

SUPERTEXTON-BASED SEGMENTATION IN EARLY *DROSOPHILA* OOGENESIS

Rodrigo Nava and Jan Kybic

Czech Technical University in Prague, Czech Republic

ABSTRACT

Studies concerning gene expression patterns of *Drosophila* are of paramount importance in basic biological research because many genes are conserved across organisms providing information of fundamental activity. However, mapping a gene requires analyzing hundreds of objects that have been segmented previously. Hence, a reliable segmentation is a crucial step. Here, we introduce the concept of supertextons and propose a novel segmentation procedure for localized *Drosophila* ovaries. First, a pre-segmentation step is performed using superpixels; each superpixel that belongs to a single class is transformed into a feature vector. Then, a dictionary is built by clustering representative feature vectors per class, such clusters are called supertextons. Finally, during the classification stage, new superpixels are assigned to certain classes using the k -NN classifier and the supertexton dictionary. This proposal has been applied to segmentation of cells in *Drosophila* oogenesis where the results have shown the effectiveness of our approach.

Index Terms— *Drosophila*, Gene expression patterns, Oogenesis, Superpixels, Textons, Visual dictionary.

1. INTRODUCTION

Gene expression represents a means to explore and understand genetic interactions that are involved in the process of biological development because it is highly regulated, namely, genes turn on in specific cell types and at a specific time. A common technique to determine such expressions is RNA *in situ* hybridization that makes visible RNA sequences based on the ability of acids to hybridize with complementary sequences [1].

An excellent model for studying this phenomenon is the fruit fly, *Drosophila melanogaster*, which was introduced as an experimental animal almost a century ago due to its interesting attributes such as: rapid generation time, ease culturing, and low maintenance cost [2].

The authors extend their gratitude to Prof. Dr. Pavel Tomancák from Max Planck Institute of Molecular Cell Biology and Genetics for providing the images. This publication was supported by the European social fund within the framework of realizing the project “Support of inter-sectoral mobility and quality enhancement of research teams at Czech Technical University in Prague,” CZ.1.07/2.3.00/30.0034. R. Nava thanks Consejo Nacional de Ciencia y Tecnología (CONACYT). The work of J. Kybic was supported by the Czech Science Foundation project 14-21421S

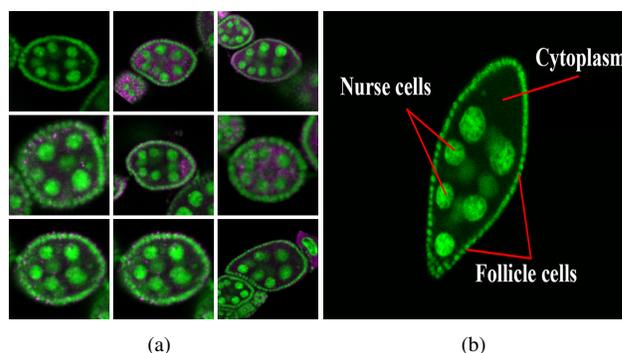


Fig. 1. *Drosophila* ovaries. (a) examples of localized ovaries during oogenesis. Green signal shows nuclei of cells during early stages of oocyte development, whereas magenta signal comes from the specific RNA expression. (b) annotated ovary showing nurse and follicle cells and cytoplasm.

Drosophila oogenesis provides outstanding conditions for genetic traceability. The process functions as an assembly line. It begins in a structure called germarium that contains somatic and germline stem cells. The latter are divided asymmetrically to produce nurse cells that are surrounded by somatic follicle cells (border cells) to form an ovary or egg chamber [3]. This process constitutes the first stage in the development of the ovary (see Fig. 1).

However, gene expression mapping requires analysis of hundreds of ovaries that have been segmented previously. We identify four classes: **i)** nurse cells; **ii)** follicle cells; **iii)** cytoplasm; and **iv)** background. This task is time-consuming and depends on trained observers. Hence, a reliable automated segmentation method would be useful to reduce time and bias.

