

3D Localization of Pronuclei of Human Zygotes Using Textures from Multiple Focal Planes

A. Giusti^{1,*}, G. Corani¹, L. Gambardella¹, C. Magli², and L. Gianaroli³

¹ Dalle Molle Institute for Artificial Intelligence (IDSIA), Lugano, Switzerland

² International Institute for Reproductive Medicine (IIRM), Lugano, Switzerland

³ INFERGEN, Lugano, Switzerland

Abstract. We propose a technique for recovering the position and depth of pronuclei of human zygotes, from Z-stacks acquired using Hoffman Modulation Contrast microscopy. We use Local Binary Pattern features for describing local texture, and integrate information from multiple neighboring areas of the stack, including those where the object to be detected would appear defocused; interestingly, such defocused areas provide very discriminative information for detection. Experimental results confirm the effectiveness of our approach, which allows one to derive new 3D measurements for improved scoring of zygotes during In Vitro Fertilization.

1 Introduction

During In Vitro Fertilization (IVF), *fertilization* is the union of a sperm with an oocyte; the resulting *zygote* contains genetic material (DNA) from both the father and the mother. Usually, several zygotes per patient are successfully fertilized, but only a limited number can be cultured (no more than three in many countries). It is therefore necessary to select the zygotes to be cultured. Currently, embryologists perform zygote selection on the basis of a morphological evaluation of the zygotes [4], which are observed through an optical microscope equipped with Hoffmann Modulation Contrast (HMC) optics.

Several scoring systems have been designed in order to recognize zygotes with high potential of implantation; yet, none seems to be strong enough to reliably predict whether an IVF cycle results in a pregnancy [11]. Within such scoring systems, the relative position of the *pronuclei* (see Fig. 1) is a factor of major importance, but so far, only 2D information has been available for measurement. Yet, the zygote has a spherical shape and thus represents a thick sample with respect to the microscope depth of field: it is therefore possible to acquire several images at different, equally-spaced focus planes. This is a common practice in automated microscopy, and the resulting set of images is named focus stack (or Z-stack): in each of the images of the stack, structures lying at the corresponding depth appear in-focus, whereas other structures appear blurred.

* Contact: alessandro@idsia.ch – Additional material at <http://www.idsia.ch/~giusti/IN3>

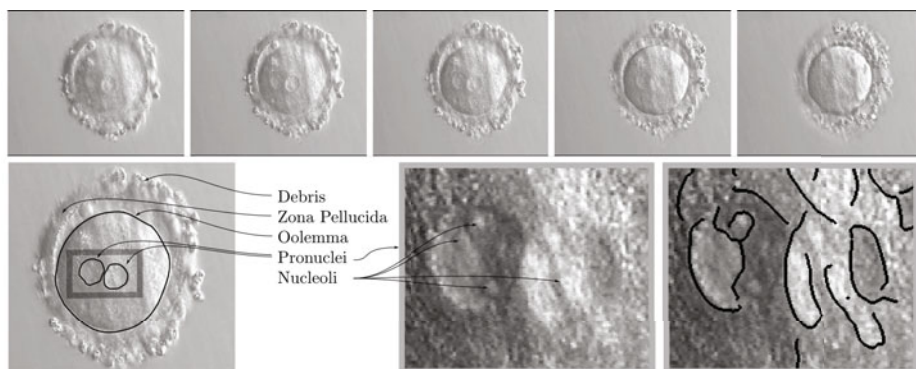


Fig. 1. A Zygote. *Top row:* 5 of the 24 images in the source focus stack. Note that different structures are visible in different images, depending on their depth. *Bottom:* parts of the zygote, and detail of pronuclei appearance, when they are both in sharp focus: this is possible because they lie at the same depth in this specific case. *Right:* edge detection, even with manually-tuned parameters, fails due to weak borders.

In this paper we present a technique to localize pronuclei in a focus stack, determining both their 2D position and depth; this allows us to perform a partial 3D reconstruction of the zygote morphology, from which a larger and more consistent array of measurements can be performed.

Detecting pronuclei in HMC image stacks is challenging because of their wide appearance variability (see Figure 2) and large amounts of clutter. We exploit Local Binary Patterns (LBP) [7] as a texture feature for pronuclei detection, thus leveraging on the peculiar (albeit complex and unpredictable) appearance of the pronucleus' interior. Additionally, our detector integrates information from several neighboring focal planes: our experimental findings show that even texture data from focal planes where the pronucleus appears defocused provide informative features.

