Semi-automatic Segmentation of Brain Tumors Using Population and Individual Information

Yao Wu · Wei Yang · Jun Jiang · Shuanqian Li ·
Qianjin Feng · Wufan Chen

Published online: 15 January 2013
© Society for Imaging Informatics in Medicine 2013

Abstract Efficient segmentation of tumors in medical images is of great practical importance in early diagnosis and radiation plan. This paper proposes a novel semi-automatic segmentation method based on population and individual statistical information to segment brain tumors in magnetic resonance (MR) images. First, high-dimensional image features are extracted. Neighborhood components analysis is proposed to learn two optimal distance metrics, which contain population and patient-specific information, respectively. The probability of each pixel belonging to the foreground (tumor) and the background is estimated by the k -nearest neighborhood classifier under the learned optimal distance metrics. A cost function for segmentation is constructed through these probabilities and is optimized using graph cuts. Finally, some morphological operations are performed to improve the achieved segmentation results. Our dataset consists of 137 brain MR images, including 68 for training and 69 for testing. The proposed method overcomes segmentation difficulties caused by the uneven gray level distribution of the tumors and even can get satisfactory results if the tumors have fuzzy edges. Experimental results demonstrate that the proposed method is robust to brain tumor segmentation.

Keywords Neighborhood components analysis · k -nearest neighborhood · Graph cuts · Distance metric learning

Introduction

Magnetic resonance imaging (MRI) is a widely used technique for detecting lesions in brain. Compared to other

imaging modalities, MRI can easily obtain high-contrast images. Furthermore, different pulse sequences of MR scans (3D images) enhance specific tissues of brain. For example, fluid-attenuated inversion recovery (FLAIR) restrains the gray level of cerebrospinal fluid, while T1C (T1 contrast enhancement) highlights the abnormal regions. Information in different pulse sequences of MR scans can be fully used in medical image analysis [1–3]. Therefore, a lot of work has been done on tumor segmentation in MR images. The precise segmentation of brain tumors in MR images is necessary for clinical surgery. However, manual segmentation is time-consuming and the results are often different between radiologists, so automatic and semi-automatic segmentation methods for brain tumors in MR images are valuable for research and clinical application.

Automatic methods have been proposed in the literature. For instance, Clark and Hall [4] presented a novel automatic brain tumor segmentation algorithm. They encoded knowledge of the pixel intensity and spatial relationships in the images to create a fully automated segmentation system known as the KG (knowledge-guided) method. It was initially trained to identify brain tumors. However, KG method was limited to enhancing tumors with clear enhancing edges [5]. Fletcher et al. [6] proposed a method based on FCM (fuzzy C-mean) clustering algorithm. They removed the gray and white matters of brain tissue according to the clustering results. Training on 14 patient images, they got a knowledge-based system, and the ventricle was removed based on the shape-matching technique. The final tumor segmentation result was achieved by integrating domain knowledge and image processing approaches. But this method cannot be applied if the tumor is too large or the ventricle symmetry is not so good. Prastawa et al. [7] used an atlas-based technique to automatically detect brain tumors. The intensity properties of the different tissues were determined using a registered brain atlas model. Geometric

Y. Wu · W. Yang · J. Jiang · S. Li · Q. Feng (✉) · W. Chen
School of Biomedical Engineering, Southern Medical University,
Guangzhou 510515, China
e-mail: qianjinfeng08@gmail.com

and spatial constraints were then applied to the detected tumor regions. This method requires a large template library of brain atlas, thus can find the similar atlas model corresponding to the input image. However, when large deformation in brain structures exists, the used atlas model may lead to incorrect sampling and its algorithm may achieve incorrect results. Khotanlou et al. [8] presented another automatic segmentation approach. Histogram analyses and fuzzy classification were used to get an initial tumor segmentation. And then, a deformable model and spatial relations were applied to improve the segmentation result. However, the intensity distribution in medical images is heterogeneous, and the gray level histograms cannot show the differences between tumors and the surrounding tissues very well. It seems that fully automatic and accurate segmentation approaches are not yet a reality [9]. To date, incomplete automatic segmentation method is adopted in clinical application. However, manual tumor segmentation by radiologists is a challenging and time-consuming task. This motivates the need for a semi-automatic segmentation method which alleviates the problems within automatic segmentation method, and the radiologists can also control over the segmentation process.

In recent years, many semi-automatic segmentation techniques have been proposed, such as intelligent scissors [10, 11], matting [12], random walker [13], GrowCut [14], watershed cuts [15]. Graph cuts is another novel semi-automatic segmentation method which was produced by Boykov et al. [16]. Graph cuts is a global optimization strategy which gives the best balance of region and boundary terms. It is also a flexible segmentation framework which can extract suitable image features directing at specific segmentation task. In addition, it is convenient to integrate priors into this framework. In recent years, many new techniques have been proposed to extend the graph cuts algorithm, such as GrabCut [17], lazy snapping [18], and graph cuts-based methods have been widely used for medical images [19–22]. These advantages motivate us to develop a graph cuts-based method to segment brain tumors in MR images.