To the best of our knowledge, there is little literature addressing *Drosophila* oogenesis segmentation. Most of the state-of-the-art methods are devoted to segment and process sub-cellular embryogenesis data [4]. Xiong et al. [5] proposed a level set method for segmenting nuclei and cytoplasm; this approach relies on the idea that a curve from a given image can evolve to detect objects. Graph-based methods have been used in [6, 7] to segment mitochondria, whereas Lucchi et al. [8] take advantage of superpixel segmentation to address the same problem.

However, these techniques have been developed for electron and not optical microscopy. We have therefore decided to apply texture analysis, which takes advantage of local intensity variations to distinguish shapes, regions, or objects.

In [9], Varma and Zisserman introduced the concept of “textons” as representative responses of the convolution of an image with a filter bank. The collection of such responses is clustered with k -means to produce a dictionary of visual words. The dictionary is used to calculate an average histogram that is employed as a model of a specific texture. Nevertheless, the main drawback of this approach is that textons characterize spatial properties of regions rather than pixels. A common practice is to split up the input image into blocks of size 32×32 or 64×64 px and build a histogram for each one. This procedure causes that small patches cannot represent regions in our data because the size of follicle cells is not larger than 10 px. Patches of such a size do not contain enough information to discriminate among classes, while larger patches lead to a block-like segmentation [10].

In this study, we investigate the use of a pre-segmentation step using simple linear iterative clustering (SLIC) algorithm [11]. SLIC groups perceptually-similar pixels and yields adherence to salient edges avoiding block-like shapes. For each segmented region that belongs to a single class, a group of features that corresponds to a filter bank response is computed and a feature vector is built. Then, a dictionary is constructed using representative features vectors that we call “supertextons.” Our proposal avoids building histograms and distinguishes small regions that represent follicle and nurse cells from the background and cytoplasm in *Drosophila* egg chambers successfully.

2. METHODOLOGY

We propose a methodology composed of three stages: **i)** pre-segmentation with SLIC superpixels; **ii)** construction of the supertexton dictionary; and **iii)** classification.

2.1. Superpixel segmentation

In our approach, an initial partition of the data is generated with SLIC superpixels. The algorithm uses seeds sampled on a regular grid and the regions belonging to individual seeds are then iteratively updated.

The number of seeds is determined as: $S = MN/s^2$ where M and N are the width and height of the image respectively and s is the size of the superpixel. Regularizer, r , is related to region compactness. If $r \rightarrow 0$ then the adherence to borders is better, otherwise the region tends to a square shape. Another advantage of SLIC is that the processing time is relatively small. In our experiments we set the initial superpixel size, $s = 10$, and regularizer, $r = 0.02$.

During the training stage, given an input image, $I(x, y)$, we generate its corresponding segmented image, $T(x, y)$, by

labeling superpixels according a ground truth. If all pixels within the superpixel belong to a single class then the superpixel is labeled with that class.

2.2. Supertexton dictionary

In order to build a supertexton dictionary, we used a reduce set of the Leung-Malik [12] filter bank by keeping only the maximum responses of first and second Derivatives of Gaussian across six orientations $\{0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$ and at three dyadic scales $\{\sqrt{2}, 2, 2\sqrt{2}\}$. Furthermore, we used eight Laplacian of Gaussian and four Gaussian filters with standard deviation $\sigma = \{\sqrt{2}, 2, 2\sqrt{2}, 4\}$. The 18 resulted filters constitute a rotation invariant representation, (LM_{18}) .

Let $I(x, y)$ be an input image and given a filter bank, LM_{18} , then the responses F_i are computed as follows:

$$F_i = I(x, y) \star LM_i \quad (1)$$

where \star is the convolution.

