Segmentation and Classification of 3-D Spots in FISH Images

Sundaresh Ram, Jeffrey J. Rodriguez
Dept. of Electrical and Computer Engineering
The University of Arizona
Tucson, AZ, USA

Giovanni Bosco
Dept. of Molecular and Cellular Biology
The University of Arizona
Tucson, AZ, USA

Abstract—In this paper, we propose an automated method to segment and classify the 3-D spots in fluorescence in-situ hybridization images from ovarian germine nurse cells of Drosophila melanogaster. The spot segmentation consists of a smoothing step followed by top-hat filtering and 3-D region growing. After the spots are segmented, a number of features such as volume, texture, and contrast are measured so as to classify the real spots from the noise with the help of a Bayesian classifier. Experimental results demonstrate the effectiveness of the proposed scheme in terms of segmentation and classification accuracy.

Keywords-FISH images, spot detection, segmentation, top-hat filtering, anisotropic diffusion, ROC analysis.

I. INTRODUCTION

Selective staining of DNA sequences using fluorescent in situ hybridization (FISH) allows for detection, analysis and quantification of specific sequences within the nucleus [12]. The use of probes that fluoresce at different wavelengths when imaged using confocal microscopy allows for simultaneous detection of multiple DNA sequences, which tend to appear as localized spots that are spheroidal in shape within a single nucleus. Detection and segmentation of these spots have been crucial for spatial analysis and quantification of numerical and structural abnormalities. In the presence of noise, the edges of the spots appear very fuzzy to a human observer, and since a single spot spans across multiple images, manual segmentation and counting of these spots is extremely difficult, fatiguing and error prone, requiring automated image analysis techniques.

Several techniques have been proposed in the past for segmentation and classification of biological images. Lerner et al. [7] propose a scheme to segment the spots based on a multiresolution image pyramid and Ostu’s thresholding. This method is not well suited for very small spots, and thus at a low resolution image in the pyramid they may not be detected. Active contours [14] have been used extensively for segmentation, but suffer from the need for initialization. People have used top-hat filtering and template matching [10], which are good for localization of the object but not for quantitative measurements and suffer in the presence of noise. Model-based methods have also been proposed for segmentation [8], which concentrate on segmentation of large regions of interest and are not well suited for small regions such as our spots. Some have also attempted to segment spots using signal enhancement and thresholding techniques [3], which do not segment the exact boundaries of the objects of interest.

In this paper we propose an automated technique to segment the 3-D spots using a smoothing step and top-hat filtering to remove the effect of noise and enhance the signal, followed by 3-D region growing. After the segmentation, a number of features such as volume, contrast, and texture are measured so as to classify real spots from noise using a Bayesian classifier.

II. SPOT SEGMENTATION

A. Smoothing

We formulate smoothing as a diffusive process as in [2], which is suppressed or stopped at boundaries by selecting locally adaptive diffusion strengths. The process can be formulated mathematically as

\[ \frac{\partial}{\partial t} I(x,y,z,t) = \text{div}(c(x,y,z,t)\nabla I(x,y,z,t)) \] (1)

The diffusion strength is controlled by the diffusion function \( c(x,y,z,t) \), which depends on the magnitude of the gradient of the image intensity \( I(x,y,z,t) \). The variable \( t \) is the process ordering parameter, and in discrete implementations it is used to enumerate iteration steps. We use the diffusion function proposed by [2] given below

\[ c(x,y,z,t) = \frac{1}{1 + \left( \frac{\nabla I(x,y,z,t)}{\kappa} \right)^\alpha} \quad |\alpha > 0 \] (2)

The parameter \( \kappa \) is chosen according to the noise level and the edge strength. A proper choice of diffusion function not only preserves, but also enhances edges while being numerically stable. In our experiments we saw a significant noise reduction in the images after twenty iterations with \( \kappa = 8 \) and \( \alpha = 3/44 \). This technique mainly diffuses within regions and does not affect region boundaries at locations of high gradients.

B. Top-Hat Filtering

The top-hat filter is a signal enhancing filter. It uses the opening operation from mathematical morphology [4, 13]. After smoothing the whole data set as described above we calculate the gray-scale 3-D top-hat transformation using

\[ g = I - \Gamma_g(I) \] (3)

where \( \Gamma_g(I) \) is the gray-scale opening by flat structuring element \( B \), which is taken to be a 5×5×3 ellipsoidal neighborhood. Note: the structuring element should be sized according to the expected spot size. The window size for the filter spans more pixels in x and y (intra-frame) than in z dimension since the intra-frame resolution is greater than the inter-frame resolution.
C. Region Growing

Region growing is one of the widely explored region-based image segmentation methods. It can also be classified as a pixel-based image segmentation technique as it requires the selection of initial seed pixels. In our method we select initial seed regions of the spots by performing histogram thresholding on the top-hat filtered images. The gray-level histogram of the top-hat output is unimodal, with a steep monotonic increase until the peak is reached, followed by a Gaussian decay. We apply two methods [9,11] described for unimodal thresholding of the histogram of the top-hat output of the 3-D data set and take the average of the two methods as the single threshold for the whole data set. The threshold could be changed so as to get the desired false positive rate.

