Image Segmentation using superpixel based MRF and shape priors

Fanny Yang
University of California
Berkeley, CA 94720, USA
fanny-yang@eecs.berkeley.edu

Siqi Wu
University of California
Berkeley, CA 94720, USA
siqi@stat.berkeley.edu

1   Introduction

Gene expression analysis on fruitfly embryos relies on dyed images taken by a DIC microscope with generally little intensity contrast. However for each gene, a different embryo had to be used. Especially in later stages of the development the organs like mouth, gut and the yolk all lie in slightly different locations and have different shapes. In order to achieve comparability of the gene expression images one needs to map the embryos to a standard template. One way to achieve this is to segment the different organs to and derive transformations between the images and the template.

In the problem that we try to tackle, the embryos are already considered to be aligned in a ellipse in the center of the image. The code had already been written to achieve that. Therefore one is left with the image segmentation task within the embryo, where the organs are very hard to discern even for the human eye. This is for example due to almost very little intensity variation and almost no color information.

In this paper I will outline a few commonly used image segmentation algorithms and their drawbacks when applied to the current problem. This will then lead to the motivation to propose a new algorithm which combines the different frameworks to form an algorithm which is relatively fast and accurate at the same time. The experiments are run on a toy example of separating the center embryo from other unwanted embryo parts on the border. I am aware this very example can be solved using much simpler algorithms such as ellipse fitting and edge detection, however the algorithm had to be ensured to work on this relatively simple example first before attempting to solve the original and much harder organ segmentation task.

2   Previous works

On a high level, the general segmentation task can be solved by minimization of an energy cost function, with each of them having an equivalent interpretation in terms of maximizing a posterior distribution. A common problem for all approaches is that they often find local structures but cannot incorporate global information. In order to overcome this, all methods also have some possibility to incorporate prior probabilities obtained by training data or prior information.

Deformable models   For example, one line of method is based on the idea of Snakes [13, 5, 4], also called "Active Contours". The decision variable they optimize over is a parameterized contour $c(s)$ with parameterization s. In the simplest model, the energy consists of the length and smoothness of the curve which is the internal energy (image independent), and a term called the external energy which measures how much it is consistent with the edge image, i.e.

$$E(\tilde{I}, c) = E_{ext}(c) + E_{int}(I, c) = \int_0^1 \tilde{I}(c(s))\mathrm{d}s + \int_0^1 \alpha(s)|c'(s)|^2 + \beta(s)|c''(s)|^2\mathrm{d}s$$

where $\tilde{I}$ here is the edge image. Priors can be incorporated using point distribution models [6].

Since the parameterization of the curve is usually cumbersome, people have proposed modelling the contour c as the zero level set of a continuous function $\phi(x, y)$ [16, 2, 17]. There, the energy minimization decision variable is a function which thus requires techniques from functional analysis, in particular partial differential equations (PDE) such as Euler-Lagrange equations. Later, this framework was extended to enable incorporation of shape and intensity priors [7, 1, 3].

In terms of incorporating a prior, this framework allows much more flexibility and robustness since we do not need to define landmarks on a contour as in [6]. Furthermore, since the contour is the zero level set of a function it can detect regions with the same label which are not connected. Although this can be a desirable for certain tasks, in our particular problem of organ segmentation this is certainly an unwanted behavior.

Graphical models Another line of work has concentrated on the probabilistic interpretation of energy minimization which is maximum likelihood estimation. The labels y are treated as unknown variables while the image features x (e.g intensity, texture etc.) are treated as observables (more details in section 3). There are mainly two different approaches: The discriminative approach (conditional random fields [14, 12]) models $p(y|x)$ directly, whereas the generative approach (Markov random fields) models $p(x|y)$ as a Gaussian distribution and uses pairwise potentials $p(y_i, y_j)$ for neighboring pixels. CRF methods generally need more training data to estimate the parameters of the distributions whereas MRF are fairly self-sufficient.

Several unsupervised and supervised extensions have been brought forth for obtaining a globally useful segmentation using prior information such as training a hidden layer as a Restricted Boltzmann machines [12] or multiscale representations of the feature space [9]. For an MRF model it is also possible to incorporate shape priors directly using a large collection of exemplars and GraphCut [15].

Hybrid models and superpixels There have already been attempts to combine the two frameworks [10, 19]. However [10] use contour priors and [19] require MCMC sampling which is computationally costly. Furthermore, pixelwise MRF/CRF are computationally inefficient to solve for high resolution images as in our segmentation task. In this case we exploit the work by Malik et al. [18] and pre-process the high-resolution images using the superpixel algorithm. It clusters regions according to their texture and brightness similarity as well as contour cues using trained data from natural images.

