**REPUBLIC OF TURKEY**
**YILDIZ TECHNICAL UNIVERSITY**
**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**

# A SYSTEM IMPLEMENTATION FOR ANALYZING AND TRACKING MOTILE OBJECTS IN BIOMEDICAL IMAGES

**HAMZA OSMAN İLHAN**

**PhD. THESIS**
**DEPARTMENT OF COMPUTER ENGINEERING**
**PROGRAM OF COMPUTER ENGINEERING**

**ADVISER**
**PROF. DR. NİZAMETTİN AYDIN**

**İSTANBUL, 2017**

**REPUBLIC OF TURKEY**

**YILDIZ TECHNICAL UNIVERSITY**

**GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**

**A SYSTEM IMPLEMENTATION FOR ANALYZING AND TRACKING MOTILE OBJECTS IN BIOMEDICAL IMAGES**

A thesis submitted by Hamza Osman İLHAN in partial fulfillment of the requirements for the degree of **DOCTOR OF PHILOSOPHY** is approved by the committee on 16.01.2018 in Department of Computer Engineering, Computer Engineering Program.

**Thesis Adviser**

Prof. Dr. Nizamettin AYDIN

Yıldız Technical University

**Approved By the Examining Committee**

Prof. Dr. Nizamettin AYDIN

Yıldız Technical University                    _____

Assoc. Prof. Dr. Songül ALBAYRAK, Member

Yıldız Technical University                    _____

Assist. Prof. Dr. Tolga ENSARİ, Member

İstanbul University                    _____

Assist. Prof. Dr. Görkem SERBES, Member

Yıldız Technical University                    _____

Prof. Dr. Osman Nuri UÇAN, Member

Altınbaş University                    _____

# ACKNOWLEDGEMENTS

# TABLE OF CONTENTS

# LIST OF SYMBOLS

$\angle^k$      The angels of the slope $k$

$\mu_c$      Mean of the Cluster $c$

$\sigma_c$      Standard Deviation of the Cluster $c$

# LIST OF ABBREVIATIONS

| | |
|---|---|
| BRISK | Binary Robust Invariant Scalable Key-points |
| CAS | Computer Aided System |
| CASA | Computer Aided Sperm Analysis |
| CB | Computer Based |
| CNN | Convolutional Neural Network |
| CSCTAS | Computerized Sperm Counting and Analyzing Software |
| DWT | Discrete Wavelet Transform |
| FAST | Features from Accelerated Segment Test |
| FCM | Fuzzy C-Means |
| FN | False Negative |
| FP | False Positive |
| FPS | Frame Per Second |
| GT | Ground Truth |
| HT | Hough transform |
| JPDAF | Joint Probability Data Acquisition Filter |
| KNN | K-Nearest Neighbors |
| LoG | Laplacian of Gaussian |
| LSD | Line Segment Detector |
| PCA | Principle Component Analysis |
| PDAF | Probability Data Acquisition Filter |
| PDF | Probability Density Function |
| RKF | Recursive Kalman Filter |
| RMSE | Root Mean Square Error |
| ROI | Region of Interest |
| SIFT | Scale Invariant Feature Transform |
| SURF | Speed-Up Robust Features |
| SVM | Support Vector Machines |
| TFR | Total Fertility Rate |
| TN | True Negative |
| TP | True Positive |
| VA | Visual Assessment |
| WHO | World Health Organization |

# LIST OF FIGURES

# LIST OF TABLES

# ABSTRACT

## A SYSTEM IMPLEMENTATION FOR ANALYZING AND TRACKING MOTILE OBJECTS IN BIOMEDICAL IMAGES

Hamza Osman İLHAN

Department of Computer Engineering

PhD. Thesis

Adviser: Prof. Dr. Nizamettin AYDIN

Spermiogram is the first step of the infertility diagnosis. Computer Aided Sperm Analysis (CASA) and Visual Assessment (VA) are two evaluation techniques employed in spermiogram analyses. The VA is carried out by manually observing the sperm on counting chambers. Hence, analysis and diagnosis strongly depend on the skills and experiences of the observers. On the other hand, the CASA is a more advanced technology due to the improved computerized techniques and minimization of human intervention. However, it is more expensive than VA since it is an integrated computer based system and requires exhaustive parameter setting process.

In this thesis, we aim to develop a combinational approach using the smartphone and computer for the sperm concentration and motility analysis. Smartphone was utilized to obtain images similar to the VA technique. The acquired samples were analyzed by using computerized methods to eliminate the observer variability. In this thesis, a software named as Computerized Sperm Counting and Trajectory Analyzing Software (CSCTAS) for automatically counting and tracking the sperm over one of the commonly using counting chamber, is proposed. Proposed software consists of seven modules executed sequentials: (1) Data Acquisition and Organization, (2) Automatic Grid and Region of Interest (ROI) detection and extraction, (3) Video Stabilization, (4) Motile/Immotile Spermatozoon Detection, (5) Spermatozoa Counting, (6) Motile spermatozoa tracking, and (7) Trajectory Classification. Each module consists of various combination of image processing techniques. Firstly, data acquisition and organization were performed using a novel approach to provide inexpensive design contrary to traditional CASA systems. Secondly, Region of Interest (ROI) extraction was realized by a combinational approach of line detection and segmentation methods. Then, feature matching based video stabilization was introduced to eliminate the vibrations occurred during the data acquisition step. In this respect, different descriptors

were tested. The fourthly, Background and foreground extraction techniques were employed in immotile and motile spermatozoon detection process, respectively. Additionally, active contour, dual thresholding and clustering were implemented to enhance the segmentation of immotile spermatozoon in this step as well. Thereafter, detected sperms were counted by pixel based blob analysis. Motile spermatozoa were tracked by the Mean Shift and the Kalman Filters for the motility analysis. Various motility features were extracted from the trajectories to classify them into four classes. As the final step, results were reported to the users.

Each module of the proposed software was individually tested. Two approaches of the automatic ROI detection and extraction steps were tested and compared on 80 videos of 20 subjects. Video stabilization idea was evaluated on 42 videos of 14 subjects. Two techniques were performed for the sperm concentration analysis. The performance of the Fuzzy C-Means based segmentation was measured on 15 videos of 5 subjects. A more advanced technique, dual thresholding and active contour based segmentation, was evaluated on 32 videos of 8 subjects. Finally, 32 videos of 8 subjects were used for the verification of the tracking technique. As a result, totally 201 semen videos obtained at different times from 55 subjects were included for the determination of the proposed spermiogram analysis approach.

In the clinical research, we initilaly compared the counting results of CASA system, VA, and proposed CSCTAS with the proper concentration calculation. Normally, experts separately and manually count the motile and immotile spermatozoon within 16 and 10 squares to generalize the result as million per ml in the VA technique, respectively. According to the concentration analysis, proposed CSCTAS resulted in similar outputs as VA. It has less variation for immotile spermatozoa counting and is more efficient than the VA for the determination of specific diseases such as Asthenospermia. It is known that conventional SQA-Vision CASA is useless in the case of less than 5 million sperm cells. Therefore, presented approach is more efficient in the infertility diagnosis. In the motility analysis, CSCTAS gives the similar outputs when compared to SQA-Vision CASA. The SQA-Vision is more reliable technique when compared to VA in motility analysis because it is impossible to track single spermatozoa by eye for a period of time within other spermatozoa. Therefore the similarities between SQA-Vision and proposed CSCTAS is more meaningfull than the comparision with VA technique for the motility analysis. According to the motility analysis results, CSCTAS is also efficient, cheaper and easier to use in labs for the motility analysis when compared to the conventional CASA systems.

According to the obtained concentration and motility results, the proposed smartphone based sperm analysis can be adapted with the developed with the proposed CSCTAS in laboratories. Our proposed system stands out by its modularity, functionality, accuracy and low cost. Additionally, it eliminates the human factor in VA and CASA.

