2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Fuzzy C-Means Clustering With A New Regularization Term for Image Segmentation

Guangpu Shao*, Junbin Gao†, Tianjiang Wang*, Fang Liu*, Yucheng Shu* and Yong Yang*
*School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, 430074, China
Email: {shaoguangpu, tjwang, fang.liu, shuyucheng}@hust.edu.cn, yangyong825825@163.com
†School of Computing and Mathematics, Charles Sturt University, Bathurst, NSW 2795, Australia
Email: jbgao@csu.edu.au

***Abstract*—We present a new fuzzy c-means algorithm for image segmentation by introducing a novel spatially constrained Student's t-distribution and a new regularization term. Firstly, considering that conventional distribution models lack spatial information and the multivariate Student's t-distribution is heavily tailed, we propose a new way to incorporate spatial information between neighboring pixels into the Student's t-distribution based on Markov random field (MRF) in order to enhance robustness. Secondly, the new regularization term, inspired by the geodesic active contour (GAC) with a strong ability in capturing boundary, can preserve the details of edges and further enhance its robustness to noise and outliers by capitalizing on the local context information and edge information. Finally, in comparison to other Markov random fields that are complex and computationally expensive, the parameters are easily optimized with the EM algorithm in our proposed method. The proposed algorithm demonstrates the robustness and effectiveness, compared with other state-of-the-art methods on synthetic and real images.**

I. Introduction

Image segmentation is fundamental to image/vision analysis and pattern recognition, regarded as an image pre-processing technique and process. Because of the diversity and complexity of image, it has been a challenging problem. The objective of image segmentation is to partition pixels in an image into disjoint sets that correspond to objects or parts of objects. There are many different approaches to image segmentation, such as graph based methods [1]-[4], partial differential equations (PDE) based segmentation techniques [5]-[7], and clustering approaches [8]-[9]. In this paper, we focus on segmentation based on fuzzy clustering.

The fuzzy c-means algorithms (FCMs) [10] have been well studied and successfully and widely applied to image segmentation. Based on fuzzy set theory [11] and introducing the idea of partial membership, FCM allows pixels to belong to multiple classes with certain similarity measure and makes a final classification decision. It is very important in applications in many real situations, such as limited spatial resolution, poor contrast, overlapping intensities, noise and intensity inhomogeneities, see [15]. Although the conventional FCM algorithm works well on most noise-free images, it is still sensitive to noise and outliers, since it does not consider any spatial information.

This work is partially supported by the National Natural Science Foundation of China under Grant 61073094 and U1233119. Junbin Gao's work is supported by Australian Research Council (ARC) through Discovery Project (DP) grant DP140102270.

To overcome this drawback of FCM, many extensions of the FCM algorithm have been explored and proposed [12]-[19]. Ahmed et al. [12] proposed FCM_S, which modifies the objective function of the standard fuzzy c-means (FCM) algorithm by incorporating regularized term on spatial neighborhood in order to compensate intensity inhomogeneity and to encourage the labeling of a pixel to be influenced by the its immediate neighborhood labels. However this approach results in a very time-consuming scheme as the neighborhood labelling was computed in each iteration step. Chen and Zhang [13] presented two variations of the original FCM_S, named, FCM_S1 and FCM_S2. These two new models incorporate the spatial contextual information by using the extra mean and median-filtered images, respectively, which can be calculated in advance. Hence the computation time is relatively reduced. Sotirios and Varvarigou [14] incorporated spatial constraints into the fuzzy objective function by Hidden Markov Random Field Models (HMRF). In this model, a fuzzy objective function is regularized by Kullback-Leibler divergence information and a mean-field-like approximation of the Markov field prior is adopted. Krinidis and Chatzis [15] presented a novel robust fuzzy local information c-means clustering algorithm (FLICM), which is free of any parameter selection. In FLICM, the novel fuzzy factor, incorporated with the local spatial information and pixel gray level information, was defined to improve noise insensitiveness and image detail preservation. This algorithm was fully free of the empirically adjusted parameters. In [16], a variant of FLICM was presented, named RFLICM. The important contribution is that the local coefficient of variation replacing the spatial distance was represented as a local similarity measure. The authors of [19] further improve the performance of FLICM by exploiting a new factor concerning on spatial distance of all neighbor pixels and their gray level difference simultaneously. Furthermore they used the kernel distance measure to replace the traditional Euclidean distance in the objective function. However, the experiments have demonstrated that these models still suffer from such as limited robustness to outliers and noise and the negligence of image edge information.

In this paper, we combine the benefits of FCM, Student's t-distribution, MRF and the regularization term. As real image objects are often nonlinearly various and heterogeneous, and Student's t-distribution is heavily tailed and more robust than Gaussian, we exploit multivariable Student's t-distribution to describe each phase rather than single constant in the conventional FCM. To enhance robustness to noise and outliers, local context information is fused into the probability distribution model with MRF. Integrated with the advantage of edge-based methods (GCA), the new regularization term is presented in the new algorithm in order to meet the criterion of the FCM algorithm. Finally, the proposed method adopts expectation maximization (EM) algorithm to optimize parameters.

II. Preliminary Theory

The Student's t-distribution has been widely used as a robust modeling method [20]-[23]. A D-dimensional random variable x is said to follow a multivariate t-distribution $S(x)$ with mean μ, a real covariance matrix Σ (positive definite) and $\nu \in [0, \infty)$ degrees of freedom, if its density function is given by

$$\begin{aligned} S(x) &= \frac{\Gamma(\nu/2 + D/2)|\Sigma|^{(-1/2)}}{(\nu\pi)^{D/2}\Gamma(\nu/2)} \\ &\times \left(1 + \nu^{-1}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)^{-(\nu+D)/2}, \end{aligned}$$

where $\Gamma(.)$ is the Gamma function:

$$\Gamma(n) = \int_0^\infty t^{n-1}e^{-t}dt.$$

As shown in Fig.1 (a) and (b), the Student's t-distribution is more heavily tailed than the Gaussian distribution. Compared with the Gaussian distribution, the Student's t-distribution has one more parameter ν which is viewed as a robustness tuning parameter. As $\nu \to \infty$, this distribution tends to the Gaussian distribution. Hence, the Student's t-distribution provides a more powerful and flexible approach for probabilistic data clustering compared with the Gaussian distribution.

