New Hybrid Approach for Identification of Spermatozoa in Human Semen Sample using Microscope Image Processing Techniques

[1] **Nandini M. Chaudhari** and [2]**B. V. Pawar**
North Maharashtra University, Jalgaon, Maharashtra, India
[1]nandini113@yahoo.co.in, [2]bvpawar@gmail.com

ABSTRACT

Due to very hectic, stressful and exhaustive life style, Today, Infertility is a common problem in most of the couples and they have to undergo for intrauterine insemination (IUI) or In-vitro fertilization (IVF). Semen analysis is fundamental test in the clinical work up of the infertile male. Computerized semen test will give many more & accurate parameters that are useful to the fertility specialist. This paper focuses on finding the numbers of spermatozoa contain in the semen sample with new hybrid approach using MATLAB. Use of background subtraction is the first stage of algorithm. Region based segmentation finds the numbers of sperm in the semen sample. The comparative analysis of segmentation shows that region base segmentation detects more objects. So it is possible to find the suppressed (low intensity) sperm in given frame.

Keywords: Semen analysis, region segmentation, spermatozoa, WHO scan line algorithm

1   Introduction

A Semen analysis is a routine pathological test which has an important role in the human reproduction. This test finds some of physical and chemical parameters. A semen analysis evaluates certain characteristics of a male's semen and the sperm contained in the semen. A routine analysis of a given ejaculate includes the evaluation of sperm concentration, motility, morphology, and vitality etc.

Some of the parameters that are calculated through microscope manually which include:

- Sperm count (technically it is the concentration) is the number of sperm in the semen sample
- Motility is the percentage of sperm those are moving in the sample. According to speed of the sperm they are classified into Type A (fast and progressive), Type B (slow progressive), Type C (Non-progressive) , Type D (immobile)
- Morphology is the shape or structure of spermatozoa

The normal values of semen parameters are referred from the manual of World Health Organization (WHO). As per WHO manual 4th & 5th edition, normal values of semen are slight variable and are shown in Table 1[1][2].

Table 1: Standard values of Semen analysis as per WHO

	WHO 1999 4th Edition	WHO 2011 5th Edition
Volume	2 ml	1.5 ml
Concentration	20 million /ml	15 million /ml
Progressive motility	50%	32%
Normal form	14%	4%

Computer assisted semen analysis (CASA) plays a vital role for the assessment of sperm concentration (counts) and motility (movement) characteristic. Most of the CASA analyzers are based on different techniques on image analysis. Various CASA system like Cell Soft (Cryo Resources Ltd, NY), Hamilton Thorne (Hamilton Thorne Research Inc, Danvers, MA) etc are used widely all over. Different technology, platforms and algorithms were embedded into different CASA systems.

Early CASA system uses signal processing technology to find the density of the semen. Lukasz Witkowski, Poland, in 2002, examine the density and trajectories of the cell movement using dilation and erosion techniques and the algorithm was developed in PASCAL [3] [4]. Since sperm finding is the first step in semen analysis it is needed to detect the individual sperm which are identified even in low contrast images. Vahid et. al. had proposed the tracking & fitting algorithm in MATLAB to identify the individual complete (head & tail) spermatozoa as an elliptical shape in 2005 [5]. Henry Corrilo et.al. [6] was focused on finding the features of spermatozoa. The method called nth-fusion was introduced in his work, the framework of the segmentation algorithm for region of interest head, Acrosome, Nucleus and Mid-piece was segmented.

J.M.Pascual-Gaspar et. al. [7] in 2008 had proposed a simple and effective system for computer assisted semen analysis took lot of efforts to evaluate the density of semen and motility of individual spermatozoa by using simple but effective computer vision techniques.

