Segmentation of Developing Human Embryo in Time-lapse Microscopy

Aisha Khan[1]*, Stephen Gould*[1]

[1]College of Engineering and Computer Science
The Australian National University
Canberra, AU
{aisha.khan, stephen.gould}@anu.edu.au
aisha.sj.khan@gmail.com

Mathieu Salzmann[1,2]

[2]CVLab
EPFL
Switzerland
mathieu.salzmann@epfl.ch

ABSTRACT

Being able to efficiently segment a developing embryo from background clutter constitutes an important step in automated monitoring of human embryonic cells. State-of-the-art automatic segmentation methods remain ill-suited to handle the complex behavior and morphological variance of non-stained embryos. By contrast, while effective, manual approaches are impractically time-consuming. In this paper, we introduce an automated approach to segment human embryo in early-stage development from a sequence of dark field microscopy images. In particular, we express segmentation as an energy minimization problem, which can be solved efficiently via graph-cuts or dynamic programming. Our experiments on twenty embryo sequences demonstrates that our method can successfully segment complex and irregular embryo structures in time-lapse microscopy (TLM) sequences.

1. INTRODUCTION

The success of *in vitro* fertilization (IVF) treatment is relatively poor (depending on the woman's age, only 10–30% of implanted embryos result in a successful pregnancy). This is mainly due to the lack of reliable methods to select viable embryos. Traditionally, embryo selection relies on manual morphology analysis and is subject to inter and intra observer variance [2]. By contrast, in VerMilyea et al. [16], it was shown that computer-automated time-lapse analysis could improve embryo selection by providing quantitative and objective information to supplement manual analysis, and could therefore increase the success rate of IVF.

Automated analysis involves detection, tracking and classification of large volumes of cellular image data. A major requirement for these tasks is an efficient method to segment embryo images. The segmentation step is critical because it serves as a basis for all subsequent tasks, such as the extraction of shape features, and ultimately the viability assessment of the embryo. In this paper, we tackle the problem of fully automated segmentation (i.e., contour extraction) of non-stained developing human embryos in TLM images.

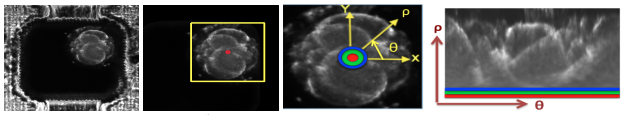
The authors thank Auxogyn, Inc. for their valuable support.

The difficulty in extracting the contour of an embryo arises from various artifacts: irregular embryo shape, weak or missing embryo boundaries, fragments attached and internal to the embryo, intensity and texture variations in foreground, background and fragments, continual contrast variation of the embryo boundary due to motion and poor image quality.

While cell segmentation has attracted a lot of attention, these difficulties make most standard techniques inapplicable to the human embryo case. For example, threshold-based methods [12] cannot cope with strong background variations, and fail as soon as one gray-value can belong to both foreground and background. The complex appearance of the embryonic cells limits the success of region-based techniques, such as watersheds-based methods [19]. Other techniques such as active contours [18] and level sets [20] are more suitable, but the large amount of clutter and artifacts in the image cause them to easily get trapped in local minima. This also hinders the use of simple edge-based algorithms [17], since many spurious contours are detected. To overcome these issues, most of the above-mentioned methods work with fluorescent stained cells. For human embryonic cells, however, such a staining procedure cannot be used.

Human embryonic cell segmentation involves noisier data and more complex structures to segment, such as multiple highly-overlapping cells. Several directions have nonetheless been investigated to address these challenges, such as using different image modalities (e.g., Hoffman Modulation Contrast [4]), alternative acquisition procedures (e.g., multiple focus planes [5]) and simpler assumptions (e.g., zona pellucida segmentation [6]). The resulting methods, however, rely on non-standard acquisition procedures which are not widely available. Furthermore, most techniques are semi-automated [3]. Recently, Markov random field based methods [7, 9, 13] were proposed to detect and localize individual cells. While these methods make use of more standard images, and would thus generalize more easily, they rely on an initial embryo segmentation to generate cell hypotheses. Improving this initial step would therefore be highly beneficial.