In the standard graph cuts algorithm, one-dimensional color information (gray value) is used as the image feature and the gray level histograms of the foreground and background are established to construct the cost function, which is based on the assumption that the intensity distribution of the foreground is distinct from that of the background. However, from Fig. 1, we can see that the gray level distributions of the tumor

Fig. 1 The intensity distributions of the tumor and its surrounding region in two MR images. The *left two* show the brain MR images. The intensity distributions of the tumors and its surrounding region are drawn in the *right side* of figure

and background have overlap area, and they are not completely separated, thus only using gray value differences is sometimes difficult to handle medical images. According to paper [21], we can see that high-dimensional features can improve the segmentation accuracy in medical images. But there are still some problems to introduce the high-dimensional features into graph cuts segmentation framework: (1) In high-dimensional feature space, it is hard to estimate the probability density of the tumor and the background (F_{obj} , F_{bkg} in Eq.(2)) if the training samples are few, which needs a new probability estimating method. (2) In high-dimensional feature space, the feature scales are often not consistent with their importance; however, Euclidean distance treat all feature components equally. Using Euclidean distance to measure the high-dimensional features between samples ($G(x_p, x_q)$ in Eq.(3)) will severely decrease the segmentation accuracy. One popular approach is to scale features by information of the training data to improve segmentation accuracy [21, 23]. In addition, for medical image segmentation, a large number of segmented images can be utilized. These already-segmented images can construct population training set, which can solve the segmentation problem caused by insufficient training samples.

Here, a novel semi-automatic method based on graph cuts optimization framework for brain tumor segmentation is presented. Our method is novel in the following ways.

1. High-dimensional features are adopted as image features to identify the tumors from the images. The k -nearest neighbors (k NN) modeling technique is used here to construct the region term in Eq. (1), predicting the probability density of the tumor and the background, which is different from using the gray level histograms usually adopted by traditional graph cuts-based methods [16, 19].
2. In high-dimensional feature space, neighborhood components analysis (NCA) [23] is proposed to learn optimal distance metric, which cannot only improve the accuracy of k NN classifier, but also be used to construct the boundary term (Eq.(3)).
3. A new segmentation model is proposed which takes both the patient specific and the population information into consideration. In addition, population training set contains much information near the tumor edges, thus, if the test image has fuzzy tumor boundary, the proposed algorithm still can get a satisfactory result.

Materials and Methods

Framework of Graph Cuts Algorithm

In graph cuts-based segmentation framework, segmentation problem is formulated in terms of energy minimization. Let X denote the set of image pixels, and N is a set of

neighboring pixels. $l = \{l_p | l_p \in L\}$ denotes a labeling, and the $L = \{0, 1\}$ is the segmentation label set, where 0 corresponds to the background and 1 corresponds to the foreground. The cost function is defined as following:

$$l^* = \underset{l}{\operatorname{argmin}} \left(\beta \sum_{p \in X} D_p(l_p) + \sum_{p, q \in N} V_{p, q}(l_p, l_q) \right) \quad (1)$$

On the right-hand side of Eq. (1), the first term is the region term, which measures the sum of penalty for assigning each individual pixel p to l_p . In the standard graph cuts algorithm [16], the region term is defined as:

$$\begin{cases} D_p(l_p = l^{\text{obj}}) = -\ln P(l_p = l^{\text{obj}} | x_p, F_{\text{obj}}) \\ D_p(l_p = l^{\text{bkg}}) = -\ln P(l_p = l^{\text{bkg}} | x_p, F_{\text{bkg}}) \end{cases} \quad (2)$$

where l^{obj} is the label of object, l^{bkg} is the label of background, x_p is the image features, and F_{obj} , F_{bkg} are probability density functions of the object and background, respectively. $P(l_p = l^{\text{obj}} | x_p, F_{\text{obj}})$ denotes the conditional probability of assigning pixel p to l^{obj} and $P(l_p = l^{\text{bkg}} | x_p, F_{\text{bkg}})$ denotes the conditional probability of assigning pixel p to l^{bkg} .

The second term in Eq. (1) is the boundary term, which can be interpreted as a penalty for a discontinuity between p and q . As introduced in paper [16], this term is defined as

$$V_{p, q} = \begin{cases} \exp \left[-\frac{G(x_p, x_q)}{2\sigma^2} \right] \cdot \frac{1}{\text{dist}(p, q)} & l_p \neq l_q \\ 0 & l_p = l_q \end{cases} \quad (3)$$

where $G(x_p, x_q)$ denotes the distance between features of pixel p and q . $\text{dist}(p, q)$ is the Euclidean distance between the locations of p and q . Parameter σ is a constant relating to the gray level which measures noise standard deviation. This penalty term is introduced when pixels p and q have different labels.