Hiromasa Oku et al. [8] reported a new automatic visual tracking system for small, fast-moving objects using high-speed visual feedback. He achieved stable tracking of fast, small ascidian spermatozoa for more than 180 s using high-speed visual feedback. The two stage algorithm proposed by Abhirramy focused on the motion detection and motion tracking of spermatozoa with some preprocessing operations like image enhancement using Laplacian filter & median filtering. Most basic morphological operations like dilation, erosion and thresholding using gradient operators Sobel & Edge. Features of identified objects was stored and labeled them accordingly. Identified spermatozoa were tracked in each frame [9]. The review on techniques of CASA for semen analysis indicates that Pathologist and microbiologist need a manual cum standard computer assisted, cost effective semen analyzer with optimal result, since semen analysis is now routine checkup for infertile couples. [10]

2 Microscope Image Processing Techniques

Techniques employed for processing microscopy images are in general rather different from those encountered in consumer image processing. Most applications will have a suite of standard operations such as sharpening, contrast and brightness enhancement, and color balancing.

One key requirement is to remove the distortion in optical path microscope & blurring caused by instrumentation.

2.1 Image subtraction

Image subtraction is an image correction techniques and calibration methods for both image visualization and quantitative analysis.

Image subtraction is usually performed when two images of the same object are obtained under different conditions. The image subtraction will highlight whatever has changed between the two images. Another application is background correction. In microscopic imaging the image is often affected by a slowly varying background shading pattern. One can move the microscope stage to an empty field and acquire an image of the background. When the background image is subtracted from the image containing the specimen, it removes the shading.

$$Imwork(x, y) = Imspeciman(x, y) - Imback(x, y)$$

Where Imspeciman(x, y) is a digital image produced by the microscope, Imback(x, y) is a background image produced by low-pass filtering the specimen image to smooth, where the smoothing effect is larger than the size of objects in the image [11].

2.2 Region based segmentation

The main aim of segmentation is the division of digital image into group of connected pixels with identical properties i.e. region which is easier to analyze. There are three general approaches for segmentation: Edge, Boundary and Region. In edge based segmentation it finds all pixels on region on boundary i.e. divide into edge and non-edge pixels. Region based segmentation, works repetitively by grouping pixel together which are neighbors (4-connectivity or 8-connectivity) and have same values. Boundary based segmentation pixels are marked on the basis of range of value in which pixel lies.

3 Proposed Hybrid Image Processing Approach to Detect the Spermatozoa

Since the sperms are rapid motile, it is very difficult to count the number of sperm manually, but by capturing them in video and then count is easy way. So we have proposed a new system in which the frames extracted from recorded video is used as an input to the proposed system. We have called it a hybrid approach since after identifying the moving object, scan line algorithm is applied for confirmation of the objects having tails, so that we can count the actual number of spermatozoa. The standard data set was prepared from the different semen samples. For testing the efficiency of proposed algorithm, same frames were applied to the existing algorithm.

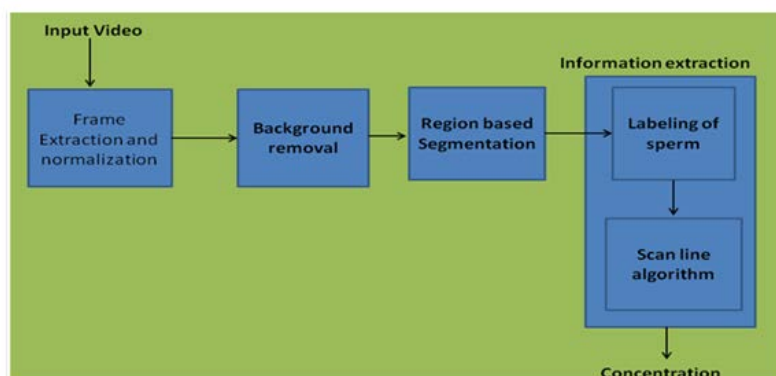The architectural framework for the propose hybrid system is shown in figure 1.

Figure 1: Proposed hybrid system

The propose algorithm contains the following steps-

- Frame extraction & Normalization
- Background creation & subtraction for background removal
- Region based sperm detection
- Information extraction
- Find concentration i.e. No of sperms

By using the background subtraction technique the algorithm find the objects. By avoiding the preprocessing steps we have improved the computation of the algorithm. Region based segmentation with region to be connected as 4-connect is used to find the area of interest that may an object which has sperm data. Actual spermatozoa are identified by detecting the tail of the sperm. So before finding the concentration we need to check whether the identified object had tail with help of scan line algorithm.

