Spot Extraction in Low-Contrast Images

TUAN D. PHAM
School of Computing and Information Technology
Griffith University, Nathan Campus, QLD 4111
AUSTRALIA

Abstract: Methods for extracting spots in images play an important role in the field of image analysis, with a particular emphasis on applications to biological images. An algorithm for extracting and isolating image spots which are subject to low-contrast or poor-image illumination is presented in this paper. The computational framework is based on the implementation of fuzzy c-means, fuzzy entropy, and the aspect-ratio criterion.

Key Words: Fuzzy c-means, image segmentation, spot images, fuzzy sharpening

1 Introduction

Non-trivial extraction of image spots can be considered as an image segmentation problem, which is one of the most difficult task in image processing [7]. In recent years, there are a number of methods developed for spot extraction such as the analysis of DNA microarray spots [1, 6, 8]. However, methods for extracting DNA microarray spots are designed to deal with spots having similar sizes and gridded structures. An associated method for spot extraction has been developed by Xu *et al.* [13], which is based on double thresholding and curve fitting to segment the images of skin cancer. This method is suitable for the segmentation of isolated spots, and its curve fitting technique can only approximate the spot areas.

In this paper we present a segmentation method for dealing with image spots which are size-variable, unstructurely located, and subject to low contrast. We discuss an effective implementation of the fuzzy c-means algorithm when its straight-forward application is uneffective for the particular problem under study. The strategy for the segmentation process includes the selection of the number of clusters, sharpening of fuzzy sets, and spot isolation. We test our proposed segmentation algorithm for extracting spots with real image data and compare the results with those obtained by other standard segmentation methods as well as a current medical image analysis software for spot extraction.

2 Implementation

In this section we presents implementation of the proposed approach for spot extraction. After a brief descrition of the computational procedure of the fuzzy c-means, the discussion for the selection of clusters is addressed. We then seek to avoid the problem of over segmentation by applying the concept of fuzzy entropy. Finally, a simple strategy for isolating touching spots is described by considering the aspect ratio of the spots.

2.1 Fuzzy c-means algorithm

As a clustering method, the fuzzy c-means (FCM) algorithm [2] seeks to partition a dataset $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M\}$, where $\mathbf{x}_m = (x_{m1}, x_{m2}, \ldots, x_{mk})$, $m = 1, 2 \ldots, M$, into a specified number of fuzzy regions which

are represented by the corresponding cluster centers. The degrees of each $\mathbf{x}_m$ that belongs to different clusters are characterized by the corresponding fuzzy membership grades taking real values between 0 and 1.

In principle, the FCM maximizes the following objective function:

$$J(\mathbf{U}, \mathbf{c}_1, \ldots, \mathbf{c}_N) = \sum_{y=1}^{N} \sum_{m=1}^{M} \mu_{ym}^{\alpha} \, d_{ym}^2 \quad (1)$$

where M is the number of data points, N is the number of clusters, $\mathbf{U}$ is the $N \times M$ fuzzy membership matrix, $\mu_{ym} \in [0, 1]$ is the fuzzy membership grade that indicates the degree $\mathbf{x}_m$ belongs to the fuzzy region y, d_{ym} is a distance measure between cluster center $\mathbf{c}_y$ and data point $\mathbf{x}_m$, and $\alpha \in [1, \infty)$ is the fuzzy exponential weight.

The computations of the cluster centers and the partition matrix $\mathbf{U}$ are updated by an iterative procedure which is described as follows.

1. Given the degree of fuzziness α and initial membership matrix $\mathbf{U}$ with random values of $\mu_{ym} \in [0, 1]$ subjected to

$$\sum_{y=1}^{N} \mu_{ym} = 1, \forall m = 1, \ldots, M$$

2. Update initial cluster centers

$$\mathbf{c}_y^{j+1} = \frac{\sum_{m=1}^{M} \mu_{ym}^{\alpha} \mathbf{x}_m}{\sum_{m=1}^{M} \mu_{ym}^{\alpha}} \quad (2)$$

3. Update fuzzy membership functions

$$\mu_{ym} = \frac{1}{\sum_{z=1}^{N} \left(\frac{d_{ym}}{d_{zm}} \right)^{2/(\alpha-1)}} \quad (3)$$

where, using the L_2 norm, d_{ym} is given by

$$d_{ym} = ||\mathbf{x}_m - \mathbf{c}_y||_2$$

4. Compute the objective function according to (1). If it converges or its improvement over the previous iteration is below a certain threshold then stop the iterative process. Otherwise, go to step 2.