Currently, one of the most effective methods to segment



(a) (b) (c)

Fig. 1. Image pre-processing. (a) Microscopy image of a four-cell embryo. (b) Centroid and bounding-box. (c) Polar transformation.

a human embryo from microscopy images was introduced in Giusti et al. [4] and relies on the graph-cuts algorithm [1]. This method, however, was designed to segment zygotes (i.e., one-cell embryos), and thus relies on fairly simple shape priors. By contrast, here, we address the problem of segmenting multi-cell human embryos. As an embryo grows beyond the one-cell stage, its shape becomes very irregular. Furthermore, the individual cells form a complex 3D structure, which, in a 2D projection, overlap immensely. As a consequence, in an image, the cell membranes may cross and cause bright contours within the embryo. Similarly, the interior of the embryo can be greatly non-homogeneous and contain intensities similar to those of the background. Finally, fragments with texture and intensity similar to that of the embryo often attach to the embryo boundary.

In this paper, we introduce shape priors and contextual cues specifically designed to address the challenges of multi-cell human embryo segmentation. We then incorporate these priors as soft and hard constraints both in a graph-cut and Markov chain inference framework. We demonstrate the effectiveness of our approach and compare it against the state-of-the-art work of Giusti et al. [4] on a set of twenty sequences of developing embryos.

2. IMAGE PRE-PROCESSING

Given a dark field microscopy image depicting a human embryo in early-stage development, we perform the following pre-processing steps. First, we automatically find a bounding box that roughly encloses the embryo. To this end, we convert the gray-scale image into a binary image using Otsu's threshold [14]. Since pixels inside the embryo can have intensities similar to those of background, the resulting binary image can contain holes within the foreground region. We fill these holes by using the flood-fill technique of [15], which connects the nearby disconnected components. We then take the largest connected region to be the embryo, since each image only contains one embryo, and compute its centroid as the point within the component with maximum shortest distance to the region boundary. We also extract a bounding-box around the region, which excludes a large part of the background, as well as many debris and fragments, from further segmentation (see Fig. 1(a)–(b)).

Within this bounding-box, we reduce noise by applying a median filter, which smoothes the image while preserving the edges. The dark field modality of our images and the nature of embryo growth (i.e., compactness of the cells) results in the additional challenge that the interior of the embryo can have both very low intensities and very high ones due to some cell membranes projecting within the embryo via the imaging process. This makes it difficult to differentiate the true embryo contour from these high-intensity interior membranes. To reduce this problem we apply non-linear intensity mapping. Specifically, we use the power law ($s = cI^{\gamma}$, where I is the intensity image, and c and γ are positive constants). A fractional value of gamma ($\gamma = 0.04$ in our experiments) maps a narrow range of dark input values to a wide range of output values, and conversely for high input values.

Following the observation of previous mask-generating methods [4, 10, 11] that images with radial symmetry should be converted into non-Cartesian representations before image processing, we transform the image to polar coordinates (see Fig. 1(c)). Below, given this image representation, we introduce our approach to segmentation and our shape priors.

3. EMBRYO SEGMENTATION

Segmentation can be formulated as a pixel labeling or contour drawing problem. Here, we study both approaches under a Markov random field (MRF) formalism.

Pixel Labeling Formulation: First we formulate embryo segmentation as a binary labeling problem. For each pixel *i* in a given image in polar coordinates (after the pre-processing of Section 2), we define a random variable y_i taking value from the label space $\mathcal{L} = \{0, 1\}$. We then construct a graph $G = \langle V, E \rangle$ with vertices V representing the pixels and edges E connecting the neighboring vertices. In contrast to Giusti et al. [4], our graph defines bidirectional edges with an eight-neighborhood structure (see Fig. 2). Additional edges are defined between the first and last column of the polar image to ensure smoothness when the segmented image is converted back to Cartesian coordinates.