3.1   Frame extraction & Normalization

3.1.1   Experimental video setup

Most of the pathological laboratories are doing semen analysis by measuring the sperm concentration, motility, vitality by manual method as shown in figure 2a. In most of the cases, sperm concentration and motility are improved manually using nebular counting chamber, Makler Chamber. Makler chamber is a counting chamber Medical instrument for microscope which allows you to count the sperm chemberwise. Setting up the video setup with microscope the frame can be extracted with the recorded video as shown in figure 2b with help of video to s-video adapter with the block diagram as shown in figure 2c. Microscope is LABOMED make LED cluster with UPS & SMF battery and AVC 561V Digital Color CCD camera with WinPvr. A video is a sequence of images or frames with time difference in second. The frames with recorded videos are converted into frames as a bitmap file with file size 352 X 576.
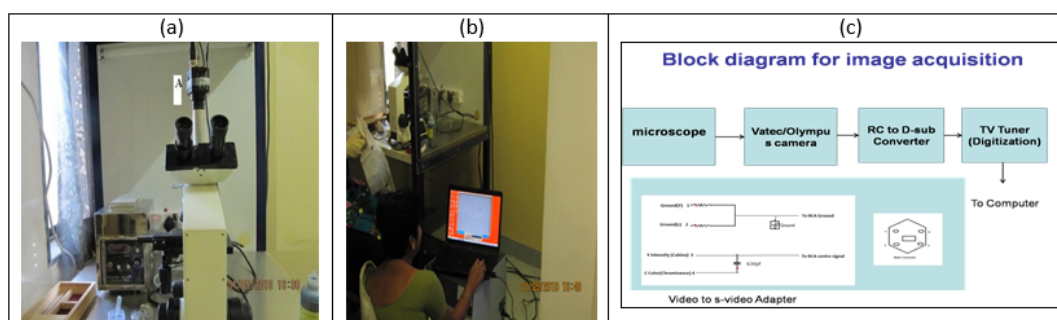

Figure 2: manual method & video setup with microscope

3.2   Background creation & subtraction for background removal

This is a first stage of the algorithm, instead of preprocessing stages and frame enhancement; we have created the background and remove the background from all frames, which is created from the first frame. Background subtraction is done pixel by pixel. The background image is created from the first frame of the video. Process for extraction if the difference Dvalue is satisfying predefined threshold.

```
Algorithm:
Background removal using background subtraction
[Input]: Frames Fn. {F1, F2, F3 …Fn}
[Output]: Region nm obtained from each frame
Where n is frame no   & m is a no of sperms in individual
frames.
Pick the background image say Bk
For i= 1 to n
Dvalue ← Fi – Bk   Where n is the no. of frames to be
processed
If (Dvalue <= Threshold)
Dvalue=0
else
Extract the region of interest (ROI) whose min area is> 4.
End
Stop
```

3.3   Region based sperm detection

3.3.1   Why region based segmentation?

The numbers of objects detected in region based segmentation technique are more due to the intensity difference of region as shown in figure 3. But here we can find the sperms which are at lower level may be not having very good intensity and cannot visualize by naked eyes. So the chances of finding number of sperm can be more by checking the region of interest as shown in table 2, by verifying whether the ROI contains sperm data we can label the object.
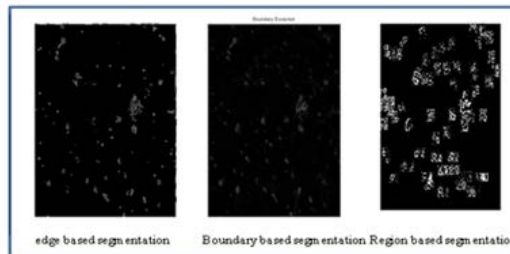

Figure 3: result of different segmentation techniques

Table 2: no of ROI generated by segmentation techniques

Frame number	Methods type		
	Edge based	Boundary based	Region based
1	800	423	1339
2	812	362	4983
3	818	350	2999
4	813	377	5255
5	806	374	5520
6	808	355	4749
7	817	386	6033
8	816	371	6140
9	806	354	5664
10	836	369	6946