For each labeled superpixel in the corresponding $T(x, y)$ the following features are computed:

- Mean: $\mu_i = \frac{1}{N_I} \sum I_{F_i}$, where I_{F_i} are the intensity values $\in F_i$ within the corresponding superpixel region and N_I the number of intensity values.
- Standard deviation: $\sigma_i = \sqrt{\frac{1}{N_I} \sum (I_{F_i} - \mu_i)^2}$
- Energy: $E_i = \frac{1}{N_I} \sum (I_{F_i})^2$
- Average gradient: $G_i = \sum \|\nabla I_{F_i}\|$

We included a normalization process to transform the feature vectors x to a random variable normally distributed [13]

$$\tilde{x} = \frac{x - \mu_x}{2(3\sigma_x + 1)} \quad (2)$$

where m_x and σ_x are the mean and standard deviation of the feature vector, respectively.

Hence, the rescaled feature vectors are built as follows:

$$\bar{f} = [\tilde{\mu}_1, \tilde{\sigma}_1, \tilde{E}_1, \tilde{G}_1, \dots, \tilde{\mu}_L, \tilde{\sigma}_L, \tilde{E}_L, \tilde{G}_L] \quad (3)$$

where L corresponds to the L -th filter response.

Then, the k -means clustering is applied on each of the four classes to obtain k representative cluster centers, called supertextons. The dictionary will consist of $4 \times k$ supertextons. In our experiments, k is the 10% of feature vectors per class.

2.3. Classification

During classification stage, an input image is segmented. Then, for each superpixel, a feature vector is calculated as we describe in the previous subsection.

We used k -NN classifier with $k = 6$ which represents less than 1% of the vectors in our training dataset and assigned the input feature vector to the class with the highest number of

Table 1. Classification rates. The values represent mean rates after the classification of 54 images. All data are expressed in (%). The values between parenthesis represent standard deviations. Bold values represent the best rates.

Class	Supertextons			Pixel-based			Textons		
	PR	SE	F_1 -Score	PR	SE	F_1 -Score	PR	SE	F_1 -Score
Background	95.98 (\pm 3.96)	90.99 (\pm 8.25)	93.15 (\pm 4.90)	96.83 (\pm 3.32)	80.80 (\pm 14.41)	87.22 (\pm 9.50)	99.82 (\pm 0.82)	45.19 (\pm 23.35)	59.18 (\pm 23.24)
Follicle cells	82.90 (\pm 6.50)	79.14 (\pm 7.17)	80.66 (\pm 4.66)	72.06 (\pm 6.76)	64.92 (\pm 8.19)	67.76 (\pm 4.60)	37.50 (\pm 7.37)	85.85 (\pm 7.38)	52.20 (\pm 7.14)
Nurse cells	79.64 (\pm 9.52)	85.22 (\pm 5.03)	81.87 (\pm 4.99)	64.33 (\pm 10.12)	71.69 (\pm 7.36)	67.11 (\pm 5.77)	74.55 (\pm 13.84)	63.04 (\pm 9.25)	67.59 (\pm 7.28)
Cytoplasm	80.18 (\pm 10.06)	86.68 (\pm 4.46)	82.84 (\pm 5.75)	67.11 (\pm 10.03)	84.26 (\pm 5.95)	74.01 (\pm 6.04)	72.18 (\pm 20.57)	47.41 (\pm 20.73)	56.28 (\pm 18.20)

neighbors using the Euclidean distance between feature vectors \vec{f} as a metric. We experimented with $k = 10$ and $k = 15$, however, the variable did not affect the results significantly.

3. EXPERIMENTAL RESULTS

For this study, we selected 54 images from 3D-stacks of *Drosophila* oogenesis of size $647 \times 1024 \times 24$. They were captured using an optical sectioning fluorescence microscope. Each stack contains two channels, the green one shows nuclei of cells that constitute egg chambers providing a common reference. Note that in every slice may appear different stages of the oocyte development. The magenta channel comes from specific RNA probes labeled with a fluorescent dye and was not used.

