A segmentation method for nuclei identification from sagittal images of *Drosophila melanogaster* embryos

Daniela Justiniano de Sousa
Faculdade de Computação
Universidade Federal de Uberlândia
daniela@mestrado.ufu.br

Maira Arruda Cardoso
Instituto de Biofísica Carlos Chagas Filho
Universidade Federal do Rio de Janeiro
mairaarrudacardoso@gmail.com

Paulo Mascarello Bisch
Instituto de Biofísica Carlos Chagas Filho
Universidade Federal do Rio de Janeiro
pmbisch@biof.ufrj.br

Francisco José Pereira Lopes
Instituto de Biofísica Carlos Chagas Filho
Universidade Federal do Rio de Janeiro, Polo de Xerém and Inmetro - National Institute of Metrology, Quality and Technology
flopes@ufrj.br

Bruno Augusto Nassif Travençolo
Faculdade de Computação
Universidade Federal de Uberlândia
travencolo@gmail.com

ABSTRACT

This paper proposes a segmentation method for sagittal images obtained from *Drosophila melanogaster* embryos at early stages of development. The proposed method operates in the spatial domain and uses traditional filters and segmentation techniques for image processing. However, one common problem encountered after the segmentation of sagittal images is the merged nuclei. In order to split these nuclei, our approach uses the curvature of the shape of the merged nuclei to find the regions of division between the nuclei. The proposed method was compared to other techniques and it achieved better results. In general, the obtained results have shown good performance and were able to identify individual nuclei, providing an efficient and accurate solution for segmentation of nuclei in images obtained from the sagittal plane of *Drosophila* embryos.

Keywords

Image segmentation, Curvature, Nuclei segmentation, *Drosophila melanogaster*.

1 INTRODUCTION

The segmentation is an important area of study for image analysis and computational vision. The term image segmentation refers to the partition of an image into a set of regions that cover it, aiming identifying or extracting regions of interest, in other words, all relevant semantic content to the application. Regions of interest should be uniform and homogeneous with respect to some characteristic, such as gray level, color, or texture. The segmentation step is usually considered critical, since, through it is possible to identify, recognize and classify digital objects. In general, it represents one of the most important steps, considering that its result may to determine the success or failure of computerized analysis procedures [Shap 01, Gonz 08].

Image segmentation performs a crucial role in bioimage informatics [Peng 08, Zhan 12]. In some cases, the complexity or particular characteristics of biological data hampers the process of segmentation, thus leading to the development of a wide range of segmentation methods addressing specific problems in biological applications. Normally, such methods make use of prior knowledge for the particular objects of interest and other possible structures in the image. In this context, we propose a new segmentation method for application to nuclei segmentation from sagittal images obtained from early *Drosophila melanogaster* embryos (e.g., Fig. 1(a-d)). The proposed method is based on analysis of curvature of the shape of the nuclei and operates in the spatial domain. In this case, aiming to achieve the desired segmentation it uses the knowledge about the geometric characteristics of objects contained in this type of image.

The *Drosophila melanogaster*, commonly known as the "fruit fly", is an important model in biological research, particularly in the study of gene expression regulatory networks, spatial expression pattern and cellular differentiation [Crow 12, Gilb 03]. Recent advances in

molecular biology and microscopy techniques substantially increased the amount of images generated for investigations on *Drosophila* development, especially images of gene expression patterns. In order to extract information from these images, several computational methods has been proposed [Pous 04, Pan 06, Jans 05, Rbel 06, Pisa 09, Kozl 08], followed by an increasing number of databases of images available on Internet as, for example, Flyex [Pisa 09], BDGP [Toma 07] and BDTNP [Fowl 08]. A common feature observed in the images in those databases is that the data (anatomical or gene expression patterns) are obtained from the surface of the embryo (or 3D in the case of DBTNP). None of these works deals with images obtained from sagittal planes passing through the embryo.

Sagittal images currently gained great interest to *Drosophila* community. By allowing the visualization of cells and tissues, localization of genes and proteins, as well the visualization of the dynamics of gene expression in embryos *in vivo* [Greg 07], this type of image can unveil new gene interactions and new findings about the relationship between anatomy and gene expression during development. Therefore, the image segmentation method proposed in this paper refers to sagittal images of early *Drosophila* embryos. More specifically, we are interesting in the identification of nuclei of the embryos, i.e., to define nuclei masks. This type of analysis increases the amount of features that can be extract from the images, as the obtained masks can be used to characterize several morphological properties, as well serve of basis to characterize gene expression patterns. For example, the nuclear mask can be used to quantify data on gene expression in each embryo nucleus, or to characterize the amount of gene products inside the embryo nuclei [Pisa 09].

In order to obtain the nuclear mask, the proposed method add a post-processing step on the result of Ostu segmentation method [Otsu 79]. This extra step is necessary because the result of initial segmentation results in many blocks of merged nuclei, which should be separated in order to obtain a correct identification of the nuclei. For separation of merged nuclei, we propose an analysis of the curvature of the shape of the merged blocks. In order to evaluate the efficiency of the method, it was performed several experiments with real images, which were collected by using confocal microscopy. The obtained results show efficiency and good performance in the identification of cell nuclei contained in the images processed.

2 PREVIOUS WORKS

Most of the works involving *Drosophila* embryos refer to the surface of the embryos [Pous 04, Jans 05, Pisa 09, Surk 08, Jaeg 04] and the nuclear mask is usually obtained through of technique proposed by Kosman [Kosm 99]. More sophisticate techniques deals with 3D data [Huan 08, Luen 06]. However, the content of both 3D data and surface images are different from sagittal images.

In addition, to the best of our knowledge, it has not been proposed any works related to the sagittal image of *Drosophila* embryos. Few works deal with computational methods in order to extract relevant information from sagittal images. Houchmandzadeh et al. [Houc 02] used a sliding rectangle of fixed size perpendicular to the edges of the embryo while probing the average pixel intensity under the rectangle, while Gregor et al. [Greg 07] used the same approach, but a circular mask were used instead of a rectangle. However, none of these approaches can identify individual nuclei.