Images from human embryology have only recently been subject to attention from the image processing community, because of their complex appearance, and the limited need of processing large amounts of data, as samples were observed only once per day. The availability of new systems integrating microscopes in incubators [10] will produce much more frequent observations, and increase the need of automated processing: moreover, significant advantages of performing objective, quantitative measurements have been shown by Beuchat et al. in [2], where various 2D zygote features, including pronuclei positions and sizes, are measured from a single image (where all such features are assumed to be in focus) with user assistance.

In [6,5] we presented algorithms for processing image stacks, with the goal of segmenting the outer shape of the zygote, or the blastomeres in an embryo. The present work fits in the same approach of also determining the object's depth from the image stack. Unfortunately, pronuclei lie at unpredictable positions and depth, and do not provide sufficiently visible outer edges, nor apparent

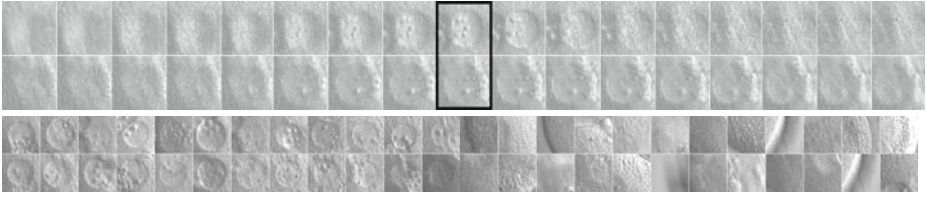


Fig. 2. *Top:* Defocused appearance of pronuclei (left and right), vs best focused plane (center). *Bottom:* positive (left) and negative (right) examples.

area features; this prevents the direct application of segmentation techniques such as active contours [3] or level sets [12], unless the pronuclei 3D location is first determined, which is the problem we are dealing with in this paper. Then, having determined a reasonable initialization, a simple snake-based approach can provide a precise segmentation and size measurement, exploiting weak large-scale gradients through strong shape priors.

Several recent works deal with the detection of cell nuclei or other subcellular structures; the problem is radically different in fluorescent-microscopy approaches or when nuclei are stained, both of which are not viable options for human zygote observation; ray features [13] have recently proven useful in capturing the irregular shapes of neuron nuclei. In our scenario their application would not be justified, as pronuclei lack clear edges but have regular shapes.

2 Texture Features at Multiple Focal Planes

HMC is a complex optical microscopy technique, which visualizes differences in optical density as variations in light intensity: it is routinely used in IVF labs for observing zygotes, as it provides a large amount of contrast for transparent specimens and eases human observation as the objects appear three-dimensional and side-lit, as if a virtual light source was illuminating them from a side (apparent lighting direction). In particular, images of pronuclei share the following appearance traits.

1. Pronuclei do not generally exhibit clear edges at any focal plane, which limits the utility of approaches based on edge detection – such as hough transform, which would perform well due to the predictable circular overall shape; they also lack a predictable large-scale appearance, which could be captured by haar-like features.
2. The interior of pronuclei has a peculiar, subtle texture when in focus; moreover, nucleoli are often visible, but have a very variable size, distribution and appearance in the pronucleus.
3. Different zygote images are subject to large photometric variations, even when imaging parameters are unchanged.
4. Zygote images are often cluttered by large amounts of high-contrast debris, as well as vacuoles and other structures mimicking the large-scale pronuclei appearance.

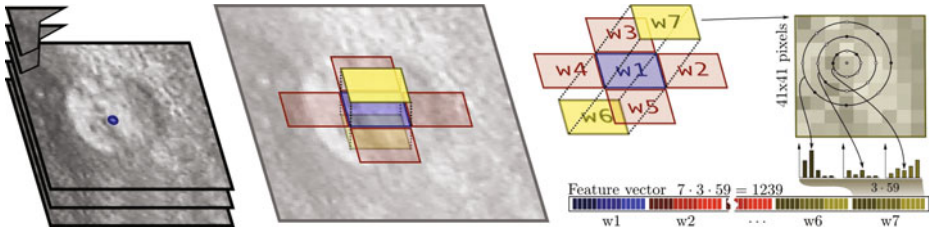


Fig. 3. Composition of the feature vector, which is obtained from 8-neighbor LBP features computed in 7 different windows at 3 scales

We use Local Binary Patterns (LBP) features for dealing with this challenging scenario [7]. Such descriptor is amongst the most powerful options for texture analysis and has the advantage of simple implementation and fast performance, and has been successfully used in several biomedical applications [14,8]. LBP features allow us to represent the characteristic interior texture of the zygote, and, as the descriptor is histogram-based, it captures the presence of nucleoli regardless of their actual position in the pronucleus (point 2); moreover, the invariance of the descriptor to grayscale transformations make it robust to the marked unpredictability in photometric characteristics between images described in point 3 above.