Given this graph, the distribution over the joint assignment of all random variables Y is defined by an MRF, whose energy function can be written as

$$E(Y) = \sum_{i \in V} \psi_i(y_i) + \lambda \sum_{i,j \in E} \psi_{i,j}(y_i, y_j), \qquad (1)$$

where the unary (i.e., local) term ψ_i is a prior encoding the cost of assigning pixel *i* to label y_i, the pairwise (edge) term maps joint variable assignments to a cost (in our work this assigns a contrast dependent penalty whenever the pair of variables disagree), and λ is a weighting factor determined using a validation set. A pixel labeling, and thus embryo segmentation, is achieved by finding an assignment to Y that minimizes the energy (Eqn. 1). Here, we design potentials that allow us to rely on graph-cuts to perform this minimization efficiently.

In particular, we obtain the prior ψ_i from training data by computing the histogram of occurrence of a pixel being foreground in frame *t*. In other words, this prior penalizes the assignments too far away from the training ground-truth. Furthermore, we also design hard-constraints for seed pixels strongly believed to be either foreground or background. These constraints can be expressed in a unary potential as

$$\psi_i(y_i = 1) = \begin{cases} -\infty, & \text{for } i = \text{foreground seed} \\ +\infty, & \text{for } i = \text{background seed.} \end{cases} \qquad (2)$$

To automatically choose the seeds, we rely on the following observations, illustrated in Fig. 2. First, in polar coordinates, the top row of the image (i.e., the Cartesian image boundary) is always background. Second, the lower part of the image (i.e., a disk around the centroid in the Cartesian image) always belongs to the embryo and should thus be foreground. Note that the latter observation also allows us to be robust to the bright contours that, as mentioned before, appear within the embryo because of the projection of the 3D embryonic structure to a 2D image plane, or because of the presence of fragments and pronuclei inside the embryo (see Fig. 3 (b)–(h)). The width of the band that we force to be assigned to foreground is computed from the training data as follow. We first mark the pixels that belong to foreground at time *t* for the complete training data and select the band width as the location of the marked pixel closest to the centroid.

For the pairwise term $\psi_{i,j}$, we rely on the fact that contours in dark field images are most likely to coincide with large changes of intensity. To capture this, we define the edge cost as

$$\psi_{i,j}(y_i, y_j) = \begin{cases} \frac{1}{w_{ij}} e^{\frac{-\|x_i - x_j\|^2}{2\zeta^2}}, & \text{if } y_i \neq y_j \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where x_i is the intensity of pixel *i* and ζ is the mean intensity difference between adjacent pixels. In other words, our edge cost penalizes neighboring vertices to take on different labels if they have similar intensity. In Eqn. 3, w_{ij} accounts for the spatial (Cartesian) distance between neighboring pixels, such that closer pixels have more influence. In polar coordinates, this weight can be computed as

$$w_{ij} = \sqrt{(\rho_i^2 + \rho_j^2 - 2\rho_i\rho_j \cos(\theta_j - \theta_i))}\,, \quad (4)$$

where ρ is the distance from the origin to the point and θ is the counterclockwise angle relative to the x-axis.

With the definitions of our potential given above (in particular the pairwise potential), it can easily be verified that the energy of Eq. 1 can be minimized with graph-cuts. In practice, we use the efficient max-flow implementation of [1], which gives us the optimal labeling in polynomial time.

While our data consists of sequences, the previous potentials work on individual images. Applying this technique independently to each frame may result in inconsistencies because, even though embryos in consecutive frames have similar appearance, motion makes the contrast between the cell boundary and background vary. To overcome this, when segmenting one frame, we combine shape and intensity information from its neighboring frames. To this end, we first register the neighboring frames to the current frame using the Matlab functions (`imregtform()`, `imwarp()`). We then compute the average image after registration, and perform segmentation on this average image. As evidenced by our results, this strategy has proven robust to overcome temporal inconsistencies.

