \alpha \cdot (\delta \cdot B(\boldsymbol{x}) + P_I(\boldsymbol{x})), \ \ F_c = -\beta \cdot P_I(\boldsymbol{x}) \cdot \kappa(\boldsymbol{x}), \ \ F_a = \gamma \cdot \nabla Q(\boldsymbol{x}) \cdot \frac{|\nabla \Psi|}{\nabla \Psi}, \tag{8}$$

where $\kappa(\boldsymbol{x})$ is the curvature of the current front $(\nabla \frac{\nabla \Phi}{|\nabla \Phi|})$, P_I is an edge potential map and Q is the advection force, defined as:

$$P_I(\boldsymbol{x}) = e^{-|\nabla G * I(\boldsymbol{x})|}, \qquad Q(\boldsymbol{x}) = -|\nabla (G_\sigma * I(\boldsymbol{x}))|, \tag{9}$$

where G_σ is the Gaussian filter. F_p, F_c, F_a are propagation, curvature, and advection speed terms, respectively. Our prior knowledge specific contribution is the enhancement of the propagation term with the *belief map*, $B(\boldsymbol{x})$.

Initialization. The algorithm was initialized with the set of points rather than the contours. The advantage of level set propagation to evolve in separated fronts that could be later merged allows this kind of initialization. Voxels with prior PDF larger than a predefined threshold and prior PDF of all neighbors in the 2x2x2 neighborhood above the threshold were selected as seed points. The threshold was empirically set to 0.8. Level set distance map was initialized with the spheres (r=3) around seed points.

2.3   Experimental Set-Up

Five patient datasets were available $(A \div E)$, four thereof with diagnosed meningioma and one with skull metastasis. All patients were scanned with Siemens Somatom Plus CT, resolution: 0.43 x 0.43 x 3mm. Tumors were manually delineated by an expert. Performance estimation of learning models is done with leave-one-out cross-validation. The method results in four EM training sets.

To address the issues introduced in section 1 following experiments were conducted:

1. Assessment of the optimal number of Gaussian classes in Eq. 2.
2. Accuracy and robustness analysis of the segmentation algorithm for different PDF modeling approaches, Eq. 2 and Eq. 4.

An open source software system Insight Segmentation and Registration Toolkit [**?**] (version 2.4.1.) was used to implement all image processing and segmentation filters needed for this study.

3   Results

3.1   Intensity Priors

For each training sample, one, two, and three-modal GMMs have been optimized, i.e. GMM1, $\{(\mu_0^1, \sigma_0^1, \pi_0^1)\}$, GMM2, $\{(\mu_0^2, \sigma_0^2, \pi_0^2), (\mu_1^2, \sigma_1^2, \pi_1^2)\}$, GMM3, $\{(\mu_0^3, \sigma_0^3, \pi_0^3), (\mu_1^3, \sigma_1^3, \pi_1^3), (\mu_2^3, \sigma_2^3, \pi_2^3)\}$. More than 3 classes could not be found.

Quality of GMM estimation and the comparison between the models have been performed using Kolmogorov-Smirnov (KS) method in the CDF (Cumulative Distribution Function) domain. KS differences (difference of CDF values of original and estimated sample for all intensities) have shown that for more classes the GMM modeling is more realistic, Table 1, Fig. 1. The first class has been found around HU=80, i.e. soft tissue, second class around HU=200, i.e. trabecular bone, and third class around HU=800, i.e. cortical bone. The soft tissue portion was about 80% of the entire distribution.

In accordance to previously modeled intensity distributions, belief maps for each patient dataset have been generated, Fig. 2. In the case of one class GMM, it is obvious that there is no sharp distinction in soft tissue between tumorous and healthy tissue. This behavior is a consequence of the wide intensities range, raising standard deviation of the Gauss distribution. Less visible differences between other three models (GMM2, GMM3, and delta priors Eq. 4) have been further analyzed considering their influence on the segmentation accuracy.