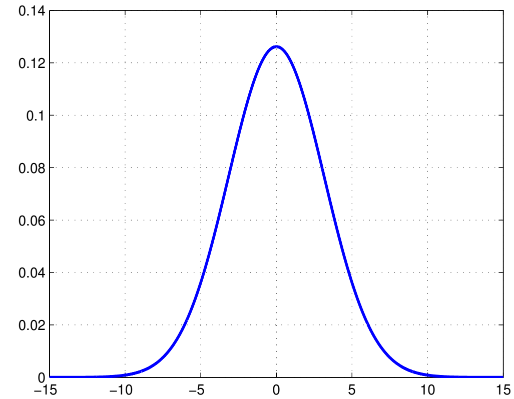
The FCM clustering algorithm was introduced by Dunn [24] and later extended by Bezdek [10]. The segmentation process of an image can be defined as the minimization of the following objective function:

$$J_m = \sum_{i=1}^{K}\sum_{j=1}^{N} u_{ij}^m d(x_i, \mu_j) \tag{1}$$

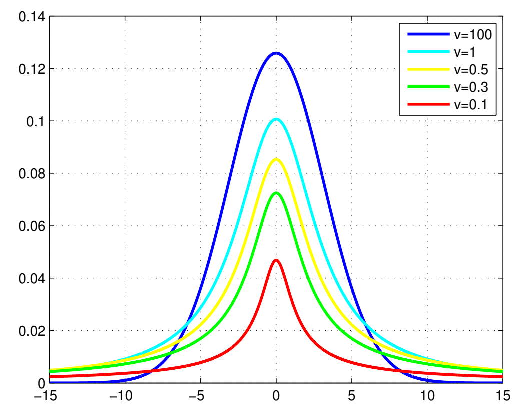
where N is the total number of pixels, K represents cluster number, x_i is the intensity value of the i-th pixel in the image, u_{ij} stands for the certain similarity degree of i-th pixel in the j-th clustering, m is a weighting exponent on each fuzzy membership u_{ij}, μ_j is the mean value of j-th clustering and $d(x_i, \mu_j)$ is the Euclidean measure between object x_i and cluster centre μ_j.

Caselles et al. [25] propose a geodesic active contour (GAC) model by defining an energy function along a curve C and minimizing the evolving curve in the normal direction. A new length definition in a different Riemannian space is given:

$$L_g(C) = \int_0^{L(C)} g(|\nabla I(C(s))|)ds \tag{2}$$



(a) Gaussian distribution



(b) Student's-t distribution for various values of v

Fig. 1. Plot of the 1-D distributions

where $L(C)$ is the length of the curve C, s is arc length parameters, $|\nabla I|$ is length of the gradient and $g : [0, +\infty] \to R^+$ is a strictly decreasing function such that $g(r) \to 0$ as $r \to \infty$.

Nguyen et al. [26] present the Spatially Constrained Gaussian Mixture Model based on MRF. The log-likelihood function can be expressed as:

$$L = \sum_{i=1}^{N} \log\left\{\sum_{j=1}^{K} \pi_{ij} p(x_i|\theta_j)\right\} + \sum_{i=1}^{N}\sum_{j=1}^{K} G_{ij}^{(t)} \log\pi_{ij} \tag{3}$$

where

$$G_{ij}^{(t)} = \exp\left\{\frac{\beta}{N_i}\sum_{q\in\partial_i}\left(z_{qj}^{(t)} + \pi_{qj}^{(t)}\right)\right\}$$

$z_{qj}^{(t)}$ is the posterior probability, $\pi_{qj}^{(t)}$ is the prior probability of the pixel x_q in the j-th label, β is the temperature value, ∂_i is the neighborhood of the i-th pixel, N_i is the number of neighboring pixels around the pixel x_i in the neighborhood ∂_i and t indicates the iteration steps.

III. PROPOSED METHOD

We exploit the advantages of (1)-(3) to present a new objective function:

$$J = \sum_{i=1}^{N}\sum_{j=1}^{K} u_{ij}^{m} d_{ij} + \sum_{i=1}^{N}\sum_{j=1}^{K} \hat{G}_{ij}^{(t)} \log \pi_{ij} + \lambda \sum_{i=1}^{N}\sum_{j=1}^{K}\sum_{q\in\partial_i} g(|\nabla I_\sigma|_q)(1-u_{qj})^m u_{ij}^m \tag{4}$$

where the distance function d_{ij} in (4) is different from that in (1), which is defined by the student's t-distribution $S(x)$ and the prior information π_{ij} as follows:

$$d_{ij} = -2\log(\pi_{ij} S(x_i|\phi_j))$$

where $\phi_j = \{\nu_j, \mu_j, \Sigma_j\}, (j = 1, 2, ..., K)$ and π_{ij} satisfies the constraints:

$$\pi_{ij} \geq 0, j = 1, 2, ..., K \quad \text{and} \quad \sum_{j=1}^{K} \pi_{ij} = 1 \tag{5}$$

$$\begin{aligned} S(x_i|\phi_j) &= \frac{\Gamma(\nu_j/2 + D/2)|\Sigma_j|^{(-1/2)}}{(\nu_j \pi)^{D/2}\Gamma(\nu_j/2)} \\ &\times \left(1 + \nu_j^{-1}(x_i - \mu_j)^T \Sigma_j^{-1}(x_i - \mu_j)\right)^{-(\nu_j + D)/2}. \end{aligned}$$

In the 2nd term of J, a novel factor $\hat{G}_{ij}^{(t)}$, inspired by anisotropic diffusion, is introduced in (6), in order to incorporate the local context information and preserve the details of image edge.

$$\hat{G}_{ij}^{(t)} = -\left\{ \exp\left(\frac{\beta}{N_i} \sum_{p\in\partial_i} \omega_{pj} (\pi_{pj}^{(t)} + u_{pj}^{(t)} - \pi_{ij}^{(t)} - u_{ij}^{(t)}) \right) \right\}^m \tag{6}$$

where

$$\omega_{pj} = \frac{1}{1+\rho_{pi}} \times \frac{1}{1 + |\nabla(\pi_{pj}^{(t)} + u_{pj}^{(t)})|}$$

Here, ρ_{pi} is the Squared Euclidean distance between the point(p) and the point(i). The factor defined in (6) is different from that in (3).

