Available online at www.sciencedirect.com
ScienceDirect
Progress in Natural Science 19 (2009) 1373–1382
Progress in Natural Science
www.elsevier.com/locate/pnsc

An adaptive spatial clustering method for automatic brain MR image segmentation

Jingdan Zhang, Daoqing Dai *

Center for Computer Vision and Department of Mathematics, Sun Yat-Sen (Zhongshan) University, Guangzhou 510275, China

Received 13 February 2009; received in revised form 10 March 2009; accepted 13 March 2009

Abstract

In this paper, an adaptive spatial clustering method is presented for automatic brain MR image segmentation, which is based on a competitive learning algorithm – self-organizing map (SOM). We use a pattern recognition approach in terms of feature generation and classifier design. Firstly, a multi-dimensional feature vector is constructed using local spatial information. Then, an adaptive spatial growing hierarchical SOM (ASGHSOM) is proposed as the classifier, which is an extension of SOM, fusing multi-scale segmentation with the competitive learning clustering algorithm to overcome the problem of overlapping grey-scale intensities on boundary regions. Furthermore, an adaptive spatial distance is integrated with ASGHSOM, in which local spatial information is considered in the clustering process to reduce the noise effect and the classification ambiguity. Our proposed method is validated by extensive experiments using both simulated and real MR data with varying noise level, and is compared with the state-of-the-art algorithms.

Keywords: Self-organizing map; Growing hierarchical SOM; Adaptive spatial distance; Image segmentation; MR images

1. Introduction

In medical images, the brain has a complicated structure, and segmentation of brain tissues in MR images plays a crucial role [1]. In this paper, we partition brain MR images into three main tissue types: white matter (WM), grey matter (GM), and cerebrospinal fluid (CSF), which is a topic of great importance.

Clustering [2,3] is the most popular method for medical image segmentation, with artificial neural network, expectation-maximization (EM), and Fuzzy *c*-Means (FCM) algorithms being the typical methods. A common disadvantage of the EM algorithm is that the intensity distribution of brain images is modeled as a normal distribution, which is not good for noisy images. The FCM algorithm is over sensitive to noise, and many extensions to FCM [1,4] have been proposed to overcome its drawback.

In artificial neural network techniques, SOM is often used in MR image segmentation [5–7] as an unsupervised method. However, the segmentation results are deteriorated by the noise and overlapping grey-scale intensities for different tissues in MR images.

In this paper, we present an adaptive spatial clustering method for brain MR image segmentation. Firstly, we construct a multi-dimensional feature vector using local spatial information. Then, we propose an adaptive spatial growing hierarchical SOM (ASGHSOM). It is a hierarchical SOM, where neurons at the lower level can adaptively grow child SOMs at the higher level under some conditions. Moreover, an adaptive spatial similarity distance is proposed to integrate with ASGHSOM, in which local spatial information is considered in the clustering process to reduce the noise effect and the classification ambiguity.

* Corresponding author. Tel.: +86 20 8411 0141; fax: +86 20 8403 7978.
E-mail address: stsddq@mail.sysu.edu.cn (D. Dai).

1002-0071/$ - see front matter
doi:10.1016/j.pnsc.2009.03.012

The key features of the proposed method are:

(1) Combining global intensity with local neighboring information to provide effective feature representation.
(2) Fusing multi-scale segmentation with the competitive learning clustering approach to solve the problem of overlapping grey-scale intensities for different tissues.
(3) An adaptive spatial distance is proposed for ASGHSOM, considering local spatial information between image voxels to reduce the noise effect and the classification ambiguity.
(4) The map structure and size of the growing layer are adaptively determined according to the image data.
(5) Segmentation is entirely unsupervised without using any probability map to obtain the prior information. Therefore, registration processes are not required and the applicability of our proposed method can be extended to diseased brains and neonatal brains.

The rest of this paper is organized as follows. Section 2 briefly describes Kohonen's self-organizing map. In Section 3, our proposed adaptive spatial clustering method is described for MR image segmentation in detail. Experimental results are presented in Section 4 and we conclude this paper in Section 5.

2. Preliminaries

In this section, we will briefly review Kohonen's self-organizing map (SOM) [8]. It is a two-layer feedforward competitive learning network, consisting of an input layer and a single output layer of neurons, which usually form a two-dimensional array. In the output layer, each neuron i has a d-dimensional weight vector w_i. Let f_v denote the input vector of voxel v in 3D image domain I. The iterative SOM training algorithm is described as follows:

Step 1: set iteration $t = 0$ and the weights of neurons in the output layer with random values.

Step 2: randomly select a sample data vector f_v of voxel v from a training set.

Step 3: compute the distance $d_{v,i}$ between f_v and each neuron i in the output layer.

$$d_{v,i} = \|f_v - w_i\|, \quad i \in \{1, \ldots, N\}, \ \ v \in I \tag{1}$$

where N is the number of neurons in the output layer. The winning neuron, denoted by c, is the neuron with the weight vector closest to f_v.

$$c = \arg\min_i d_{v,i}, \quad i \in \{1, \ldots, N\}$$

Step 4: update the winner neuron and its neighbor neurons to move its feature vector towards the input vector. The weight-updating rule in the sequential SOM algorithm can be written as

$$w_i(t+1) = \begin{cases} w_i(t) + \alpha(t) N_t(c,i)(f_v - w_i) & \forall i \in N_c, \ \ v \in I \\ w_i(t) & \text{otherwise} \end{cases}$$

where N_c denotes the set of neighboring neurons of winner c, and $N_t(c,i)$ is the neighborhood kernel function. $\alpha(t)$ is a monotonically decreasing learning rate.

Step 5: stop if the maximum iteration is reached. Otherwise, set $t = t + 1$ and go to *step* 2.

Because of the complicated structure of the brain, most brain MR images always present overlapping grey-scale intensities for different tissues, particularly on transitional regions. Thus, there are some regions that are not partitioned accurately in SOM segmentation results. Moreover, unknown noise also influences the segmentation results.

Several improved SOM algorithms and SOM-related algorithms have been proposed in recent years. Lee and Peterson [9] proposed a self-development neural network with two levels of adaptation – namely, structure and parameter levels. However, maintaining a regular connectivity has been proven to be a difficult task during structure adaptation [9]. Hierarchical SOM (HSOM) is another variation of SOM [10]. The basic idea of HSOM is to use multiple SOMs from the low-resolution level to the high-resolution level, but the number of neurons in each layer is predefined. Growing hierarchical SOM (GHSOM) [11] is another HSOM. It can grow neurons horizontally at each level or vertically for the whole structure under some condition. In the next section, we propose an adaptive spatial growing hierarchical self-organizing map to segment T1-weighted brain MR images.