Figure 1. Original image A

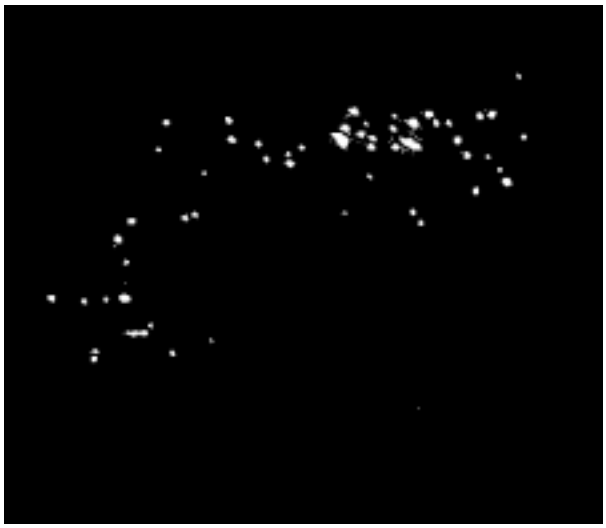

Figure 2. Segmentation of image A by Otsu's thresholding

2.2 Estimating the number of clusters

Taking a first look at the image (412×357) as shown in Figure 1, there seems to be two classes to be segmented. These two classes are the background pixels and the peroxisome pixels. If we apply the well-known Otsu's thresholding method [9] and the FCM, with the number of clusters N=2, to segment the gray image of Figure 1, we obtain Figures 2 and 3 which are the results given by Otsu method and the FCM respectively. It can be seen that both results overestimate the spot sizes and highlight

noise and outliers. These are due to the low contrast of the image and particularly the flourescent stains around the peroxisome spots. We therefore need to add another cluster to represent the flourescent-shadow pixels, i.e., the number of classes is now three instead of two. Because Otsu method only works for gray-scale images with two classes, we now apply the FCM with N=3 and obtain another result as shown in Figure 4. This result shows some improvement over that obtained by the FCM with N=2. However, overestimation of spot areas and touching spots still remain at some extent. We will tackle these problems by a strategy for sharpening the fuzziness of the peroxisome cluster, an aspect-ratio criterion, and quadtree decomposition, which are presented in the following subsections.

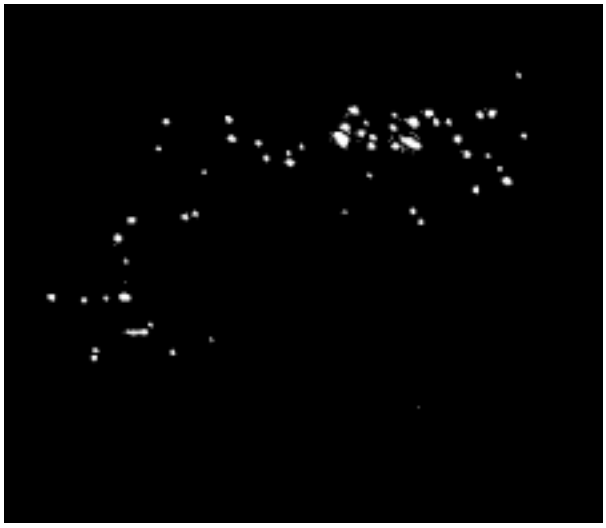

Figure 3. Segmentation of image A by FCM with two clusters

2.3 Sharpening fuzzy regions

Based on the concept of a fuzzy set [14] and the notion of the Shannon's entropy [12], the measure of fuzziness of a fuzzy set was initially defined by [4] as follows:

1. The fuzziness of $A = 0$ if A is a crisp set, that is, $\mu_A(x) \in \{0, 1\}$, $\forall x \in X$.

2. The fuzziness of A is maximum when $\mu_A(x) = 0.5$, $\forall x \in X$.

3. The fuzziness of A is greater than or equal to that of A^* if A^* is a *sharpened* version of A, that is, $\mu^*_A(x) \geq \mu_A(x)$ if $\mu_A(x) \geq 0.5$; and $\mu^*_A(x) \leq \mu_A(x)$ if $\mu_A(x) \leq 0.5$.

Let $\mu_P(\mathbf{x})$ be the fuzzy membership grade that indicates how possible a pixel $\mathbf{x}$ belongs to the set containing all the peroxisome images, we then apply the notion of the measure of fuzziness to sharpen the fuzzy region of interest (peroxisome) by defining

$$\mu^*_P(\mathbf{x}) = \begin{cases} 1 & : \quad \mu_P(\mathbf{x}) \geq \delta_\mu \\ 0 & : \quad \mu_P(\mathbf{x}) < \delta_\mu \end{cases} \tag{4}$$

where $0.5 < \delta_\mu < 1$ is a fuzzy membership threshold.