For the 3rd term of J, we introduce the novel regularization term. According to the paper [27], the length of the cluster boundaries is defined as

$$E_L = \sum_{i=1}^{N}\sum_{j=1}^{K}\sum_{p\in\partial_i} \chi_j(x_p, x_i)$$

where the characteristic function χ_j is defined as

$$\chi_j(x_p, x_i) = \begin{cases} 1 & x_p \in \Omega_j \text{ and } x_i \in \Omega_j. \\ 0 & \text{otherwise.} \end{cases}$$

where Ω_j stands for the region corresponding to the j-th clustering, $j = 1, 2, ..., K$. Let us introduce a vector-valued characteristic function $\hat{u} = (\hat{u}_1, \hat{u}_2, ..., \hat{u}_K)$. Each $\hat{u}_j$: $\Omega \to \{0, 1\}$, where Ω is defined on a bound domain of image, is defined as

$$\hat{u}_{ij} = \begin{cases} 1 & x_i \in \Omega_j. \\ 0 & \text{otherwise.} \end{cases}$$

So E_L can be rewritten as

$$\begin{aligned} E_L &= \sum_{i=1}^{N}\sum_{j=1}^{K} \left\{ \sum_{p\in\partial_i} \sum_{k=1,k\neq j} \hat{u}_{pk} \right\} \hat{u}_{ij} \\ &= \sum_{i=1}^{N}\sum_{j=1}^{K}\sum_{p\in\partial_i} (1-\hat{u}_{pj})\hat{u}_{ij} \end{aligned}$$

It was shown in [27] that: $E_L \propto L(C)$, where $\propto$ represents a proportional relation. It is well-known that GCA is the essentially weighted arc length, expressed as:

$$\begin{aligned} L_g(C) &= \int_0^{L(C)} g(|\nabla I(C(s))|)ds \\ &\propto \sum_{i=1}^{N}\sum_{j=1}^{K}\sum_{p\in\partial_i} g(|\nabla I|_p)(1-\hat{u}_{pj})\hat{u}_{ij} \end{aligned}$$

$L_g(C)$ is modified in order to satisfy the criterion of the FCM algorithm. We present the novel regularizer named improved GCA, which is

$$\hat{L}_g = \sum_{i=1}^{N}\sum_{j=1}^{K}\sum_{p\in\partial_i} g(|\nabla I_\sigma|_p)(1-\hat{u}_{pj})^m \hat{u}_{ij}^m$$

I_σ stands for the image preprocessed by Gaussian filter (0 mean, σ variance). $g(r) = (1 + r/M)^{-1}$ where M is a constant. Finally, we shall consider the constrained optimization problem (4) under the condition:

$$\sum_{i=1}^{N} u_{ij} = 1, u_{ij} \geq 0, \quad \text{for every } j \in \{1, 2, .., K\} \tag{7}$$

IV. PARAMETER LEARNING

We use the Lagrange multiplier to minimize the objective function (4) over parameters π_{ij} under the condition (5), so we have

$$\pi_{ij}^{(t+1)} = \frac{(u_{ij}^{(t)})^m - G_{ij}^{(t)}}{\sum_{k=1}^{K} \left((u_{ik}^{(t)})^m - G_{ik}^{(t)} \right)} \tag{8}$$

As to u_{ij}, in the same way, the u_{ij} is given as

$$u_{ij}^{(t+1)} = \frac{1}{\sum_{k=1}^{K} \left\{ \frac{d_{ij} + \lambda \sum_{p\in\partial_i} g(|\nabla I_\sigma|_p)(1-u_{pj}^{(t)})^m}{d_{ik} + \lambda \sum_{p\in\partial_i} g(|\nabla I_\sigma|_p)(1-u_{pk}^{(t)})^m} \right\}^{1/(m-1)}} \tag{9}$$

The parameter set $\phi_j = \{\nu_j, \mu_j, \Sigma_j\}, (j = 1, 2, ..., K)$ are optimized by minimizing the objective function (4). Note that there is no closed form solution under a Student's-t distribution. In order to make use of the EM algorithm and overcome this problem, we use the fact, refer to [28]

