# Fuzzy Clustering using Local and Global Region Information for Cell Image Segmentation

Amin Gharipour School of Information and Communication Technology Gold Coast campus, Griffith University, QLD4222, Australia amin.gharipour@griffithuni.edu.au

Abstract—In high-throughput applications, accurate segmentation of biomedical images can be considered as an important step for recognizing cells that have the phenotype of interest. In this paper, while conventional fuzzy clustering is not able to implement the local and global spatial information, a novel spatial fuzzy clustering cell image segmentation algorithm is proposed. The segmentation procedure is divided into two stages: the first stage involves processing the local and global spatial information of the given cell image and a final segmentation stage which is based on the idea of conventional fuzzy clustering. Our idea can be considered as a sequential integration of region based methods and fuzzy clustering for cell image segmentation. Experimental results show that the proposed model yields significantly better performance in comparison with several existing methods.

#### Keywords—spatial fuzzy clustering; Chan-Vese model; local Chan-Vese model; Split Bergman method

## I. INTRODUCTION

Fluorescence staining in combination with fluorescence microscopy provides useful information about cells in image cytometry. However, visual assessment of fluorescence microscopy images is time consuming and inter- and intraobserver variability is often observed. Fast automated high throughput detection could be achieved by computer analysis of images produced by a digital camera attached to a fluorescence microscope. The first step of all automatic image analysis in high throughput cell screening applications is cell image segmentation. In general, in high-throughput tasks, manual and semi-automatic procedures are not feasible due to the massive amount of image data. Therefore, automatic approaches are required which efficiently handle different cell types and image artifacts such as intensity inhomogeneity and noise.

The introduction of fuzzy segmentation methods, especially the fuzzy c-means algorithm (FCM) has advanced the field of image segmentation [1-7, 31, 32]. Their popularity and effectiveness is essentially related to the concept of fuzziness in the membership degree of each image pixel. Because of this ability, FCM and its variants have the ability to use more information from the image compared to the crisp or hard segmentation methods [20] which restrict each pixel to exactly one cluster. The purpose of clustering is to partition image pixels into clusters of similar characteristics. Generally, the number of clusters is determined based on some cluster

Alan Wee-Chung Liew School of Information and Communication Technology Gold Coast campus, Griffith University, QLD4222, Australia a.liew@griffith.edu.au

validity methods or prior knowledge. However, the main drawback of FCM is that it only considers the spectral properties of the pixels, and ignores the local and global spatial information, making FCM very sensitive to noise and intensity inhomogeneity. Recently, the idea of incorporating spatial information into clustering methods has been widely investigated, and several algorithms have been reported [3-7, 31, 32].

On the other hand, region-based level set methods [8-11] have been proposed and applied to image segmentation. Based on the Mumford-Shah functional [8] for image segmentation, the Chan–Vese (CV) model is proposed [11] to identify objects whose boundaries are not accurately identified by the gradient. The minimal partition problem of the Mumford–Shah model is solved in the CV model by implementing level set functions, which use the regional image information inside and outside the evolving curve, rather than the boundaries' gradients information. Although, by using the regional information, the CV model is generally robust to noise, it often converges to poor segmentation results for images with intensity inhomogeneity. Furthermore, since the CV model and its derivatives are based on the minimal partition problem and are not convex, the global convex segmentation technique (GCS) is proposed [12] to show that the non-convex energy of the CV model can be reformulated as a convex optimization problem. In [13] the GCS technique and the split Bergman algorithm are utilized to propose a convex and fast segmentation method based on the CV model but it is unable to use local information and handle images with intensity inhomogeneity.

In this study, since reliable cell image segmentation requires considering the intensity of each pixel, as well as local and global spatial information in the image, the spatial fuzzy clustering segmentation is decomposed into two stages. The idea behind this decomposition is based on the idea of sequential integration [14, 15] in which the sequential separability assumption is implemented to ensure reliable segmentation. In this approach, the final solution is defined as the set of solutions acquired from solving the objectives in a cascade way.

The rest of this paper is organized as follows: the CV approach, the Local Chan-Vese (LCV) and the FCM are described in the background section. Section three introduces the proposed Globally Convex Local Chan-Vese model (GCLCV) as well as our two stages spatial fuzzy clustering

algorithm. Comparison experiments and validation are provided in Section 4. Finally, Section 5 concludes this paper and highlights possible future work.

## II. BACKGROUND

#### A. Chan-Vese model

Let  $\Omega \subset \Re^2$  denotes the image domain and  $I_0: \Omega \to R$ denotes a given grey level image. The main idea of the CV models, as an alternative solution to the Mumford–Shah problem [8], is to find the smooth curve  $C \subset \Omega$  which segments the image into non-overlapping regions, and two constant functions  $c_1$  and  $c_2$  that estimate the mean image intensities inside and outside the curve C, respectively. The CV energy functional can be written as:

$$E^{CV}(c_1, c_2, C) = \int_{\text{inside}(c)} |I_0(\mathbf{x}) - c_1|^2 d\mathbf{x} + \int_{\text{outside}(C)} |I_0(\mathbf{x}) - c_2|^2 d\mathbf{x} + \mu. \text{ Lengt}$$
(1)

where  $\mu$  denotes a positive constant and inside(C) and outside(C) denote foreground and background regions, respectively, and Length(C) denotes the length of curve C.

A main disadvantage of the CV model is that since it assumes that the image intensity is piecewise constant, it generally considers global image information in the segmentation procedure. Therefore, the CV model yields poor segmentation results for images with intensity inhomogeneity, due to wrong movement of evolving curves guided by global image information.

