Classification of human sperm heads using elliptic features and LDA

Fariba Shaker[1], S. Amirhassan Monadjemi[2]
[1]PhD Student, [2]Associate Professor
Department of Computer Engineering
University of Isfahan
Isfahan, Iran, 81746
Email: f.shaker@eng.ui.ac.ir, monadjemi@eng.ui.ac.ir

Javad Alirezaie
Professor, Senior Member IEEE
Department of Electrical and Computer Engineering
Ryerson University
Toronto, Canada, M5B 2K3
Email: javad@ryerson.ca

Abstract—**For diagnosis of infertility in men semen analysis is conducted in which sperm morphology i.e. the size and shape of the sperm, is one of the factors that are evaluated. Since manual assessment of sperm morphology is time consuming and subjective, automatic classification methods are being developed. Automatic classification of sperm heads is a complicated task due to the "within class" differences and "between class" similarities. To automatically classify the sperms, appropriate features should be extracted from their microscopic images. In this research, a set of previously proposed features is extracted and examined in an automatic framework in order to evaluate their discriminating capacity in classifying sperms into four classes of shapes (Normal, Tapered, Pyriform and Amorphous). Also, a new set of features called elliptic features is proposed and added to the original features to improve the classification results. Both sets of features are used with Linear Discriminant Analysis (LDA) classifier. It is shown that adding these new features, significantly improves the discrimination between those classes of sperm shapes.**

Keywords—sperm head classificatin; sperm abnormality; sperm morphology; infertility; eliptic features

I. INTRODUCTION

Infertility affects almost ten percent of the population and close to half of the cases are related to men [1]. To diagnose infertility in men, semen examination is conducted in which one of the steps is the assessment of sperm morphology. Morphology assessment of sperms involves finding the percentage of morphologically abnormal sperms and their type of abnormality. The manual assessment of sperm abnormalities is known to be subjective as well as time consuming [2], so automatic methods are sought after.

Automatic detection of sperm shape categories not only is helpful in diagnosing infertility but it will also facilitates more research in the causes of male infertility and helps to develop methods for its treatment and also its prevention [3].

A sperm consists of three main parts namely head, midpiece and tail (Figure 1) in which the head is further divided into acrosome and nucleus. Each of these three parts should have certain characteristics to be considered morphologically normal. For example, the head should be smooth, regularly contoured and generally oval in shape. Also, there should be a well-defined acrosomal region comprising 40-70% of the head area. There should be no large vacuole or more than two small vacuoles on the acrosomal region and there should be no vacuoles on the post-acrosomal region [4]. Sperm head abnormalities are the result of one or more of these characteristics being disrupted. Sperm abnormalities are divided into three main categories: head defects, midpiece defects and tail defects among which head defects are considered more important [4]. In this research, we focus on head defects.

According to WHO [1] 2010, there are ten classes of abnormalities related to the head of the sperms which include Small, Large, Amorphous, Tapered, Pyriform, Large Acrosome, Small Acrosome, Round, Two Headed and Vacuolated. These classes are characterized by the specific shape, size, and texture of the head and its constituent parts. These classes of abnormalities present a vast and at the same time continuous difference in size, shape, and texture which make their classification a complicated task. In addition to within class variability, there are between class similarities as well. For example, there are Amorphous sperms that are elongated like Tapered sperms.

Most of the work in the area of sperm classification is in the veterinary field and the works related to human sperm is scarce. Although there are similarities between these two areas but the types of abnormalities and the motivation for assessing different classes of abnormalities in these two fields are different. Most of the works in the veterinary field are focused

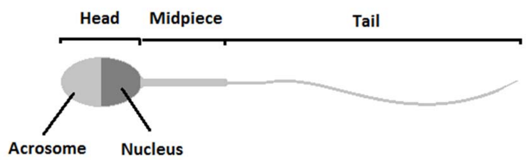


Figure 1. Different parts of a sperm [7]

[1] World Health Organization

978-1-5090-6454-0/17/$31.00© IEEE

on classifying sperms in two classes of normal/abnormal or dead/alive based on the state of their acrosome.