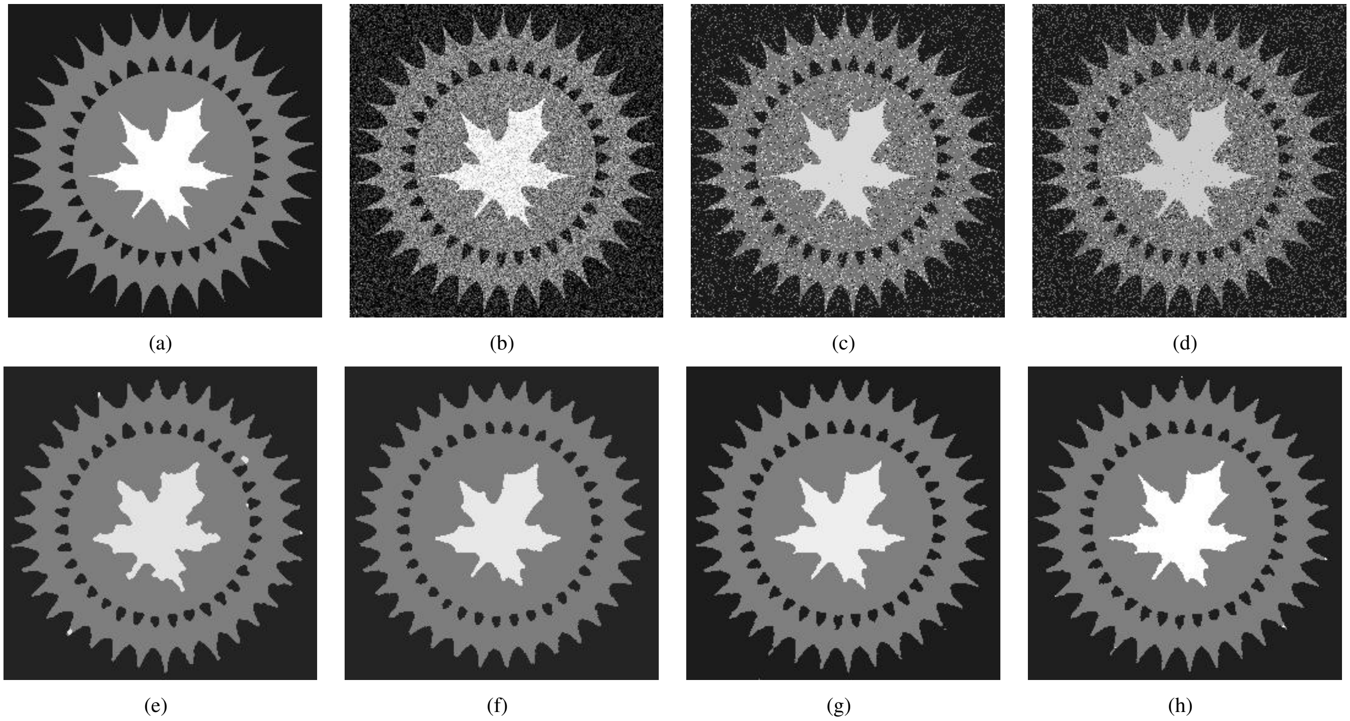


Fig. 2. (a) Original image (b) Corrupted original image with Gaussian noise (0mean,0.03variance) (c) SVFMM (MCR=13.77%) (d) DCA-SVFMM (MCR=19.81%) (f) HMRF-FCM (MCR=2.32%) (e) ST-FCM (MCR=2%) (g) KWFLICM (MCR=1.37%) (h) Proposed method (MCR=1.35%)

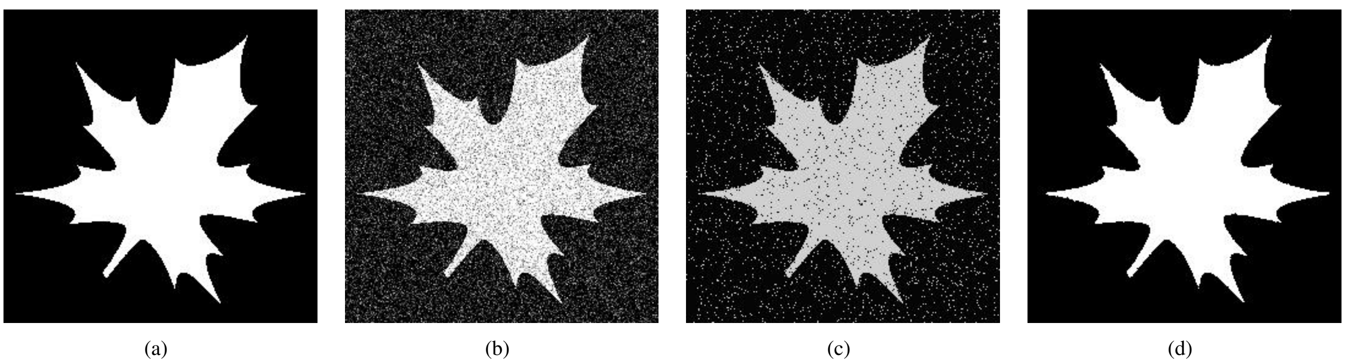


Fig. 3. (a) Original image (b) Corrupted original image with Gaussian noise (0mean,0.08variance) (c) FCM based on spatially constrained Student's t-distribution (d) FCM based on spatially constrained Student's t-distribution and the new regularization term

and [29], that the Student's t-distribution can be decomposed into a superposition of infinite number of Gaussians with the precision scaling factor S_{ij} given by:

$$S(x_i|\phi_j) = \int \Phi(x_i|\mu_j, \Sigma_j/S_{ij}) d\wp(S_{ij}|\nu_j/2, \nu_j/2).$$

where $\Phi(x_i|\mu_j, \Sigma_j/S_{ij}) =$

$$\frac{1}{(2\pi)^{D/2}} \frac{S_{ij}^{D/2}}{|\Sigma_j|^{1/2}} \exp\left\{-\frac{1}{2} S_{ij}(x_i-\mu_j)^T \Sigma_j^{-1}(x_i-\mu_j)\right\}$$

and the Gamma distribution $\wp(S_{ij}|\nu_j/2, \nu_j/2)$ is given by:

$$\wp(S_{ij}|\nu_j/2, \nu_j/2) = \frac{1}{\Gamma(\nu_j/2)}(\nu_j/2)^{\nu_j/2}(S_{ij})^{\nu_j/2-1} \mathrm{e}^{-\nu_j S_{ij}/2}$$

Considering S_{ij} as an introduced latent variable, following [30], we can updated the scaling factor as follows

$$S_{ij}^{(t)} = \frac{\nu_j^{(t)} + D}{\nu_j^{(t)} + (x_i - \mu_j^{(t)})^T (\Sigma_j^{(t)})^{-1} (x_i - \mu_j^{(t)})} \tag{10}$$