Contour Extraction Formulation: Embryo segmentation can also be framed as a contour extraction problem and

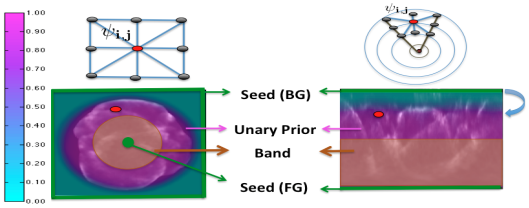


Fig. 2. Graph neighborhood structure, unary prior heatmap and topological constraints in Cartesian (left) and polar (right) coordinates.

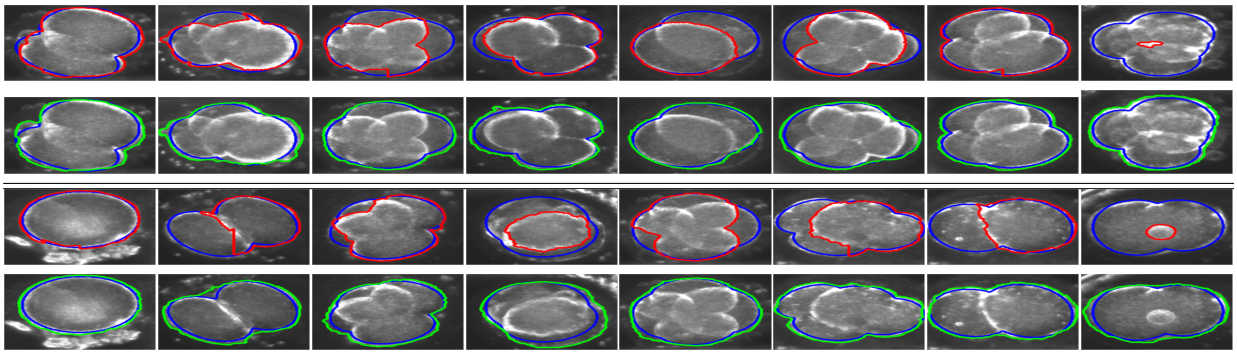
can be formulated as inference in a Markov chain by dynamic programming. A simple change of variables from the above formalism allows us to achieve this. More specifically, instead of defining one binary random variable for each pixel, we can make use of one discrete (but non-binary) random variable per column in the polar image. Such random variables take labels from the set $\mathcal{L} = \{1, \ldots, R\}$, where R represents the number of rows in the polar image. In other words, and considering the meaning of the columns and rows in the polar image, for each angle, we search for the distance to the embryo boundary. In this formulation, we define the unary term ψ_i as the absolute intensity difference between neighboring pixels in column i, which captures the evidence of a pixel being part of the contour. The pairwise term $\psi_{i,j}$ encourages spatial smoothness of the contour by penalizing sudden changes in the contour location (i.e., $\psi_{i,j} = |y_i - y_j|$). The seed constraints and temporal image averaging defined above easily transfer to this formulation.

4. EXPERIMENTAL RESULTS

We evaluated the proposed approach on twenty time-lapse image sequences of developing embryos consisting of a total of 7,000 frames. The images were captured with the integrated time-lapse imaging System $Eeva^{TM}$ developed by Auxogyn, Inc. The system fits into an incubator and includes a dish that holds the embryos. The image acquisition software captures a single-plane image once every five minutes. The sequences capture the embryos of six different patients and show a certain degree of variation, such as fragments and missing boundaries. To obtain the ground-truth masks, we manually segmented all 7,000 frames.

We report results obtained using the following variants of our method: i) graph-cuts with (topological) band constraints (GC) (Eqn. 2), ii) GC with band and unary term (GC+U), and iii) GC with band, unary and temporal smoothness (GC+U+S). We compare these variants against the following methods: i) Giusti et al. [4], ii) Giusti et al. [4] with Eqn. 3 as edge cost (Giusti et al. [4]+Enr), iii) Giusti et al. [4] with topological band constraints (Giusti et al. [4]+Band), iv) our chain MRF formulation (Chain), and (v) GC with image and constraints in Cartesian coordinates (GC+U+S(C)) .

