

NEW HUMAN SEMEN ANALYSIS SYSTEM (CASA) USING MICROSCOPIC IMAGE PROCESSING TECHNIQUES

N.M. Chaudhari¹ and B.V. Pawar²

¹Department of Computer Engineering, KCES's College of Engineering & Information Technology, India
E-mail: ¹nandini113@yahoo.co.in

²School of Computer Science, North Maharashtra University, India
E-mail: ²bvpawar@hotmail.com

Abstract

Computer assisted semen analysis (CASA) helps the pathologist or fertility specialist to evaluate the human semen. Detail analysis of spermatozoa like morphology and motility is very important in the process of intrauterine insemination (IUI) or In-vitro fertilization (IVF) in infertile couple. The main objective for this new semen analysis is to provide a low cost solution to the pathologist and gynecologist for the routine raw semen analysis, finding the concentration of the semen with dynamic background removal and classify the spermatozoa type (grade) according to the motility and structural abnormality as per the WHO criteria. In this paper a new system, computer assisted semen analysis system is proposed in which hybrid approach is used to identify 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. For removal of background initially the dynamic background generation algorithm is proposed to create a background for background subtraction stage. The standard data set is created with 40× and 100× magnification from the different raw semen s. For testing the efficiency of proposed algorithm, same frames are applied to the existing algorithm. Another module of the system is focused on finding the motility and Type classification of individual spermatozoa.

Keywords:

Semen Analysis, CASA, Microscope Image Processing, Spermatozoa, Region Based Segmentation, Hybrid Algorithm, WHO

1. INTRODUCTION

Now a day's infertility is a common problem observed in human reproduction. So, human semen analysis is a routine pathological test for the analysis of semen. This test finds two parameters of semen. One is physical and another is chemical parameters like fructose and pH levels. Physical parameter examines the sperms contained in the semen. Many Factors influence the result of semen. The way the patient is advised to collect the specimen, the method and equipment used for testing in different laboratories decided the result of the semen. A routine analysis of semen includes the evaluation of sperm concentration, motility, morphology and vitality [1].

Most of the laboratory technician does this by analyzing the under microscope and physically counting the parameter values as shown in Fig.1.

Manually we can evaluate the physical parameter of the like sperm count, motility and morphology. The number of sperm present in the sample is called sperm count. The number of moving sperm decides the motility of the. Moving sperms are classified into Type A which are fast and progressive, Type B which are slow progressive, Type C sperms are Non-progressive and Type D sperms could not move at all i.e. they immobile. Here morphology is nothing but the shape of the spermatozoa [2].

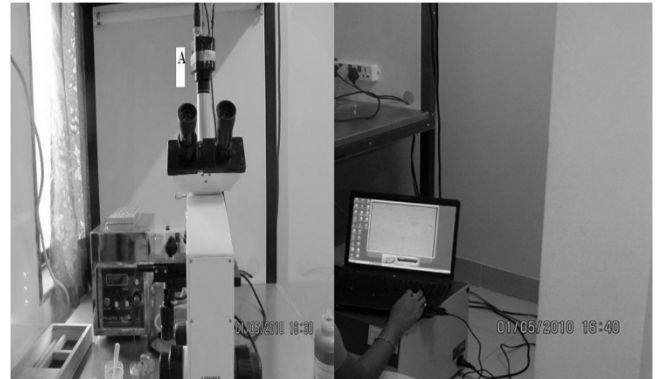


Fig.1. Semen analysis by manual method

2. LITERATURE SURVEY

Earlier most of the semen analyzers were used for this routine analysis. Due to fast moving sperm or medium velocity or a slow moving sperm it is very difficult to find the concentration of the semen and the other parameters of the semen manually.

Computer Assisted Semen Analysis (CASA) plays a vital role by providing ability to study and understand sperm function. CASA is very useful in assessment of sperm concentration (counts) and motility (movement) characteristic. Variations of the result of Manual method using Makelar Chamber with the computer automated sperm analysis ATS 20 and SQA system and the correlation for concentration and motility which ranges between 0.4 - 0.7. These techniques are widely used for calculating number of sperms in given known as sperm concentration (counts) and mobility (movement) characteristic [3]. Most of the CASA analyzers are based on different techniques on image analysis. The commercial introduction of CASA systems are facilitated the analysis of semen in many andrology laboratories. Various CASA system like Cell Soft (Cryo Resources Ltd, NY), Hamilton Thorne (Hamilton Thorne Research Inc, Danvers, MA) etc are used widely all over the world. (PATRICIA OLDSCLARKE, January/February 1990) [4].

Most of the CASA analyzers are based on different techniques on image analysis. Still there are certain limitations due to poor technical support. Due to technological advancement and different image processing techniques, new algorithms are proposed day by day to improve the accuracy of the results. We had done the extensive survey of semen analysis techniques and methods [5]. Microscope image processing is an emerging field in image processing where background removal is the first stage of the system as it is in any generic microscope image processing system [6]. The increasing use of Microscope image processing

(MIP) has seen increasingly used in the field of biomedical research and clinical medicine as well as in many other fields. Microscope image processing allows large scale statistical evaluation. It is used in various fields from medical to physical structures of the things. This helps the stakeholders to take decisions in better fashion [7].

The main objective is to track the object i.e. moving spermatozoa in the consecutive frames. Image enhancement is always the first stage observed of any image processing algorithm. In preprocessing stage, many authors had used various techniques like Image binarization, filtering, morphological operations and power law transformation for image enhancement. Witkowski had suggested operations of image enhancement such as logical, arithmetic, morphologic, neighborhood and point-to-point operations [8].

Image enhancement was used to reduce the noise and to remove the extra parts present in the image before starting the sperm determination step. The two step enhancement method produces background image and union of enhanced images. The background image needs to be enhanced to eliminate the non-uniform lighting distribution throughout the image. For this purpose closing-opening filter was used. Image I only includes motile sperms while Image B includes all other cells, immotile sperms, and agglutinated region [9]. Hiromasa Oku et al. used binarization for preprocessing. The captured image was binarized by detecting pixels whose intensity was lower than an intensity threshold set by an operator, because the heads are darker than the background with the phase-contrast method [10].

The image of each spermatozoon was processed with a power law transformation over each component of the RGB cube. The red and green components are the most important components for the transformation. The red component holds concentrated information associated with the darkest color, which domains the head, and with its enhancement a better differentiation of the nucleus from the others parts could be obtained. The transformation of the green component gives uniformity to the background which discriminates the spermatozoon's contour from the seminal plasma [11]. Laplacian filters were also used to enhance the image emphasizing on edges and to sharpen it. Then to remove the impulsive noise from an image Median filter was applied. It preserves the sharp edges. The images were transformed to Gray scale intensity images from RGB format. The low contrast gray images were then enhanced by remapping the data values which fills the entire intensity range in [0, 255] [12].