Other works not directly related to *Drosophila* image could be used for sagittal images, as the works by Costa et al. [Ferr 97], Malpica et al. [Malp 97] and Bala [Bala 12]. In this paper we analysed the results of these three methods, along with the method of Kosman [Kosm 99], in order to compare to the results of our method. All of these proposals address variations of watershed segmentation algorithm [Vinc 91, Beuc 79], which is considered a powerful technique to isolate and recognize cell nuclei in biological images [Chen 12, Chan 12].

3 METHODOLOGY

Image segmentation consists in the extraction and identification of regions of interest contained in a image [Gonz 08]. In the case of sagittal images of early *Drosophila melanogaster* embryo, the regions of interest are those concerning to the cell nuclei. In this case, the objective of the segmentation is to obtain a nuclear mask image, i.e., a binary image where the pixels with value equal to 1 represent cell nuclei and the pixels with a value 0 represents the background of the image [Pisa 09]. For each embryo, a binary nuclear mask must to be constructed. This mask shows where the individual nuclei of an embryo are located [Pous 04, Pisa 03].

This section describes the proposed segmentation method to obtain the nuclear mask image. The method consists of three main steps: pre-processing, image binarization and a post-processing step to split merged nuclei.

3.1 Pre-processing

Sagittal images from *Drosophila* embryos were obtained using confocal microscopy. These images present some degree of noise, which should be attenuated in order to optimize the results of subsequent operations, such as image binarization. In the pre-processing stage, a simple spatial filtering can

be considered for noise removal. In addition, it is interesting to include some sharpen operations in order to highlight the edges of the nuclei presented in the image.

In the images analyzed in this work the noise are located on the extremities of cell nuclei, characterized by the random occurrence of lightness values significantly different from pixel values referring to the foreground of image. Gaussian smoothing was applied in order to reduce these noise [Gonz 08].

With regard to highlighting techniques, two operations are proposed, the linear intensity transformation and the morphological highlighting [Solo 10, Gonz 08]. The linear highlighting was applied in order to emphasize and improve the appearance of images. It is the simplest way to modify contrast and the brightness of a image, where the mapping function is represented by equation $r = ax + b$, such that x is the original pixel value and r is the pixel value after the transformation. The parameter a (angular coefficient) affects jointly the contrast and brightness of the resulting image and b (linear coefficient) adjusts only the brightness [Solo 10]. After this linear mapping function, it was applied in the images the morphological highlighting [Shih 09, Soil 99], expressed by Eq. (1).

$$I(x,y) = (G(x,y) + T_{hat}(G)) - B_{hat}(G), \qquad (1)$$

where $I(x,y)$ represents the sharpen image, $G(x,y)$ the smoothed input image and $T_{hat}(G)$ and $B_{hat}(G)$ the result of the application of "top-hat" and "bottom-hat" operators, respectively.

The technique of morphological highlighting (Eq. (1)), constructed from the combination of the operations "top-hat" and "bottom-hat" [Shih 09], aims to increase the visual discrimination between foreground and background of processed images, through addition of the "top-hat" result to the smoothed image, followed by the subtraction of "bottom-hat" image. The operations involved in morphological highlighting include binary opening and closing. The binary opening consists in applying in the binary image the operation of erosion followed by a dilation, while the binary closing consists of a dilatation followed by an erosion.

The operation "top-hat" detects the peaks (local maxima) of smoothed image $G(x,y)$, and is defined by the difference between $G(x,y)$ and the prior result of its morphological opening [Shih 09], as defined in Eq. (2). The symbol SE defines the structuring element, and it is usually chosen according to the characteristics of the image.

$$T_{hat}(G) = G(x,y) - (G(x,y) \circ SE). \qquad (2)$$

On the other hand, the operation "bottom-hat" detects the valleys (local minima) of the smoothed image $G(x,y)$, and is defined by difference between the result of its morphological closure [Shih 09] and the image $G(x,y)$, as is expressed by Eq. (3).

$$B_{hat}(G) = (G(x,y) \bullet SE) - G(x,y). \qquad (3)$$

The application of morphological highlighting emphasizes the pixels belonging to cell nuclei of sagittal images, while it attenuates the intensity values of the pixels between the nuclei.

3.2 Image Binarization

In order to binarize the pre-processed images it was verified the performance of traditional binarization techniques existing in the literature, as thresholding, edge detection and region growing [Gonz 08]. It was verified that, when applied separately, such techniques are insufficient for obtaining a precise nuclear mask [Huan 08]. These techniques were not able to provide full separation of the nuclei (i.e., they have many merged nuclei, as illustrated in Fig. 1(e-h)). This factor is influenced by morphological features (the nuclei may indeed be merged) and by small variations of light intensity existing between the nuclei.

After several tests, the Otsu thresholding [Otsu 79] was chosen for image binarization. This is an unsupervised segmentation method which relies on discriminant analysis and searches for an optimal threshold $T*$ which maximizes the variance between classes C_0 and C_1, which, in turn, represent the objects – the nuclei – and the background image (or vice versa) [Solo 10, Gonz 08]. Many approaches for cell nuclei detection in fluorescence images rely of the Otsu segmentation method [Xion 06b, Quel 10, Xion 06a, Yan 08]. In addition, in order compensate the non-uniformity of the lightning in some images, the images were subdivided into rectangles and, for each subimage, the Otsu method was applied individually. This subdivision considerable improved the results of the binarization.

However, as illustrated in Fig. 1(e-h), the result of Otsu binarization also resulted in an image with a significant presence of two or more nuclei unified. These nuclei are identified in yellow color in Fig. 1(e-h). So, after binarization it is necessary an additional step to split the merged nuclei.