By avoiding the preprocessing steps we can improve the computation of the algorithm. Region based segmentation with region to be connected as 4-connect is used to find the area of interest that may an object which has sperm data.

```
Algorithm: Region based sperm detection
Convert current instance of frame to grayscale
Label 4-connect component
After Applying region based segmentation measure
the properties of contiguous regions.
Extract the region of interest (ROI) whose min area
is 4.
Extracting all the parameter into the detector
region into the form of file & in cell data structure a
structural array.
```

3.4   Information extraction

There are two stages in this step, one is labeling of the object and check if tail exists in the detected region so as it is a sperm. First Detect the bounding box value of each detector region, crop the image & Check whether it contain sperm data. The identified objects from first five frames are labeled and are shown in figure 4.

```
[Input]: Information of detected region in individual
frames
[Output]: Structural array contains sperm information

Part1 Labeling of sperm

Here the tail information is extracted and the sperm is
labeled and detect its variations in all frames.
For i= 1 to m
  r ← Ri     Where m is the total no. of region
obtained
For j = 1 to n    Where n is no. of frames
  r ' ← Rij Where Rij is a region of its order in frame j.

  If (r-r' <threshold)
  Accept it as a same region
else
     reject  r'  and  calculate  its  variations(v)  &
distance(d)
     [v,d] ← calculate (r')
  r[v,d] ← [v,d]
End
End
Sc < r then c++
Stop
```

3.5   Find concentration i.e. No of sperms

After extraction of information and storing it in mat file we need to check for the redundancy to find the actual number of sperms exists in the semen sample. Load the .mat file and compare the current file to next file. If the parameters are same then the same sperm is detected, we need to find the variation of the centroid otherwise store it as another sperm and count.

```
[Input] .mat file generated for individual sperm in each
frame
[Output] Count the number of sperm as a concentration
For i=1 to total .mat file
Load current .mat file
Load next .mat file file1
Compare sperm parameters of file and file1
If equal
   The sperm is detected in next frame, find its variation
Else
Count=count+1
End
```

4 Experiments, Results & Discussion

4.1   Comparative study

With setting up the microscope with video setup as shown in figure1 & extracting the frames at the rate of 25 FPS, we are trying to find the concentration of the raw semen sample. A proposed hybrid approach shows that detected labeled objects are more accurate than the existing system. The sperm which are not detected in the current frame gets detected in the next frame as shown in figure 4, So the chances for finding the more number of correct sperms. The number of object labeled in Table 3 shows that the accuracy for labeling of sperm by our proposed approached is near by the manual method. Figure5 indicates that the percentage of accuracy for labeling of the object is nearly 93.18%.
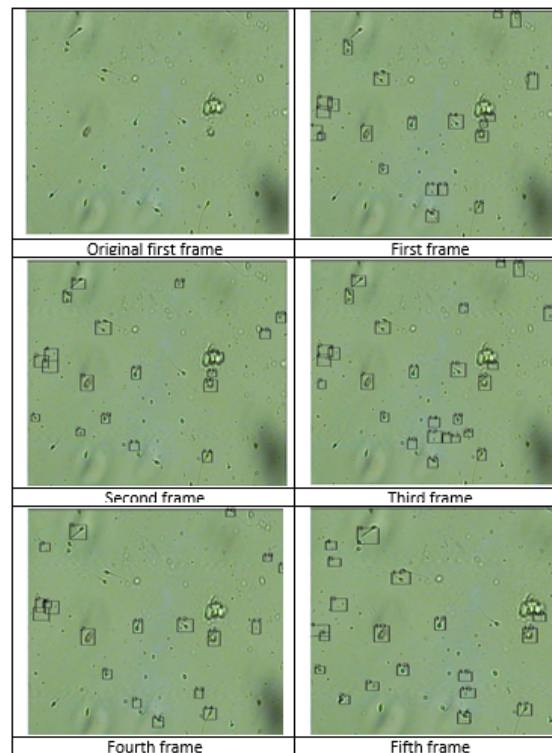

