Soft Segmentation of CT Brain Data

Alexandra Lauric and Sarah Frisken

Tufts University, Halligan Hall Room 102, 161 College Ave, Medford MA 02155, USA
alauri02,frisken@cs.tufts.edu

Abstract. Unlike research on brain segmentation of Magnetic Resonance Imaging (MRI) data, research on Computed Tomography (CT) brain segmentation is relatively scarce. Because MRI is better at differentiating soft tissue, it is generally preferred over CT for brain imaging. However, in some circumstances, MRI is contraindicated and alternative scanning methods need to be used. We have begun to explore methods for soft tissue segmentation of CT brain data with a goal of enhancing the utility of CT for brain imaging. In this study, we consider the effectiveness of existing algorithms for segmenting brain tissue in CT images. Three methods (Bayesian classification, Fuzzy c-Means and Expectation Maximization) were used to segment brain and cerebrospinal fluid. While these methods outperformed the commonly used threshold-based segmentation, our results show the need for developing new imaging protocols for optimizing CT imaging to differentiate soft tissue detail and for designing segmentation methods tailored for CT.

Key words: Computed tomography, brain imaging, soft segmentation

1 Introduction

Image segmentation is the process of assigning labels to individual pixels in a volumetric data set, based on some criteria in the image. In medical image segmentation, pixels are labeled by tissue type. Image segmentation is required in many medical applications ranging from the education and assessment of medical students to image-guided surgery and surgical simulation.

Appropriate segmentation methods are highly dependent on image acquisition modality and the tissue of interest. For example, bone can be segmented in CT data using simple thresholding techniques because of the high contrast between bone and surrounding tissues. In contrast, soft tissues are not well differentiated in CT images and thresholding is inadequate.

Because it provides superior contrast of soft tissue structures (Fig. 1), MRI is the method of choice for imaging the brain and most research on brain segmentation focuses on MRI. There are, however, numerous situations when MRI is contraindicated and other alternatives, usually CT, need to be explored [1]. These situations include patients who are too large for the MRI scanner, claustrophobic patients, patients with metallic or electrical implants and patients unable to remain motionless for the duration of the examination due to age, pain, or

medical condition. Additional advantages of using CT include good detection of calcification, hemorrhage and bony detail, plus lower cost, short imaging times, and widespread availability. For these reasons, we have begun to explore methods for segmenting soft tissue in CT images. As a first step, we have evaluated the effectiveness of several segmentation methods developed for MRI for segmenting brain tissue in CT images.

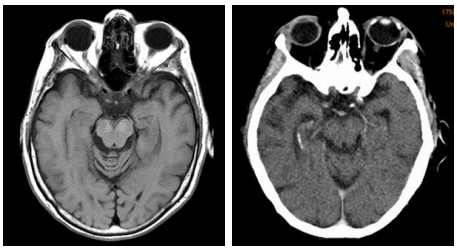

Fig. 1. Comparison between MRI(left) and CT(right) images of the brain. Gray and white matter are clearly distinguishable in the MRI image, but not in the CT image, while bone is easy to detect in the CT image.

In this paper, we compare three soft segmentation methods applied to CT images of the brain. These methods (Bayesian classification, Fuzzy c-Means, and Expectation Maximization) have been used on MRI data with good results, and are representative of the class of soft segmentation algorithms. Soft segmentation allows each pixel to belong to multiple classes, with varying degrees of membership. More common segmentation methods use hard segmentation where a pixel is assigned exclusively to one class. Hard decisions about the membership of a pixel involve throwing away some of the information contained in the data; soft segmentation methods have the advantage of avoiding these decisions until later in the process, thereby keeping more options available for post-processing steps. Soft segmentation can be converted to hard segmentation by using the maximum membership classification rule. Under this rule a pixel is assigned to the class with which it has the highest membership value.

The first method presented is based on Bayesian classification. The method uses Bayes rule to predict class membership as a conditional probability that a given pixel belongs to a particular class.

