2015 11th International Conference on Signal-Image Technology & Internet-Based Systems

Automatic detection and tracking of animal sperm cells in microscopy images

Ouadi Beya *ouadi.beya@u-bourgogne.fr*
Mazen Hittawe *mazen.hittawe@gmail.com*
Désiré Sidibé *desire.sidibe@u-bourgogne.fr*
Fabrice Mériaudeau *fmeriau@u-bourgogne.fr*

Le2i UMR CNRS 6163, Université de Bourgogne, 12 rue de la Fonderie, 71200 Le Creusot, France
{ouadi.beya, mohamad-mazen.hittawe}@u-bourgogne.fr

***Abstract*—Sperm tracking-and-analysis is one of the interesting topics in biological research and reproductive medicine, as it helps to assess the quality of the sperm for the male infertility. Computer-Assisted Sperm Analysis (CASA) systems provide a rapid and automated assessment of the parameters of sperm motion, together with improved standardization and quality control. In this paper, we propose a method to detect and track animal sperms automatically. First, we detect the sperms in the first frame of all the sequences using a bag-of-words approach and SVM classifier. Then, the detected sperm cells are tracked in the rest of all sequences using mean-shift. The proposed algorithm is evaluated on three videos in our datasets which have sperms as groundtruth. The experimental results show that our method achieves a precision of 0.94, 0.93 and 0.96, and a recall of 0.96, 0.92, and 0.97 for the three videos respectively in terms of sperm detection. RMSE (Root mean square error) is calculated to evaluate our results in terms of sperms tracking. The results show that we achieve high performance with RMSE of 8.06, 9.01,and 7.09 pixels for three different videos.**

***Keywords*-microscopy; classification; segmentation; tracking; sperm.**

I. Introduction

Sperm is a male reproductive cell which consists of an ellipsoidal or spherical head, a short midpiece and a thin motile tail. The sperm head contains the nucleus with genetic materials, surrounded anteriorly by a cp-like acrosome which contains digestive enzymes. Tracking of human sperm cells has a great importance in clinical applications of male infertility. According to biological studies, infertility is one of the common clinical problem which causes considerable morbidity, including stress, depression and sexual dysfunction [4]. The infertility for the male partner is assessed based on conventional criteria of semen quality such as semen volume, sperm concentration, motility percentage and morphology percentage [1]. The conventional assessment of human semen, especially sperm movement or sperm motility characteristics, is a highly subjective assessment, with considerable intra- and inter-technician variability. Proper characterization of sperm motility is an important goal in reproductive health studies. A motility grade is often used as a specified measure and it is classified into four grades: sperm with fast progressive movements, sperm with slow progressive movements, sperm with slow non-progressive movements and sperm are immotile and fail to move at all. In addition to this, the more standard measure of semen quality such as sperm concentration and sperm morphology are used for the assessment. However the conventional assessment are limited prognostic value of predicting the pregnancy achievement [1].

To overcome these limitations computer assisted sperm analysis methods have been introduced to measure the sperm morphology and sperm motility. The computer-assisted sperm analysis (CASA) system provides a rapid and automated assessment of sperm motion parameters [8]. Many algorithms have been developed to track sperm trajectories, measure sperm velocities and analyze sperm morphology. The authors in [9] used digital microscopes with phase-contrast accessories which make the sperm's head appear brighter than the other parts. To identify and track the sperm in the sequence of the frames, its brightest point is considered as the center of the sperm's head. The main limitation of this method is that the images lack proper contrast and sharpness. Shi *et al.* [11] proposed a robust single-sperm tracking algorithm based on a four-class thresholding methods to extract a single sperm in a region of interest (ROI). The nearest neighbour method is complemented with a speed-check feature to aid tracking in the presence of additional sperm or other particles. Nafisi *et al.* [10] proposed a template matching algorithm for sperm tracking which is insensitive to image acquisition conditions. In this method image enhancement is done to remove the background and extra particles and finally, the sperm tracking is done using template matching method. Friedrich *et al.* [5] used the resistive force theory which is based on tail movement of sperm cell to track then in 250 frame rate video. Zhou *et al.* [13] proposed an efficient and effective algorithm for sperm cells tracking which attempts to capture the motion uncertainty of the target object. The authors incorporated an orientation adaptive mean shift optimization into particle filter framework to enhance the efficiency of particle filter in sperm tracking. Imani *et al.* [7] proposed a method for tracking ple sperm cells. The authors used frame difference algorithm for background subtraction as a first step and proposed an improved non-linear diffusion filtering in the time domain to avoid the dependency of output accuracy with the selected threshold value and finally, for multiple tracking an optimal matching strategy is introduced which is based on the optimization of a new cost function over time domain. All the these previous methods are limited to sperm head tracking which has high intensity and neglects the sperm tails which has size of $(\leq 1\mu m)$ in thickness. Tracking the sperm cells with both head and low contrast

978-1-4673-9721-6/15 $31.00 © 2015 IEEE
DOI 10.1109/SITIS.2015.111

tail is one of the most challenging task under optical microscopy.