Figure 4: labeling of the sperm

Table 3: Actual no of labeled objects

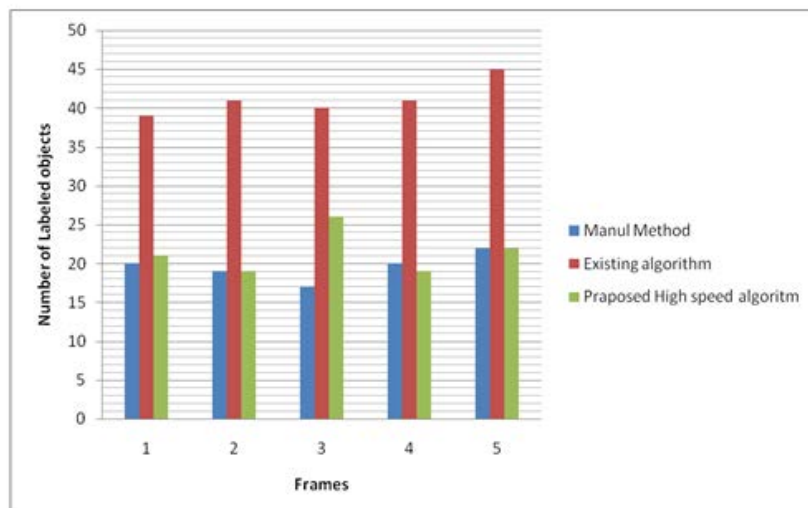| Frame No	Manual Method	Existing algorithm	Proposed Hybrid algorithm
1	20	39	21
2	19	41	19
3	17	40	26
4	20	41	19
5	22	45	22

Figure 5: Graph for Comparison graph of labeled objects in different method

4.2 Performance Evaluation

Evaluation of the proposed system is based on the performance measure for the rate of false positive, false negative, sensitivity, specificity & accuracy. Accuracy allows us to measure how well binary classification of detected sperms and non-detected sperms tests correctly identified or excludes. It is a proportion of true results for both true positive and true negative in the identified spermatozoa. The classification of TP, TN, FP, and FN are shown according to table 4.

Table: 4 Classification criteria for labeled objects

Labeled objects	Sperms	Debris/Artifacts
Detected object	TP	FP
Not detected object	TN	FN

According to above table the labeled spermatozoa, the accuracy, sensitivity and specificity can be calculated as

$$Accuracy = \frac{True\ Positive + True\ negative}{True\ Positive + True\ negative + False\ Positive + False\ negative}$$

$$Sensitivity = \frac{True\ Positive}{True\ Positive + False\ negative}$$

$$Specificity = \frac{True\ negative}{True\ negative + False\ Positive}$$

The no of true positive labeled is considered as a spermatozoa , As compared with the area and perimeter of the true spermatozoa the artifacts or debris are classified as a false negative and false positive. According the outcome of percentage of accuracy, sensitivity and specificity are tabularized in Table 5. The overall labeled sperms average accuracy is up to 70%. So as per discussion with experts to find semen analysis it is considered as a 100 percentile. So according to our proposed method the percentile varies in the range of 80% to 100%. The evaluation of these performance measures are shown in figure 6 in graphical format.

