Computer Vision and Image Understanding xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation

Zhimin Wang [a,1], Qing Song [b,*], Yeng Chai Soh [b], Kang Sim [c]

[a] Department of Radiology, University of Pittsburgh, 3362 5th Avenue, Room 104, Pittsburgh, PA 15213, United States
[b] School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore
[c] Institute of Mental Health, Woodbridge Hospital, 10 Buangkok View, Singapore 539747, Singapore

ARTICLE INFO

Article history:
Received 3 June 2010
Accepted 5 May 2013
Available online xxxx

Keywords:
Image segmentation
Fuzzy c-means
Spatial information
Information clustering
MRI brain image

ABSTRACT

This paper presents an adaptive spatial information-theoretic fuzzy clustering algorithm to improve the robustness of the conventional fuzzy c-means (FCM) clustering algorithms for image segmentation. This is achieved through the incorporation of information-theoretic framework into the FCM-type algorithms. By combining these two concepts and modifying the objective function of the FCM algorithm, we are able to solve the problems of sensitivity to noisy data and the lack of spatial information, and improve the image segmentation results. The experimental results have shown that this robust clustering algorithm is useful for MRI brain image segmentation and it yields better segmentation results when compared to the conventional FCM approach.

1. Introduction

Image segmentation is one of the most important steps in many computer vision applications. The goal of image segmentation is to partition a given image into regions or parts that have a strong correlation with patterns or areas corresponding to objects in the real world. In the last few decades, data clustering techniques had been widely used in image segmentation and had achieved great successes in the areas of medical image analysis [1,2], face recognition [3,4], and machine vision [5,6]. The fuzzy clustering method, especially the fuzzy c-means (FCM) algorithm [7], is one of the most popular data clustering algorithms for image segmentation. It has been shown to be more advantageous than crisp clustering in that the crisp assignment of a data point to cluster center is not reasonable in real applications. Further, when compared to the crisp approaches, the FCM method is more tolerant to variations and noise in the input data [8].

Like many other standard unsupervised clustering methods, the FCM method suffers the lack of spatial information when applying to image segmentation applications. They rely only on the intensity distribution of the pixels, and disregard their geometric information, which make them very sensitive to noise and other artifacts introduced during the imaging process. In order to make the standard FCM algorithm [7] more robust to noise and outliers for image segmentation, many modified fuzzy clustering approaches had been reported in [9–15]. These methods are better than the conventional fuzzy clustering algorithms because the clustering decision of a single pixel is now reconsidered by the influence from the neighborhood patterns. This improvement tends to smooth out the isolated noise or image artifacts and produces more homogeneous segmentation results.

Most of these improved methods were developed in the last decade. The major contribution to the modifications of the FCM algorithm through the incorporation of the spatial information can be tracked back to 1998. Mohamed et al. [9] used a modified FCM algorithm for medical image segmentation. They introduced the spatial information into the similarity measure. The similarity measure is modified such that the pixel can be dragged closer to the cluster center if it is in homogenous regions. Although there is no more model assumption in the FCM algorithm, the drawbacks of the FCM algorithm [8] are its sensitivity to the non-descriptive initial centers and its massive computational load. Therefore, Ahmed et al. [10] used another similarity measure in a Bias-Corrected FCM (BCFCM) algorithm for bias field estimation and segmentation, which achieved a better performance when an additional term is added to the new objective function. Many researchers subsequently modified the objective functions and developed several robust FCM variants for image segmentation [11–15]. These algorithms were shown to have better performance than the standard FCM algorithm. However, some of these methods [11,13,14] depend on a fixed spatial factor which needs to be

* Corresponding author. Tel.: +65 67904530.
E-mail address: eqsong@ntu.edu.sg (Q. Song).
[1] Present address: Maccine Pte. Ltd., 10 Science Park Road, #01-05 The Alpha, Singapore Science Park II, Singapore 117684, Singapore.

1077-3142/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cviu.2013.05.001

adjusted according to the real applications. The shortcoming of using a fixed spatial parameter is evident. It makes the segmentation algorithm less robust to various images and causes the problem of over-smoothed edges. In order to overcome this, clustering algorithms that have an adaptive selection mechanism of the spatial parameters have been proposed recently [16,15]. The new similarity measure makes use of local and spatial intensity information, and so it performs better and is able to reduce the blurring effect. However, the need for experimentally adjusted parameter and the blurring problems still exist.

In this paper, we propose a new adaptive spatial fuzzy clustering algorithm, called the Adaptive Spatial Information-Theoretic Fuzzy Clustering Algorithm (ASIFC), to address the noise problem and the lack of spatial information for the conventional FCM algorithms with two-level robustness enhancements. We extend the idea in [17] to specifically enhance the robustness of FCM algorithms. The lack of spatial information problem is taken into consideration by using a novel adaptive similarity measure, while the noise and outliers are to be identified through the mutual information (MI) maximization. This new information framework allows us to add robust capability to the basic FCM algorithm while inheriting its advantages such as fast processing time and less sensitive to the density of clusters.

2. The conventional FCM algorithm and the modified spatial FCM algorithms

2.1. Conventional FCM algorithm

One approach to fuzzy clustering, probably the most commonly used, is the FCM algorithm, which was proposed by Dunn [18] in 1973 and improved by Bezdek [7] in 1981. FCM is formulated as the minimization of the following objective function with respect to the membership functions u and the centers w:

$$J_{fcm} = \sum_{i=1}^{N}\sum_{k=1}^{K} u_{ik}^{q}\|x_i - w_k\|^2, \quad 1 \leqslant q < \infty \tag{1}$$

where q is any real number greater than 1, u_{ik} is the degree of fuzzy membership of x_i in the kth cluster, and $\|\cdot\|$ is any norm expressing the similarity measure. Typically, the Euclidean distance measure is used in this paper.

Intuitively, the objective function (1) is minimized when the large membership values are assigned to input patterns are close to their nearest cluster centers and low membership values are assigned when they are far from other cluster centers.

Minimizing the objective function (1) with the hidden constraint $\sum_{k=1}^{K} u_{ik} = 1$, we have

$$\frac{\partial J_{fcm}}{\partial u_{ik}} = 0, \quad \frac{\partial J_{fcm}}{\partial w_k} = 0 \tag{2}$$

These lead to the iterative solutions

$$u_{ik} = \frac{\|x_i - w_k\|^{-2/(q-1)}}{\sum_{c=1}^{K}\|x_i - w_c\|^{-2/(q-1)}} \tag{3}$$

and

$$w_k = \frac{\sum_{i=1}^{N} u_{ik}^{q} x_i}{\sum_{i=1}^{N} u_{ik}^{q}} \tag{4}$$

Iterating through these two necessary conditions leads to the minimum of J_{fcm}. This iteration will stop when $|u_{ik}^{(p+1)} - u_{ik}^{(p)}| < \varepsilon$, where ε is a termination condition between 0 and 1, p is the iteration steps. This procedure converges to a local minimum or a saddle point of J_{fcm}.

2.2. Modified spatial FCM algorithms

In image segmentation, the neighboring pixels are highly correlated in the spatial domain. If the segmentation algorithm fails to take into account this neighborhood correlation, the segmentation result will have flaws and errors. In other words, the data clustering algorithm should incorporate the spatial information in such a way that if the neighboring pixels share similar characteristics, the center pixel should have higher probability of grouping to the same cluster as the neighboring pixels.

