Automatic classification of the acrosome status of boar spermatozoa using digital image processing and LVQ

Enrique Alegre[1], Michael Biehl[2], Nicolai Petkov[2], Lidia Sánchez[3]
[1]*Department of Electrical, Systems and Automatic Engineering, University of León, Spain*
[2]*Institute of Mathematics and Computing Science, University of Groningen, The Netherlands*
[3]*Department of Mechanical, Computing and Aerospace Engineerings, University of León, Spain*

Abstract. We consider images of boar spermatozoa obtained with an optical phase-contrast microscope. Our goal is to automatically classify single sperm cells as acrosome-intact (class 1) or acrosome-damaged (class 2). Such classification is important for the estimation of the fertilization potential of a sperm sample for artificial insemination. We segment the sperm heads and compute a feature vector for each head. As a feature vector we use the gradient magnitude along the contour of the sperm head. We apply learning vector quantization (LVQ) to the feature vectors obtained for 320 heads that were labelled as intact or damaged using stains. A LVQ system with four prototypes (two for each class) allows us to classify cells with a overall test error of 6.8%. This is considered to be sufficient for semen quality control in an artificial insemination center.

Keywords: acrosome assessment, digital image processing, boar sperm cells, classification, LVQ, gradient magnitude contour, phase-contrast microscope image.

1 Introduction

Several computer assisted approaches have been developed to assess the quality of semen samples [27, 15]. They are used in Computer Aided Sperm Analysis (CASA) systems to automatically evaluate the fertilization potential of semen samples [25]. Approaches that had initially been developed for human semen analysis have been meanwhile adapted to other species. There are works that address the problem in multiple species [3]. CASA systems that deploy image processing techniques consider usually sperm characteristics like motility, concentration or sperm cell head shape. However, the use of these features has several drawbacks. For instance, sperm motility is sensitive to temperature changes and, furthermore, its relation to fertility is not clear [13, 19]. High sperm concentration does not guarantee fertilization either: a sample can have a high number of spermatozoa but most of them can be dead, hence infertile. Regarding sperm morphology, assessment is focused on computing morphometric measures like area and perimeter or shape descriptors in order to detect head shape abnormalities [17, 20, 4, 5] but normal shape also gives no guarantee that a sperm cell is alive and fertile. There are also works on the detection of droplets in tails [9]. An assessment of sperm vitality and fertilization potential has also been done using analysis of the grey level patterns of sperm cell heads in images obtained with a phase-contrast microscope [22, 21, 6]. The current work is a further elaboration of this latter approach.

Specifically, in this article we investigate the relation of the grey level pattern of a sperm cell head to the state of the cell acrosome. The acrosome is a cap-like structure that develops over the anterior half of a spermatozoon head. It has its own membrane and contains enzymes. As the sperm approaches the oocyte, an acrosome reaction takes place during which the anterior head plasma membrane fuses with the outer membrane of the acrosome, exposing the contents of the acrosome. The released enzymes are required for the penetration of sperm through a layer of follicular (cumulus) cells that encase the oocyte. The acrosome reaction also renders a sperm cell capable of penetrating through the zona pellucida, an extracellular coat surrounding the oocyte, and fusing with the egg. While spermatozoa have to present acrosome integrity at the time of collecting a sperm sample, at insemination time a certain fraction of spermatozoa that undergo acrosome reaction is required. This is due to the fact that the acrosome reaction produces a chain effect that increases the possibility of fertilization. A semen sample is potentially fertile if it has a certain sperm cell concentration and most of the spermatozoa are alive and present acrosome integrity. Otherwise, such

a sample is considered of low quality. Hence, a method for the automatic evaluation of the fraction of acrosome intact cells in a sperm sample would be of practical importance. Despite the broadly recognized importance of acrosome integrity evaluation in semen quality assessment, we are not aware of any image processing work on this topic. In the following we will use the term ”acrosome-damaged” (as an opposite to ”acrosome-intact”) for sperm cells that are either undergoing or have already undergone an acrosome reaction.

The acrosome states of spermatozoa can be determined objectively by staining or subjectively (and less reliably) by visual inspection by a veterinary expert of grey level images of sperm cell heads obtained with a phase-contrast microscope. Fig.1a shows the result of staining with PNA and iodure propidium using the staining method described by Harrison and Vickers [12]. The anterior parts of acrosome-damaged spermatozoa are colored green in an image obtained with a fluorescence microscope. However, staining, similar to visual inspection by veterinary experts, is time consuming and has a relatively high cost. In this work we propose a method to assess acrosome integrity based on the automatic analysis of grey level images acquired with a phase contrast microscope, Fig.1b. The idea is to eliminate staining and the use of a fluorescence microscope. Our approach is based on the observation that there are some characteristic differences in the grey level profiles along the contours of acrosome-intact vs. acrosome-damaged sperm heads. We extract a feature vector from the gradient magnitude image of a cell head and use this vector to classify the cell as acrosome-intact or acrosome-damaged by comparing it with prototype feature vectors. We determine these prototype vectors by applying Learning Vector Quantization (LVQ) to a training set of feature vectors labelled as acrosome-intact or acrosome-damaged using the mentioned staining technique.