In this paper, we proposed a method to address the issue of sheep sperm cell tracking. The proposed method is based on classification and has two stages training and tracking. In the training stage, we build a dictionary based on low-level features using the bag-of-words approach. A multiple feature based dictionary created using SURF, HOG and LBP features. In the tracking stage, we first detect potential regions in the first frame of test video using pre-processing step. Then SURF, HOG and LBP features are extracted for each candidate region in the test video frame, and are transformed into a histogram of words occurrences using the learned dictionary from the training stage. These histograms are given as input to the SVM classifier for classification and detection of sperms. The detected sperm are then tracked using mean-shift tracking in the rest of the video frames. The rest of the paper is organized as follows. In Section 2, we describe the proposed method for sperm detection and tracking. Section 3 shows the performance evaluation of the proposed method on sperm video dataset. Finally, Section 4 gives concluding remarks.

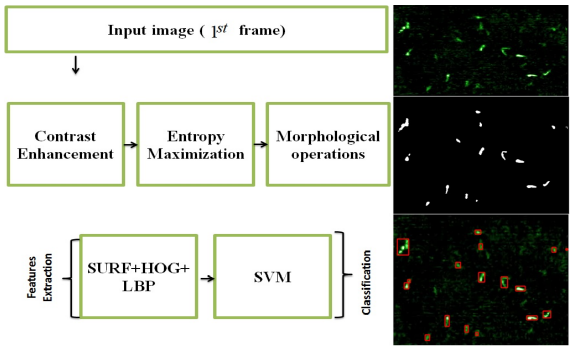

Figure 1. Flowchart of the sperms detection method.

II. Proposed method

In this section we introduce the proposed method for sperm detection and tracking. The method consists of steps such as pre-processing, feature extraction, classification and finally tracking the sperm. An overview of proposed sperms detection method is presented in Fig. 1 and the different steps are described in the following subsections.

A. Initialization and sperm detection

1) Pre-processing: Sperms appear as regions of dark intensities in RGB color images, and they can be detected by thresholding the image. First, we perform contrast stretching to enhance sperm regions. Contrast stretching is the process of transforming the intensity characteristics of the image such that dark regions are further darkened and bright regions are further brightened [6]. The stretching operator is defined as:

$$I_s = \frac{1}{1 + (\frac{\bar{I}}{I+\epsilon})^n} \quad (1)$$

where I is the input image, $\bar{I}$ its mean intensity value, and I_s is the output of the transformation. ϵ is a term added to avoid division by 0, and n controls the slope of the applied function. Then, potential regions are detected by thresholding the contrast enhanced image using a threshold obtained by entropy maximization. Furthermore, morphological operations (such as opening and closing) with a square structure element of size 10-by-10 are performed subsequent to thresholding in order to form connected components comprising the sperms. Opening detaches poorly connected regions, while closing forms connected components. The pre-processing step finds the sperms but also produces many false detections which correspond to image regions having similar intensity characteristics with the sperms. Therefore, a final classification step using a SVM is applied to reject false positive and discriminate between images containing sperms. In summary, to detect the sperms in a given test image, we apply the pre-processing step (contrast stretching and thresholding) to find potential sperm regions in the image. We then extract the low-level features from each potential region in the color test image and find the closest word to each individual feature in the dictionary obtained during the training stage. The frequency of occurrences of the visual words form the histogram representation of the region. Finally, the histograms are used as input feature vectors to the SVM classifier which predicts the labels of the regions. While the thresholding method finds the sperms in the input image (first frame in video), it also produces many false detection in cases where the intensity characteristics of the sperms in the image are not distinguishable or in case of noise in the image. In the next subsection, we propose a classification method to reduce the number of false detections.