Algorithms on embryo organ segmentation The first attempts which were made to segment the gut for example used superpixels at different scales and applied a random forest algorithm on these superpixels. Similarly a superpixel CRF and Restricted Boltzmann machines were tested on the data. However, albeit being accurate for some of the images, they could not provide a very reliable segmentation result throughout the range of different embryos.

Since the embryos all have a similar structure for a given life stage and the superpixel segmentations usually include the desired boundary, the main idea which motivated this work is to incorporate shape priors from only few training images into the superpixel MRF framework. As we do not have a large training data set with lots of variations in the priors, the Branch-and-Bound framework proposed in [15] is not the most appropriate choice for our purposes. Instead we will incorporate the level set segmentation framework with shape prior into the superpixel MRF. The probability distribution over the shape priors is calculated using Gaussian kernel density estimation.

3   Model

In our novel model we use an MRF similar to the one in [10] on superpixels, however opposed to their model we do not use contours which carry the burden of pointwise parameterization, but take advantage of the level set functions.

Let's denote y as our hidden labels, x as our observed superpixel features (e.g. intensities, texture, relative location etc.). We index the superpixels by i and our signed distance function is $\phi(i)$. Our model for the joint probability distribution corresponds to use an undirected graphical model as in Figure 1 with node set V and edge set E.

Note that in general each superpixel can have an arbitrary number of neighbors so that the graphical model in 1 can have maximal cliques of size bigger than two. However, in our approach we only
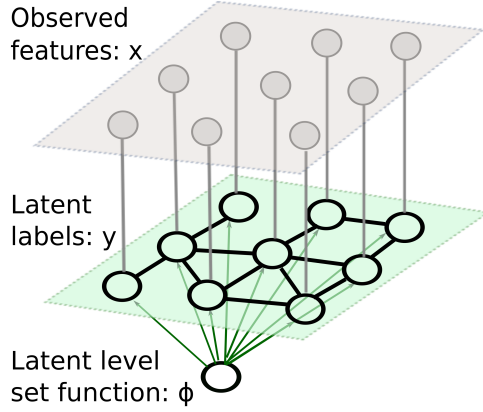
Figure 1: Markov random field model

consider pairwise potentials. This leads to the following factorization

$$p(x, y, \phi) = \prod_{i \in V} \psi(x_i, y_i) \prod_{(i,j) \in E} \psi(y_i, y_j) \prod_{i \in V} p(y_i|\phi)p(\phi) \tag{1}$$

$$=: \prod_{i \in V} p(x_i|y_i)p(y_i, y_j)p(y_i|\phi)p(\phi|\tilde{\phi})$$

where $\tilde{\phi} = \{\tilde{\phi}_k\}_{k \in I}$ denote prior shapes with I being the index set of shape priors.

We model $p(x_i|y_i)$ as a multivariate Gaussian mixture with means μ_k and diagonal covariance matrices Σ_k for $k = 0, \ldots, K - 1$ where K is the number of labels (in this paper $K = 2$ for binary segmentation), i.e.

$$p(x_i|y_i = k) = \frac{1}{\sqrt{|2\pi\Sigma_k|}} \exp(-\frac{1}{2}(x - \mu_k)\Sigma_k^{-1}(x - \mu_k))$$

Our pairwise potentials penalize neighbors with different labels and read

$$p(y_i, y_j) \propto \theta_1 \mathbb{I}_{y_i \neq y_j} + \theta_2 \mathbb{I}_{y_i = y_j} \tag{2}$$

For the consistency measure of the labels y on ϕ we use the logistic function

$$p(y_i|\phi) = \left(\frac{1}{1 + e^{-\lambda\phi(i)}}\right)^{y_i} \left(\frac{1}{1 + e^{\lambda\phi(i)}}\right)^{1-y_i}$$

where $\phi(i)$ is usually a signed distance function and λ controls the sharpness of the slope.

For the probability distribution on the level set function we use the well-known energy cost function [7, 2, 17]

$$p(\phi|\tilde{\phi}) = E_{int}(\phi) + E_{prior}(\phi) \tag{3}$$

$$= \lambda_1 \int_\Omega \delta_\epsilon(\phi(x,y))\|\nabla\phi(x,y)\|\mathrm{d}x\mathrm{d}y + \lambda_2 \int_\Omega H_\epsilon(\phi(x,y))\mathrm{d}x\mathrm{d}y$$

$$- \log \sum_{i=1}^N \exp\left(-\frac{1}{2\sigma^2}d^2(H(\phi), H(\tilde{\phi}_i))\right)$$

where d is the symmetric distance between the level set functions defined as

$$d^2(H_\epsilon(\phi_1), H_\epsilon(\phi_2)) = \int_\Omega (H_\epsilon(\phi_1) - H_\epsilon(\phi_2))^2 \mathrm{d}x\mathrm{d}y$$

Instead of using the non-differentiable Heaviside function we apply a smoothed version, i.e.

$$H_\epsilon(z) = \frac{1}{2}(1 + \frac{2}{\pi}\arctan\left(\frac{z}{\epsilon}\right))$$

3

as in [2] with derivative $\delta_\epsilon(z) = \frac{\epsilon}{z^2 + \epsilon^2}$. The ϵ is dropped in the following for simpler notation.