In order to capture larger-scale, structural characteristics of pronuclei, we concatenate LBP descriptors computed on multiple neighboring windows; a similar approach has been used for face description [1]. For a given point $p = (x, y, z)$, we compute a feature vector by considering data from 7 different 41×41 windows of the image stack. Window w_1 is centered on p , whereas windows w_2 to w_5 are positioned around it (see Figure 3). We also consider windows w_6 and w_7 , centered on the same (x, y) position as p , but containing data from focal planes below and above p , respectively. In each window, we compute LBP descriptors using 8 neighbors and uniform mapping, which originates a 59-bin histogram; features are computed for 3 different scales, at radii 1, 2 and 3. Feature vectors for each of the 3 scales at each of the 7 windows are concatenated, thus creating a global feature vector for point p composed by $7 \cdot 3 \cdot 59 = 1239$ variables, which summarizes data from multiple focal planes. We will then use statistical classification to predict, on the basis of these 1239 variables, whether point p belongs or not to a pronucleus.

Extensions of LBP features to a third dimension have been successfully proposed for several applications [17], mostly dealing with dynamic content. One common approach is to consider the texture characterizing $x-t$ and $y-t$ space-time slices (LBP-TOP). In our evaluations, texture in $x-z$ or $y-z$ slices proved uninformative in dealing with focus stacks, regardless of the scale factor along the z axis; this is probably due to marked anisotropy in such slices because of the effects of defocus: therefore, we only include descriptors for $x-y$ planes in our final feature vector.

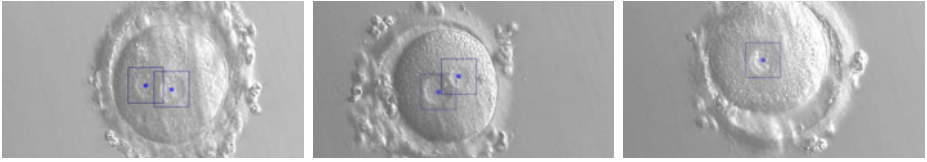


Fig. 4. Detection results

3 Experiments

We experimented our technique on 193 focus-stacks representing zygotes, acquired during routine IVF observations using an IX51 microscope equipped with 20x HMC optics and a 720x576 camera. Each stack is composed by 24 frames, acquired at focal planes with a z spacing of roughly $5\mu m$. Zygotes and pronuclei have an apparent diameter of approximately 300 and 50 pixels, respectively. The approximate position and depth of pronuclei in each image stack is labeled by an operator, by means of an ad-hoc focus-stack viewer. In total, our dataset contains 347 labeled pronuclei (1.8 per image stack on average).

In order to detect pronuclei in a new image, feature vectors are computed for each point on a 5×5 grid, on all the images in the stack; the process is not computationally demanding, because LBP codes can be precomputed on the whole image. Processing can be significantly sped up by limiting the evaluation to points inside the zygote oolemma, which can be robustly found using the technique in [6], thus limiting to 50 seconds the time for processing a full stack in our MATLAB implementation. Each point is classified for the presence of a pronucleus by means of the classifier described in the following section. Positive responses, which constitute a small fraction of the tested points, are used to initialize snake-based detectors; this automatically handles multiple positive responses from a single pronucleus, as multiple snakes initialized near a pronucleus converge and finally overlap.

From the previously described labeled data, we built a data set of 347 positive examples (i.e., points belonging to the pronucleus) and 11927 negative examples, which are centered on random points farther than $15\mu m$ from any pronucleus center.

Classification. First, we split the data into two halves, each containing some 6000 instances; we used the *selection set* to perform feature selection and to select the classification algorithm; we used the *test set* to eventually measure the performance of the chosen classifier. We performed all classification experiments using WEKA [16].