### B. Local Chan–Vese model

The local CV model [16] is proposed to handle images with intensity inhomogeneity. It uses global regional image information and local intensity information in the LCV energy functional as follows:

$$E^{LCV}(c_1, c_2, d_1, d_2, C) = E_1 + E_2$$
(2)

where  $E_1$  and  $E_2$  are defined as follows:

$$E_{1} = \int_{\text{inside}(c)} |I_{0}(\mathbf{x}) - c_{1}|^{2} d\mathbf{x} + \int_{\text{outside}(C)} |I_{0}(\mathbf{x}) - c_{2}|^{2} d\mathbf{x}$$
  
and

$$E_{2} = \int_{\text{inside(c)}} (g_{k} * I_{0}(\mathbf{x}) - I_{0}(\mathbf{x}) - d_{1})^{2} d\mathbf{x}$$
$$+ \int_{\text{outside(C)}} (g_{k} * I_{0}(\mathbf{x}) - I_{0}(\mathbf{x}) - d_{1})^{2} d\mathbf{x}$$

where  $g_k$  is the averaging operator, with a k×k size window which is implemented to consider local intensity information.  $d_1$  and  $d_2$  are constant functions and approximate the mean intensities difference image inside(C) and outside(C), respectively.

By applying the level set method and replacing the unknown curve C with signed distance function,  $\emptyset(\mathbf{x})$ , (2), can be reformulated as follows:

$$E^{LCV}(c_1, c_2, d_1, d_2, \emptyset) =$$

where

$$\begin{cases} F_1 = (|I_0(\mathbf{x}) - c_1|^2 + (g_k * I_0(\mathbf{x}) - I_0(\mathbf{x}) - d_1)^2) \\ F_2 = (|I_0(\mathbf{x}) - c_2|^2 + (g_k * I_0(\mathbf{x}) - I_0(\mathbf{x}) - d_2)^2) \end{cases}$$

and H(z),  $\mathcal{L}(\emptyset(\mathbf{x}))$  and  $\mathcal{P}(\emptyset(\mathbf{x}))$  denote Heaviside function, length functional and level set regularization term, respectively. They are defined as follows:

$$\begin{aligned} H(z) &= \begin{cases} 1 & \text{if } z \ge 0\\ 0 & \text{if } z \le 0 \end{cases} \\ \mathcal{L}(\emptyset(\mathbf{x})) &= \int_{\Omega} \delta(\emptyset(\mathbf{x})) |\nabla(\emptyset(\mathbf{x}))| d\mathbf{x} \\ \mathcal{P}(\emptyset(\mathbf{x})) &= \int_{\Omega} \frac{1}{2} (|\nabla \emptyset(\mathbf{x})| - 1)^2 d\mathbf{x} \end{aligned}$$
(4)

in which  $\delta(z)$  denotes the Dirac delta function  $\delta(z) = \frac{\partial H(z)}{\partial z}$ .

The Heaviside function H(z) can be approximated by a smooth function  $H_{\epsilon}(z)$ , which is defined as [16]:

$$H_{\varepsilon}(z) = \frac{1}{2} \left[ 1 + \frac{2}{\pi} \arctan\left(\frac{z}{\varepsilon}\right) \right]$$
(5)

where  $\varepsilon$  is a positive constant.

According to (5), the energy functional (3) can be approximated as follows:

$$\begin{split} & E_{\epsilon}^{\ LCV}(c_{1},c_{2},d_{1},d_{2},\varnothing) = \left(\int_{\Omega}H_{\epsilon}(\emptyset(\mathbf{x}))F_{1}d\mathbf{x}\right) + \\ & \left(\int_{\Omega}(1-H_{\epsilon}(\emptyset(\mathbf{x}))F_{2}d\mathbf{x}\right) + \mu\mathcal{L}_{\epsilon}(\emptyset(\mathbf{x})) + \mathcal{P}_{\epsilon}(\emptyset(\mathbf{x})) \end{split}$$
(6)

where  $\mathcal{L}_{\varepsilon}(\emptyset(\mathbf{x}))$  and  $\mathcal{P}_{\varepsilon}(\emptyset(\mathbf{x}))$  denote the approximated version of  $\mathcal{L}(\emptyset(\mathbf{x}))$  and  $\mathcal{P}(\emptyset(\mathbf{x}))$ , respectively.

To minimize  $E_{\varepsilon}^{LCV}$  with respect to  $\emptyset$  for fixed  $(c_1, c_2, d_1, d_2)$ , an artificial time variable  $t \ge 0$  is used to obtain the following gradient descent equation:

$$\begin{split} &\frac{\partial \emptyset}{\partial t} = \delta_{\varepsilon}(\emptyset) [-((I_0(\mathbf{x}) - c_1)^2 - (I_0(\mathbf{x}) - c_2)^2) - \\ &((g_k * I_0(\mathbf{x}) - I_0(\mathbf{x}) - d_1)^2 - (g_k * I_0(\mathbf{x}) - I_0(\mathbf{x}) - d_2)^2)] \\ &+ \mu \delta_{\varepsilon}(\emptyset) \operatorname{div} \left(\frac{\nabla \emptyset}{|\nabla \emptyset|}\right) + \left(\nabla^2 \emptyset - \operatorname{div} \left(\frac{\nabla \emptyset}{|\nabla \emptyset|}\right)\right) \end{split}$$
(7)

where  $\delta_{\varepsilon}(z) = \frac{1}{\pi} \left( \frac{\varepsilon}{\varepsilon^2 + z^2} \right)$ .