Table 5: Evaluation of the proposed hybrid approach

Frame No	No. of labeled object	Total sperm in the frame	detected sperm (TP)	Not detected sperm (TN)	non detected artifacts(FN)	detected artifacts as sperm (FP)	Accuracy	sensitivity	Specificity	Labeled Sperms Accuracy
1	21	12	9	3	4	7	0.5217	69.2308	30.0000	75.0000
2	19	12	7	5	3	5	0.6000	70.0000	50.0000	58.3333
3	26	10	8	2	4	4	0.5556	66.6667	33.3333	80.0000
4	19	8	5	3	3	3	0.5714	62.5000	50.0000	62.5000
5	22	8	6	2	4	4	0.5000	60.0000	33.3333	75.0000
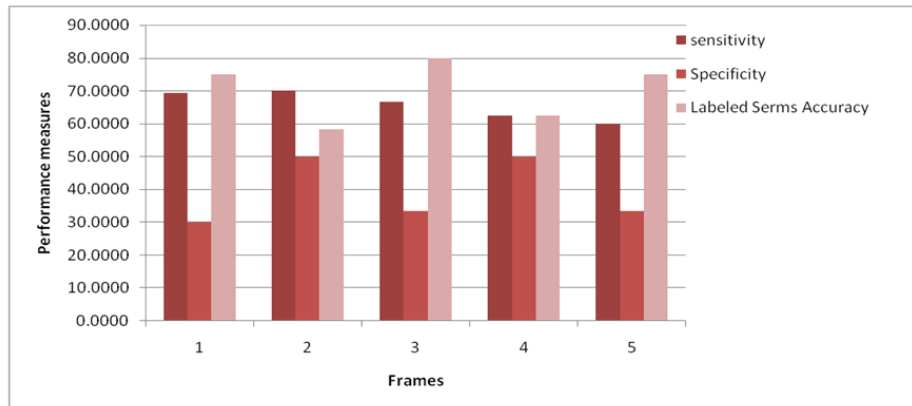
Figure 6: The evaluation of these performance measures

5 Conclusions

The number of objects (spermatozoa) detected by our hybrid approach for spermatozoa detection algorithm are at par as per the manual method with 83% of frame wise labeling. In this paper region based segmentation and scan line algorithm helps in finding the object as a sperm which are at lower level with low intensity and so the concentration. This is achieved without using any preprocessing operation like image enhancement. The overall labeled sperms with tails the accuracy percentage is 70%.

Acknowledgement

REFERENCES

[1]. Kevin Coetzee, Thinus F.Kruge and Carl J.Lombard, "Predictive value of normal sperm morphology: a structured literature review", Human Reproduction Update 1998, Vol. 4, No. 1 pp. 73–82

[2]. WHO Laboratory manual for the examination of human semen and sperm-cervical mucus interaction. World Health Organization, 4th ed., Cambridge University Press, 1999.

[3]. Lukasz Witkowski , " Exam of the density of semen and analysis of sperm cell movement ",Journal of medical informatics & technology, Vol 3/2002 , ISSN 1642-6037 .

[4]. Lukasz Witkowski , " An automatic system for calculating basic semen parameter" , TASK QUARTERLY 8 No 2 (2004), 231–236 .

[5]. Vahid Reza Nafisi, Mohammad Hasan Moradi, and Mohammad Hosain Nasr-Esfahani , " Sperm Identification Using Elliptic Model and Tail Detection", World Academy of Science, Engineering and Technology ,6 2005.

[6]. Henry Carrillo, Jorge Villarreal, Miguel Sotaquirá ,colombia , " Spermatozoon Segmentation Towards an Objective Analysis of Human Sperm morphology", Proceedings of the 5th International Symposium on image and Signal Processing and Analysis (2007).

[7]. J. M.Pascual-Gaspar, H. Olmedo, A.I. Exposito , A. Exposito3 and J.Finat1 , Spain, "A simple and effective system for Computer Assisted Semen Analysis ",Advances in Medical signal and information processing , 2008, 4th International conference, Santa Margherita Ligure, Italy, ISBN:9780 86341 934 8, 14-16 July 2008 .

[8]. Hiromasa Oku and Masatoshi Naoko Ogawa Kogiku Shiba andManabu , japan , "How to Track Spermatozoa using High-Speed Visual Feedback ",30th Annual International IEEE EMBS Conference , Vancouver, British Columbia, Canada, August 20-24, 2008 .

[9]. V.S.Abbiramy et. Al.,"Spermatozoa Detection, Counting and Tracking in Video Streams to Detect Asthenozoospermia ", 2010 International Conference on Signal and Image Processing.

[10]. Chaudhari, N.M.; Pawar, B.V., "Light scattering study on semen analysis methods/techniques," Engineering (NUiCONE), 2013 Nirma University International Conference on Engineering, vol., no., pp.1,4, 28-30 Nov. 2013.

[11]. Qiang Wu, Fatima A. Merchant, Kenneth R. Castleman, Microscope Image Processing, Academic Press, United States of America, 2008.