In Section 2, we present the methods used to extract a feature vector from the grey level image of a cell head. Learning and classification are described in Section 3. The achieved results are presented in Section 4. The paper is concluded with a summary and discussion in Section 5.

2 Vectorization

2.1 Pre-processing and segmentation

An optical phase-contrast microscope Nikon Eclipse with magnification $\times 100$ and a digital camera Nikon Coolpix 5000 were used to acquire boar semen images of resolution/ size 2560×1920 pixels. Such an image typically contains a number of spermatozoa that can vary widely across images and samples (Fig. 1b).

Sperm head images were cropped manually from such a boar semen sample image, Fig.2a. In each sperm head image we segment automatically the sperm head by binarization using Otsu's [18] method and applying several morphological operations (dilations and erosions) [1], Fig. 2b. We use the contour of the obtained sperm head binary mask in the following. We also automatically localize the point where the middle piece, from which the tail develops, connects to the head. This point is used as a reference point in the following.

Each sperm head was manually labelled as acrosome-intact or acrosome-damaged according to its stain color in the corresponding fluorescence image (Fig.1a). The staining method described by Harrison and Vickers [12] was used. Automatic labelling is also possible [2].

2.2 Scale-dependent gradient computation and vectorization

We use a scale-dependent approach to gradient computation [24] that has been shown to reduce noise and discretisation effects in image processing [7, 26, 10]. Let $f(x, y)$ be a grey level distribution in an input image and let $g_\sigma(x, y)$ be a two-dimensional (2D) Gaussian function of standard deviation σ which we consider as a scale or resolution parameter. The x- and y-components of the scale-dependent gradient of $f(x, y)$ are defined as convolutions of $f(x, y)$ with the x- and y-derivatives of $g_\sigma(x, y)$, respectively:

$$\nabla_{\sigma,x} f = f * \frac{\partial g_\sigma}{\partial x} , \quad \nabla_{\sigma,y} f(x, y) = f * \frac{\partial g_\sigma}{\partial y} . \tag{1}$$

In the following we use the magnitude $M_\sigma(x,y)$ of the scale-dependent gradient:

$$M_\sigma(x,y) = \sqrt{\left(\nabla_{\sigma,x} f(x,y)\right)^2 + \left(\nabla_{\sigma,y} f(x,y)\right)^2}\,. \tag{2}$$

Fig.2c shows the result of such a computation together with a superimposed cell head contour as determined in the segmentation procedure described above.

Fig. 3 shows intracellular density distribution images typical for acrosome-intact and acrosome-damaged boar spermatozoa, respectively, and Fig. 4 shows the corresponding gradient magnitude images.

Next we determine the gradient magnitude along the cell head boundary contour as a 1D function of the boundary curve length from the reference point to a given contour point in a clock-wise direction. The cell head boundary is defined by the segmentation process presented above. As illustrated by Fig.2, the cell head boundary obtained in this way does not coincide everywhere with the local gradient maximum. To compensate for this effect, for each contour point we take the local maximum of the gradient in a 5×5 neighborhood of that point. The resulting discrete function is a vector that we re-size by nearest-neighbor interpolation to a uniform length of 40 elements. We also normalize this vector by dividing it by its largest element, Fig. 5. The vectors obtained in this way for different sperm heads are used for LVQ.

3 Gradient Vectors Classification

3.1 LVQ training

In the following we analyse the performance of several classifiers obtained from the available example data. We are mainly interested in prototype based systems because they are easy to implement and provide insights into typical properties of the data. For comparison, we furthermore use the k-nearest-neighbor (knn) approach.

In the following, the construction or training of classifiers is based on a (sub-) set of the labelled training data, denoted as $I\!D = \{\boldsymbol{\xi}^\mu, S_T^\mu\}_{\mu=1}^P$. Here, the $\boldsymbol{\xi}^\mu \in I\!R^N$ $(N = 40)$ are the vectors of gradient magnitudes along the contour as described in the previous section. In total, 320 examples are available, 176 of those belong to class 1 (acrosome-intact) while 144 represent class 2 (acrosome-damaged). The class membership will be denoted as $S_T^\mu \in \{1,2\}$ accordingly. Fig. 5 displays three example profiles from each of the two classes.

Throughout the following we will use the quadratic Euclidean distance measure when comparing two N-dimensional vectors $\boldsymbol{x}, \boldsymbol{y} \in I\!R^N$:

$$d(\boldsymbol{x},\boldsymbol{y}) \;=\; (\boldsymbol{x}-\boldsymbol{y})^2 \;=\; \sum_{j=1}^{N} (x_j - y_j)^2\,. \tag{3}$$

While this choice is not necessarily optimal, we restrict the analysis to the simplest metric here and deal with the selection of potentially more suitable measure in a forthcoming project.