Table 1. Maximal KS differences

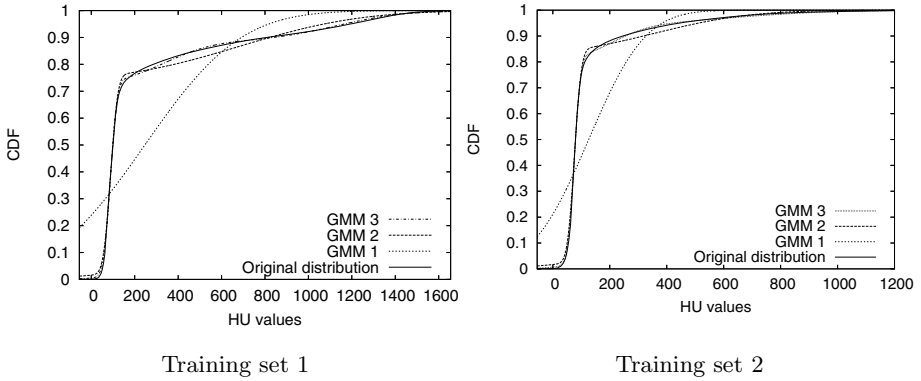|	Training set 1	Training set 2	Training set 3	Training set 4	Training set 5
GMM1	0.34	0.36	0.37	0.34	0.33
GMM2	0.036	0.043	0.042	0.038	0.046
GMM3	0.017	0.025	0.026	0.024	0.026

Training set 1

Training set 2

Fig. 1. Cumulative density functions of original image and three Gaussian mixture models for two different training sets

(A)   (B)

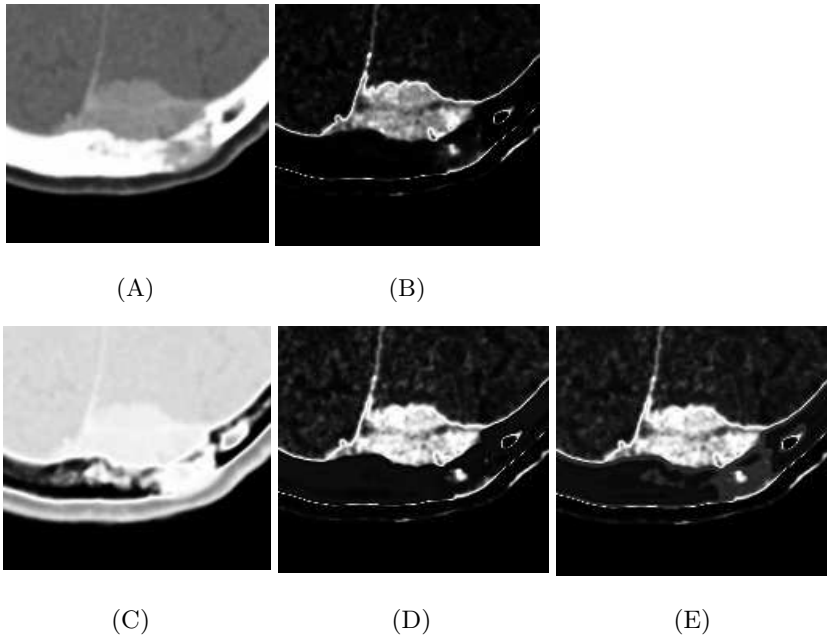(C)   (D)   (E)

Fig. 2. Belief maps generated for different priors modeling; (A) original image (B) delta priors, Eq. 4 (C) GMM with one class (D) GMM with two classes (E) GMM with three classes

3.2 Segmentation Accuracy

The segmentation accuracy has been assessed combining analysis in the Receiver Operating Curve (ROC) space and the spatial overlap metrics, Dice similarity coefficient (DSC), [12]. Level set propagation parameters $[\alpha, \beta, \gamma, \delta]$ range has been defined prior to leave-one-out procedure. The range has been selected wide enough to observe the best segmentation qualities for all cases in the middle of the parameter space.