3. Adaptive spatial clustering method

In this section, a detailed description about our proposed segmentation method is presented for automatic T1-weighted brain MR image segmentation. Firstly, we extract the feature vector from each voxel. Then, we propose the adaptive spatial growing hierarchical self-organizing map (ASGHSOM) as the classifier.

3.1. Feature extraction

Generally, using multispectral images (T1-, T2-, PD-weighted images) will usually generate segmentation results superior to those using single modality images [12,13]. However, extracting enough information from a single modality image could make the segmentation based on it potentially comparable to a multimodal approach [7,14]. In this paper, we exploit T1-weighted MR images and expand each voxel into a multi-dimensional feature vector, characterizing the image data beyond simple voxel intensities. In the low-contrast medical images, more features should be extracted from local spatial information. Let I_v denote the intensity of voxel v, and the set of neighbor voxels in up, down, left, right, front, and back directions is defined as the 3D neighborhood N_v. Usually, average gradient magnitude could provide most of the local characteristics. When a voxel is on a smooth region, the average gradient value would be small enough. Otherwise, if the voxel is on the boundary region of different tissues, the

average gradient would exceed a given threshold. Therefore, the average gradient magnitude G_v is chosen as a local feature, and computed by

$$G_v = \frac{1}{|N_v|} \sum_{v' \in N_v} \Phi_{v,v'} \tag{2}$$

where $\Phi_{v,v'} = \|I_v - I_{v'}\|$ is the distance between voxel v and its neighbor v' in N_v, and $|N_v|$ is the cardinality of the neighborhood set. However, in some conditions, the high value of the average gradient also may be generated by the noise and blur in brain MR images. To overcome this problem, the mean value M_v of voxel v in its neighborhood is computed as another local feature.

$$M_v = \frac{1}{|N_v|} \sum_{v' \in N_v} I_{v'} \tag{3}$$

If the voxel is similar with most of its neighbors, the high value of the average gradient is caused by noise. Otherwise, it can be concluded that the voxel with high average gradient value is on the boundary region. Hence, for each voxel v, the feature vector is extracted and has three components: the intensity I_v, average gradient G_v, mean value M_v in its neighborhood, denoted by $f_v = [I_v, G_v, M_v]$.

3.2. ASGHSOM

In this subsection, we will introduce the adaptive spatial distance, growing algorithm, training algorithm, and classification algorithm of ASGHSOM step by step.

3.2.1. An adaptive spatial distance

SOM contains a set of neurons that construct a feature map, and each neuron i has a weight vector w_i. The input is compared with the neurons to find the winner that is most similar, using the Euclidean distance (Eq. (1)). During this process, each input voxel is assumed to be independent of every other voxel. However, for real image data, many voxels are ambiguous and cannot be classified consistently based on feature attributes alone. In this subsection, an adaptive spatial distance is proposed for the ASGHSOM algorithm, considering local spatial information between image voxels to reduce the noise effect and the classification ambiguity.

In the similarity matching phase of SOM, if voxel v is similar with its neighboring voxel v', we would like $d_{v,i}$ to be greatly influenced by $d_{v',i}$ so that they eventually have higher similarity and excite the same neuron in the output layer. Otherwise, $d_{v,i}$ should be largely independent of $d_{v',i}$. So, the actual image feature could be preserved during the noise smoothing process.

Taking all voxels in the 3D neighborhood N_v into account, we define a new adaptive spatial distance equation instead of Eq. (1) in the SOM algorithm that measures the similarity between the input feature vector f_v and neuron i in the output layer, as

$$d^*_{v,i} = \frac{1}{|N_v|} \sum_{v' \in N_v} [d_{v,i}\lambda_{v,v'} + d_{v',i}(1 - \lambda_{v,v'})] \tag{4}$$

where N_v is defined as the 3D neighborhood with the set of neighbor voxels in up, down, left, right, front, and back directions. $\lambda_{v,v'}$ is a weighting factor controlling the influence degree of the neighboring voxel v' on the center voxel v, with ranges between zero and one, defined by

$$\lambda_{v,v'} = \frac{1}{1 + \exp(-\beta(\Phi_{v,v'} - G_v))} \tag{5}$$

The parameter G_v is the average gradient of center voxel v, computed by Eq. (2), and $\Phi_{v,v'}$ is the distance between voxel v and its neighbor v'. When the difference between voxel v and its neighbor v' is much larger than the average gradient G_v ($\Phi_{v,v'} \gg G_v$), $\lambda_{v,v'} \approx 1$, the influence v' on the center voxel v is suppressed in $d^*_{v,i}$, and v and v' are less likely to excite the same neuron. On the contrary, when v' is similar with v, i.e., $\Phi_{v,v'} \ll G_v, \lambda_{v,v'} \approx 0, v'$ greatly influences v in $d^*_{v,i}$, and v and v' are more likely to excite the same neuron.

The constant parameter β specifies the steepness of the sigmoid curve, which controls the influence degree of the neighboring voxels on the center one. We hope that the choice of β could make the clustering process not only smooth out the noise voxels but also preserve the image feature on the boundary regions. In our experiment, the default value of β is set to 0.18, and its sensitivity analysis will be discussed in Section 4.3.

The distance $d^*_{v,i}$ effectively smoothes the cluster result of voxel v by that of its neighbors. When v falls on a boundary region, it is only affected by those neighboring voxels in the same class. When v is on a smooth region and is affected by all its neighbors, the influence of each neighbor v' on v is affected by the distance $\Phi_{v,v'}$, through the weighting $\lambda_{v,v'}$. Hence, $d^*_{v,i}$ enables spatial interaction between neighboring voxels and is adaptive to the image content.

It is important that adaptive spatial distance has a noise suppression capability due to the adaptive smoothing operation. Random noise would either increase or decrease the distance of the center voxel and the distances of its neighboring voxels to neurons in the output layer randomly. When the weighted average of these distances is taken, the effect of random noise is smoothed out. The segmentation results of our method using adaptive spatial distance (ASGHSOM) and without it (termed as GHSOM in the following) are compared in Section 4.2.1. When the noise level is increased, the segmentation results of ASGHSOM are better than those of GHSOM.