E_{int} denotes the regularization on length and area of the region (boundary), E_{prior} is the energy given by the shape prior which gives the probability of ϕ given the training data set $\hat{\phi}$. The shape distribution is estimated using nonparametric kernels with width $\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} \min_{j \neq i} d^2(H(\phi_i), H(\phi_j))$.

3.1 Variational inference approach

The inference task now is to maximize $p(y, \phi|x) \propto p(x, y, \phi)$. However finding the maximum of (1) exactly is computationally intractable. First of all the graphical model is not a tree and secondly the domain of ϕ is infinite dimensional. A very nice framework to overcome this difficulty is to use shape priors directly in the model as described in [15] since submodular quadratic pseudo-boolean function of the form

$$E(y) = a + \sum_i a_i y_i + \sum_{i,j} a_{ij} y_i y_j$$

can be computationally efficiently. However as mentioned above their approach for incorporating priors is not suitable for our task and the general GraphCut approach does not work for infinite dimensional variables. The aim is thus to approximate the conditional distribution by adding variational parameters a, b which render the labels and the level set function conditionally independent.

This approach is called variational inference [11, 20] and our ansatz for the structure of Q is the following factorization

$$Q(y, \phi|x, a, b) = Q_M(y|x, a)Q_D(\phi|b) \tag{4}$$

with $Q_M(y|x, a) = \prod_{i,j} \psi(y_i, y_j) \prod_i \xi(x_i, y_i)p(y_i|a_i)$ and $Q_D(\phi|b) \propto \prod_i p(b_i|\phi)p(\phi)$. The best approximation Q with this structure is found by minimizing the Kullback Leibler divergence $KL(Q||P) = \sum_\phi \sum_y Q(y, \phi|x, a, b) \log \frac{Q(y,\phi|x,a,b)}{p(y,\phi|x)}$ over a and b.

Using this factorization the KL divergence can be rewritten as

$$
\begin{aligned}
KL(Q||P) =~ & \mathbb{E} \, \log Q_M(y|x, a) + \mathbb{E} \, \log Q_D(\phi|b) - \mathbb{E} \, \log p(y, \phi|x) \\
=~ & \mathbb{E} \, [\log Q_M(y|x, a)|\phi] + \mathbb{E} \, [\log Q_D(\phi|b)|y] - \mathbb{E} \, [\, \mathbb{E} \, \log p(y|\phi)|\phi] - \mathbb{E} \, \log p(\phi) \\
=~ & \mathbb{E}_{Q_M} [\log Q_M(y|x, a) - \mathbb{E}_{Q_D} \log p(y|\phi)] + \mathbb{E}_{Q_D}(\log Q_D(\phi|b) - \log p(\phi))
\end{aligned}
$$

where c is a constant independent of a and b. The first equality follows from the conditional independence of y, ϕ given a, b.

For the minimization of this term with respect to a, b we use Lagrange equations of first kind to incorporate the constraint of $\mathbb{E} \, [Q_D(\phi|b)|b] = 1$ and $\mathbb{E} \, [Q_M(y|x, a)|x, a] = 1$ and obtain for the Lagrange multiplier $\lambda = 1$. Then we set the partial derivatives with respect to the distributions to zero

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial Q_M} =~ & - \mathbb{E}_{Q_D} \log p(y|\phi) + \log Q_M(y|x, a) = 0 \\
\implies & \sum_i (\, \mathbb{E}_{Q_D} \log p(y_i|\phi) - \log p(y_i|a_i)) = \sum_{j \in N(i)} \log p(y_i, y_j) - \mathbb{E}_{Q_D} \log(p(y_i, y_j)) \\
\frac{\partial \mathcal{L}}{\partial Q_D} =~ & - \mathbb{E}_{Q_M} \log p(y|\phi) + \log p(b|\phi) = 0
\end{aligned}
$$

from which it follows that the following fixed point equations must hold for optimal a, b:

$$\log p(y_i|a_i) = \mathbb{E}_{Q_D} \log p(y_i|\phi), \qquad \log p(b_i|\phi) = \mathbb{E}_{Q_M^i} \log p(y_i|\phi) \tag{5}$$

where $Q_M^i = Q_M(y_i|x, a)$.