Modifying the objective functions is one of the most commonly used methods to incorporate the spatial information into the FCM algorithm. Many researcher modified the objective function to include the spatial constraint imposed on the center pixel. The generalized modified objective function can be expressed

$$J_s = \sum_{i=1}^{N}\sum_{k=1}^{K} u_{ik}^{q}\|x_i - w_k\|^2 + \alpha\sum_{i=1}^{N}\sum_{k=1}^{K} u_{ik}^{q}\sum_{r\in N_i}\|x_r - w_k\|^2 \tag{5}$$

where N_i is the subset of neighborhood pixels of x_i and the parameter α reflects the preference of neighboring effect. In real applications, α can be fixed or adaptive to the image content. As can be seen from (5), the modified objective function has two components. One puts the constraints that the pixels are to be close to the cluster centers. The other imposes spatial continuity, and is weighted by α. It favors a piecewise-homogenous clustering solution.

With this modification, the spatial information can be incorporated into the clustering process. However, the above modified objective function may smooth out the important structures (such as regional borders or edges) in the image, because it always imposes a significant amount of spatial constraint on the center pixel. In other words, the center pixel will be influenced by the neighboring pixels no matter it is in a homogenous region or along an edge. Hence, the blurring effect will occur when the center pixel is along an edge [19]. Some other significant modifications can be found in [14,15]. Nevertheless, the need for experimentally adjusted parameter and the blurring problems still exist.

3. The adaptive spatial information-theoretic fuzzy clustering algorithm

3.1. Introducing spatial information – level 1

Besides the lack of spatial information problem, the conventional FCM algorithm also suffers from the lack of robustness against outliers [8].

In order to overcome the outlier problem of FCM algorithm, Krishnapuram et al. [20] proposed the possibilistic clustering algorithm to achieve membership values that are possibilistic. Here, one input data point has little effect on good clusters if its possibility as noise is high, so that the clustering results are less distorted as compared to the FCM method. The modified objective function of PCM algorithm is

$$J_{PCM} = \sum_{i=1}^{N}\sum_{k=1}^{K} t_{ik}^{q}\|x_i - w_k\|^2 + \sum_{k=1}^{K}\eta_k\sum_{i=1}^{N}(1 - t_{ik})^q \tag{6}$$

where η_k is the estimated noisy cluster hyper-parameter, and there is no constraint on the possibility membership value t_{ik} other than the requirement that it should be in $[0,1]$.

The PCM algorithm utilizes different hyper-parameters η_k to form K different noisy clusters and has the capability to identify noisy data through the possibility membership values, i.e. the sum of possibilistic values t_{ik} is not equal to one in a particular cluster as compared to the FCM algorithm. However, PCM is much more sensitive to the initialization and η_k's selection as compared

Please cite this article in press as: Z. Wang et al., An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Understand. (2013), http://dx.doi.org/10.1016/j.cviu.2013.05.001

to the FCM-type algorithms, and this has limited its applications [8].

To circumvent these shortcomings in the conventional FCM algorithm and to take into consideration the incorporation of spatial information, we introduce input data distribution and an improved similarity measure $d_s(x_{i,j}, w_k)$ into the objective function of FCM algorithm. This can be done by modifying the objective function (1) as follows:

$$J_{ASIFC} = \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{K} e_{i,j} u_{k|i,j}^{q} d_s(x_{i,j}, w_k), \quad 1 \leqslant q < \infty \tag{7}$$

where

$$d_s(x_{i,j}, w_k) = \lambda_{i,j}\|x_{i,j} - w_k\|^2 + \frac{(1-\lambda_{i,j})}{N_R}\sum_{x_{r,c}\in N_{i,j}} \|x_{r,c} - w_k\|^2 \tag{8}$$

and $e_{i,j}$ is the estimated input data distribution (with constraint $\sum_{i=1}^{M}\sum_{j=1}^{N} e_{i,j} = 1$), which is set to $1/M \times N$ in the initial stage of the ASIFC algorithm. The $N_{i,j}$ is the subset of neighborhood pixels of (i,j) in a 3×3 window. N_R is the window size. $u_{k|i,j}$ is actually another form of notation for the membership value as we try to make the notations consistent with information theory. The choice of the spatial weighting factor $\lambda_{i,j}$ will be discussed later.

Note that the estimated input data distribution plays an important role in identifying the noise data and outliers. During the initial stage, since we have no preference for any input data points, the initial values for $e_{i,j}$ are usually set to $1/M \times N$ unless specified user preference is available. In the pruning phase, the data reliability/preference is assessed by MI maximization. The identified unreliable data points can be deleted (for common data clustering applications) or relabeled (for image segmentation applications) to achieve better clustering results. Embedded in the new objective function (7), the estimated input data distribution $e_{i,j}$ also plays a similar role of the possibility value as in the PCM algorithm to eliminate the outliers. Furthermore, the improved similarity measure $d_s(x_{i,j}, w_k)$ offers an optimal way to tackle the lack of contextual information from the image coordinates. Except for the fuzzifier q, there is no pre-fixed parameter in the specific objective function (7) for the clustering phase, which makes our proposed ASIFC algorithm relatively simpler to use when compared to other robust versions of fuzzy clustering algorithms like the PCM-type algorithms.

By introducing the hidden constraint $\sum_{k=1}^{K} u_{k|i,j} = 1$, and minimizing the objective function (7), we have

$$\frac{\partial J_{ASIFC}}{\partial u_{ik}} = 0, \quad \frac{\partial J_{ASIFC}}{\partial w_k} = 0 \tag{9}$$

These lead to the iterative solutions

$$u_{k|i,j} = \frac{d_s(x_{i,j}, w_k)^{-1/(q-1)}}{\sum_{c=1}^{K} d_s(x_{i,j}, w_c)^{-1/(q-1)}} \tag{10}$$

and

$$w_k = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} e_{i,j} u_{k|i,j}^{q}(\lambda_{i,j} x_i + \frac{(1-\lambda_{i,j})}{N_R}\sum_{x_{r,c}\in N_{i,j}} x_{r,c})}{\sum_{i=1}^{M}\sum_{j=1}^{N} e_{i,j} u_{k|i,j}^{q}} \tag{11}$$

As can be seen from (8), the improved similarity measure has two components. One puts the constraint that the pixels are to be close to the cluster centers. The other imposes spatial continuity. These two terms are weighted by $\lambda_{i,j}$. With this new objective function, we can choose $\lambda_{i,j}$ to be small so that the pixel $x_{m,n}$ is greatly influenced by its neighbors if it is in a homogeneous region. On the other hand, a big $\lambda_{i,j}$ will be used to reduce the neighboring effect, if $x_{m,n}$ is along the borders or edges. Thus, the over-smoothed edges problem can be solved by this adaptive weighting scheme.