2) Bag-of-words representation and classification: Our method uses the bag-of-words (BoW) representation of images for classification. BoW introduced in [12] is a powerful image representation method that has been used in different applications such as object recognition and image category classification. In our work, we adopt a supervised classification method. We first form a training set by collecting a set of images of sperms regions in different images. The training set also contains images of noisy regions (these are regions not corresponding to sperms). Our training set contains 50 sperm images and 50 clear noisy images respectively. We then extract a set of low-level features from the training images and build a visual vocabulary, or codebook, by quantization of the low-level features. SURF, HOG and LBP features are extracted from the three channels of the RGB color space. The features extracted from the entire set of training examples are then used to create a codebook using K-means clustering. If we define K clusters in the feature space, then the visual dictionary or codebook will contain K words each one being the center of one cluster. After creating the codebook, each of the training example is represented as a histogram of size K obtained by calculating the frequency of occurrences of each of the K words in

the features extracted from the image.

Finally, the obtained histogram representations are used as input features to train a linear SVM classifier [2] for distinguishing between sperm areas and other regions in a our images. The SVM finds a linear hyper-plane which maximizes the margin in this higher dimensional space. The training vectors x are mapped into a higher dimensional space by a kernel function $k(x, y)$, here we use linear kernel defined as following:

$$K(x, y) = x^T y. \quad (2)$$

For detecting sperm in a new test image, the pre-processing method explained in section II-A1 is first applied which results in a set of potential sperms regions in the image. Then, we extract SURF, HOG and LBP features from each potential region and find the closest word in the codebook to each individual feature. The frequency of occurrences of the visual words forms the histogram representation of the region. Finally, the SVM classifier uses this histogram representation as input feature vector to predict the label of this region.

B. Sperm tracking and trajectory

Once the sperms are detected in first frame, then the initial locations are used as input for mean-shift tracking [3].

1) Sperm mobility and vitality: *sperm mobility* and *sperm vitality* are used for sperm analysis. The sperm mobility and vitality measures help to decide the ability of sperm to move properly through the female reproductive tract to reach the egg. In order to measure the sperm mobility and sperm vitality from the video in the frame i to the next frame $(i + 1)$, the instantaneous velocity v_i and instantaneous acceleration a_i are calculated with sampling period $\Delta t = 0.033s$ (corresponding to capture frequency $f = 30Hz$). The speed of movement for each sperm are calculated according to the two axis directions (x,y) as $v_x(i)$, $v_y(i)$ for frame i using Eq(3). Then we can determine the instantaneous velocity v_i for each sperm as in Eq(4).

$$v_x(i) = \frac{x(i+1) - x(i)}{\Delta t}, v_y(i) = \frac{y(i+1) - y(i)}{\Delta t} \quad (3)$$

$$v_i = \sqrt{v_x^2(i) + v_y^2(i)} \quad (4)$$

where (x_i, y_i) is measured position of sperm in the frame i. The analysis of the average speed for each sperm which allows to compare the threshold set by the medical expert, to assess the mobility of sperms and determine the fertility of an individual. In the same way, we determine the acceleration $a_x(i)$ and $a_y(i)$ for each sperm moving from frame i to next frame $(i + 1)$.

$$a_x(i) = \frac{v(i+1) - v(i)}{\Delta t}, a_y(i) = \frac{v(i+1) - v(i)}{\Delta t} \quad (5)$$

The accelerations of movement of each sperm are calculated according to the two axis directions (x,y) as $a_x(i)$, $a_y(i)$ for frame i as defined in Eq(5). Then we can determine the instantaneous acceleration a_i for each sperm as follows:

$$a_i = \sqrt{a_x^2(i) + a_y^2(i)} \quad (6)$$

The average acceleration of each sperm which allows to compare the threshold set by the medical expert, to assess the vitality of sperms and distinguish the true concentration of active sperms. Through this approach, we can track and measure the speed and acceleration of sperms, to assess their mobilities and vitalities. So we can supply a medical expert with physiological parameters, allowing to analyze objectively the fertility of an individual, and measure the influence of taking medication.

III. Experimental results

In this section, the proposed sperm-detection-and-tracking method is evaluated on three complex sperm videos dataset with image resolution of 336×446 pixels. The acquisition of the video sequence is performed using fibered confocal fluorescence microscopy. The real dimension of each frame in our dataset is $425mm \times 310mm$ =131750 mm^2 and the image resolution is $464 \times 336 =$ 155904 pixels. This means the area of one pixel is equal to 0.84 mm^2. The experimental evaluation is divided into two parts and are explained in Section. III-A and Section. III-B

A. Performance evaluation of sperm detection

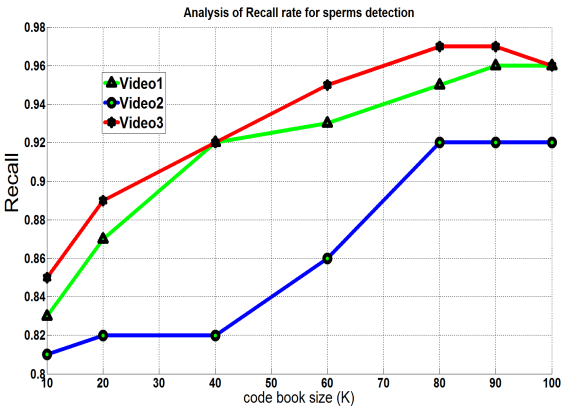

Figure 2. Recall analysis for our dataset.