3.3 Splitting merged nuclei using curvature

After the binarization of the image using Otsu method, there were still several merged nuclei. In order to split these nuclei, we propose the use of the curvature of the shape of the merged regions. The curvature quantifies the variations in the contour of a shape. It is very useful in defining of relevant geometrical properties, such

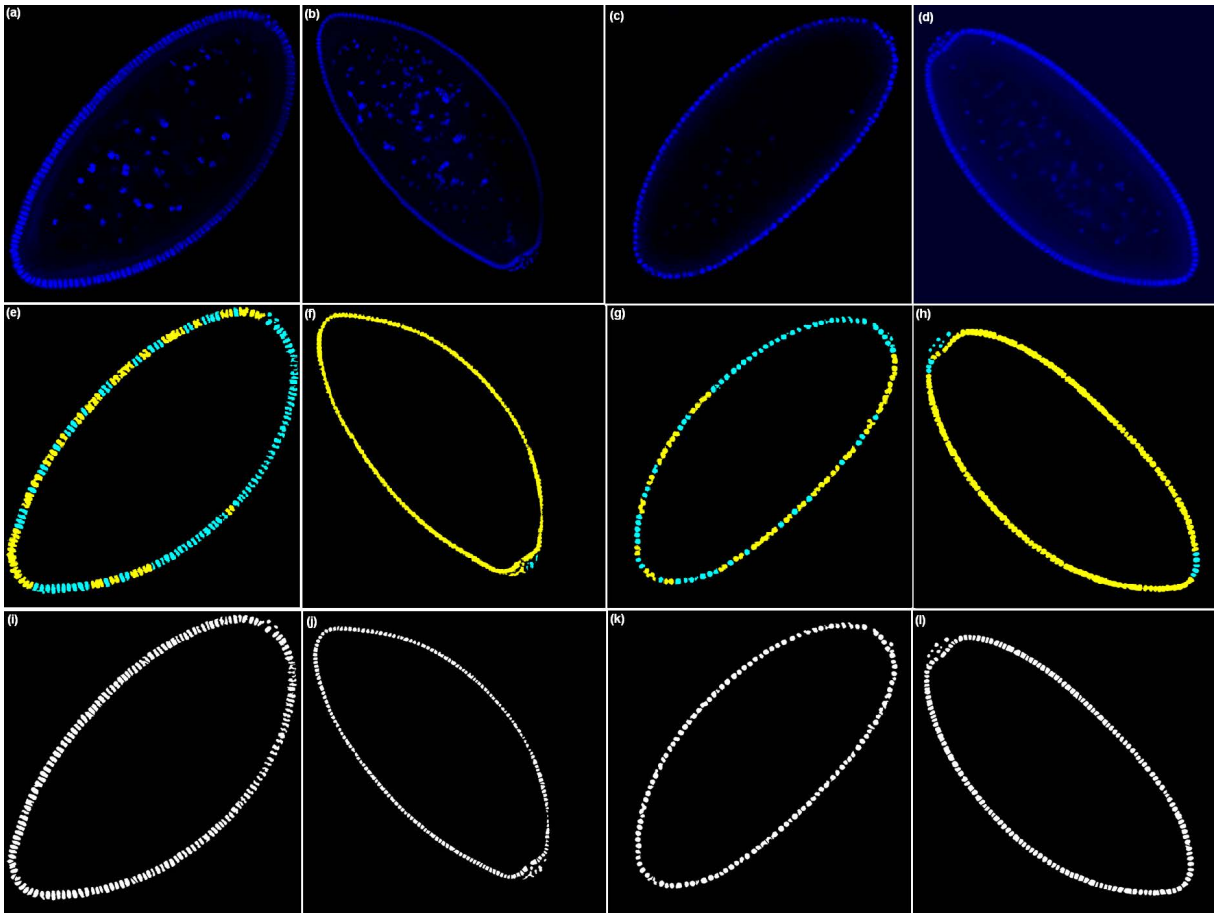

Figure 1: Original and segmented images. (a-d) Original images, obtained by confocal microscopy at the sagittal plane of *Drosophila melanogaster* embryos. (e-h) Results after of Otsu binarization. Merged nuclei are identified in yellow while rightly detected nuclei are shown in cyan. It can be seen that several nuclei remain unified after the binarization. The components outside the border of the embryo (e.g. cells inside the embryo) were removed manually after the binarization. (i-l) Results obtained after the application of the proposed methodology. Note that the merged nuclei are now separated.

as corners, valleys, concave regions and convex, and straight lines [Attn 54, Levi 85].

In digital images, the curvature of an object can be estimated from its parametric contour $C(t)$ (Eq. (4)), i.e., the representation in which the object contour is described in terms of a single parameter t, where $t = 1,...,n$ is the index of the positions of the sequential pixels $(x(t),y(t))$ defining the contour of the object and n is the number of pixels of the contour of the object.

$$C(t) = [x(t_1),y(t_1)],[x(t_2),y(t_2)],...,[x(t_n),y(t_n)]. \quad (4)$$

The curvature k for each element of a discrete contour $C(t)$ can be obtained according to Eq. (5), where $x'(t)$ and $y'(t)$ are the first order derivative with respect to x and y, and $x''(t)$ and $y''(t)$ are the second order derivative in relation to x and y, respectively [Cost 09].

$$k(t) = \frac{x'(t)y''(t) - y'(t)x''(t)}{(x'(t)^2 + y'(t)^2)^{\frac{2}{3}}}. \quad (5)$$

In this work the discrete derivatives necessary for curvature estimation was obtained using the derivative property of the Fourier Transform ($\mathfrak{F}$) [Cost 09]. Let $X(f)$ and $Y(f)$ be the Fourier Transform of the signals $x(t)$ and $y(t)$ respectively, the first and second derivatives of these signals are defined by Eqs. 6a-6d.