Since the tails of the spermatozoa was very thin and in some cases it is transparent also, it gets lost in the process of binarization or in edge detection segmentation. Also due to noise there was a problem of over segmentation. So our proposed system applies an image subtraction algorithm as a preprocessing stage.

Image subtraction technique is used in motion segmentation. Initially the background image is created by averaging the images over time. Pixel by pixel image subtraction of averaged image is done from current image to detect the moving region. It gives the status of the pixel at that particular condition. From which one can observe the location details by verifying the same object in the current image.

Background subtraction is a very popular approach for detecting moving objects from a still scene [13]. Combining shape and color feature background subtraction is robustly detecting

foreground objects in various illumination conditions [14]. Hati, K.K et al. proposes two different algorithm for intensity range based object detection. The first one model the background from initial few frames and the second algorithm extracts the objects based on local thresholding. The scheme has a strong potential for applications in real time surveillance [15]. Fan-Chieh Cheng et al. had proposed Illumination-Sensitive Background Modeling Approach to analyze the illumination change and detect moving objects. For the sudden illumination change, an illumination evaluation is used to determine two background candidates, including a light background image and a dark background image. Based on the background model and illumination evaluation, the binary mask of moving objects had generated by the thresholding function provides a promising detection outcome and low computational cost [16]. Background subtraction method is simple and easy to realize, and accurately extracts the characteristics of target data, but it is sensitive to the change of external environment, so it is applicable to the condition that the background is known [17]. There are several approaches to implement the BSA (Background Subtraction Algorithm), like temporal differencing, statistical approaches, optical flow which are implemented in the context of video surveillance system [18].

The accuracy of detection of spermatozoa was well analyzed with background subtraction technique. The background was created with certain threshold value of the first frame. This was the only preprocessing operation done on the raw semen to detect the spermatozoa [19].

3. ARCHITECTURAL FRAMEWORK OF PROPOSED SYSTEM

The proposed system consists of three modules, In Phase-I the concentration of the semen is evaluated. After extraction of the individual spermatozoa, the variations of each spermatozoon in a series of frames are calculated and the motility of the sperm is observed in Phase-II. According to the spermatoc motility they are classified as TYPE A, TYPE B, TYPE C, TYPE D. Phase-III is purely designed for the abnormality check of each spermatozoa which is captured on 100 \times . According to the WHO criteria, individual spermatozoa are classified as normal or abnormal sperm.

Major steps in overall proposed semen analysis system are,

1. Generate the frames from video of raw semen
2. Background removal by subtracting the created dynamically background
3. Use Region based segmentation to detect the sperm
4. Information extraction & Check of (ROI) contains the tail data with Scan line algorithm
5. Calculate number of sperms i.e. concentration (Phase-I)
6. Find the variation of individual spermatozoa in each frame and calculate motility. (Phase-II)
7. Classify it according to variations.
8. Information like head, tail, and acrosome is extracted from 100 \times frames for abnormality check.(Phase-III)
9. Classification

The Fig.2 shows the proposed system for human semen analysis. The first step of the system is to capture the video from microscope and extract frames. In preprocessing only background creation and subtraction is applied on the extracted frames.

Region based segmentation is used to generate the region of interest (ROI) with specific data. To label the each sperm first checked whether the ROI contains the tail data by applying scan line algorithm. Motility and abnormality check are done after the information extraction phase.

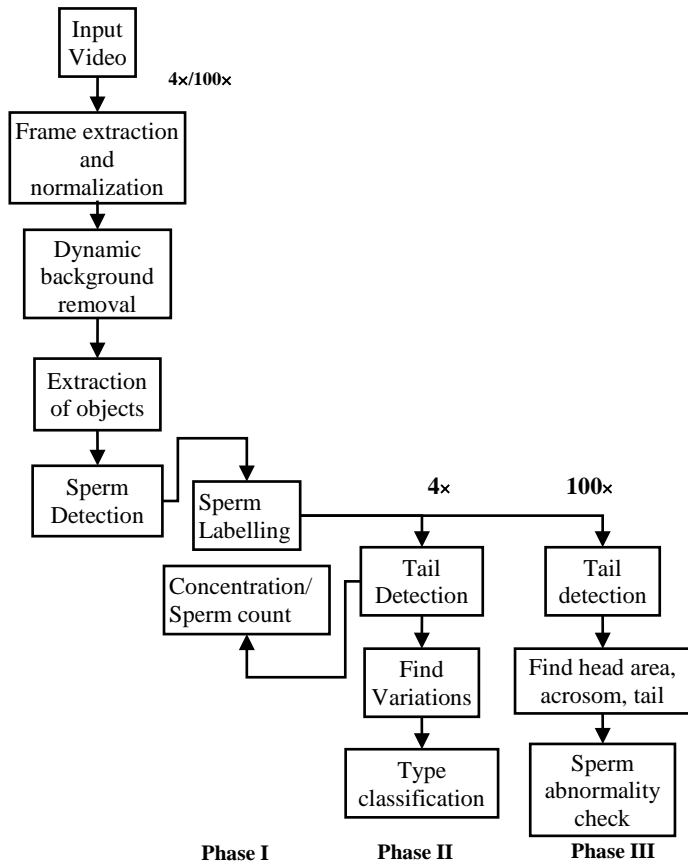


Fig.2. Framework of new semen analysis system

3.1 PREPROCESSING

Background subtraction is the first stage of the overall system, so we had introduced a run time dynamic background creation algorithm to create the background for background subtraction. As lab technician shall access the system, so to make the system user friendly we had created run time background frame. To create the dynamic background we had proposed the following algorithm 1.

3.1.1 Algorithm 1: Dynamic Background Creation:

1. Select the region over the image i.e. extracted frame.
2. Clip the input image as per the rectangle obtained from the diagonal points of the rectangle.
3. Calculate the average mean of the color values from the clipped image.
4. Create an empty bitmap.
5. Regenerate the empty bitmap with the obtained average mean value.
6. Store it as a background image.

3.2 PHASE I: OBJECT ANALYSIS

Object analysis is done through two stages, one is detection of the object and check if the object consists of tail in the detected

region so as to label it as a sperm. Region based segmentation with region to be connected as 4-connect is used to find the area of interest that may contain an object which has sperm data. Initially detect the bounding box value of each detector region whose area is greater than 4. Crop the image & Check whether it contain sperm data. By applying scan line algorithm on cropped region helps us to detect the tail.

To find concentration i.e. No. of sperms in each frame, After extraction of information and storing it in mat file we need to check for the redundancy if same sperm is appear in next frame or any new sperm occurs. To find the actual number of sperms exists in the semen. Load the .mat files and compare data from mat files from previous frames to files from current frame. 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. The labeled sperms in different frames are shown in Fig.3.