As for the estimate of μ_j at the $(t+1)$ step, we have

$$\mu_j^{(t+1)} = \frac{\sum_{i=1}^{N} S_{ij}^{(t)} (u_{ij}^{(t)})^m x_i}{\sum_{i=1}^{N} S_{ij}^{(t)} (u_{ij}^{(t)})^m} \tag{11}$$

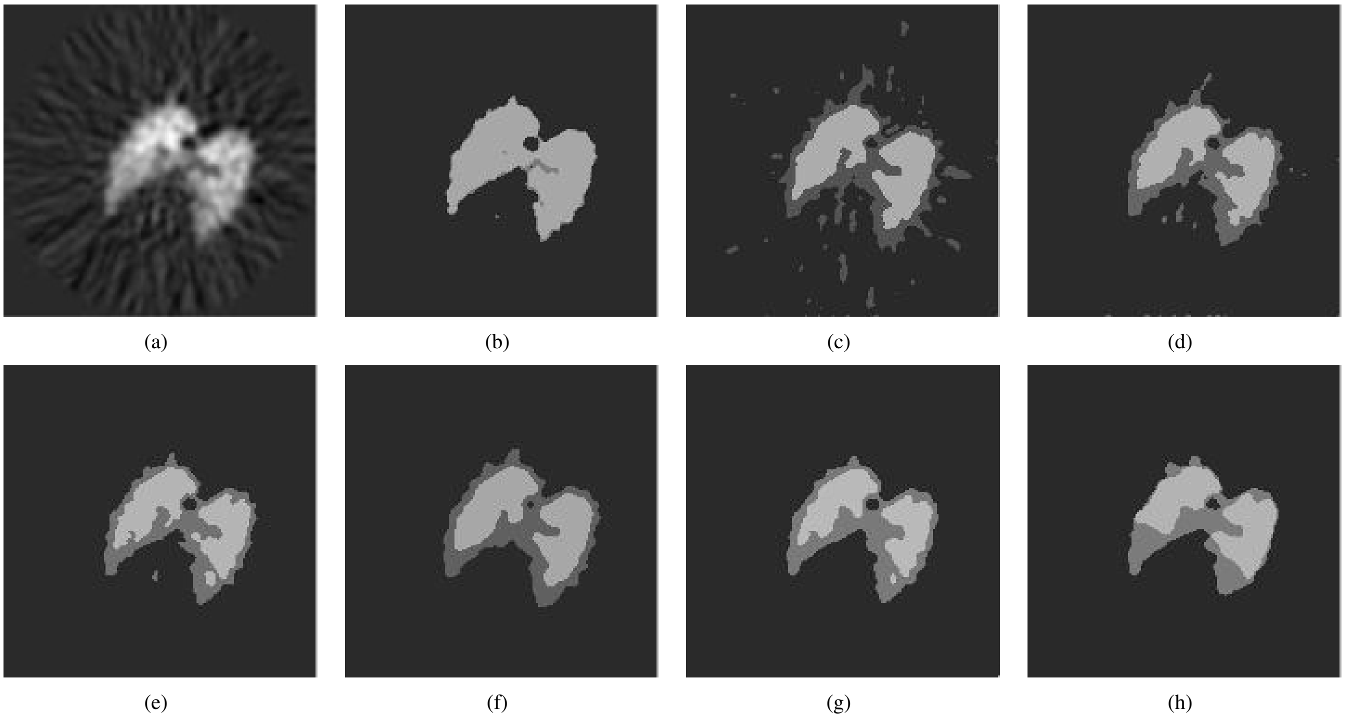


Fig. 4. (a) Original image (b)KL-FCM (c) SVFMM (d) DCA-SVFMM (f) HMRF-FCM (e) ST-FCM (g) KWFLICM (h) Proposed method

Regarding Σ_j at the $(t+1)$ iteration step, we have

$$\Sigma_j^{(t+1)} = \frac{\sum_{i=1}^{N} S_{ij}^{(t)} (u_{ij}^{(t)})^m (x_i - \mu_j^{(t+1)})(x_i - \mu_j^{(t+1)})^T}{\sum_{i=1}^{N} (u_{ij}^{(t)})^m} \tag{12}$$

The estimates of the degrees of freedom (ν_j) are given by the solution of the equation:

$$\log\left(\frac{\nu_j}{2}\right) - \Psi\left(\frac{\nu_j}{2}\right) + 1 + \frac{\sum_{i=1}^{N} \left(u_{ij}^{(t)}\right)^m \left(\log(S_{ij}^{(t)}) - S_{ij}^{(t)}\right)}{\sum_{j=1}^{N} \left(u_{ij}^{(t)}\right)^m} + \Psi\left(\frac{\nu_j^{(t)} + D}{2}\right) - \log\left(\frac{\nu_j^{(t)} + D}{2}\right) = 0 \tag{13}$$

To summarize it, we have the following algorithm procedure for the proposed method as follows:

Step 1 Initialize the parameters: μ_j, Σ_j, ν_j, π_{ij}, β and λ.

Step 2 (E step):
- * Evaluate fuzzy membership function u_{ij} by (9).
- * Compute the factor G_{ij} by (6)
- * Compute the scaled precision S_{ij} by (10).