$$x'(t) = \mathfrak{F}^{-1}(i2\pi fX(f)), \quad (6a)$$

$$x''(t) = \mathfrak{F}^{-1}(-(2\pi f)^2X(f)), \quad (6b)$$

$$y'(t) = \mathfrak{F}^{-1}(i2\pi fY(f)), \quad (6c)$$

$$y''(t) = \mathfrak{F}^{-1}(-(2\pi f)^2Y(f)). \quad (6d)$$

In the case of discrete functions, the Eqs. 6a-6d normally generates noisy derivatives. The solution in this case is to apply a Gaussian filter in the function by including in the Eqs. 6a-6d the term $G(f,\sigma) = exp(-2(\sigma\pi f)^2)$, a Fourier Transform of a Gaussian function with zero mean and standard deviation σ, defined in the frequency space. Thus,

the effect of this filter is smoothing, where it reduces (or even eliminated) the influence of noise and small undesired details in the contour of the analyzed shape.

$$x'(s) = \mathfrak{F}^{-1}(i2\pi f X(f)G(f,\sigma)), \quad (7a)$$

$$x''(s) = \mathfrak{F}^{-1}(-(2\pi f)^2 X(f)G(f,\sigma)), \quad (7b)$$

$$y'(s) = \mathfrak{F}^{-1}(i2\pi f Y(f)G(f,\sigma)), \quad (7c)$$

$$y''(s) = \mathfrak{F}^{-1}(-(2\pi f)^2 Y(f)G(f,\sigma)). \quad (7d)$$

The estimation of smoothed curvature is calculated for all connected components with two or more nuclei unified. These components are identified by their size, been analyzed the components that are larger than a preestablished value. This value is defined based on the average size of the nuclei. The main idea in using the curvature is to find points in the image where we can trace a straight line separating the merged nuclei. The same idea was employed in the work of Beletti et al. [Bele 05], which used the curvature in the analysis of sperm images in order to separate head from the tail. The process of using the curvature can be better understood by analyzing the illustration in Fig. 2, which represents a connected component formed by intersection of four nuclei, and shows the respective points of relative minimum P_1 to P_6. These points corresponds to the relative minima (or local minima) of the curvature $k(t)$ of the contour $C(t)$. The inset in Fig. 2 depicts the curvature $k(t)$ and the corresponding relative minima P_1 to P_6.

The change of sign in the values of curvature $k(t)$ identifies the alternation between a point of relative maximum and a point of relative minimum of the curve. Between two points of relative maximum there is always a relative minimum, and vice versa. By using this analysis of maximum and minimum is possible to find the all the relative minima of the contour, and these points corresponds to the location of where the nuclei should be separated.

After the identification of the points of relative minimum, it is necessary to establish the correspondence between then in order trace the line which will separate the nuclei. In the proposed method, to select the pairs candidates to correspondence, the points of relative minimum are separated into two groups, accordingly to which side of the merged nuclei they belong. In order to define in which side of the merged block the points are located, the contour was divided into two regions. The two points of the contour which define these regions corresponds to the two peaks of the curvature of a oversmoothed version of the contour (i.e., the contour was smoothed by applying a Gaussian filter with $\sigma = 50$). This approach allows to identify the extremities of a merged block and it is translation and rotation invariant. For example, in Fig. 2, points P_1, P_2 and P_3 are located on one side of the form, while P_4, P_5 and P_6 are located on the other side.

The next step is to consider only the points located at one of the sides of the contour. For each of these points, the vectors $\overrightarrow{Vp_i}$ are constructed considering the current point and all the points on the opposite side. Then, the angles θ_i formed between the normal vector $\overrightarrow{V_n}$ at the current point with the respect to the vectors $\overrightarrow{Vp_i}$ are computed. The smallest θ_i indicates to which point of the opposite side the line that splits the nuclei should be drawn. Consider, for example, the point P_1 in Fig. 2. Then, the angles θ_i formed between the normal vector $\overrightarrow{V_n}$ at the point P_1 with the respect to the vectors $\overrightarrow{Vp_1}$, $\overrightarrow{Vp_2}$ and $\overrightarrow{Vp_3}$ are computed. The vectors $\overrightarrow{Vp_1}$, $\overrightarrow{Vp_2}$ and $\overrightarrow{Vp_3}$ are formed considering P_1 and its opposites points P_6, P_5 and P_4, respectively. For the considered example, the smallest θ_i corresponds to the angle formed between $\overrightarrow{V_n}$ and $\overrightarrow{Vp_1}$, hence P_6 is identified as the corresponding point of P_1 and a line between these points is drawn in order to split the nuclei.

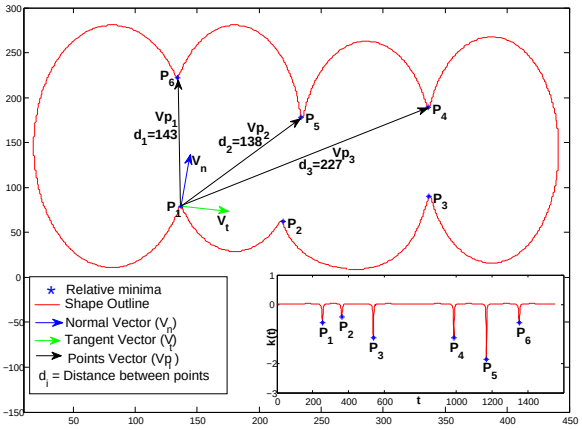

Figure 2: Illustration of a merged nuclei and the normal and tangent vector of one relative minimum point of the contour of the object. Points P_1 to P_6 are the relative minima of the contour. The vectors $\overrightarrow{Vp_1}$, $\overrightarrow{Vp_2}$ and $\overrightarrow{Vp_3}$ are formed considering the point of relative minimum P_1 and its opposite points P_6, P_5 and P_4, respectively. The distances d_i between P_1 and the opposite points are also shown. Note that the minimum distance cannot be used as a criteria to find the corresponding opposite point, as $d_2 < d_1$. In the inset it is depicted the curvature $k(t)$ of the shape of the object, along with the respective relative minimum peaks.