In order to make the smoothing parameter $\{\lambda_{i,j}\}$ adaptive to the image content, the value of $\{\lambda_{i,j}\}$ should be adjusted according to the homogeneousness of the neighboring window. Consider a 3×3 window, for the pixel in location (i,j), we can calculate the dispersion of intensity differences between its intensity value and its neighbors':

$$d(x_{i,j}, x_{i+r,j+c}) = \|(x_{i,j} - x_{i+r,j+c})\|^2 \tag{12}$$

with $r = \{-1, 0, 1\}$, and $c = \{-1, 0, 1\}$, but $(r,c) \neq (0,0)$. Thus, the homogeneousness of this 3×3 neighboring window can be measured by the standard derivation of the intensity difference between the center pixel and its neighboring pixels, leading to

$$\begin{aligned} \eta_{i,j} &= \left(\frac{1}{N_R}\sum_{r=-1}^{1}\sum_{c=-1}^{1}(d(x_{i,j}, x_{i+r,j+c}) - \mu)^2\right)^{\frac{1}{2}} \\ \mu &= \frac{1}{N_R}\left(\sum_{r=-1}^{1}\sum_{c=-1}^{1} d(x_{i,j}, x_{i+r,j+c})\right) \end{aligned} \tag{13}$$

It is apparent from (13) that the standard derivation of the intensity difference $\{d(x_{i,j}, x_{i+r,j+c})\}$ does not take into account the magnitude of intensity transition in the edge region, and hence will over smooth obscure edge regions. Therefore, a further modification is required to eliminate the unbalanced effect on the weighting parameters between the obvious and obscure edge regions.

Since (13) measures the local variance of the intensity difference between the center pixel and its neighbors, a straightforward solution is to divide $\eta_{i,j}$ by the local variance of intensity of all the pixels over the 3×3 window, i.e.

$$\hat{\lambda}_{i,j} = \eta_{i,j}/\zeta_{i,j} \tag{14}$$

where

$$\begin{aligned} \zeta_{i,j} &= \left(\frac{1}{9}\sum_{r=-1}^{1}\sum_{c=-1}^{1}(x_{i+r,j+c} - \tau)^2\right)^{\frac{1}{2}} \\ \tau &= \frac{1}{9}\left(\sum_{r=-1}^{1}\sum_{c=-1}^{1} x_{(i+r,j+c)}\right) \end{aligned}$$

Finally, we normalize $\hat{\lambda}_{i,j}$ and obtain $\lambda_{i,j}$

$$\lambda_{i,j} = \hat{\lambda}_{i,j}/\max(\hat{\lambda}_{i,j}) \tag{15}$$

Apparently, this approach is totally adaptive to the local image content, and can be pre-computed before the clustering iteration. Furthermore, there is no experimentally adjusted parameters in the whole process. Detailed discussions about the adaptive weighting factor $\lambda_{i,j}$ can be found in [19].

3.2. Outlier further elimination by the mutual information maximization – level 2

As interpreted in Section 3.1, $e_{i,j}$ acts as an extra weighting factor or typical value to eliminate the outliers. Note that these values are chosen to be fixed for all input data points during the normal fuzzy clustering process, which implies that initially we do not have any preference about the input data points. After generating the estimated cluster center w_k and membership value $u_{k|i,j}$, we actually can reassess the input data reliability, i.e. via the input data distribution $e_{i,j}$, under the estimated w_k and $u_{k|i,j}$. After some proper iterations between these two steps, we can produce a set of new robust input data reliability values $e_{i,j}$ for every input data points. These values can be used to label/eliminate the remaining

Please cite this article in press as: Z. Wang et al., An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Understand. (2013), http://dx.doi.org/10.1016/j.cviu.2013.05.001

unreliable data points with the improved similarity measure, and thus improve the clustering/segmentation results.

To derive the robust estimation of input data distribution $e_{i,j}$, we first define the mutual information between the set of optimal membership distributions of input data points obtained via (10) and (11) and the input data distribution as follows:

$$I(\mathbf{X},\mathbf{W}) = \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{K} e_{i,j}u_{k|i,j}\log\frac{u_{k|i,j}}{e_{i,j}\sum_{i=1}^{M}\sum_{j=1}^{N}(e_{i,j}u_{k|i,j})} \tag{16}$$

Using similar procedure in the information theory [21], for fixed w_k and $u_{k|i,j}$, we can obtain the robust estimate for $e_{i,j}$ by the following function

$$e_{i,j} = \frac{e_{i,j}\exp\left(\sum_{k=1}^{K}u_{k|i,j}\log\frac{u_{k|i,j}}{\sum_{i=1}^{M}\sum_{j=1}^{N}(u_{k|i,j}e_{i,j})}\right)}{\sum_{i=1}^{M}\sum_{j=1}^{N}\exp\left(\sum_{k=1}^{K}u_{k|i,j}\log\frac{u_{k|i,j}}{\sum_{i=1}^{M}\sum_{j=1}^{N}(u_{k|i,j}e_{i,j})}\right)} \tag{17}$$

The ASIFC is basically a two-step clustering algorithm for image segmentation. First, it uses a conventional FCM algorithm together with the improved similarity measure as the starting point to explore the details of the input features from the image and group them effectively into several clusters subject to the adaptive spatial weighting factors $\lambda_{i,j}$. Given the estimated cluster centers w_k and the optimal membership values $u_{k|i,j}$ from ASIFC, the reliability of the input data points are assessed by the MI maximization. Finally, these identified unreliable data points are clustered again with a bigger neighboring window.

4. Implementation of ASIFC algorithm

The implementation of the ASIFC algorithm is described below. Suppose that we are given the feature vectors of input image $X = \{x_{1,1}, x_{1,2}, \ldots, x_{M,N}\}$ with the resolution of $L = M \times N$.

Fuzzy clustering with adaptive spatial weighting factors:

(1) Set the pre-specified number of clusters K, convergence parameter $\epsilon = 0.001$, fuzzifier $q > 0$, and the input data distribution $e_{i,j} = 1/(M \times N)$ for $i = 1$–M, $j = 1$–N.
(2) Initialization: calculate the weighting function $\lambda_{i,j}$ as discussed in Section 3.1.
(3) Perform fixed point iteration according to (10) and (11).
(4) Convergence test: If not satisfied, then go to (3).
MI Maximization for identifying the unreliable data points:
(5) Select $e_{i,j} = 1/(M \times N)$, for $i = 1, \ldots, M$, $j = 1, \ldots, N$, and $\varepsilon > 0$, then start the fixed point iteration of the robust density estimation according to steps (6)–(7).
(6) Iteratively update the input data distribution by using

$$c_{i,j} = \exp\left(\sum_{k=1}^{K}u_{k|i,j}\log\frac{u_{k|i,j}}{\sum_{i=1}^{M}\sum_{j=1}^{N}(u_{k|i,j}e_{i,j})}\right)$$

and

$$e_{i,j} = e_{i,j}\frac{c_{i,j}}{\sum_{i=1}^{M}\sum_{j=1}^{N}c_{i,j}} \tag{18}$$

(7) If

$$\ln\sum_{i=1}^{M}\sum_{j=1}^{N}e_{i,j}c_{i,j} - \ln\max_{\{i,j\}}c_{i,j} < \varepsilon \tag{19}$$

then go to (9) where $\varepsilon > 0$.
(8) Label unreliable data points, and obtain the membership of all data points for a hard clustering solution.
(9) Create a 3×3 window centered at the outlier, obtain the clustering labels for all neighboring pixels within this neighborhood window from the fuzzy clustering step.
(10) Store all label counts for each cluster into ξ_k, $k = 1, 2, \ldots, K$.
(11) Find the maximum value of ξ_k and its corresponding cluster index, i.e. $\xi_m = max(\xi_k)$.
(12) If $\xi_1 = \xi_2 = \ldots = \xi_K$, increase the current window edge by 2 more pixels, (for example, from 3×3 to 5×5), go to (12); Otherwise, update the outlier label with m.
(13) Iterate for all outlier pixels until there is no further change, then stop.

