CELL SEGMENTATION BASED ON SPATIAL INFORMATION IMPROVED INTUITIONISTIC FCM COMBINED WITH FOPSO

Chuxiong Sun[1], *Xiangzhi Bai*[1, 2,*]

[1]Image Processing Center, Beihang University, Beijing, China;
[2]State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China

ABSTRACT

Fuzzy c-means clustering (FCM) algorithm has been proved to be effective for image segmentation. However, it is sensitive to the noises and initialization. FCM could not effectively segment cell images with inhomogeneity and complicate adhesives. Aimed to overcome these disadvantages, this paper proposes a cell image segmentation algorithm using spatial information improved intuitionistic fuzzy c-means clustering (SI-IFCM) combined with fractional-order velocity based particle swarm optimization (FOPSO). SI-IFCM and FOPSO will iterate alternately with different object functions to obtain the clustering result. Experimental results demonstrate the advantages of our algorithm for cell segmentation comparing with state-of-arts algorithms.

Index Terms— *cell image segmentation, fuzzy clustering, particle swarm optimization, spatial information*

1. INTRODUCTION

Cell segmentation is important for biomedical image analysis. Many algorithms have been proposed for cell image segmentation. These algorithms could be divided into several types: intensity threshold, edge detection, region accumulation, deformable models, fuzzy c-means clustering (FCM) based algorithms etc.

The threshold based algorithms (for example: Otsu method [1]) is simple and effective for computation but could not work well on the images with inhomogeneity. The edge detection algorithms [2,3,4] detect the possible cell contours for segmentation. They could have disconnected contours and are sensitive to the initialization. The region accumulation algorithms such as watershed transform [5,6,7,8] may result in inaccurate cell boundaries caused by over-segmentation or by the inaccurate classification of the cell region [9]. The deformable models based algorithms [10,11,12] use the deformable contours to match the actual cell contours. They are sensitive to the initial contours.

FCM [13] clusters the pixels to achieve the segmentation. It is sensitive to the initialization and could not effectively segment the targets with low gray levels. To overcome these disadvantages, many FCM based algorithms have been proposed. For example, to better describe the data with more uncertainty, the intuitionistic fuzzy c-means clustering (IFCM) [13,14,15] algorithm was proposed. IFCM considers the non-membership and intuitionistic degree [14,15]. Some other FCM based algorithms [16,17,18] have also been proposed. They could be divided into following types: improvements by incorporating spatial information, improvements by importing kernel function, improvements by developing fuzzy set theory and so on.

Intelligent optimization algorithms could be combined with FCM based algorithms to make them more flexible. Among them, the particle swarm optimization (PSO) algorithm [19,20] has advantages of considerable convergence speed and extensive applicability. However, PSO may easily converge into the local extreme.

This paper proposes a spatial information improved IFCM (SI-IFCM) combined with fractional-order velocity based particle swarm optimization (FOPSO) [21] to achieve cell segmentation. SI-IFCM is based on IFCM with an improved intuitionistic membership and local spatial information. FOPSO uses the fractional-order velocity to enhance the convergence speed [21] and has local spatial information incorporated into the fitness function. They optimize alternately.

2. PROPOSED METHOD

2.1. Spatial information improved FCM (SI-IFCM)

2.1.1. Improved intuitionistic membership optimized for inhomogeneity

To improve the performance on cell images with inhomogeneity, based on IFCM, we propose an extended fuzzy generator without conventional constraint [22] and an improved intuitionistic membership. The fuzzy generator is a function to calculate the non-membership from the membership.

Firstly, we propose the extended fuzzy generator:

$$v_{pq} = (1 - U_{pq}{}^{\alpha})^{\frac{1}{\alpha}}, \quad (1)$$

* Corresponding author: Xiangzhi Bai (jackybxz@buaa.edu.cn)

where $para3$ meets the constraint $\alpha \in (0,+\infty)$. When $\alpha \geq 1$, $\nu_{pq} + U_{pq} \geq 1$ and $\pi_{pq} = 1 - U_{pq} - \nu_{pq} < 0$. IFCM algorithm with the negative intuitionistic degree is still executable experimentally. So we extend the constraint of the fuzzy generator.