3.2 K-Nearest-Neighbor classifier

We evaluate the performance of this standard approach (see e.g. [8]) using leave-one-out estimation of the classification error. For each of the available profiles $\boldsymbol{\xi}^\nu \in I\!D$ we compute its distances $d(\boldsymbol{\xi}^\nu, \boldsymbol{\xi}^\mu)$, to all other profiles $\boldsymbol{\xi}^\mu \in I\!D, \mu \neq \nu$, and determine the k profiles closest to $\boldsymbol{\xi}^\nu$. The majority of labels S_T^μ of these k closest profiles determines to which class, denoted as $S_{knn}^\nu \in \{1,2\}$, the profile $\boldsymbol{\xi}^\nu$ is assigned by the knn classifier.

The leave-one-out estimation of the over-all classification error, denoted by ε_{knn}, is defined as the fraction of profiles $\boldsymbol{\xi}^\nu$ to which knn labels are assigned that are different from their respective true labels, $S_{knn}^\nu \neq S_T^\nu$. Analogously, the errors $\varepsilon_{knn}^{(1)}$ and $\varepsilon_{knn}^{(2)}$ quantify the performance regarding data from class 1 or class 2 only, respectively.

3.3 Prototype based classification

The above discussed knn approach requires the explicit storage of all example data and the frequent evaluation of many distances. Hence, it is an attractive concept to represent the example set by only a few *prototype* vectors and classify novel data according to a closest prototype scheme, instead.

The conceptually simplest set of prototypes is obtained from a training set $I\!D$ by evaluating the class-conditional means

$$\boldsymbol{m}_1 = \frac{1}{P_1} \sum_{\mu=1}^{P} \boldsymbol{\xi}^\mu \, \delta(S_T^\mu, 1) \quad \text{and} \quad \boldsymbol{m}_2 = \frac{1}{P_2} \sum_{\mu=1}^{P} \boldsymbol{\xi}^\mu \, \delta(S_T^\mu, 2)$$

where $\delta(k,l)$ is the Kronecker-delta and $P_S = \sum_\mu \delta(S_T^\mu, S)$ is the number of examples from class $S = 1, 2$, respectively. The resulting classifier assigns a vector $\boldsymbol{\xi}$ to class 1 if $d(\boldsymbol{m}_1, \boldsymbol{\xi}) \leq d(\boldsymbol{m}_2, \boldsymbol{\xi})$ and to class 2 else.

We furthermore apply Learning Vector Quantization (LVQ) for the identification of sets of class prototypes. LVQ was originally proposed by Kohonen and has been used in a variety of problems due to its flexibility and conceptual clarity [14, 16].

In a set of prototypes $\left\{\boldsymbol{w}^1, \boldsymbol{w}^2, \ldots, \boldsymbol{w}^M\right\}$ a vector $\boldsymbol{w}^j \in I\!R^N$ is supposed to represent data with class membership $S^j \in \{1, 2\}$. These assignments, as well as the number of prototypes are specified prior to training and remain unchanged.

At each time step t of the iterative training procedure, one example $\{\boldsymbol{\xi}^\mu, S_T^\mu\}$ is selected randomly from $I\!D$. We evaluate the distances $d(j,\mu) = d(\boldsymbol{\xi}^\mu, \boldsymbol{w}^j(t))$ of the presented profile from all current prototype vectors. Next we identify the minimal distance $d(J,\mu)$ among all prototypes and the corresponding *winnner*

$$\boldsymbol{w}^J(t) \;\text{ with }\; d(J,\mu) = \min_k \left\{d(k,\mu)\right\}. \tag{4}$$

In LVQ1, only this winner is updated according to

$$\boldsymbol{w}^J(t+1) = \boldsymbol{w}^J(t) + \eta(t) \left[1 - 2\,\delta(S_T^\mu, S^J)\right] \left(\boldsymbol{\xi}^\mu - \boldsymbol{w}^J(t)\right), \tag{5}$$

where the update is towards (away from) the actual input $\boldsymbol{\xi}^\mu$, if the class labels of winner and example agree (disagree).

In the following studies, prototype vectors are initialized close to the corresponding class conditional mean. In order to avoid exactly coinciding $\boldsymbol{w}^j(0)$, we set

$$\boldsymbol{w}^j(0) = c\; \boldsymbol{m}_{S^j} + (1-c)\; \boldsymbol{\xi}$$

where $\boldsymbol{\xi}$ is a randomly selected training example from class S^j. Throughout the following, the coefficient c was fixed to 0.9. Our results depend only very weakly on details of the initialization.