**Key words:** Spermiogram, Makler Counting Chamber, Semen Analysis, Sperm Counting, Computer Aided Diagnosis Systems, Clinical Research

# ÖZET

## BİYOMEDİKAL GÖRÜNTÜLERDE HAREKETLİ NESNELERİN ANALİZİ VE TAKİBİ İÇİN BİR SİSTEM GERÇEKLEMESİ

Hamza Osman İLHAN

Bilgisayar Mühendisliği Anabilim Dalı

Doktora Tezi

Tez Danışmanı: Prof. Dr. Nizamettin AYDIN

Sperm Sayımı infertilite teşhisinin ilk aşamasıdır. Bilgisayar Destekli Sperm Analizi (BDSA) ve Görsel Değerlendirme (GD) sperm analizlerinde kullanılan iki değerlendirme tekniğidir. GD, sayım çemberi üzerinden sperm gözlemleyerek gerçekleştirilir. Bu nedenle, teşhis genelde beceri ve deneyime bağlıdır. Öte yandan, BDSA, gelişmiş bilgisayar teknikleri ve insan etkilerinin izolesi sayesinde daha gelişmiş bir teknolojidir. Ancak, tümleşik bir sistem olduğundan GD tekniğinden daha pahalı bir analizleme çözümüdür. Buna ek olarak, BDSA sistemlerinde, uğraştırıcı bir parametre belirleme işlemi gereklidir. Sistemin başarımı, belirlenen parametrelerle bağıntılıdır.

Bu tez çalışmasında akıllı telefon ve bilgisayar kullanarak sperm konsantrasyonu ve motilite analizi için kombinasyonel bir yaklaşım geliştirmeyi amaçlıyoruz. Akıllı telefon, GD tekniğine benzer görüntüler elde etmek için kullanılmıştır. Sonrasında, insan etkilerini ortadan kaldırmak için bilgisayarla analiz yöntemleri kullanılarak numune analiz edilir. Çalışma kapsamında, sperm sayım çemberleri üzerinden elde edilen görüntüleri analiz etmek amacıyla Bilgisayar Destekli Sperm Sayma ve Hareket Analiz Yazılımı (BSSHAS) adlı bir yazılım önerildi. Yazılım, sıralı olarak koşturulan yedi modülden oluşmaktadır: (1) Veri Toplama ve Organizasyon, (2) Otomatik Izgara Tespiti ve İlgi Alanı çıkartımı, (3) Video Sabitleme, (4) Hareketli / Hareketsiz Sperm Hücresi Algılama, (5) Spermlerin Sayımı, (6) Hareketli Sperm Takibi ve (7) Yörünge Sınıflandırması. Her adım, görüntü işleme tekniklerinden oluşan çeşitli kombinasyondan oluşur. Öncelikle, veri toplama ve organizasyonu, geleneksel BDSA sistemlerinin aksine, daha az maliyetli bir tasarım sağlamak için yeni bir yaklaşım kullanılarak gerçekleştirildi. Sonrasında, yazılım tarafında, İlgi Alanları (İA) çıkarma işlemi, çizgi algılama ve bölütleme yöntemlerinin kombinasyonel bir yaklaşımı ile

gerçekleştirildi. Ardından, veri toplama aşamasında oluşan titreşimleri ortadan kaldırmak için özellik eşleme tabanlı video sabitlemesi üzerine algoritma geliştirildi. Özelliklerin çıkartılması anlamında, farklı tanımlayıcılar test edildi. Dördüncü olarak, hareketsiz ve hareketli sperm hücresinin algılamasında sırasıyla arka plan ve ön plan çıkartma teknikleri kullanıldı. Ek olarak, bu aşamada hareketsiz sperm bölütleme başarımını arttırmak için aktif kontur, çift eşikleme ve kümeleme yöntemleri uygulandı. Bundan sonra tespit edilen spermler piksel tabanlı blob analizi ile sayıldı. Hareketlilik analizi için hareketli olarak tesbit edilen spermler, Mean Shift ve Kalman Filtreleri ile takip edildi. Yörüngelerden çeşitli hareket özellikleri çıkarılarak dört sınıfa ayrıldı. Son adım olarak, sonuçlar kullanıcılara bildirildi.

Önerilen yazılımın her modülü ayrı ayrı test edildi. Otomatik İA saptama ve çıkartma adımında, iki yaklaşım incelendi ve 20 hastadan elde edilen 80 video üzerinden denendi. Sonuçlar karşılaştırıldı. Video sabitleme çalışması, 14 hastadan elde edilen 42 video üzerinde değerlendirildi. Sperm konsantrasyon analizi için iki teknik test edildi. Fuzzy C-Means tabanlı bölümlemenin performansı 5 denekten elde edilen 15 video üzerinde ölçüldü. Daha gelişmiş bir teknik, çift eşikleme ve aktif kontur tabanlı bölütleme, 8 deneğin 32 videosunda değerlendirildi. Son olarak, 8 hastadan elde edilen 32 video, izleme tekniğinin doğrulanması için kullanıldı. Sonuç olarak, önerilen spermiogram analiz yaklaşımının test aşamasına, 55 denekten farklı zamanlarda elde edilen toplam 201 semen videosu dahil edilmiştir.

Klinik araştırmada, başlangıç olarak BDSA sistemi, GA ve önerilen BSSHAS sperm konsantrasyon sayım sonuçlarını uygun yoğunluk hesaplama tekniği kullanarak karşılaştırdık. GA tekniğinde, uzmanlar manuel olarak hareketli ve hareketsiz spermleri 10 ve 16 kare genellemesini kullarak saymaktadırlar. Konsantrasyon analiz sonuçlarına göre, önerilen BSSHAS, GA ile benzer çıktılar üretmiştir. Önerilen yöntem, hareketsiz spermların sayımında daha tutarlı ve az dalgalanma göstermektedir. Ayrıca, Asthenospermi gibi spesifik hastalıkların belirlenmesi için GA'den daha etkilidir. Geleneksel SQA-Vision BDSA sisteminin ise 5 milyondan az sperm hücresi bulunan örneklerde analiz yapamayacağı bilinmektedir. Bu nedenle, sunulan yaklaşım, kısırlık tanısında daha etkilidir. Hareketli sperm analizinde ise, önerilen BSSHAS, SQA-Vision CASA ile karşılaştırıldığında benzer çıktılar vermektedir. Tek bir spermatozanın diğer spermler içerisinde bir süre boyunca manuel olarak gözle izlenmesinin güç olmasından dolayı, SQA-Vision, hareketli hücre sayımı analizinde, GA tekniğine göre daha güvenilir bir tekniktir. Dolayısıyla, SQA-Vision ile önerilen BSSHAS arasındaki benzerlikler, hareketlilik analizi için GA tekniği ile karşılaştırmadan daha anlamlıdır.

Elde edilen konsantrasyon ve hareket analizi sonuçlarına göre, önerilen akıllı telefon temelli sperm analiz yaklaşımı laboratuarlarda uyarlanabileceği kanıtlanmaktadır. Önerilen sistemimiz modülerlik, işlevsellik, doğruluk ve düşük maliyet ile dikkat çekmektedir. Buna ek olarak, GA ve BDSA'deki insan faktörününe bağlı hataları ortadan kaldırmaktadır.