It is worth to note that the minimum distance between the opposite points could also be thought as a way of finding correspondence between opposite points. However, we found some examples where the smallest distance does not represent the desired correspondence. For example, in Fig. 2 the distance between the point P_1 and P_5 is lower than the distance between P_1 and P_6. On the other hand, analysis of θ is a safe choice, since

the criterion of lowest θ precisely defines the opposite point corresponding to the current point.

Finally, after the correspondence between the relative minima points are established, a straight line between then is drawn in the image in order to separate the nuclei, i.e., the pixels under the line segment unifying opposite relative minima are set to zero, separating the nuclei. In this stage, it is possible to obtain the fully separated nuclei. This procedure is applied for all the points located on the same side of the first analyzed point (i.e., the points P_2 and P_3 in the example of Fig. 2). However, if after this first iteration two or more nuclei remains merged, the procedure can be reapplied in order to obtain the expected result.

4 EXPERIMENTS

In order to evaluate the proposed methodology, 10 sagittal images from different embryos were collected. All embryos are at the cleaved cycle 14*A* [Gilb 03]. These images were processing using Matlab routines. The results obtained in our experiments were compared with other proposed techniques [Ferr 97, Malp 97, Kosm 99, Bala 12].

The parameters used in the pre-processing and image binarization steps are described as follows: the optimal size of the Gaussian filter usually depends on the size of the objects in the image. We found that a σ=3 and window size of 5x5 was adequate for subsequent analysis. The linear mapping function has been applied considering an angular coefficient equal to 9 and a linear coefficient -35 in the cases which were necessary to adjust the brightness and contrast of the image. For the Top-hat and Bot-hat transformations it was used a disk structuring element with radius 15 and 5, respectively. For the Gaussian smoothing used in curvature estimation (in the frequency space), a standard deviation equal to 5 was used. The following sections details the analysis performed.

4.1 Dataset

The images used in this work were obtained in a Leica TCS SP5 confocal microscopy, at the *Inmetro - National Institute of Metrology, Quality and Technology*. To visualize the Drosophila nuclei, we used the blue-fluorescent DAPI nucleic acid stain from Invitrogen. Each image has 1024x1024 pixels and 8 bits depth. Fig. 1(a-d) shows four of the analyzed images.

4.2 Results

In order to evaluate the accuracy of the proposed method, all 10 collected images were tested. These images contains about 61 to 183 nuclei, being 1452 in total.

Among the images collected, Fig. 1(i-l) shows the nuclear masks obtained for four images. In all images analyzed most of the nuclei were properly detected, producing a nuclear mask with a good quality.

A quantitative analysis was performed based in the number of detected nuclei. This was initially accomplished by counting this number based on visual inspection of original images. Posteriorly, the number of nuclei identified from the nuclear mask was compared to the original images. The results are summarized in Table 1, where it can be noted that the proposed method has, with reference to the number of nuclei, 94,4% of successfully detected nuclei.

Rightly detected nuclei	94,4%
Merged nuclei	3,9%
Absent nuclei	0,9%
Over-segmentation	0,8%

Table 1: Results of the application of the proposed method.

Most of the nuclei were properly identified, since only 3,9% of nuclei have not been properly separated (Merged nuclei). These were the case for blocks of merged nuclei without regions of well-defined valleys. This characteristic prevents the localization of the points of minimum curvature, therefore the segments cannot be separated.

A small portion of cell nuclei, about 0,9%, weren't accurately detected (Absent nuclei). This was the case for image regions with a very low contrast, where the Otsu's algorithm were unable to precisely identify the nuclei.

The rate of over-segmentation (i.e. individual nuclei that were split in more than one component) is very low (0,8%) and shows the efficiency of method in segmenting complex nucleus configurations. It is important to note that most of the errors verified were generated from images with low contrast and/or distinctness. Such images show nuclei with contours poorly defined (blurry), in addition to little variation of light intensities between regions of background and foreground (nuclei). These images are even difficult to perform manual segmentation of nuclei.

In addition, in order to assess the quality of the proposed segmentation method, we compared our results with the results obtained by techniques proposed by Costa et al. [Ferr 97], Malpica et al. [Malp 97], Kosman [Kosm 99] and Bala [Bala 12], which are all variations of watershed segmentation algorithm. The results of these methods are shown in Fig. 3 and are summarized in Tab. 2.

The choice to compare this work with the proposal of Kosman D. [Kosm 99] was motivated by its application in the analysis of gene expression patterns of

