An Iterative Classification Method of 2D CT Head Data Based on Statistical and Spatial Information

Feng Li
IGST, Shanghai Jiaotong University
windman021@hotmail.com

Dirk Bartz
ICCAS, University of Leipzig
dirk.bartz@iccas.de

Lixu Gu
IGST, Shanghai Jiaotong University
gu-lx@cs.sjtu.edu.cn

Michel Audette
ICCAS, University of Leipzig
michel.audette@medizin.uni-leipzig.de

Abstract

An iterative classification method developed for 2D CT head data classification problem and using both statistical and spatial information is introduced in this paper. The method reduces the chance of misclassification, preserving the contiguity of tissue classes. This method is minimally supervised so that it enforces a relation between tissues and classes. In later iterations high-confidence points are used to help classify nearby ambiguous points, based on the assumption that points in close proximity and of comparable intensities are probably representing the same tissue class.

1. Introduction

The classification of CT head data has many useful applications in the planning and simulation of ENT and neuro-surgical interventions. Our objective is a method that yields a clinically useful tissue map, in which air, soft tissues, bone of both high and low density, as well as possible surgical landmarks are distinguished. For a method to be successful, it must overcome the ambiguities caused by image noise, partial volume effects, as well as overlap in intensities between different tissue classes. Moreover, any assumption used to alleviate these ambiguities must hold for typical patient data. One such assumption is that tissue is in general contiguous. A single-pass classification method based on solely statistical information will in general not produce contiguous tissues. In contrast, an iterative classification can benefit from blobs of high-confidence points that influence neighboring ambiguous points that are in close proximity and of comparable intensity.

In this paper, an iterative classification method, adapted from Bayesian Level Sets [6] and incorporating both statistical and spatial information, is introduced. One difficulty of such a method is a suitable choice of high-confidence points in the first iteration. Our solution is a minimally supervised method using a small number of training points together with a large number of unlabeled points. The method aims to reduce the chance of misclassification, preserve the contiguity of tissue classes, and retain small and narrow structures.

Thereafter, distances between high-confidence points and ambiguous points are computed and the 1D problem is cast as an iterative classification in N+1-dimensional feature space consisting of both intensity and N distances to nearest blobs of high-confidence points, for an N-class problem. New high-confidence points are selected at the end of each iteration due to a membership function based on both intensity and distances and the rest ambiguous points become closer to the high-confidence points. The process ends when there are no ambiguous points left.

2. High-Confidence Points Initialization

2.1 Minimally Supervised Expectation Maximization of Gaussian Mixture Model

In the first iteration of our iterative classification method, high-confidence (HC) points are selected for each tissue class. The choice of HC points should be of sufficiently high density to be useful thereafter, while not penalizing small or thin tissue structures, and still excluding misclassified points of each tissue class. The consideration of the CT histogram as a Gaussian Mixture Model[1], whose modes coincide with relevant tissues, suggests an interpretation that leads to a definition

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

of HC points. Combined with a carefully chosen size threshold applied through a component labeling algorithm, the method leads to an intuitive and robust definition of HC points in the first iteration.

Firstly, we focus on the intensity histogram of the given images. In practice, due to the inhomogeneity of the tissues, the occurrence of image noise, partial volume effect or overlap between different tissue classes, the histogram looks like a combination of N Gaussian lobes for an N-class problem. The density function of a 1D Gaussian Mixture Model has such a form:

$$f(g|\theta) = \sum_{k=1}^{N} \alpha_k p(g|\mu_k, \sigma_k) \quad (1)$$

$$p(g|\mu_k, \sigma_k) = (2\pi\sigma_k)^{-1/2} e^{-\frac{1}{2\sigma_k}(g-\mu_k)^2}, \quad (2)$$

where g denotes a grey-level of the histogram. The estimation of the parameters θ consisiting of average intensities μ_k, standard deviations σ_k and mixing proportions α_k can be solved by using training points or the Expectation Maximization (EM) algorithm[2, 1]. A small number of training points is available in our method. The parameters μ_k and σ_k are initialized from the training points, while the parameter α_k is initialized from the amplitudes and areas of the relevant lobes. The EM algorithm will then use the preliminary estimation as an initialization and iteratively refine the estimation to make it more precise.(Fig. 1(a))