To illustrate intensity priors influence, two characteristic patient datasets have been selected: one with substantial bone involvement (patient C, skull metastasis

Table 2. Mean and standard deviation of DSC metric for two characteristic patient datasets. Patient C is diagnosed with skull metastasis and patient E with meningioma, WHO grade II.

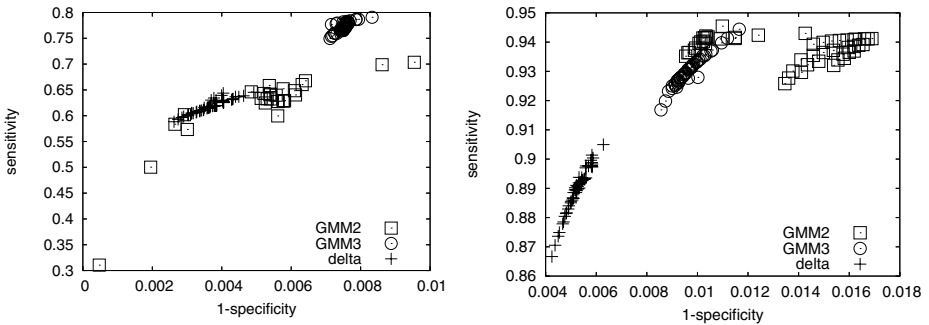| Patient	GMM2	GMM3	Delta
C	0.622 ± 0.1308	0.801 ± 0.0045	0.730 ± 0.0061
E	0.873 ± 0.0171	0.893 ± 0.0016	0.902 ± 0.0016

Fig. 3. Analysis of the segmentation accuracy in ROC space for 3 different approaches in generation of intensity priors

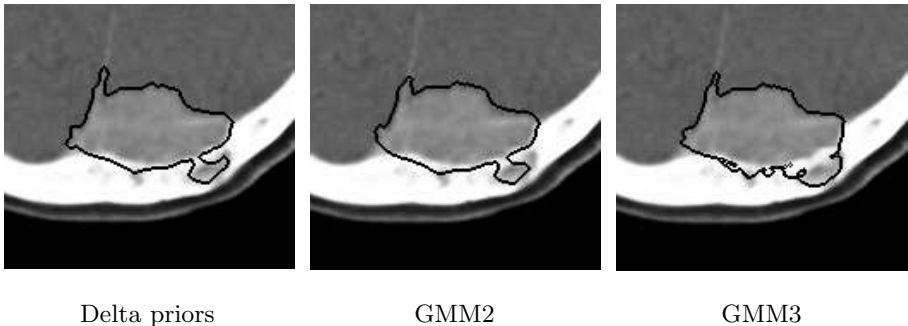Delta priors   GMM2   GMM3

Fig. 4. Example segmentation results for patient C

from mama carcinoma) and one with with minor bone involvement (patient E, meningioma, WHO grade II).

Table 2 shows results of DSC evaluation for these two patients. Figure 3 depicts accuracy results in the ROC space.

4   Discussion and Conclusion

Assessment of modeling accuracy with Gaussian mixture models, using Kolmogorov-Smirnov method, Table 1, has unambiguously shown that the most accurate model could be obtained by using three classes GMM. This statistical result, giving modeling accuracy only, has been further validated considering its influence on the segmentation algorithm outcome. Whereas in the case of the substantial bone involvement, GMM3 has significantly outperformed other methods, Table 2 and Fig. 3, in the case of the minor bone involvement, difference between delta priors and GMM is less obvious (mean DSC 0.89 and 0.9, respectively). This indicates that delta priors method performs effective in predominantly single class problems, e.g. for kidney segmentation [7]. GMM2 has shown inadequate robustness towards segmentation parameter change, Table 2.

Conclusively, GMM with three classes is the most suitable model, of those utilized in this study, for a complex multi-class problem of the modeling intensity priors of calvarial tumors. Segmentation accuracy of the automated algorithm presented here could be regarded as *very good* (DSC > 0.7) according to [13]. Our further work will focus on integration of the spatial information in the intensity prior modeling and level set segmentation framework proposed.