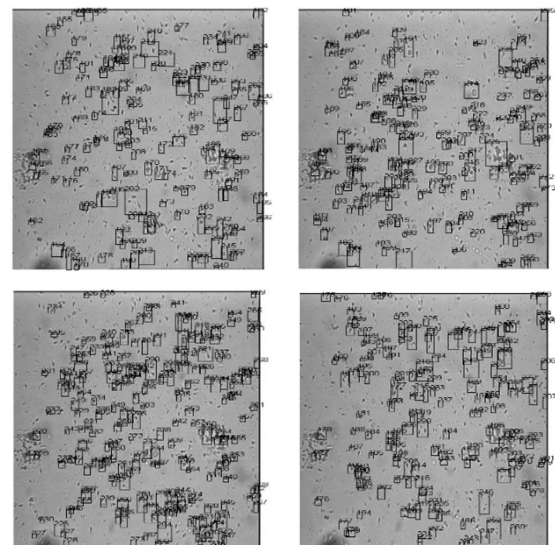


Fig.3. Object detections in different

3.3 PHASE II: MORTILITY AND TYPE CLASSIFICATION

Motility is nothing but the distance travelled by the spermatozoa per second. For this motility we need to find the motion of sperm and systematically count and classify the sperms into four categories namely, A, B, C and D according to the motion specified by WHO i.e. World health Organization manual.

3.3.1 WHO Suggested Criteria for Classification:

- a. Type A (Fast progressive) travelled in a straight line with speed at least $25\mu\text{m/s}$.
- b. Type B (Slow progressive) travelled in curved or crooked motion with speed between $5\text{--}25\mu\text{m/s}$.
- c. Type C (Non-progressive) move $<5\mu\text{m/s}$ & move their tails.
- d. Type D (Immobile) immotile and cannot move at all.

3.3.2 Motion Estimation:

Generally the sperms travels in straight lines and they may not take abrupt turn in 180. Even if they collide with each other they may not change their direction. As they swim during the test, they may not change their speed. Different CASA systems uses different

mathematical algorithms to calculate the movement of the sperms and the comparability of measurement is yet unknown [20].

Some standard parameters measured in motility analysis as per WHO manual for examination and processing of human sperm are:

VCL, Curvilinear Velocity (µm/s): It is time-averaged velocity of a sperm with index i along its actual curvilinear path, over the N frames.

$$VCL_i = \frac{\sum_{j=1}^N \sqrt{(x_{j+1} - x_j)^2 + (y_{j+1} - y_j)^2}}{(N - 1)\nabla t}$$

where, x_{j+1}, y_{j+1} is centroid of i in the current frame, x_j, y_j is centroid of i is the previous frame. ∇t is the time difference between consecutive frames

VSL, Straight-Line (rectilinear) Velocity (µm/s): It is velocity of the sperm between its location in first frame and its last N^{th} frame calculated directly along with straight line.

$$VSL_i = \frac{\sqrt{(x_N - x_1)^2 + (y_N - y_1)^2}}{(N - 1)\nabla t}$$

where, (x_1, y_1) is centroid of the sperm in first frame, (x_N, y_N) is the centroid of the same sperm in last N^{th} frame.

VAP, Average Path Velocity (µm/s): It is time-averaged velocity of a sperm head along its average path. This path is computed by smoothing the curvilinear trajectory according to algorithms in the CASA instrument; these algorithms vary between instruments, so values may not be comparable among systems.

$$VAP_i = \frac{\sum_{j=1}^N \sqrt{(\bar{x}_{j+1} - \bar{x}_j)^2 + (\bar{y}_{j+1} - \bar{y}_j)^2}}{(N - 1)\nabla t}$$

where,

$$\bar{x}_{j+1} = \frac{1}{5} \sum_{k=j-1}^{j+1} \bar{x}_k \text{ and } \bar{y}_j = \frac{1}{5} \sum_{k=j-1}^{j+1} \bar{y}_k, \text{ respectively}$$

These above variables are illustrated in Fig.4.

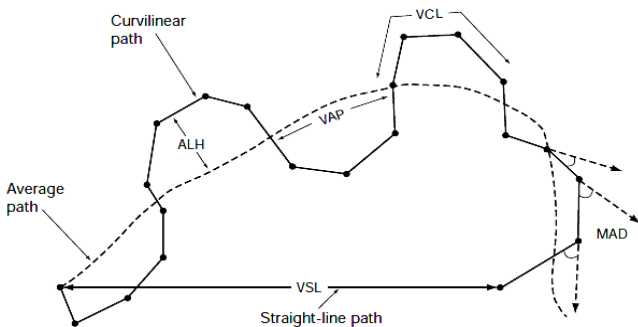


Fig.4. Terminology for variables defined in WHO manual

3.4 PHASE II: MOTILITY ANALYSIS

In Phase I we had already stored the information of the spermatozoa in .mat file. Now check if the same sperm is observed in the next frame. Save the information of the current instance in .mat file. Find out the variations of the sperms in

different consecutive frames and then using Euclidian distance formula; find the mean variation with respect to their centroid. Chances if addition of the new sperm considers the current frame as the initial frame for that sperm. To find the motility of individual spermatozoa and classify it according the estimated motion we had proposed the following algorithm 2.

3.4.1 Algorithm 2: To Find the Motility of the Individual Sperm:

1. Frame extraction
2. Background creation and subtraction
3. Region based segmentation
4. Label the sperm with the help of scan line algorithm which detect the tail of the sperms.
5. Store the extracted information of the each individual sperm as a structural array into the .mat file.
6. Repeat the steps 1 to 5 for each extracted frame and store the current instance of the sperm.
7. Match the labeled object from the first frame into the next frame.
8. Compare the instances and find the distance travelled and angle of the detected sperms.
9. Add the new occurred sperms into the structural array from the next frame.
10. Find the variation of the distances with the help of distance formula and find the mean distance.
11. Find the curvilinear path and straight line path with the help of calculate path algorithm.
12. Classify the sperm into the four categories as Type A, Type B, Type C and Type D. as per the classification algorithm.
13. Quantification, Determine the distribution of the spermatozoa in each class i.e. find the number of sperms in Type A, Type B, Type C and Type D.

Spermatic motility is dependent on the movement of the individual spermatozoa and the typical distance travelled by the sperms with their trajectories. For tracking of the path the centroid of the sperm is observed from initial frame to final frames i.e. from 1,2,3,4,...,N frames. The total distance travelled in individual frame is nothing but the curvilinear distance travelled by the individual sperm. So the curvilinear velocity is calculated with the frame difference of 25FPS. The new function Calculate Path determines the curvilinear path tracking of centroid (X,Y) of the sperm. The curvilinear tracking of some spermatozoa in 50 frames is shown graphically in Fig.5.