#### C. Split Bergman Method

The general form of  $L_1$ -regularized minimization problems can be written as follows:

$$\min_{\mathbf{u}} |\mathbf{A}(\mathbf{u})|_1 + \mathbf{H}(\mathbf{u}) \tag{8}$$

where  $|.|_1$  denotes  $L_1$ -norm, and both  $|A(.)|_1$  and H(.) are convex functionals.

To solve (8), the split Bergman method [17, 18] utilizes splitting to "de-couple" the  $L_1$  and  $L_2$  terms in (8). Therefore (8) can be reformulated as follows:

$$\min_{\mathbf{u},\mathbf{d}} |\mathbf{d}|_1 + H(\mathbf{u}) \tag{9}$$

where  $\mathbf{d} = A(\mathbf{u})$ .

Using a quadratic penalty function, (9) can be converted to an unconstrained minimization problem as follows:

$$\min_{\mathbf{u},\mathbf{d}}(|\mathbf{d}|_1 + \mathbf{H}(\mathbf{u}) + \frac{\nu}{2}\|\mathbf{d} - \mathbf{A}(\mathbf{u})\|^2)$$
(10)

where v and  $\|.\|$  denote a positive constant and the L<sub>2</sub>-norm, respectively. To enforce the constraint  $\mathbf{d} = A(u)$ , a vector  $\mathbf{b}$  is added inside the quadratic penalty function and Bergman iteration is performed. Then (10) can be considered as a sequence of unconstrained optimization problems and Bergman updates:

$$(u^{k+1}, \mathbf{d}^{k+1}) = \arg\min_{\substack{u, \vec{d} \\ u, \vec{d}}} (|\mathbf{d}|_1 + H(u) + \frac{v}{2} \|\mathbf{d} - A(u) - \mathbf{b}^k\|^2)$$
  
$$\mathbf{b}^{k+1} = \mathbf{b}^k + A(u^{k+1}) - \mathbf{d}^{k+1}$$
(11)

The existence of convergence for the above iterations (11) which was named the Split Bergman algorithm has been considered in [13, 17, 18].

#### D. The fuzzy c-means clustering algorithm

The fuzzy c-means clustering algorithm was first suggested by Dunn [19], and further developed by Bezdek [20]. FCM aims to partition  $\xi = (\xi_1, ..., \xi_n)$ , the set of voxel or pixel locations in  $\Omega$ , into  $\eta$  clusters  $2 \le \eta \le n$  that are determined by prototypes  $v = (v_1, ..., v_\eta)$ . With fuzzy clustering, every pixel  $\xi_j$ is a member of all clusters at the same time, but with different membership degrees. The FCM algorithm performs clustering by solving:

$$\begin{aligned} \text{Minimize } J_{\text{FCM}}(U, V) &= \sum_{i=1}^{\eta} \sum_{j=1}^{n} u_{ij}^{m} \left\| \xi_{j} - v_{i} \right\|^{2} \end{aligned} \tag{12} \\ \text{Subject to } U \in \mathcal{M} \end{aligned}$$

Where  $\mathcal{M} =$ 

$$\left\{ U = \left[ u_{ij} \right]_{\substack{i=1,\dots,k\\j=1,\dots,n}} \middle| u_{ij} \in [0,1], \sum_{i=1}^{\eta} u_{ij} = 1, \sum_{i=1}^{\eta} u_{ij} > 0 \right\}$$

and d is the distance metric based on the Euclidean norm and the parameter m,  $1 \le m \le \infty$ , influences the fuzziness of the partition. As can be seen from (12) every pixel has the same overall weight in the image's data set. The fuzzy clustering method under the mentioned constraint is often named probabilistic clustering and  $u_{ij}$  can be considered as the posterior probability  $p(\eta_i | \xi_j)$  [21]. Alternating optimization, which alternates between optimizations of  $\tilde{J}_{FCM}(U|\dot{v})$  over U with fixed  $\dot{v}$  and  $\tilde{J}_{FCM}(v|\dot{U})$  over v with fixed  $\dot{U}$ , would converge to a minimizer or a saddle point of  $J_{FCM}$ [22].

## III. SPATIAL FUZZY CLUSTERING AS A TWO-STAGE PROCEDURE

Our segmentation algorithm involves the following two stages.

Stage 1: a) Image pre-segmentation based on the globally convex local Chan–Vese model (GCLCV)

The GCS and split Bergman methods are used to presegment the cell image based on the local and global region information.

## b) Finding the prior memberships of each pixel

The Classical Bayes classifier is used to estimate the membership degrees of each pixel based on the segmentation result in (a).

Stage 2: Spatial FCM using local and global region information for final segmentation

A cell image and the prior membership degrees of each pixel computed in stage 1 are taken as inputs to the spatial FCM for final segmentation.

### A. Globally convex local Chan–Vese model (GCLCV)

Since the LCV energy functional is non-convex, it might tend towards undesirable local minima. Chan et al. [12] proposed the GCS technique to show that the non-convex energy of the CV model can be reformulated as a convex optimization problem. Goldstein et al. [13] utilized the GCS technique and the Split Bergman algorithm to propose a globally convex and fast segmentation model based on the CV model (GCCV) which is unable to use local information and handle images with intensity inhomogeneity. This section proposes the globally convex energy functional (GCLCV) based on the LCV model, the GCS technique and the split Bergman algorithm.