Perez-Sanchez et al. studied ten shape features to find out their relevance to classification of human sperm heads into 14 classes of shapes [5]. These features comprise of two sets of head measurements; the first set is called basic features (length, width, area, perimeter and mass) and the second set is called derived features (ratio, the difference of length and width, ellipticity, form and total mass). Their method is manual and they used statistical tools to show these features are relevant in discriminating some of the shape classes.

In another study by Yi et al., the sperm heads are classified into four classes of Normal, Small, Elongated and Megalo [6]. They used the first ten coefficients of Fourier transform of the head contour points to reconstruct the head contour. Then they applied the wavelet transform to the enclosed area of the head. The root mean square error in transform space between the image and each class of abnormalities was used to classify the sperm head.

In a recent study Jiaqian et al. used Principal Component Analysis (PCA) combined with K-nearest neighbor algorithm to classify human sperm heads into two classes of Normal and Abnormal. Their results show there is a bias towards the Normal class which makes the usefulness of the method questionable.

In this research we first examine the set of previously proposed features by Perez-Sanchez et al. and their discriminating capacity in classifying the sperm heads into four different classes of Normal, Tapered, Pyriform and Amorphous. Then we introduce a new set of features called elliptic features to improve the classification results. It should be noted that unlike Perez-Sanchez work, our method is completely automatic and instead of manual investigation of features we used Linear Discriminant Analysis (LDA) method to classify the sperm heads. It is shown that our proposed set of features is able to significantly improve the classification results. To the best of our knowledge, this is the first time that these four classes of human sperm heads are classified automatically. The significance of this research is the underlying difficulties in classifying these four groups of sperm heads due to the variability within each class and at the same time similarities existing between classes.

II. MATERIALS AND METHODS

A. Data Set

Semen samples were collected from patients at the Isfahan Fertility and Infertility center (IFIC). After preparing the semen smears, the slides were air dried and fixed. The slides were stained using the Diff-Quik method as described in WHO 2010 manual [4]. Images of the slides were taken using Sony (SSC-DC58AP) digital camera attached to the third eyepiece of Olympus CX21 bright field microscope with a ×100 objective and ×10 eyepiece. The format of the images is Bitmap and their size is 720×576 pixels. Sperm heads in each image were cropped and classified by experienced professionals at the IFIC. The resulting dataset consists of 66 Normal, 39 Pyriform, 42 Amorphous, and 23 Tapered sperm heads. Figure 2 shows one example from each class.

B. Segmentation of sperm heads

In this research, sperm heads that are cropped are segmented automatically using active contours. The RGB images of sperms are first transformed into gray-level and then are thresholded using Otsu method to locate the sperm heads. The edges are extracted using Canny edge detector. The extracted head masks, together with extracted edges, are used in Gradient Vector Flow (GVF) active contour method to extract the outer contour of sperm heads. The details of the segmentation method can be found in reference [7]. These contours are later used to extract different head features.

C. Perez-Sanchez(PS) Features

As mentioned earlier there are ten features proposed by Perez-Sanchez et al. for sperm heads, five of which are basic features and the rest are derived features. These features are summarized in Table 1. In this table, A is the area of the sperm head in pixels and P is its perimeter. L and W are the length and the width of the sperm head in pixels. M is the average gray level of the sperm head and is calculated using (1). In this equation, N is the number of pixels and $g(x_i, y_i)$ is the gray level of the pixel (x_i, y_i) of the head (H).

$$M = \frac{1}{N}\sum_{i\in H} g(x_i, y_i) \quad (1)$$

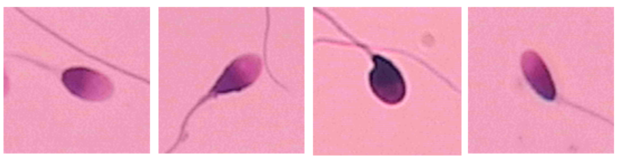
Derived features are calculated using the basic features, and their formulas are summarized in Table 1.

In order to measure the length and the width of the sperm heads, after extracting the contour of the head using active contours, an ellipse is fitted to the head contour. The length and the width of the ellipse are used as the length and the width features. The perimeter of the head is measured by counting the number of pixels comprising the head contour. Area of the head is also measured by summing the number of pixels enclosed by the head contour.