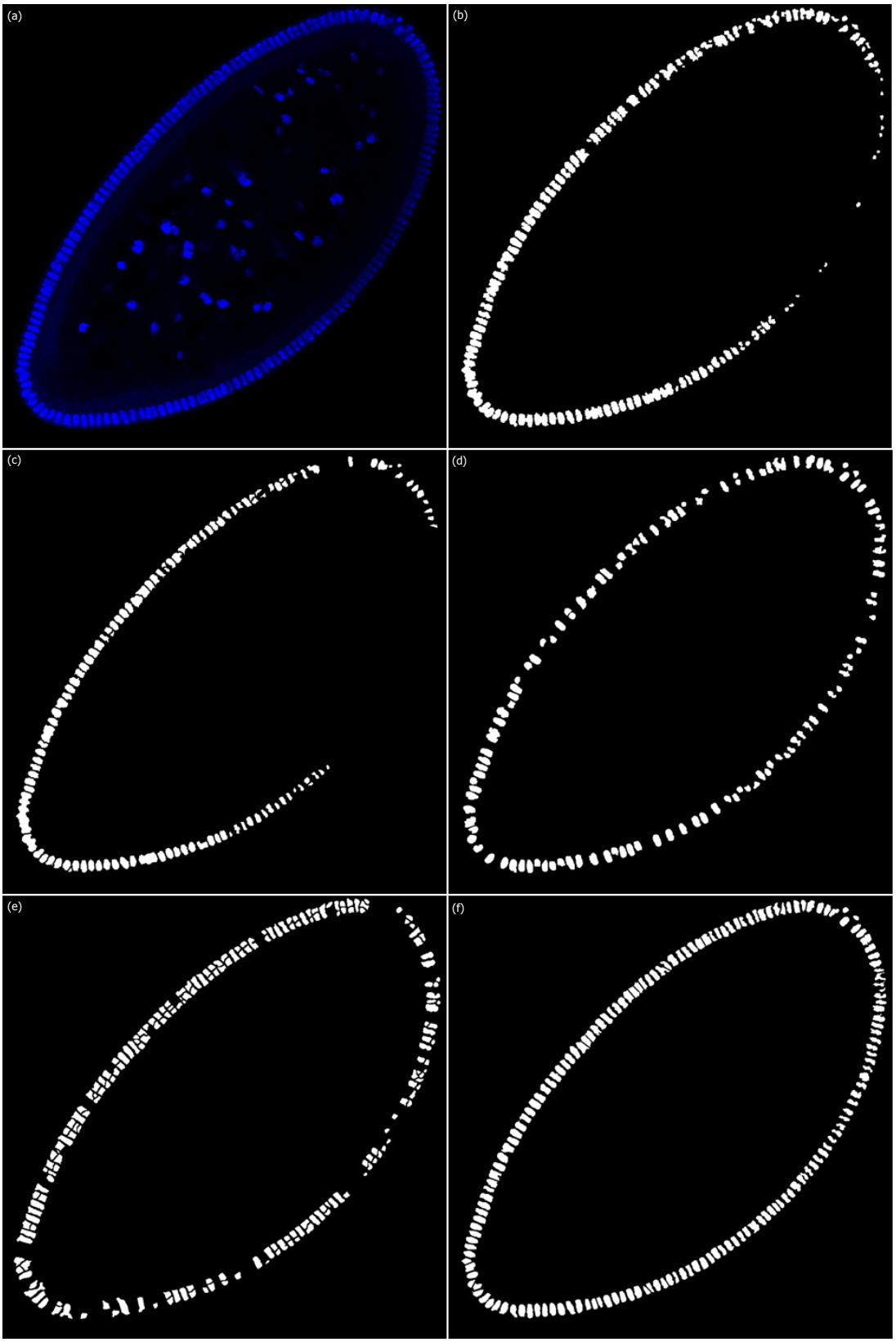

Figure 3: Comparative results of the segmentation obtained by different techniques. (a) Original image. (b) Segmentation result obtained by Costa et al. [Ferr 97]. (c) Segmentation result obtained by Malpica et al. [Malp 97]. (d) Segmentation result obtained by Bala [Bala 12]. (e) Segmentation result obtained by Kosman D. [Kosm 99]. (f) Segmentation result obtained by the proposed method.

	1-Our method	2- [Kosm 99]	3- [Ferr 97]	4- [Malp 97]	5- [Bala 12]
Rightly detected nuclei	1371 (94,4%)	937 (64,5%)	680 (46,8%)	982 (67,6%)	1043 (71,8%)
Merged nuclei	56 (3,9%)	48 (3,3%)	613 (42,2%)	192 (13,2%)	122(8,4%)
Absent nuclei	13 (0,9%)	200 (13,8%)	138 (9,5%)	247 (17,0%)	263 (18,1%)
Over-segmentation	12 (0,8%)	267 (18,4%)	21 (1,4%)	31 (2,1%)	24 (1,7%)

Table 2: Comparison of nuclei segmentation for different methods.

Drosophila . The method was used in the creation of FlyEx database [Pisa 09], and works with superficial images of *Drosophila* embryos.

The comparison of the method proposed here with the works of Costa et. al [Ferr 97], Malpica et al. [Malp 97] and Bala [Bala 12] are justified because they address revisions of Watershed algorithm, a method widely used to segment globular objects (such as nuclei/cells) in fluorescence images (2D or 3D) [Peng 08, Chen 12, Chan 12, Du 10, Hukk 10, Clop 10].

The Costa's method [Ferr 97] (Fig. 3(b)), compared to methods 2, 4 and 5 (Table 2) showed the lowest rate of over-segmentation. However, it generated the largest number of merged nuclei, mainly in regions with weak borders between the nuclei.

The Bala's method [Bala 12] (Fig. 3(d)), after our method, obtained good results in the metric 1 (Rightly detected nuclei (Table 2)). On the other hand, similarly to the Malpica's method [Malp 97] (Fig. 3(c)), it showed large portion of nuclei absent (18,1% - Table 2). These binarization errors were largely due to variations in the nuclear signal intensity, specifically, the weak signal made it difficult the definition of markers for the Watershed algorithm.

The kind of errors mentioned above are, in some extent, caused by the choice of parameter settings. Thus, a clear difficulty with such algorithms is the effort required to tune them by selecting appropriate parameter settings to different images.

In general, the experimental results show that the proposed method is more effective and highly competitive, both qualitatively and quantitatively. Furthermore, we report differences in the quality of performance of the algorithms 2–5 (Table 2).

5 DISCUSSION AND CONCLUSION

In this paper it is proposed a new segmentation method for nuclei identification from sagittal images of *Drosophila* embryos. The differential of the method is the analysis of the curvature of binary shapes in order to segment complex nuclei configuration, such as merged nuclei.

Many approaches have been used to solve this problem in the last years, especially employing segmentation methods based on variations of watershed algorithm. However, some problems are recurrent when these algorithms are used in different applications, such as the difficulty to define the markers, the efforts required to define parameters and the over-segmentation.

The results with the proposed segmentation method demonstrated that it can segment nuclei from images of *Drosophila* embryos with higher efficiency when compared to other methods. The technique proposed also can be applied to segment similar configurations of cell nuclei in other kinds of images, such as nuclei/cells commonly found in fluorescence images, where no satisfactory solution was found yet.

The proposed method is efficient and accurate, and can be integrated in a database of sagittal images of *Drosophila* embryos, contributing to gain new insights from this biological model.