**Anahtar Kelimeler:** Spermiogram, Makler Sayma Çemberi, Sperm Analizi, Sperm Sayımı, Bilgisayar Destekli Tanı Sistemleri, Klinik Araştırma

---

# CHAPTER 1

# INTRODUCTION

## 1.1    Literature Review

Fertility is necessary and substantial function for human being to sustain the generation. In every segment of society, it is accepted as the key factor for the future. On the other hand, infertility is scientifically described in literature as being unable to conceive after 1-year relationship having regular sexual intercourse without any contraceptive substances [1]. Currently, it has become an increasingly significant worldwide problem according to one of the recently published statistical report [2] and affected the people, mentally and physically. Roser significantly emphasized in the report that the global fertility rates are decreased from 4.97 to 1.99 between 1950 and 2015 [2]. Figure 1.1 shows the total fertility rates (TFRs), which represents the number of children born to a woman, in Turkey, and the cumulative variability over the continents and world by years.
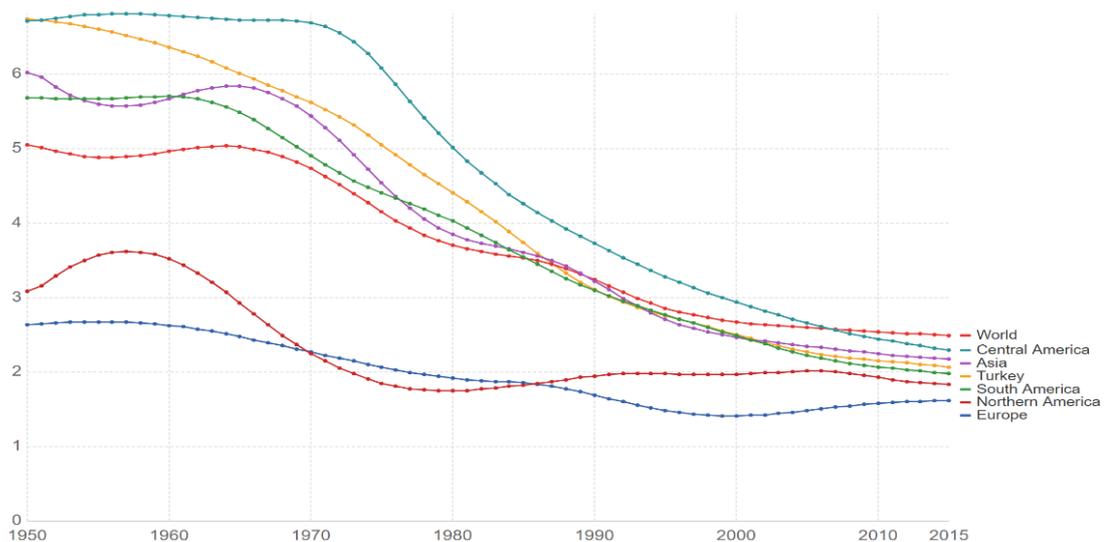
Figure 1.1 Total Fertility Rates (TFRs)

There might be many reasons behind the infertility problem which can be categorized as male, female, both side based or unexplained reasons. Figure 1.2 indicates the infertility rates of occurrence for specific reasons according to World Health Organization (WHO) [3].
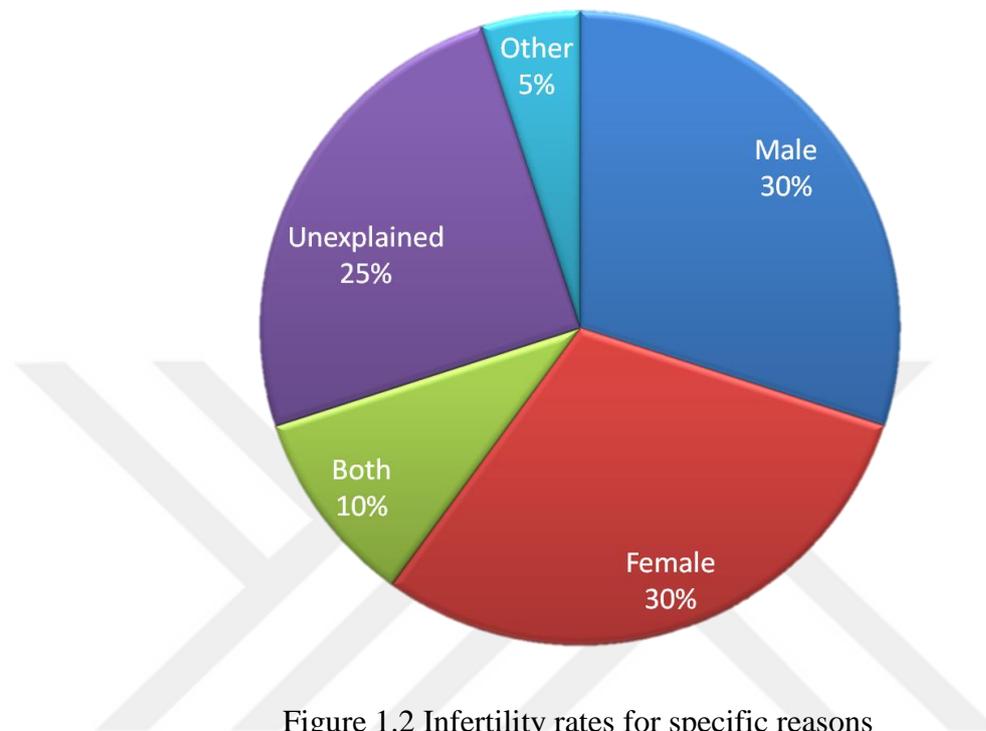


Figure 1.2 Infertility rates for specific reasons

Men are solely responsible for the difficulty in 20-25% of the time, and are involved in 20-25% of the time along with their female partners [3]. Consequently, male factor equals to the female based problems and effective in 30% of all infertility problems.

More specifically to the male infertility, another published report about the sperm quality and reproduction emphasized that the sperm concentration is decreasing to critical level, gradually. Carlsen et al. revealed that the number of sperm (sperm/ml) in the world surprisingly decreased by 50% in last 50 years [4]. In 1940, 113 million/ml was accepted as normal sperm concentration, but in 1990, it was described as 66 million/ml. After 8 years, Swan et al. indicated that 37 million/ml is a new normal sperm concentration [5]. The World Health Organization (WHO) also drew attention to the trouble on sperm quality. They published several manuals by updating over years to establish or validate reference limits for semen parameters such as volume, concentration, total sperm count, motility etc. 15 million/ml is stressed out as the normal sperm concentration in the last edition of the manual [3]. As a result, the sperm concentration is decreased 100 million from 1940 to 2010 and the reasons are not

clearly and fully described. The changes of motility rate is also reported as decreasing from 60% to 30% between 1970 and 2010. The investigation of the reasons of decrease has become the essential and crucial research topic for multidisciplinary researchers. Many researchers seek diagnostic semen analysis to reveal the causes of the male factor infertility. In this respect, medical doctors recommend the semen analysis tests as a first assessment.

Typical semen analysis named as spermiogram include the parameters as sperm concentration, total sperm number, percentage of motile sperm, percentage of forward progression, and percentage of normal morphological sperm shape. Tests are carried out using two evaluation techniques based on computerized or visual assessment [6]. In visual assessment technique, different standardization equipment should be employed such as Makler [7] or Hemocytometers [8] [9], to generalize the cell numbers in per ml. According to [10], Makler is one of the promising materials and the most employed chamber in laboratories. Today, visual assessment techniques is fast, heuristic and the most used technique in laboratories. But, the results strongly depend on the observer variability. Even if there are several practical guides to standardize the procedures such as [3, 6 and 11], the experience and the expertise are still the key parameters in the analysis. On the other hand, some laboratories utilize an automated analysis system named as Computer Aided Sperm Analysis (CASA) in computerized techniques. It is costly due to being an integrated system including embedded camera, microscope and computer. Additionally, the parameters cannot be readjusted and results cannot be verified by experts/doctors [12].