3.2.2. Growing algorithm of ASGHSOM

Though the SOM algorithm with adaptive spatial distance could deal with the noise and blur in brain MR images, the partial volume effect originating from the low sensor resolution deteriorates the segmentation results of the SOM algorithm. Therefore, using a single scale for an entire image may not be appropriate. An adaptive growing hierarchical SOM integrated with adaptive spatial distance

(ASGHSOM) is proposed to solve the overlapping grey-scale intensities problem on the boundary regions of different tissues (partial volume effect). We construct the ASGHSOM network with the first output layer size 4×20, Gaussian neighborhood, and linear decrease of the learning ratio. Different from Kohonen's SOM, neuron-clustered voxels on the boundary regions in the first output layer (lower level) can adaptively generate its child SOM at the higher level for re-segmentation. Therefore, the growing conditions of ASGHSOM are defined to find which neurons cluster voxels on the boundary regions.

At each layer of ASGHSOM, each neuron i has a three-dimensional weight vector $w_i = [w_{iI}, w_{iG}, w_{iM}]$ corresponding with the input feature vector $f_v = [I_v, G_v, M_v]$. So $w^n_{iI}, w^n_{iG}, w^n_{iM}$ denote the intensity centroid, average gradient centroid, and neighborhood mean centroid of voxels clustered into the ith neuron of the nth output layer of ASGHSOM, respectively. The growing conditions are denoted as follows:

If $w^1_{iG} > G\_Thresh$ and $\left|w^1_{iM} - w^1_{iI}\right| > M\_Thresh$ are all satisfied, neuron i in the first output layer spawns two neurons as a child SOM in the second output layer. *G_Thresh* and *M_Thresh* are constant thresholds taken according to the experimental results, and we will discuss the choice of them in Section 4.3.

Satisfying the growing conditions means that the high value of the average gradient is not influenced by noise, and the voxels clustered into neuron i in the first output are on the boundary regions. So, generating a child SOM in the second output layer implies that the voxels on the boundary regions could be segmented again in a higher scale.

Combining the concepts of self-organization and topographic mapping with multi-scale segmentation, ASGHSOM could deal with the overlapping grey-scale intensities problem from the low-resolution level to the high-resolution level. Comparing with other GHSOM, ASGHSOM has one vertically growing layer and does not grow neurons horizontally to simplify the growing process and train the network faster. In the growing layer of ASGHSOM, the network structure and size are adaptively determined by the image data themselves. Fig. 1 shows the structure chart of ASGHSOM.

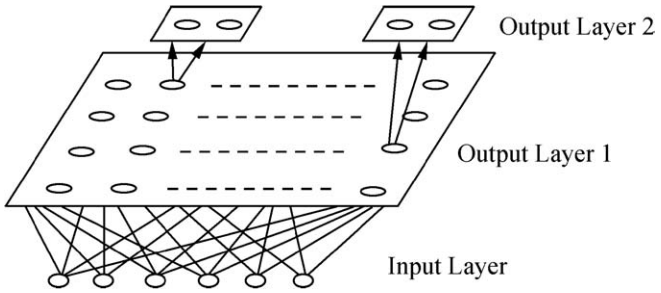

Fig. 1. The structure chart of ASGHSOM that grows neurons hierarchically when needed.

3.2.3. Training algorithm of ASGHSOM

In the training phase, 30% voxels (not including voxels of the background and extra-cranial tissues) are chosen randomly from the tested T1-weighted brain MR image data as the training samples. ASGHSOM will be re-trained on each individual 3D MR image data rather than trained only once. So rigid standardization and tight quality control of data acquisition, coupled with explicit intensity, are not needed in our experiments.

When training each SOM in ASGHSOM, we implement the original SOM algorithm except for substituting the distance $d_{v,i}$ with adaptive spatial distance $d^*_{v,i}$. The child SOMs are trained with data associated with their mother neurons. ASGHSOM completes the training of SOMs at the first output layer and then proceeds to train SOMs at the second output layer. Fig. 2 shows the flow chart of the ASGHSOM training algorithm.

The ASGHSOM training algorithm is summarized as follows:

Step 1: Set level $n = 1$, iteration $t_1 = 0$, and initialize the weights at the first output layer $w^1_i = \left[w^1_{iI}, w^1_{iG}, w^1_{iM}\right]$ with random values and the maximum training times T_1 of the first output layer.

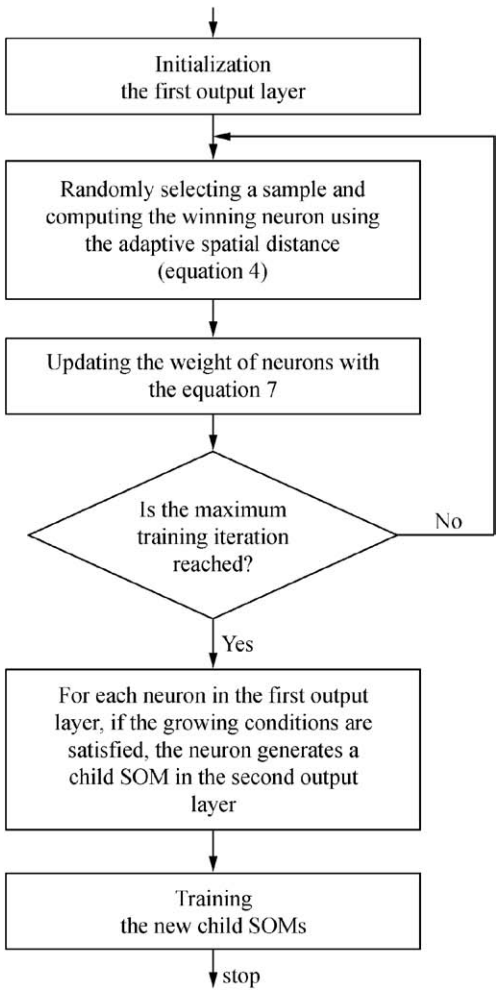

Fig. 2. The flow chart of the ASGHSOM training algorithm.

Step 2: Randomly select a sample feature vector f_v of voxel v from the training set.

Step 3: Compute the adaptive spatial distance $d^*_{v,i}$ between v and each neuron i in the first output layer according to Eq. (4). Obtain the winning neuron c.

$$c = \arg\min_i d^*_{v,i}, \quad i \in \{1, \ldots, N\} \tag{6}$$

where N is the number of neurons in the first output layer.

Step 4: Update the winner neuron and its neighbor neurons to move its weight vector towards the input vector. The weight-updating rule in the sequential ASGHSOM algorithm, which is different from that of SOM, is rewritten as

$$w^1_i(t_1+1) = \begin{cases} w^1_i(t_1) + \alpha(t_1) N_{t_1}(c,i)\mathrm{Sgn}(f_v - w^1_i(t_1)) d^*_{v,i} & \forall i \in N_c(t_1),\ v \in I \\ w^1_i(t_1) & \text{otherwise} \end{cases} \tag{7}$$

where $N_{t_1}(c,i) = \exp\left(-\frac{\|r_i - r_c\|^2}{2N^2_c(t_1)}\right)$ is the Gaussian neighborhood kernel function. r_i is the coordinate of neuron i on the first output layer and $N_c(t_1) = N_0\left(1 - \frac{t_1}{T}\right)$ is the kernel width. The parameter $\alpha(t_1) = \alpha_0\left(1 - \frac{t_1}{T}\right)$ is a monotonically decreasing learning rate.