Overview of the Proposed Algorithm

We introduce high-dimensional features into graph cuts segmentation framework. NCA is proposed to learn two optimal distance metrics, which contain population and patient-specific information, respectively. The probability of each pixel belonging to the foreground (tumor) and the background is estimated by the k NN classifier under the learned optimal distance metrics. A new cost function for segmentation is constructed through these probabilities and is optimized using graph cuts. Construction of the graph is based on Kolmogorov et al. [24], and minimization of the cost function is conducted using the min-cut/max-flow algorithm [25]. Finally, some morphological operations are performed to improve the achieved segmentation results. Figure 2 shows a flowchart of the proposed brain tumor segmentation algorithm.

Fig. 2 A flowchart of the algorithm

Image Feature Extraction

The texture features which captures the spatial correlation among adjacent pixels such as gray level co-occurrence matrix (GLCM) [26] and the moments of the gray-level histogram (MGH) [27] which represent the local structure of the image have achieved satisfactory results [28, 29]. In our paper, these two methods are used to extract the image features.

Before extracting the image features, we should normalize the intensity value. Normalization is a vital step in MR image analysis, since the intensity values of MR images are sensitive to acquisition conditions. For each slice, intensity values at the 1 and 99 % quantiles are computed for the brain region. Then, these two values are used to normalize intensity to [0, 1] through the min–max method. After intensity normalization, MGH- and GLCM-based features are extracted.

1. MGH-Based Features

Gray level histogram is a summary of the statistical information. MGH is based on the statistical properties of intensity distribution in a local region, and its effectiveness for the classification of tissues has been evaluated in [27, 28]. The feature extraction method adopted here is similar to that of [28]. Nine statistics, including the mean, variance, smoothness, uniformity, entropy, and so on, are calculated and compose the image features. The following is the list of

these features:

$$\text{Mean} : f^1 = \sum_{i=0}^{H-1} z_i p(z_i) \quad (4)$$

$$\text{Variance} : f^2 = \sum_{i=0}^{H-1} (z_i - m)^2 p(z_i) \quad (5)$$

$$\text{Smoothness} : f^3 = 1 - \frac{1}{1 + \sigma^2} \quad (6)$$

$$\text{Third moment} : f^4 = \sum_{i=0}^{H-1} (z_i - m)^3 p(z_i) \quad (7)$$

$$\text{Uniformity} : f^5 = \sum_{i=0}^{H-1} p^2(z_i) \quad (8)$$

$$\text{Entropy} : f^6 = - \sum_{i=0}^{H-1} p(z_i) \log_2 p(z_i) \quad (9)$$

$$\text{Others} : f^7 = \sum_{i=0}^{H-1} p^3(z_i), f^8 = \sum_{i=0}^{H-1} p^4(z_i), f^9 = \sum_{i=0}^{H-1} p^5(z_i) \quad (10)$$

where z_i indicates the pixel intensity, $p(z_i)$ is the gray level histogram in a region, which is calculated from a $W \times W$ window with the pixel of interest in the center. H is the number of possible intensity levels. m is the mean intensity of the pixels in the $W \times W$ window, and σ^2 is the variance. They are defined as

$$m = \sum_{i=1}^{W \times W} z_i / (W \times W); \sigma^2 = \sum_{i=1}^{W \times W} (z_i - m)^2 \quad (11)$$

2. GLCM-Based Features

Only using intensity distribution-based features is not sufficient because they do not take spatial information into consideration. Features extracting from the GLCM are based on the joint probability distribution of pairs of pixels, which have been proved to be effective for texture image classification [29].

The dimension of the co-occurrence matrix is equal to the number of gray levels. The gray levels in medical images are usually reduced to M discrete levels to save computational time. GLCM is a $W \times W$ dimensional matrix. Distance d and angle θ are used to calculate the joint probability distribution between pixels. The (i, j) th entry $p(i, j)$ in the matrix gives the number of times that the gray level j follows the gray level i for a distance with an angle.

Fourteen well-known coefficients were introduced as the GLCM-based features [26]. But only six of these features are the most relevant [29]. These are energy, entropy, contrast, variance, correlation, and inverse difference moment. Here, we extract these six effective features. The following equations define these features:

$$\text{Energy(Angular second moment)} : \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} [p(i, j)]^2 \quad (12)$$

$$\text{Entropy} : - \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} p(i, j) \log_2 [p(i, j)] \quad (13)$$

$$\text{Contrast(Inertia)} : \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} (i - j)^2 p(i, j) \quad (14)$$

$$\text{Variance} : \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} (i - \mu)^2 p(i, j) \quad (15)$$

$$\text{Correlation} : \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{ijp(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (16)$$

$$\text{Inverse difference moment} : \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{p(i, j)}{1 + (i - j)^2} \quad (17)$$

where M is the dimension of GLCM. Coefficients μ_x , μ_y and σ_x , σ_y denote the mean and standard deviations of the row and column sums of the matrix, respectively, and μ is the mean of μ_x and μ_y . They are defined as follows:

$$\begin{aligned} \mu_x &= \sum_{i=0}^{M-1} ip_x(i); \quad \mu_y = \sum_{i=0}^{M-1} ip_y(i); \quad \mu = \frac{\mu_x + \mu_y}{2}; \\ \sigma_x &= \sqrt{\sum_{i=0}^{M-1} p_x(i)(i - \mu_x)^2}; \quad \sigma_y = \sqrt{\sum_{i=0}^{M-1} p_y(i)(i - \mu_y)^2} \\ p_x(i) &= \sum_{j=0}^{M-1} p(i, j); \quad p_y(j) = \sum_{i=0}^{M-1} p(i, j); \end{aligned} \quad (18)$$