Computer based sperm analysis mainly utilizes two microscope imaging techniques; a) Phase-Contrast and b) Fluorescence dying [12, 13]. The Phase-Contrast microscopic images are obtained by using optical-microscopy. Studies in this kind of images mostly focus on shape analyses such as acrosome classification [14, 15, 16], tail/head detection [17, 18, 19], sperm counting [20, 21, 22], and morphological abnormality detection [23]. Additionally, several studies are reported about computerized analyzing motility of the sperms using Phase-Contrast imaging technique [24, 25]. However, challenging issues such as occlusion, illumination changes, different image sets, camera effects, and noise are presented in the related papers as the crucial problems that must be overcame. Therefore, another microscope imaging technique called fluorescence dying is useful to extract the objects from background and to form a noise free system. In this technique,

the field of view is more extensive, altough the details disappear due to the low resolution of wide area. In this imaging technique, motile sperm counting can be easily performed due to the reflection of light over semen as presented in [26]. However, semen sample also includes several different particles such as blood cells, debris, or leucocyte besides the sperm cells. In this kind of images, segmentation of immotile spermatozoa and other particles may not be perfectly achieved since there is no textural details. Only the spatial features (area, eccentricity etc.) of foreground objects detected by the motility can be used in segmentation. Figure 1.3 shows example images obtained from both techniques.



a                                          b
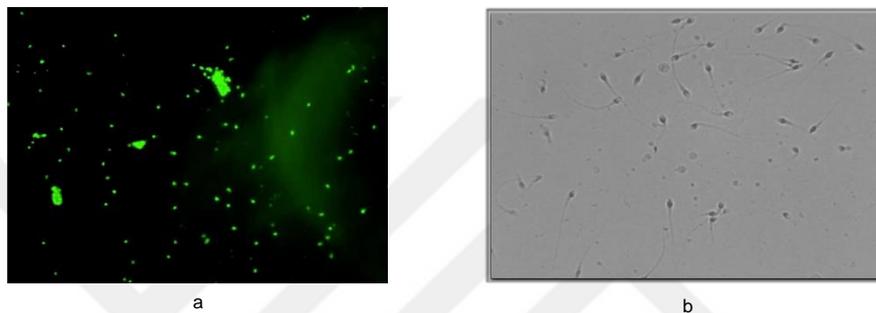
Figure 1.3 Microscope Imaging Techniques: a) Fluorescence dying b) Phase-Contrast

In medical literature, the substances used in the dying process is reported as spoiling the pure movement of the motile spermatozoon [27]. Therefore, fluorescence-dying technique generally excluded for the comprehensive analysis. Recently, computer based systems are preferred to perform all kind of analyses over images obtained by a camera on phase-contrast microscopy. But, the images obtained by the camera over the phase contrast microscopy might be in different scale, color space etc. In this respect, it is impossible to develop all-in-one software that presents the analysis for each camera due to the differences in images. Figure 1.4 indicates four different output images of different camera settlement on the phase-contrast microscopy.



Figure 1.4 Different camera outputs of phase-contrast microscopy

Addition to the microscopic images based methods, Gillan et al. proposed flow cytometry experimental techniques for the evaluation of sperm parameters on fertility rates [28]. However, it is more expensive and not practical technique for the daily laboratory analysis. It is reported in the study that the technique should be employed for the sperm analysis related with the gene mutation or inherence based problems which constitutes 25% unexplained reasons in Figure 1.2.

Today, many CASA systems such as SQA-Vision [29], SCA CASA [30], Hamilton Thorne [31] etc. are developed. Some of them provide only shape analyses named as morphological assessment while advanced systems setting up with high level computational power functionality focus on motility analyses as well as morphological assessment [32]. They are the compact systems including integrated camera with appropriate microscope and software. Microscopes with an embedded camera system are far more expensive than regular microscopes. Therefore the systems are extra costly for laboratories. Additionally, it is widely recognized that the image processing and sperm tracking algorithms employed in the today's CASA should be improved [12, 33, and 34]. Systems operate with predefined reference points; thereby it strongly requires more human intervention. Comparative studies such as [12, 27, 32 - 34] signify the importance of parameter settings for high accuracy and correct results. More doctor intervention to setting up the system is the essential part to obtain more trustful results. Well-regulated CASA over predefined limitations and requirements will provide more accurate and trustful results than manually observed analyses, but the settings will be suitable for the specific one case [35]. In another case, doctors/experts should re-arrange the parameters. In particular, most CASA cannot reconstruct reliably paths of two or more sperm in close distance or in intersect problem for motility analysis. They solved that problem by excluding cell-to-cell and near-misses sperms from the trajectory analysis in many CASA software. But, this exclusion ignores the higher velocities which is more likely to be involved in collisions and occlusions and reports the motility with slower results [36]. In another approach to solve the tracking problem, dilution was applied on the semen sample [37]. They reduced the collision by analyzing the diluted sample. But, this does not reflect the original sample due to the used substances for dilution process [37].

According to the pros and cons of two techniques (Visual and Computerized Assessment), computerized analyses should be employed due to the higher reliability

5

and observer invariance, but the idea should be also improved in several aspects such as practicability, modularity, intelligibility, cost, accessibility, capability etc. In this thesis, we intent to improve the CASA. According to all abovementioned microscopic imaging techniques, we also focus on the Phase-Contrast microscopic images with different image acquisition steps to form cheap and more modular sperm analyzing system. Similar to our approach, different studies emerged in literature to form CASA-like systems. Witkowski developed an integrated system and published the results describing the methods clearly in [39]. Another paper published by Wilson-Leedy and Ingermann. They referred an automated zebrafish motility analyses software [40]. Authors emphasized that the software is open source. Shi et al. presented an original design and prototype in [41] to analyze sperms in terms of motility with extra details than regular CASA systems. Differently, we focus on the computer based analysis of Makler images obtained from the ocular part of the phase contrast microscopes. We proposed a novel approach to the data acquisition step and analyzing techniques. Detailed literatures will be given in each chapter of the utilized and employed techniques and methods.

## 1.2    Objective of the Thesis

Sperm counting for the concentration analysis and tracking is generally carried out using two techniques: (1) Visual assessment by experts; (2) Automatic analysis by Computer Aided Systems (CAS). Visual assessment is the manual examination of samples over a counting chamber under microscopes by experienced laboratory assistants, biologists or medical doctors. It is more practical and heuristic method, but counting results can vary according to personal experiences and skills. Missing of even one sperm in manually counting process results in different concentration output depending on the counting chambers employed in the analysis. The error rate directly related with the expertise and the experiences of laboratory assistant. Problem is defined in the literature as the observer variability. Additionally, in the case of high sperm concentrations, it is impossible to count entire region, manually. Samples should be appropriately diluted into a fixative medium. But this affects the motility of sperms due to disrupting the viscosity of the semen sample. In this scenario, one of the CAS specified for sperm analyses, Computer Aided Sperm Analyze (CASA) systems, gives a better sight when compared to manual observation. It is one of the promising

technologies. There are several CASA systems in medical market. However, most laboratories still carry out spermiogram tests with visual assessment technique because it is cheap and no parameter setting is required in contrast to current CASA systems.

In order to decrease the observer variability problem in the VA diagnosis, we aimed to develop an automated semen analysis system using counting chambers. In this respect, we modelled the VA technique in the computerized analysis.

In terms of the VA modelling, we intended to develop a system that use the idea of visual assessment technique and consider to the current disadvantages of microscopy imaging techniques in CASA based sperm analysis. Basically, the combination of the data acquisition step of visual assessment and computerized analysis step of CASA systems is the main goal in the proposed system. Makler counting chamber as in visual assessment technique is utilized in our approach on computer side. Additionally, in this system, laboratory assistant can still observe the samples and count the sperm manually over the screen if necessity. In this respect, we also purposed to eliminate the mistakes, oversight or any other user dependent errors. Additionally, we aimed to provide the system as cheap, re-configurable, having wide field of view compared to existing CASA systems by using the visual assessment based data acquisition idea.

## 1.3   Hypothesis

In this thesis, our hypothesis is that to count and track the sperm cells by more accurately, objectively and sensitively by using modular, portable and inexpensive systems when compared to conventionally CASA systems. In this respect, we proposed a biomedical image analyzing software for spermiogram tests including the sperm concentration evaluation and motility analysis with a novel image acquisition method. Proposed approach named as Computerized Sperm Counting and Trajectory Analyzing Software (CSCTAS) stands out with its hardware independency, easy to implement on any kind of phase-contrast microscopy and cost by using a smartphone based data acquisition approach in the image acquisition step. Shortly, visual assessment technique is performed by the computer software with advanced image processing techniques. Proposed system provides more convenient and inexpensive way to count the sperm automatically into two categories as motile and immotile spermatozoa than the traditional CASA systems and VA technique. Additionally, motility analysis is easily

and better performed by the trajectory analysis after successful tracking of spermatozoa than the most CASA systems.