The learning rate $\eta(t)$ controls the step width of the iteration. It is gradually decreased in the course of learning following a schedule of the form

$$\eta(t) = \eta_o \,/(1 + \, a\, t) \quad \text{with} \;\; a \;\; \text{such that} \;\; \eta(t_f) = \eta_f. \tag{6}$$

Results presented in the next section were obtained with a schedule that decreases the learning rate from $\eta_o = 0.1$ to $\eta_f = 0.001$ in 1000 sweeps through the training data, i.e. $t_f = 1000\, P$ in Eq. (6) where P is the number of examples in the training set.

After training, the system parameterizes a distance-based classification scheme: a vector $\boldsymbol{\xi}$ is assigned to the class S^J which is represented by the closest prototype.

3.4 Cross-validation

In order to obtain estimates of the performance after training we employ eight-fold cross-validation: We split the set of 320 available training data randomly into disjoint subsets $I\!D_i, i = 1, 2 \ldots 8$, of equal size.

For a given number of prototypes, each of eight identically designed LVQ systems, $n = 1, 2 \ldots 8$, is trained from the set $\cup_{i \neq n} I\!D_i$ containing $P = 280$ examples. Then, $I\!D_n$ serves a test set to evaluate the performance on novel data.

In addition, this procedure is repeated $n_{rep} = 20$ times, each time with a randomly shuffled set of data. The additional average performed over the n_{rep} cross-validation results suppresses the influence of random fluctuations due to *lucky set* compositions.

In the following, ε_{train} denotes the fraction of misclassified example data, obtained after training and on average over the eight systems and 20 runs. Correspondingly, the test error ε_{test} quantifies the averaged performance with respect to the test set. We furthermore evaluate the class-specific test errors $\varepsilon_{test}^{(1)}$ and $\varepsilon_{test}^{(2)}$ as well as the training errors $\varepsilon_{train}^{(1)}$ and $\varepsilon_{train}^{(2)}$ with respect to only class 1 or class 2 data, respectively.

The main purpose of the cross-validation scheme is to compare the performance of different LVQ schemes, i.e. systems with different numbers of prototypes. We also apply it to the classifier constructed from the class-conditional means.

4 Results

4.1 KNN classifier

The dependence of the performance on the number of neighbors k turns out to be rather weak, Table 1. The knn-classifier performs relatively well with respect to class 1 data with $\varepsilon_{knn}^{(1)} < 0.03$, while $\varepsilon_{knn}^{(2)}$ is significantly larger ($\varepsilon_{knn}^{(2)} > 0.13$). This reflects the fact that gradient profiles of acrosome-damaged cells display a greater variability than those of the acrosome-intact spermatozoa; we elaborate on this aspect in the discussion below. The smallest overall test error $\varepsilon_{knn} \approx 0.075$ is obtained for the range $43 \leq k \leq 98$, while even larger values of k result in an inferior performance.

4.2 Prototype based classifiers

A simple two prototype system based on class-conditional means yields the estimates for test and training error (denoted in $\%$) is showed in Table. 2.

Clearly, the different runs of the training procedure are statistically dependent as the training sets $I\!D_i$ overlap. Nevertheless, the observed standard deviations can serve as a rough measure of statistical variations and we estimate the accuracy of the above results to be on the order of 0.003, i.e. 0.3%. Within this margin, the simple scheme based on class-conditional means competes well or even outperforms the more involved knn-classifier. It furthermore yields a more balanced classification in the sense that the difference between $\varepsilon^{(1)}$ and $\varepsilon^{(2)}$ is smaller.

Figure 6, left panel, shows the class conditional mean profiles as obtained from one of the training sets in the cross-validation scheme. The main distinctive feature appears to be a pronounced maximum (minimum) in the gradient magnitude for class 1 (class 2) data, respectively. The right panel of Fig. 6 displays a scatter plot of distances from the class conditional mean vectors. The observed statistical variance of distances is somewhat larger in class 2 (approx. 0.75) than in class 1 (approx. 0.66), reflecting the greater variability of acrosome-damaged cells.

We have furthermore performed LVQ1 training following the above described scheme for systems with m_1 (m_2) prototypes representing class 1 (class 2) data, respectively. The Table.3 summarizes the observed classification errors in different LVQ configurations. In each entry, the upper row corresponds to the test errors (overall, class 1, and class 2) while the training errors are given in the lower row (all results in $\%$).

Also here, the statistical errors of the results are estimated to be on the order of 0.003 (i.e. 0.3%).

In the minimal setting with $m_1 = m_2 = 1$, the performance of LVQ1 differs only slightly from that of using the class-conditional means as prototypes. However, the introduction of several prototypes allows for representation of more details and, hence, better classification.