Secondly, we propose an improved intuitionistic membership. It is the combination of the membership and non-membership as below:

$$U_{pq}^{*} = U_{pq} + \beta \cdot \nu_{pq}, \tag{2}$$

where $para2$ controls the difference between the improved intuitionistic membership and the original membership.

For further explanation, we firstly define a new membership U_{pq}^{1} as below:

$$U_{pq}^{1} = \frac{U_{pq}^{*m}}{\sum_{p=1}^{c} U_{pq}^{*m}}, \tag{3}$$

obviously $0 \leq U_{pq}^{1} \leq 1$ and

$\sum_{p=1}^{c} U_{pq}^{1} = 1, \forall q \in \{1,2,3...n\}$ (usually $m = 2$) .

And, we give the formulas for updating centroids and membership matrix:

$$P_p = \sum_{q=1}^{n} U_{pq}^{1} x_q, \tag{4}$$

$$U_{pq} = \sum_{k=1}^{c} \left[\frac{D(x_q, P_p)}{D(x_q, P_k)} \right]^{\frac{-2}{m-1}}. \tag{5}$$

We calculate the object function in the form as below:

$$obj = \sum_{p=1}^{c} \sum_{q=1}^{n} (U_{pq})^m d^2(x_q, P_p). \tag{6}$$

Now we could compare the newly defined memberships with IFCM and FCM.

From **Fig. 1**, we get the following conclusions:

1. Our algorithm makes the membership more uncertain: when the original membership is around 0.5, the newly defined membership will be closer to 0.5 compared with FCM. For IFCM it is true only when $0 < p_2 < 1$ (it exactly meets the conventional constraint [22]).

2. In our algorithm, the newly defined membership does not belong to $[0,1]$ seriously, in fact:

$$U_{pq}^{1} \in [q_1, q_2] (0 \leq q_1 \leq q_2 \leq 1). \tag{7}$$

This is because:

$$\begin{aligned} U_{pq} \to 1 \Rightarrow U_{pq}^{*} &= U_{pq} + \beta \cdot (1 - U_{pq}^{\alpha})^{\frac{1}{\alpha}} \\ &\to 1 + 0 = 1, \end{aligned} \tag{8}$$

$$\begin{aligned} U_{pq} \to 0 \Rightarrow U_{pq}^{*} &= U_{pq} + \beta \cdot (1 - U_{pq}^{\alpha})^{\frac{1}{\alpha}} \\ &\to 0 + \beta = \beta, \end{aligned} \tag{9}$$

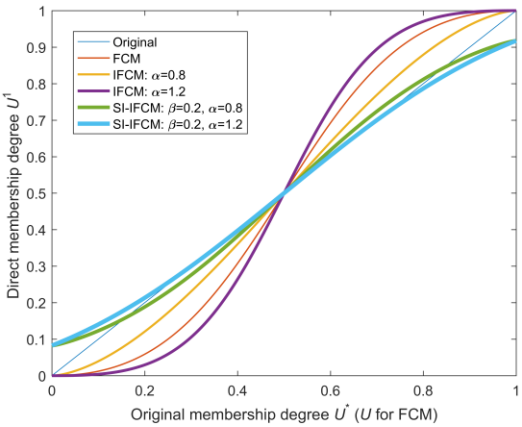

Fig. 1. Comparison between SI-IFCM, IFCM and FCM. For cell segmentation in monochrome cell images, 2 is experimentally taken as the number of cluster centroids for these algorithms. For IFCM in curve 5 and curve 6, we use the generator $\nu_{pq} = (1 - U_{pq}^{\alpha})^{\frac{1}{\alpha}}$.

$$U_{pq} \to 1 \Rightarrow U_{pq}^{1} \to \frac{1}{1+\beta^m}, \tag{10}$$

$$U_{pq} \to 0 \Rightarrow U_{pq}^{1} \to \frac{\beta^m}{1+\beta^m}. \tag{11}$$

This makes the value domain more compact. It means that the calculation of the centroids in each iteration gets fuzzier globally. It is intended for the segmentation of dim cells in cell images, which requires robustness for the variance of gray levels in cells.