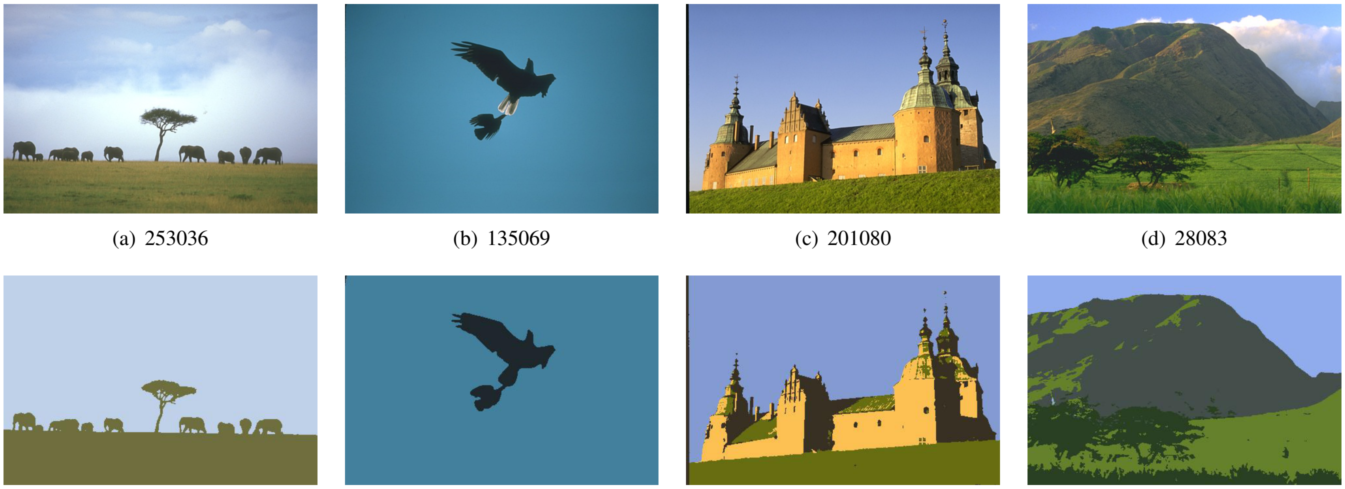
Step 3 (M step):
- * Update the means μ_j by (11)
- * Update the covariance matrices Σ_j by (12).
- * Update the degree of freedom ν_j by (13).
- * Update prior distributions π_{ij} by (8).

Step 4 Evaluate the objective function in (4) and check for the convergence of either the objective function or the parameter values. If the convergence criterion is not satisfied, then go to step 2

In the next section, we will demonstrate the robustness, accuracy and effectiveness of the proposed model, as well as comparing with other approaches.

V. EXPERIMENTS

In this section, some comprehensive experiments are conducted to evaluate the performance of our proposed model, compared with some state-of-the-art image segmentation models. These models include the spatially constrained mixture model (SVFMM) [31], the class-adaptive spatially variant mixture model (DCA-SVFMM) [32], the fuzzy c-means clustering with regularization by K-L information (KL-FCM) [33], the fuzzy clustering approach toward hidden Markov random field models for enhanced spatially constrained image segmentation (HMRF-FCM) [14], the fuzzy C-Means Clustering With Local Information and Kernel Metric for Image Segmentation (KWFLICM) [19] and the Robust Fuzzy Algorithm Based on Student's t-Distribution and Mean Template for Image Segmentation Application (ST-FCM) [34]. The source code of SVFMM and DCA-SVFMM can be downloaded from http://www.cs.uoi.gr/kblekas/sw/MAPsegmentation.html and http://www.cs.uoi.gr/cnikou/, respectively. HMRF-FCM code and KWFLICM code are available at the website: https://www.cut.ac.cy/eecei/staff/sotirios.chatzis/languageId=2 and http: //see.xidian.edu.cn/faculty/mggong/publication.htm, respectively.



(a) 253036

(b) 135069

(c) 201080

(d) 28083

(e)

(f)

(g)

(h)

Fig. 5. Segmentation examples of our method based on the Berkeley database. First row: original image. Second row: the segmentation results of image in first row by our proposed method.

The proposed method was implemented and tested on synthetic and real-world images. To quantify segmentation results, the misclassification ratio (MCR) [36], which is the number of misclassified pixels divided by the total number of pixels, is adopted to evaluate the segmentation accuracy of synthetic images. The probabilistic rand (PR) index [37] is used to evaluate the image segmentation performance of the natural images. All methods are initialized by the K++ means algorithm [35]. Parameters β, m, M and λ are assigned to 20, 2, 100 and 60 respectively. The proposed method is implemented in Matlab and tested on a PC with Intel Pentium D running at 2.8 GHz and 2 GB of RAM.

TABLE I
COMPARISON OF IMAGE SEGMENTATION RESULTS BY PROBABILISTIC RAND (PR) INDEX

	SVFMM	DCA-SVFMM	HMRF -FCM	ST -FCM	Proposed method
28083	0.839	0.841	0.849	0.865	0.866
201080	0.884	0.874	0.873	0.925	0.926
135069	0.976	0.989	0.987	0.977	0.991
253036	0.923	0.923	0.922	0.923	0.924
mean	0.9055	0.9067	0.9077	0.9225	0.9267

A. Synthetic Images

In the first experiment, a synthetic image (256x256 image resolution) as shown in Fig. 2(a), is used to compare the proposed algorithm with others. The image has three classes with luminance values [0.1,0.5,1]. The image shown in Fig. 2(b) is obtained by corrupting the original image with Gaussian noise with mean 0 and variance 0.03. Segmentation results shown in Fig. 2(c)-(g) are obtained by using SVFMM, DCA-SVFMM, HMRF-FCM, ST-FCM, KWFLICM, respectively. Misclassification ratio (MCR) [36] has been used to compare the results. As can be seen, a lot of noise exist in the segmentation results of SVFMM and DCA-SVFMM. From a quantitative perspective, the MCR of the SVFMM and DCA-SVFMM is much higher than that of others, respectively. Under a given level of noise, the two algorithms (SVFMM and DCA-SVFMM) are sensitive to noises. As shown in Fig. 2(e), the HMRF-FCM algorithm reduces the effect of noise significantly. However, there still exist a little catchy noise. Obviously segmentation results of ST-FCM algorithm shown in Fig. 2(f) have further reduced the effect of noise. As MCR of ST-FCM is lower 0.32% than that of HMRF-FCM, ST-FCM algorithm obtains a better result than the previous algorithms. But after careful observation, some details are lost in this segmented images. We can see that the proposed method obtains the best performance and the edge information of the image is kept very well in Fig. 2(h). From the objective perspective, our algorithm has lower MCR compared with the other methods.