4 Algorithm

The algorithm now has to find $p(y_i|a_i)$, $p(b_i|\phi)$ and ϕ which satisfy (5) before maximizing Q as in (4). This is done by using a non-informative initialization for $p(y_i|a_i)$ and an alternating minimization approach for the update steps. Instead of using (5) in its exact form, we approximate $Q_D(\phi|b)$ by a

dirac distribution $Q_{D'}(\phi|b) = \delta_{\phi=\phi^*}$ where $\phi^* = \operatorname{argmax}_\phi Q_D(\phi|b)$ is found by running a level set segmentation algorithm. Then the conditional probability reads

$$p(y_i|a_i) \approx \exp(\,\mathbb{E}_{Q'_D}\log p(y_i|a_i)) = p(y_i|\phi^*) \tag{6}$$

Given $p(y_i|a_i)$ we can calculate $Q_M(y_i|x_i, a)$ by estimating the parameters for $p(x_i|y_i)$ using the EM algorithm. At the same time we are interested in $Q_M(y_i|x, a)$ for computing $p(b_i|\phi)$ for which we need to compute the marginal. Since it is not a tree, we cannot obtain the exact marginal via simple message passing but have to rely on loopy belief propagation or other alternative approximate inference methods. In fact, as will be seen in section 4.1, the marginalization will have to be computed as the E-Step of the EM algorithm.

The algorithm now executes the following steps

- Initialize $p(y_i|a_i)$ by a uniform distribution.
- Step k.
  1. In order to calculate $Q_M(y_i|x_i, a)$ estimate the parameters μ_k, Σ_k of $p(x_i|y_i)$
  2. Given $Q_M(y_i|x_i, a)$ use approximate inference methods to compute $Q_M(y_i|x, a)$.
  3. Calculate $\log p(b_i|\phi) = \mathbb{E}_{Q_M^i}\log p(y_i|\phi)$.
  4. Calculate $\phi^* = \operatorname{argmax}_\phi Q_D(\phi|b)$ using (7) and obtain $p(y_i|a_i)$ as in (6)
- Upon termination: $y_i = \max_k Q_M(y_i = k|x, a)$ which will have already been solved in step 2.

In the following the different steps in the algorithm will be elaborated in more detail.

4.1   EM algorithm using loopy belief propagation

In the first step, we need to estimate the parameters for $p(x_i|y_i)$ by maximizing the log likelihood $\log p(x|\theta, a) = \log \sum_y p(x|y, \theta)p(y|a)$ with respect to θ which is the same as

$$\max_\theta \log p(x|\theta, a) \propto \max_\theta \log \sum_y Q_M(y|x, a)$$

This is a classical case to employ the EM algorithm. Note that

$$\mathcal{L}(q^{(t+1)}, \theta) = \sum_y Q_M(y|x, \theta^{(t)})(\log p(x|y, \theta) + \log p(y|a))$$

$$= \sum_i \sum_{y_i} \log p(x_i|y_i, \theta) \sum_{y_j:j\neq i} Q_M(y|x, \theta^{(t)}, a) + c$$

$$= \sum_i \sum_{y_i} Q_M(y_i|x, \theta^{(t)}, a)\mathbb{I}_{y_i=l}(-\log\sigma_l - \frac{(x_i-\mu_l)^2}{2\sigma_l^2}) + c$$

$$= \sum_i \sum_l Q_M(y_i = l|x, \theta^{(t)}, a)(-\log\sigma_l - \frac{(x_i-\mu_l)^2}{2\sigma_l^2}) + c$$

where c is independent of θ. Thus for the expectation step we need to compute the marginal

$$Q_M(y_i|x, \theta^{(t)}, a) \propto \sum_{y_j:j\neq i} \prod_i p(x_i|y_i) \prod_{i,j} p(y_i, y_j) \prod_i p(y_i|a_i)$$

given the current parameter estimates which is used in step 3 of the algorithm.

For the maximization step we maximize $\mathcal{L}(q^{(t+1)}, \theta)$ over θ which yields

$$\mu_k^{(t+1)} = \frac{\sum_i Q_M(y_i = k|x, \theta^{(t)}, a)x_i}{\sum_i Q_M(y_i = k|x, \theta^{(t)})}$$

$$\Sigma_k^{(t+1)} = \frac{\sum_i Q_M(y_i = k|x, \theta^{(t)}, a)(x_i - \mu_k^{t+1})(x_i - \mu_k^{t+1})^T}{\sum_i Q_M(y_i = l|x, \theta^{(t)}, a)}$$

for $k = 0, \ldots, K-1$.

4.2 Level Set Segmentation with prior coupled with MRF

In this section we will briefly sketch the classical level set segmentation approach which we will then augment by the coupling with the MRF. In general the task is to maximize $p(x, \phi) = p(x|\phi)p(\phi)$ with observed x with respect to ϕ. The term $p(\phi)$ enables us to incorporate priors and regularizations. Note that instead of $p(x|\phi)$ in level set formulations we now have $p(y, \phi)$, or in our variational inference framework $Q_D(\phi|b) \propto p(b|\phi)p(\phi|\tilde{\phi})$.