Coefficient $p(i, j)$ is the (i, j) th entry in GLCM. After calculating six GLCM-based features for each $p(i, j)$, the mean and variance of the four values (for $\theta=0^\circ, 45^\circ, 90^\circ, 135^\circ$) were taken as the final features, and then, 12 GLCM-based features are present.

kNN-Based Probability Model

In high-dimensional feature space, it is hard to estimate the probability density of the tumor and the background. k NN is a probability estimating method, which can be used for classification. This algorithm is implemented by computing the distances of feature vectors from the test sample to all training samples. In this model, a test sample is classified into the majority class among its k -nearest neighbors in the training dataset. It is simple to implement without any assumption on the shape of the class distributions. In this paper, a k NN-based probability model is introduced to construct the region term, estimating the probabilities of assigning each pixel to tumor or background.

Let $S = \{x_i, l_i \mid i=1, \dots, n\}$ denote a training set, x_i is the feature vector of training sample i , and l_i is the corresponding label. Under a distance metric G , the K -nearest training samples of a test pixel are determined. C_{obj} , C_{bkg} denote the number of these K training samples belonging to the object and background, respectively. The probability of pixel p belonging to the object and background can be denoted as

$$P(l_p = l^{\text{obj}} \mid k\text{NN}(x_p, S, G)) = \frac{C_{\text{obj}}}{K} \quad (19)$$

$$P(l_p = l^{\text{bkg}} \mid k\text{NN}(x_p, S, G)) = \frac{C_{\text{bkg}}}{K} \quad (20)$$

Given the training set S and distance metric G , $k\text{NN}(x_p, S, G)$ denotes the K -nearest training samples of x_p . $P(l_p = l^{\text{obj}} \mid k\text{NN}(x_p, S, G))$ is the conditional probability of assigning pixel p to l^{obj} and $P(l_p = l^{\text{bkg}} \mid k\text{NN}(x_p, S, G))$ denotes the conditional probability of assigning pixel p to l^{bkg} . Using the k NN model to predict the probability, the gray level

histograms need not be constructed. In our method, the region term in Eq. (1) is defined as

$$\begin{cases} D_p(l_p = l_{\text{obj}}) = -\ln \frac{C_{\text{obj}}}{K} \\ D_p(l_p = l_{\text{bkg}}) = -\ln \frac{C_{\text{bkg}}}{K} \end{cases} \quad (21)$$

Optimal Distance Metric Learning

In the current paper, k NN classifier is introduced to construct the region term, and it can be significantly improved by learning a distance metric from labeled seeds via NCA [23]. NCA is a method for learning a Mahalanobis distance measure for k NN, and it can also reduce high-dimensional linear space that can be used for immediate classification. In this section, we briefly describe the basic principle of NCA algorithm.

Let $S = \{x_i, l_i \mid i = 1, \dots, n\}$ denote a training set with sample $x_i \in R^{n1}$ and its corresponding label $l_i \in \{0, 1\}$. The goal of NCA is to find a transformation matrix $T \in R^{n1 \times n2}$ ($n2$, which projects the samples to space R^{n2}). In this new space, k NN classifier can reach the highest classification accuracy:

$$T^* = \arg \max_T O(T) \quad (22)$$

where $O(T)$ is the sum of classification accuracies for all training samples in S .

In this new space, the distance between transformed samples x_i and x_j is measured by

$$G(x_i, x_j) = (x_i - x_j)^T (TT^T)(x_i - x_j) \quad (23)$$

And the similarity between x_i, x_j can be defined as

$$s_{ij} = \exp(-G(x_i, x_j)) \quad (24)$$

For one sample x_i , the probability of being correctly classified according to k NN classification rule can be written as

$$\Pr(x_i = l_i) = \frac{\sum_{j \neq i, l_j = l_i} s_{ij}}{\sum_{j \neq i} s_{ij}} \quad (25)$$

Maximizing the leave one out error on the training dataset, a transformation matrix T can be achieved. The objective function of NCA is defined as

$$\begin{aligned} O(T) &= \sum_{i=1}^n \Pr(x_i = l_i) = \sum_{i=1}^n \frac{\sum_{j \neq i, l_j = l_i} s_{ij}}{\sum_{j \neq i} s_{ij}} \\ &= \sum_{i=1}^n \frac{\sum_{j \neq i, l_j = l_i} \exp(-G(x_i, x_j))}{\sum_{j \neq i} \exp(-G(x_i, x_j))} \\ &= \sum_{i=1}^n \sum_{j \neq i, l_j = l_i} \frac{\exp(-G(x_i, x_j))}{\sum_{k \neq i} \exp(-G(x_i, x_k))} \end{aligned} \quad (26)$$