In order to utilize the GCS technique for the gradient flow equation (7), we drop the last term in the gradient flow equation. Dropping the last term guarantees the global minima of the GCS technique and the efficiency of applying Split Bergman Algorithm for the minimization procedure [13]. We get:

$$\frac{\partial \emptyset}{\partial t} = \delta_{\varepsilon}(\emptyset) \left[ -\lambda_1 ((I_0(\mathbf{x}) - c_1)^2 - (I_0(\mathbf{x}) - c_2)^2) -\lambda_2 ((g_k * I_0(\mathbf{x}) - I_0(\mathbf{x}) - d_1)^2 - (g_k * I_0(\mathbf{x}) - I_0(\mathbf{x}) - d_2)^2) \right] + \mu \delta_{\varepsilon}(\emptyset) \operatorname{div}\left(\frac{\nabla \emptyset}{|\nabla \varphi|}\right)$$
(14)

With no loss of generality, we take  $\mu = 1$ . Therefore (14) can be written as follows:

$$\frac{\partial \emptyset}{\partial t} = \delta_{\varepsilon}(\emptyset) \left[ P_1 + P_2 + \operatorname{div}\left(\frac{\nabla \emptyset}{|\nabla \emptyset|}\right) \right]$$
(15)

where

~~

$$P_{1}(\mathbf{x}) = -\lambda_{1}((I_{0}(\mathbf{x}) - c_{1})^{2} - (I_{0}(\mathbf{x}) - c_{2})^{2})$$
  
and  
$$P_{2}(\mathbf{x}) = -\lambda_{2}((g_{k} * I_{0}(\mathbf{x}) - I_{0}(\mathbf{x}) - d_{1})^{2} - (g_{k} * I_{0}(\mathbf{x}) - I_{0}(\mathbf{x}) - d_{2})^{2})$$

Following the idea of the GCS technique, we find that the simplified flow of (15) has the same stationary solution as the

original gradient flow equation. Note that  $H_{\varepsilon}$  is a noncompactly supported, smooth approximation of H. Therefore by removing  $\delta_{\varepsilon}$ , the following gradient descent equation which has the same stationary solution as (15) results:

$$\frac{\partial \emptyset}{\partial t} = \left[ P_1 + P_2 + \operatorname{div}\left(\frac{\nabla \emptyset}{|\nabla \emptyset|}\right) \right]$$
(16)

The equation above is the gradient descent for the following energy functional:

$$E^{GCLCV}(\emptyset) = \int_{\Omega} |\nabla \emptyset(\mathbf{x})| d\mathbf{x} + \int_{\Omega} \emptyset(\mathbf{x}) \Phi(\mathbf{x}) d\mathbf{x}$$
(17)

where  $\Phi(\mathbf{x}) = -(P_1(\mathbf{x}) + P_2(\mathbf{x}))$ . Now (17) can be reformulated as follows:

$$E^{GCLCV}(\emptyset) = |\emptyset|_1 + \langle \emptyset, \Phi \rangle$$
(18)

Since  $E^{GCLCV}$  is homogeneous of degree 1 in  $\emptyset$  and its representation with level sets is not unique, the minimization procedure does not yield a global minimum. To guarantee the global minimum [12], the minimization of  $\emptyset$  can be restricted such that:

$$\alpha \le \emptyset(\mathbf{x}) \le \beta \qquad \forall \mathbf{x} \in \Omega \tag{19}$$

Hence, the minimization problem for GCLCV can be written as follows:

$$\min_{\emptyset \in [\alpha,\beta]} \mathbb{E}^{\text{GCLCV}}(\emptyset) = \min_{\emptyset \in [\alpha,\beta]} (|\emptyset|_1 + \langle \emptyset, \Phi \rangle)$$
(20)

The Split Bergman algorithm can be utilized to solve the minimization problem given above. We use splitting for "decoupling" the  $L_1$  and  $L_2$  terms, so (20) becomes:

$$(\emptyset^*, \mathbf{d}^*) = \arg\min_{\emptyset \in [\alpha, \beta], \vec{\mathbf{d}}} (|\mathbf{d}|_1 + \langle \emptyset, \Phi \rangle + \frac{\nu}{2} ||\mathbf{d} - \nabla \emptyset||^2) \quad (21)$$

where **d** is an auxiliary variable and  $\mathbf{d} = \nabla \emptyset$ ,  $(\emptyset^*, \mathbf{d}^*)$  is the solution,  $\|.\|$  is L<sub>2</sub>-norm and  $\nu$  denotes the positive constant parameter.

Since the constraint  $\mathbf{d} = \nabla \phi$  is weakly enforced, a vector **b** is involved inside the quadratic penalty function and the Bergman iteration is used to enforce the constraint. Therefore (21) can be written as sequence of unconstrained minimization:

$$\begin{pmatrix} \emptyset^{k+1}, \mathbf{d}^{k+1} \end{pmatrix} = \arg\min_{\emptyset \in [\alpha, \beta], \vec{d}} (|\mathbf{d}|_1 + \langle \emptyset, \Phi \rangle + \frac{\nu}{2} \| \mathbf{d} - \nabla \emptyset - \mathbf{b}^{k+1} \|^2)$$

$$\mathbf{b}^{k+1} = \mathbf{b}^k + \nabla \emptyset^{k+1} - \mathbf{d}^{k+1}$$

$$(22)$$

Based on the calculus of variations, (22) can be minimized over  $\emptyset$  with fixed **d** by applying the following Euler-Lagrange equation:

$$\Delta \emptyset = \frac{\Phi}{\nu} + \nabla . \left( \mathbf{d} - \mathbf{b} \right) \quad \text{for } \emptyset \in [\alpha, \beta]$$
(24)