2.1.2. The local spatial information incorporated in SI-IFCM to optimize adhesive cells

Our proposed algorithm would further improve the performance through importing local spatial information.

To obtain the local spatial information, the possible boundaries of adhesive cells should be extracted. The proposed algorithm searches the intersections of cell contours from the temporary result image. This image will be obtained by the global best particle in FOPSO (The definition of the particle will be introduced in next part).

For example:

$$a \in \zeta_i. \tag{12}$$

a is a point belonging to a contour ζ_i.

$$l_N(a)_1 \in \zeta_i, l_N(a)_2 \in \zeta_i. \tag{13}$$

$l_N(a)_1$ and $l_N(a)_2$ are two adjacent points of A on ζ_i.

They are on two sides of A and meet the condition as below:

$$\left| D(l_N(a)_k, a) - D_0 \right| < \varepsilon_D, k \in \{1,2\}. \tag{14}$$

Here we usually set D_0 to 2 and set ε_D to 0.5.

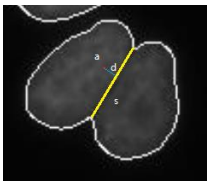

Fig. 2. Example of calculating the matrix M. The algorithm will check the neighborhood around the segment s. We take point a for example. Its distance from the target segment s is d, which will be used in matrix M.

$l_N(a)_1$ and $l_N(a)_2$ form the adjacent segment $l_N(a)$. We check if this segment meets the condition:

$$\forall c \in l_N(a), c \in P_2. \tag{15}$$

P_2 refers to the centroid representing cells.

If it meets the condition, then we classify A as a possible intersection. This rule is based on a hypothesis: The cell contours are all convex. It is considerable because the most of cells are ellipse-like.

Among these intersections, we search each point in random order, the redundant points near to the check point (with distances less than $\varepsilon_2 = 4$) will be removed.

Then we may give possible intersections. They could form possible boundaries of adhesive cells.

To import this local spatial information into our proposed algorithm, we define the improvement matrix $M_{c\times n}$.

Firstly, the improvement matrix is initialized as the matrix of ones. Then we search the points with the distances to the segment less than existed threshold ($\varepsilon_3 = 3$).Their corresponding terms in M will be updated by these distances.

$$M_{2q} = \frac{D(I_q, l_s)}{length_{ref}}, \tag{16}$$

where l_s is a boundary, I_q is the actual pixel location in the image of the q_{th} data point we are searching, 2 is the number of centroid representing cells, and $length_{ref}$ is the reference length ($length_{ref} = 15$ usually).

When the distance is smaller, I_q is closer to the segment. Then, the point will have more possibility of being in the boundary. We incorporate this possibility into the membership. It means that smaller $M_{2q}(M_{2q} < 1)$ will be multiplied to the associated membership of the centroid representing cells:

$$(U_{2q})_{ij}^{*} = (U_{2q})_{ij}^{*} \cdot M_{2q}. \tag{17}$$

2.2. FOPSO combined with SI-IFCM

The proposed algorithm defines a set of cluster centroids in IFCM as a particle in FOPSO (the population of particles is usually 10~20). We define the group of particles as below:

$$\begin{aligned} &(P_p)_{ij}, i \in \{1,2,3...pop\},\ \mathrm{j} \in \{1,2,3...iter\}, \\ &p \in \{1,2,3...c\}. \end{aligned} \tag{18}$$

i is the number of the particle, j is the number of the iteration, p is the number of the centroid.