Images of isolated egg chambers were obtained manually by expert biologists at Max Planck Institute of Molecular Cell Biology and Genetics in Dresden, Germany. As a result, the dimensions of the egg chambers varied from one to another. For each image, four classes were annotated: background, follicle cells, nurse cells, and cytoplasm, (see Fig. 2).

We compared our proposal against two methods: first, a pixel-based classification method, where the feature vectors were built using pixel-wise filter bank outputs rather than superpixels and; second, the original Varma and Zisserman’s texton approach. We used textons patches of size 10×10 px. (see a summary of the results in Table 1.

Furthermore, we evaluated the discrimination ability of the feature vectors by using the non-parametric Kruskal-Wallis test. This test computes the occurrence of significant differences between two distributions and provides an idea of how well-separated are. Since this procedure is a binary test, we computed one-class-versus-the-rest p-values. The results reflect that there are about 90% of features having p-values < 0.001 .

In order to assess the segmentation, leave-one-out cross-validation was used. The precision (PR) and sensitivity (SE) were computed and compared with a manual segmentation in a pixel-wise fashion (see Fig. 2). We also present F_1 -Score = $2 \frac{PR \times SE}{PR + SE}$, which is a measure of accuracy and reaches its best value at 1 and worst score at 0.

4. CONCLUSIONS

We proposed a novel supervised segmentation method based on superpixels and textons. For each superpixel that belongs to a single class, a feature vector was computed using mean and standard deviation, energy, and the gradient of the corresponding filter responses. Since nurse and follicle cells are pretty similar in intensities, we used gradient information to characterize superpixel contours. In addition, energy increases differences between background and cytoplasm; thus this feature led to an improvement in the classification rate.

This procedure outperforms the traditional approach based on patches that leads to a block-like segmentation. Our proposal does not use histograms to describe regions; hence, it is not subject to the size of the patch. Square-like regions including those used in textons cannot describe follicle cells as they are narrow, so, it is necessary to use small patches that may hamper the discriminative task. Furthermore, pixel-based classification is highly sensitive to noise which results in regions with holes. Nevertheless, texton-based approaches need large dictionaries in order to classify new images. Here, we segmented only egg chambers from an early stage of the *Drosophila* oogenesis. Since egg chambers are highly variables in shape and intensity values, they represent a challenge for automated segmentation methods. Further improves must include global features that discriminate further stages and avoids to merge individual nurse cells.

5. REFERENCES

- [1] P. Tomancak, A. Beaton, R. Weiszmann, E. Kwan, S. Shu, S. Lewis, S. Richards, M. Ashburner, V. Hartenstein, Celniker S., and G. Rubin, “Systematic determination of patterns of gene expression during *Drosophila* embryogenesis,” *Genome Biology*, vol. 3, no. 12, pp. 88.1–88.14, 2002.
- [2] A.M. Arias, “*Drosophila melanogaster* and the development of biology in the 20th century,” in *Drosophila*, Christian Dahmann, Ed., vol. 420 of *Methods in Molecular Biology*, pp. 1–25. Humana Press, 2008.
- [3] D.J. Montell, “Border-cell migration: the race is on,” *Nat Rev Mol Cell Biol*, vol. 4, no. 1, pp. 13–24, 2003.