We started by computing the information gain (IG) of each feature; the IG is a standard index of predictivity for features: better features have a higher IG. As shown in Figure 5 (central plot), the distribution of IG is highly skewed,

Keeping the test set separate from the data used for selecting features and classification algorithm is necessary to avoid optimistic measures of accuracy [9].

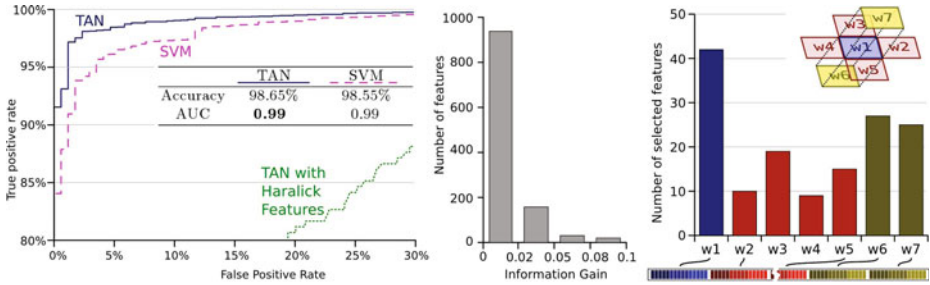


Fig. 5. *Left:* ROC curve, accuracy and area under curve (AUC) measured on the test set. *Center:* distribution of the IG across the 1239 features. *Right:* number of selected features in each window, after IG analysis.

namely few features have high IG, while most features have a very low one. We retained the features having $IG > IG_{\max}/3$, where IG_{\max} is the maximum IG among all features. In this way, we selected 152 features; the average IG of the selected features is about five times larger than the average IG of the non-selected features. Most selected features are from regions $w1$, $w6$ and $w7$, as shown in Figure 5 (right plot). This shows that focal planes where the pronucleus appears defocused ($w6$ and $w7$) provide useful texture information.

The IG analysis, although useful for a preliminary screening, analyzes only one feature at a time and cannot capture interaction between features. To identify and remove correlated features, we then used the correlation-based feature selector (also available within WEKA), further reducing the features to 45.

By experiments on the selection set we select SVM with linear kernel and TAN (Bayesian network with tree-augmented structure) as the best performing classifiers; they are slightly but significantly better than ADABOOST with decision stumps [15]. Eventually, we trained the classifiers on the selection set and evaluated them on the test set, measuring for both accuracy of 99% and AUC (area under the ROC curve) higher than 0.99.

Further Classification Experiments. An usually desirable property of texture features is rotation invariance, which can be obtained in LBP features by using a different mapping between LBP codes and histogram bins [7]. However, rotation-invariant features are much less powerful than those obtained with uniform mapping; the IG of the best 10 rotation-invariant features is on average 2.5 times smaller than that of the best 10 features with uniform mapping. Moreover, classification accuracy drops to 96% (both for SVM and TAN) when rotation invariant features are used; such an accuracy corresponds to that of a trivial classifier, which always returns the label “negative”.

We compared the LBP-based approach described in this paper to Haralick texture features computed in a single 40×40 window around the candidate point; features are computed on normalized intensities using 8 equally-spaced intensity bins, and considering two orthogonal displacements; such naive approach yielded significantly worse results in terms of both IG (about four times smaller for the

10 best Haralick features compared to best 10 features with our LBP-based descriptor) and classification accuracy (which drops to 96%); figure 5 (left) reports a dotted line for the resulting ROC curve.

We investigated to what extent the high accuracy of the classifiers shown in Figure 5 was due to the unbalance of the data set; in fact, a classifier always returning the class “negative” would have achieved an accuracy of 96%. We therefore built a data set of some 700 examples, evenly distributed between positive and negative. Both TAN and SVM achieved 95% accuracy, thus confirming the high predictivity of LBF from multiple focal planes (recall that in this case, simply returning “negative” would only achieve 50% accuracy).

Localization and Segmentation. In order to precisely localize pronuclei and determine their size, from each positively-classified point we initialize a single circular active contour, with three DOF for position and one additional DOF for size. Such naive low-dimensional model is evolved using gradient descent in order to minimize an energy measure, determined from large-scale image gradients near its border, which are expected to have directions consistent with the apparent lighting due to HMC, as described in [6]. Candidate pronuclei are determined by solutions not overlapping in 3D space more than 30% of their volume, and amount to an average of 3.4 per image.