Then the central difference for the Laplace operator and a backward difference for the divergence operator are performed to obtain the following numerical scheme:

$$\begin{cases} u_{i,j} = \mathbf{d}_{i-1,j}^{x} - \mathbf{d}_{i,j}^{x} - \mathbf{b}_{i-1,j}^{x} + \mathbf{b}_{i,j}^{x} + \mathbf{d}_{i,j-1}^{y} - \mathbf{d}_{i,j}^{y} - \mathbf{b}_{i,j-1}^{y} + \mathbf{b}_{i,j}^{y} \\ w_{i,j} = \frac{1}{4} \Big( \emptyset_{i-1,j} + \emptyset_{i+1,j} + \emptyset_{i,j-1} + \emptyset_{i,j+1} - \frac{\Phi}{\nu} + u_{i,j} \Big) \\ \theta_{i,j} = \max\{\min\{w_{i,j}, \beta\}, \alpha\} \end{cases}$$
(25)

Vector-valued shrinkage operator is used to minimize (22) with respect to  $\mathbf{d}$ :

$$\mathbf{d}^{k+1} = \max\{\|\mathbf{b}^{k} + \nabla \emptyset^{k+1}\| - \nu, 0\}\frac{\mathbf{b}^{k} + \nabla \emptyset^{k+1}}{\|\mathbf{b}^{k} + \nabla \emptyset^{k+1}\|}$$
(26)

Note that at each iteration before updating  $\emptyset$ , for a fixed function  $\emptyset$ , the constant functions are needed to be updated as follows:

$$\begin{cases} c_{i} = \frac{\int_{\Omega} I_{0}(\mathbf{x})_{i}^{c}(\boldsymbol{\emptyset}(\mathbf{x}))d\mathbf{x}}{\int_{\Omega} \frac{c}{i}(\boldsymbol{\emptyset}(\mathbf{x}))d\mathbf{x}} \\ d_{i} = \frac{\int_{\Omega} (g_{k}*I_{0}(\mathbf{x}) - I_{0}(\mathbf{x}))_{i}^{c}(\boldsymbol{\emptyset}(\mathbf{x}))d\mathbf{x}}{\int_{\Omega} \frac{c}{i}(\boldsymbol{\emptyset}(\mathbf{x}))d\mathbf{x}} \end{cases} \quad i = 1,2,$$

$$(27)$$

where  ${}_{1}^{\varepsilon}(\emptyset) = H_{\varepsilon}(\emptyset)$  and  ${}_{2}^{\varepsilon}(\emptyset) = (1 - H_{\varepsilon}(\emptyset))$ .

# B. Prior membership degree computation

Assume the part of interest in GCLCV results is  $\emptyset(x, y)$ .  $\mathfrak{C}_1 = \{\xi | \emptyset(x, y) \ge 0\}$  and  $\mathfrak{C}_2 = \Omega \setminus \mathfrak{C}_1$  are defined with cardinality  $\tau_1$  and  $\tau_2$ , respectively. Thus for every pixel  $\xi$ , it can be supposed that

$$\xi \in \mathfrak{C}_i \Leftrightarrow p(\mathfrak{C}_i | \xi) \ge p(\mathfrak{C}_j | \xi) \ \forall \ i \neq j$$

Based on Bayesian rule,  $p(\mathfrak{C}_i|\xi)$  can be calculated as follows for i = 1, 2:

$$p(\mathfrak{C}_{i}|\xi) = \frac{p(\xi|\mathfrak{C}_{i})P(\mathfrak{C}_{i})}{\sum_{j=1}^{2} p(\xi|\mathfrak{C}_{j})P(\mathfrak{C}_{j})}$$
(28)

In practice, priors  $P(\mathfrak{C}_i)$  are estimated on the basis of the proportion of pixels in class  $\mathfrak{C}_i$ . To represent the  $p(\xi|\mathfrak{C}_i)$ , it can be assumed that distribution generated by the class  $\mathfrak{C}_i$  is characterized using a Gaussian kernel as follows:

$$p(\xi|\mathfrak{C}_{i}) = \frac{1}{|2\pi F_{i}|^{\frac{n}{2}}} \exp\left(-\frac{1}{2}(\xi - \mu_{i})^{T}F_{i}^{-1}(\xi - \mu_{i})\right) \text{ for } i = 1,2$$

where  $\mu_i$  and  $F_i$  denote the mean value and the covariance matrix of pixels in  $\mathfrak{C}_i$  respectively.

The posterior probability  $p(\mathfrak{C}_i|\xi)$  can thus be formulated as

$$p(\mathfrak{C}_{i}|\xi) = \frac{p(\xi|\mathfrak{C}_{i})P(\mathfrak{C}_{i})}{\sum_{j=1}^{2} p(\xi|\mathfrak{C}_{j})P(\mathfrak{C}_{j})}$$
(29)

In which

$$p(\xi|\mathfrak{C}_{i})P(\mathfrak{C}_{i}) = \frac{\frac{\tau_{i}}{\Sigma_{k=1}^{2}\tau_{k}}}{|2\pi F_{i}|^{\frac{n}{2}}} \exp\left(-\frac{1}{2}(\xi-\mu_{i})^{T}F_{i}^{-1}(\xi-\mu_{i})\right)$$

In the same way the membership values of the pixels are organized in matrix form:

$$\boldsymbol{\aleph} = \left\{ \begin{bmatrix} \boldsymbol{\aleph}_{ij} \end{bmatrix}_{\substack{i=1,2\\j=1,\dots,n}} \middle| \, \boldsymbol{\aleph}_{ij} = p\left(\boldsymbol{\mathfrak{C}}_i \middle| \boldsymbol{\xi}_j\right) \right\}$$
(30)

# *C. Spatial fuzzy clustering*

To improve the performance of FCM, prior membership degrees  $\aleph$  are implemented in the objective function (12) as follows:

$$J_{i=1} (u, v) = \sum_{i=1}^{2} \sum_{j=1}^{n} u_{ij}^{m} \|\xi_{j} - v_{i}\|^{2} + \gamma (u_{ij} - \aleph_{ij})^{m} \|\xi_{j} - v_{i}\|^{2}$$
(31)

where  $\gamma$  is called control parameter whose role is to keep a balance between the original FCM and GCLCV method in the optimization procedure.