The second segmentation method is based on Fuzzy c-Means (FCM) clustering. The method alternates between partitioning the pixels into clusters and updating the parameters of each cluster. Like many clustering algorithms, FCM favors large clusters and, as a result, pixels belonging to small clusters are often misclassified. To compensate for this behavior, we use the Population-Diameter Independent (PDI) algorithm, which was introduced in [2] as a variation of FCM. The PDI algorithm uses cluster weights to balance the influence of large clusters.

The third segmentation method is the Expectation Maximization (EM) algorithm, which partitions pixels into clusters by determining the maximum likelihood parameters of a mixture of known distributions.

All three methods perform segmentation by constructing statistical models but they have complimentary strengths and limitations. Bayesian classification is simple, fast and robust, but it requires training and is sensitive to the accuracy of training data. FCM is an efficient, self-organizing clustering algorithm which does not require training. However, unlike the Bayesian classifier, FCM does not explicitly incorporate prior knowledge. Expectation Maximization combines the strengths of both algorithms: it is based on Bayes rule, so it incorporates prior knowledge, but it is an unsupervised segmentation method, so it does not require any training.

2 Related Work

The literature on brain segmentation from MRI data is extensive and soft segmentation methods are well researched. In [3], Spence gives an overview of the research on MRI brain segmentation. We will only refer here to work on MRI that uses one of the three methods presented in this paper.

Laidlaw et al. [4] used Bayesian classification to identify the distribution of different materials in MRI volumetric datasets of the brain. In this approach, voxels are allowed to contain more than one material and information from neighboring pixels is incorporated into the classification process.

The FCM algorithm was used by Li et al. [5] to segment MR brain images. Pham and Prince [6] extended the traditional FCM to address the inhomogeneities due to the MR acquisition process. PDI was introduced by Shihab in [2]. Shihab proposed a new objective function and normalized each cluster contribution by dividing it by a value representing the strength of its contribution.

EM segmentation for MR brain images was first proposed by Wells et al. in [7]. Wells used EM algorithm to simultaneously estimate the classification of the image and a corrupting bias field characteristic to MR imaging. Kapur [8] introduced a variation of the EM algorithm which replaced the statistically independent priors from classic EM with a Gibbs prior, thereby incorporating spatial information into the segmentation process.

Unlike literature on brain segmentation for MRI, the literature on brain segmentation on CT data is relatively sparse. The most common approaches use active contours [9], thresholding and region growing [10]. However all of these methods require manual input.

In [11], Hu et al. segmented brain matter from 3D CT images by first applying Fuzzy C-means and thresholding to 2D slices of the volume to create 2D masks, and then propagating the 2D masks between neighboring slices.

A method for automatic CT brain segmentation of intracerebral brain hemorrhage (ICH) is proposed by Loncaric and Kovacevic in [12]. Five regions of interest (background, skull, brain, ICH and calcifications) were segmented using a combination of K-Means clustering and neural networks.

3 Classification

Classification is a supervised learning technique which seeks to partition a feature space into groups (classes). Classifiers are trained on labeled (pre-classified) data and then applied to samples with unknown labels. The pre-classified data is used to learn the characteristics of the classes which are in turn used to label a new data sample.

In image segmentation, classification is used to label each image pixel. Pixels are described by a set of features (i.e., a feature vector) such as a pixel's intensity, the average intensity of its neighbors, or its spatial relationship to known structures. The pre-classified data is usually obtained by manually segmenting related image data.

3.1 Bayesian Classification

Bayesian classification is a probabilistic algorithm based on Bayes rule:

$$P(A|B) = \frac{P(B|A)\,P(B)}{P(A)} \tag{1}$$

In image segmentation, Bayes rule can be applied to determine the a posteriori conditional probability for each class for a given pixel. In mathematical terms, Bayes rule for pixel i, with feature vector $\mathbf{x}_i$, and class k is

$$P(k|\mathbf{x}_i) = \frac{P(\mathbf{x}_i|k)\,P(k)}{P(\mathbf{x}_i)}. \tag{2}$$

$P(k|\mathbf{x}_i)$ is the a posteriori conditional probability that pixel i is a member of class k, given its feature vector $\mathbf{x}_i$. $P(k|\mathbf{x}_i)$ gives the membership value of pixel i to class k, where k takes values between 1 and c (the number of classes), while i takes values between 0 and N-1 (the number of pixels in the image). Note that $\sum_{j=1}^{c} P(j|\mathbf{x}_i) = 1$.