6 ACKNOWLEDGMENTS

B.A.N.T. thanks PROPP-UFU, CNPq and FAPEMIG (PEE-00436-13 and *Rede 52/11*) for financial support. D.J.S. thanks CNPq (process 153870/2011-7) for the financial support. F.J.P.L. thanks FAPERJ. P.M.B. would like to thank FAPERJ and CNPq. M.A.C thanks CAPES and CNPq.

7 REFERENCES

[Attn 54] F. Attneave. "Some Informational Aspects Of Visual Perception". *Psychological Review*, Vol. 61 (3), pp. 183–193, 1954.

[Bala 12] A. Bala. "An Improved Watershed Image Segmentation Technique Using MATLAB". *International Journal Of Scientific and Engineering Research*, Vol. 3, No. 6, pp. 1206–1206, 2012.

[Bele 05] M. E. Beletti, L. F. Costa, and M. P. Viana. "A spectral framework for sperm shape characterization". *Computers in Biology and Medicine*, Vol. 35, No. 6, pp. 463–473, 2005.

[Beuc 79] S. Beucher and C. Lantuejoul. "Use of watersheds in contour detection.". *In: Internat. Workshop on Image Processing, Real-time Edge and Motion Detection/Estimation*, Vol. 132, pp. 1–12, 1979.

[Chan 12] Y.-K. Chan, P.-Y. Pai, C.-C. Liu, Y.-S. Wang, C.-W. Li, and L.-Y. Wang. "Fluorescence Microscopic Image Cell Segmentation". *International Journal of Future Computer and Communication*, Vol. 1, pp. 72–75, 2012.

[Chen 12] S. Chen, M. Zhao, G. Wu, C. Yao, and J. Zhang. "Recent Advances in Morphological Cell Image Analysis". *Computational and Mathematical Methods in Medicine*, Vol. 2012, p. 10, 2012.

[Clop 10] F. Cloppet and A. Boucher. "Segmentation of complex nucleus configurations in biological images". *Pattern Recogn. Lett.*, Vol. 31, No. 8, pp. 755–761, 2010.

[Cost 09] L. F. Costa and R. Cesar Jr. *Shape classification and analysis: theory and practice*. CRC Press, 2 Ed., 2009.

[Crow 12] D. C. Crowther and K.-F. Chen. "Functional genomics in Drosophila models of human disease". *Briefings in Functional Genomics*, Vol. 11(5), pp. 405–4115, 2012.

[Du 10] X. Du and S. Dua. "Segmentation of fluorescence microscopy cell images using unsupervised mining". *The open medical informatics journal*, Vol. 4, pp. 41–49, 2010.

[Ferr 97] J. A. Ferreira Costa, N. D. A. Mascarenhas, and M. L. de Andrade Netto. "Cell nuclei segmentation in noisy images using morphological watersheds". *Proc. SPIE, Applications of Digital Image Processing XX*, Vol. 3164, pp. 314–324, 1997.

[Fowl 08] C. Fowlkes, C. L. Luengo Hendriks, S. Keranen, G. Weber, O. Rubel, M.-Y. Huang, S. Chatoor, A. DePace, L. Simirenko, C. Henriquez, A. Beaton, R. Weiszmann, S. Celniker, B. Hamann, D. Knowles, M. Biggin, M. Eisen, and J. Malik. "A Quantitative Spatiotemporal Atlas of Gene Expression in the Drosophila Blastoderm". *Cell*, Vol. 133, No. 2, pp. 364–374, 2008.

[Gilb 03] S. Gilbert. *Developmental Biology*. Sinauer Associates, 7 Ed., 2003.

[Gonz 08] R. Gonzalez and R. Woods. *Digital Image Processing*. USA, 3 Ed., 2008.

[Greg 07] T. Gregor, E. F. Wieschaus, A. P. McGregor, W. Bialek, and D. W. Tank. "Stability and Nuclear Dynamics of the Bicoid Morphogen Gradient". *Cell*, Vol. 130, pp. 141–152, 2007.

[Houc 02] B. Houchmandzadeh, E. Wieschaus, and S. Leibler. "Establishment of developmental precision and proportions in the early Drosophila embryo". *Nature*, Vol. 415, No. 6873, pp. 798–802, 2002.

[Huan 08] M.-Y. Huang, O. Rübel, G. Weber, C. Hendriks, M. Biggin, H. Hagen, and B. Hamann. "Segmenting Gene Expression Patterns of Early-stage Drosophila Embryos". In: L. Linsen, H. Hagen, and B. Hamann, Eds., *Visualization in Medicine and Life Sciences*, pp. 313–327, Springer Berlin Heidelberg, 2008.

[Hukk 10] J. Hukkanen, A. Hategan, E. Sabo, and I. Tabus. "Segmentation of cell nuclei from histological images by ellipse fitting". In: *Proc. of the European Signal Processing Conference, Aalborg, Denmark*, pp. 1219–1223, 2010.

[Jaeg 04] J. Jaeger, M. Blagov, D. Kosman, K. N. Kozlov, Manu, E. Myasnikova, S. Surkova, C. E. Vanario-Alonso, M. Samsonova, D. H. Sharp, and J. Reinitz. "Dynamical Analysis of Regulatory Interactions in the Gap Gene System of Drosophila melanogaster". *Genetics*, Vol. 167, No. 4, pp. 1721–1737, 2004.

[Jans 05] H. Janssens, D. Kosman, C. Vanario-Alonso, J. Jaeger, M. Samsonova, and J. Reinitz. "A high-throughput method for quantifying gene expression data from early Drosophila embryos". *Development genes and evolution*, Vol. 215, No. 7, pp. 374–381, 2005.

[Kosm 99] D. Kosman, J. Reinitz, and D. H. Sharp. "Automated assay of gene expression at cellular resolution". In: *Proceedings of the 1998 Pacific Symposium on Biocomputing*, pp. 6–17, 1999.