3.5 PHASE-III: MORPHOLOGY ANALYSIS

Sperm morphology measurement consists of Length and Width of spermatozoon's head, Area and Acrosome percentage occupied by the head, Shape parameters [11].

In our proposed system Phase-III is designed to find the morphology of individual spermatozoa. Obviously to find the structural details of the sperm we need to acquire the frame on 100× with stable background. Here also we are considering the raw semen. The first stage is obviously the detection of the spermatozoa contained in the semen as in phase-I.

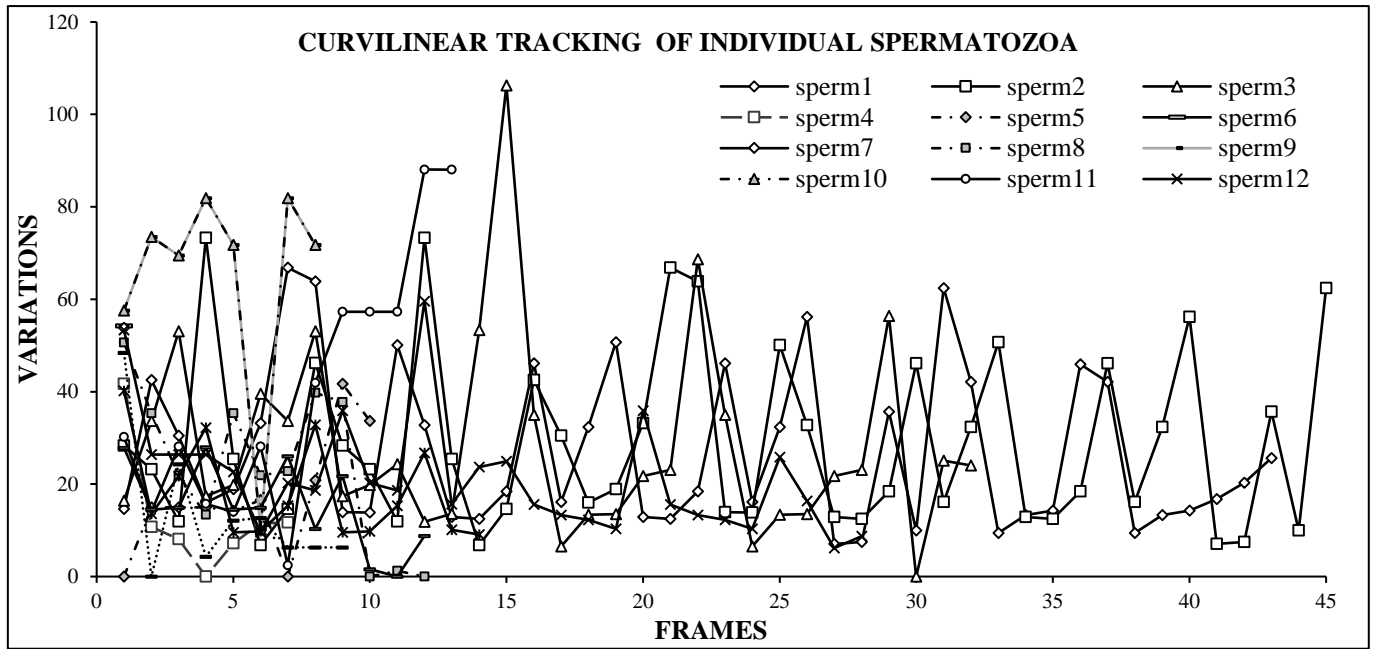


Fig.5. Tracking of spermatozoa in 50 consecutive frames

3.5.1 Morphology Classification to Check Structural Abnormality:

With the help of rule based classification and sequential forward feature selection, the morphological classification to check the structural abnormality is proposed. The first feature to select the spermatozoa for classification is its area. So first we need to verify Sperm {index}. Area > 200. Other parameters are verified as per the WHO criteria of classification for normal and abnormal spermatozoa.

3.5.2 WHO Rules for Morphology Check:

According to the WHO manual the followings classifications conditions are considered to classify the sperm as normal or abnormal

1. Acrosome area is 40% to 70% of head area.
2. Head width should be 2.5 To 3.5 μ m , head height should be 4 to 5 μ m
3. Head length to width ratio should be 1.5
4. Tail length should be 45 μ m

So we are classifying the detected sperm as a normal or abnormal according to the shape measurements. Since we had manufactured the conditions considering the measurements are in the spatial form. And the actual measurement of the spermatozoa is in μ m. so considering the magnification factor and conversion factor for pixel to mm (millimeter) the classification conditions is used. Standard unit conversion of 1 pixel = 0.264583333 mm is considered while producing the rules. For the classification purpose we need to form the rules R1, R2, R3 etc.

Rule 1: If (Headheight/Headwidth is 1.5:1)

Then Class = "Normal"

Rule 2: If (Head/Acrosome is greater than 4)

Then Class = "Normal"

Rule 3: If (Tail is greater than 45)

Then Class = "Normal"

Failing to one of the above rule the Class = "Abnormal"

3.5.3 Algorithm 3: Proposed algorithm for morphological classification to check the structural abnormality:

[Input]: frame acquired at 100 for the same semen

[Output]: Classification of spermatozoa and the parameters detection

1. Apply region based segmentation and scan line algorithm to detect the individual spermatozoa and extract the information same is in Phase-I.
2. Load the extracted sperm file and convert into the appropriate data structure.
3. Calculate the global sperm index by comparing the existing sperms into the next frame.
4. Find the motility of the individual spermatozoa and classify it in the Type of A, B, C and D same as in Phase-II.
5. Compute the parameter like head height and head width of the individual spermatozoa and evaluate the area covered by head and acrosome.
6. Cropped the specific region of interest to find the head and acrosome percentage by calculating the head and acrosome count.
7. Select the appropriate regions of which the first feature area is greater than 200, whose output for raw is shown in Fig.6.
8. Apply the region based segmentation based on thresholding on the cropped region and with the help of scan line algorithm compute the tail length.
9. Classification

Input Head, Acrosome, Head width, Head height and Tail for this stage and classify the spermatozoa as a normal or abnormal category according to the Rule based classification.

The prominent features of the spermatozoa which are computed in above steps 5, 6, 7 and 8 are store in the text file in which all conditions for classifications are used to categorize the sperm as a normal or abnormal sperm.