The velocity $v_{i(j+1)}$ and the position $P_{i(j+1)}$ of the i^{th} particle in $(j+1)^{th}$ iteration are updated as below [21]:

$$\begin{aligned} v_{i(j+1)} &= \lambda \cdot v_{ij} + \frac{1}{2}\lambda \cdot v_{i(j-1)} \\ &+ \frac{1}{6}\lambda \cdot (1-\lambda) \cdot v_{i(j-2)} \\ &+ \frac{1}{24}\lambda \cdot (1-\lambda) \cdot (2-\lambda) \cdot v_{i(j-3)} \\ &+ c_1 \cdot ((p_b)_i - P_{ij}) + c_2 \cdot (g_b - P_{ij}), \end{aligned} \tag{19}$$

$$(P)_{i(j+1)} = (P)_{ij} + v_{i(j+1)}, \tag{20}$$

where the personal best particle $(p_b)_i$ of i^{th} particle and the global best particle g_b in j^{th} iteration are selected by fitness:

$$fitness((p_b)_i) = \max(fitness(P_{ik})), k \in \{1,2...j\}, \tag{21}$$

$$\begin{aligned} &fitness(g_b) = \max(fitness(P_{ik})), \\ &i \in \{1,2...pop\}, k \in \{1,2...j\}. \end{aligned} \tag{22}$$

To further improve the performance for cell segmentation, the convergence speed information and local spatial information are introduced into the fitness function of FOPSO.

The fitness function is defined as below:

$$fitness_{ij} = \begin{cases} (\overline{obj}_j - obj_{ij} + 2\times\sigma_j) \cdot \exp(-num / num_{ref}) \\ if\ \ \overline{obj}_j - obj_{ij} + 2\times\sigma_j > 0, \\ 0 \\ \text{otherwise} \end{cases}, \tag{23}$$

where num is the number of possible intersections of cell contours, and num_{ref} is the reference number ($num = 10^6$ usually).

The first part of fitness function [23] is calculated with the average and the standard variance of object functions of IFCM that this particle ever had. It is aimed to select the particle with better convergence speed. If the first part is larger, the particle has faster convergence speed.

The second part is the part of local spatial information using the number of possible intersections of cell contours in temporary image. When the second part is smaller, the number is larger. It means the higher adhesion degree.

The alternate iterations and the different object functions of two algorithms ensure that our algorithm could search the solution in the directions of different senses and lead to the lower adhesion degree and higher clustering degree.

3. EXPERIMENTAL RESULTS

We compare the proposed algorithm with other classical or recent algorithms: marker-based watershed algorithm [5,7,8], two-step cell splitting algorithm [25], fuzzy local information c-means clustering algorithm (FLICM) [16], kernel weighted fuzzy local information c-means clustering algorithm (KWFLICM) [17], kernel intuitionistic fuzzy c-means clustering algorithm (KIFCM) [18] and conventional fuzzy c-means clustering algorithm (FCM) [13].

The marker-based watershed algorithm [5,7,8] uses the marker image, which is the eroded binary image with Otsu method [1], to find the contours using EDT. The two-step cell splitting algorithm [25] is based on the two-step binarization and clump splitting.

All experiments are performed on two image sets containing 56 images and 6374 cells with size 1034 × 1382 (type 1) or 512 × 512 (type 2) pixels on a PC (CPU: Inter Xeon E3-1231 v3 3.4 GHz, RAM: 8.0 GB). These images are fluorescence images from DNA stained HeLa cells. They were imaged on a Zeiss fluorescence microscope at 20× magnification. The sizes of the cells range between 2 and 10 μ m.

The sample images for comparison are marked as below:

a=original image, b=proposed algorithm, c=marker-based watershed algorithm, d=two-step splitting algorithm, e=KWFLICM, f=FLICM, g=KIFCM, h=FCM.