In order to maximize

$$Q_D(\phi|b) = p(b|\phi)p(\phi) = \exp\left(-(E_{int}(\phi) + E_{prior}(\phi, \tilde{\phi}) + E_{MRF}(\phi, Q_M))\right)$$

which is the same as minimizing the energies E_{int}, E_{prior} as in (4.2) and $E_{MRF} = \mathbb{E}_{Q_M^i} \frac{\partial}{\partial\phi} \log p(y_i|\phi)$, we use the multivariate Euler Lagrange equations [8]. They state that $F(x) = \int \mathcal{L}(x, \phi(x), \nabla\phi(x))\mathrm{d}x$ is minimized if

$$\frac{\partial}{\partial\phi}\mathcal{L}(x, \phi(x), \nabla\phi(x)) + \mathrm{div}\,\frac{\partial}{\partial\nabla\phi}\mathcal{L}(x, \phi(x), \nabla\phi(x)) = 0$$

In words, E_{MRF} is the term which couples the MRF and level segmentation framework.

In order to find the ϕ which solves this equation we initialize the level set function and perform gradient steps obeying

$$\frac{\partial\phi}{\partial t} = -\frac{\partial E_{int}(\phi)}{\partial\phi} - \frac{\partial E_{prior}(\phi)}{\partial\phi} + \mathbb{E}_{Q_M^i} \frac{\partial}{\partial\phi}\log p(y_i|\phi) \tag{7}$$

$$= \delta_\epsilon(\phi)(\lambda_1\left[\mathrm{div}(\frac{\nabla\phi}{\|\nabla\phi\|})\right] + \lambda_2) - \frac{\sum_{i\in I}\alpha_i\frac{\partial}{\partial\phi}d^2(H(\phi), H(\tilde{\phi}_i))}{2\sigma^2\sum_i\alpha_i}$$

$$+ \mathbb{E}_{Q_M^i}\lambda\left(\frac{y_i}{1 + \mathrm{e}^{\lambda\phi}} - \frac{1 - y_i}{1 + \mathrm{e}^{-\lambda\phi}}\right)$$

with $\alpha_i = \exp(-\frac{d^2(H\phi, H\phi_i)}{2\sigma^2})$. For the second term, since E_{prior} is a direct function of ϕ, we can readily take the partial derivative with respect to ϕ. $\frac{\partial}{\partial\phi}d^2(H(\phi), H(\phi_i))$ is a Gateaux derivative and yields

$$\frac{\partial}{\partial\phi}d^2(H(\phi), H(\tilde{\phi}_i)) = \delta_\epsilon(\phi)(H(\phi) - H(\tilde{\phi}_i))$$

The continuous first term in (**??**) can be expanded to

$$\mathrm{div}\,\frac{\nabla\phi}{\|\nabla\phi\|} = \frac{\frac{\partial^2\phi}{\partial x^2}(\frac{\partial\phi}{\partial y})^2 - 2\frac{\partial^2\phi}{\partial x\partial y}\frac{\partial\phi}{\partial x}\frac{\partial\phi}{\partial y} + \frac{\partial^2\phi}{\partial y^2}(\frac{\partial\phi}{\partial x})^2}{(\frac{\partial\phi}{\partial x}^2 + \frac{\partial\phi}{\partial y}^2)^{\frac{3}{2}}}$$

which we implement in our algorithm using finite differences.

5 Implementation and Results on a toy example

5.1 Toy Example Setting

For a much simpler example than organ segmentation to test our algorithmic framework, we have chosen the following problem: The task is to find the exact boundary of the center embryo in an image where parts of other embryos are sticking in and might actually share boundaries with the embryo of interest. The aim is to automatically find the boundary with only little training data.

In the results shown below, a total of 3 training images (from touching and non-touching embryos) was used for the shape priors, where each embryo was segmented manually and its shape shifted and resized leading to a pool of 27 priors in total.

5.2 Implementation details

We set the number of desired superpixels to 50. For building the edge structure after running the superpixel algorithm, we add an edge between two superpixels if their common boundary is above a certain length. We also merge very small superpixels if their number of pixels is below a threshold of 100.

The smoothing constant ϵ in the Heaviside function was chosen to be 10 and $\lambda_2 = 0$ for simplicity. For all studied cases $\lambda_1 = 5$ yielded a good result. In order to overcome numerical issues (since the exponential rendered all values even of $d^2/\sigma^2 \sim 100$ to zero) the kernel width σ also had to be chosen as $\sigma^2 = \max_i d^2(H(\phi), H(\tilde{\phi}_i))$. The level set segmentation part of the algorithm was iterated for a fixed number of 100.