Acknowledgement

References

1. Popovic, A., Engelhardt, M., Wu, T., Portheine, F., Schmieder, K., Radermacher, K.: CRANIO - Computer Assisted Planning for Navigation and Robot-assisted Surgery on the Skull. In Lemke, H., Vannier, M., Inamura, K., Farman, A., Doi, K., Reiber, J., eds.: Proceedings of the 17th International Congress and Exhibition (CARS). Volume 1256 of International Congress Series., Elsevier (2003) 1269–1276
2. Grunert, P., anf J. Espinosa, K.D., Filippi, R.: Computer-aided Navigation in Neurosurgery. Neurosurg Rev **26** (2003) 73–99
3. Grover, S., Aggarwal, A., Uppal, P.S., Tandon, R.: The CT Triad of Malignacy in Meningioma - Redefinition, with a Report of Three New Cases. Neuroradiology **45** (2003) 799–803
4. Popovic, A., Engelhardt, M., Radermacher, K.: Segmentation of Skull-inflatrated Tumors Using ITK: Methods and Validation. In: ISC/NA-MIC/MICCAI Workshop on Open-Source Software at MICCAI 2005. (2005)

5. Popovic, A., Engelhardt, M., Radermacher, K.: Knowledge-based segmentation of calvarial tumors in computed tomography images. In: Bildverarbeitung für Medizin, BVM 2006. Informatik-Aktuell, Springer-Verlag (2006) 151–155
6. Ruf, A., Greenspan, H., Goldberger, J.: Tissue classiffcation of noisy mr brain images using constrained gmm. In Duncan, J., Gerig, G., eds.: Proceedings of Eight International Congerence on Medical Image Computing and Computer Assisted Intervention, MICCAI 2005, Lecture Notes on Computer Science. Volume 3750 of Lecture Notes on Computer Science., Springer-Verlag (2005) 790–797
7. Touhami, W., Boukerroui, D., Cocquerez, J.P.: Fully automatic kidneys detection in 2d ct images: A statistical approach. In Duncan, J., Gerig, G., eds.: Eight International Conference on Medical image Computing and Computer-Assisted Intervention, MICCAI. Volume 3750 of Lecture Notes on Computer Science., Springer (2005) 262–269
8. Ho, S., Bullitt, E., Gerig, G.: Level Set Evolution with Region Competition: Automatic 3-D Segmentation of Brain Tumors. In: Proceedings of 16th International Conference on Pattern Recognition (ICPR). Volume 1. (2002) 532–535
9. Colliot, O., Mansi, T., Bernasconi, N., Naessens, V., Klironomos, D., Bernasconi, A.: Segmentation of focal cortical dysplasia lesions using a feature-based level set. In Duncan, J., Gerig, G., eds.: Eight International Conference on Medical image Computing and Computer-Assisted Intervention, MICCAI. Volume 3749 of Lecture Notes on Computer Science., Springer (2005) 375–382
10. Leventon, M., Grimson, E., Faugeras, O., Kikinis, R., III, W.W.: Knowledge-based Segmentation of Medical Images. In: Geometric Level Set Methods in Imaging, Vision, and Graphics. Springer-Verlag (2003) 401–420
11. Marin, J.M., Mengersen, K., Robert, C.P.: Chapter 16: Bayesian modelling and inference on mixtures of distributions. Number 25. In: HANDBOOK OF STATISTICS. Elsevier (2005)
12. : Statistical Validation of Image Segmntation Quality Based on a Spatial Overlap Index. Acad Radiol **11** (2004) 178–189
13. Zijdenbos, A., Dawant, B., Margolin, R., Palmer, A.: Morphometric Analysis of White Matter Lesions in MR Images: Method and Validation. IEEE Transactions on Medical Imaging **13** (1994) 716–724