Step 5: Set $t_1 = t_1 + 1$. If $t_1 = T_1$, go to *step* 6. Otherwise, go to *step* 2.

Step 6: Recursive loop: for each neuron i in the first output layer, if the growing conditions $w^1_{iG} > G\_Thresh$ and $|w^1_{iM} - w^1_{iI}| > M\_Thresh$ are all satisfied, neuron i spawns two neurons representing a child SOM in the second output layer.

Step 7: Set level $n = 2$, iteration $t_2 = 0$, and the weights at the growing level $w^2_i = \left[w^2_{iI}, w^2_{iG}, w^2_{iM}\right]$ with random values.

Step 8: The child SOMs are trained with data associated with their mother neurons, and the training method is the same as the neurons in the first output layer.

Step 9: Stop if the maximum iteration times of the growing layer is reached. Otherwise, set $t_2 = t_2 + 1$ and go to 8.

3.2.4. Classification algorithm of ASGHSOM

When SOM is used for clustering, finding clusters becomes a crucial task. A neuron is iteratively updated during training based on the learning vectors so a well-trained SOM represents a distribution of the input data over a two-dimensional surface preserving topology. In this context, a cluster can be defined as a group of neurons with short distances between them and long distances to the other neurons [15].

In ASGHSOM, the leaf neurons representing the final clustering results are scattered among different output layers. We should reconstruct a new map with the leaf neurons, so the algorithm 1 in [15] can be adopted to cluster the leaf neurons in ASGHSOM with three tissue classes.

The structure of the new map is the same as that of the first output layer, and the locations of leaf neurons in the new map depend on the locations of themselves (they are in the first output layer without any child SOM) or their mother neurons in the first output layer. For instance, let neuron $(2,3)$ represent the neuron in the second column, third row of the first output layer. If this neuron has a child SOM, its location will be substituted by its child neurons with locations $(1.75,3)$ and $(2.25,3)$ in the new map. Otherwise, its location in the new map is $(2,3)$, the same as that in the first output layer. By this way, all the leaf neurons are set in a new map. Fig. 3 shows a simulated reconstructed map, where the dots represent the well-trained leaf neurons, and dashed lines divide the regions associated with each neuron of the first output layer.

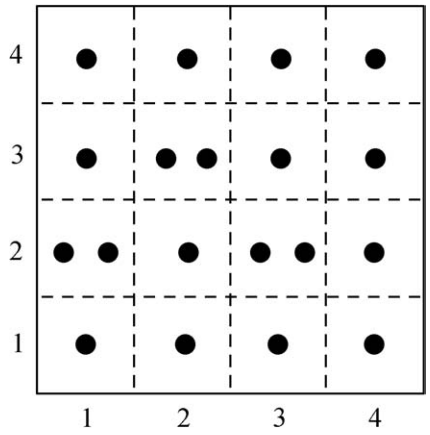

Fig. 3. A simulated new map with the dots represent the well-trained leaf neurons and the dashed lines dividing the regions associated with each neuron in the first output layer.

In ASGHSOM, the leaf neurons related to the same tissue have an approximate weight component w_{iI}. In Algorithm 1 proposed by Samsonova [15], we use w_{iI} as the attribute of leaf neuron i to label its tissue class.

4. Experimental results

In this section, we present a set of experiments to validate our proposed clustering method. At first, we introduce the database and two volumetric overlap metrics exploited in our experiments. Then, the simulation test is presented to evaluate the efficiency of our clustering method and followed with the discussion on the parameters selection. At last, our proposed clustering method is compared with state-of-the-art methods on real brain MR image data.

4.1. Database and volumetric overlap metrics

4.1.1. Database

Our proposed method is tested on both simulated MR images and real MR image data. The simulated MR image database is obtained from the BrainWeb Simulated Brain Database[1] at the McConnel Brain Imaging Center of the Montreal Neurological Institute (MNI), McGill University. In our experiments, we exploit the simulated MR

[1] http://www.bic.mni.mcgill.ca/brainweb

brain image data, with a slice thickness of 1 mm, volume size of $217 \times 181 \times 181$ voxels (pixels), 3–9% noise levels and 20% intensity inhomogeneity. The real MR image database is taken from the Center for Morphometrics Analysis, Massachusetts General Hospital Repository (IBSR).[2] Twenty normal T1-weighted real MR data sets are exploited from IBSR with a slice thickness of 3.1 mm. These databases all provide the ground truth (expert's segmentation) for quantitative validation.

In the experiments, the number of tissue classes for segmentation is set to three, which corresponds to CSF, GM and WM. Background voxels are ignored in the computation, and extra-cranial tissues are removed from all images prior to segmentation. For real image data, this can be done using any of the techniques reported in Refs. [16,17].

4.1.2. Two volumetric overlap metrics

To quantify the overlap between the segmentation result and the ground truth for each tissue, two volumetric overlap metrics are used in our experiments. The first one ρ_1 is described by Dice [18], and the other one is the Tanimoto coefficient ρ_2 [19]. For a given brain tissue i, $i = 1, 2, 3$ for CSF, GM, and WM respectively, suppose that A_i and B_i represent the sets of voxels labeled into class i by the ground truth and by the segmentation result respectively. $|A_i|$ denotes the number of voxels in A_i. The Dice metric ρ_1 is an intuitive and plain one to consider the matching area between A_i and B_i, defined as

$$\rho_1 = \frac{2|A_i \cap B_i|}{|A_i| + |B_i|}. \tag{8}$$

The other one, the Tanimoto metric ρ_2 is computed by

$$\rho_2 = \frac{|A_i \cap B_i|}{|A_i \cup B_i|}, \tag{9}$$

where $|A_i \cap B_i|$ denotes the number of voxels classified as class i by both the ground truth and the segmentation result, and $|A_i \cup B_i|$ denotes the number of voxels classified as class i by either the ground truth or the segmentation result. Usually, for a given segmentation result and the ground truth, the Dice metric ρ_1 is no less than the Tanimoto metric ρ_2.

4.2. Simulation test

In this subsection, we test our method with T1-weighted simulated brain MR data, slice thickness of 1 mm, volume size of $217 \times 181 \times 181$, 3–9% noise levels and 20% intensity inhomogeneity.