The second experiment is used to test and verify the effect of our innovations. As is shown in Fig. 3(a), the synthetic image with 256x256 pixels includes two classes with two luminance values $\{0,1\}$. This image can be downed from the website: https://sites.google.com/site/nguyen1j/home/10-code. Fig. 3(b) shows this synthetic image corrupted by Gaussian noise with mean 0 and variance 0.08. Next, we test our proposed algorithm performance in this noised image. We conducted an experiment in order to test the effect of spatially constrained Student's t-distribution in our proposed algorithm. Fig. 3(c) shows the segmentation result of FCM based on spatially constrained Student's t-distribution. We can see that the effect of noise is reduced clearly and the details of the image is well preserved. To test and verify effect of the novel regularizer terms in our method, another experiment is conducted. Fig. 3(d) shows the image segmentation result by the FCM based on spatially constrained Student's t-distribution and the new regularization terms. It can be seen that the influence of the noise is significantly reduced with a better image segmentation result than the previous results.

B. Natural Images

It is very difficult and challenging to segment natural images. In this series of experiments, we tested the performance of our proposed algorithm on the real medical image and the images in the Berkeley image segmentation database [40]. This image set includes natural images along with their ground truth segmentation results provided by human subjects. So performance of the various algorithms can be quantitative estimated by PR index [37]. The value of PR is bounded by the interval [0, 1], where the larger PR values indicate better segmentation results.

In order to further test the accuracy of the proposed algorithm, the third experiment was conducted. A positron emission tomography (PET) image (128x128 image resolution) of a dog lung [38] [39], as shown in Fig. 4(a), is used to compare the proposed algorithm with other algorithms. These methods include KL-FCM, SVFMM, DCA-SVFMM, HMRF-FCM, ST-FCM, and KWFLICM. The objective is to segment the image into three classes ($K = 3$). Fig. 4(b) to Fig. 4(h) show the results obtained by implementing KL-FCM, SVFMM, DCA-SVFMM, HMRF-FCM, ST-FCM, KWFLICM and the proposed method. From a visual perspective, segmented result of KL-FCM in Fig. 4(b) keep the image's edges well. However, there are a few noise in this image. From Fig. 4(c)-(d), we can see that the results of SVFMM, DCA-SVFMM, HMRF-FCM, ST-FCM, and KWFLICM lost a lot of edge information. As we observe, the proposed algorithm manages to derive more homogeneous regions and reserves more the edge details, compared with other algorithms.

In this experiment, we evaluate the performance of the proposed algorithm based on a subset of the Berkeley image dataset, which consists of 300 natural color and grayscale images. Each image in the image database has 481 $\times$ 321 pixels. Four images (253036, 135069, 201080, 28083) are used to test the effectiveness of the algorithm. In order to facilitate the segmentation performance estimation visually and quantitatively, the proposed method is compared with many latest algorithms. These methods includes SVFMM, DCA-SVFMM, HMRF-FCM and ST-FCM. Fig. 5(a)-(d) show the original image, respectively. Segmentation results of Fig. 5(a)-(d) by our method are shown in Fig. 5(e)-(h). As we observe, our method effectively retain the local homogeneity and preserve more details at the same time. To quantitatively evaluate the performance of the proposed method against SVFMM, DCA-SVFMM, HMRF-FCM and ST-FCM, PR index [37] values are provided in Table 1 for the comparing methods on the given set of real world images from the Berkeley image segmentation dataset. As can be easily seen, on average, the proposed method outperforms other methods with a higher PR index.

VI. Conclusion

In this paper, we propose a novel fuzzy-clustering algorithm that integrates Student's t-distribution, MRF and the novel regularizer term. The Student's t-mixture model is heavily tailed and more robust for outliers, but it lacks local context information. In order to unite Student's t-distribution with local information, we present the new spatially constrained Student's t-distribution by the MRF. With the aim of further enhancing robustness to the noise and capturing boundary, the novel regularizer term is presented by capitalizing on the local context information and the edge information. At last, they are integrated into the framework of FCM. In comparison to other MRFs that are complex and computationally expensive, our proposed method is easier to optimize the parameters with the EM algorithm.

References

[1] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, 2000.

[2] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1124-1137, 2004.

[3] V. Kolmogorov and R. Zabin, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 26, no. 2, pp. 147-159, 2004.

[4] Y. Yang, S. Han, T. Wang, W. Tao, and X. Tai, "Multilayer graph cuts based unsupervised color-texture image segmentation using multivariate mixed student's t-distribution and regional credibility merging," *Pattern Recognition*, vol. 46, no. 4, pp. 1101-1124, 2013.

[5] N. Paragios and R. Deriche, "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 3, pp. 266-280, 2000.

[6] L. Vese and T. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271-293, 2002.

[7] H. Zhou, J. Zheng, and L. Wei, "Texture aware image segmentation using graph cuts and active contours," *Pattern Recognition*, vol. 46, no. 6, pp. 1719-1733, 2013.

[8] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image segmentation using expectation-maximization and its application to image querying," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 24, no. 8, pp. 1026-1038, 2002.

[9] W. Cai, S. Chen, and D. Zhang, "Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation," *Pattern Recognition*, vol. 40, no. 3, pp. 825-838, Mar. 2007.

[10] J. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms," New York: Plenum, 1981.

[11] L. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, pp. 338-353, 1965.

[12] M. Ahmed, S. Yaman, N. Mohamed, A. Farag, and T. Moriarty, "A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data," *IEEE Trans. Med. Imag*, vol. 21, pp. 193-199, 2002.

[13] S. Chen and D. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," *IEEE Trans. Syst., Man, Cybern*, vol. 34, pp. 1907-1916, 2004.