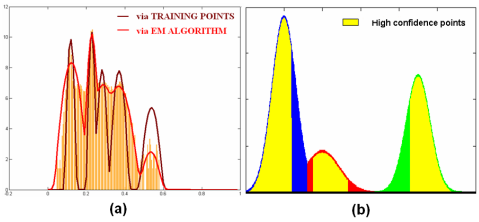

Figure 1. GMM & HC Points Selection

The EM algorithm refines the estimation by iteratively solving a maximum-likelihood estimation problem. In each iteration there are two steps: the E(xpectation)-step and the M(aximization)-step. Let set $G = \{g\}$ denote the intensity of points and set $Y = \{y\}$ denote the relationship between points and tissue classes, i.e. $y_i = c$ implies that point i belongs to tissue class c. In the E-step, the expected value Q of the log-likelihood $\log p(G, Y|\theta)$ with respect to set Y given the intensity set G and the current estimation of the parameters θ is computed. Thus, we have

$$Q(\theta, \theta_{t-1}) = E_Y[\log \prod_{i=1}^{M} p(g_i, Y|\theta)|g_i, \theta_{t-1}] \quad (3)$$

$$s.t. \sum_{k=1}^{N} p(y_i = k|\theta) = 1, \quad (4)$$

where M is the number of the points in the CT image, while N is the number of tissue classes.

The M-step aims to find the parameter estimation that leads to the maximum expected value $Q(\theta, \theta_{t-1})$. That is

$$\theta_t = \arg\max_{\theta} Q(\theta, \theta_{t-1}), \quad (5)$$

A numerical method for the estimation is given in [1].

2.2 High-Confidence Points Selection

High-confidence points are those points which have the least chance to be misclassified for each class. Two thresholds are used for HC points selection, a statistical intensity based threshold and a size threshold for contiguous blobs. The statistical threshold is based on the GMM model and the size threshold is then applied on the result of using the previous threshold.

For each individual Gaussian component in the GMM model, we first choose the points with statistically the least chance to be misclassified. For the left-most Gaussian distribution, the intensities in the left-most lobe are selected, since very few points from other distributions fall in that area. For the same reason, the intensities in the right-most lobe of the right-most Gaussian distribution are selected. For the middle Gaussian distributions, intensities near the peaks are selected. (Fig. 1(b))

When the image is very noisy or the intensity ranges of different classes overlap significantly with each other, it is still possible that some misclassified points are selected as HC points by using the statistical threshold. This problem can not always be solved by raising the threshold of probability. A size threshold is introduced to solve this problem. According to the assumption of tissue contiguity, the misclassified points remaining after the statistical threshold are more likely to be separated. A size threshold, which discards the blobs of smaller size than the given threshold, can get rid of those misclassified points. The component labeling[4] algorithm is used to compute the size of blobs.

3. Iterative Classification

After the first iteration, we obtain the HC points set for each tissue class. In the subsequent iterations, the HC points are used to affect the nearby ambiguous points according to intensity similarity and distances. The distances are computed through the Fast Marching method[5], a numerical algorithm for simulating a monotonically propagating front by solving the Eikonal Equation $F|\nabla T| = 1$ with emphasis on efficiency.

In the Eikonal Equation, F indicates the speed at which a 2D front passes a point, while T stands for the time that the front arrives at a point. The Fast Marching

method starts with *an initial contour coinciding here with the boundary of HC blobs*. A speed image stores the speed for every point in the image. The result of the Fast Marching method is an Arrival Time map, while the arrival time of the points on the initial contour is set as zero. When we set the speed function to be related to the intensity similarity between an ambiguous point and HC points, the arrival time is impacted by two factors: the intensities of the points and Euclidean distances from HC blobs of each class to each ambiguous point. In other words, *the arrival time indicates an intensity-weighted distance instead of the Euclidean distance*.

In an N-class problem, we set N speed images, one for each tissue class. The speed function is set as

$$F(\boldsymbol{x}, c) = \frac{1 + \min_k (I(\boldsymbol{x}) - \overline{I(k)})}{1 + |I(\boldsymbol{x}) - \overline{I(c)}|}, \quad (6)$$

where $\boldsymbol{x}$ denotes the point, c indicates a tissue class and $\overline{I(c)}$ is the average intensity of the HC points in class c. The function reaches its maximum, when $\overline{I(c)}$ is closest to $I(\boldsymbol{x})$. The other N-1 speeds at point $\boldsymbol{x}$ will be smaller. The larger the difference between an $\overline{I(c)}$ and the $I(\boldsymbol{x})$ is, the slower the speed is.

An improvement for computing the distances is made according to such a consideration that *blobs consisting of HC points of other classes should not be passed through when computing distances*, which means that the distance computation should not cross the HC blobs of the other tissue classes. The intention of this consideration is to preserve the contiguity of tissues.