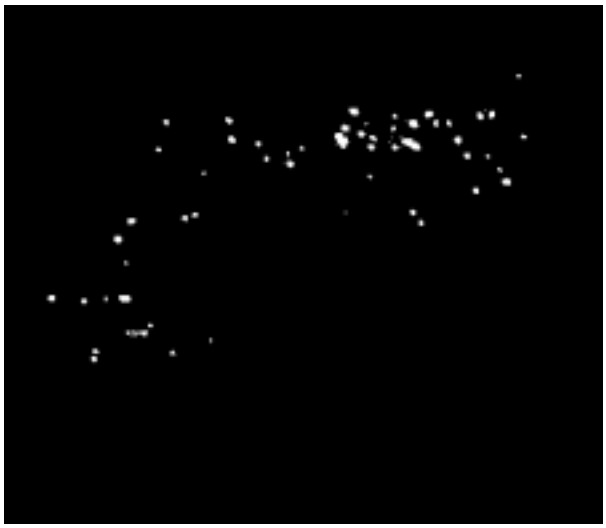

Figure 4. Segmentation of image A by FCM with three clusters

Figure 5. Segmentation of image A by sharpening FCM with three clusters

What we discuss next is how to get an appropriate value for δ in oder to obtain good sharpened peroxisome spots which can make the task of isolating touching spots easier. To fix a concrete idea, let $\mu_{\mathbf{c}^*}(\mathbf{x})$ be the fuzzy membership grade of a pixel $\mathbf{x}$ belonging to the peroxisome cluster $\mathbf{c}^*$. We can say that an optimal value of $\mathbf{c}^*$ must be some value between the *least*, denoted by $f_{\min}(\mathbf{x}|\mathbf{c}^*)$, and the *most*, denoted by $f_{\max}(\mathbf{x}|\mathbf{c}^*)$, bright intensities which are to be assigned to $\mathbf{c}^*$. Of course, it is difficult to determine $f_{\min}(\mathbf{x}|\mathbf{c}^*)$ straigth away; however, finding $f_{\max}(\mathbf{x}|\mathbf{c}^*)$ is immediately available, that is, by checking the membership grade of the brightest pixel of the whole image assigned to $\mathbf{c}^*$ given by the FCM. We therefore select $\delta = \mu_{\mathbf{c}^*}(\mathbf{x}^*)$, where $f(\mathbf{x}^*)$ is the maximum intensity value, because $\mu_{\mathbf{c}^*}(\mathbf{x}^*)$ represents the brightest and the least bright pixels which are to be assigned to $\mathbf{c}^*$. Finally, each segmented peroxisome region will be filled up in case if there are any holes in the region. This is because there exist some low-intensity pixels within the regions.

Figure 5 shows an improved segmentation version, in comparison with the result as shown in Figure 4, by applying the sharpening procedure defined in (4) – the segmented spot areas are sharpend and closer to the real spot areas than the former segmented results; in addition, more outliers are also removed in this sharpened version.

2.4 Isolating touching spots

We define an aspect ratio of a spot image p, based on which touching spots can be isolated, as

$$r(p) = \frac{w_{\min}(p)}{w_{\max}(p)}$$

where $w_{\min}(p)$, and $w_{\max}(p)$ are the minimum and maximum widths of the spot area, and $w_{\min}(p) \geq$ the maximum width of the estimated smallest spot size.

The procedure for splitting touching spots is described as follows.

1. Given a spot image p^i, $i = 1, \ldots I$, where I is the number of segmented spots which are greater than an estimated smallest spot image.

2. If $r(p^i) < 0.5$, then split p^i into two subimages p_1^i and p_2^i at the location of $w_{\min}(p^i)$.

   (a) If p_g^i, $g = 1, 2$, is greater than an estimated smallest spot size and $r(p_g^i) < 0.5$, then separate p_j^i into two subimages $p_{g,1}^i$ and $p_{g,2}^i$ at the location of $w_{\min}(p_g^i)$.

   (b) Repeat step (a) for all subimages $p_{g,\ldots,G}^i$ where each subscript takes the values from 1 to 2.

3. Repeat steps (1) and (2) for all p^i.

3 Experimental Results

In addition to the illustrations, which have been presented in the foregoing sections, showing some advantages of our FCM-based segmentation approach, we further test our proposed method, where $\alpha = 2$ for all cases, for extracting peroxisome image spots on several real images and also compare with other methods for image spot extraction.