Two measures namely *recall* and *precision* are used to perform experimental evaluations for sperms detection and classification. The recall (or true positive rate) gives the proportion of actual sperms correctly identified. The precision gives the proportion of positively identified cases that are actual sperms as in Eq(7). We first create a training dataset which contains 50 sperm images and 50 non-sperm images from the our dataset. These training images are used to create one dictionary based on SURF, HOG and LBP features. The images in our datasets not used for training are used for testing. The test dataset contains 50 images in total for each video. We analyze the proposed method with different dictionary size. Fig. 2 illustrates the

Dataset	Recall	Precision
Video 1	0.96	0.94
Video 2	0.92	0.93
Video 3	0.97	0.96

Table I
QUANTITATIVE RESULTS OF SPERMS DETECTION.

Dataset	Frames	RMSE
Video 1	151	8.06 px
Video 2	145	9.01 px
Video 3	142	7.09 px

Table II
QUANTITATIVE RESULTS OF SPERMS TRACKING.

influence of the size parameter K in detecting the sperms on three videos.

$$Recall = \frac{TP}{TP+FN}, Precision = \frac{TP}{TP+FP} \quad (7)$$

We evaluate the performance for dictionary size varying from 10 to 100. From Fig. 2, we can clearly observe that, as the size of the dictionary increases the recall values significantly increases. However, the method reaches to its best performance for $K = 90$ and the results remain stable for larger values. Therefor, we use $K = 90$ in all our experiments. The results also confirm that a multiple features based dictionary (Hog, Surf and LBP) achieves high level of sperm classification accuracy. Table. I summarizes the results of sperms detection of proposed method. From this table we can clearly observe that the proposed method achieves average *precision* of $0.94, 0.93, 0.96$ respectively and average *recall* of $0.96, 0.92, 0.97$ for the three videos.

B. Performance evaluation of sperm tracking

Root Mean Square Error (RMSE) is used to evaluate the performance of the sperms tracking for the proposed method. The RMSE measure gives the error rate between the ground truth path and estimated path by the proposed method. The smaller the RMSE value, the higher the accuracy at each path point will be. The RMSE measure are defined as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i - x_i^*)^2 + (y_i - y_i^*)^2}{N}} \quad (8)$$

where (x_i, y_i) is measured position of sperm in frame i, (x_i^*, y_i^*) is ground truth position of sperm in the frame, and N is the number of frames.

Fig. 3 illustrates the velocity for sperms contained in each video to assess their mobilities. These mobilities analysis help to locate the fastest sperm. In a similar way, Fig. 4 illustrates the acceleration for sperms contained in each video to evaluate their vitalities. We can clearly notice that these two parameters allow to extract the most fertile sperm for insemination in vitro (medical laboratory). Fig. 5 shows some qualitative tracking results on color images of the video 3 in our dataset, where the results are compared against their ground-truth trajectories. Quantitative results are summarized in table. II for sperms tracking of the

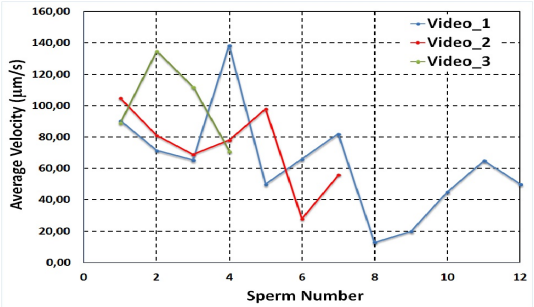

Figure 3. Output of sperms velocity.

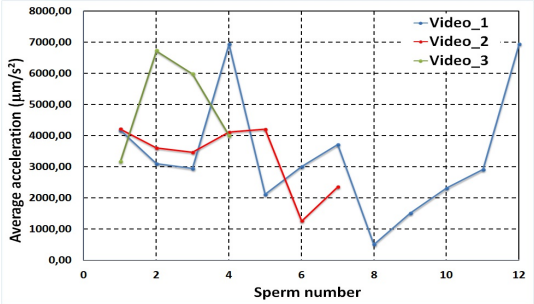

Figure 4. Output of sperms acceleration.

proposed method. We can observe that the proposed method achieves average RMSE of $8.06px$ for video 1 which contains 12 sperms, $9.01px$ for the video 2 of seven sperms, and of $7.09px$ for the video 3 of four sperms. We can clearly observe that the proposed method is accurate in tracking the sperms with low RMSE values.