Since we need to grow multiple regions in 3-D – one for each spot – we need to connect the foreground pixels obtained from the thresholding step to determine the seed region for each 3-D spot. This is done using 3-D connected component labeling [4].

Next, we apply region growing to each 3-D seed region, using similarity and discontinuity measures proposed by [5]. We start to grow the object region by adding one pixel to the region at a time. All of the pixels that neighbor the initial seed region are examined in terms of their gray level, and the pixel that has the highest gray level is added to the current region. This induces a directional growing such that the pixels of highest gray level will be absorbed first. When several pixels having the same gray level become candidates to join the region, we employ the first-come first-served strategy to choose one of them and add it to the region.

The growing process is controlled by making use of two different properties for each region, called the “average contrast and peripheral contrast” [5]. We define the outer boundary (OB) of a 3-D region as the set of exterior pixels that neighbor an interior pixel of the region. The average contrast measure for a 3-D region is defined as the difference between the mean of the region and the mean of the OB pixels. As long as the growth of the region does not exceed the true 3-D boundary of the spot, the rate of decrease of the mean of OB pixels is greater than the rate of decrease of the mean of the region. Once the region starts to grow outside the spot into the background, the rate of decrease of the mean of OB pixels is less than that of the mean of the region. Thus, the average contrast measure is a curve that initially increases and then decreases, and the maximum of this curve is the point where the process starts to grow into the background. The inner boundary (IB) of a 3-D region is defined as the interior pixels of the region that neighbor an exterior pixel. The peripheral contrast measure for a region is defined as the difference between the mean of the IB pixels and the mean of the OB pixels. The peripheral contrast measure is similar to the average gradient magnitude of the IB and/or OB pixels of the evolving region. The reason for choosing such a measure instead of the actual gradient magnitude is because it is less sensitive to noise. In case of a relatively homogeneous region, the peripheral contrast measure has a global maximum, but for a noisy and textured region the peripheral contrast measure exhibits many local peaks. We select the last local maximum of the peripheral contrast occurring just before the maximum of the average contrast to determine the boundary of the 3-D segmented region [5]. Fig. 1 shows images at the various stages of the segmentation algorithm.

III. SPOT CLASSIFICATION

It is likely that many valid spots have intensity much lower than that of some false signals, e.g. false spots due to accidental or non-specific staining. Therefore, the best threshold choice is not enough to isolate all of the true spot signals from false ones using only the intensity information. The thresholding of the top-hat output yields not only the low intensity spots but also the higher intensity false spots. Applying 3-D region growing to the seed regions leads to false positives in the images. Thus, a classification step is required in order to remove any false spots.

A. Bayesian Classifier

Given a set of variables, \( X = \{x_1, x_2, \ldots, x_d\} \), we construct the set of posterior probability densities for the class \( C_j \) among the set of possible classes \( C = \{c_1, c_2, \ldots, c_d\} \). Using Bayes rule,

\[
p(C_j | x_1, x_2, \ldots, x_d) \propto p(x_1, x_2, \ldots, x_d | C_j) p(C_j)
\]

The naïve Bayesian classifier assumes that the variables are statistically independent, and decomposes the likelihood to a product of terms:

\[
p(X | C_j) \propto \prod_{k=1}^{d} p(x_k | C_j)
\]

The posterior probability is rewritten as:

\[
p(C_j | X) \propto p(C_j) \prod_{k=1}^{d} p(x_k | C_j)
\]

We label a new case \( X \) with a class level \( C_j \) that maximizes

![Fig. 1. Single slice of a 3-D data set containing spots. (a) Original image. (b) Image after smoothing by diffusion process. (c) Image after top-hat transform. (d) Image after thresholding. (e) Image after 3-D region growing. (f) Detected boundaries overlaid on the original image.](image-url)
the posterior probability in (5). In our work we classify the objects into two classes, the spots and the background. Although the assumption that the variables $X$ are always independent is not always accurate, it simplifies the classification task, since it allows the class conditional densities $p(x|C_i)$ to be calculated separately for each variable by reducing the multidimensional task to many one-dimensional ones.

As in earlier work [6], we assume that the data is generated by a parametric model based on a single Gaussian as it has convenient analytical and statistical properties and it is suitable for representing measurements of many natural phenomena. The mean and standard deviation of the Gaussian density are estimated using the maximum likelihood procedure.