As features of the superpixels we choose the intensity mean and standard deviation. For the edge potential we use information about the probability boundary (which the probability of each pixel to be an edge and is also computed as a side product in the superpixel algorithm) in the following way in (2)

$$\theta_2 = 1 - \tilde{e}(i,j) + \epsilon, \ \theta_1 = 1$$

where $\tilde{e}(i,j)$ is the normalized average probability along the boundary between superpixel i and j.

In order to speed up computation we run the superpixel algorithm on a rescaled image (by 30%) and apply the level set segmentation with shape prior on an even smaller scale (in this example case 7%). This way we can make use of the sophisticated superpixel algorithm to encode high resolution information in the feature representation of the superpixels and run both MRF and level set segmentation efficiently in much smaller dimensions.

For all of the 9 test pictures, the algorithm is found to converge after less than 4 iterations. For each picture that amounts to at most 60 seconds on a ThinkPad, including running the super pixel algorithm, which is the computationally most expensive and takes 35 seconds on its own.

5.3 Results and Discussion

An example of an original image of touching embryos is shown in Figure 2(a). One can see how the shared boundary of the two embryos is fairly long which makes the separation harder. The superpixel algorithm gives the pre-segmentation and edge structure depicted in Figure 2(b). As mentioned before one can observe that the ultimate correct boundary of the center embryo is correctly "found" by the superpixel algorithm and edges are only added if the shared boundary between two superpixels is long enough.
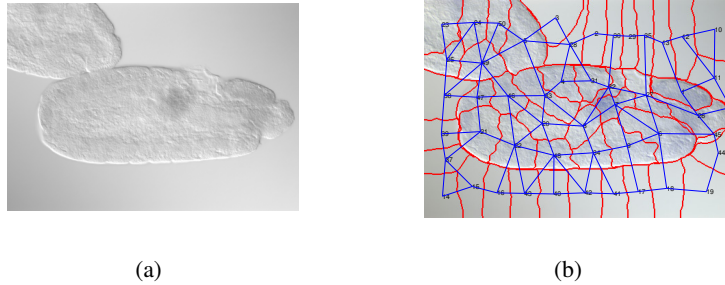

(a)   (b)

Figure 2: (a) Original picture. (b) Superpixel segmentation with edge structure for MRF (on a resized image)

We now show some results of the proposed algorithm. We first compare the algorithm on a test set with the individual approaches in Figure (3). The superpixel MRF approach directly applied on the superpixel level detects both embryos as the object of interest since they are distinct from the homogeneous background. The level set segmentation alone concentrates around areas with dark intensity but misses the global structure. Adding the prior term as in [7] adds global structure but

it still cannot find the exact edges (see also Figure 4). Another problem is that it is much harder to balance between the influence of the shape prior and the image given intensity cues.

Note that this picture sh/could be in case because of too strong shape prior. These issues can be resolved by combining the advantages of all three approaches to find the perfect contour in Figure 3(d). The pixelation of the contour is due to the image resizing for computational efficiency.
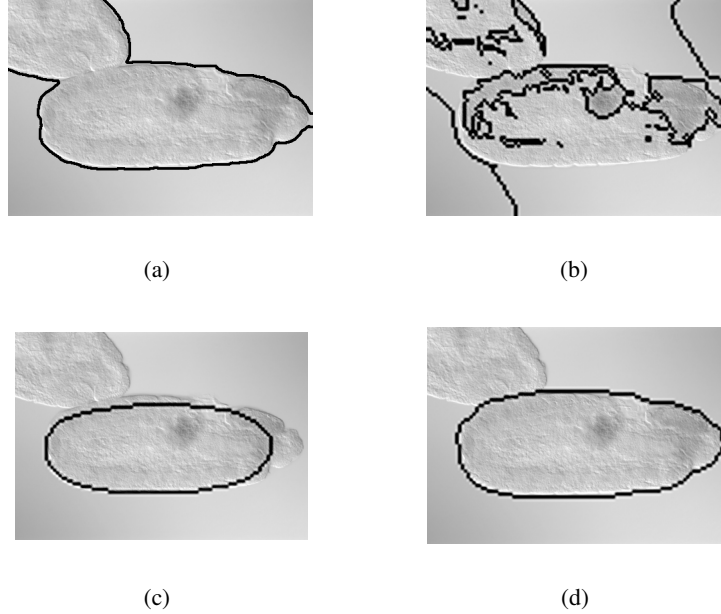

(a)

(b)

(c)

(d)

Figure 3: (a) MRF on superpixels. (b) Level set segmentation. (c) Level set segmentation and prior. (d) MRF with level set segmentation

All of the depicted results are obtained by the mean field maximum likelihood inference on the Markov random field. An interesting observation is that the level set segmentation does not always find the boundary entirely correctly, i.e. the set $\{i \in V : \phi(i) >= 0\}$ does not entirely match with the set $\{i \in V : y_i = 1\}$ found by the MRF maximum likelihood.