$P(\mathbf{x}_i|k)$ is the conditional density distribution, i.e., the probability that pixel i has feature vector x_i given that it belongs to class k. If the feature vectors of class k have a Gaussian distribution, the conditional density function has the form

$$P(\mathbf{x}_i|k) = \frac{1}{(2\pi)^{\frac{n}{2}}\, det^{\frac{1}{2}}\,(\Sigma_k)} e^{-\frac{1}{2}(\mathbf{x}_i-\mathbf{p}_k)^T \Sigma_k (\mathbf{x}_i-\mathbf{p}_k)}, \tag{3}$$

where $\mathbf{p}_k$ and Σ_k are the mean feature vector and the covariance matrix of class k. The mean and the covariance of each class are estimated from training data. n is the dimension of the feature vector.

The prior probability of class k is

$$P(k) = \frac{|\omega_k|}{\sum_{j=1}^{c} |\omega_j|}, \tag{4}$$

where $|\omega_k|$ is a measure of the frequency of occurrence of class k and $\sum_{j=1}^{c} |\omega_j|$ is a measure of the total occurrence of all classes. In image segmentation, $|\omega_k|$ is

usually set to the number of pixels which belong to class k in the training data, and $\sum_{j=1}^{c} |\omega_j|$ to the total number of pixels in the training data.

$P(\mathbf{x}_i)$ is the marginal likelihood of pixel i which measures the likelihood that feature vector $\mathbf{x}_i$ will occur in the dataset. This term is used as a normalizing factor according to formula

$$P(\mathbf{x}_i) = \sum_{j=1}^{c} P(\mathbf{x}_i|j) P(j). \tag{5}$$

The output of Bayesian classification is a soft segmentation of the input image. To convert from soft to hard segmentation, the label of pixel i is decided using the maximum a posteriori (MAP) rule: $class(i) = k$ such that $P(\mathbf{x}_i|k) = \max_j P(\mathbf{x}_i|j)$.

4 Clustering

Clustering is an unsupervised classification of data samples into groups (clusters). While classifiers are trained on pre-labeled data and tested on unlabeled data, clustering algorithms take as input a set of unlabeled samples and organize them into clusters based on similarity criteria. The algorithm alternates between dividing the data into clusters and learning the characteristics of each cluster using the current division [13].

In image segmentation, a clustering algorithm iteratively computes the characteristics of each cluster (e.g. mean, standard deviation) and segments the image by classifying each pixel in the "closest" cluster according to a distance metric. The algorithm alternates between the two steps until convergence is achieved or a maximum number of iterations is reached.

4.1 Fuzzy C-Means

Fuzzy clustering algorithms partition an input data set into fuzzy clusters, such that each element in the data can belong to multiple clusters. Fuzzy c-means (FMC) [14] is such an algorithm which provides a method to segment the pixels into c fuzzy clusters, while simultaneously updating the locations of these clusters and the degrees of membership associated with each pixel. Intuitively, the degree to which a pixel belongs to a cluster is inversely proportional to the distance from the pixel to the center of that cluster. The objective function used to optimally partition the data is:

$$Minimize: J_{FCM} = \sum_{k=1}^{c} \sum_{i=1}^{N} (u_{ik})^m d_{ik}^2 (\mathbf{x}_i, \mathbf{p}_k) \tag{6}$$

$$Subject\ to: \sum_{k=1}^{c} u_{ik} = 1, \forall i \in \{0..N-1\}, \tag{7}$$

where c is the number of clusters, N is the number of pixels in the image, $m \geq 1$ is an exponent controlling how fuzzy the result should be, $\mathbf{x}_i$ is the feature vector associated with pixel i, u_{ik} is the membership value for pixel i with cluster k, $\mathbf{p}_k$ is the mean vector of cluster k and d_{ik} is the distance from $\mathbf{x}_i$ to $\mathbf{p}_k$.