The weighting factors $\{\lambda_{i,j}\}$ do not change in each fixed point iteration, so we can pre-compute them at step (2). After this, the algorithm begins the fixed point iteration.

5. Experimental results

In this section, the results of applying ASIFC algorithm to synthetic and real images are presented. In the comparative study, we also included other competing techniques from the literature like bias-corrected fuzzy c-means (BCFCM) [10], enhanced fuzzy c-means (EnFCM) [14], Fast Generalized fuzzy c-means (FGFCM) [15], and our own Adaptive Spatial Information-Theoretic Clustering (ASIC) [19].

In all examples, the cooling factor $\alpha = 0.95$ was used unless otherwise stated. The default values for fuzzifier $q = 2$, convergence value $\epsilon = 0.001$ and maximum iteration number of 100, are used unless otherwise stated. The neighborhood window size is 3×3. The spatial parameter α in the BCFCM algorithm is obtained by searching the interval [0.5, 0.9] with a step size of 0.05 in terms of the optimal SA value. Similarly, the α in EnFCM is obtained by searching the interval [0.28] as suggested in [15]. The λ_g in FGFCM is selected from 0.5 to 6 with an increment of 0.5.

5.1. Synthetic images

In order to compare the ASIFC algorithm with the other four methods, we created a synthetic test image as shown in Fig. 1a, and the resolution of this image is 128×128 pixels and the gray-scale range is 0–255. Gaussian noise is used in this experiment, with noise levels of 3, 5, 10, and 20.

To validate the segmentation results, we use the segmentation accuracy (SA) which is computed by

$$\text{SA} = \frac{\text{number of misclassified pixels}}{\text{total number of pixels}} \tag{20}$$

Fig. 1a displays the synthetic image with Gaussian noise level of 20 added. We applied the five techniques and obtained the segmentation results as shown in Fig. 1c–g. Overall, the ASIC and ASIFC yield better results than the rest. Visually, we can observe that ASIFC performs slightly better than the ASIC algorithm. This can be further illustrated by the SA values listed in Table 1. Table 2 shows the average execution time for each algorithm implemented with MATLAB. It is obtained by averaging 10 optimal trials. From the comparison of the computational time among these five techniques, we can see that the ASIC algorithm has the largest computation load among them due to its annealing-like optimization process. The robust estimation step also contributes some computational burden to this. On the other hand, EnFCM and FGFCM perform much faster than ASIC as they are derived from fuzzy c-means algorithm. They also have faster convergence rates than the BCFCM's as they have better spatial constraints. For the proposed ASIFC algorithm, although it performs slower than the other FCM

Please cite this article in press as: Z. Wang et al., An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Understand. (2013), http://dx.doi.org/10.1016/j.cviu.2013.05.001

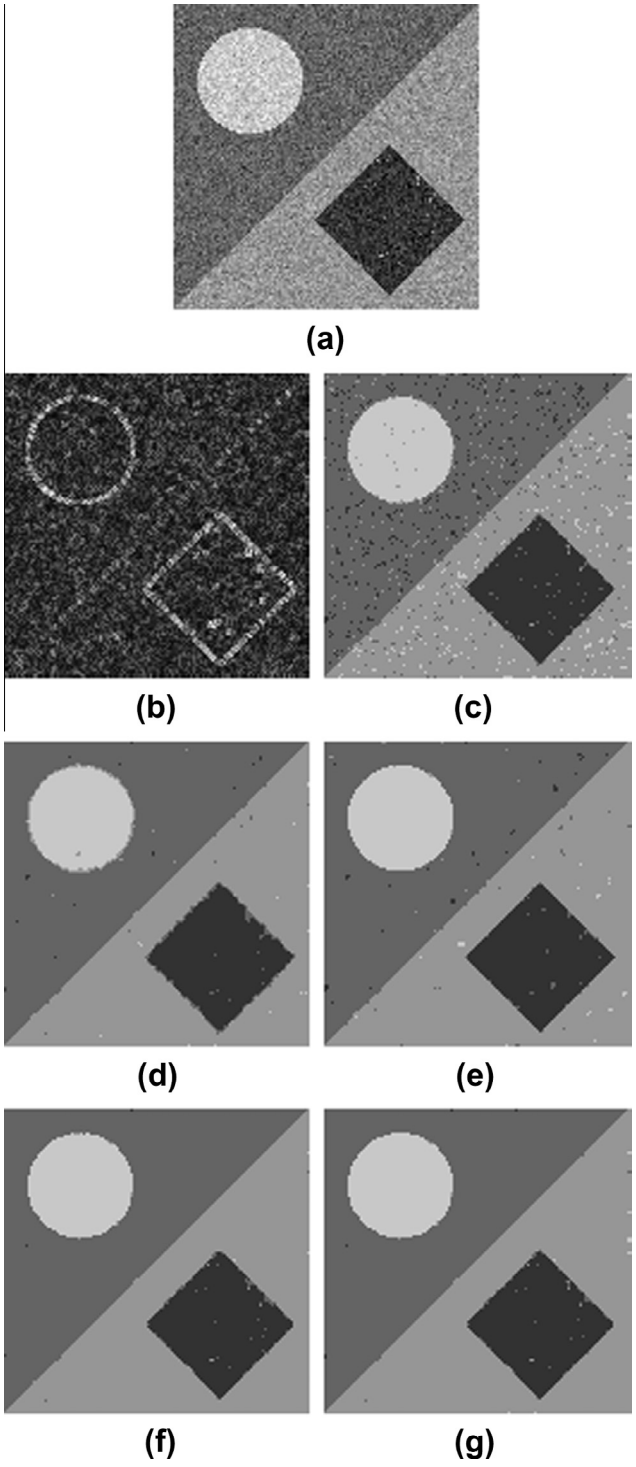


Fig. 1. Comparison of BCFCM, EnFCM, FGFCM, ASIC and ASIFC approaches on a synthetic image: (a) original synthetic image, (b) weighting image for ASIFC algorithm, (c) BCFCM segmentation result, (d) EnFCM segmentation result, (e) FGFCM segmentation result, (f) ASIC segmentation result, and (g) ASIFC segmentation result.

variants, it did improve significantly on the computational complexity by using the fuzzy clustering framework. This gap can be further narrowed by using more efficient method to compute the weighting image and estimate the outliers.

Table 1
SA values of BCFCM, EnFCM, FGFCM, ASIC and ASIFC methods for the synthetic image.

Noise levels (%)	BCFCM (%)	EnFCM (%)	FGFCM (%)	ASIC (%)	ASIFC (%)
Gaussian 3	0.03	1.25	0.01	0	0
Gaussian 5	0.04	1.39	0.02	0.02	0.02
Gaussian 10	0.27	1.53	0.01	0.22	0.22
Gaussian 20	6.14	2.05	0.95	0.73	0.64

Table 2
Average computational time (in seconds) with MATLAB on synthetic image for BCFCM, EnFCM, FGFCM, ASIC and ASIFC methods.