In Figure 5 and 6 we show results of our algorithm on our test and training images. In most of the cases, the edges can be found very precisely. In practice it turned out that the choice of the constant λ_1 needs to be carefully tuned, as it determines the influence of the MRF on the level set segmentation which should neither be too strong such that the shape prior distribution can "drive" the contour to the center, nor too small such that the level set algorithm does not get stuck on the "mean" prior. However the coupled model is much more robust to inappropriate choices of the constant as we observe that the set $\{i \in V : \phi(i) >= 0\}$ does not always entirely match with the set $\{i \in V : y_i = 1\}$ found by the mean field inference on the Markov random field. In fact, all results shown in Figures 6 are obtained by the MRF. In practice this means that if the level set segmentation does not find the perfect edge e.g. because of too strong priors, there is a chance that the superpixels will use the high-resolution information to yield the correct result.

Figure 7 shows in detail how the algorithm evolves over the iteration steps. In the first iteration the MRF selects parts of the picture with relatively darker intensity and high standard deviation. This particularly includes the half-embryo which sticks in from the right. After one iteration of the level set segmentation with prior we see that it focuses the attention to the center and away from the very strong intensity cue on the right. The MRF then incorporates this new information while joint effort of prior and intensity cues from the superpixel MRF pushes the contour to the boundary of the embryo until it converges.

8
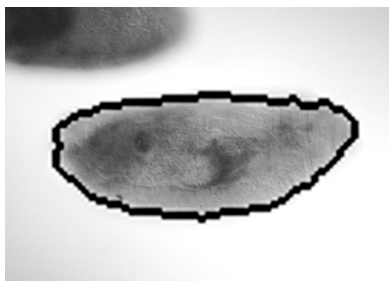
(a)
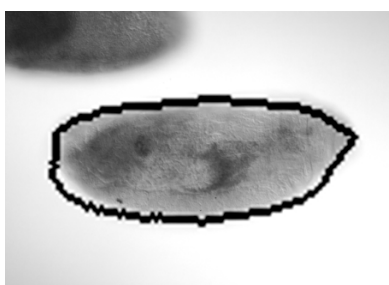
(b)

Figure 4: (a) shows the final result found by the MRF, (b) depicts the zero level set found by the level set segmentation algorithm
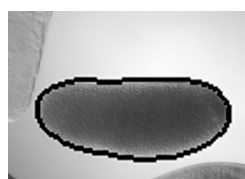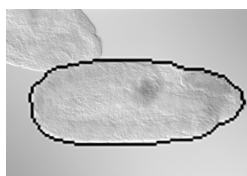

Figure 5: Segmentation results on the three training images used for the set of priors
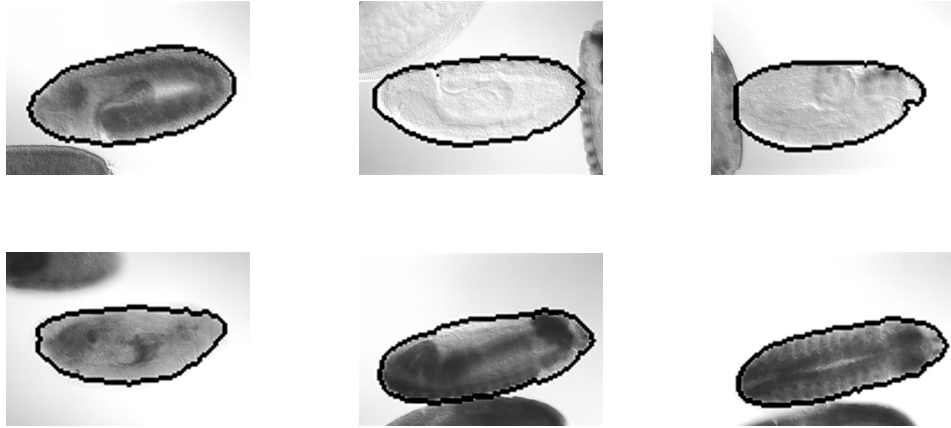
Figure 6: Segmentation results on 6 different test embryos

6 Summary and Outlook

We have proposed an algorithm combining the advantages of the Markov Random Field model applied on superpixel features, and the level set segmentation framework which allows an easy way to incorporate shape priors. It shows very accurate behavior in the segmentation of touching embryos for which the edge and intensity cues alone are not enough for correct segmentation.

The next step would be to test it on the actual data set for the segmentation of the gut. Preliminary results looked quite encouraging, but have to be validated on more data. The ultimate goal of simultaneous organ segmentation could then be tackled by providing a multilabel extension to this framework which works for $K = 2$ (because of the single level set function).