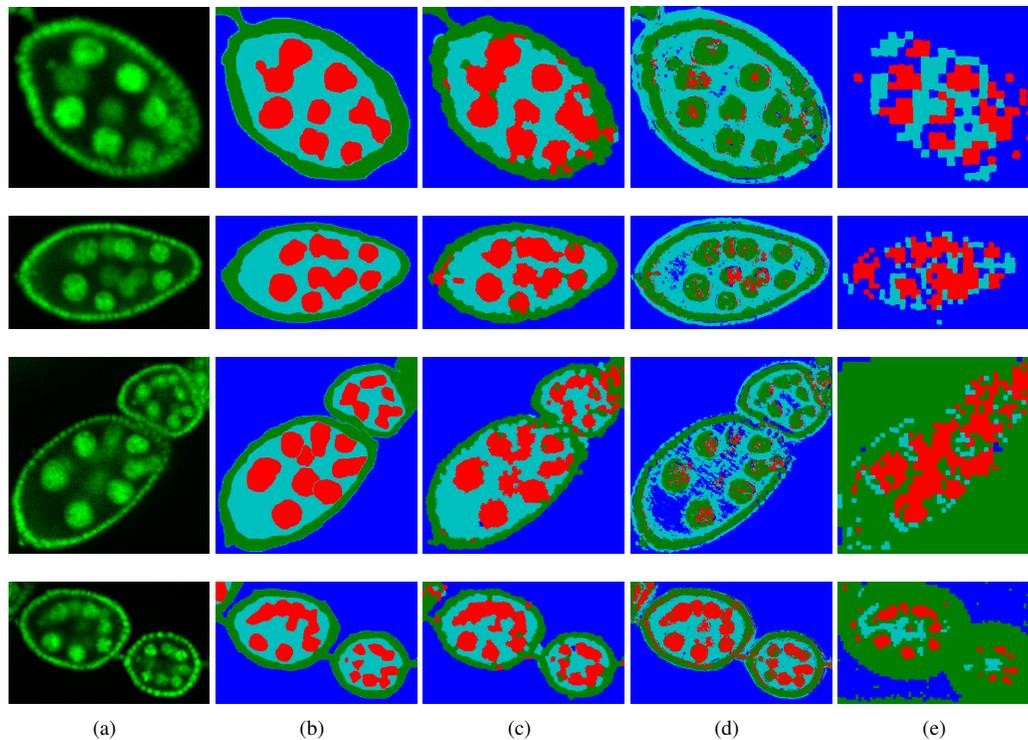


Fig. 2. Examples of segmentation in *Drosophila* egg chambers. (a) localized egg chamber; (b) the corresponding ground-truth; (c) supertexton segmentation; (d) pixel-based segmentation, and; (e) texton segmentation.

- [4] C. Castro, M.A. Luengo-Oroz, L. Douloquin, T. Savy, C. Melani, S. Desnoullez, P. Bourguine, N. Peyrieras, M.J. Ledesma-Carbayo, and A. Santos, "Image processing challenges in the creation of spatiotemporal gene expression atlases of developing embryos," in *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2011, pp. 6841–6844.
- [5] G. Xiong, X. Zhou, and L. Ji, "Automated segmentation of drosophila RNAi fluorescence cellular images using deformable models," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 11, pp. 2415–2424, 2006.
- [6] C. Chen, H. Li, and X. Zhou, "Automated segmentation of drosophila RNAi fluorescence cellular images using graph cuts," in *Advances in Multimedia Modeling*, vol. 4351 of *Lecture Notes in Computer Science*, pp. 116–125. Springer Berlin Heidelberg, 2006.
- [7] O. Ghita, J. Dietlmeier, and P.F. Whelan, "Automatic segmentation of mitochondria in EM data using pairwise affinity factorization and graph-based contour searching," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4576–4586, 2014.
- [8] A. Lucchi, K. Smith, R. Achanta, G. Knott, and P. Fua, "Supervoxel-based segmentation of mitochondria in EM image stacks with learned shape features," *IEEE Transactions on Medical Imaging*, vol. 31, no. 2, pp. 474–486, 2012.
- [9] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *International Journal of Computer Vision*, vol. 62, no. 1-2, pp. 61–81, 2005.
- [10] Y. Javed and M.M. Khan, "Image texture classification using textons," in *7th International Conference on Emerging Technologies (ICET)*, 2011, pp. 1–5.
- [11] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [12] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *Int. J. Comput. Vision*, vol. 43, no. 1, pp. 29–44, 2001.
- [13] Selim Aksoy and Robert M. Haralick, "Feature normalization and likelihood-based similarity measures for image retrieval," *Pattern Recognition Letters*, vol. 22, no. 5, pp. 563–582, 2001.