4.2.1. The efficiency of adaptive spatial distance in ASGHSOM

To illustrate the efficiency of adaptive spatial distance in ASGHSOM, we compare the segmentation results of SOM, GHSOM (our method without using adaptive spatial distance), and ASGHSOM with the ground truth. In our experiments, the feature vector of SOM and GHSOM is the same as that of ASGHSOM, and parameter selection in the growing conditions of GHSOM is also the same as that of ASGHSOM.

Fig. 4 demonstrates segmentation results on a single slice ($z = 68$) with varying noise levels and 20% intensity inhomogeneity. The SOM method is sensitive to noise and obtains inaccurate segmentation results, particularly on the boundary regions of different tissues, which can be observed in the second column of Fig. 4. While taking multi-scale segmentation, the results of GHSOM, shown in the third column of Fig. 4, outperform those of conventional SOM. But, it is still sensitive with high level noise, demonstrated in the third and fourth rows of Fig. 4c, where some voxels in GM regions are classified into WM. Integrated with adaptive spatial distance, the ASGHSOM algorithm could suppress the noise influence and obtains more robust segmentation results even with the 9% noise level exhibited in the fourth column of Fig. 4.

ASGHSOM is tested on simulated image data, not only visually but also quantitatively. Fig. 5 shows the comparison segmentation results of the SOM, GHSOM, and ASGHSOM algorithm with the ground truth (expert's segmentation results) on each tissue (CSF, GM, and WM) according to the Dice performance metric ρ_1. In these three segmentation algorithms, we exploit the same input feature vector and the same parameters in the network. The parameters used in growing child SOMs are set as $G\_Thresh = 14, M\_Thresh = 4$, and the parameter β used in ASGHSOM is taken as 0.18. These parameters were empirically found to give a good performance, and set as the default parameters. It is clearly evident that the ASGHSOM method is robust to noise, and it can obtain smooth segmentation results with the Dice performance metric $\rho_1 > 90\%$ for each tissue even with 9% noise level and 20% intensity inhomogeneity. From Fig. 5, it can be found that for any of the three brain tissues, the maximum difference between the segmentation results of ASGHSOM with two noise levels (3% and 9%) is no more than 3% in the Dice performance metric. But, the maximum difference between two noise levels (3% and 9%) is more than 5% for GHSOM and 8% for SOM.

4.2.2. The efficiency of hierarchical structure in ASGHSOM

ASGHSOM is a hierarchical SOM, which could overcome the problem of overlapping grey-scale intensities for different tissues. To illustrate the efficiency of the hierarchical structure of ASGHSOM, we compare the segmentation results of ASSOM (our method without hierarchical structure), and ASGHSOM with the ground truth on the simulated brain MR images with 9% noise level and 20% intensity inhomogeneity. Table 1 demonstrates quantitative segmentation results according to the Dice overlap metric. The comparison results show that ASGHSOM is obviously superior to ASSOM.

[2] http://www.cma.mgh.harvard.edu/ibsr

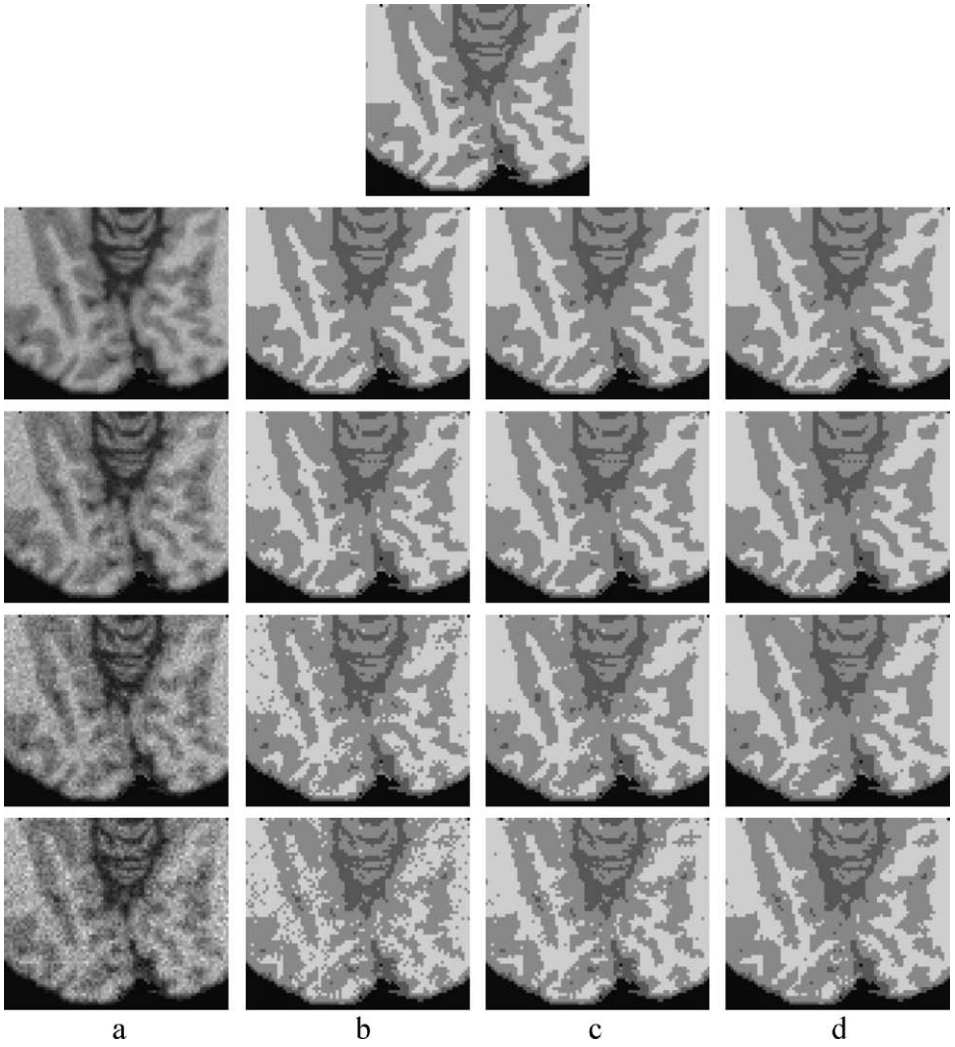

Fig. 4. Comparing the segmentation results of simulated T1-weighted MR slice 68 with different noise levels and 20% intensity inhomogeneity using SOM, GHSOM, and ASGHSOM. Upper image: ground truth. Upper row: 3% noise level. Second row: 5% noise level. Third row: 7% noise level. Fourth row: 9% noise level. Columns: (a) original image; (b) SOM algorithm; (c) GHSOM algorithm; (d) ASGHSOM algorithm.