To compare these methods, we report the following error metrics. Area of overlap (AoL): intersection over union with ground-truth; True positive rate (TPR): intersection with ground-truth over ground-truth; False negative rate (FNR):

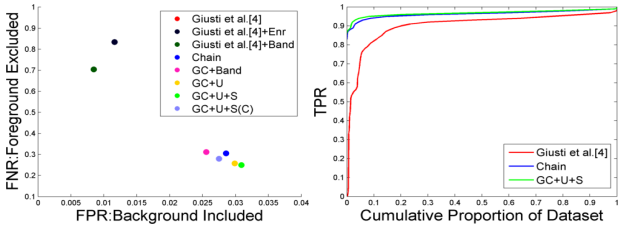


(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 3. Embryo segmentation results: Giusti et al. [4] (red contour), GC + U+ S (green contour) and ground-truth (blue contour).

Methods	AoL	FPR	FNR	ME	Pred.N
GC	0.9494	0.0022	0.0273	0.0148	83.10
GC+U	0.9502	0.0026	0.0226	0.0126	82.18
GC+U+S	0.9500	0.0027	**0.0219**	**0.0123**	**84.85**
Chain	0.9481	0.0024	0.0273	0.0148	83.33
GC+U+S(C)	**0.9504**	0.0024	0.0245	0.0135	81.36
Giusti et al. [4]	0.9063	**0.0006**	0.0877	0.0441	71.00

Table 1. Methods evaluation: Average AoL, average mean error (mean of avergae FPR and average FNR) and prediction on number of the cells [8] (overall %). Polar image size is 52×210 and Cartesian image size is 100×100.



(a) FPR vs. FNR (b) TPR cumulative distribution

Fig. 4. Quantitative evaluation.

excluded foreground over ground-truth; and False positive rate (FPR): included background over background. Furthermore, to evaluate the impact of segmentation on further embryo analysis, we use the segmentation results of the different algorithms as input to our previous work Khan et al. [8], which predicts the number of cells in each frame. We then report the cell stage prediction accuracy, i.e., the percentage of frames where the correct number of cells was predicted (Pred.N).

Table 1 compares the results of all the algorithms. We can see that, with the exception of FPR, all variants of our approach perform better than the method of Giusti et al. [4]. In many applications, however, and in human embryo analysis in particular, FNR is typically more important than FPR. Indeed, if a cell is removed by the segmentation process, it will be excluded from further analysis, which would affect the embryo selection. In Fig. 4(a), we focus more specifically on these two measures. Note that the method of Giusti et al. [4] yields a high FNR, which visual inspection revealed was due to the method's sensitivity to high intensity contours appearing within the embryo. While introducing band constraints and the edge cost of Eq. 3 in the method of Giusti et al. [4] reduces this error, it remains higher than that of our approach. In Fig. 4(b), the analysis of the TPR shows that both our approach and Chain also outperform Giusti et al. [4] using this metric. In particular, Giusti et al. [4] failed to correctly segment several embryos (TPR < 0.8), which visual inspection revealed was also due to the presence of contours within the embryo, or of bright spots and pronuclei. Fig. 3 shows some of these cases.

Among the different variants of our method, we can see that the error is reduced by adding a unary term and using temporal smoothness (GC+U+S). The alternative formulation, Chain, however, yields results similar to our basic GC. Furthermore, performing graph-cuts in Cartesian coordinates also yields slightly higher errors than our polar-coordinate approach. We conjecture that this is due to the different neighborhood structures induced by these two approaches. We leave a more thorough analysis of the effect of neighborhood structure on segmentation for future work.

Finally, and importantly, the last column of Table 1 shows the importance of having good segmentations for further embryo analysis. This result clearly evidences that our approach leads to much better prediction of the number of cells that of Giusti et al. [4], with an improvement of 13.8%.