Noise levels (%)	BCFCM (s)	EnFCM (s)	FGFCM (s)	ASIC (s)	ASIFC (s)
Gaussian 3	0.4141	0.2360	0.2594	2.2766	1.2078
Gaussian 5	0.4891	0.2344	0.3047	3.4641	1.3563
Gaussian 10	0.5891	0.2500	0.3065	3.6172	1.6484
Gaussian 20	0.8391	0.3547	0.3766	3.3172	2.4781

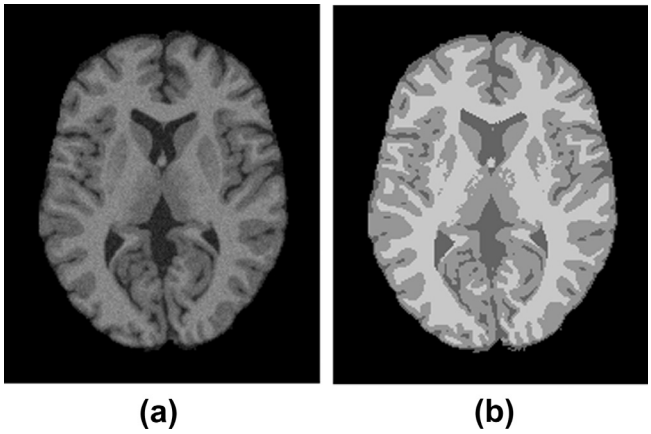


Fig. 2. Example of simulated MRI brain image segmentation: (a) sample slides of a simulated MR brain image with 5% noise, and no intensity inhomogeneity, (b) discrete anatomical model for WM, GM, and CSF.

5.2. Phantom MRI brain image

To demonstrate the performance of our method on medical images, we will test our algorithm with the Brainweb MRI Brain Database [22] and compare the results with other competing methods. In this example, we will focus on the segmentation of white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), i.e. the three main tissues of the human brain. Before we apply the clustering methods, we will perform the extracranial tissues removal by using the discrete model provided by [23]. This will help to remove the negative effect from the extracranial tissues during segmentation, and it will also make the validation easier to manipulate. The images we chose in this example are simulated T1-*weighted* MR brain images with 1-mm spacing, with 1%, 3%, 5% noise levels, and 0%, 20%, 40% inhomogeneities.

To validate the segmentation results, we use the same evaluation measure as in [24], namely, tissue segmentation accuracy (TSA):

$$TSA = \frac{2N_{CTk}}{N_{CITk} + N_{GTk}} \tag{21}$$

where N_{CTk} denotes the number of pixels that were correctly (inside the mask of ground truth) assigned to tissue k by a given method, N_{CITk} is the total (inside and outside the mask of ground truth) number of pixels assigned to tissue k, and N_{GTk} is the number of pixels

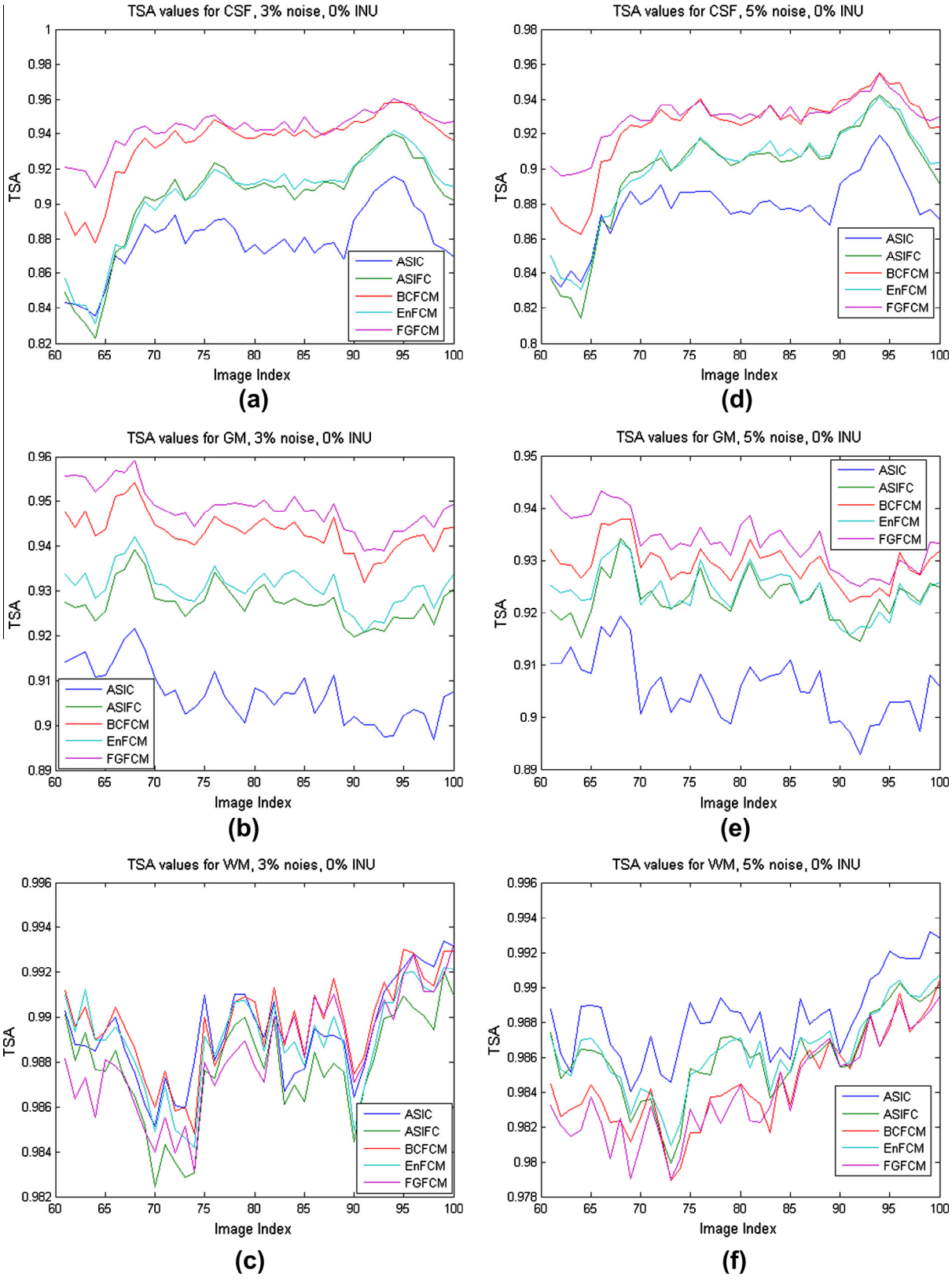
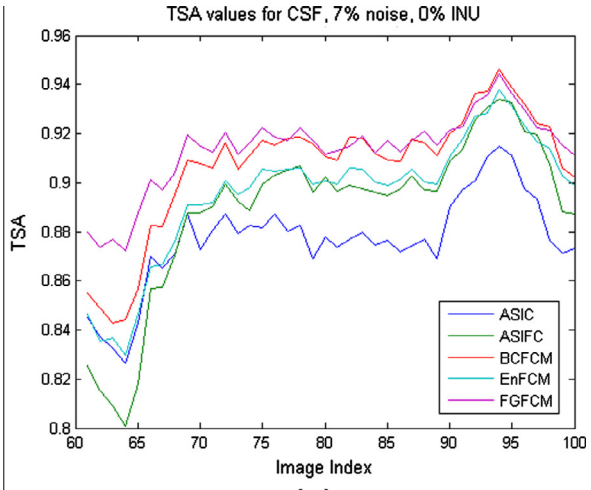