There are three touching cells and one single cell in comparison 1. The marker-based watershed algorithm segments only three cells with the biggest one segmented incompletely. The two-step algorithm could not split two touching cells below. FLICM has similar result with smaller correctly segmented area. KWFLICM has inaccurate boundary of two cells. KIFCM could identify and split these four cells. However, it has inaccurately segmented area, which leads to inaccurate shape of contours. Our proposed algorithm could identify and split these cells while reserving the most accurate shapes of contours. This indicates that our proposed algorithm could segment the bright cells when keeping the accuracy of the cell area. Because our algorithm is optimized for the segmentation of such area with vague boundary.

In comparison 2, there are two touching cells (one is much brighter than another one). The two-step algorithm, FLICM, KWFLICM and FCM could not split them. The marker-based watershed algorithm could split them with inaccurate area of brighter cell. Only our proposed algorithm and KIFCM algorithm could identify and split them. Because we use improved intuitionistic membership, which is intended to improve the performance on the image with inhomogeneity.

In comparison 3, there are two dim adhesive cells. FLICM and KWFLICM could not identify them. FCM is able to segment only small area of them. The marker-based watershed algorithm and two-step cell splitting algorithm could not split them. Our proposed algorithm could identify and split them. Our introduction of improved intuitionistic membership ensures the identification of dim cells. The local spatial information incorporated into the algorithm leads to better segmentation of adhesive cells. In comparison 4, it becomes three dim touching cells. Only our proposed algorithm and KIFCM could split them, while KIFCM only obtains the noise.

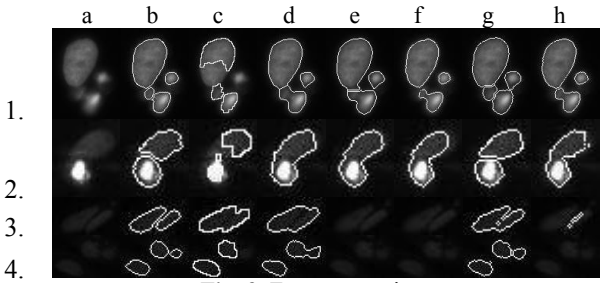

Fig. 3. Four comparisons

The statistical comparison is as below:

Measures / Algorithms	PCD (%)	PND (%)	PFA (%)
Proposed algorithm	**90.25**	**9.75**	**0.28**
Watershed	84.20	15.80	1.00
Two-step	86.70	13.30	0.50
KWFLICM	86.03	13.97	0.40
FLICM	86.18	13.82	0.62
KIFCM	88.40	11.60	0.90
FCM	87.76	12.24	0.79

Table. 1 We define the number of correctly detected cells as N_c , the number of missing cells as N_n , the number of wrongly detected cells as N_f and the total number of cells as N_{all} . $PCD = N_c / N_{all}$, $PND = N_n / N_{all}$ and $PFA = N_f / N_{all}$.

Our proposed algorithm owns the highest PCD and PND with the lowest PFA. It could segment the most accurate cells with fewest errors. The statistical results also demonstrate the advantage of our proposed algorithm.

4. CONSLUSION

We propose SI-IFCM with an improved intuitionistic membership and local spatial information to improve the performance on cell images with inhomogeneity, much noise, and adhesive cells. The alternate iteration of SI-IFCM and FOPSO with different object functions effectively avoids the local extreme. Sometimes we need simple postprocessing to improve the primary result with noise, voids and unusual boundaries. Then we perform the distance transform and watershed transform to finish the final segmentation.

Compared with other algorithms based on threshold, concavity detection and FCM, our algorithm has advantages in PCD, PND and PFA, while keeping considerable shapes of contours.

5. REFERENCE

[1] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, 11, 23-27, 1975.

[2] D. Marr, and E. Hildreth, "Theory of edge detection," *Proceedings of Royal Society of London. Series B: Biological Science*, 207 (1167), 187-217, 1980.

[3] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8 (6), 679-98, 1986.