5. CONCLUSION

Embryo segmentation is crucial for further image analysis, and, ultimately, to be able to select viable embryos in IVF. In this work, we have introduced a graph-cuts based approach to segmenting a developing human embryo in time-lapse microscopic images. In particular, we have introduced a shape prior that lets us overcome the noise and artifacts of dark field embryo images. Our results have shown that good segmentation can only be achieved if sufficient prior knowledge about the shape of the embryo is taken into account. We have also demonstrated that better segmentation results could improve subsequent analysis, such as cell number prediction. In the future, we intend to study the impact of our results on other tasks, such as cell localization and tracking, as well as cell lineage extraction.

References

[1] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2004.

[2] A. A. Chen, L. Tan, V. Suraj, R. R. Pera, and S. Shen. Biomarkers identified with TL imaging: discovery, validation, and practical app. *Fertility and Sterility*, 2013.

[3] E. S. Filho, J. Noble, and D. Wells. A review on automatic analysis of human embryo microscope images. 2010.

[4] A. Giusti, G. Corani, L. M. Gambardella, C. Magli, and L. Gianaroli. Lighting-aware segmentation of microscopy images for in vitro fertilization. In *ISVC (1)*, 2009.

[5] A. Giusti, G. Corani, L. Gambardella, C. Magli, and L. Gianaroli. Blastomere segmentation and 3d morphology measurements of early embryos from hoffman modulation contrast image stacks. *ISBI*, 2010.

[6] J. Hoey and S. McKenna. Automatic Segmentation of Zona Pellucida in Human Embryo Images Applying an Active Contour Model. In *Medical Understanding and Analysis*, 2008.

[7] A. Khan, S. Gould, and M. Salzmann. A linear chain markov model for detection and localization of cells in early stage embryo development. In *WACV*, 2015.

[8] A. Khan, S. Gould, and M. Salzmann. Automated monitoring of human embryonic cells up to the 5-cell stage in time-lapse microscopy images. In *ISBI*, 2015.

[9] A. Khan, S. Gould, and M. Salzmann. Detecting abnormal cell division patterns in early stage human embryo development. *6th International Workshop on Machine Learning in Medical Imaging (MLMI)*, 2015.

[10] M. A. Luengo-Oroz and J. Angulo. Cyclic mathematical morphology in polar-logarithmic representation. *Image Processing, IEEE Transactions on*, 2009.

[11] M. A. Luengo-Oroz, J. Angulo, G. Flandrin, and J. Klossa. Mathematical morphology in polar-logarithmic coordinates. application to erythrocyte shape analysis. In *Pattern Recognition and Image Analysis*. Springer, 2005.

[12] E. Meijering, O. Dzyubachyk, I. Smal, and W. A. van Cappellen. Tracking in cell and developmental biology. *Seminars in Cell and Developmental Biology*, 2009.

[13] F. Moussavi, W. Yu, P. Lorenzen, J. Oakley, D. Russakoff, and S. Gould. A unified graphical models framework for automated mitosis detection in human embryos. *IEEE Trans. Med. Imaging*, pages 1551–1562, 2014.

[14] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 1975.

[15] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer-Verlag New York, Inc., 2003.

[16] M. D. VerMilyea, L. Tan, J. T. Anthony, J. Conaghan, K. Ivani, M. Gvakharia, R. Boostanfar, V. L. Baker, V. Suraj, A. A. Chen, et al. Computer-automated time-lapse analysis results correlate with embryo implantation and clinical pregnancy: A blinded, multi-centre study. *Reproductive biomedicine online*, 2014.

[17] Q. Wu, F. Merchant, and K. Castleman. *Microscope Image Processing*. Academic Press, 2008.

[18] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 1998.

[19] P. Yan, X. Zhou, M. Shah, and S. T. Wong. Automatic segmentation of high-throughput rnai fluorescent cellular images. *Information Technology in Biomedicine, IEEE Transactions on*, 2008.

[20] H.-K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multiphase motion. *Journal of computational physics*, 1996.