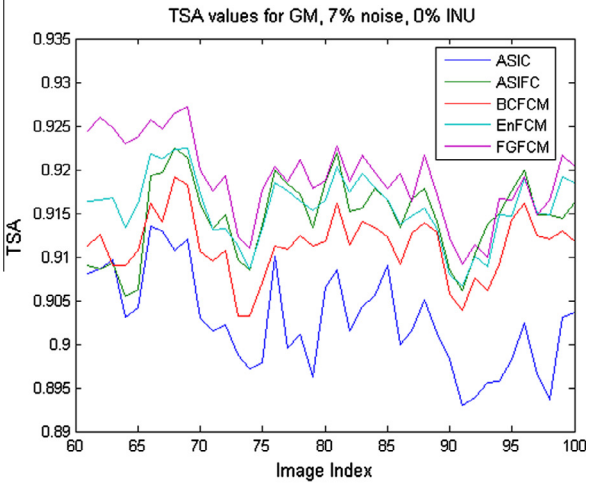
Fig. 3. Tissue Segmentation Accuracy (TSA) value of CSF, GM, and WM with different noise levels for 40 simulated MR brain images.

belonging to tissue k in the discrete anatomical model (the ground truth mask). The ground truth mask can be obtained by using the discrete model provided by [23].

Fig. 2 shows the sample slides of simulated MRI brain image, the true partial volume model of that was used to generate the simulated image. To perform a more comprehensive comparison, we selected 40 MR brain images (from index 61 to index 100) from the Brainweb database, with a variety of shapes and sizes of the three main brain tissues. Fig. 3 shows the TSA values of three main brain tissues obtained when BCFCM, EnFCM, FGFCM, ASIC and the

Please cite this article in press as: Z. Wang et al., An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Understand. (2013), http://dx.doi.org/10.1016/j.cviu.2013.05.001





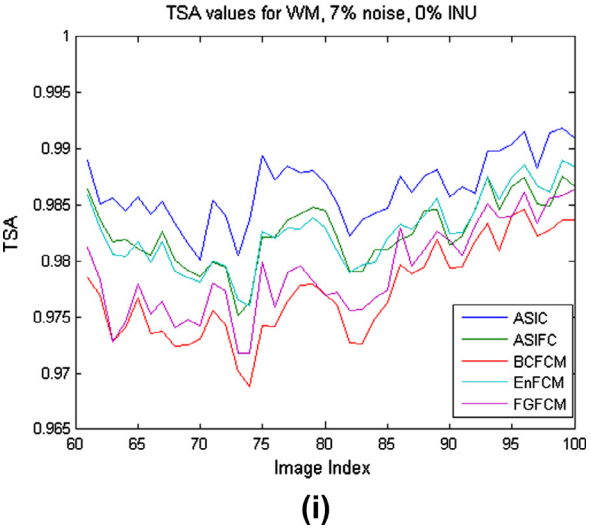


Fig. 3 (*continued*)

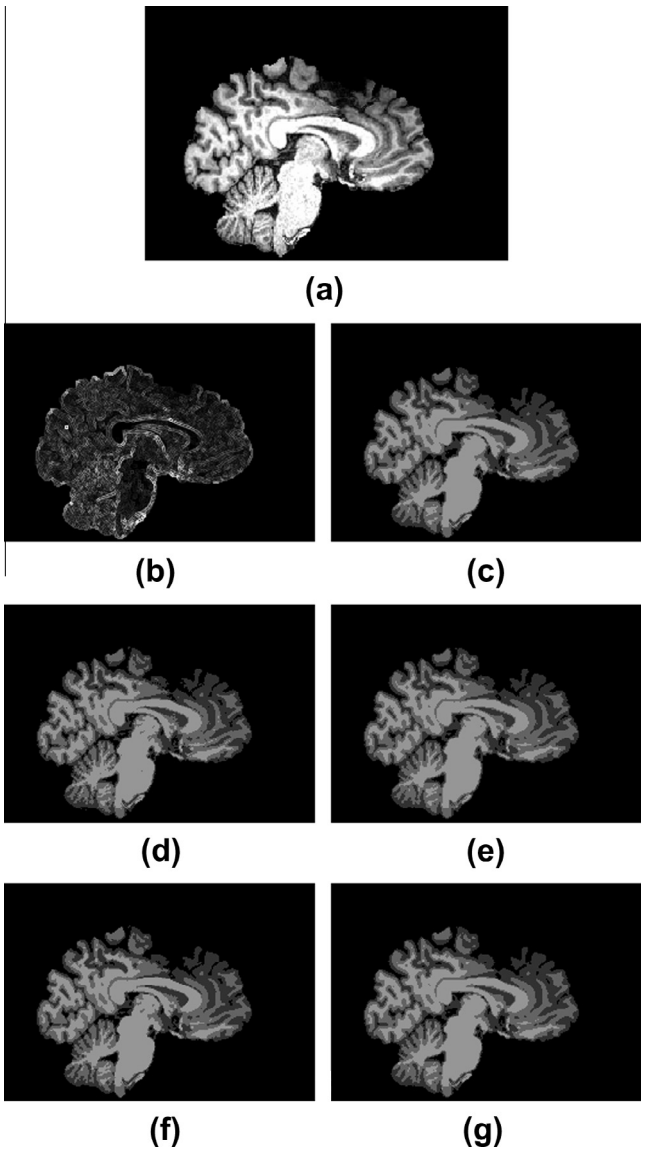


Fig. 4. Example 1: Segmentation on a sagittal view slice. (a) Sample slides from real 3D MRI brain image volume. (b) Adaptive weighting parameters of our ASIFC algorithm. (c) ASIFC segmentation. (d) BCFCM segmentation. (e) EnFCM segmentation. (f) FGFCM segmentation. (g) ASIC segmentation.

proposed ASIFC algorithms are applied to the selected images. The 3%, 5%, and 7% noise cases are tested. Fig. 3a, d and g show that the FGFCM and BCFCM algorithms perform quite well for all the noise levels on CSF tissue segmentation. EnFCM, ASIC and our proposed ASIFC algorithm seem to yield bigger error rates. From the sample MRI image in Fig. 2a, we can see that CSF tissue consists of many small regions with very low intensity values. Some even smaller than a 3×3 window. Therefore, for the ASIC and our ASIFC algorithm, the dissimilarity measure tends to trust more on the center pixel itself and imposes little spatial constraint on it. This is the main reason for the relatively lower TSA values. This problem can also be found in the TSA values for the GM tissue, although some of the misclustering pixels are contributed by the CSF tissue. For the segmentation of WM tissue, it is apparent that the ASIFC algorithm performs better. This is improved by the influence of stronger neighborhood effect imposed by the adaptive selection of the weighting functions. This is more obvious by looking at the TSA values obtained by ASIC algorithm. Although the ASIC algorithm outperforms the ASIFC algorithm for the WM tissue, the proposed ASIFC has better segmentation for the GM and CSF tissues when compared to the ASIC algorithm as the fuzzy c-means algorithm is more robust to the deterministic annealing algorithm [8].