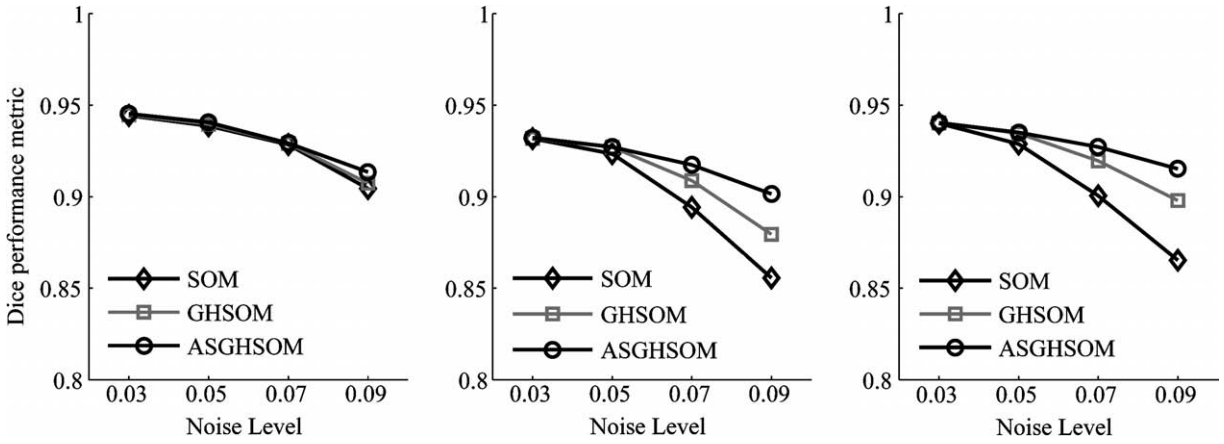

Fig. 5. Comparison of conventional SOM, GHSOM, and ASGHSOM with the ground truth on simulated MR data with different noise levels and 20% spatial inhomogeneity, measured with the Dice performance metric ρ_1. (a) CSF; (b) GM; (c) WM.

4.3. Parameter sensitivity analysis

In ASGHSOM, child SOMs have been generated to partition the voxels on the boundary regions in a higher scale. During this process, two parameters (*G_Thresh* and *M_Thresh*) need to be determined that describe the characteristic of the boundary regions. Another parameter β is also important in ASGHSOM. It specifies the steepness

Table 1
The quantitative comparison results of ASSOM and ASGHSOM with the ground truth on simulated brain MR data with 9% noise level and 20% intensity inhomogeneity, according to the Dice overlap metric.

	CSF	GM	WM
ASSOM	0.9187	0.8824	0.8977
ASGHSOM	0.9134	0.9015	0.9152

of the sigmoid curve, which controls the influence degree of the neighboring voxels on the center voxel. In this subsection, a sensitivity analysis on these parameters is discussed.

G_Thresh and *M_Thresh*: these parameters control the growing process of ASGHSOM. Higher values of *G_Thresh* and *M_Thresh* mean less child SOMs will be generated from the first output layer. This leads to some voxels on the boundary regions not being partitioned again in a higher scale to obtain an accurate segmentation result. While, lower values mean more neurons in the first output layer will generate their child SOMs. So the time for the growing and training process will be longer.

β: when β is smaller, the noise-smoothing effect will be suppressed. On the contrary, the image details would be smoothed out. In ASGHSOM, we hope that the choice of β could make the clustering process not only smooth out the noise voxels but also preserve the image feature in the local regions.

Segmentation performance is measured across several cases on the simulated image data with 9% noise level and 20% intensity inhomogeneity with the Dice performance metric ρ_1. We keep one parameter fixed while changing the others, which is similar with [2,14]. Table 2 lists the segmentation performance in variation parameters *G_Thresh* and *M_Thresh* with $\beta = 0.18$. The default parameter set with $G\_Thresh = 14$ and $M\_Thresh = 4$ are empirically found to provide good results on most of the tested image data. The rows and columns exhibit gradual modification of *G_Thresh* and *M_Thresh*, respectively. When increasing the *G_Thresh* threshold from 12 to 18 and increasing the *M_Thresh* threshold from 3 to 6, minor monotonic performance deterioration is generated with about 0.6% for both GM and WM. The effect of β variation with $G\_Thresh = 14$ and $M\_Thresh = 4$ is exhibited in Table 3. The performance deterioration with varying β is no more than 3% for both GM and WM. As can be seen, the ASGHSOM algorithm is insensitive to these parameters.

Table 2
Sensitivity of the ASGHSOM algorithm performance in variation parameters *G_Thresh* and *M_Thresh* measured on the simulated brain MR image with 9% noise level and 20% intensity inhomogeneity, according to the Dice performance metric.

G_Thresh	*M_Thresh*			
	3	4	5	6
12	0.902, 0.915	0.902, 0.915	0.901, 0.913	0.896, 0.909
14	0.902, 0.915	0.902, 0.915	0.901, 0.913	0.896, 0.909
16	0.896, 0.909	0.896, 0.909	0.896, 0.909	0.896, 0.909
18	0.896, 0.909	0.896, 0.909	0.896, 0.909	0.896, 0.909

4.4. Performance on real MR data

Twenty normal T1-weighted real data sets with a slice thickness of 3.1 mm from IBSR are taken to validate the efficiency of our proposed method. These data sets contain varying levels of difficulty, with the worst scans consisting of low-contrast and large spatial inhomogeneities. Six different segmentation algorithms are provided as part of the IBSR website for comparison analysis, which are the adaptive map (amap), the biased map (bmap), the fuzzy *c*-means (fuzzy), the maximum a posteriori probability (map), the tree-structure *k*-mean (tskm), and the maximum-likelihood (mlc) algorithm. The overlap metric used by the IBSR repository is the Tanimoto performance metric ρ_2. Fig. 6 demonstrates the comparison results of our method and these six methods with the ground truth (expert's segmentation results) on 20 real data sets (*x*-axis) from IBSR, using the Tanimoto performance metric (*y*-axis). The bold line in Fig. 6 corresponds to our algorithm. The segmentation results indicate our clustering method outperforms six different methods provided by the IBSR website for both GM and WM on most of the data sets.