[14] S. P. Chatzis and T. A. Varvarigou, "A fuzzy clustering approach toward hidden Markov random field models for enhanced spatially constrained image segmentation," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 5, pp. 1351-1361, Oct. 2008.

[15] S. Krinidis and V. Chatzis, "A robust fuzzy local information C-means clustering algorithm," *IEEE Trans. Image Process*, vol. 19, no. 5, pp. 1328-1337, May 2010.

[16] Maoguo Gong, Zhiqiang Zhou, and Jingjing Ma, "Change Detection in Synthetic Aperture Radar Images based on Image Fusion and Fuzzy Clustering," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2141-2151, 2012.

[17] Prabhjot Kaur, A.K. Soni, and Anjana Gosain, "A robust kernelized intuitionistic fuzzy c-means clustering algorithm in segmentation of noisy medical images," *Pattern Recognition Letters*, vol. 34, no 2, pp. 163-175, 2013.

[18] Maoguo Gong, Linzhi Su, Meng Jia, and Weisheng Chen, “Fuzzy Clustering with Modified MRF Energy Function for Change Detection in Synthetic Aperture Radar Images,” *IEEE Transactions on Fuzzy Systems*, vol. 22, no. 1, pp. 98-109, 2014.

[19] Maoguo Gong, Yan Liang, Jiao Shi, Wenping Ma, and Jingjing Ma, “Fuzzy C-Means Clustering With Local Information and Kernel Metric for Image Segmentation,” *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 573-583, 2013.

[20] D. Peel and G. McLachlan, “Robust mixture modeling using the t distribution,” *Stat. Comput.*, vol. 10, no. 4, pp. 339-348, Oct. 2000.

[21] S. Shoham, “Robust clustering by deterministic agglomeration EM of mixtures of multivariate t-distributions,” *Pattern Recognition*, vol. 35, no. 5, pp. 1127C1142, May 2002.

[22] G. Sfikas, C. Nikou, and N. Galatsanos, “Robust image segmentation with mixtures of Student’s t-distributions,” in *Proc. IEEE Int. Conf.Image Process.*, vol. 1, pp. 273-276, 2007.

[23] M. N. Thanh and Q. M. J. Wu, “Robust student’s-t mixture model with spatial constraints and its application in medical image segmentation,” *IEEE Trans. Med. Imag.*, vol. 31, no. 1, pp. 103-116, Jan. 2012.

[24] J. Dunn, “A fuzzy relative of the ISODATA process and its use in detecting compact well separated clusters,” *J. Cybern.*, vol. 3, pp. 32-57, 1974.

[25] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic active contours,” *International Journal of Computer Vision.*, vol. 22, no. 1, pp. 61-79, 1997.

[26] T. M. Nguyen and Q. M. Jonathan Wu, “Fast and Robust Spatially Constrained Gaussian Mixture Model for Image Segmentation,” *IEEE Transactions on Circuits and Systems for Video Technology.*,vol. 23, no. 4, pp. 621 - 635, 2013.

[27] J. Wang, L. L. Ju, and X. Q. Wang, “An edge-weighted centroidal Voronoi tessellation model for image segmentation,” *IEEE Trans. Image Process.*, vol. 18, no. 8, 2009.

[28] P. Dempster, N. M. Laird, D.B.Rubin, “Maximum likelihood from incomplete data via EM algorithm,” *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1-38, 1977.

[29] C. Liu and D. Rubin, “ML estimation of the t distribution using EM and its extensions, ECM and ECME,” *Statistica Sinica*, vol. 5, no. 1, pp. 19-39, 1995.

[30] M. N. Thanh and Q. M. J. Wu, “Robust student’s t-mixture model with spatial constraints and its application in medical image segmentation,” *IEEE Trans. Med. Imag.*, vol. 30, no. 1, pp. 1-14, Jan. 2011.

[31] K. Blekas, A. Likas, N. P. Galatsanos, and I. E. Lagaris, “A spatially constrained mixture model for image segmentation,” *IEEE Trans. Neural Netw.*, vol. 16, no. 2, pp. 494-498, 2005.

[32] C. Nikou, N. Galatsanos, and A. Likas, “A class-adaptive spatially variant mixture model for image segmentation,” *IEEE Transactions on Image Processing.*, vol. 16, no. 4, pp. 1121-1130, 2007.

[33] H. Ichihashi, K. Miyagishi, and K. Honda, “ Fuzzy c-means clustering with regularization by K-L information,” in *Proc. 10th IEEE Int. Conf. Fuzzy Syst.*, pp. 924-927, 2001.

[34] Hui Zhang, Q. M. JonathanWu, and Thanh Minh Nguyen, “A Robust Fuzzy Algorithm Based on Student’s t-Distribution and Mean Template for Image Segmentation Application,” *IEEE Signal Process. Lett.*, vol. 20, no. 2, pp. 117-119, 2013.

[35] D. Arthur and S. Vassilvitskii, “K-means++: the advantages of careful seeding,” *Proceedings of the ACM Symposium on Discrete Algorithms.*, pp.1027-1035, 2007.

[36] Y. Zhang, M. Brady, and S. Smith, “Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm,” *IEEE Trans. Med. Imag.*, vol. 20, no. 1, pp. 45-57, Jan. 2001.

[37] R. Unnikrishnan, C. Pantofaru, and M. Hebert, “A measure for objective evaluation of image segmentation algorithms,” in *IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 3, pp. 34-41, 2005.

[38] F. Forbes and N. Peyrard, “Hidden Markov random field model selection criteria based on mean field-like approximations,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.25, no. 9, pp. 1089-1101, 2003.

[39] N. T. Minh and Q. M. Wu, “Dirichlet Gaussian mixture model: Application to image segmentation,” *Image and Vision Computing*, vol.29, no. 12, pp. 818-828, 2011.

[40] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proc. Int. Conf. Comput. Vision*, pp. 416-423, Jul. 2001.