D. Elliptic Features

In order to improve the classification performance, we propose a new set of features which we name elliptic features. As stated before, after segmenting the sperm head an ellipse is fitted to the extracted head contour. The minimum Euclidian distance of each pixel of the ellipse from the contour of the sperm head will depend on how close the shape of the head is to a perfect ellipse. The mathematical expression of the fitted ellipse is presented in (2). In this equation, x and y are the indices of the points in a two dimensional space and the coefficients a to f are the parameters defining the ellipse.

$$ax^2 + bxy + cy^2 + 2dx + 2ey + f = 0 \quad (2)$$



(a) (b) (c) (d)

Figure 2. Examples of each class of sperm shapes a) Normal, b)Pyriform, c) Amorphous, and d) Tapered

TABLE 1. FEATURES PROPOSED BY PEREZ-SANCHEZ ET AL [5].

Basic Features	Derived Features
Area (A)	Ratio (L/W)
Perimeter (P)	Length–Width (L–W)
Length (L)	Ellipticity (L–W)/(L+W)
Width (W)	Form $(4\pi A/P^2)\times 10^3$
Mass (M)	Total Mass (A×M)

Figure 3(a) shows one of the sperm heads from the data set and Figure 3(b) shows its extracted contour in white color and the fitted ellipse in red color.

Then we use the ellipse coefficients to construct (3). In this equation, x_h and y_h are the coordinates of the pixels located on the head contour. Coefficients a to f are the same coefficients as the fitted ellipse. D is a value showing how close is the location of the pixel to the fitted ellipse and in which direction. We refer to D as *ellipse distance*. For pixels located on the ellipse, D is zero and for the pixels located inside/outside the ellipse, D is a negative/positive number.

$$ax_h^2 + bx_h y_h + cy_h^2 + 2dx_h + 2ey_h + f = D \qquad (3)$$

Starting from one point on the head contour and moving clockwise, ellipse distance is calculated for each pixel and a vector of ellipse distances are constructed for each head contour. As an example, Fig. 3(c) shows the graph of the ellipse distances of a head contour. As is seen in this figure the graph is not smooth and the reason is that the head contour consists of discreet points (pixels) rather that being a smooth curve. To smooth the vector of ellipse distances, the vector is convolved twice with a uniform filter of size 1×3. Fig. 3(d) shows the smoothed vector.

From this vector, three features are extracted. The first and the second features are the maximum and the minimum values

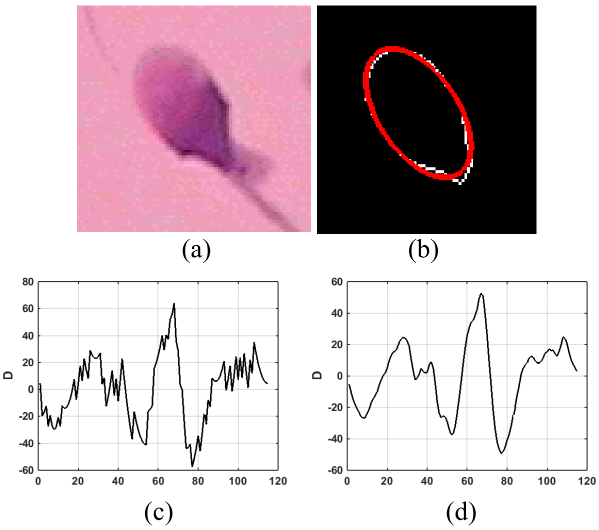


Figure 3. a) The original sperm, (b) The extracted head contour (white) and the fitted ellipse (red), (c) Noisy ellipse distance, and (d) Smoothed ellipse distance

of the ellipse distances. The third feature is the number of times that the ellipse distances change their sign from positive to negative and vice versa along the vector. These three features are added to the PS features and used in classification.