For comparison with the state-of-the-art segmentation algorithm, we quote the comparison analysis results of Marroquin's algorithm and the constrained GMM algorithm in [2]. Marroquin's algorithm presents a fully automatic Bayesian segmentation algorithm [14], and the constrained GMM (CGMM) is an automatic segmentation method using the constrained Gaussian Mixture Model [2]. To coincide with the experimental data sets in [2,14], 18 real MR data sets of 20 normal T1-weighted real data sets are exploited except the data sets of 202_3 and 4_8. The comparison results of various segmentation methods on 18 real brain MR data are shown in Table 4, according to the mean and standard deviation of the Tanimoto performance metric. Note that a higher mean value indicates a better correspondence to the ground truth, and a lower standard deviation value indicates robust segmentation results on data sets with varying levels of difficulty. Our method produces better segmentation results than six different methods provided from the IBSR website and obtains similar results with Marroquin and CGMM methods on mean value.

Table 3
Sensitivity of the ASGHSOM algorithm performance in variation parameter β measured on the simulated brain MR image with 9% noise level and 20% intensity inhomogeneity, according to the Dice performance metric.

β	GM	WM
0.16	0.8726	0.8888
0.17	0.8959	0.9090
0.18	0.9015	0.9152
0.19	0.8915	0.9018
0.20	0.8925	0.8944

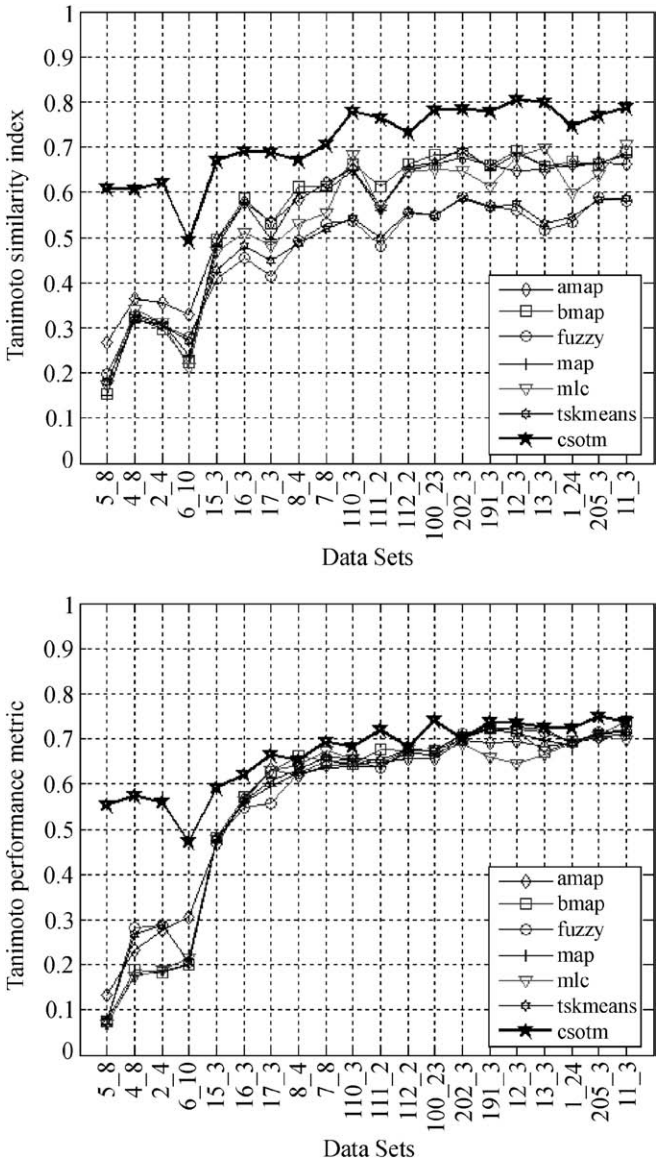

Fig. 6. Comparing the ASGHSOM algorithm and six different algorithms with ground truth on 20 real MR data sets from IBSR with the Tanimoto performance metric ρ_2. Up: GM; down: WM.

In our work, the ASGHSOM method is designed to overcome the noise influence and solve the overlapping grey-scale intensities problem, but not to deal with the intensity inhomogeneity problem. However, integrating the feature vector-combining local information and adaptive spatial distance, our proposed method can deal with the low level intensity inhomogeneity, such as the simulated brain MR data with 20% intensity inhomogeneity. But, as an unsupervised method without any prior information, it can not work well with high level intensity inhomogeneity, such as the real data set 6_10. Thus, the standard deviation of our segmentation results is higher than that of CGMM.

Table 4
Mean and standard deviation of the Tanimoto performance metric for various segmentation methods, calculated over 18 real brain MR data from the IBSR repository.

Method	Source	GM		WM	
		Mean	Std.Dev.	Mean	Std.Dev.
Adaptive MAP	IBSR	0.57	0.13	0.58	0.17
Biased MAP	IBSR	0.56	0.17	0.58	0.58
Fuzzy c-means	IBSR	0.47	0.11	0.58	0.19
Maximum a posteriori probability	IBSR	0.55	0.16	0.57	0.20
Tree-structure k-means	IBSR	0.48	0.12	0.58	0.19
Maximum-likelihood	IBSR	0.54	0.16	0.57	0.20
Marroquin	Ref. [14]	0.66	0.10	0.68	0.10
Constrained GMM	Ref. [2]	0.68	0.04	0.66	0.06
Our proposed method (ASGHSOM)		0.69	0.08	0.66	0.07

Fig. 7 shows two slice images ($z = 19$ and $z = 29$) of real MR data 6_10 with high level intensity inhomogeneity. According to the ground truth shown in Fig. 7b and d, the voxels in the red squares of Fig. 7a and c belong to the same tissue class (WM), but the difference of their intensity values is more than 30. Moreover, the intensity values of GM in Fig. 7a are similar with the intensity values of WM in Fig. 7c.