Let

$$\tau_{ij}(T) = \frac{\exp(-G(x_i, x_j))}{\sum_{k \neq i} \exp(-G(x_i, x_k))} \quad (27)$$

Then, the objective function of NCA can be denoted as

$$O(T) = \sum_{i=1}^n \sum_{j \neq i, l_j = l_i} \tau_{ij}(T) \quad (28)$$

The sum of the correct classification probabilities of all training samples is calculated via Eq. (26). Using the gradient-based approaches, the maximum of the objective function can be obtained, thus, achieves the distance transformation matrix T and the optimal distance metric G (Eq.(23)) is finally used to construct the region term (Eq. (21)) and boundary term (Eq. (3)).

The Combination Model

Using graph cuts segmentation framework, we need to provide “object” and “background” seeds through manual interactions, and these training samples are very important for the construction of k NN model and the optimal distance metric learning.

There are two drawbacks in traditional interactive modes. First, the interactive users often obtain the seeds away from the boundary of the tumor, in order to avoid that the brush straddles the edge, thus neglecting the information near the edge. The k NN classifier may have difficulty near the edge of tumor because of lacking seeds in this margin region. Second, in traditional interactive modes, the “object” and “background” seeds are only taken from one patient-specific image with manual interaction, and sometimes the tumor is hard to distinguish from other tissues if the seeds are not sufficient.

To solve these problems, two groups of training sets are constructed here: (1) 68 brain MR images of other patients are used to build the population training set, denoted as S_{Pop} . The manual segmentations of the tumor in each image are also provided. To solve the first problem mentioned above, which is that the interactive users often obtain the seeds away from the boundary of the tumor thus neglecting the information near the edge, this population dataset is built by extracting seeds near the tumor boundaries, where the edge information can be emphasized; (2) The individual training set S_{Ind} is made up of the manual interactive seeds on the test image.

Based on the population and individual sets, $k\text{NN}_{\text{Pop}}$ and $k\text{NN}_{\text{Ind}}$ classifiers are constructed, and two transformation matrices G_{Pop} , G_{Ind} are learned from via NCA, respectively. Two segmentation models based on population and individual information are constructed, and their cost functions are

Table 1 A summary of parameter setting in the current paper

Parameter	Description	Setting
β	Cost function (Eq. (1))	5~10
σ	Boundary term (Eq. (3))	500
H	Gray levels in MGH (Eq. (4)~(10))	256
d	Distance in GLCM	1
θ	Directions in GLCM	0°, 45°, 90°, 135°
M	Used in GLCM (Eq.(12)~(18))	16
W	Size of window in GLCM and MGH	15
K	k NN classifier (Eq. (21))	7
δ	Parameter in Eq. (30)	200

denoted as E_{Pop} , E_{Ind} , respectively. In our method, a weighted combination of these two cost functions is finally used:

$$E = \alpha E_{\text{Ind}} + (1 - \alpha) E_{\text{Pop}} \quad (29)$$

The weight of E_{Pop} and E_{Ind} is dynamically adjusted by factor α according to the position of each test pixel during the segmentation process. The choice of α is based on the following assumption: If the test pixel is closer to the manual interactive seeds, the image information between the test pixel and the individual set is more similar, thus increasing the weight of E_{Ind} . The weight of E_{Pop} increases as the test pixel gets far to the individual seeds. To restrict the weight factor α between 0 and 1, it is defined as:

$$\alpha = \exp(-d_i^2/\delta) \quad (30)$$

where d_i is the shortest distance from the test pixel i to the manual interactive seeds. The value of parameter δ is determined through experiments.

Results and Discussion

Data Acquisition and Workstation

Our dataset consists of 21 patients (10 for training and 11 for testing) with total 137 T1C MR images, including 68 for training and 69 for testing. All these images were extracted from 42 scans with brain gliomas. The data were acquired at Nanfang Hospital, Guangzhou, China and General Hospital, Tianjin Medical University, China from September 2005 to October

2010 using spin echo weighted images with 512×512 matrix. The manual segmentations of the tumors in each image were also provided to evaluate the performance of the segmentation.

A PC with Intel P4 3.0 GHz processor and 4 GB RAM was used as the workstation for the segmentation. The software environment was a Visual Studio C++2008 MFC under Windows XP.

Parameter Choices

Many parameters exist in the current paper. In this section, we discuss how to choose the suitable values and a summary of the settings is listed in Table 1.

The coefficient β in Eq.(1) denotes the relevance of the region term to the boundary term in the cost function, computing the ratio of the region term to the boundary term, and setting the reciprocal of the ratio to β . Normally, β ranges from 5 to 10 in our dataset.

Parameter σ in the boundary term (Eq. (3)) measures the noise standard deviation. To set this parameter properly, σ was varied from 1 to 1,000 with 50 for interval in our experiments. Finally, σ was set as 500.