The distance metric is defined as

$$d_{ik}^2\left(\mathbf{x}_i, \mathbf{p}_k\right) = \left\|\mathbf{x}_i - \mathbf{p}_k\right\|_A^2 = \left(\mathbf{x}_i - \mathbf{p}_k\right)^T A \left(\mathbf{x}_i - \mathbf{p}_k\right). \tag{8}$$

A is a positive definite matrix. If A is the identity matrix, d_{ik} is the Euclidian distance. In this work, A is the covariance matrix, Σ_k, so that d_{ik} is the Mahalanobis distance [13].

In order to minimize J_{FCM}, we use Lagrange multipliers to update estimates of $\mathbf{p}_k$, Σ_k and u_{ik} according to:

$$u_{ik}^{t+1} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}^2}{d_{ij}^2}\right)^{\frac{1}{(m-1)}}}, \tag{9}$$

$$\mathbf{p}_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik}^m \mathbf{x}_i}{\sum_{i=0}^{N-1} u_{ik}^m}, \text{and} \tag{10}$$

$$\Sigma_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik} \left(\mathbf{x}_i - \mathbf{p}_k^{t+1}\right)\left(\mathbf{x}_i - \mathbf{p}_k^{t+1}\right)}{\sum_{i=0}^{N-1} u_{ik}}. \tag{11}$$

FCM iterates through the equations (9)-(11) until convergence is achieved. The steps of the algorithm are:

1. Fix c to the number of classes and m typically to 2. Initialize the mean vector $\mathbf{p}_k$ and the covariance matrix Σ_k for each class to a random vector and the identity matrix respectively. For each pixel initialize the membership values u_{ik} to $\frac{1}{c}$.
2. Update membership values u_{ik}, weighted means $\mathbf{p}_k$, and covariance matrices Σ_k, using (9)-(11).
3. Compare the change in the membership values between times t and $t+1$. If $\max_{ik}\{|u_{ik}^{t+1} - u_{ik}^t|\} \leq \epsilon$ stop. Otherwise, repeat steps 2 to 5.

The strengths and weaknesses of FCM are analyzed in detail in [2] with the conclusion that the standard objective function is suboptimal when clusters are close to each other, but differ in size and diameter. The size of a cluster is defined in terms of its population, while its diameter is defined as the diameter of a hyper-sphere which contains the entire cluster. FCM favors clusters with large sizes and/or large diameters, resulting in misclassification of pixels belonging to small clusters (Fig. 2).

To compensate for these shortcomings in FCM, Shihab [2] proposed the Population-Diameter Independent algorithm (PDI) which uses a new objective function in which each cluster contribution is normalized and divided by a number ρ_k representing the strength of its contribution. If the clusters make equal contributions to the objective function, the normalizer ρ_k is set to $\frac{1}{c}$. Otherwise

clusters with small contributions are assigned small normalizers and clusters with large contributions are assigned large normalizers, thereby increasing the contribution of smaller clusters and decreasing the contribution of larger clusters. The new optimization function is:

$$Minimize : J_{FCM} = \sum_{k=1}^{c} \frac{1}{\rho_k^r} \sum_{i=1}^{N} (u_{ik})^m \, d_{ik}^2 \, (\mathbf{x}_i, \mathbf{p}_k) \tag{12}$$

$$Subject\ to : \sum_{k=1}^{c} u_{ik} = 1, \forall i \in \{0..N-1\}, \text{and} \tag{13}$$

$$\sum_{k=1}^{c} \rho_k = 1, \tag{14}$$

where ρ_k is the normalizer for cluster k and r allows control over the influence of ρ. When r is set to 0, PDI collapses to FCM. The updates for u_{ik}, $\mathbf{p}_k$, Σ_k and ρ_k are derived, using Lagrange multipliers, as follows:

$$u_{ik}^{t+1} = \frac{\rho_k^r}{\sum_{j=1}^{c} \left(\frac{d_{ik}^2}{d_{ij}^2} \right)^{\frac{\rho_j^r}{(m-1)}}}, \tag{15}$$

$$\rho_k^{t+1} = \frac{\left(\sum_{i=0}^{N-1} u_{ik}^m d_{ik}^2 \right)^{\frac{1}{r+1}}}{\sum_{j=1}^{c} \left(\sum_{i=0}^{N-1} u_{ij}^m d_{ij}^2 \right)^{\frac{1}{r+1}}}, \tag{16}$$

