Detection and localization of Drosophila egg chambers in microscopy images

MLMI P−3

Jiří Borovec & Jan Kybic & Rodrigo Nava

Biomedical Imaging Algorithms Group, CMP, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

jiri.borovec@fel.cvut.cz
http://cmp.felk.cvut.cz/∼borovji3/

Abstract

Drosophila melanogaster is a well-known model organism that can be used for studying oogenesis (egg chamber development) including gene expression patterns. Standard analysis methods require manual segmentation of individual egg chambers, which is a difficult and time-consuming task. We present an image processing pipeline to detect and localize Drosophila egg chambers that consists of the following steps:

1. superpixel-based image segmentation into relevant tissue classes;
2. detection of egg center candidates using label histograms and ray features;
3. clustering of center candidates and;
4. area-based max. likelihood ellipse model fitting.

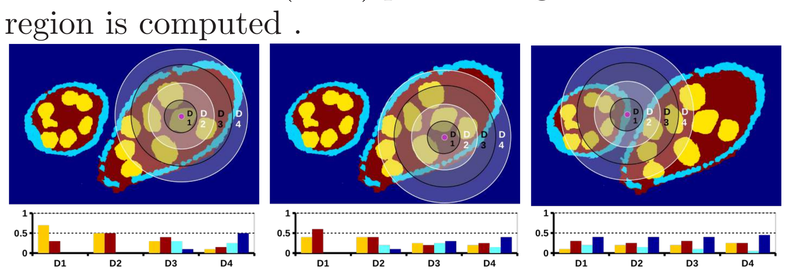
Our proposal is able to detect 96% of human-expert annotated egg chambers at relevant developmental stages with less than 1% false-positive rate and improves the mean adjusted Rand score (ARS) from 0.75 using common watershed technique to 0.86, which is adequate for the further analysis.

Superpixel segmentation

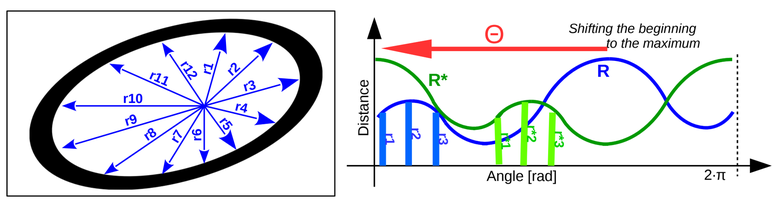
We use superpixel segmentation proposed in [1]: First, SLIC superpixels are calculated [2] with an initial size of 15 pixels. For each superpixel, color and texture features are computed. Then, the superpixels are assigned to one of four classes (background, follicle cells, cytoplasm, or nurse cells) using a random forest classifier with GraphCut [3] regularization.

Center features

Label histograms. Around a given point, a set of N annular regions D_i is defined. For each region, a normalized label (class) pixel histogram within each region is computed .

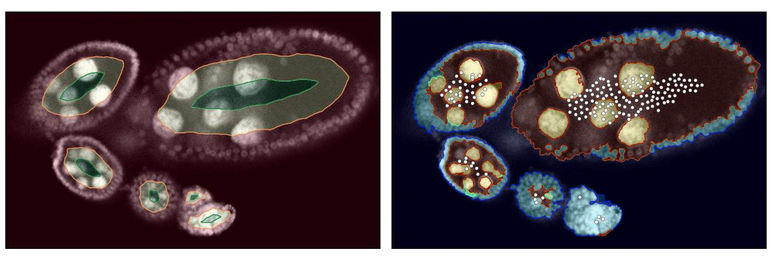


Ray features. [4] For each ray i, we measure the distance r_i to the first background-class point in the given direction. To obtain rotational invariance, the vector is circularly shifted to start with the largest element.