The best overall performances, marked by bold-face figures in the table, are obtained when two or more prototypes are used in each class, i.e. $m_1, m_2 \geq 2$. Within the estimated precision all these configurations yield an overall test error of $\varepsilon_{test} \approx 6.8\%$

Figure 7 displays the six LVQ prototypes as obtained in one of the training runs with $m_1 = m_2 = 3$ as an example. Note how the class 1 prototypes have *specialized* to represent different types of acrosome-intact profiles with a more or less pronounced maximum in the center. Within class 2, all prototypes display a minimum close to the center of the profile, i.e. the *head* of the spermatozon. They differ, however, in their symmetry properties.

The use of a larger number of prototypes, i.e. a more complex LVQ system, yields a decrease of training errors. However, test erros remain unchanged or even increase moderately which signals *over-fitting*: While the particular training set can be represented in greater detail with many prototypes, the generalization ability may deteriorate.

Based on the currently available data, we conclude that the LVQ classifier with $m_1 = m_2 = 2$ gives the best possible performance, already.

5 Summary and conclusions

We use grey level images of boar sperm samples obtained with a phase-contrast microscope. We segment the sperm cell heads and for each cell we compute a vector that comprises the values of the gradient magnitude along the cell contour. The vector obtained for a cell is the information that we use for the classification of the cell as acrosome-intact or acrosome-damaged.

For classifier training we use vectors for which its is known from staining results whether they stem from cells with intact or damaged acrosome. We deploy prototype based LVQ classifiers with different numbers of prototypes. We analyse the fraction of example data ε_{train} and test data ε_{test} misclassified for different LVQ schemes: considering up to four prototypes for each class. Obtained errors show that already the simplest scheme (with just one prototype of each class) can compete with a costly k-nearest-neighbor classifier. LVQ systems with two or more prototypes per class can realize lower training and test errors. We find the best overall test error $\varepsilon_{test} \approx 0.068$ for $m_1 = m_2 = 2$, already. This performance is sufficient for practical purposes. Note that the ultimate goal is not the classification of single spermatozoa, but the reliable estimation of the fraction of acrosome-damaged cells in a large sample.

In future investigations we intend to apply more sophisticated cost function based schemes such as *Generalized Learning Vector Quantization* (GLVQ) as suggested in [23]. Modified distance measures and *relevance learning*, see e.g. [11], could be applied in order to obtain a better understanding of this classification task and to extract the most relevant features.

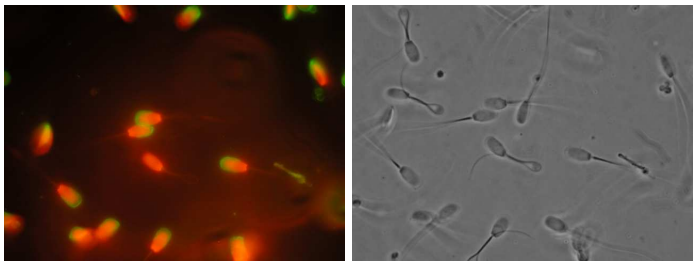
6 Acknowledgment

This work was supported by the Spanish Ministerio de Educación y Ciencia under grant No DPI2006-02550.

References

1. E. Alegre, R. Fernández, L. Sánchez, V. Rodríguez, and J. Domínguez. Digital image segmentation methods for automatic quality evaluation of boar semen. *Reproduction in Domestic Animals*, 40:392–392, 2005.
2. T. Alonso, E. Alegre, L. Sánchez, F. Tejerina, and J. Domínguez. Segmentación y etiquetado automático de cabezas de espermatoides de verraco. In *XXVII Jornadas de Automática. Almería*, Septembre 2006.
3. M. Beletti and L. Costa. A systematic approach to multi-species sperm morphometrical characterization. *Analytical and Quantitative Cytology and Histology*, 25(2):97–107, 2003.
4. M. Beletti, L. Costa, and M. Viana. A comparison of morphometric characteristics of sperm from fertile Bos taurus and Bos indicus bulls in Brazil. *Animal Reproduction Science*, 85:105–116, 2005.
5. M. Beletti, L. Costa, and M. Viana. A spectral framework for sperm shape characterization. *Computers in Biology and Medicine*, 35(6):463–473, 2005.
6. M. Biehl, P. Pasma, M. Pijl, L. Sanchez, and N. Petkov. Classification of boar sperm head images using learning vector quantization. In M. Verleysen, editor, *Proc. European Symposium on Artificial Neural Networks (ESANN), Brugge, April 26-28, 2006*, pages 545–550. d-side, Evere, Belgium, 2006.