Parameter H in Eq. (4) is the number of possible intensity levels. However, in brain MR images, the intensity levels vary a lot from different images. The maximal gray value even exceeds 3,000 in many cases. We decrease this highest level and set H as 256 after several trials.

Parameter d , θ , and M are used to calculate features based on GLCM which is a classical feature extracting method. Normally, only the distances $d=1$ and 2 with angles $\theta=0^\circ$, 45° , 90° , and 135° are used. To get a fine textural segmentation, d was set to a small value ($d=1$) in our experiments. As suggested in paper [30], we set M as 16 in Eq. (12) to compute GLCM.

MGH and GLCM features were computed in a $W \times W$ window with test pixel at the center. A 15×15 window was finally chosen after several trials.

Parameter K is the number of nearest training samples of each test pixel in k NN model. Larger value of K can reduce the effect of noise, but make boundaries between tumor and background less distinct. A good K can be selected by cross-validation. It is helpful to choose K to be an odd number in segmentation problem, and it was set as 7 in our experiments.

Parameter δ in Eq.(30) is used to adjust factor α . To set this parameter properly, δ was varied from 1 to 1,000 with 50 for interval in our experiments. Finally, δ was set as 200.

Table 2 Statistical results of 69 test images using different segmentation models

Different segmentation models	Mean $\pm$ std (%)	Min (%)	Max (%)
Using Euclidean distance metric	83.5 $\pm$ 11.2	42.7	94.8
Population model	87.4 $\pm$ 6.8	62.2	94.9
Individual model	92.3 $\pm$ 3.7	76.8	96.6
Combination model	94.1 $\pm$ 3.0	79.6	97.1

Fig. 3 DSCs of 69 test images using different segmentation models

Effectiveness of the Combination Model

Some experiments were carried out to evaluate the effectiveness of the combined segmentation model. Dice similarity coefficient (DSC) was used to measure the similarity between segmentation results and ground truth.

$$DSC = \frac{2|A_G \cap A_A|}{|A_G| + |A_A|} \quad (31)$$

where A_G represents the area of the ground truth of target and A_A represents the area of segmented target by the testing method. DSC is expressed as mean $\pm$ standard deviation. Higher mean DSC value denotes better segmentation.

Table 2 displays the performance results of 69 test images using different segmentation models. Compared to the

model using Euclidean distance metric (using k NN without NCA), the mean DSC of our method is improved from 83.5 to 94.1 % with 10.6 % increase (t test p value less than 0.0001), which shows the necessity of using NCA to learn a new distance metric. Compared to population and individual segmentation models, the mean DSC of our approach increases by 6.7 % (t test p value less than 0.0001) and 1.8 % (t test p value of 0.0003), respectively. These results indicate that the proposed method using the combination model is effective.

All segmentation results of the 69 images are shown in Fig. 3. After investigating all the results, we find that even the minimum DSC of our method is 79.6 %, indicating that the combined segmentation model is robust. These results can be explained as follows. First, as introduced above, k NN

Fig. 4 Some typical segmentation results showing the effectiveness of the combination model (all segmentation results show the magnified tumor regions): **a** test images with seeds: *internal lines* (foreground),

external lines (background), **b** model using Euclidean distance metric, **c** population model, **d** individual model, **e** combination model, and **f** the manual segmentation results by a radiologist

Table 3 Statistical results of 69 test images using different features

Using different features	Mean $\pm$ std (%)	Min (%)	Max (%)
Gray value	86.9 $\pm$ 6.9	62.8	94.8
GLCM features	81.7 $\pm$ 8.6	58.2	93.3
MGH features	92.1 $\pm$ 3.9	73.5	94.9
GLCM–MGH features	94.1 $\pm$ 3.0	79.6	97.1

classifier is used to predict the probability density of the tumors and the background. When the feature scales are not consistent with their importance, the accuracy of k NN classifier decreases using Euclidean distance metric. Thus, using NCA to learn a Mahalanobis distance metric is effective. Second, if we use individual model, taking seeds from one patient-specific image, it might be hard to distinguish the tumors from other tissues if the seeds are few. Additionally, the interactive users often obtain the seeds away from the boundary of the tumor and thus neglect the information near the edge, which leads to difficulty for k NN classifier near the boundary of the tumor. Therefore, the combination model can use the advantage of both population and individual information, thus providing better segmentation results. Two samples are shown in Fig. 4.

Effectiveness of the Combination Features

The performance of the GLCM-MGH combination features is also evaluated on the 69 test images in our experiments. Table 3 shows the DSC computed between the segmentation results and the ground truth.

Compared to method using gray value as image feature, the mean DSC of the GLCM-MGH features is improved

from 86.9 to 94.1 %. This is because just using intensity changes of foreground and background is sometimes insufficient to identify the tumors from medical images. From Fig. 5, we can see the high-dimensional features based on GLCM and MGH produce more accurate segmentation results.