The optimal values of  $v_i$  and  $u_{ij}$  is solved as followed. Considering the constraint (13) and the Lagrangian multiplier  $\lambda$ , we define the following Lagrangian function  $\mathcal{L}$ :

$$\begin{aligned} \mathcal{L}(U, V, \lambda) &= \sum_{i=1}^{2} \sum_{j=1}^{n} u_{ij}^{m} \left\| \xi_{j} - v_{i} \right\|^{2} \\ &+ \gamma \left( u_{ij} - \aleph_{ij} \right)^{m} \left\| \xi_{j} - v_{i} \right\|^{2} - \lambda (\sum_{i=1}^{\eta} u_{ij} - 1) \end{aligned}$$

Assume that  $u_{kj}$  is the membership degree of  $\xi_j$  belonging to the cluster  $\eta_k$  whose centroid is  $v_k$ . The stationary point of the optimized functional can be defined as  $(u_{kj}, v_k, \lambda)$  [23] if and only if  $\frac{\partial \mathcal{L}}{\partial \lambda} = 0$ ,  $\frac{\partial \mathcal{L}}{\partial v_k} = 0$  and  $\frac{\partial \mathcal{L}}{\partial u_{kj}} = 0$ . Taking these derivatives returns the relationships as follows:

$$\begin{split} \frac{\partial \mathcal{L}}{\partial \lambda} &= \sum_{i=1}^{2} u_{kj} - 1 = 0\\ \frac{\partial \mathcal{L}}{\partial v_{k}} &= 2 \sum_{j=1}^{n} u_{kj}^{m} \left( \xi_{j} - v_{k} \right) + \gamma \left( u_{kj} - \aleph_{kj} \right)^{m} \left( \xi_{j} - v_{k} \right) = 0 \end{split}$$

The optimal solution of  $v_k$  is given by:

$$\mathbf{v}_{\mathbf{k}} = \frac{\sum_{j=1}^{n} u_{kj}^{m} \xi_{j} + \gamma \left( u_{kj} - \boldsymbol{\aleph}_{kj} \right)^{m} \xi_{j}}{\sum_{j=1}^{n} u_{kj}^{m} + \gamma \left( u_{kj} - \boldsymbol{\aleph}_{kj} \right)^{m}}$$
(32)

Similarly:

$$\begin{split} \frac{\partial \mathcal{L}}{\partial u_{kj}} &= m u_{kj}^{m-1} \left\| \xi_j - v_k \right\|^2 \\ + \gamma m \left( u_{kj} - \aleph_{kj} \right)^{m-1} \left\| \xi_j - v_k \right\|^2 - \lambda = 0 \end{split}$$

~ ~

Let m = 2 we have

$$2u_{kj} \left\| \xi_{j} - v_{k} \right\|^{2} + 2\gamma \left( u_{kj} - \aleph_{kj} \right) \left\| \xi_{j} - v_{k} \right\|^{2} = \lambda$$
$$u_{kj} + \gamma \left( u_{kj} - \aleph_{kj} \right) = \frac{\lambda}{2 \left\| \xi_{j} - v_{k} \right\|^{2}}$$
$$u_{kj} = \frac{1}{1 + \gamma} \left( \frac{\lambda}{2 \left\| \xi_{j} - v_{k} \right\|^{2}} - \gamma \aleph_{kj} \right)$$

 $u_{ki} \in \mathcal{M}$  indicates

$$\boldsymbol{\lambda} = 2 \left( 1 + \gamma \big( 1 - \sum_{i=1}^2 \boldsymbol{\kappa}_{ij} \big) \right) \left( \sum_{i=1}^2 \frac{1}{\left\| \boldsymbol{\xi}_j - \boldsymbol{v}_i \right\|^2} \right)^{-1}$$

Therefore, the optimal solution of  $u_{kj}$  is given by:

$$\mathbf{u}_{kj} = \frac{1}{1+\gamma} \left( \frac{1+\gamma(1-\sum_{i=1}^{2} \aleph_{ij})}{\sum_{i=1}^{2} \frac{\left\| \xi_{j} - \mathbf{v}_{k} \right\|^{2}}{\left\| \xi_{j} - \mathbf{v}_{i} \right\|^{2}}} + \gamma \aleph_{kj} \right)$$
(33)

## IV. EXPERIMENTAL RESULTS

In this section, we compare the performance of our proposed algorithm GCLCV and our spatial fuzzy clustering based on the global and local region information(SFCGL) with conventional FCM [19, 20], CV [11], SFCM [24], SFLS [25] and region-scalable fitting energy (RSFE) [26], LCV [16], Otsu thresholding (OT)[27], watershed algorithm(WA)[28], GCCV[13]. The data set consists of human HT29 colon cancer cells images with a size of  $512 \times 512$  pixels from image set BBBC008v1 [29, 30]. The ground truth is available for this data set. An example of an original image and segmentation result based on the proposed method is shown in Fig. 1. To evaluate the performance of our algorithm, recognition error rate (RER) and Dice coefficient were calculated. RER is defined as:

$$\operatorname{RER}(\%) = \frac{T_{\text{cell}} + T_{\text{background}}}{n} \times 100$$
(34)

where n denotes the total number of pixels in the given image,  $T_{cell}$  and  $T_{background}$  denote the number of pixels incorrectly classified as the cell and background, respectively. Dice coefficient is defined as:

$$Dice(R,S) = \frac{2|R \cap S|}{|R| + |S|}$$
(35)

where R denotes the binary reference image, and S denotes the segmented image.