CSCTAS consists of 7 sequential modules; Data acquisition and organization, Region of Interest (ROI) detection and extraction, Video stabilization, Motile/Immotile Spermatozoon Detection, Spermatozoa Counting, Motile Spermatozoa Tracking, Trajectory Classification. Flow diagram is presented in Figure 1.5.
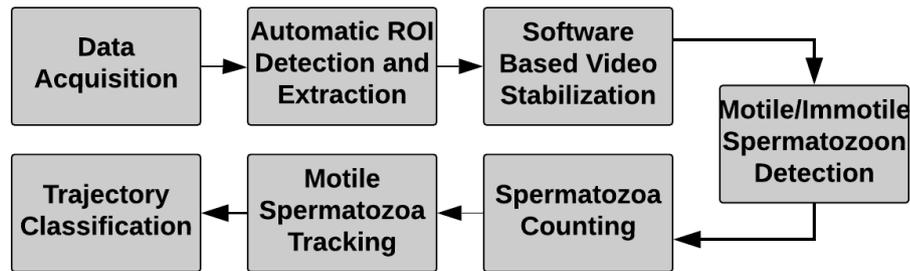


Figure 1.5 Flow diagram of Computerized Sperm Counting Software methodology

Each step will be explained with details in the corresponding chapters. Briefly, in the data acquisition step, the novel approach to the current spermiogram techniques will be explained. The images obtained from the novel approach includes different parts. Therefore, images are required to be split into the sub-regions which will be analyzed in the sperm concentration analysis part of the software. Two techniques performed in this part will be explained in the automatic ROI detection and extraction chapter. Ocular part of microscopy is a sensitive part to the vibrations which will directly affect to the images. Therefore, we used a software based video microscopy stabilization idea which will be introduced in the next chapter. Stabilization eliminated the miss detections of the spermatozoa. The concentration evaluation will be evaluated in the Motile/Immotile spermatozoon detection chapter. Makler counting chamber was utilized in our spermiogram analysis approach. Therefore, the suitable counting process based on the Makler counting standardization rule will also explained in the same chapter. As a last analysis, motility analysis will be given in the last chapter with the evaluated two tracking algorithms.

## DATA ACQUISITION AND ORGANIZATION

Analyses of visual assessment technique are performed by using different counting chambers to standardize the counting process [10]. Makler is one of the commonly used chambers in laboratories due to the easy implementation. In Makler chamber, the number of counted sperm cells in one strip of grid can be generalized in millions per ml as the sperm concentration [7]. Makler counting chamber and the accepted rule for the spermatozoa concentration analysis is demonstrated in Figure 2.1.

Figure 2.1 Makler Counting Chamber and the generalization standard in sperm concentration analysis

In this thesis, developed CSCTAS works on the Makler images to generalize the results as in visual assessment (VA) technique. Ocular images of optical phase-contrast microscopes are similar, therefore, a smartphone based data acquisition was implemented. Videos were recorded from the ocular part of microscopy by using a designed apparatus. As a result, extra improvements, modification or adaptation for each microscope type are not needed in the software. Apparatus provides more portable and easy attachable system. Designed apparatus and the data acquisition by the smartphone over the ocular part of microscopy are illustrated in Figure 2.2.
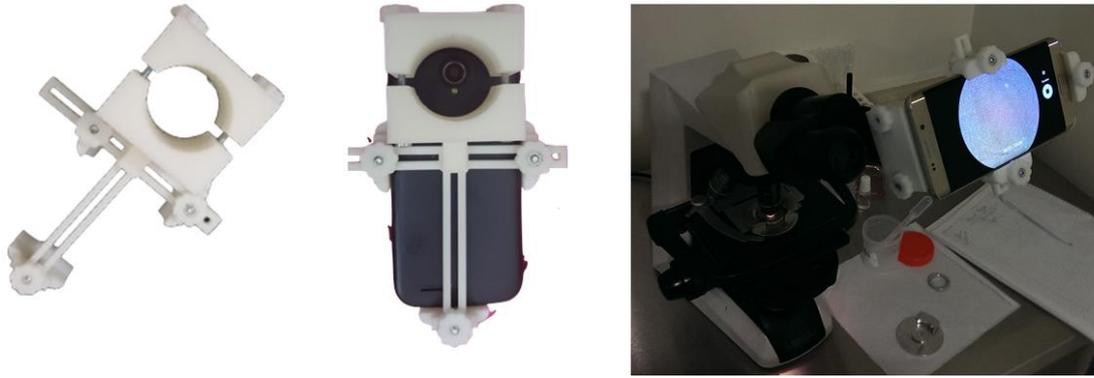
Figure 2.2 Smartphone based data acquisition approach

Makler videos were recorded by the mounted smartphone and transferred to the central storage over local network for computer side analysis. Because of high speed motility and ability change direction, sperm is known as maneuver target. Hence, the videos of the samples were recorded by high-resolution camera with 1920x1080 resolution and 30 Hz frame per second for better and more detailed analysis. Duration of video can be different. Each video were automatically split into 30 seconds lengths sub-videos after transferring to the computer software. Concentration and trajectory analysis were individually performed over the each sub-videos by the CSCTAS. The counting result of entire video was then calculated by averaging results of each sub-sample counting process.
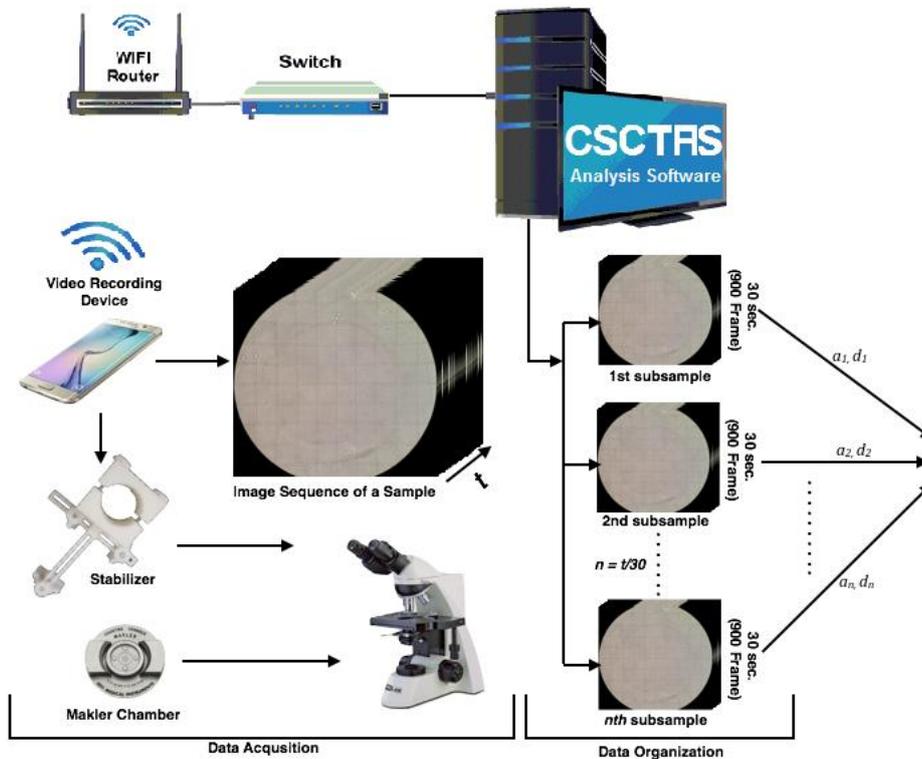