Comparison with Other Semi-automatic Segmentation Methods

In the introduction, many semi-automatic methods are mentioned. To demonstrate the performance of our method, other five semi-automatic methods were tested in our experiments. Figure 6 shows the segmentation results of six methods: (a) our method, (b) Random walker [13], (c) grow cut [14], (d) graph cuts [19], (e) GrabCut [17] and (f) SLIC superpixels [31].

Superpixel is another novel segmentation method. In superpixel-based methods, an image is divided into many adjacent patches (see Fig. 6f). A superpixel is an image patch where the pixel intensity is well aligned with each other. The paper [31] supplies a novel technique which can output a desired number of superpixels with regular shapes and sizes, called simple linear iterative clustering (SLIC) superpixels. Unlike other semi-automatic methods, this approach need not provide seeds, but the user should choose the patches which look like the desired object with mouse. Because manual interactions exist in this approach, we also regarded it as a semi-automatic method and compared it with other algorithms.

From Fig. 6, it can be seen that our method and GrabCut appear to outperform the other methods. In GrabCut algorithm, if we only use a single initial user interaction

Fig. 5 Some typical segmentation results showing the feature influence: **a** test images with seeds, **b** using gray value as image feature, **c** using GLCM features, **d** using MGH features, **e** combining GLCM and MGH features, and **f** the manual segmentation results by a radiologist

Fig. 6 Comparison of different semi-automatic segmentation methods: **a** our method, **b** Random walker, **c** grow cut, **d** graph cuts, **e** GrabCut, **f** SLIC Superpixels. Images **a~d** in the *top row* show the original images with seeds. Image **e** (*top row*) shows GrabCut initial

user interaction (*rectangle*) and further user edits (nine foreground brushes). Image **f** (*top row*) is the result after classifying all pixels of original image into 1,024 superpixels. The *bottom row* illustrates the segmentation results

(rectangle in Fig. 6e), the segmentation result is not satisfying. Additionally, some brain tumors such as glioma often have high gray value near the tumor edges. This specific characteristic increases the difficulty for GrabCut to succeed on the edge of the tumor. In this cases, more foreground brushes need to be added. After adding further user edits (nine brushes in Fig. 6e), the result can be better. Considering the characteristic of brain tumor, our method seems more acceptable than GrabCut due to the simplicity of user input.

In our experiments, not all results of the six methods were investigated, but four comparable methods (Fig. 6a~d) were used to test on the 69 images, for they had the same manual interactive seeds. Here, we also use standard segmentation measure DSC to evaluate the four methods. Figure 7 shows the performance results for these methods.

Fig. 7 Statistical performance evaluation by comparing different segmentation methods using DSC (mean±std)

The statistical results show that our method and grow cut have better results. Grow cut algorithm can give satisfactory segmentation results but also needs the most computing time because of iterative calculating. For one image in our dataset, the computing time of grow cut was about 30 s, which is not desired by the user. Random walker shows good results when the seeds are few or placed near the center of the tumor. It is fast and less affected by the number of the seeds, but is seriously influenced by the seed location, thus leading to a higher std DSC. Traditional graph cuts algorithm needed the least computing time (no more than 1 s) but gave the worst results in our experiment. Compared to graph cuts, the mean DSC of our method is improved from 88.1 to 94.1 %. Our method has the lowest std DSC of the four methods, showing robustness to data changes. Additionally, the computing time for one image is about 4~8 s (smaller tumor needs less time). This speed can also meet the clinical demand. Considering both the computing efficiency and the result accuracy, our method seems more acceptable by clinical use.

Conclusions

This paper presents a novel brain tumor segmentation method. The proposed algorithm uses GLCM and MGH to extract image features. *k*NN is introduced to estimate the probability of assigning each pixel to foreground or background instead of using the gray level histogram. NCA is used here to learn optimal distance metrics which contain population and individual information. A new cost function is constructed through these probabilities and is optimized using graph cuts. The proposed method handles the problems caused by the uneven gray level distribution of the

tumors and even can achieve good results with few manual interactive seeds. Experimental results demonstrate that the proposed method is robust to brain tumor segmentation considering both efficiency and accuracy.

NCA is a learning algorithm, which spends more computing time. Future work includes speeding up our algorithm using graphic processing unit (GPU). Compute unified device architecture on GPU allows the design and implementation of single instruction and multiple data parallel algorithms on a much more simple and fast way. Using this technique, our algorithm will speed up significantly. And we know that information in different pulse sequences of MR scans can be fully used in brain tumor segmentation, our future work also includes extracting features from T1, T2, T1C, and FLAIR sequences of MR scans to handle the segmentation problems.

Acknowledgment The research was supported by the grants from the National Basic Research Program of China (973 Program, no. 2010CB732505), National Science and Technology Pillar Program of China (no. 2012BAI14B02), National Natural Science Funds of China (NSFC, no. 81101109, no.81000613), the National Major Scientific Instrument Equipment Development Program of China under Grant (2011YQ03011404).