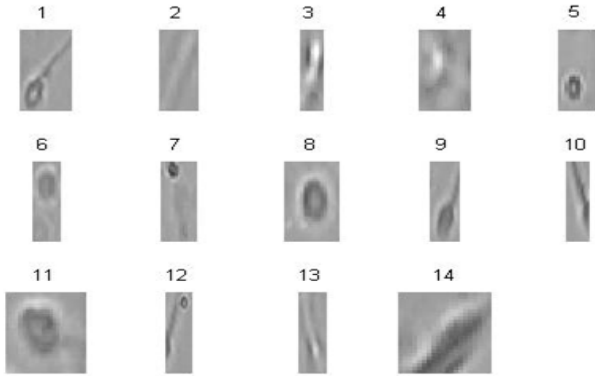


Fig.6. Selected regions of 2 for further stage classification

4. PROPOSED APPROXIMATE GENERALIZED MATHEMATICAL MODEL IN THE PROCESS OF SEMEN ANALYSIS

This section represents the details about the formulation of field data based mathematical model to analyze the impact of various parameters on the quality of sperms in semen. In present scenario where majority of total operations are still executed using generalized methods which needs to be focused and develop a mathematical relation which simulate the real input and output data. The findings indicate that the topic under study is of great importance as no such approach of field data based mathematical modeling is adopted for the formulation of mathematical model.

Data sets contain information and the behavior of the process variables, often much more than can be learned from just looking at plots of those observed data. Mathematical models based on observed input and output data from real life situation (Semen analysis Process) help us to gain new information and understanding from these data. Thus, it is not possible to plan such activities on the lines of design of experimentation When one is studying any completely physical phenomenon but the phenomenon is very complex to the extent that it is not possible to formulate a logic based model correlating causes and effects of such a phenomenon, then one is required to go in for the field data based models. Hence the approach of formulating a field data based model is suggested to analysis of semen.

The general equation form of dependent and independent variables are as follows

$$Y = f\{V, F, B, R, S, M, C, Z\}$$

where,

Symbol	Independent Variables
V	Set of input videos
F	Set of input frames
B	Set of output of background subtraction
R	Set of region selection
S	Set of detected sperms
M	Set of motility information
C	Set of morphological information

Z Set of other parameters

Following is the dependent variable

Y Semen Analysis (Quality of the Semen)

Considering the above dependent and independent variable the proposed mathematical model of new semen analysis system is having following form

$$\Pi_9 = f\{\Pi_1, \Pi_2, \Pi_3, \Pi_4, \Pi_5, \Pi_6, \Pi_7, \Pi_8\} \quad (1)$$

where, $V = \Pi_1, F = \Pi_2, B = \Pi_3, R = \Pi_4, S = \Pi_5, M = \Pi_6, C = \Pi_7, Z = \Pi_8$ and $Y = \Pi_9$

The mathematical relation between inputs and outputs could be of any form may be polynomial, exponential or log linear. The Buckingham Pi theorem found suitable for developing the model. As it states that if the inputs and outputs represented in dimensionless pie terms by dimensional analysis then they can be represented as,

$$\Pi_9 = k_1 \cdot (\Pi_1)^{a_1} \cdot (\Pi_2)^{b_1} \cdot (\Pi_3)^{c_1} \cdot (\Pi_4)^{d_1} \cdot (\Pi_5)^{e_1} \cdot (\Pi_6)^{f_1} \cdot (\Pi_7)^{g_1} \cdot (\Pi_8)^{h_1} \quad (2)$$

where, Π_1 to Π_8 are pi-terms related with independent variables and Π_9 is the dependent variable of new semen analysis system.

The model has various indices as $k_1, a_1, b_1, c_1, d_1, e_1, f_1, g_1$ and h_1 whose values are unknown. This model represents the field data base mathematical model for new semen analysis system. There are 8 unknown terms and one curve fitting constant k_1 in the Eq.(2). To get the values of these unknowns we need minimum a set of all unknown dimensionless pi terms. To simplify the model Consider the following relation,

$$Z = A + bX + CY \quad (3)$$

The Eq.(3) represents equation of a curve fitting technique. Eq.(2) and Eq.(3) can be brought in the form of Eq.(4) as follows by taking log on both sides,

$$\begin{aligned} \log \Pi_9 &= \log k_1 + a_1 \log(\Pi_1) + b_1 \log(\Pi_2) + c_1 \log(\Pi_3) \\ &+ d_1 \log(\Pi_4) + e_1 \log(\Pi_5) + f_1 \log(\Pi_6) + g_1 \log(\Pi_7) \\ &+ h_1 \log(\Pi_8) \end{aligned} \quad (4)$$

Let,

$$\begin{aligned} \log \Pi_9 &= Z, \log k_1 = K_1, \log(\Pi_1) = A, \log(\Pi_2) = B, \log(\Pi_4) = D, \\ \log(\Pi_5) &= E, \log(\Pi_6) = F, \log(\Pi_7) = G, \log(\Pi_8) = H \end{aligned}$$

Then, the equation can be written as,

$$\begin{aligned} Z &= K_1 + a_1 \cdot A + b_1 \cdot B + c_1 \cdot C + d_1 \cdot D + e_1 \cdot E \\ &+ f_1 \cdot F + g_1 \cdot G + h_1 \cdot H \end{aligned} \quad (5)$$

The Eq.(5) is a regression equation of Z_1 on A, B, C, D, E, F, G in a dimensional co-ordinate system and again Eq.(5) is written as,

$$\begin{aligned} \Sigma Z &= n \cdot K_1 + a_1 \cdot \Sigma A + b_1 \cdot \Sigma B + c_1 \cdot \Sigma C + e_1 \cdot \Sigma E \\ &+ f_1 \cdot \Sigma F + g_1 \cdot \Sigma G + h_1 \cdot \Sigma H \end{aligned} \quad (6)$$

$$\begin{aligned} \Sigma ZA &= K_1 \cdot \Sigma A + a_1 \cdot \Sigma A \cdot A + b_1 \cdot \Sigma B \cdot A + c_1 \cdot \Sigma C \cdot A \\ &+ e_1 \cdot \Sigma E \cdot A + f_1 \cdot \Sigma F \cdot A + g_1 \cdot \Sigma G \cdot A + h_1 \cdot \Sigma H \cdot A \end{aligned} \quad (7)$$

$$\begin{aligned} \Sigma ZB &= K_1 \cdot \Sigma B + a_1 \cdot \Sigma A \cdot B + b_1 \cdot \Sigma B \cdot B + c_1 \cdot \Sigma C \cdot B \\ &+ e_1 \cdot \Sigma E \cdot B + f_1 \cdot \Sigma F \cdot B + g_1 \cdot \Sigma G \cdot B + h_1 \cdot \Sigma H \cdot B \end{aligned} \quad (8)$$