From this comprehensive comparison we can conclude that, the FGFCM and BCFCM algorithms are useful in the segmentation of

Please cite this article in press as: Z. Wang et al., An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Understand. (2013), http://dx.doi.org/10.1016/j.cviu.2013.05.001

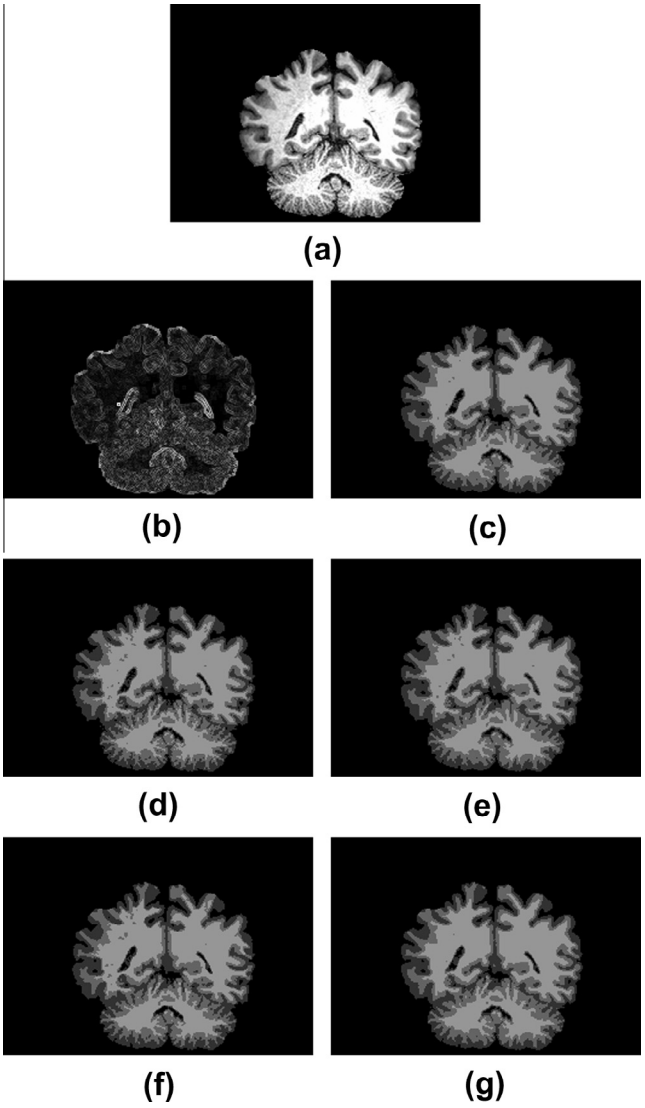


Fig. 5. Example 2: Segmentation on a coronal view slice. (a) Sample slides from real 3D MRI brain image volume. (b) Adaptive weighting parameters of our ASIFC algorithm. (c) ASIFC segmentation. (d) BCFCM segmentation. (e) EnFCM segmentation. (f) FGFCM segmentation. (g) ASIC segmentation.

CSF and GM tissues, while the ASIC and our proposed algorithm would provide better segmentation results for the WM tissue, and probably the GM tissue as well.

5.3. MRI brain image

In this experiment, we shall evaluate the performance of our proposed clustering algorithm on real MRI brain image data. The real MRI brain images are obtained using a Magnetization Prepared Rapid Gradient Echo (MPRAGE) imaging sequence. The size of the image volume is $256 \times 256 \times 180$ voxels. The voxel size is $0.9 \times 0.9 \times 0.9$ mm.

We chose 3 sample slices under different anatomical views (sagittal, coronal and transverse) as our test images. Visually, all the competing algorithms are able to produce pretty good segmentation results for these three MRI brain images. When we try to examine the results closely, we can see that there are some noticeable differences among them. For example, the segmentation results produced by the BCFCM is a bit 'noisier' when compared to other methods. You can see many isolated small regions in

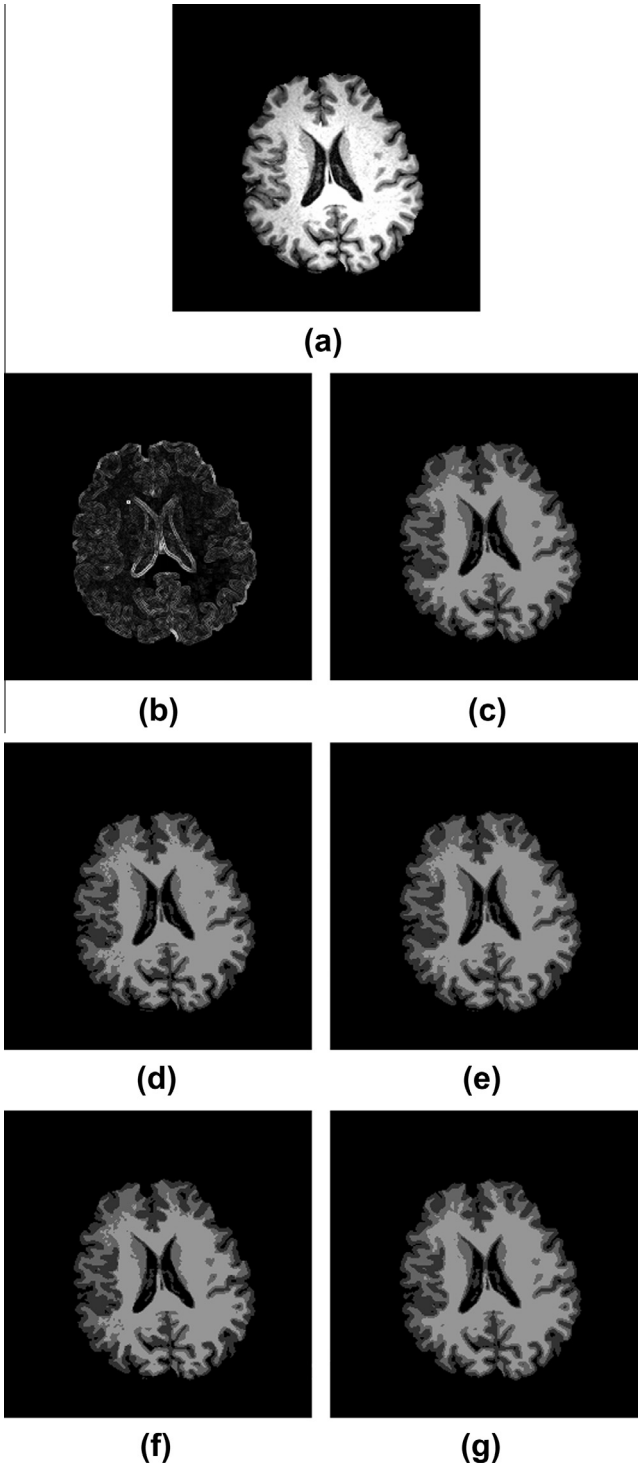


Fig. 6. Example 3: Segmentation on a transverse view slice. (a) Sample slides from real 3D MRI brain image volume. (b) Adaptive weighting parameters of our ASIFC algorithm. (c) ASIFC segmentation. (d) BCFCM segmentation. (e) EnFCM segmentation. (f) FGFCM segmentation. (g) ASIC segmentation.