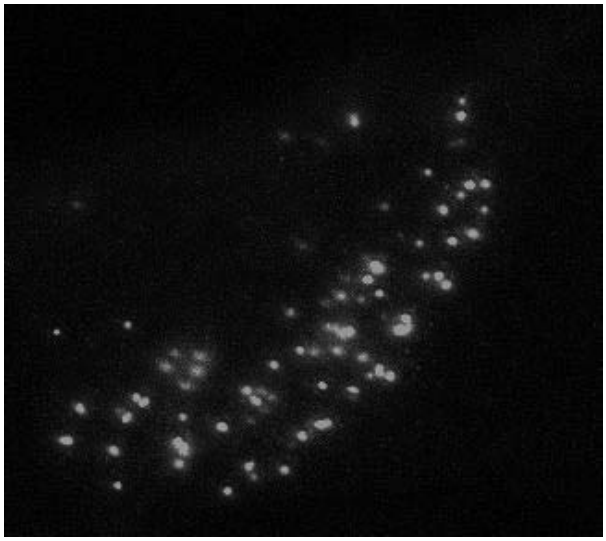

Figure 6. Original image B

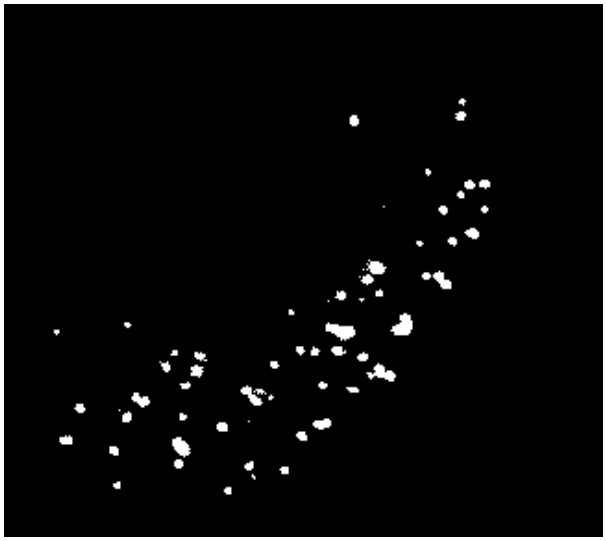

Figure 8. Segmentation of image B by FCM with three clusters

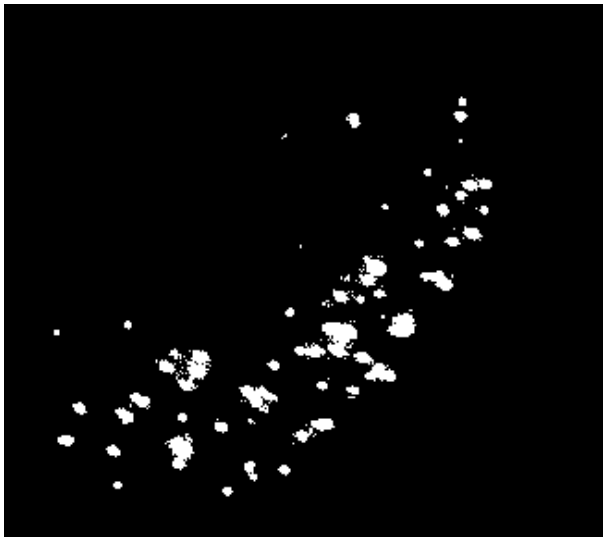

Figure 7. Segmentation of image B by Otsu's thresholding

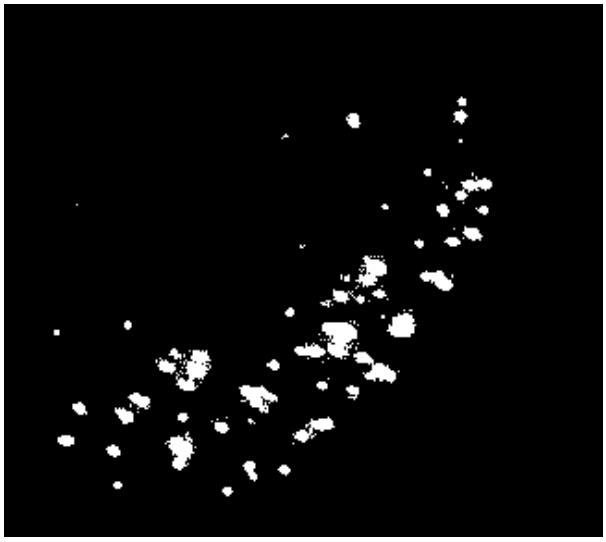

Figure 9. Segmentation of image B by ImageJ (iterative thresholding)