References

- Weizman L, et al: Automatic segmentation, internal classification, and follow-up of optic pathway gliomas in MRI. *Med Image Anal* 16:177–188, 2012
- Akselrod-Ballin A, et al: Automatic segmentation and classification of multiple sclerosis in multichannel MRI. *IEEE Trans Biomed Eng* 56:2461–2469, 2009
- Garcia-Lorenzo D, Prima S, Arnold DL, Collins DL, Barillot C: Truncated-likelihood estimation for focal lesions and tissue segmentation in multisequence MRI for multiple sclerosis. *IEEE Trans Med Imaging* 30:1455–1467, 2011
- Clark MC, Hall LO, Goldgof DB, Velthuizen R, Murtagh FR, Silbiger MS: Automatic tumor segmentation using knowledge-based techniques. *IEEE Trans Med Imaging* 17:187–201, 1998
- Mazzara GP, Velthuizen RP, Pearlman JL, Greenberg HM, Wagner H: Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. *International Journal of Radiation Oncology, Biology, Physics* 59:300–312, 2004
- Fletcher-Heath LM, Hall LO, Goldgof DB, Murtagh FR: Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artif Intell Med* 21:43–63, 2001
- Prastawa M, Bullitt E, Ho S, Gerig G: A brain tumor segmentation framework based on outlier detection. *Med Image Anal* 8:275–283, 2004
- Khotanlou H, Colliot O, Atif J, Bloch I: 3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models. *Fuzzy Set Syst* 160:1457–1473, 2009
- Andrews S, Hamarneh G, Saad A: Fast random walker with priors using precomputation for interactive medical image segmentation. *Med Image Comput Comput-Assist Interv: MICCAI Int Conf Med Image Comput Comput-Assist Interv* 13:9–16, 2010
- Mortensen EN, Barrett WA: Intelligent scissors for image composition. *International Conference on Computer Graphics and Interactive Techniques*:191–198, 1995
- Mishra A, Wong A, Zhang W, Clausi D, Fieguth P: Improved interactive medical image segmentation using Enhanced Intelligent Scissors (EIS). *Conf Proc: Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Conf* 2008:3083–3086, 2008
- Ruzon M, Tomasi C: Alpha estimation in natural images. *Conference proceedings: IEEE Conference on Computer Vision and Pattern Recognition* 2000
- Grady L: Random walks for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 28:1768–1783, 2006
- Vezhnevets V, Konouchine V: “GrowCut”—Interactive multi-label N-D image segmentation by cellular automata. *Graphicon*: 150–156, 2005
- Cousty J, Bertrand G, Najman L, Couprie M: Watershed cuts: minimum spanning forests and the drop of water principle. *IEEE Trans Pattern Anal Mach Intell* 31:1362–1374, 2009
- Boykov Y, Jolly MP: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. *Int Conf Comput Vis* 1:105–112, 2001
- Rother C, Kolmogorov V, Blake A: “GrabCut”—Interactive foreground extraction using iterated graph cuts. *ACM Trans Graph* 23:309–314, 2004
- Li Y, Sun J, Tang CK, Shum HY: Lazy snapping. *ACM Siggraph*: 303–308, 2004
- Boykov Y, Funka LG: Graph cuts and efficient N-D image segmentation. *Int J Comput Vis* 70:109–131, 2006
- Corso JJ, Sharon E, Dube S, El-Saden S, Sinha U, Yuille A: Efficient multilevel brain tumor segmentation with integrated Bayesian model classification. *IEEE Trans Med Imaging* 27:629–640, 2008
- Feng QJ, Li SQ, Yang W, Chen WF: Tumor segmentation using the learned distance metric. *IEEE International Symposium on Biomedical Imaging*: 1958–1961, 2011
- Li SQ, Feng QJ, Chen WF, Lin YZ: A Graph cuts based interactive segmentation method of meningioma in MR brain images. *J South Med Univ* 31:1164–1168, 2011
- Goldberger J, Roweis S, Hinton G, Salakhutdinov R: Neighbourhood components analysis. *Advances in Neural Information Processing Systems*: 513–520, 2005
- Kolmogorov V, Zabih R: What energy functions can be minimized via graph cuts? *IEEE Trans Pattern Anal Mach Intell* 26:147–159, 2004
- Boykov Y, Kolmogorov V: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans Pattern Anal Mach Intell* 26:1124–1137, 2004
- Haralick RM: Textural features for image classification. *IEEE Trans Syst Man Cybern* 3:610–621, 1973
- Sheshadri HS, Kandaswamy A: Experimental investigation on breast tissue classification based on statistical feature extraction of mammograms. *Comput Med Imaging Graph* 31:46–48, 2007
- Iskan Z, Yüksel A, Dokur Z: Medical image segmentation with transform and moment based features and incremental supervised neural network. *Digit Signal Process* 19:890–901, 2009
- Mokji MM: Adaptive thresholding based on co-occurrence matrix edge information. *Journal of Computers* 2(8):44–52, 2007
- Gelzinisa A, Verikas A, Bacauskiene M: Increasing the discrimination power of the co-occurrence matrix-based features. *Pattern Recogn* 40:2367–2372, 2007
- Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Susstrunk S: SLIC Superpixels Compared to State-of-the-art Superpixel Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012