[4] J. Sivakamasundari, *et al.*, "Proposal of a Content Based Retinal Image Retrieval system using Kirsch template based edge detection," *International Conference on Informatics, Electronics & Vision*, 1-5, 2014.

[5] F. Meyer, "Topographic distance and watershed lines," *Signal Processing*, 38, 113-125, 1994.

[6] H. Song, *et al.*, "Splitting touching cells based on concave-point and improved watershed algorithms," *Frontiers of Computer Science*, 8 (1), 156-162, 2014.

[7] A. Mouelhi, *et al.*, "Hybrid segmentation of breast cancer cell images using a new fuzzy active contour model and an enhanced watershed method," *International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 382-387, 2013.

[8] H. Zhao, and K.Z. Mao, "Adaptive successive erosion-based cell image segmentation for p53 immunohistochemistry in bladder inverted papilloma," *Proceedings of the IEEE Annual Conference on Engineering in Medicine and Biology*, pp. 6484-6487, 2005.

[9] D. Sotiris, *et al.*, "Accurate cell segmentation in microscopy images using membrane patterns," *Bioinformatics*, 30 (18), 2644-2651, 2014.

[10] M. Kass, *et al.*, "Snakes: Active contour models," *International Journal of Computer Vision*, 1, 321-331, 1987.

[11] V. Caselles, *et al.*, "Geodesic active contours," *International Journal of Computer Vision*, 22 (1), 6l-79, 1997.

[12] TF. Chan, and LA. Vese, "Active contour without edges," *IEEE Transactions on Image Processing.*, 10 (2), 266-277, 2001.

[13] J.C. Bezdek, *et al.*, "Review of MR image segmentation technique in pattern recognition," *Medical Physics*, 20 (4), 1033-1048, 1993.

[14] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, 20, 87-96, 1986.

[15] T. Chaira, "A novel intuitionistic fuzzy c means color clustering on human cell images," *Proceedings of IEEE, World congress on Nature and Biologically inspired computing*, 736-741, 2009.

[16] S. Krinidis, and V. Chatzis, "A Robust Fuzzy Local Information C-Means Clustering Algorithm," *IEEE Transactions on Image Processing*, 19 (5), 1328-1337, 2010.

[17] M. Gong, *et al.*, "Fuzzy C-Means Clustering with Local Information and Kernel Metric for Image Segmentation," *IEEE Transactions on Image Processing*, 22 (2), 573-584, 2013.

[18] K. Lin, "A Novel Evolutionary Kernel Intuitionistic Fuzzy C-Means Clustering Algorithm," *IEEE Transactions on Fuzzy Systems*, 22 (5), 1074-1087, 2014.

[19] J. Kennedy, and R.C. Eberhart, "Particle swarm optimization," *Proceedings of the IEEE International Conference on Neural Networks*, Vol. 4. IEEE, pp. 1942-1948, 1995.

[20] A. Banks, *et al.*, "A review of particle swarm optimization. ii: Hybridisation, combinatorial, multicriteria and constrained optimization, and indicative applications," *Natural Computing*, 7 (1), 109-124, 2008.

[21] EJS. Pires, *et al.*, "Particle swarm optimization with fractional-order velocity," *Journal on Nonlinear Dynamics*, 61, 295-301, 2010.

[22] H. Bustince, *et al.*, "Intuitionistic fuzzy generators: application to intuitionistic fuzzy complementation," *Fuzzy Sets and Systems*, 114, 485-504, 2000.

[23] W. Chen, *et al.*, "An Improved Genetic FCM Clustering Algorithm," *International Conference on Future Computer and Communication*, Vol. 1. IEEE, pp. 45-48, 2010.

[24] Q. Chen, *et al.*, "Watershed segmentation for binary images with different distance transforms," *IEEE International Workshop on Haptic*, 111-116, 2004.

[25] A. LaTorre, *et al.*, "Segmentation of neuronal nuclei based on clump splitting and a two-step binarization of images," *Expert Systems with Applications*, 40, 6521–6530, 2013.