Figure 6 shows the intensity version of an RGB colour image (412×357) that contains flourescent-stained peroxisome spots [3]. Edges of these spots are fuzzy due to low constrast, also some of the spots are connected to each other. Some flourescent stains may misrepresent spots (false spots) for simple segmentation methods. Figures 7-10 shows the segmented versions using Otsu thresholding method, FCM with three clusters, ImageJ (http://rsb.info.nih.gov/ij/), and our proposed FCM-based segmentation method. It can be seen from these figures that the results obtained from both Otsu's thresholding, straight-forward FCM, and ImageJ that uses a thresholding method developed by Ridler and Calvard [11], show false as well as overestimated peroxisome spots; whereas our proposed method yields the segmentation results that are quite close to the actual spots and can also isolate touching spots.

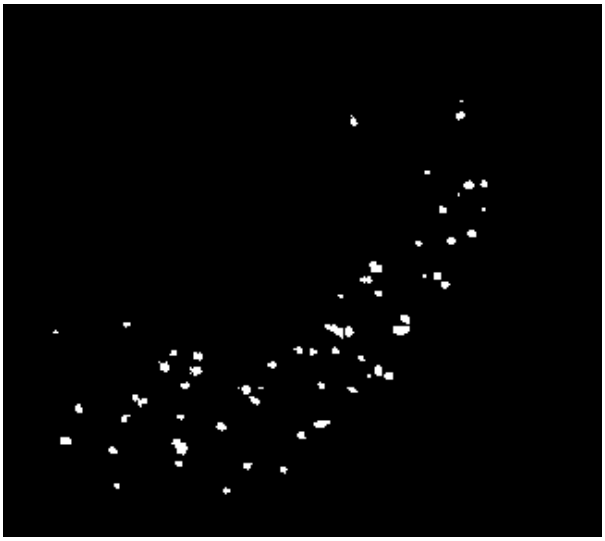

Figure 10. Segmentation of image B by proposed FCM-based method

4 Conclusion

We have presented an effective algorithm for extracting spots in fuzzy images where the contrast is low and spots are touching, which make standard techniques for image segmentation or edge detection ineffective. We have tested our proposed FCM-based algorithm with real image data and obtained favourable results in all cases in comparison with other methods. We also conclude that appropriate implementations of the fuzzy c-means algorithms and the theory of fuzzy sets are very useful for dealing with biological and medical data [5, 10].

References

[1] Angulo, J., and Serra, J. (2003) Automatic analysis of DNA microarray images using mathematical morphology. *Bioinformatics*, **19**, 553-562.

[2] Bezdek, J.C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.

[3] Chang, C. C., South, S., Warren, D., Jones, J., Moser, A.B., and Moser, H.W. (1999) Metabolic control of peroxisome abundance. *Journal of Cell Science*, **112**, 1579-1590.

[4] DeLuca, A. and Termini, S. (1972) A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory. *Information and Control*, **20**, 301-312.

[5] Dembele, D., and Kastner, P. (2003) Fuzzy c-means method for clustering microarray data. *Bioinformatics*, **19**, 973-980.

[6] Glasbey, C.A., and Ghazal, P. (2003) Combinatorial image analysis of DNA microarray features. *Bioinformatics*, **19**, 194-203.

[7] Gonzalez, R.C., and Woods, R.E. (2002) *Digital Image Processing*. Prentice-Hall, New Jersey.

[8] Liew, A.W.C., Yan, H., and Yang, M. (2003) Robust adaptive spot segmentation of DNA microarray images. *Pattern Recognition*, **36**, 1251-1254.

[9] Otsu, N. (1979) A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man, and Cybernetics*, **9**, 62-66.

[10] Chi, Z., Yan, H., and Pham, T. (1996) *Fuzzy Algorithms: With Applications to Image Processing and Pattern Recognition*. World Scientific Publishing, Singapore.

[11] Ridler, T.W. and Calvard, S. (1978) Picture thresholding using an iterative selection method. *IEEE Trans. Systems, Man, and Cybernetics*, **8**, 630-632.

[12] Shannon, C., and Weaver, W. (1948) The mathematical theory of communication. *The Bell System Technical Journal*, **27**, 379-423, 623-656.

[13] Xu, L., Jackowski, M., Goshtasby, Roseman, D., Bines, S., Yu, C., Dhawan, A., and Huntley, A. (1999) Segmentaion of skin cancer images. *Image and Vision Computing*, **17**, 65-74.

[14] Zadeh, L.A. (1965) Fuzzy sets. *Information and Control*, **8**, 338-353.