$$\mathbf{p}_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik}^m \mathbf{x}_i}{\sum_{i=0}^{N-1} u_{ik}^m}, \text{and} \tag{17}$$

$$\Sigma_k^{t+1} = \frac{\sum_{i=0}^{N-1} u_{ik} \left(\mathbf{x}_i - \mathbf{p}_k^{t+1} \right) \left(\mathbf{x}_i - \mathbf{p}_k^{t+1} \right)}{\sum_{i=0}^{N-1} u_{ik}}. \tag{18}$$

Figure 2 shows a comparison between FCM and PDI segmentation for one input image. Regions of the brain matter are misclassified by FCM as cerebrospinal fluid and the brain matter appears very fragmented. PDI balances the contributions of all clusters and results in a better segmentation with significantly less fragmentation.

4.2 Expectation Maximization

The Expectation-Maximization algorithm is an estimation method used to determine the maximum likelihood parameters of a mixture of known distributions in the presence of hidden or unobserved data [8]. In image segmentation, the observed data are the feature vectors associated with pixels, while the hidden variables are the expectations for each pixel that it belongs to each of the given clusters.

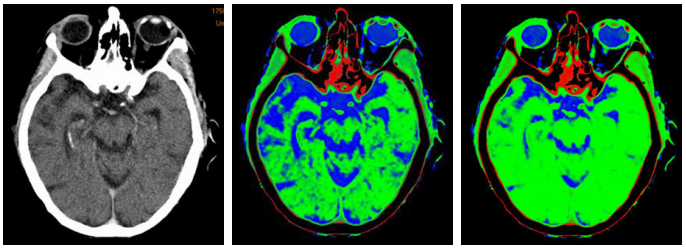

Fig. 2. Comparison between FCM and PDI. a) Original CT image. b) FCM results. c) PDI results. In the RGB labels, red is for low intensity bone, green for brain matter and blue for cerebrospinal fluid (CSF). CSF cluster has a large diameter (high variance) causing FCM to assign more pixels to this cluster and misclassify regions of brain matter as CSF. As a result, the brain matter appears very fragmented. PDI corrects this by decreasing the strength of CSF cluster.

The algorithm starts with an initial guess at the model parameters of the clusters and than re-estimates the expectations for each pixel in an iterative manner. Each iteration consists of two steps: the expectation (E) step and the maximization (M) step. In the E-step, the probability distribution of each hidden variable is computed from the observed values and the current estimate of the model parameters (e.g. mean, covariance). In the M-step, the model parameters are re-estimated assuming the probability distributions computed in the E-step are correct.

Using the notation of Sections 3.1 and 4.1 and assuming Gaussian distributions for all clusters, the hidden variables are the expectations E_{ik} that pixel i belongs to cluster k. The model parameters to estimate are the mean, the covariance and the mixing weight corresponding to each cluster. The mixing weight is a measure of a cluster's strength, representing how prevalent the cluster is in the data.

E-step:

$$E_{ik}^t = P\left(k|\mathbf{x}_i, \Theta_k^t\right) = \frac{P\left(\mathbf{x}_i|\omega_k, \Theta_k^t\right)\pi_k^t}{\sum_{j=1}^{c} P\left(\mathbf{x}_i|\omega_j, \Theta_j^t\right)\pi_j^t}. \tag{19}$$

M-step:

$$\pi_k^{t+1} = \frac{1}{N}\sum_{i=0}^{N-1} E_{ik}^t, \tag{20}$$

$$\mathbf{p}_k^{t+1} = \frac{1}{N\pi_k^{t+1}}\sum_{i=0}^{N-1} E_{ik}^t \mathbf{x}_i, \text{and} \tag{21}$$

$$\Sigma_k^{t+1} = \frac{1}{N\pi_k^{t+1}}\sum_{i=0}^{N-1} E_{ik}^t \left(\mathbf{x}_i - \mathbf{p}_k^{t+1}\right)\left(\mathbf{x}_i - \mathbf{p}_k^{t+1}\right), \tag{22}$$

where Θ_k^t are the model parameters of class k at time t and π_k^t is the mixing weight of class k at time t. Note that $\sum_{j=1}^{c} \pi_j^t = 1\ \forall t$.