B. Features

To isolate the true valid spots from the false ones we use the Bayesian approach described above. The features used for classification must be carefully chosen to fully characterize the objects in question, i.e. the true spots, and must be capable of discriminating them from the false spots. In addition, the features must be general enough to account for object diversity and image variability. The set of features used for this study are summarized as follows:

1) **Average Intensity ($I$):** Average intensity of all the voxels in the object.
2) **Volume ($V$):** Total number of voxels present in the object multiplied by the resolution of each voxel in $x,y,$ and $z$-directions.
3) **Shape Factor ($S$):** Defined as
   \[ S = \frac{|b|}{64\pi V^2} \] (7)
   where $b$ is the surface area of the object.
4) **Standard Deviation ($\sigma$):** The usual standard deviation of all the voxels in the object.
5) **3-D Contrast ($C$):** Defined as
   \[ C = \frac{\mu_c - \mu_d}{\mu_c + \mu_d} \] (8)
   where $\mu_c$ is the average intensity of the eroded 3-D spot using a $2\times2\times1$ structuring element, and $\mu_d$ is the average intensity of the dilated 3-D spot using a $6\times6\times2$ structuring element. Note: the structuring element should be sized according to the expected spot size.
6) **Convexity ($\zeta$):** The ratio of the object volume to the volume of the convex hull of the object.
7) **Edge Gradient Magnitude ($\upsilon$):** The gradient magnitude of the edge voxels of the object, is calculated by using the first-difference operator to compute the derivatives in the $x, y,$ and $z$ directions. A histogram with four bins is formed and the maximum of difference between two bins is taken as edge gradient magnitude [1].

8) **Average Gradient Magnitude ($G$):** The average gradient magnitude of all the voxels in the object, where we use the first-difference operator to compute the derivatives in the $x, y,$ and $z$ directions.

IV. RESULTS

The images were acquired using the Zeiss LSM 510 Meta confocal microscope. The intra-slice sampling interval is 0.31 $\mu$m and the inter-slice sampling interval is 1 $\mu$m. The values of all the parameters used in the algorithm were determined by varying only one parameter at a time while keeping others fixed and setting the parameter to a value that maximizes the area under the ROC curve. We tested our method on 3 data sets, with each data set containing 3 channels, one for each FISH probe. A total of 103 true spots (34 channel 1 spots, 37 channel 2 spots and 32 channel 3 spots) were manually segmented and used for training and testing the segmentation and classification accuracy of the proposed method. We compare the proposed method with two other methods described in [7] and [10].

A. Segmentation

For performance evaluation, we used Dice evaluation [13] based on the amount of mutual overlap between the proposed segmentation method and manual segmentation. The performance metric for comparing the accuracy of the algorithms is given by

\[ O = \frac{\#(R_1 \cap R_2)}{\#(R_1) + \#(R_2)} \] (9)

where $R_1$ and $R_2$ are the automatically segmented region and the manually segmented region, respectively. Here, $\#(\cdot)$ represents the number of voxels in the region. Fig. 2 shows an example of the results obtained for the three methods. Table 1 (sub-column 1) shows the segmentation accuracy as a percentage for various data sets.

B. Classification

We evaluated the accuracy of the classification algorithm using receiver operating characteristic (ROC) analysis. We trained one data set at a time, estimated the

![Fig. 2. Detected contours of a single slice overlaid on the original image. (a) Manual segmentation. (b) Lerner’s method. (c) Raimondo’s method. (d) Proposed method. (e) Zoomed view of the spots with all methods’ detected boundaries overlaid.](image-url)
analysis. We trained one data set at a time, estimated the features from this data set and tested them on the remaining two data sets. The false positive rate is defined as the ratio between the number of detected spots not present in the manual segmentation over the total number of spots present in the manual segmentation. Similarly, the true positive rate is defined as the ratio between the number of correctly detected spots over the total number of spots present in the manual segmentation. ROC curves were constructed by computing the true positive rate and false positive rate for 256 different thresholds. Fig. 3 shows the ROC curves when data set 2 is taken as the training data. Table 1 (sub-column 2) shows the detection accuracy for the various data sets for the three methods.

V. CONCLUSION

We have proposed an automated scheme for the segmentation and classification of 3-D spots in FISH images. Our method combines anisotropic smoothing, top-hat filtering and region growing along with Bayesian classification using features such as volume, texture, and contrast. A false positive rate of 3% was the acceptable level for the genetic application under consideration. The average segmentation accuracy of the proposed method is 7.9 percentage points greater than that of Lerner’s method and 13.1 percentage points greater than that of Raimondo’s method. At a 3% false positive rate, the true positive rate of the proposed algorithm was 7.7 percentage points greater than that of Raimondo’s method.

REFERENCES


TABLE I. PERCENTAGE SEGMENTATION ACCURACY (SUB-COLUMN 1) AND CLASSIFICATION ACCURACY (SUB-COLUMN 2)

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Proposed Method</th>
<th>Lerner’s Method</th>
<th>Raimondo’s Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>90.20</td>
<td>92.66</td>
<td>84.87</td>
</tr>
<tr>
<td>2</td>
<td>89.63</td>
<td>95.20</td>
<td>86.05</td>
</tr>
<tr>
<td>3</td>
<td>91.40</td>
<td>90.17</td>
<td>82.91</td>
</tr>
</tbody>
</table>