$$\begin{aligned} \Sigma ZC &= K_1 \cdot \Sigma C + a_1 \cdot \Sigma A \cdot C + b_1 \cdot \Sigma B \cdot C + c_1 \cdot \Sigma C \cdot C \\ &+ e_1 \cdot \Sigma E \cdot C + f_1 \cdot \Sigma F \cdot C + g_1 \cdot \Sigma G \cdot C + h_1 \cdot \Sigma H \cdot C \end{aligned} \quad (9)$$

$$\Sigma ZD = K_1.\Sigma D + a_1.\Sigma A.D + b_1.\Sigma B.D + c_1.\Sigma C.D + e_1.\Sigma E.D + f_1.\Sigma F.D + g_1.\Sigma G.D + h_1.\Sigma H.D \tag{10}$$

$$\Sigma ZE = K_1.\Sigma E + a_1.\Sigma A.E + b_1.\Sigma B.E + c_1.\Sigma C.E + e_1.\Sigma E.E + f_1.\Sigma F.E + g_1.\Sigma G.E + h_1.\Sigma H.E \tag{11}$$

$$\Sigma ZF = K_1.\Sigma F + a_1.\Sigma A.F + b_1.\Sigma B.F + c_1.\Sigma C.F + e_1.\Sigma E.F + f_1.\Sigma F.F + g_1.\Sigma G.F + h_1.\Sigma H.F \tag{12}$$

$$\Sigma ZG = K_1.\Sigma G + a_1.\Sigma A.G + b_1.\Sigma B.G + c_1.\Sigma C.G + e_1.\Sigma E.G + f_1.\Sigma F.G + g_1.\Sigma G.G + h_1.\Sigma H.G \tag{13}$$

$$\Sigma ZH = K_1.\Sigma H + a_1.\Sigma A.H + b_1.\Sigma B.H + c_1.\Sigma C.H + e_1.\Sigma E.H + f_1.\Sigma F.H + g_1.\Sigma G.H + h_1.\Sigma H.H \tag{14}$$

In the above set of equations the values of the multipliers K_1, A, B, C, D, E, F, G and H are substituted to compute the values of $a_1, b_1, c_1, d_1, e_1, f_1, g_1$ and h_1 . After substituting these values in the equations one will get a set of nine equations, which are mutinously to get the values of $k_1, a_1, b_1, c_1, d_1, e_1, f_1, g_1$ and h_1 . The above equations can be verified in the matrix form and further values of $k_1, a_1, b_1, c_1, d_1, e_1, f_1, g_1$ and h_1 can be obtained by using matrix analysis.

$$[P_1] = [W_1] [X_1] \tag{15}$$

$$X_1 = \text{inv}(W_1) \times P_1 \tag{16}$$

where, $W_1 = 9 \times 9$ matrix multipliers of $k_1, a_1, b_1, c_1, d_1, e_1, f_1, g_1$ and h_1 and $P_1 = 9 \times 1$ matrix of the terms on LHS and $X_1 = 9 \times 1$ matrix of values of $k_1, a_1, b_1, c_1, d_1, e_1, f_1, g_1$ and h_1 .

5. EXPERIMENTAL RESULTS AND ANALYSIS

For experimental purpose results are verified on 25 different semen s. Initially the frame wise labeling with proposed hybrid approach and existing method was implemented. We are able to detect the sperm without losing tails of sperms with 93.18% accuracy. The concentration of different raw semen s are evaluated by the proposed system and compared with traditional semen analysis manual method which is shown in Table.1.

laboratories. A graph of concentration evaluated by expert from manual method and the proposed system is shown in Fig.3.

Table.1. Concentration and the evaluation time of proposed system

Semen	Concentration		Evaluation Time (s)
	Proposed system	Technician	
1	106	160	82.53
2	85	90	100.88
3	67	80	64.54
4	72	20	68.42
5	105	220	85.77
6	105	170	133.67
7	85	180	87.25
8	56	80	86.65
9	59	60	84.44
10	80	90	99.9
11	89	70	101.98
12	81	170	103.61
13	165	160	165.73
14	84	110	98.81
15	55	70	65.8
16	69	80	81.47
17	91	120	84.69
18	101	130	203.53
19	67	120	89.19
20	93	100	93.38
21	110	140	138.34
22	95	110	131.52
23	77	70	72.33
24	105	110	139.56
25	87	100	146

While analyzing standard deviation for spermatozoa detection in the Table.1 and graph shown in Fig.3, it is observed that , the average standard deviation for detection of total sperm i.e. concentration is 19.79 excluding one or two outlier. The minimum standard deviation is 0.707 and that indicates that the system is able to detect the concentration of which shows the accuracy of the proposed system.

The mortality of the individual spermatozoa is evaluated according to the mean distance travelled by the spermatozoa from the initial frames to final frame. As per WHO standard criteria, the velocity of the sperms greater than $25\mu\text{m/s}$ is categorized as TYPE A. Considering constant multiplying factor for pixel conversion and the zoom factor of the microscope, motility is evaluated with respect to the mean distance travelled. The other distances line straight-line path and curvilinear path is also evaluated for each individual spermatozoa and classify them. The parameters evaluated by Algorithm 2 are stored in the text file is as shown in Table.2.

Mortality analysis i.e. classification and quantification of 25 different raw semen s are shown in Table.3.

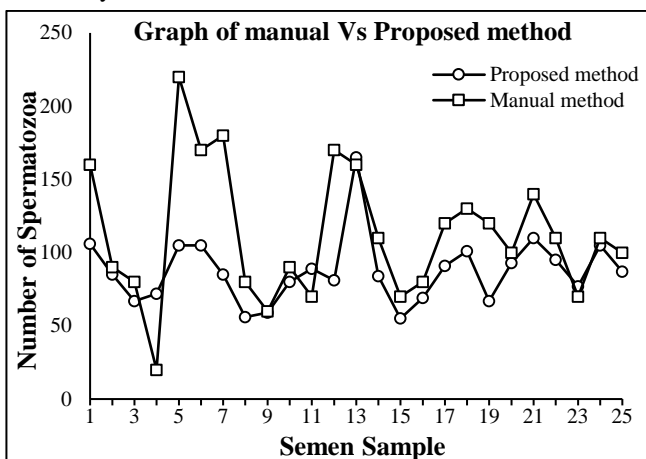


Fig.3. Concentration observed by experts and proposed system

According to the survey and experts opinion 30% of variations in semen parameters exist while evaluating it in different