[Kozl 08] K. Kozlov. "ProStack: a new platform for image analysis". In: *CSHL Conference "Computational Cell Biology"*, 2008.

[Levi 85] M. D. Levine. *Vision in man and machine*. McGraw–Hill, 1985. NY.

[Luen 06] C. L. Luengo Hendriks, S. V. E. Keränen, C. C. Fowlkes, L. Simirenko, G. H. Weber, A. H. DePace, C. Henriquez, D. W. Kaszuba, B. Hamann, M. B. Eisen, J. Malik, D. Sudar, M. D. Biggin, and D. W. Knowles. "Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution I: data acquisition pipeline.". *Genome Biol*, Vol. 7, No. 12, p. R123, 2006.

[Malp 97] N. Malpica, C. O. de Solórzano, J. J. Vaquero, A. Santos, I. Vallcorba, J. M.

García-Sagredo, and F. D. Pozo. “Applying watershed algorithms to the segmentation of clustered nuclei”. *Cytometry*, Vol. 28, No. 4, pp. 289–297, 1997.

[Otsu 79] N. Otsu. “A threshold selection method from gray-level histograms”. *IEEE Transactions on Systems*, Vol. 9, pp. 62–66, 1979.

[Pan 06] J.-Y. Pan, A. G. R. Balan, E. P. Xing, A. J. M. Traina, and C. Faloutsos. “Automatic mining of fruit fly embryo images”. In: *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 693–698, ACM, New York, NY, USA, 2006.

[Peng 08] H. Peng. “Bioimage informatics: a new area of engineering biology”. *Bioinformatics*, Vol. 24, No. 17, pp. 1827–1836, 2008.

[Pisa 03] A. Pisarev, E. Poustelnikova, M. Samsonova, and P. Baumann. “Mooshka: a system for the management of multidimensional gene expression data in situ”. *Information Systems*, Vol. 28(4), pp. 269–285, 2003.

[Pisa 09] A. Pisarev, E. Poustelnikova, M. Samsonova, and J. Reinitz. “FlyEx, the quantitative atlas on segmentation gene expression at cellular resolution”. *Nucleic Acids Research*, Vol. 37, No. suppl 1, pp. D560–D566, 2009.

[Pous 04] E. Poustelnikova, A. Pisarev, M. Blagov, M. Samsonova, and J. Reinitz. “A database for management of gene expression data in situ”. *Bioinformatics*, Vol. 20, No. 14, pp. 2212–2221, 2004.

[Quel 10] P. Quelhas, M. Marcuzzo, and A. M. Mendonça. “Cell Nuclei and Cytoplasm Joint Segmentation Using the Sliding Band Filter”. *IEEE Transaction on Medical Imaging*, Vol. 29, No. 8, pp. 1463–1473, 2010.

[Rbel 06] O. Rübel, G. H. Weber, S. V. E. Keränen, C. C. Fowlkes, C. L. L. Hendriks, L. Simirenko, N. Y. Shah, M. B. Eisen, M. D. Biggin, H. Hagen, D. Sudar, J. Malik, D. W. Knowles, and B. Hamann. “Pointcloudxplore: Visual analysis of 3d gene expression data using physical views and parallel coordinates”. In: *Eurographics/IEEE-VGTC Symposium on Visualization Proceedings*, pp. 203–210, 2006.

[Shap 01] L. G. Shapiro and G. C. Stockman. *Computer Vision*. Prentice Hall, New Jersey, 2001.

[Shih 09] F. Y. Shih. *Image Processing and Mathematical Morphology: Fundamentals and Applications*. CRC Press, 2009.

[Soil 99] P. Soille. *Morphological Image Analysis: principles and applications*. Springer-Verlag Berlin Heidelberg, 1999.

[Solo 10] C. Solomon and T. Breckon. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*. Wiley-Blackwell, 2010.

[Surk 08] S. Surkova, E. Myasnikova, H. Janssens, K. N. Kozlov, A. A. Samsonova, J. Reinitz, and M. Samsonova. “Pipeline for acquisition of quantitative data on segmentation gene expression from confocal images.”. *Fly (Austin)*, Vol. 2, No. 2, pp. 58–66, March 2008.

[Toma 07] P. Tomancak, B. P. Berman, A. Beaton, R. Weiszmann, E. Kwan, V. Hartenstein, S. E. Celniker, and G. M. Rubin. “Global analysis of patterns of gene expression during Drosophila embryogenesis.”. *Genome Biol*, Vol. 8, No. 7, p. R145, 2007.

[Vinc 91] L. Vincent and P. Soille. “Watersheds in digital spaces: An efficient algorithm based on immersion simulations”. *IEEE Trans. Pattern Anal. Machine Intell*, Vol. 13, No. 6, pp. 583–598, 1991.

[Xion 06a] G. Xiong, X. Zhou, and L. Ji. “Automated segmentation of drosophila RNAi fluorescence cellular images using deformable models”. *IEEE Transactions on Ciruits and Systems*, Vol. 53, No. 11, pp. 2415–2424, 2006.

[Xion 06b] G. Xiong, X. Zhou, L. Ji, P. Bradley, N. Perrimon, and S. Wong. “Segmentation of Drosophila RNAI Fluorescence Images Using Level Sets”. In: *Image Processing, 2006 IEEE International Conference on*, pp. 73–76, 2006.

[Yan 08] P. Yan, X. Zhou, M. Shah, and S. T. Wong. “Automatic Segmentation of High-Throughput RNAi Fluorescent Cellular Images”. *IEEE Transactions on Information Technology in Biomedicine*, Vol. 12, No. 1, pp. 109–117, 2008.

[Zhan 12] C. Zhang, C. Sun, R. Su, and T. D. Pham. “Segmentation of clustered nuclei based on curvature weighting”. In: *Proceedings of the 27th Conference on Image and Vision Computing New Zealand*, pp. 49–54, ACM, New York, NY, USA, 2012.