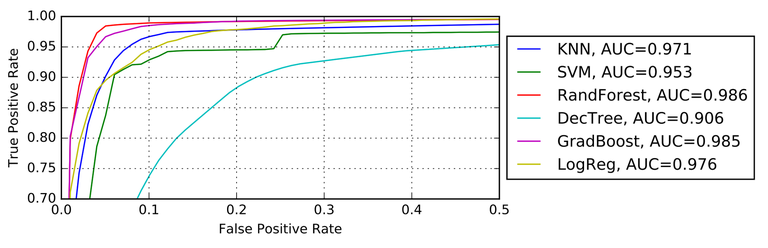


Center classification

The classifier is trained using positive central examples in green and negative far away examples in red, ignoring the intermediate zone in yellow.

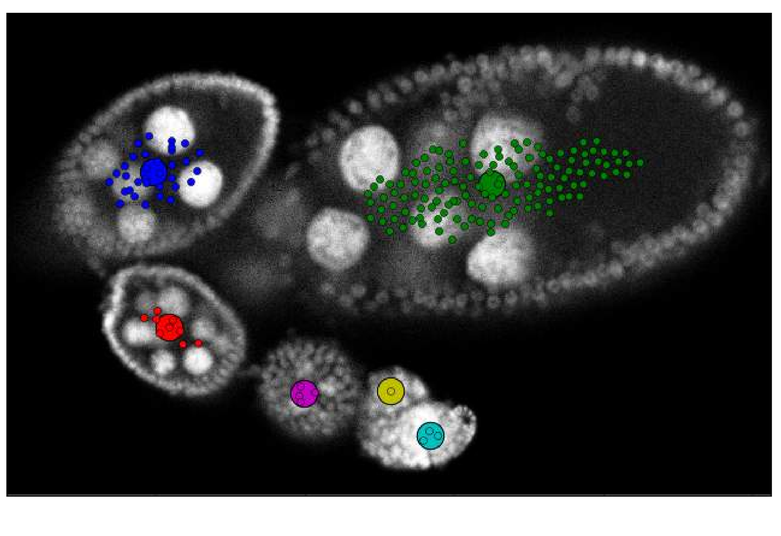


Performances of different classifiers for the center candidate detection task. Random Forest classifier was chosen.