Table.2. Motility analysis of a semen with mean distance, velocity of sperms

w	Centroid (X,Y)	Theta	Average Path	Mean Distances	Straight Line Path	Curvilinear path	Type
1	(92.0417,51)	23.3111	3959	33.8376	395.9	4231.1818	a
2	(121.4348,169.2698)	36.7956	2916	35.1325	291.6	3763.9863	a
3	(125.4,138.4211)	22.4939	1862	30.0323	186.2	1886.6296	a
4	(94.1064,94.1064)	74.4987	751	34.1364	75.1	488.2828	a
5	(190.7692,190.7692)	89.0693	288	26.1818	28.8	178.7512	a
6	(155.5484,173.9216)	45.1811	1068	26.7	106.8	1360.8349	a
7	(171.875,206.5263)	6.1611	367	26.2143	36.7	471.8254	a
8	(171.875,497.7917)	NaN	0	0	0	0	d
9	(221.5862,221.5862)	NaN	98	8.9091	9.8	62.2404	b
10	(176.6939,515.4082)	NaN	0	0	0	0	d
11	(186.75,186.75)	56.9974	1507	37.675	150.7	1545.4643	a
12	(223.3158,223.3158)	82.24	102	9.2727	10.2	74.2502	b
13	(237.0625,237.0625)	79.6952	406	36.9091	40.6	98.4712	a
14	(249.069,251.8571)	88.7832	494	44.9091	49.4	288.7978	a
15	(220.284,220.284)	29.9124	977	29.6061	97.7	1041.5253	a
16	(257.125,263.1077)	48.7472	352	32	35.2	199.0577	a
17	(261.7391,261.7391)	72.497	911	41.4091	91.1	702.9112	a
18	(282.2,281.5714)	78.6781	998	45.3636	99.8	639.1721	a
19	(269.0222,291.5814)	66.5735	1961	40.0204	196.1	1781.1866	a
20	(276.4,276.4)	NaN	96	8.7273	9.6	50.4386	b
21	(303.916,301.7292)	77.7902	562	43.2308	56.2	280.6428	a
22	(296.9,301.3333)	57.9275	463	30.8667	46.3	339.0347	a
23	(39.2222,50.1698)	73.8034	2173	31.0429	217.3	2706.7192	a
.							
.							
.							
.							
.							
206	(346.35,336.95)	NaN	0	0	0	0	d
207	(136.7083,303.625)	NaN	0	0	0	0	d

According to the graph of motility analysis for Type A, Type B and Type C spermatozoa, small variations are observed in comparative analysis. Large variations are evaluated in finding the motility of Type D sperms. Since Type D sperms are dead sperms and the debris in the semen are added while evaluating the Type D sperms which is shown in Fig.4.

The prominent features of the spermatozoa are calculated with the help of nested region based segmentation and scan line algorithm from the 100x frames of the semen. Rule based classification method classifies the individual spermatozoa as a normal or abnormal. We are also successful in finding the percentage of head and acrosome and tail of the spermatozoa which is stored in file and is shown in Table.4.

The features selected for the classifications are area, head, head width, head height and tail of the spermatozoa. The proposed algorithm's coverage to classify the sperms is 82.35% with accuracy of 78.57%.

The experimental result shows with the selected features and rule based classifier helps to yield better output result. These

results are very useful for decision making to the gynecologist or embryologist in the treatment of IUI and IVF. The result shows that the proposed algorithm for morphology evaluation is efficient and sensitive and having the ability to detect the specific abnormal or debris 83.33%.

6. CONCLUSION

In this paper a new computer automated semen analysis system is proposed for the analysis of raw semen. Algorithms for dynamic background creation, spermatozoa detection, Mortality analysis and morphology analysis using region based segmentation, scan line algorithm were implemented. Existing algorithm is also implemented for testing purpose on the database of semen created by us.

Our hybrid approach helps us in finding concentration, motility and morphology of raw semen. The tail of spermatozoa which was lost in existing preprocessing operations was also evaluated.

Table.3. Motility of individual semen by proposed method

Semen Sample	Type A		Type B		Type C		Type D	
	Proposed	Technician	Proposed	Technician	Proposed	Technician	Proposed	Technician
1	52	60	37	20	9	10	35	10
2	45	40	28	30	12	10	94	20
3	35	30	22	10	10	20	50	40
4	35	10	26	30	11	20	43	50
5	52	70	41	20	12	5	37	5
6	54	60	42	20	9	10	102	10
7	40	50	32	30	13	10	64	10
8	28	50	24	20	4	10	97	20
9	29	20	23	20	7	20	86	40
10	38	50	33	20	9	10	89	20
11	47	30	34	20	8	20	79	30
12	38	40	36	30	7	20	94	10
13	98	40	48	20	19	10	136	30
14	39	40	35	30	10	20	97	10
15	23	40	23	20	9	10	69	30
16	31	40	33	20	5	20	66	20
17	43	50	42	20	6	20	61	10
18	55	40	36	20	10	30	125	10
19	35	40	25	20	7	20	90	20
20	54	50	35	20	4	10	78	20
21	62	60	39	20	9	10	137	10
22	47	30	40	20	8	10	139	40
23	42	40	25	40	10	10	46	20
24	54	50	42	30	9	10	102	10
25	45	40	34	30	8	20	114	10

Table.4. Output of Classification stage

Label	Centroid	Area	Head	Head Height	Head width	Acrosome	Tail	Grade classification	Type classification
1	(75.2521,75.2521)	238	97.723	33.5	15	2.277	73.75	1	normal
2	(198.1505,210.4327)	372	93.9909	20.5	10	6.0091	47	1	normal
3	(197.1073,448.5107)	233	93.9542	33.5	8.5	6.0458	66.25	4	normal
5	(210.4327,206.151)	245	97.0588	16.5	11.5	2.9412	46.75	4	normal
6	(207.5121,451.4831)	207	94.6023	32.5	7.5	5.3977	55.75	4	normal
7	(222.8618,252.6829)	246	99.374	35	8.5	0.62598	61	4	normal
8	(160.4963,165.5309)	405	50	0	0.5	50	51.25	4	abnormal