7. J. F. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.
8. R. Duda, P. Hart, and D. Stork. *Pattern Classification*. J. Wiley & Sons, New York, 2001.
9. R. Fernández, E. Alegre, L. Sánchez, F. Tejerina, R. González, and J. Domínguez. Automatic detection and measuring of distal droplets in boar spermatozoa using computer vision techniques. In *Proc. 9th Annual Conference European Society for Domestic Animal Reproduction, ESDAR 2005, Reproduction in Domestic Animals*, volume 40, page 390, 2005.
10. C. Grigorescu, N. Petkov, and M. A. Westenberg. Contour and boundary detection improved by surround suppression of texture edges. *Image and Vision Computing*, 22(8):609–622, 2004.
11. B. Hammer and T. Villmann. Generalized relevance learning vector quantization. *Neural Networks*, 15:1059–1068, 2002.
12. R. Harrison and E. Vickers. Use of fluorescent probes to assess membrane integrity in mammalian spermatozoa. *Journal of Reproducton and Fertility*, 88:343–352, 1990.
13. M. Hirai, A. Boersma, A. Hoeflich, E. Wolf, J. Foll, T. Aumuller, and J. Braun. Objectively measured sperm motility and sperm head morphometry in boars (Sus scrofa): relation to fertility and seminal plasma growth factors. *J. Androl*, 22:104–110, 2001.
14. T. Kohonen. *Self-organizing maps*. Springer, Berlin, 1995.
15. C. Linneberg, P. Salamon, C. Svarer, and L. Hansen. Towards semen quality assessment using neural networks. In *Proc. IEEE Neural Networks for Signal Processing IV*, pages 509–517, 1994.
16. Neural Networks Research Centre. *Bibliography on the Self-Organizing Map (SOM) and Learning Vector Quantization (LVQ)*. Helsinki University of Technology, 2002.
17. G. Ostermeier, G. Sargeant, T. Yandell, and J. Parrish. Measurement of bovine sperm nuclear shape using Fourier harmonic amplitudes. *J. Androl.*, 22:584–594, 2001.
18. N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9:62–66, 1979.
19. A. Quintero, T. Rigaub, and J. Rodríguez. Regression analyses and motile sperm subpopulation structure study as improving tools in boar semen quality analysis. *Theriogenology*, 61:673 – 690, 2004.
20. T. Rijsselaere, A. V. Soom, G. Hoflack, D. Maes, and A. de Kruif. Automated sperm morphometry and morphology analysis of canine semen by the Hamilton-Thorne analyser. *Theriogenology*, 62(7):1292–1306, 2004.
21. L. Sánchez, N. Petkov, and E. Alegre. Classification of boar spermatozoid head images using a model intracellular density distribution. In M. Lazo and A. Sanfeliu, editors, *Progress in Pattern Recognition, Image Analysis and Applications: Proc. 10th Iberoamerican Congress on Pattern Recognition, CIARP 2005, Lecture Notes in Computer Science*, volume 3773, pages 154–160. Springer-Verlag Berlin Heidelberg, 2005.
22. L. Sánchez, N. Petkov, and E. Alegre. Statistical approach to boar semen head classification based on intracellular intensity distribution. In A. Gagalowicz and W. Philips, editors, *Proc. Int. Conf. on Computer Analysis of Images and Patterns, CAIP 2005, Lecture Notes in Computer Science*, volume 3691, pages 88–95. Springer-Verlag Berlin Heidelberg, 2005.
23. A. Sato and K. Yamada. Generalized learning vector quantization. In G. Tesauro, D. Touretzky, and T. Leen, editors, *Advances in Neural Information Processing Systems*, volume 7, pages 423–429. MIT Press, 1995.
24. L. Schwartz. *Théorie des Distributions*. Vol. I, II of Actualités scientifiques et industrielle. L'Institute de Mathématique de l'Université de Strasbourg, 1950-51.
25. T. Suzuki, H. Shibahara, H. Tsunoda, Y. Hirano, A. Taneichi, H. Obara, S. Takamizawa, and I. Sato. Comparison of the sperm quality analyzer IIC variables with the computer-aided sperm analysis estimates. *International Journal of Andrology*, 25:49–54, 2002.
26. H. Tagare and R. de Figueiredo. On the localization performance measure and optimal edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(12):1186–1190, 1990.
27. J. Verstegen, M. Iguer-Ouada, and K. Onclin. Computer assisted semen analyzers in andrology research and veterinary practice. *Theriogenology*, 57:149–179, 2002.



(a) (b)

Fig. 1. Images of a boar semen sample acquired with (a) a flourescence microscope and (b) a phase-contrast microscope.

(a) (b) (c)

Fig. 2. (a) Grey level image of a sperm head and a part of the middle piece protruding from the head. (b) Binary image of the sperm head obtained by thresholding and subsequent morphological processing. (c) Contour of the binary head mask (white line) superimposed on the gradient magnitude image.