Figure 2.3 Sample Acquisition and Organization

10

Figure 2.3 demonstrates the data acquisition and organization schema. $a_k$ and $d_k$ represents the number of motile and immotile spermatozoon counted in each sub-samples, respectively. *n* refers to total number of created sub-samples from the main sample video *(t)* with respect to frame rate of video recording device. *A* and *D* denote the final average numbers as motile and immotile spermatozoa according to sample numbers of corresponding subjects. An example image that is obtained by the presented data acquisition step and transferred to the server side analysis software is shown in Figure 2.4.
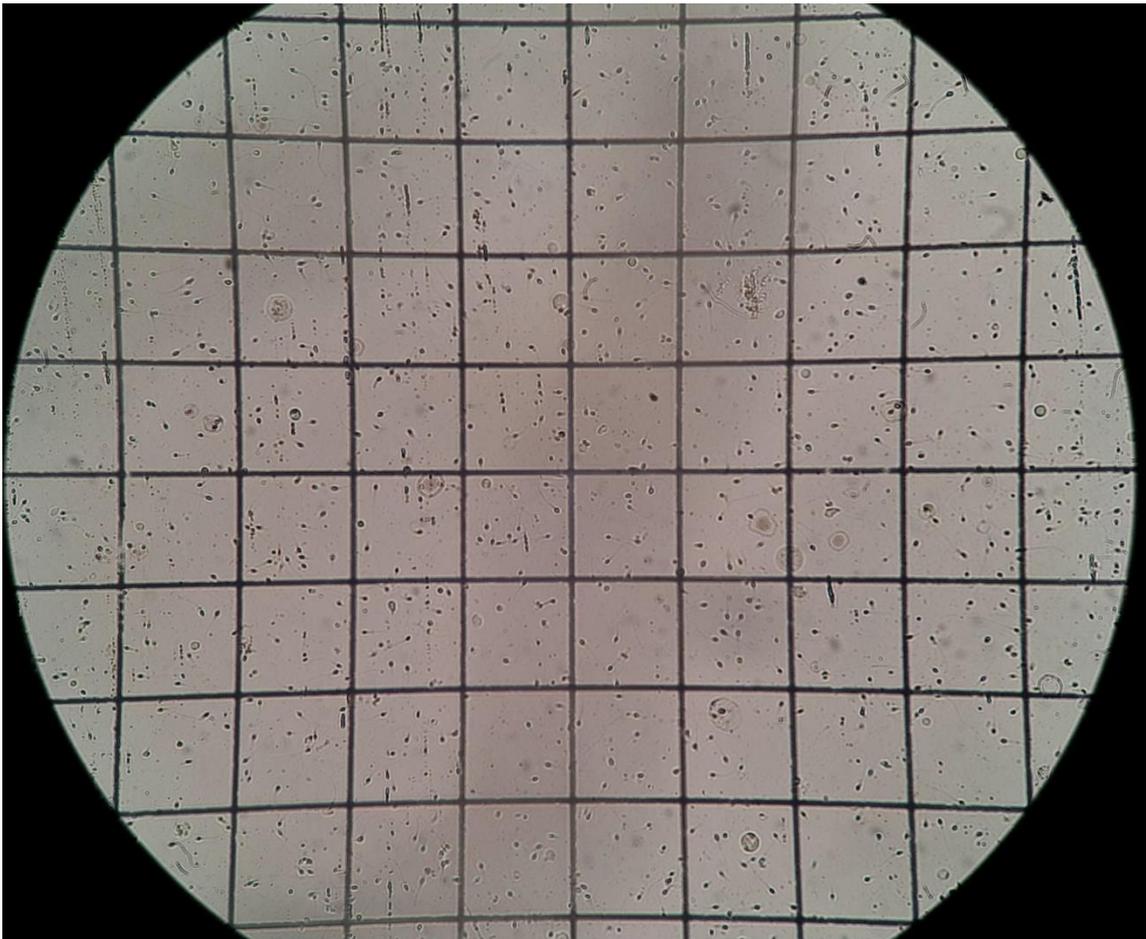


Figure 2.4 Example ocular image obtained by the proposed smartphone based data acquisition technique

Each sequential steps of the server side analysis program, CSCTAS, will be explained with details in the following chapters.

# REGION OF INTEREST DETECTION AND EXTRACTION

## 3.1 Introduction

Spermiogram tests include a manual counting process using a Makler counting chamber. The automatic detection of ROI in Makler images constitutes the first phase to use the advantages of the Makler chamber in a computerized counting system. ROI is defined between grids, hence, another challenging issue, that of exact grid detection, is examined. In the proposed CSCTAS, initially several line detection algorithms with their applications and possible usage on the grid-detection problem of Makler images were reviewed. Next, a combinational grid-detection technique, particularly for Makler images, was introduced.

Makler images include a square grid structure with horizontal and vertical lines. Grid detection is mainly a line detection problem. However, it requires eliminating lines out of grid structure and line crossings into a complete grid structure form. Each line is normally in straight form, but there might be curvature or skewness occurring through lens distortion of the camera. Additionally, there might be extra lines that overlap the grid structure or partly constitute grid-like lines on images caused by microscope or Makler chamber effects. Furthermore, immotile spermatozoa positioned next to lines affect line detection by ruining the linearity of the lines. In this context, current line detection and medical image processing studies are explored to overcome challenges.

Line detection is theoretically a segmentation problem and various segmentation techniques in medical imaging have been conducted so far. Massood et al. gathered different approaches on the segmentation of medical images in their survey [42]. McLean et al. [43] used hierarchical clustering for line detection, where the clustering of pixels is maintained by the spatial contiguity and similarity of average gradient

orientation. Segmentation is used in the estimation of line equation. They compared four different strategies on synthetic and real images. The study emphasized the modularity of the proposed approach, but complained about a distortion effect of lines on results. Therefore, the pixel-based line detection algorithm is considered an unstable technique on the Makler grid structure because of the lens distortion effect. Schuster and Katsaggelos reported the importance of an edge detection process for problems in line detection [44]. They tried to develop a line tracking robot in nature. However, edge detection is easily affected by noise that causing false line detection. The authors used minimum mean square error (MSE) to weight the edges. Consequently, the lines were detected with less edges and noise. However, they also reported that the reduced edges caused interrupts on detected lines. Therefore, their approach resulted in less success for short and weak lines. The adaptation of this technique for grid-detection would result in cropped lines because of the effect of immotile semen cells near the grids.

The Hough transform (HT) is a form of the Radon transform technique. It is mainly used for line detection applications by connecting edges in different angles. Herout et al. published a review chapter about the usage and the comparison of various HT techniques. He emphasized in the study that the selected parameters in techniques have crucial importance on line detection [45]. Different shapes require different parameterization and transform. Therefore, it is hard to form a more adaptive HT solution on the detection of different shapes. Tchinda et al. used HT to detect circles [46]. After, they employed the circles as the initial contour for an active contour technique to detect parasites, reporting that the proposed schema was efficient for the presented dataset. However, single HT implementations fail in the case of more complex backgrounds such as Makler images and require a sufficient number of points on straight lines. It is efficient in the detection of long lines, but useless for short lines. Our grid structure, however, is mostly made of short lines. Additionally, lines are not fully straight because of lens distortion. In other words, regular HT is not fully adaptable on Makler grids. A modified HT technique is presented in [47]. Li and Tsai utilized regular HT with a distance tolerance calculation.

The grids in Makler images are similar to the white lines on the road detection studies. However, lines on the road provide enough in terms of distance for analyzing and the feature extraction process. Additionally, this applications have no immotile spermatozoa or lens distortion effects. Hayashi et al. extracted the luminosity feature of detected lines

by a double circular operator [48]. They also performed HT to detect lines, but it was enhanced by the double circular operator in terms of luminosity features. It was reported in that paper that line detection with a luminosity feature is more robust to shadows or occlusions. It might be a solution for the immotile spermatozoa effect, but lines should be entirely straight to eliminate.

As a special case, Ahmad et al. [49] used dynamic programming and machine learning to detect horizon lines. They presented an edge-less method to detect lines. In their study, they employed support vector machine (SVM) and convolutional neural network (CNN) techniques similar to another study made by Lie et al. [50]. The results are promising, but the technique requires a training set. Another HT-based approach is presented by Chen et al. [51]. They modified HT with a curve-fitting algorithm to detect intersection points over lines. However, this solution is also designed for fully straight lines obtained by a camera without any distortion effect.