After getting the distances between HC points and ambiguous points, we try to find the new HC points. This time we form a membership function for each point with both intensity and distance as parameters. The reason for using membership functions instead of using a multivariate GMM-EM method again is that the distributions of distances are usually not Gaussian distributions. The general form of the function is

$$U(\boldsymbol{x}, c) = \beta U_I(\boldsymbol{x}, c) + (1 - \beta) U_D(\boldsymbol{x}, c) \qquad (0 \leq \beta \leq 1) \quad (7)$$

$$U_I(\boldsymbol{x}, c) = \frac{(1 + |I(\boldsymbol{x}) - \overline{I(c)}|)^{-1}}{\sum_{k=1}^{N} (1 + |I(\boldsymbol{x}) - \overline{I_k}|)^{-1}} \quad (8)$$

$$U_D(\boldsymbol{x}, c) = \frac{(1 + D(\boldsymbol{x}, c))^{-1}}{\sum_{k=1}^{N} (1 + D(\boldsymbol{x}, k))^{-1}} \quad (9)$$

$U_I(\boldsymbol{x}, c)$ implements the influence of the intensity of the point. $I(\boldsymbol{x})$ denotes the intensity of point $\boldsymbol{x}$ and $\overline{I(c)}$ denotes the average intensity of tissue class c, while N is the number of classes. $|I(\boldsymbol{x}) - \overline{I(c)}|$ reaches its minimum and $U_I(\boldsymbol{x}, c)$ reaches its maximum, when c denotes the tissue class with the closest average intensity to the intensity of point $\boldsymbol{x}$.

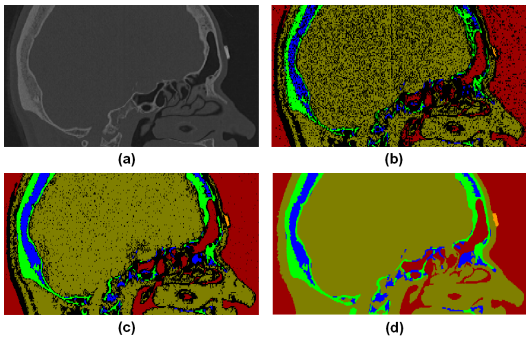

Figure 2. Iteration Process *(a) Original CT data; (b) HC points in the first iteration; (c) HC points in a middle iteration; (d) Final result*

$U_D(\boldsymbol{x}, c)$ implements the effect of the distances. $D(\boldsymbol{x}, c)$ denotes the distance between point $\boldsymbol{x}$ and the nearest HC blob of tissue class c. $D(\boldsymbol{x}, c)$ reaches its minimum and $U_D(\boldsymbol{x}, c)$ reaches its maximum when c represents the tissue class which is the closest in proximity to point $\boldsymbol{x}$.

A size threshold is also needed here to eliminate the small blobs and preserve tissue contiguity. This method is preferred to morphological erosion, which would unduly wipe out thin bone structures.

4. Results & Discussions

Fig. 2 illustrates the iteration process. Figure 2(b) shows the HC points selected by the method in Section 2. New HC points are then iteratively selected with the method in Section 3 and figure 2(c) illustrates one of the middle results. Figure 2 (d) is the result of our method.

Fig. 3 illustrates the evolution of the six features of each point (intensity and five distances to five tissue classes) during the iteration process with a parallel coordinate plot[3]. Each point is converted to a line which intersects with the six axes (one axis for one feature) according to the corresponding feature values. A distance value of -1 means the point is a HC points of the tissue class. The white lines show the ambiguous points, while the color lines shows the HC points. The number of white lines reduces and the distribution of white lines changes during the process (from top to bottom).

At first, the white lines cover most part on the five distance axes. During the iterations, the intersection points on each distance axis move in two directions: moving up means that the point is less probably representing the tissue class of that axis, while moving down means that the point is more probably representing that tissue class. This motion of the intersection points reflects the classification becoming unambiguous.