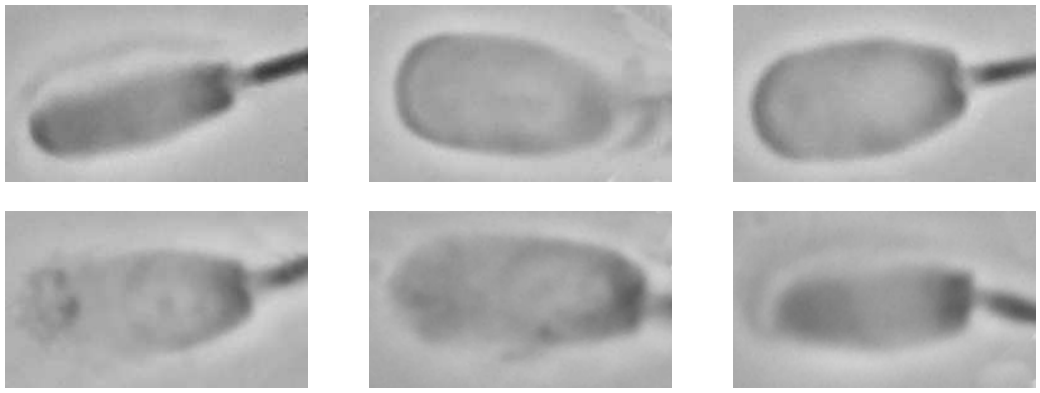


Fig. 3. Grey level images characteristic of (upper row) acrosome-intact and (lower row) acrosome-damaged boar spermatozoa.

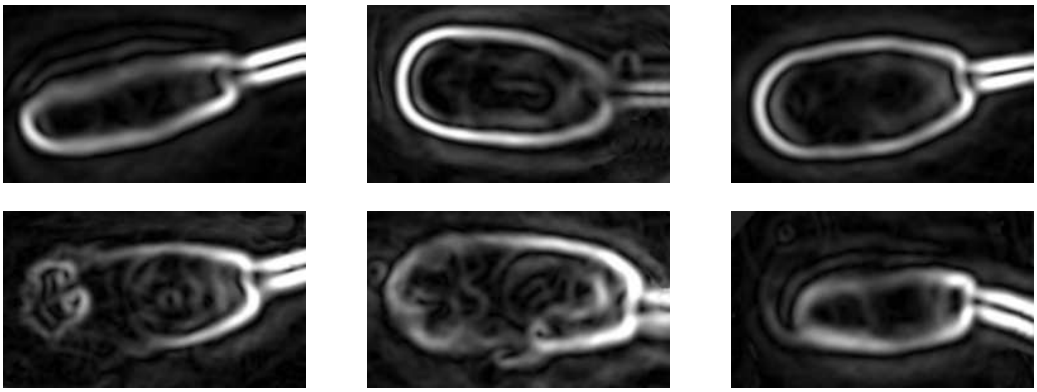


Fig. 4. Gradient magnitude images characteristic of (upper row) acrosome-intact and (lower row) acrosome-damaged boar spermatozoa. The parameter σ used to compute the scale-dependent gradient is set to 0.03 of the cell head length.

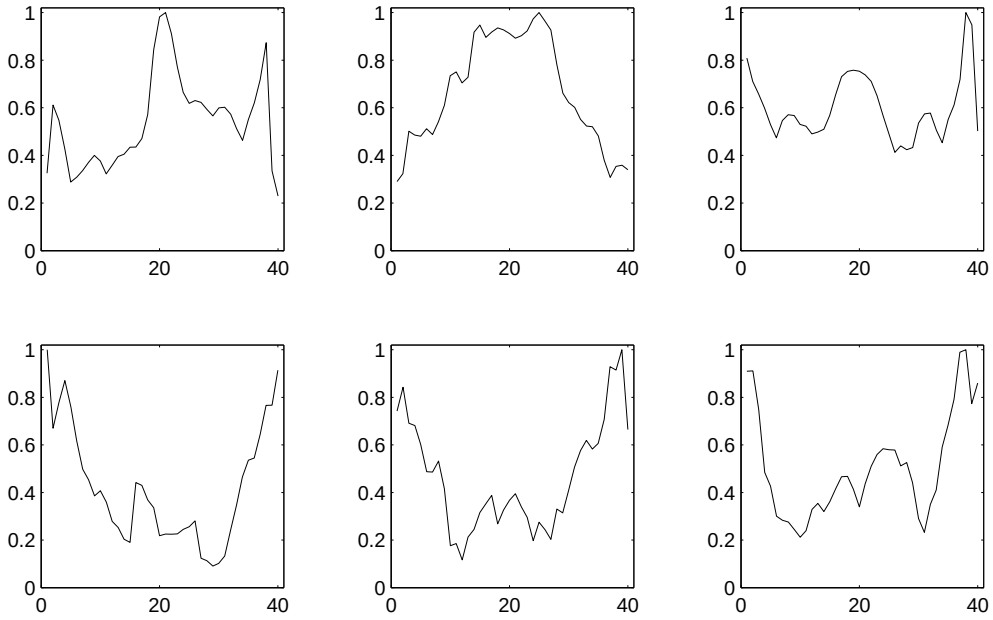


Fig. 5. Example gradient magnitude profiles along the head boundary $\boldsymbol{\xi} \in I\!R^{40}$ from the class of (upper row) acrosome-intact and (lower row) acrosome-damaged spermatozoa, respectively. The displayed profiles correspond to the images shown in Figs. 3 and 4. These discrete 1D functions represent the vectors used for LVQ.