Fig. 5d. The ASIC method produce a more consistent and smoother segmentation results. However, some details are lost in the cerebellum region as compared to the EnFCM, FGFCM and the proposed ASIFC algorithms, which can be seen in the cerebellum region in Figs. 4 and 5g. Among the EnFCM, FGFCM, and ASIFC algorithms, our algorithm yields an optimal trade-off between smoothing out noise and preserving the fine details for the whole image. These

Please cite this article in press as: Z. Wang et al., An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Understand. (2013), http://dx.doi.org/10.1016/j.cviu.2013.05.001

Table 3
Average computational time (in seconds) with MATLAB on MRI images for BCFCM, EnFCM, FGFCM, ASIC and ASIFC methods.

	BCFCM (s)	EnFCM (s)	FGFCM (s)	ASIC (s)	ASIFC (s)
Phantom image	2.90	0.67	1.42	18.62	5.01
Example 1	2.84	1.34	1.51	10.25	4.99
Example 2	2.95	1.48	1.36	10.17	5.37
Example 3	4.91	2.03	1.56	16.08	7.38

improvements can be visualized in Figs. 5 and 6c. With the adaptive similarity measure and information-theoretic framework, the newly proposed fuzzy clustering algorithm has strengths to overcome the sensitivity of noise or outlier data points and the lack of spatial information. The power of the ASIFC algorithm does not only lie in providing more homogeneous and smoother segmentation results, but also in preserving the image details and boundary information.

The average computational time for both synthetic data sets and real MRI images are listed in Table 3. As expected, due to the additional computational time for the weighting image, the ASIFC algorithm is relative slower than the BCFCM, EnFCM, and FGFCM algorithms. However, with the new adopted information fuzzy framework, the ASIFC has significant improvement in computational complexity as compared to the ASIC algorithm. This problem could be further improved by using more efficient programming tools, as the current implantations are all based on MATLAB.

6. Conclusions

In this paper, we have presented a novel robust information fuzzy clustering method that uses the information-theoretic clustering concept and adaptive spatial weighting factors to improve the image segmentation results. The proposed ASIFC algorithm addresses the two interwined problems, i.e. the identification of noisy data, and the lack of spatial contextual information, in data clustering based image segmentation. The algorithm is fully automated, except for the initialization of fuzzifier, cluster number and stop condition parameters. The experimental results based on synthetic images showed that the proposed algorithm can effectively improve the segmentation accuracy. The experiment on real MRI image also demonstrated that our ASIFC method can deliver useful segmentation results for brain image analysis.

References

[1] M.C. Clark, L.O. Hall, D.B. Goldgof, L.P. Clarke, R.P. Vethuizen, M.S. Silbiger, MRI segmentation using fuzzy clustering techniques, IEEE Eng. Med. Biol. Mag. 13 (1994) 730–742.
[2] A. Pitiot, A.W. Toga, P.M. Thompson, Adaptive elastic segmentation of brain MRI via shape-model-guided evolutionary programming, IEEE Trans. Med. Imag. 21 (2002) 910–923.
[3] X.W. Chen, T. Huang, Facial expression recognition: a clustering-based approach, Patt. Recog. Lett. 24 (2003) 1295–1302.
[4] N. Davidenko, Silhouetted face profiles: a new methodology for face perception research, J. Vis. 7 (2007) 1–17.
[5] Y. Tsaig, A. Averbuch, Automatic segmentation of moving objects in video sequences: a region labeling approach, IEEE Trans. Med. Imag. 12 (2002) 597–612.
[6] A. Millonig, K. Schechtner, Developing landmark-based pedestrian-navigation systems, IEEE Trans. Intell. Trans. Syst. 8 (2007) 43–49.
[7] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, NY, 1981.
[8] R.N. Dave, R. Krishnapuram, Robust clustering methods: a unified view, IEEE Trans. Fuzzy Syst. 5 (1997) 270–293.
[9] N.A. Mohamed, M.N. Ahmed, A.A. Farag, Modified fuzzy c-mean in medical image segmentation, in: Proc. of IEEE-EMBS, vol. 20, pp. 1377–1380.
[10] M.N. Ahmed, S.M. Yamany, A.A. Farag, T. Moriaty, Bias field estimation and adaptive segmentation of MRI data using a modified fuzzy c-means algorithm, in: Proc. of the 13th International Conf. on Computer Assisted Radiology and Surgery, CARS'99, Paris, p. 1004.
[11] Y.A. Tolias, S.M. Panas, Image segmentation by a fuzzy clustering algorithm using adaptive spatially constrained membership functions, IEEE Trans. Syst. Man Cyber.—Part A: Syst. Humans 28 (1998) 359–369.
[12] A. Liew, S. Leung, W. Lau, Fuzzy image clustering incorporating spatial continuity, IEE Proc. Vis. Image Sign. Process. 147 (2000) 185–192.
[13] D.L. Pham, Spatial models for fuzzy clustering, Comput. Vis. Image Understand. (2001) 285–297.
[14] L. Szilágyi, Z. Benyó, S.M. Sizlágyi, H.S. Adam, MR brain image segmentation using an enhanced fuzzy c-means algorithm, in: Proc. of the 25th Annual International Conference of the IEEE EMBS, Cancun, Mexico, pp. 724–726.
[15] W. Cai, S. Chen, D. Zhang, Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation, Patt. Recogn. 40 (2007) 835–838.
[16] A.W.C. Liew, Y. Hong, An adaptive spatial fuzzy clustering algorithm for 3D MR image segmentation, IEEE Trans. Med. Imag. 22 (2003) 1063–1075.
[17] Q. Song, A robust information clustering algorithm, Neural Comput. 17 (2005).
[18] J.C. Dunn, A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters, J. Cybern. 3 (1973) 32–57.
[19] Z.M. Wang, Y.C. Soh, Q. Song, K. Sim, Adaptive spatial information-theoretic clustering for image segmentation, Patt. Recogn. 42 (2009) 2029–2044.
[20] R. Krishnapuram, J.M. Keller, A possibilistic approach to clustering, IEEE Trans. Fuzzy Syst. 1 (1993) 98–110.
[21] R.E. Blahut, Computation of channel capacity and rate-distortion functions, IEEE Trans. Inform. Theor. 18 (1972) 460–473.
[22] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, G.B. Pike, A.C. Evans, Brainweb: online interface to a 3d mri simulated brain database, NeuroImage 5 (1997) 425.
[23] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes, A.C. Evans, Design and construction of a realistic digital brain phantom, IEEE Trans. Med. Imag. 17 (1998) 463–468.
[24] J.L. Marroquin, E.A. Santana, S. Botello, Hidden markov measure field models for image segmentation, IEEE Trans. Patt. Anal. Mach. Intell. 25 (2003) 1380–1387.

Please cite this article in press as: Z. Wang et al., An adaptive spatial information-theoretic fuzzy clustering algorithm for image segmentation, Comput. Vis. Image Understand. (2013), http://dx.doi.org/10.1016/j.cviu.2013.05.001