Project APC 524

Guanhua He (Molbio), Ping Wu (Molbio), Zongjun Tan (ORFE)

December 4, 2018

1 Task description

This project is about to improve some image analyzing codes and their performance related to a study on drosophila embryo conducted by one of our group members Ping Wu. She uses some fluorescent antibodies to stain for endogenous proteins to get spatial gene expression profiles at certain stage of embryonic development. Then, she will infer how the perturbation exerted on the embryos affect the transcription network and downstream gene expression base on the profile. Before making any biological analysis, she requires to extract fluorescent intensities from each cell along the whole embryo, Given the above background information, we would like to achieve two objectives in this project. The first one is to create an interface adapted to the existing codes in order to facilitate the future development in her own research. The second objective tends to improve automation. The initial image processing is still done manually currently, which is time consuming. The initial step requires recognition the morphology of the embryo and distinguish head and tail, dorsal and ventral side, and then rotate and center the object. Some noise removal steps are also required. If we have spare time, we could adapt this code to a series of time lapse images, with cell tracking function and parallel computing function for large scale image processing.

2 Organization

2.1 Goal

[detail description will be completed by Ping]

A group of drosophila embryos that lie on a plane have been photographed in order to keep track of the development of 4 genes' spatial expression profiles (say "gene" for short in the following). At a special given time, we are provided with a collection of images corresponded to different sample embryos as well as their gene profiles. Each sample embryo has 4 different gene profiles.

Question: please describe how many gene profiles do we have? different gene expression profile due to how many different situations?

Our objective is the following:

1. Rotate the sample embryo in an image so that the line from head to tail of that embryo is level.

2. Detect meaningful area of/inside a sample embryo image to extract the luminosity information of gene expression.

Optimal: Create a mask from a particular gene expression of a sample embryo, and then apply this mask to other gene expressions of the same embryo in order to detect luminosity information of other genes.

2.2 Design pattern

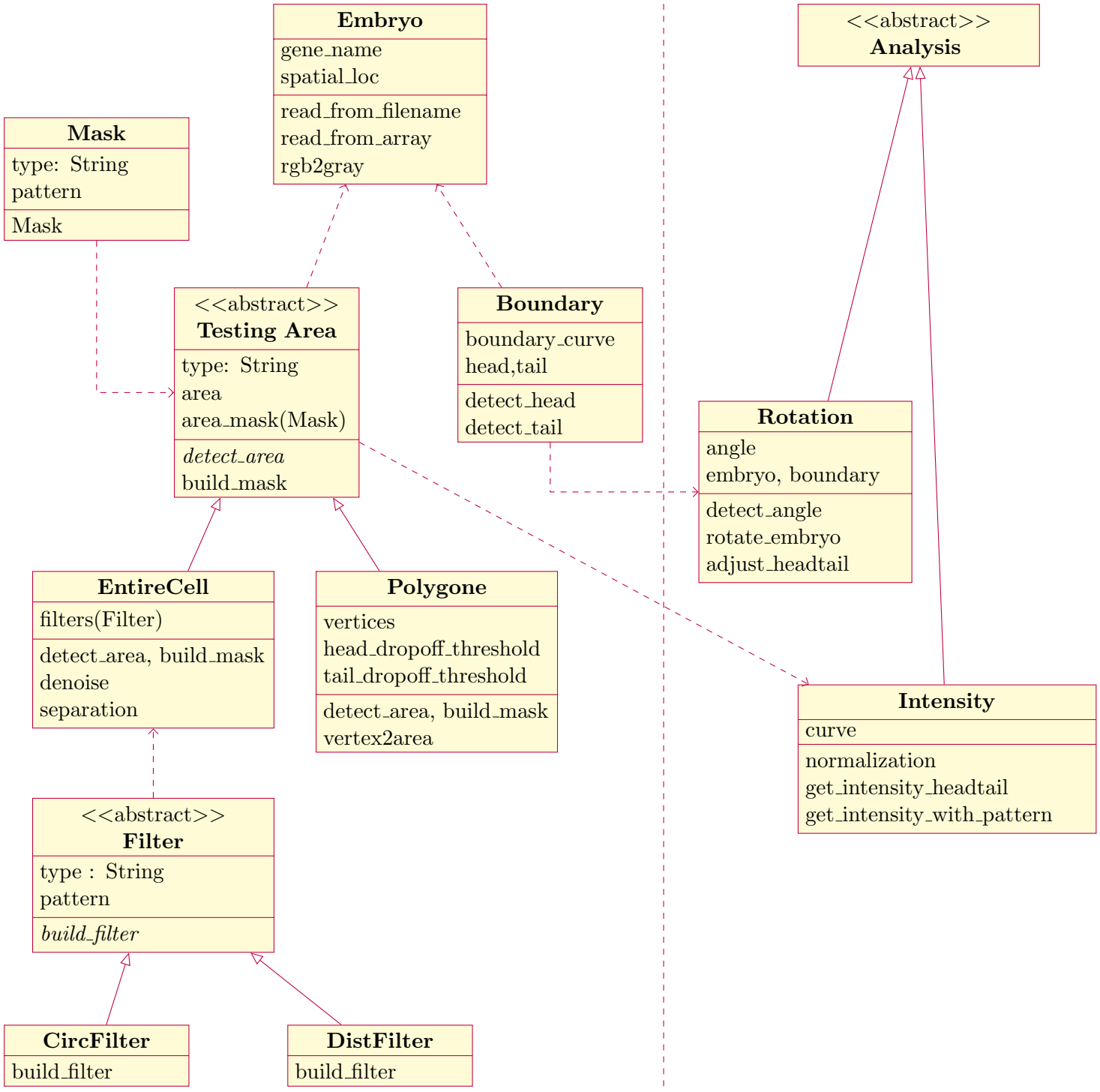
The codes will be written in Python.

2.2.1 detail description

Embryo

attribute	type	Description
gene_name	String	name of the gene
spatial_loc_rgb	$2d \times 3$ array	rgb image of the spatial gene expression in an embryo
spatial_loc	2d array	gray scale image of *spatial_loc_rgb*
method	**return type**	**description**
read_from_filename	$2d \times 3$ array	read an rgb image from a filename, store it in *spatial_loc*
read_from_array	$2d \times 3$ array	read an rgb image from a 2d array, store it in *spatial_loc*
rgb2gray	2d array	convert to gray scale image

2.2.2 Hierarchy between classes



3 Schedule

In total 8 weeks.

- Nov 18 (Sun) – Nov 24 (Sat) (Thanksgiving week):

  Discuss the project with Gabe. Question: what should be included in the prototype? To which extend should we do parallelization? How to represent profiling checking?

  Read reference papers of Ping's project.

  Understand properly the existence code structure. (potentially 1 extra meeting) Clarification of all relationships between existing functions and their input/output.

  List one or two potential requires in the future development.(Ping)

  First version of the prototype. (Graphical + coding)

- Nov 25 (Sun) – Dec 1 (Sat) :

  Second version of the prototype.

  Familiar with parallel programming.

  Familiar with location detection techniques in the existing codes.(To be done)

- Dec 2 (Sun) – Dec 8 (Sat):

  prepare for presenting the prototype in course. (third version)

  alpha version of parallelization.

- Dec 9 (Sat) – Dec 15 (Sat):

  Improve parallelization.

- Dec 16 (Sun) – Dec 22 (Sat) [Winter Break]:
- Dec 23 (Sun) – Dec 29 (Sat) [Winter Break]:
- Dec 30 (Sun) – Jan 5 (Sat) [Winter Break]:
- Jan 6 (Sun) – Jan 12 (Sat):

References