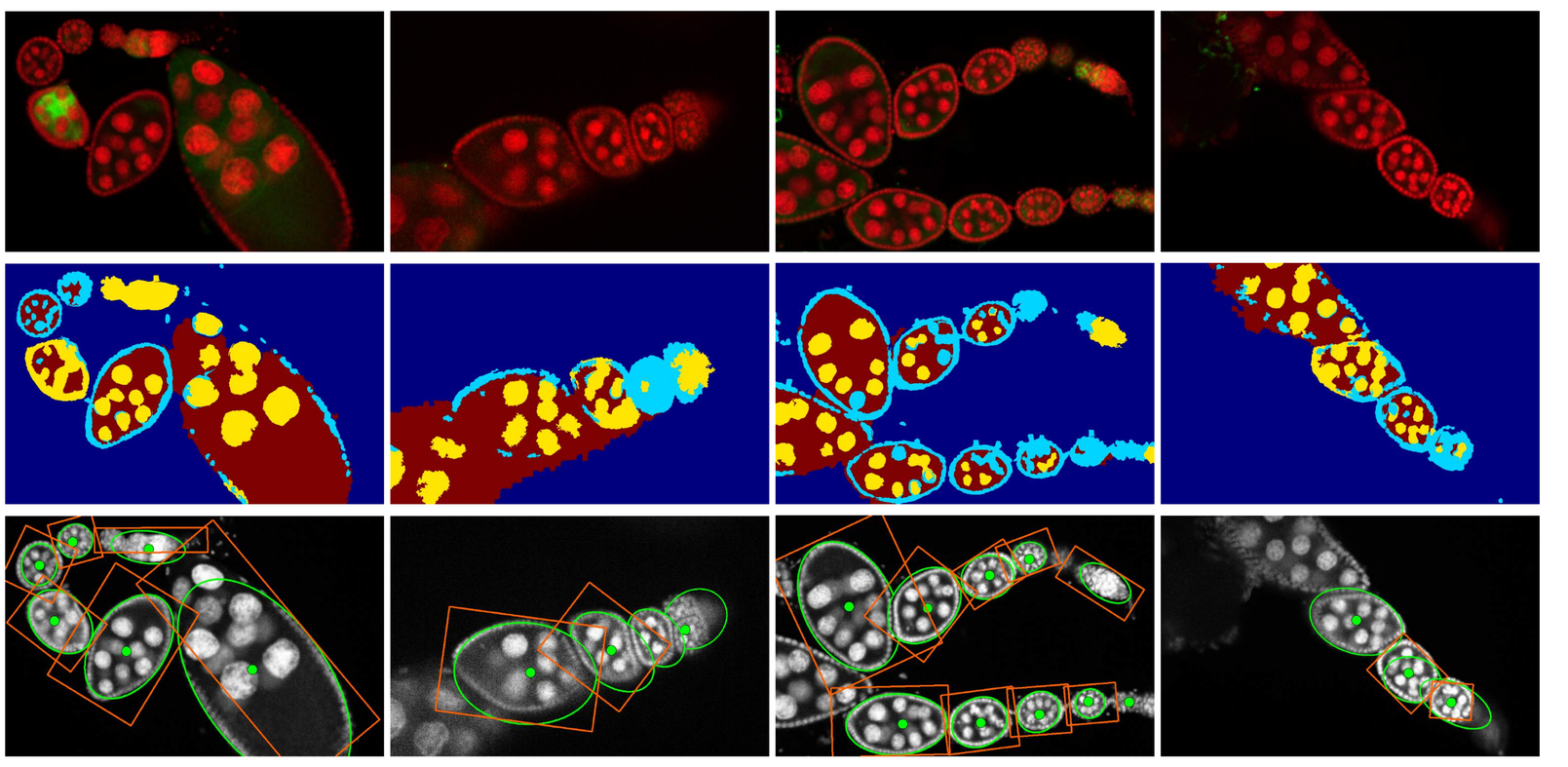


Center clustering

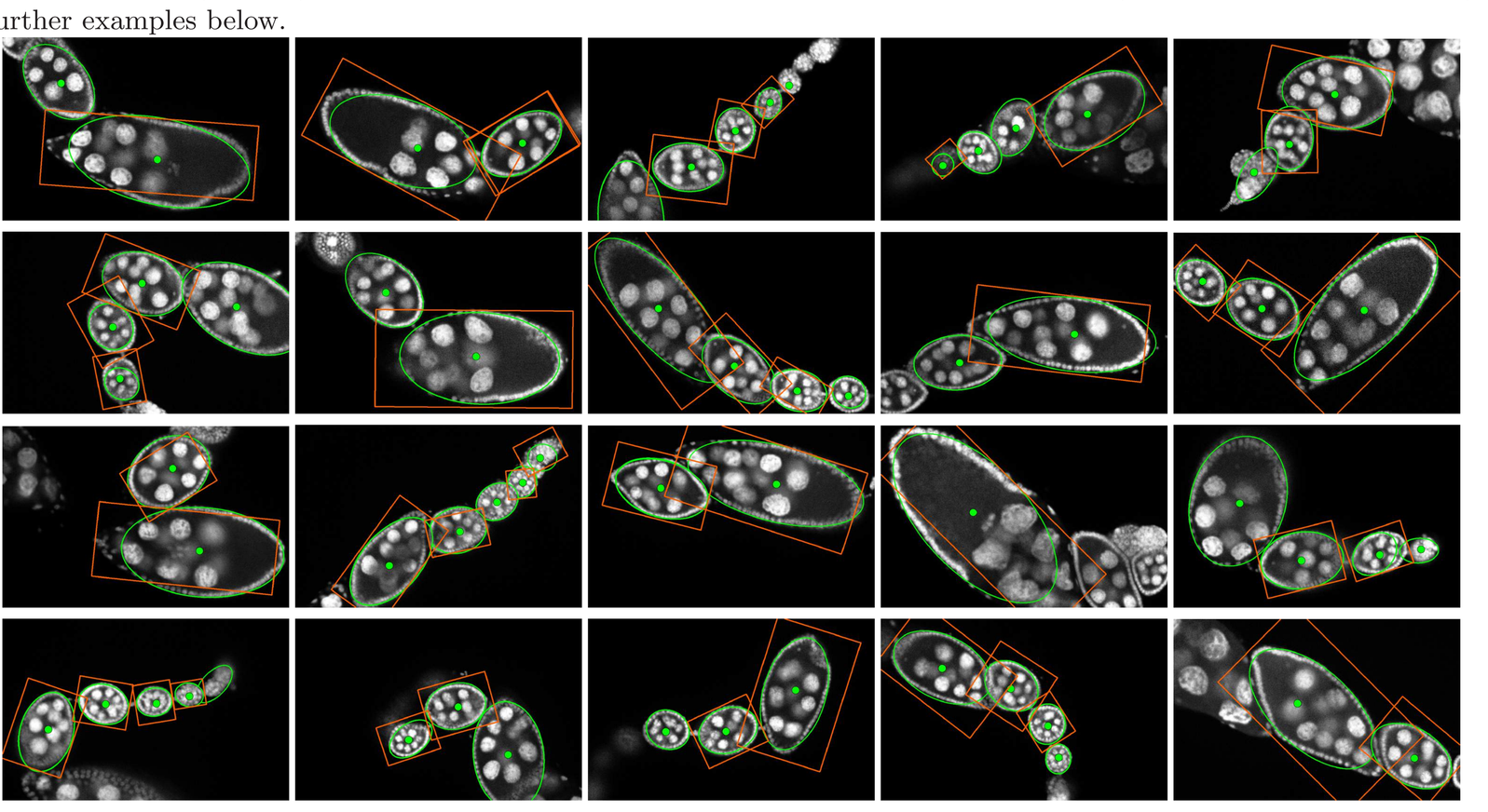
Detections corresponding to individual eggs are grouped together using density-based spatial clustering (DBSCAN) [5]. The distance threshold of DBSCAN is set to 3× the superpixel size.



Experimental results



Input images (Top), initial segmentation (middle) followed by the detected centers (cluster means) as dots and the fitted ellipses in green (bottom). Expert drawn bounding boxes are shown as red rectangles (not all eggs are annotated). Further examples below.



Detection results

Egg detection performance of the egg detection task by development stages, in terms of false positives, false negatives, and the number of multiply detected eggs before and after post-processing with ellipse fitting.

Egg chambers	Stage				
	1	2	3	4	5
number	921	1403	865	834	836
false negatives	306 (33%)	158 (11%)	6 (0.7%)	1 (0.1%)	0 (0.0%)
multiple detections (MD)	37 (4.0%)	31 (2.2%)	109 (12%)	80 (9.6%)	90 (11%)
MD after ellipse fitting	18 (2.0%)	13 (0.9%)	27 (3.1%)	20 (2.4%)	30 (3.6%)
false positives	43 (0.9%)				