The number of false positives (FP) and the number of false negatives (FN) are reported.



Fig 1. Example of (a) Original image, (b) Ground truth and (c) Segmentation result

TABLE 1. QUANTITATIVE RESULTS FOR THE BBBC008v1 AND NOISY BBBC008v1 FOR THE DIFFERENT APPROACHES

ADDED

HUMAN HT29 COLON CANCER CELLS WITH 5% GAUSSIAN NOISE

HUMAN HT29 COLON CANCER CELLS

| Method | RER(% | Dice | FP  | FN  | RER (%) Dice |      | FP  | FN  |
|--------|-------|------|-----|-----|--------------|------|-----|-----|
| CV     | 4.68  | 0.75 | 0.9 | 3.1 | 5.28         | 0.68 | 1.1 | 3.6 |
| FCM    | 5.39  | 0.69 | 1.9 | 4.1 | 6.11         | 0.61 | 2.2 | 3.9 |
| RSFE   | 3.12  | 0.78 | 1.5 | 3.6 | 4.64         | 0.73 | 1.8 | 3.2 |
| LCV    | 3.04  | 0.76 | 1.6 | 3.2 | 3.55         | 0.71 | 1.9 | 3.3 |
| SFLS   | 3.54  | 0.73 | 1.6 | 4.2 | 4.39         | 0.65 | 2   | 4.3 |
| SFCM   | 4.46  | 0.71 | 1.8 | 3.9 | 5.14         | 0.66 | 2.3 | 3.9 |
| WA     | 5.91  | 0.67 | 2.1 | 4.8 | 6.41         | 0.65 | 2.6 | 3.8 |
| OT     | 4.82  | 0.85 | 0.8 | 2.7 | 6.32         | 0.81 | 1.1 | 3.7 |
| GCCV   | 2.98  | 0.79 | 0.9 | 3.2 | 3.38         | 0.77 | 1.3 | 2.8 |
| GCLCV  | 2.6   | 0.92 | 0.7 | 3.3 | 3.12         | 0.88 | 1.1 | 2.8 |
| SFCGL  | 1.63  | 0.94 | 0.8 | 3.4 | 2.33         | 0.91 | 0.9 | 2.9 |

Table 1 displays the segmentation results of the proposed SFCGL and GCLCV approach averaged over all images in the data set for both clean and noisy images. As can be seen from Table 1 the proposed SFCGL approach produces the best results for the RER (%) and Dice coefficient. In particular, for the more challenging noisy images significantly better results is obtained in comparison with existing approaches. Furthermore, it can be seen that the proposed GCLCV and the spatial fuzzy clustering yield better results than the GCCV. Also, the proposed GCLCV approach yields better results than the convex functionals, GCCV and GCLCV, yield better results than the non-convex functionals, CV, RSFE, and LCV.

It is remarkable that, the minimization of the GCLCV functional for an image of size of  $512 \times 512$  pixels converges after about 6 iterations and the computation time is

approximately 215 s per image (using an Intel core i7 CPU 2640 M, 2.8 GHz, with 8 GB RAM, and Windows 7 64 bit). In comparison, utilizing a standard level set method for the LCV with gradient descent optimization requires about 60 iterations and the computation time is about 601 s on the same machine.

## V. CONCLUSION

A new spatial fuzzy clustering algorithm using global and local region information (SFCGL) is introduced in this paper. The proposed SFCGL algorithm included as its first stage a novel GCLCV algorithm based on the GCS method and Split Bergman algorithm. In the second stage, the spatial information captured by the GCLCV algorithm is transferred to the FCM algorithm through an induced prior membership defined by the posterior probabilities. The proposed algorithm is fully automated. The experimental results demonstrated that the proposed approaches can effectively yield significantly better segmentation accuracy in cell image analysis.

#### ACKNOWLEDGMENT

This work was supported by the Australian Research Council (ARC) Discovery Grant DP1097059 and the Griffith International Postgraduate Research Scholarship (GUIPRS).

#### REFERENCES

- J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: theory, algorithm and applications in image segmentation," Graph. Models Image Process, vol. 58, no. 3, pp. 246–261, May 1996.
- [2] D. L. Pham and J. L. Prince, "An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities," Pattern Recognit. Lett, vol. 20, pp. 57–68, Jan 1999.
- [3] A.W.C. Liew, S.H. Leung and W.H. Lau, "Segmentation of Color Lip Images by Spatial Fuzzy Clustering," IEEE Transactions on Fuzzy System, vol. 11, no. 4, pp. 542-549, August 2003.
- [4] S. Chen, D.Q. Zhang, "Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure," IEEE Trans Systems, Man and Cybernetics-Part B, vol. 34, no. 4, pp. 1907–1916, Aug 2004.
- [5] A.W.C. Liew, H. Yan," An adaptive spatial fuzzy clustering algorithm for 3-DMR image segmentation", IEEE Trans. on Medical Imaging, vol. 22, no. 9, pp. 1063–1075, Sep 2003.
- [6] J.Z. Wang, J. Kong, Y. H. Lu, M. Qi, B. X. Zhang," A modified algorithm FCM for MRI brain image segmentation using both local and non-local spatial constrains", Computerized Medical Imaging and Graphics, vol. 32, no. 8, pp. 685–698, Sep 2008.
- [7] A.W.C. Liew, S.H. Leung and W.H. Lau, "Fuzzy Image Clustering Incorporating Spatial Continuity", in Proc. Vision, Image and Signal Processing, vol. 147, no.2, pp. 185-192, 2000.
- [8] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modelling with front propagation: A level set approach," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 17, no. 2, pp. 158–175, Feb 1995.
- [9] V. Caselles, F. Catte, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," Numerische Mathematik, vol. 66, no. 1, pp.1–31,1993.