Generally, without prior information (probability map) or user interaction, it is difficult to overcome the problem of high level intensity inhomogeneity. Marroquin [14] has proposed separate parametric, smooth models for the intensity of each class to solve the high level intensity inhomogeneity problem, but the prior probabilities for each class at each voxel location are computed based on the

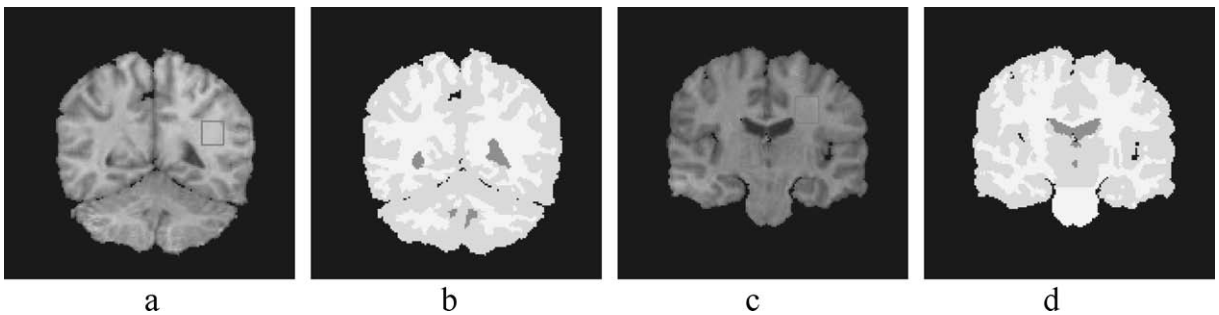

Fig. 7. Two slice images ($z = 19$ and $z = 29$) of real MR data 6_10. (a) one slice image ($z = 19$); (b) the ground truth of (a); (c) the other slice image ($z = 29$); (d) the ground truth of (c).

brain atlas (probability map). CGMM [2] is an unsupervised method without using any atlas, but the intensity inhomogeneity is corrected based on fourth-degree polynomial fitting as part of an EM algorithm [20], which obtains the prior information from the atlas. As an unsupervised clustering method without prior information, our proposed method can not solve the problem of the high level intensity inhomogeneity problem only depending on the local information (the feature vector combining local information and adaptive spatial distance).

The processing time of our method depends on the number of child SOMs on the second output layer, which is generated dynamically according to the content of image data. The average processing time is about 5 min for a single real brain volume, executed on a 1.6 GHz pentium4 processor of a PC machine, with 512 MB memory, while, it takes about 7 min for CGMM [2], executed on a single 3.0 GHz pentium4 processor, with 1 GB memory. Song et al. [7] segmented the brain MR slice with a hybrid algorithm, and their average processing time on a single slice is 2 min, running on a personal computer with a Pentium 3.0 GHz processor and 512 MB memory.

5. Conclusions

In this paper, we introduce an adaptive spatial clustering approach for brain MR image segmentation. Based on the competitive learning method – SOM, ASGHSOM is proposed as the classifier, which fuses the competitive learning clustering algorithm with multi-scale segmentation to overcome the problem of overlapping grey-scale intensities on the boundary regions of different tissues. Furthermore, an adaptive spatial distance is integrated with ASGHSOM, in which local spatial information is considered in the clustering process to reduce the noise effect and the classification ambiguity.

Our proposed clustering method is validated by extensive experiments using both simulated and real MR data. To illustrate its efficacy, we compare the segmentation results with SOM and GHSOM on simulated MR data with varying noise levels, both visually and quantitatively. The results in Fig. 4 show that our method is more robust to noise than others. In addition, quantitative comparison with state-of-the-art methods (Marroquin's algorithm and the CGMM method) and six classic approaches provided by the IBSR website is performed on real T1-weighted MR data. Table 4 demonstrates that our method outperforms six classic methods, and has similar results with Marroquin's algorithm and the CGMM method on the mean value, but higher standard deviation value than the CGMM method.

Our proposed method is designed to overcome the noise influence and solve the overlapping grey-scale intensities problem. It can also deal with the image data with low level intensity inhomogeneity, but it does not work well with higher level intensity inhomogeneity. In further research, we will extend our work to deal with the intensity inhomogeneity problem.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 60575004 and 10771220) and the Ministry of Education of China (Grant No. SRFDP-20070558043).

References

[1] Xue JH, Pizurica A, Philips W, et al. An integrated method of adaptive enhancement for unsupervised segmentation of MRI brain images. Pattern Recogn Lett 2003;24:2549–60.
[2] Greenspan H, Ruf A, Goldberger J. Constrained Gaussian mixture model framework for automatic segmentation of MR brain images. IEEE Trans Med Imag 2006;25(9):1233–45.
[3] Huang WL, Jiao LC. Artificial immune kernel clustering network for unsupervised image segmentation. Prog Nat Sci 2008;18:455–61.
[4] Liew AWC, Yan H. An adaptive spatial fuzzy clustering algorithm for 3D MR image segmentation. IEEE Trans Med Imag 2003;22(9):1063–74.
[5] Lin KCR, Yang MS, Liu HC, et al. Generalized Kohonen's competitive learning algorithm for ophthalmological MR image segmentation. Magn Reson Imag 2003;21:863–70.
[6] Lange O, Meyer-Baese A, Hurdal M, et al. A comparison between neural and fuzzy cluster analysis techniques for functional MRI. Biomed Signal Process Control 2006;1(3):243–52.
[7] Song T, Gasparovic C, Andreasen N, et al. A hybrid tissue segmentation approach for brain MR images. Med Biol Eng Comput 2006;44:242–9.
[8] Kohonen T. The selforganizing maps. Proc IEEE 1990;78(9): 1464–80.
[9] Lee TC, Peterson AM. Adaptive vector quantization using a self-development neural network. IEEE Trans Neural Netw 1997;8(1):43–53.
[10] Marsland S, Shapiro J, Nehmzow U. A self-organizing network that grows when required. Neural Netw 2002;15(8–9):1041–58.
[11] Rauber A, Merkl D, Dittenbach M. The growing hierarchical self-organizing map: Exploratory analysis of high-dimensional data. IEEE Trans Neural Netw 2002;13(6):1331–41.
[12] Alfano B, Brunetti A. Unsupervised automated segmentation of the normal brain using a multispectral relaxometric magnetic resonance approach. Magn Reson Imag 1997;37(1):84–93.
[13] Reddick WE, Glass JO, Cook EN, et al. Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks. IEEE Trans Med Imag 1997;16:911–8.
[14] Marroquin JL, Vemuri BC, Botello S, et al. An accurate and efficient Bayesian method for automatic segmentation of brain MRI. IEEE Trans Med Imag 2002;21(8):934–45.
[15] Samsonova EV, Kok JN, Ijzerman AP. TreeSOM: cluster analysis in the self-organizing maps. Neural Netw 2006;19:935–49.
[16] Goldszal AF, Davatzikos C, Pham DL, et al. An image processing system for qualitative and quantitative volumetric analysis of brain images. J Comput Assist Tomogr 1998;22:827–37.
[17] Kapur T, Grimson WEL, Wells WM, et al. Segmentation of brain tissue from magnetic resonance images. Med Image Anal 1996;1:109–27.
[18] Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26(3):297–302.
[19] Duda RO, Hart PE, Stork DG. Pattern classification. New York: Wiley; 2001.
[20] Leemput KV, Maes F, Vandermeulen D, et al. Automatic model-based tissue classification of MR images of the brain. IEEE Trans Med Imag 1999;18(10):897–908.