The algorithm iterates between the two steps until the log likelihood increases by less than some threshold or a maximum number of iterations is reached. EM algorithm can be summarized as follows:

1. Initialize the means p_k^0, the covariance matrices Σ_k^0 and the mixing weights π_k^0 . Typically, the means are initialized to random values, the covariance matrices to the identity matrix and the mixing weights to $\frac{1}{c}$.
2. E-step: Estimate E_{ik} for each pixel i and class k, using (19).
3. M-step: Estimate the model parameters for class k, using (20)-(22).
4. Stop if convergence condition $\left(\log \prod_{i=0}^{N-1} E_{ik}^{t+1} - \log \prod_{i=0}^{N-1} E_{ik}^{t}\right) \leq \epsilon$ is achieved. Otherwise, repeat steps 2 to 4.

5 Implementation

The three segmentation algorithms were implemented in C++. We used the Insight Segmentation and Registration Toolkit (ITK) library [16] for IO operations (reading and writing DICOM and RGB series).

In our approach, the feature vector for each pixel is composed of a single value, the mean value of a 3x3 window centered on the pixel. Using this mean pre-filters the image to reduce noise in the input data.

Two intensity threshold values were preset to segment air ($\mathbf{x}_i \leq -150$) and bone ($\mathbf{x}_i \geq 300$). Only pixels with intensities between these values were considered for classification. This approach is feasible for CT images as CT scanners are calibrated so that tissues of certain densities always produce the same range of CT values, clustered around the "true" CT number (Hounsfield unit) of the tissue (e.g. 0 for water and -1000 for air).

The number of classes was set to 3 and pixels were categorized as belonging to low intensity bone, brain matter, and cerebrospinal fluid (CSF). All methods require initial means and covariance matrices for each of the three classes. The initial means were provided by the user. A visualization tool was used to compute statistics on small regions in the initial data to estimate means for all classes (100 for low intensity bone, 40 for brain matter and 0 for CSF). The resulting means were used to generate a first fuzzy classification of the data using Euclidian distances. The resulting classes provided the initial means and covariance matrices for the three methods. In the case of Bayesian classification, this approach provided a reasonable alternative to the manual segmentation step.

6 Results

The three segmentation algorithms were applied to volumetric CT images of a human head in order to segment the brain tissue from bone and cerebrospinal fluid (CSF). The input scan consists of 3D CT images of size 512 x 512. In the volumetric data, the voxel size is 0.48mm x 0.48mm x 4.5mm (width x height x depth). To visualize our results, we show the output from soft segmentation as

RGB images, the corresponding hard segmentation as grayscale images, and we generate 3D models of the brain from the resultant hard segmentation.

Figure 3 presents the results for soft segmentation of the brain from the Bayesian classification, the PDI and the EM method. The probability fields are saved as RGB images. Each color component of each pixel contains the membership value for one of the three classes (red for low intensity bone, green for brain matter, and blue for CSF). As expected, the Bayesian classification (Fig. 3b) results in a very fuzzy labeling. This is because the classification is not an iterative method and there is no convergence criteria incorporated in the algorithm. PDI (Fig. 3c) gives the best results of the three methods. The EM method (Fig. 3d) appears to be more sensitive to noise in the data than PDI.

To create labels for brain matter (Fig. 4) and for CSF (Fig. 5) we converted the output from soft segmentation to hard segmentation. In the soft segmentation results, each pixel has three associated probabilities, one for each class. We classified a pixel as belonging to the class with which it has the highest probability value as long as the value is bigger than a threshold. The threshold was set to 0.5 for Bayesian classification and to 0.8 for EM and PDI. A higher threshold was used for EM and PDI because they resulted in less fuzzy segmentation. We also compared the resulting labels with threshold-based segmentation of the brain (Fig. 4b and 5b). All three methods outperformed threshold-based segmentation, which resulted in very noisy labels.