Ellipse fitting

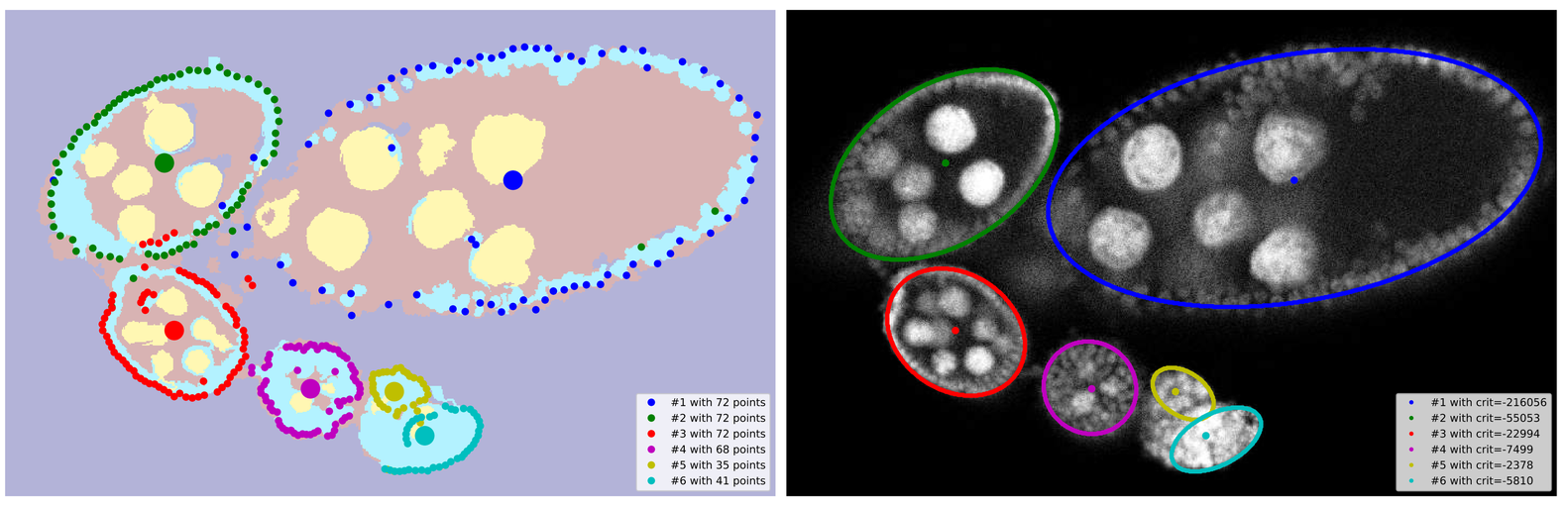
Maximize the likelihood (Ω is the entire image, Ω_F the ellipse interior)

$$\prod_{i\in\Omega_F} P_F(Y_i) \cdot \prod_{i\in\Omega\setminus\Omega_F} P_B(Y_i)$$

$P_F(Y_i)$ and $P_B(Y_i)$ are image-based foreground/background probabilities for pixel i.
Taking negative log likelihood $g_\bullet = -\log P_\bullet$ we get

$$\min \sum_{i\in\Omega_F} g_F(Y_i) - g_B(Y_i)$$

To obtain a robust fit, ellipses are fitted [6] RANSAC-like to randomly selected subsets of 40% of detected boundary points for each center.



Conclusion

We presented a complete pipeline for Drosophila egg chamber detection and localization by ellipse fitting in microscopic images. Our contributions include novel label histogram features, the rotation invariant ray features, and area-based maximum likelihood ellipse fitting. The performance is completely adequate for the desired application — it is important that the number of false positives is small but false negatives are not a problem, as long as a sufficiently high number of egg chambers is detected.

References

[1] R. Nava and J. Kybic. Supertexton-based segmentation in early Drosophila oogenesis. In *IEEE International Conference on Image Processing (ICIP)*, pages 2656–2659, 2015.

[2] R. Achanta and et al. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE PAMI*, 34(11):2274–2282, 2012.

[3] Yuri Boykov and Olga Veksler. Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE*, 23(11):1222–1239, 2001.

[4] K. Smith, A. Carleton, and V. Lepetit. Fast ray features for learning irregular shapes. In *IEEE 12th International Conference on Computer Vision*, pages 397–404, September 2009.

[5] Martin Ester, Hans P Kriegel, Jorg Sander, and Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *International Conference on Knowledge Discovery and Data Mining*, pages 226–231, 1996.

[6] Radim Halir and Jan Flusser. Numerically stable direct least squares fitting of ellipses. *Central Europe on Computer Graphics and Visualization*, 98(WSCG):125–132, 1998.