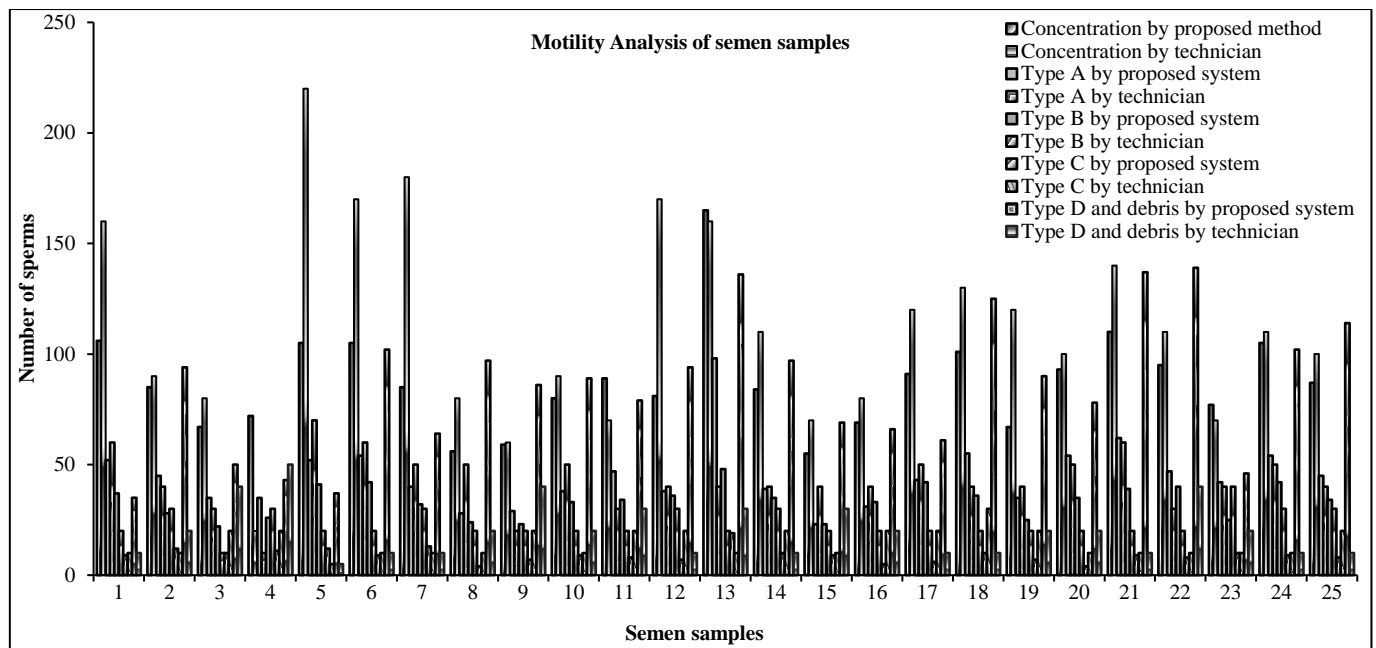


Fig.4. Evaluation of Type A, Type B, Type C and Type D sperms by proposed system and experts

ACKNOWLEDGEMENT

The authors are thankful to the University Grants Commission, New Delhi for supporting this research under the Special Assistance Programme (SAP) at the level of DRS-I (No: F.3-52/2011(SAP-II) and North Maharashtra University, Jalgaon for supporting this research to create a dataset for human semen under the Vice Chancellors Research motivation scheme (VCRMS).

REFERENCES

[1] World Health Organization, "WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction", 4th Edition, Cambridge University Press, 1999.

[2] Lars Bjorndahl et al., "A Practical Guide to Basic Laboratory Andrology", 1st Edition, Cambridge University Press, 2010.

[3] Jane E. Johnson, William R. Boone and Dawn W. Blackhurst, "Manual versus Computer-Automated Semen Analyses. Part II Determination or the Working Range of a Computer-Automated Semen Analyzer", *Fertility and Sterility*, Vol. 65, No. 1, pp. 156-159, 1996.

[4] Patricia Olds-Clarke, Harry M. Baer and Walter L. Gerber, "Human Sperm Motion Analysis by Automatic (Hamilton-Thorn Motility Analyzer) and Manual (Image-80) Digitization Systems", *Journal of Andrology*, Vol. 11, No. 1, pp. 55-58, 1990.

[5] Nandini M. Chaudhari and Bhausahab V. Pawar, "Light Scattering Study on Semen Analysis Methods/Techniques", *Proceedings of Nirma University International Conference on Engineering*, pp.1-4, 2013.

[6] Qiang Wu, Fatima A. Merchant and Kenneth R. Castleman, "Microscope Image Processing", 1st Edition, Academic Press, 2008.

[7] N.M. Chaudhari and B.V. Pawar, "Microscope Image Processing: An Overview", *International Journal of Computer Applications*, Vol. 124, No. 12, pp.23-28, 2015.

[8] L. Witkowski, "Examination of the Density of Semen and Analysis of Sperm Cell Movement", *Journal of Medical Informatics and Technology*, Vol. 3, pp. 1-8, 2002.

[9] 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*, Vol. 6, pp. 205-208, 2005.

[10] H. Oku, M. Ishikawa, N. Ogawa, K. Shiba and M. Yoshida, "How to Track Spermatozoa using High-Speed Visual Feedback", *Proceedings of 30th Annual International IEEE Conference on Engineering Medicine and Biology Society*, pp. 20-24, 2008.

[11] Henry Carrillo, Jorge Villarreal, Miguel Sotaquira, Alvaro Goelkel and Ricardo Gutierrez, "Spermatozoon Segmentation Towards an Objective Analysis of Human Sperm Morphology", *Proceedings of 5th International Symposium on Image and Signal Processing and Analysis*, pp. 522-527, 2007.

[12] V.S.Abbiramy, V.Shanthi and Charanya Allidurai, "Spermatozoa Detection, Counting and Tracking in Video Streams to Detect Asthenozoospermia", *Proceedings of International Conference on signal and Image processing*, pp. 265-270, 2010.

[13] Wonjun Kim and Changick Kim, "Background Subtraction for Dynamic Texture Scenes using Fuzzy Color Histograms", *IEEE Signal Processing Letters*, Vol. 19, No. 3, pp. 127-130, 2012.

- [14] L.Y. Liu, N. Sang, and R. Huang, "Background Subtraction using Shape and Colour Information", *Electronics Letters*, Vol. 46, No. 1, pp. 41-43, 2010.
- [15] Kalyan Kumar Hati, Pankaj Kumar Sa and Banshidhar Majhi "Intensity Range based Background Subtraction for Effective Object Detection", *IEEE Signal Processing Letters*, Vol. 20, No. 8, pp. 759-762, 2013.
- [16] Fan-Chieh Cheng; Shih-Chia Huang and Shanq-Jang Ruan, "Illumination-Sensitive Background Modeling Approach for Accurate Moving Object Detection", *IEEE Transactions on Broadcasting*, Vol. 57, No. 4, pp. 794-801, 2011.
- [17] U Manjesh kumar and T.A. Nalina, "A Fast and Robust Video Object Segmentation using Global Motion Estimation", *International Journal For Technological Research In Engineering*, Vol. 2, No. 9, pp. 2062-2064, 2015.
- [18] S.H Shaikh, "Moving Object Detection Approaches, Challenges and Object Tracking", Springer Briefs in Computer Science, pp. 1-14, 2014.
- [19] Nandini M. Chaudhari and Bhausahab V. Pawar, "New Hybrid Approach for Identification of Spermatozoa in Human Semen Sample using Microscope Image Processing Techniques", *Advances in Image and Video Processing*, Vol. 2, No. 6, pp. 15-24, 2014.
- [20] World Health Organization, "WHO Laboratory Manual for the Examination and Processing of Human Semen", 5th Edition, Cambridge University Press, 2010.