From the resultant hard segmentation, we generated 3D models of the brain (Fig. 6). Threshold-based segmentation gives a very rough brain surface while PDI produces the smoothest results. The fact that EM is sensitive to noise is also apparent in the 3D model. However, the EM algorithm has the advantage of explicitly incorporating prior knowledge and, while this characteristic was not explored in this work, it could lead to interesting directions for future research.

7 Conclusions

We are interested in brain segmentation in CT data. Because of a lack of research in this area, we started by exploring existing methods that have been successful for segmenting MRI data. Three soft segmentation methods (Bayesian classification, Fuzzy c-Means and Expectation Maximization) were applied to segment brain tissue from bone and CSF in CT brain data and the results have been presented. All three methods performed better than threshold-based segmentation, with PDI giving the best results. Unfortunately, we were unable to locate a database of pre-segmented CT brain images and the lack of ground truth data prevented us from performing an absolute comparison of the three methods.

Our results show that brain segmentation can be performed on CT data. There are, however, limitations to the level of detail that can be achieved. While most work on brain MRI data segments gray matter and white matter, this is not possible on CT images where, as shown in Fig. 1, the two tissues are not distinguishable. To address this problem we have begun exploring new imaging protocols that provide better differentiation of soft tissue detail in CT data.

The new use of dual energy CT for clinical applications has the potential of redefining the role of CT imaging for soft tissue segmentation. Dual energy CT scanners use two X-ray energy levels simultaneously [15], allowing radiologists to better differentiate, characterize and isolate body tissues and fluid. Adapting the algorithms presented here to incorporate characteristics of the dual energy CT data is a major focus of our future research.

References

1. Berquist TH, Dalinka MK, Alazraki N, Daffner RH, DeSmet AA, el-Khoury GY, Goergen TG, Keats TE, Manaster BJ, Newberg A, Pavlov H, Haralson RH, McCabe JB, Sartoris D.: Soft tissue masses. American College of Radiology. ACR Appropriateness Criteria. Radiology. **215** (2000) 255–9
2. Shihab AI.: Fuzzy Clustering Algorithms and Their Application to Medical Image Analysis. In Ph.D. thesis, University of London 2000
3. Spence KVN.: Automatic Segmentation of Magnetic Resonance Images of the Brain. In Ph.D. thesis, Luisiana State University 2005
4. Laidlaw DH, Fleischer KW, Barr AH.: Classification of Material Mixtures in Volume Data for Visualization and Modeling. 1994. Technical Report CS-TR-94-07, California Institute of Technology, Pasadena, CA 91125
5. Li CL, Goldgof DB, Hall LO.: Knowledge-based classification and tissue labeling of MR images of human brain. IEEE Trans. Med. Imag. **12** (1993) 740–750
6. Pham DL, Prince JL.: Adaptive fuzzy segmentation of magnetic resonance images. IEEE Trans. Med. Imag. **18** (1999) 737–752
7. Wells III WM, Grimson WEL, Kikinis R, Jolesz FA.: Adaptive segmentation of MRI data. IEEE Trans. Med. Imag. **15** (1996) 429–442
8. Kapur T.: Model based three dimensional medical image segmentation. In Ph.D. thesis, Massachusetts Institute of Technology 1999
9. Maksimovic R, Stankovic S, Milovanovic D.: Computed tomography image analyzer: 3D reconstruction and segmentation applying active contour models - 'snakes'. International Journal of Medical Informatics **58-59** (2000) 29–37.
10. Deleo JM, Schwartz M, Creasey H, Cutler N, Rapoport SI.: Computer-assisted categorization of brain computerized tomography pixels into cerebrospinal fluid, white matter, and gray matter. Computers and Biomedical Research **18** (1985) 79–88.
11. Hu Q, Qian G, Aziz A, Nowinski WL.: Segmentation of brain from computed tomography head images. IEEE Engineering in Medicine and Biology Society (2005) 3375–3378.
12. Loncaric S, Domagoj Kovacevic.: A method for Segmentation of CT Head Images. Lecture Notes in Computer Science **1311** (1997) 388–395
13. Jain AK, Murty MN, Flynn PJ.: Data Clustering. A review. ACM Computing Surveys **31** (1999)
14. Bezdek JC.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York, 1981.
15. Ying Z, Naidu R, Crawford CR.: Dual energy computed tomography for explosive detection. Journal of X-Ray Science and Technology **14** (2006) 235-256.
16. Insight Segmentation and Registration Toolkit. http://www.itk.org.