The idea of clustering lines is used by Lee et al. and is named the TRACLUS method [52]. They clustered the trajectories into a set of line segments. Then, another clustering was performed to group the line segments. The proposed technique formed sub-trajectories from main trajectory. Finally, the sub-trajectories were evaluated for similarity with the main trajectory. According to the results, the proposed technique correctly discovered sub-trajectories. Another clustering idea was tested in human behavior analyses. Similar to [52], Piciarelli and Foresti segmented the human trajectories obtained by surveillance camera systems [53]. They implemented a tree structure over predefined distance metric in terms of cluster similarities where the centers updated by tree.

Wu et al. aimed to find the chest wall line in magnetic resonance images of breasts [54]. They stressed that the chest wall line is a different shape in each case due to the physical differences in the human body. In these cases, using regular line detection techniques are not effective. Therefore, they conceived the idea of joining differently detected lines [54]. They mainly used the Canny edge detection technique on anisotropic diffusion and bilateral filtered images. This resulted in many line detections and they combined them in a direction to compose a full path. This process also eliminated the contrast, or device-based noise effects on detections. Another biomedical-related study was made by Uyen et al. [55]. They proposed an effective approach for automatically detecting retinal blood vessels. They implemented line detectors in various scales and 12

14

directions (angular resolution of 15°) to connect parts linearly. Current practice on this issue reports that the hardest problem is in defining the direction angle due to the many lines at different angles. Uyen et al. formed the concept known as the "winning line," which uses gray-level pixel values to define correct angles [55].

In terms of fuzzy meaning, Dave et al. adapted a c-shell clustering algorithm on ellipse detection [56]. They emphasized that the regular c-shell algorithm had a convergence problem on partial shapes. They modified the algorithm, tested over real images, and compared their improved algorithm to regular HT. It was emphasized in the paper that the proposed method is faster and requires less memory than HT. In another paper specifically on the line detection problem [57], they addressed the problem of line detection algorithms with regard to the definition of the cluster center step. In this case, they made the argument that the proposed technique automatically defines the centers of clusters, especially on polynomial shapes, better than HT.

An image processing application requires pre- and post-processing steps. Filters are the most implemented techniques [58]. Karuppathal and Palanisamy used a hybrid median filter to preserve the edges of tumor images in post-processing segmentation [59]. Another filter, the Gaussian filter, has been applied on the same segmentation problem of tumor images in several machine learning approaches [60]. Both such studies have emphasized that the filtering increases the results.

Segmentation performance of lines is evaluated by different metrics in literature. Matsopoulos and Economopoulos tested five matching algorithms on 263 medical image pairs and reported the subjectivity and reliability of the algorithms [61]. Another metric for images without ground truth (GT) has been published by Erdem et al. [62] using inter- and intra-frame information of objects. In our study, we created a GT of each grid and implemented one of the matching algorithms such as in [62] for the evaluation of a combinational approach to the grid detection problem.

Makler chamber consists of a 1 sq. mm grid structure in the center, which is subdivided into 100 squares of 0.1 x 0.1 mm each, and a 10-micron deep reservoir for the semen sample [7]. Makler is the shallowest of all the various counting chambers [10]. The Makler chamber provides a counting standardization with squares. The number of immotile spermatozoa per ml can be generalized according to sperm numbers observed in any 10 squares of the grid. In other cases, motile spermatozoa analysis is performed

by counting motile sperm over 9 or 16 squares within a defined time. Normally, squares can be examined easily by visual observation; however, in computer assessment, points should be emphasized and the region within the points should be extracted as ROI. In this respect, grids must be detected with high accuracy rates and F-measure scores in the first stage; then, the ROI clarified from the grids should be extracted by using detected grids.

Makler can be placed at different angles to the microscopes. Additionally, particles and/or substances remained in the Makler chamber can be seen in the background of the images with the halo effect occurred by the microscope light which disrupts detection of grids. Examples of several Makler images derived from the ocular videos are given in Figure 3.1.
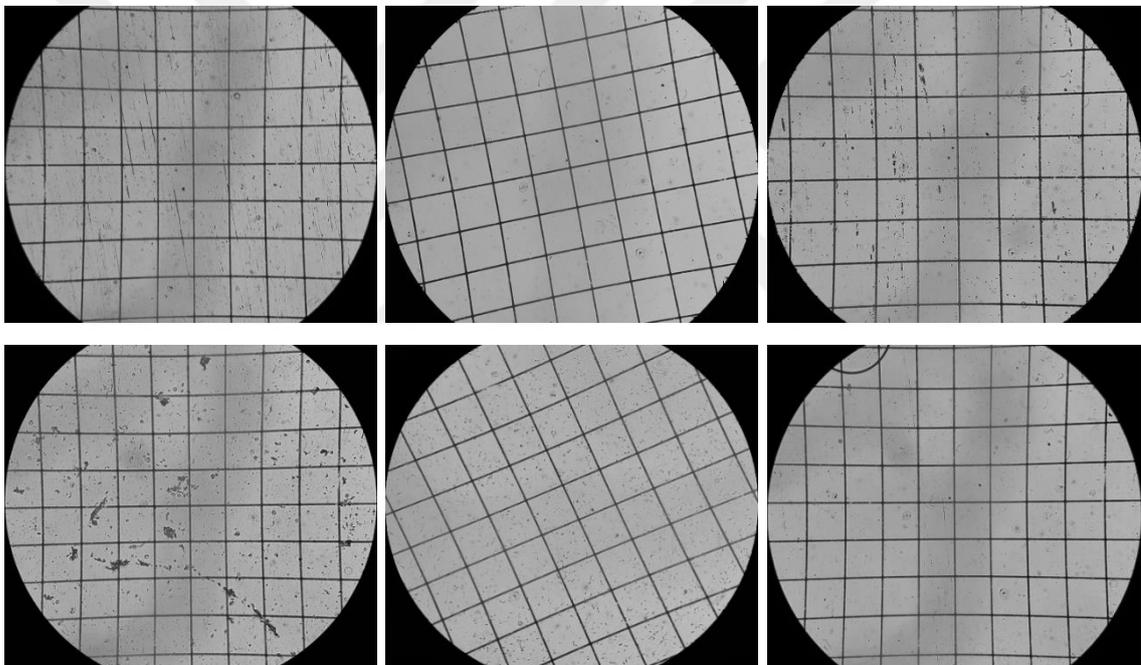


Figure 3.1 Example frames of recorded Makler videos

Two approaches for the automatic grid detection of Makler images were introduced to the literature based on the abovementioned studies [63, 64]. First one is not an efficient technique, the second approach is a robust technique. Therefore we will discuss the second approach in advance.

## 3.2 First Approach (Hough Transform + Angle Clustering)

HT and K-Means clustering were employed in our first approach [63]. We implemented a well-known algorithm; Hough Transform, after Otsu Thresholding. Then, the angle of

16

each detected line by the HT were calculated. K-Means clustering within a certain range is utilized over angles for the line elimination process as outlier values. Finally, the image is rotated by the cluster centers which indicating the weighted angle value. The flowchart of this approach is presented in Figure 3.2.
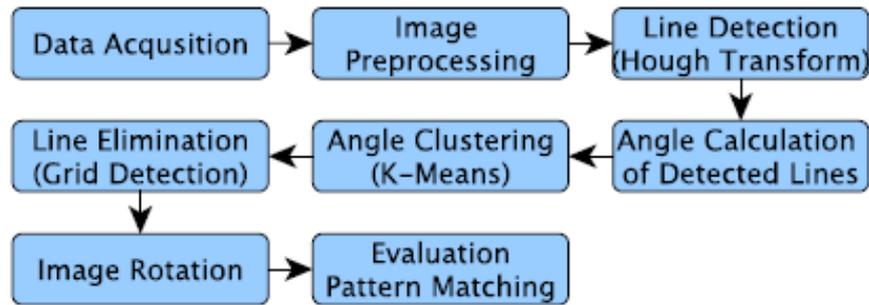


Figure 3.2 The flowchart of the first approach for the grid detection in Makler Images