E. *Classification*

In this research Linear Discriminant Analysis (LDA) for multiple classes is used as the classification method. It was shown that LDA performs better than other classifiers such as Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN) and Quadratic Discriminant Analysis (QDA) in classifying boar sperm heads [8]. The idea behind LDA is to maximize the between-class variance and at the same time minimize the within-class variance. In this way, maximum separation between classes is achieved.

In order to train the classifier, half of the images from each class are selected randomly and their features are used to train the classifier. The other half of the images is used to verify the performance of the classifier. The training and testing of the classifier are done 100 times and the results are averaged. We performed the classification first using PS features and then using both PS and elliptic features.

F. *Evaluation criteria*

In order to evaluate multiclass classification results, there are a number of metrics to consider. For an individual class C_i, the metrics are tp_i (the number of true positives for class C_i), fp_i (the number of false positives for class C_i), tn_i (the number of true negatives for class C_i), fn_i (the number of false negatives for class C_i), $Accuracy_i$, $Precision_i$ and $Recall_i$. The formulas for calculating these metrics are summarized in Table 2. In this table, M is the confusion matrix and the index i is the class number. l is the number of classes which is four in our experiments. In order to find the overall Accuracy, Precision and Recall there could be two approaches; one is to average the same measures calculated for each class C_i, (macro-averaging) and the other is to find cumulative tp, tn, fp and fn and then calculating the performance measure (micro-averaging) [9]. Macro-averaging treats all classes equally while micro-averaging favors bigger classes. Since in our experiments, the numbers of samples in different classes are not the same so macro averaging is used in this paper.

III. RESULTS

As discussed before, we used LDA classifier in order to classify sperm heads using PS features and also the combination of PS features and elliptic features. The confusion matrix for PS features and the combination of PS and elliptic features are presented in Table 3 and Table 4 respectively. As tables show, elliptic features improved the discrimination power of the classifier for Tapered and Pyriform classes. Table 5 presents the precision, recall, and accuracy of each class as well as the overall results. The table shows significant improvement in the precision of Normal, Tapered and Pyriform classes. The improvement in the precision for the Normal class is achieved by decreasing the number of false positives. The new feature set seems to have a positive effect in discrimination of Pyriform and to a less degree on Tapered class. As confusion matrix shows, PS features will result in a relatively high confusion between Pyriform and Normal sperms but elliptic features will result in more distinction between these two classes. The drawback of introducing the elliptic features is a small increase in the confusion between Pyriform and Amorphous classes. For Tapered and Pyriform classes, all three evaluation criteria i.e. Accuracy, Precision and Recall, show improvements. Also the overall Accuracy, Precision, and Recall show improvements when using elliptic features.

The confusion matrices show some interesting points. Table 3 shows that with PS features, Normal and Pyriform sperms are confused more than other classes. The reason is that the sperms from these two classes are very close in shape. The only difference is the narrowing toward the midpiece in Pyriform sperms. Actually, this narrowing can also happen in Normal shapes but to a lesser degree. As Table 4 shows the proposed elliptic features are able to distinguish this subtle difference between these two classes and as a result reduce the number of misclassified Pyriform shapes. The same thing happened for Tapered and Pyriform shapes.

IV. CONCLUSION

Morphological classification of sperm heads is a complicated task due to a large variability within a single class and at the same time, the similarities present between the samples from different classes. The aim of this research was to find optimum features for automatic classification of human sperm heads into four classes of Normal, Tapered, Pyriform and Amorphous. To this end, a set of previously proposed features were examined against a dataset of labeled sperm heads. Also, a new set of features, called elliptic features, were proposed to improve the classification results. Linear Discriminant Analysis algorithm was used as the classifier. The results show that the proposed features have improved the overall Precision, Recall, and Accuracy. Also, these measures show class level improvements for the Normal, Tapered and Pyriform classes. The most improvement has happened for Pyriform class with 7 and 10 percent increase in Precision and Recall respectively. The results show that the newly proposed elliptic features are improving the discrimination of Pyriform sperms.