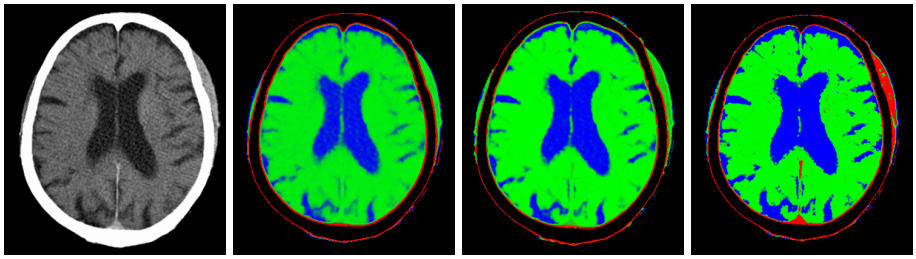

Fig. 3. CT brain segmentation. In the RGB labels, red is for low intensity bone, green for brain matter and blue for cerebrospinal fluid (CSF). (a) Original axial CT slice. (b) Output from Bayesian classification. This method results in fuzzier segmentation compared to the other two methods. (c) Output from PDI method. The method corrects FCM behavior by reducing the influence of large clusters. (d) Output from EM method. The method is sensitive to noise in the data.

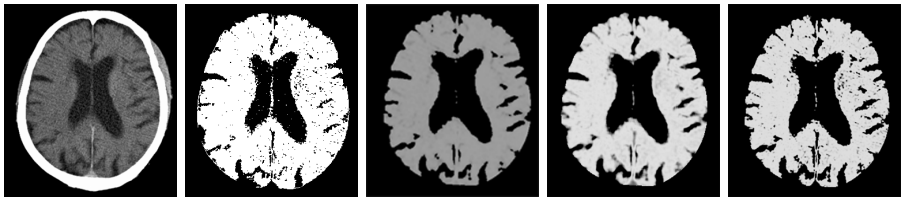

Fig. 4. CT brain segmentation. (a) Original axial CT slice. (b) Simple thresholding of the original data. (c) Brain matter from Bayesian classification. (d) Brain matter from PDI method. (e) Brain matter from EM method.

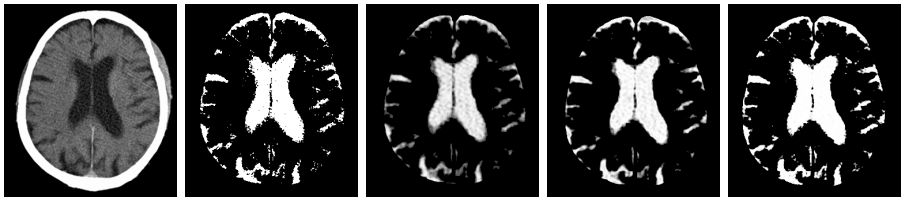

Fig. 5. CT brain segmentation. (a) Original axial CT slice. (b) Simple thresholding of the original data. (c) CSF from Bayesian classification. (d) CSF from PDI method. (e) CSF from EM method.

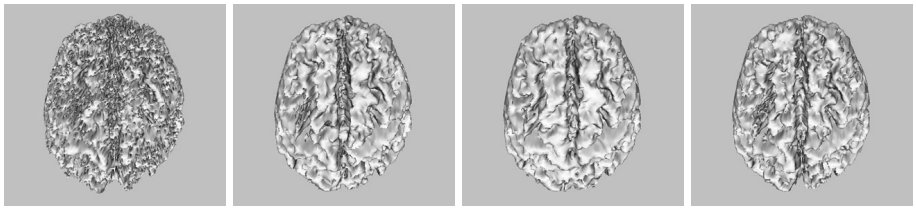

Fig. 6. 3D models of the brain (a) Simple thresholding of the original CT image (b) Bayesian classification (c) PDI (d) EM algorithm.