Given enough training data, it should also be possible to train the parameters μ, σ, θ beforehand so that the EM iteration can be boiled down to the E-Step alone, which is equivalent to the mean field approximation for the marginal $Q_M(y_i|x, a)$.

Furthermore it would be desirable to have a better feature set for which the Markov Random Field already gives better initial results to speed up the runtime of the algorithm. As a minor point we also need to find good stopping conditions for the level set part.

Last but not least it would be interesting to compare the performance of our hybrid model to the work in [10], [19].

7 Acknowledgements

I want to particularly thank Siqi Wu for his support, introducing me to the problem and helping out on the implementation part. I also want to thank Antony Joseph, Zac Zhang and Erwin Frise for valuable discussions on various results and problems which occurred using previous segmentation algorithms on this data set.

(a)
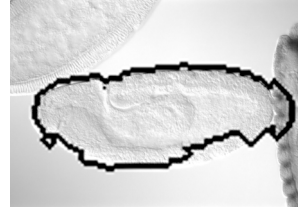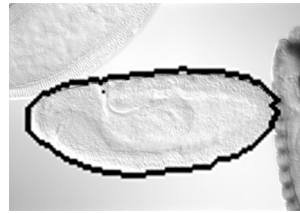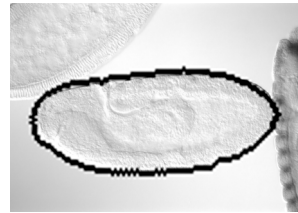

(b)

(c)

(d)

(e)

(f)

(g)

(h)

Figure 7: This series of pictures depicts the algorithm at different iteration steps for one image. Each row corresponds to one iteration where the left column depicts the MRF result and the right column depicts the level set segmentation result

References

[1] Tony Chan and Wei Zhu. Level set based shape prior segmentation. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 1164–1170. IEEE, 2005.

[2] Tony F Chan and Luminita A Vese. Active contours without edges. *Image processing, IEEE transactions on*, 10(2):266–277, 2001.

[3] Siqi Chen and Richard J Radke. Level set segmentation with both shape and intensity priors. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 763–770. IEEE, 2009.

[4] Timothy F Cootes, Gareth J Edwards, Christopher J Taylor, et al. Active appearance models. *IEEE Transactions on pattern analysis and machine intelligence*, 23(6):681–685, 2001.

[5] Timothy F Cootes and Christopher J Taylor. Active shape modelssmart snakes. In *BMVC92*, pages 266–275. Springer, 1992.

[6] Timothy F Cootes, Christopher J Taylor, David H Cooper, and Jim Graham. Training models of shape from sets of examples. In *BMVC92*, pages 9–18. Springer, 1992.

[7] Daniel Cremers, Stanley J Osher, and Stefano Soatto. Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *International Journal of Computer Vision*, 69(3):335–351, 2006.

[8] L.C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.

[9] Xuming He, Richard S Zemel, and MA Carreira-Perpindn. Multiscale conditional random fields for image labeling. In *Computer vision and pattern recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE computer society conference on*, volume 2, pages II–695. IEEE, 2004.

[10] Rui Huang, Vladimir Pavlovic, and Dimitris N Metaxas. A graphical model framework for coupling mrfs and deformable models. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–739. IEEE, 2004.

[11] Michael I Jordan, Zoubin Ghahramani, Tommi S Jaakkola, and Lawrence K Saul. An introduction to variational methods for graphical models. *Machine learning*, 37(2):183–233, 1999.

[12] Andrew Kae, Kihyuk Sohn, Honglak Lee, and Erik Learned-Miller. Augmenting crfs with boltzmann machine shape priors for image labeling. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2019–2026. IEEE, 2013.

[13] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331, 1988.

[14] John Lafferty, Andrew McCallum, and Fernando CN Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning, p.282-289, June 28-July 01, 2001*, 2001.

[15] Victor Lempitsky, Andrew Blake, and Carsten Rother. Branch-and-mincut: global optimization for image segmentation with high-level priors. *Journal of Mathematical Imaging and Vision*, 44(3):315–329, 2012.

[16] Michael E Leventon, W Eric L Grimson, and Olivier Faugeras. Statistical shape influence in geodesic active contours. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 316–323. IEEE, 2000.

[17] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685, 1989.

[18] Xiaofeng Ren and Jitendra Malik. Learning a classification model for segmentation. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 10–17. IEEE, 2003.

[19] Dmitrij Schlesinger. A continuous shape prior for mrf-based segmentation. In *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 350–361. Springer, 2013.

[20] Martin J Wainwright and Michael I Jordan. Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning*, 1(1-2):1–305, 2008.