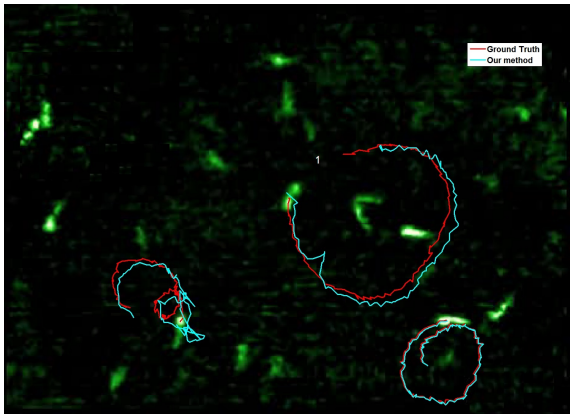

Figure 5. Trajectories of tracking results between ground-truth and our method.

IV. CONCLUSION

In this paper, we proposed a method for detecting and tracking the sperms for medical analysis. This method has two steps which are detection and tracking. The sperms detection is based on creating a dictionary using SURF, HOG and LBP with bag-of-words approach and SVM

classification. In the second step the detected sperm is tracked using mean-shift. The experimental results on sheep sperm videos dataset show that the proposed method perform better with high *precision* and *recall* and low RMSE. In future, our work will be extended to apply on a database with videos of human sperms to study human infertility and to collaborate with medical experts to analyze the resultant medical effects by the consumption of drugs, coffee and cigarettes.

REFERENCES

[1] E Bostofte, P Bagger, A Michael, and G Stakemann. Fertility prognosis for infertile men: results of follow-up study of semen analysis in infertile men from two different populations evaluated by the cox regression model. *Fertility and sterility*, 54(6):1100–1106, 1990.

[2] C-C. Chang and C-J. Lin. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, 2011.

[3] Dorin Comaniciu, Visvanathan Ramesh, and Peter Meer. Real-time tracking of non-rigid objects using mean shift. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 2, pages 142–149. IEEE, 2000.

[4] Alice D Domar, A Broome, Patricia C Zuttermeister, M Seibel, and Richard Friedman. The prevalence and predictability of depression in infertile women. *Fertility and sterility*, 58(6):1158–1163, 1992.

[5] BM Friedrich, IH Riedel-Kruse, J Howard, and F Jülicher. High-precision tracking of sperm swimming fine structure provides strong test of resistive force theory. *The Journal of experimental biology*, 213(8):1226–1234, 2010.

[6] R. C. Gonzalez, R. E. Woods, and S. L. Eddins. Digital image processing using matlab. *Upper Saddle River, NJ Jensen: Prentice Hall,*, 2004.

[7] Yoones Imani, Niloufar Teyfouri, Mohammad Reza Ahmadzadeh, and Marzieh Golabbakhsh. A new method for multiple sperm cells tracking. *Journal of medical signals and sensors*, 4(1):35, 2014.

[8] Jun Liu, Clement Leung, Zhe Lu, and Yu Sun. Human sperm tracking, analysis, and manipulation. In *Smart Materials-Based Actuators at the Micro/Nano-Scale*, pages 251–264. Springer, 2013.

[9] Sharon T Motimer. Casa practical aspects. *Journal of andrology*, 21(4):515–524, 2000.

[10] Vahid Reza Nafisi, Mohammad Hasan Moradi, and Mohammad Hosain Nasr-Esfahani. A template matching algorithm for sperm tracking and classification. *Physiological measurement*, 26(5):639, 2005.

[11] Linda Z Shi, Jaclyn M Nascimento, Michael W Berns, and Elliot L Botvinick. Computer-based tracking of single sperm. *Journal of biomedical optics*, 11(5):054009–054009, 2006.

[12] J. Sivic and A. Zisserman. Video Google: A text retrieval approach to object matching in videos. In *Proceedings of the International Conference on Computer Vision*, volume 2, pages 1470–1477, October 2003.

[13] Xiuzhuang Zhou and Yao Lu. Efficient mean shift particle filter for sperm cells tracking. In *Computational Intelligence and Security, 2009. CIS'09. International Conference on*, volume 1, pages 335–339. IEEE, 2009.