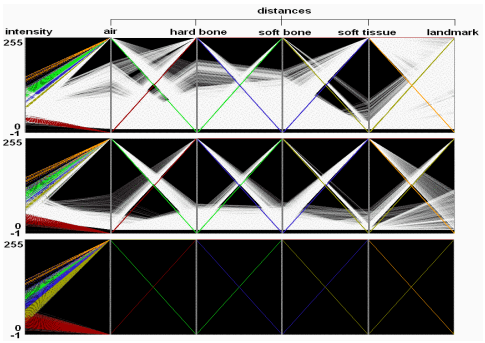

Figure 3. The Evolution of the 6 Features: *From top to bottom: iteration 1; middle iteration; final result*

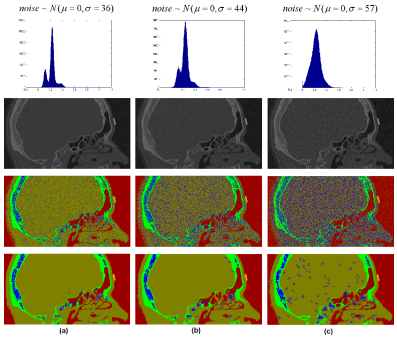

Figure 4. Comparison (with & without Spatial Information) *(a)*$\sigma = 36$*, (b)*$\sigma = 44$*, (c)*$\sigma = 57$*; row 1: histogram; row 2: CT data with noises; row 3: Single stage GMM-EM method; row 4: Our iterative method*

Table 1: Sensitivity and Specitivity of the Results

	ass.	air	bone (hard)	bone (soft)	soft tissue	landmark
$\sigma = 0$ (S)	se	0.958	0.802	0.588	0.992	0.872
$\sigma = 0$ (S)	sp	0.996	0.988	0.994	0.918	1.000
$\sigma = 0$ (I)	se	0.969	0.779	0.820	0.992	0.876
$\sigma = 0$ (I)	sp	0.994	0.996	0.990	0.938	1.000
$\sigma = 36$ (S)	se	0.936	0.729	0.498	0.917	0.533
$\sigma = 36$ (S)	sp	0.990	0.989	0.938	0.902	1.000
$\sigma = 36$ (I)	se	0.959	0.802	0.589	0.992	0.873
$\sigma = 36$ (I)	sp	0.996	0.988	0.994	0.918	1.000
$\sigma = 44$ (S)	se	0.911	0.689	0.503	0.751	0.565
$\sigma = 44$ (S)	sp	0.961	0.986	0.848	0.893	0.993
$\sigma = 44$ (I)	se	0.960	0.810	0.578	0.988	0.899
$\sigma = 44$ (I)	sp	0.995	0.987	0.995	0.917	1.000
$\sigma = 57$ (S)	se	0.839	0.560	0.437	0.530	0.624
$\sigma = 57$ (S)	sp	0.870	0.961	0.786	0.848	0.994
$\sigma = 57$ (I)	se	0.957	0.798	0.572	0.963	0.869
$\sigma = 57$ (I)	sp	0.990	0.985	0.980	0.921	1.000

S = single-stage method; I = iterative method; se = sensitivity; sp = specitivity;

In Fig. 4, artificial Gaussian noise is added to the original CT data to illustrate the robustness of our method to noise. Comparing to the single stage GMM-EM method, our method is less sensitive to the noise and basically preserves tissue contiguity. The assessment of sensitivity and specitivity is listed in table 1 by comparing our results to a manually segmented result. The results from the experiment on some other test data are showed in Fig. 5.

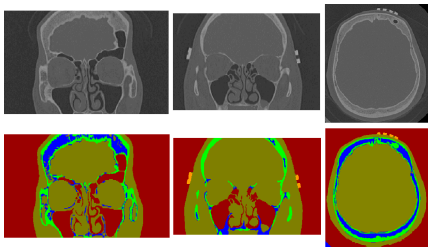

Figure 5. Results on Other Test Data

5. Conclusion

In this paper, we have developped an iterative classification method for 2D CT data. The method uses both statistical and spatial information to reduce the chance of misclassification, maintain tissue contiguity and retain the narrow structures. It improves the Bayesian Level Sets method through greater relevance to tissues, in the context of a minimally supervised classification. The method will in the future be extended to 3D CT data based on 3D Fast Marching method.

References

[1] J. A. Bilmes. A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models. *Technical Report, University of Berkeley, ICSI-TR-97-021*, 1998.

[2] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38, 1977.

[3] A. Inselberg and B. Dimsdale. Parallel coordinates: a tool for visualizing multi-dimensionalgeometry. *Proc. of IEEE Visualization*, 1990.

[4] A. Rosenfeld and J. L. Pfaltz. Sequential operations in digital picture processing. *J. ACM*, 13(4):471–494, 1966.

[5] J. A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proc. National Academy of Sciences*, 93(4):1591–1595, 1996.

[6] E. Sifakis, C. Garcia, and G. Tziritas. Bayesian level sets for image segmentation. *Visual Communication and Image Representation*, 2001.