- [10] S. Osher and J. A. Sethian, "Fronts propagating with curvaturedependent speed: Algorithms based on Hamilton–Jacobi Formulation," Journal of Computational Physics, vol. 79, pp. 12– 49, Nov 1988.
- [11] T.F. Chan and L. A. Vese, "Active Contours without edges," IEEE Trans. Image Processing, vol. 10, no.2, pp.266-277, Feb 2001.
- [12] T. F. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," SIAM Journal on Applied Mathematics, vol. 66, no. 5, pp.1632-1648, 2006.
- [13] T. Goldstein, X. Bresson, S. Osher, "Geometric applications of the split Bregman method: segmentation and surface reconstruction," SIAM Journal on Applied Mathematics, vol. 45 (1–3), pp. 272– 293, Oct 2010.
- [14] H. I. Bozma, and James S. Duncan. "A game-theoretic approach to integration of modules," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no.11, pp.1074-1086, Nov 1994.
- [15] P. Parent, S.W. Zucker, "Trace inference, curvature consistency, and curve detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.11, no. 8 pp. 823-839, Aug1989.
- [16] Wang, Xiao-Feng, De-Shuang Huang, and Huan Xu. "An efficient local Chan–Vese model for image segmentation," Pattern Recognition, vol. 43, no.3, pp. 603-618, Mar 2010.
- [17] T. Goldstein, S. Osher, "The split Bregman method for L1 regularized problems," SIAM Journal of Imaging Science, vol. 2, no. 2, pp. 323–343,2009.
- [18] X. Bresson, S. Esedoglu, P. Vandergheynst, J.-P. Thiran, S.Osher, "Fast global minimization of the active contour/snake mode", Journal of Mathematical Imaging and Vision, vol. 28, pp. 151– 167, Jun 2007.
- [19] J. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact, well-separated clusters," Journal of Cybernetics, vol. 3, no. 3, pp. 32–57, Sep 1973.
- [20] J. Bezdek, Pattern recognition with fuzzy objective function algorithms, Plenum Press, New York, 1981.
- [21] N.R. Pal, K. Pal, J. M. Keller, J. Bezdek, "A possibilistic fuzzy cmeans clustering algorithm," IEEE Trans. Fuzzy Systems, vol. 13, no. 4, pp. 517–530, Aug 2005.
- [22] F. Höppner and F. Klawonn, "A contribution to convergence theory of fuzzy c-means and derivatives," IEEE Trans. Fuzzy Systems, vol. 11, no. 5, pp. 682–694, Oct 2003.
- [23] W. Pedrycz, and J. Waletzky. "Fuzzy clustering with partial supervision," IEEE Trans. Systems, Man, and Cybernetics, Cybernetics, vol. 27, no.5 pp. 787-795, Sep 1997.
- [24] K.S. Chuang, H.L. Hzeng, S. Chen, J. Wu, T.J. Chen, 'Fuzzy cmeans clustering with spatial information for image segmentation," Computerized Medical Imaging and Graphics, vol.30, pp. 9–15, Jan 2006.
- [25] B. N. Li, C. K. Chui, S. Chang, and S. H. Ong, "Integrating spatial fuzzy clustering and level set methods for automated medical image segmentation," Computers in Biology and Medicine, vol. 41, pp.1–10, Jan 2011.
- [26] C. Li, C. Kao, J. Gore, Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," IEEE Trans. Image Processing, vol.17, pp.1940–1949, Oct 2008.
- [27] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. System, Man and Cybernetics, vol. 9, pp. 62–66, Jan 1979.
- [28] G. Lin, U. Adiga, K. Olson, J.F. Guzowski, C.A. Barnes, B. Roysam, "A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks," Cytometry, part A, vol. 56, pp.23–36, Nov 2003.

- [29] www.broadinstitute.org/bbbc
- [30] A.E. Carpenter, T.R. Jones, M.R. Lamprecht, C. Clarke, IH. Kang, O. Friman, DA. Guertin, JH. Chang, RA. Lindquist, J. Moffat, P. Golland, DM. Sabatini, "CellProfiler: image analysis software for identifying and quantifying cell phenotypes," Genome Biology, 7:R100, 2006.
- [31] A.W.C. Liew, H. Yan and N.F. Law, "Image Segmentation Based on Adaptive Cluster Prototype Estimation," IEEE Transactions on Fuzzy Systems, vol.13, no.4, pp. 444-453, August 2005.
- [32] S.L. Wang, W.H. Lau, A.W.C. Liew, and S.H. Leung, "Robust Lip Region Segmentation for Lip Images with Complex Background," Pattern Recognition, vol. 40, no. 12, pp. 3481-3491, Dec 2007.