We consider a candidate being correctly segmented if three conditions are met: (a) its center is within $3 \mu\text{m}$ (10 pixels) of the manually-labeled center; (b) its depth less than 2 stack slices away from the labeled depth – which is the depth where the pronucleus edges appear sharpest; (c) its radius not differing more than $2 \mu\text{m}$ from the actual one. We visually determined that circular models meeting these constraints manage to exactly fit the pronucleus shape in all instances, and routinely exceed the precision of manual labeling of the center of pronuclei. This prevents us from quantifying localization error, as exact ground truth is not available.

In our dataset, 93.3% of pronuclei are correctly segmented and appear in the final list of candidates; in our present implementation, candidates are displayed to users during routine evaluation of the zygote image stacks, and are manually confirmed with a single click on each – after which the related position, depth and size measurements are saved in the database and used as additional features for statistical evaluations. Naively selecting the two lowest-energy candidates provides correct results for both pronuclei in 84.0% of our image stacks, whereas in 14.3% of the instances only one of the two pronuclei is correctly detected.

4 Conclusions

We presented a technique for detecting pronuclei in focus stacks, and determining their depth, by exploiting texture information from several focal planes simultaneously. Experimental results show promising results, which enable users to perform zygote scoring with an insight of the zygote’s 3D morphology. As future works, we plan to deal with timelapse data, which poses a new set of challenges for automated processing.

Acknowledgements. Research partially supported by CTI (Confederation's innovation promotion agency) grant n. 9707.1 PFSL-LS and NSF 200021-118071/1.

References

1. Ahonen, T., et al.: Face description with local binary patterns: application to face recognition. *Trans. on PAMI* (2006)
2. Beuchat, A., Thévenaz, P., Unser, M., Ebner, T., Senn, A., Urner, F., Germond, M., Sorzano, C.O.S.: Quantitative morphometrical characterization of human pronuclear zygotes. *Human Reproduction* 23, 1983–1992 (2008)
3. Blake, A., Isard, M.: *Active Contours*. Springer, Heidelberg (1998)
4. Gianaroli, L., Magli, M.C., Ferraretti, A.P., Lappi, M., Borghi, E., Ermini, B.: Oocyte euploidy, pronuclear zygote morphology and embryo chromosomal complement. *Human Reproduction* 22(1), 241–249 (2007)
5. Giusti, A., Corani, G., Gambardella, L., Magli, C., Gianaroli, L.: Blastomere segmentation and 3d morphology measurements of early embryos from hoffman modulation contrast image stacks. In: *Proc. of ISBI* (2010)
6. Giusti, A., Corani, G., Gambardella, L., Magli, C., Gianaroli, L.: Simultaneous focusing and contouring of human zygotes for in vitro fertilization. In: *Proc. of BIOSIGNALS* (2010)
7. Ojala, T., et al.: Gray scale and rotation invariant texture classification with local binary patterns. In: *Vernon, D. (ed.) ECCV 2000. LNCS, vol. 1842, pp. 404–420*. Springer, Heidelberg (2000)
8. Qureshi, H., et al.: Adaptive discriminant wavelet packet transform and local binary patterns for meningioma subtype classification. In: *Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 196–204*. Springer, Heidelberg (2008)
9. Reunanen, J.: Overfitting in making comparisons between variable selection methods. *The Journal of Machine Learning Research* 3, 1371–1382 (2003)
10. Scott, L., et al.: Innovative techniques in human embryo viability assessment. *Reproductive Biomedicine online* (2008)
11. Scott, L., Finn, A., O'Leary, T., McLellan, S., Hill, J.: Morphologic parameters of early cleavage-stage embryos that correlate with fetal development and delivery: prospective and applied data for increased pregnancy rates. *Human Reproduction* 22(1), 230–240 (2007)
12. Sethian, J.A.: *Level Set Methods and Fast Marching Methods*. Cambridge University Press, Cambridge (1999)
13. Smith, K., et al: Fast ray features for learning irregular shapes. In: *Proc. of CVPR* (2009)
14. Sørensen, L., et al.: Texture classification in lung ct using local binary patterns. In: *Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 934–941*. Springer, Heidelberg (2008)
15. Viola, Jones: Rapid object detection using boosted cascade of simple features. In: *Proc. of CVPR* (2001)
16. Witten, I.H., Frank, E.: *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd edn. Morgan Kaufmann, San Francisco (2005)
17. Zhao, G., Pietikäinen, M.: Dynamic texture recognition using local binary patterns with an application to facial expressions. *Trans. on PAMI* (2007)