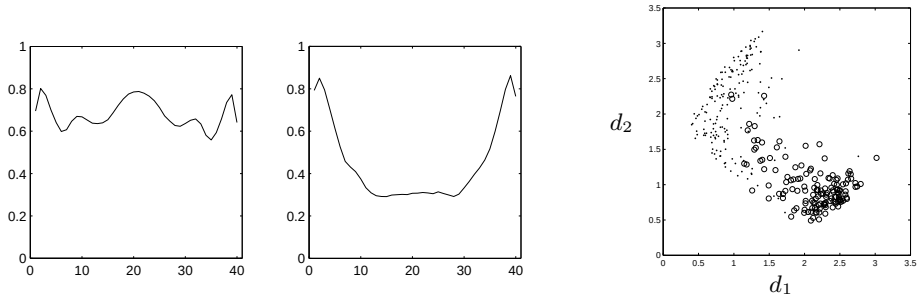


Fig. 6. Left panel: The class conditional means as obtained from one set of 280 training examples. The left and right profile represent class 1 of acrosome-intact and class 2 of acrosome-damaged cells, respectively. **Right panel:** Scatter plot of distances $d_1 = d(\boldsymbol{\xi}, \boldsymbol{m}_1)$ and $d_2 = d(\boldsymbol{\xi}, \boldsymbol{m}_2)$. Dots represent class 1 data while open circles correspond to class 2 data.

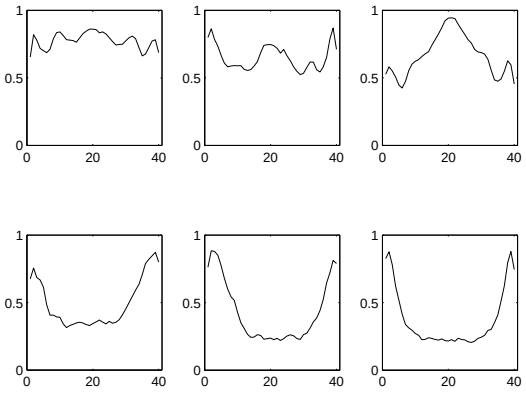


Fig. 7. Prototype profiles obtained in LVQ1 with $m_1 = m_2 = 3$ from one set of 280 training examples. The upper (lower) row of prototypes represent class 1 (class 2) data, respectively.

k	ε_{knn}	$\varepsilon_{knn}^{(1)}$	$\varepsilon_{knn}^{(2)}$
1	8.4	2.3	16.0
3	9.4	2.8	17.4
23	8.4	2.3	16.0
33	7.8	2.3	14.6
43	7.5	2.3	13.9

Table 1. Classification errors using K-NN (errors are in %).

	ε	$\varepsilon^{(1)}$	$\varepsilon^{(2)}$
ε_{test}	7.6	6.4	9.1
ε_{train}	7.6	6.4	9.0

Table 2. Test and training errors with two prototypes (errors are in %).

m_2	1			2			3			4		
m_1	ε	$\varepsilon^{(1)}$	$\varepsilon^{(2)}$	ε	$\varepsilon^{(1)}$	$\varepsilon^{(2)}$	ε	$\varepsilon^{(1)}$	$\varepsilon^{(2)}$	ε	$\varepsilon^{(1)}$	$\varepsilon^{(2)}$
1	7.8	6.0	9.8	8.1	7.7	8.5	8.4	8.9	7.9	8.6	8.9	8.1
	7.8	5.9	10.0	7.8	7.4	8.3	7.8	8.1	7.3	7.6	8.0	7.0
2	7.7	3.0	13.2	**6.8**	4.3	10.2	7.2	4.9	10.1	7.3	5.1	10.0
	7.5	2.9	13.1	6.7	4.5	10.1	6.8	4.3	9.8	6.7	4.4	9.6
3	7.8	3.0	13.8	**6.7**	3.4	10.4	**6.8**	3.8	10.3	7.1	4.4	10.1
	7.6	2.8	13.5	6.5	3.4	10.3	6.6	3.7	10.3	6.6	3.8	10.0
4	7.8	2.9	13.7	**6.7**	3.7	10.3	**6.8**	3.7	10.5	7.1	4.5	10.2
	7.7	2.8	13.6	6.5	3.4	10.2	6.4	3.3	10.2	6.5	3.7	9.9

Table 3. LVQ1 training for m_1 and m_2 prototypes (errors are in %).