TABLE 2. EVALUATION METRICS

Measure	Formula	Evaluation
tp_i	M_{ii}	True positive per class
fp_i	$\sum_{j \atop j\neq i} M_{ji}$	False positive per class
tn_i	$\sum_{j,k \atop j,k\neq i} M_{jk}$	True negative per class
fn_i	$\sum_{j \atop j\neq i} M_{ij}$	False negative per class
$Precision_i$	$\frac{tp_i}{tp_i + fp_i}$	Per class Precision (Specifity)
$Recall_i$	$\frac{tp_i}{tp_i + fn_i}$	Per class Recall (Sensitivity)
$Accuracy_i$	$\frac{tp_i + tn_i}{tp_i + tn_i + fp_i + fn_i}$	Per class Accuracy
Average Precision	$\frac{\sum_i Precision_i}{l}$	The average per class Precision
Average Recall	$\frac{\sum_i Recall_i}{l}$	The average per class Recall
Average Accuracy	$\frac{\sum_i Accuracy_i}{l}$	The average per class effectiveness of classifier

ACKNOWLEDGMENT

The authors would like to thank Isfahan Fertility and

TABLE 3. CONFUSION MATRIX USING PS FEATURES

Normal	30	0	2	1
Tapered	1	6	2	2
Pyriform	6	2	12	1
Amorphous	2	3	2	14
	Normal	**Tapered**	**Pyriform**	**Amorphous**

TABLE 4. CONFUSION MATRIX USING PS AND ELLIPTIC FEATURES

Normal	30	0	2	1
Tapered	1	7	1	2
Pyriform	3	2	14	2
Amorphous	2	3	2	14
	Normal	**Tapered**	**Pyriform**	**Amorphous**

TABLE 5. PRECISION AND RECALL FOR EACH CLASS USING PS FEATURES OR PS PLUS ELLIPTIC FEATURES

	PS features			**PS plus elliptic features**		
	Precision	***recall***	***accuracy***	***Precision***	***recall***	***accuracy***
Normal	77%	91%	86%	**83%**	91%	90%
Tapered	55%	55%	88%	**58%**	**64%**	**90%**
Pyriform	67%	57%	83%	**74%**	**67%**	**86%**
Amorphous	**78%**	67%	86%	74%	67%	86%
Over all	69%	68%	86%	**72%**	**72%**	**88%**

Infertility Centre for providing the samples and microscopic images; and their specialized staff for labeling the data. We would also wish to thank Ms. Mehraban for providing us with valuable advice in addition to her assistance in creating the dataset.

REFERENCES

[1] M. R. Maduro and D. J. Lamb, "Understanding New Genetics of Male Infertility," *J. Urol.*, vol. 168, no. 5, pp. 2197–2205, 2002.

[2] G. Barroso, R. Mercan, K. Ozgur, M. Morshedi, P. Kolm, K. Coetzee, T. Kruger, and S. Oehninger, "Intra- and inter-laboratory variability in the assessment of sperm morphology by strict criteria: impact of semen preparation, staining techniques and manual versus computerized analysis," *Hum. Reprod.*, vol. 14, no. 8, pp. 2036–2040, 1999.

[3] R. J. Aitken, "Whither must spermatozoa wander? The future of laboratory seminology," *Asian J. Androl.*, vol. 12, no. 1, pp. 99–103, 2010.

[4] *WHO laboratory manual for the Examination and processing of human semen*, 5th ed. 2010.

[5] F. Pérez-Sánchez, J. J. de Monserrat, and C. Soler, "Morphometric analysis of human sperm morphology," *Int. J. Androl.*, vol. 17, no. 5, pp. 248–255, 1994.

[6] W. J. Yi, K. S. Park, and J. S. Paick, "Parameterized characterization of elliptic sperm heads using fourier representation and wavelet transform," 1998, vol. 20, no. 2, pp. 974–977.

[7] F. Shaker, S. A. Monadjemi, and A. R. Naghsh-Nilchi, "Automatic detection and segmentation of sperm head, acrosome and nucleus in microscopic images of human semen smears," submitted for publication.

[8] E. Alegre, V. Gonzalez-Castro, S. Suarez, and M. Castejon, "Comparison of supervised and unsupervised methods to classify boar acrosomes using texture descriptors," in *ELMAR, 2009. ELMAR'09. International Symposium*, 2009, pp. 65–70.

[9] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.