Using only HT in the line detection step resulted in inefficient F-measure scores as a result of the difficulties in the obtained images as demonstrated in Figure 3.3. Blue lines indicate the Hough Transform result and the red lines show the only remained lines after K-Means clustering of angles.
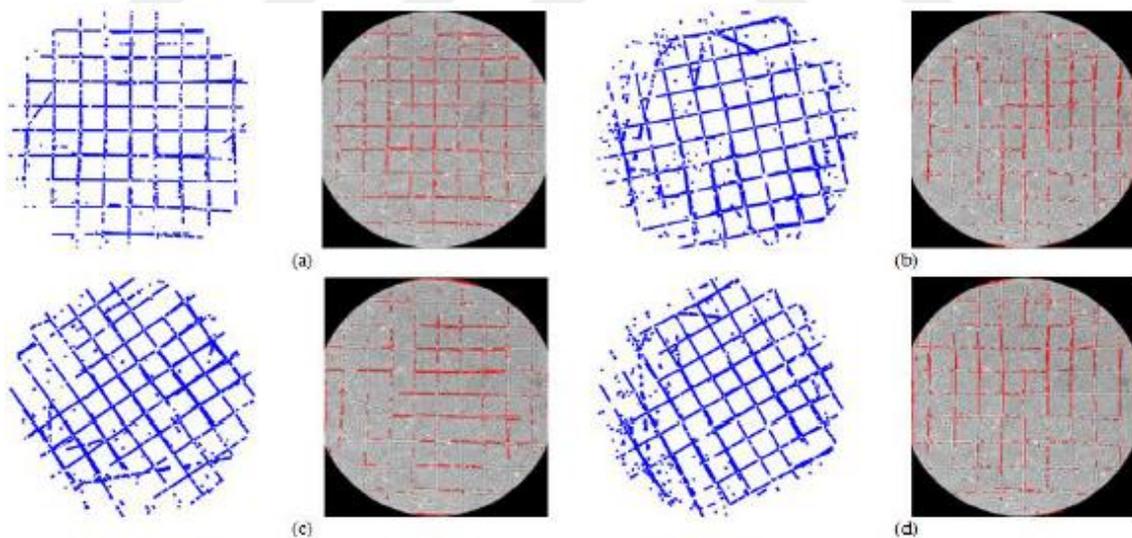


Figure 3.3 Flow chart of methodology

To overcome this problem, we used two different line detection algorithms in combination, HT and line segment detector (LSD) [65].

17

### 3.3　Second Approach (Line Segment + Angle Clustering + Hough Transform)

The idea came from the aforementioned studies in the literature review section. Conventional HT is inefficient on the detection of mixed lines because of being a parameter-dependent algorithm. It is more appropriate to use on direct lines rather than the skewed shapes. Defining optimal parameters is a hard problem. We also experienced this difficulty in our previous work. In this approach, LSD was initially utilized as line segmentation [65]. Each segmented line was, then, combined to each other by using the joining idea. We employed HT to connect segmented lines. However, HT requires an angle of direction to combine line parts. We performed another segmentation, using K-Means clustering, on each LSD line angle to form a "winning line". A filter process was then adopted as a post-processing step. Finally, evaluation was measured by a cross-correlation technique or, in other words, template-matching. According to the satisfactory results, ROI selection of the images was realized by the logical queries upon grid detections.

The flowchart of the second approach for the ROI detection is given in Figure 3.4. Initially, a background model is extracted. Then, LSD is performed on the background image to detect lines superficially. Pre-segmented lines are then clustered by K-Means algorithm into two categories using slope angles of lines. HT provides reinforced lines according to predefined parameters arranged by obtained class information of angles. If the Makler chamber is situated with a different rotation than 0° according to clustering result, the video is adjusted to a 0° state. As a final step of grid detection, filters, in terms of post-processing, are utilized to sharpen and clarify the grids. The middle zone of images including 36 squares is, then, selected as the target zone of images. Each square inside that area is accepted as ROI, individually saved as an image with as much line elimination as possible. The grid detection process is only performed on the first image of the video sequences. The remainder of the images are accepted as identical (except small vibrations) with the first frame because of the fixed Makler chamber during the video recording process. Later, ROI for each frame are extracted without grid lines upon already detected coordinates of squares derived from the first frame. The details of the first approach will not be given due to the low F-measure scores but the second proposed approach will be explained in the following subsections.
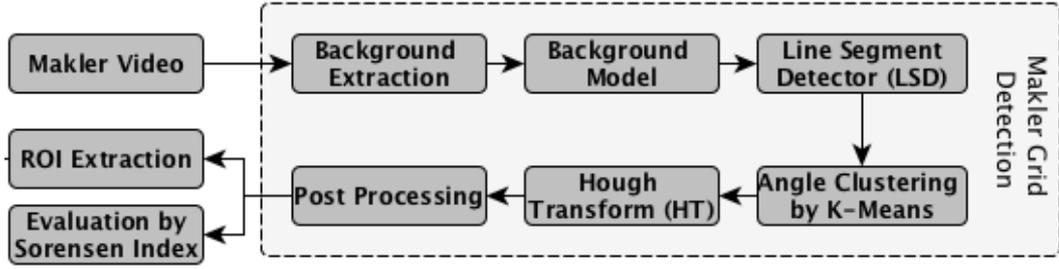
Figure 3.4 Flow chart of the second approach for Grid Detection

### 3.3.1 Background Extraction

The grid layout and immotile objects are stable during the videos but the videos also include high motility cells, which ruin the linearity of the grids during the detection process. Additionally, in some cases, motile semen can be detected as lines by HT or/and the LSD algorithm. In this case, segmentation of videos into background and foreground was deployed in the first step to avoid motile cell effects. While immotile cells and grids was left in the background image, motile semen remained into a foreground image.

The mean value of the pixels alongside some duration of the frame set will reflect the states of pixels for stable objects. In this case, an adaptive mean filter algorithm was applied to calculate the mean of images in video sequence similar to [66]. Three seconds of video (90 frames) was arranged as a down sample interval to skip 90 frames after each processed frame. Basically, the down sample interval allows ignoring the motile semen through the detection of variable pixels within a defined time interval. It allows the detection of the steady objects by observing constant pixels as well. Mean filter algorithm is mathematically presented in Equation (3.1).

$$B(x,y)^k = \frac{1}{d} \times \sum_{i=1}^{d}\left(V(x,y,t-i)\right) \qquad k = 1,2,3,...,\frac{N}{d} \qquad d < N \qquad (3.1)$$

where $d$ represents the down sample rate (90 frame), $N$ indicates the total number of frames (~1800 frame) in the video. $B$ is the $k$th background model derived from the image sequence $V$ alongside time $t$. Two example of background extraction are shown in Figure 3.5. Images are in high resolution, thus only a small part of the image is manually cropped and displayed to indicate the effectiveness of the background extraction technique.
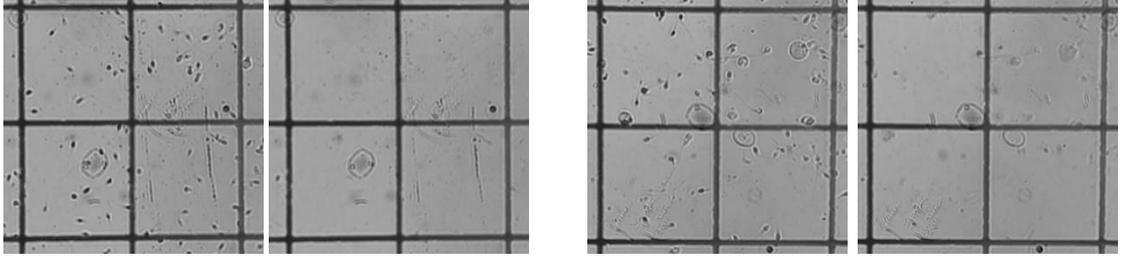
Figure 3.5 Two example image of Original (L) and composed Background model (R)

### 3.3.2 Line Segment Detector (LSD)

A linear-time Line Segment Detector (LSD) detects locally straight contours of images [65]. It is a piecewise technique that divides each contour with a verification process during their linearity and concludes the line segmentation using an adaptive and automatically arranged threshold point according to verification results. Algorithmically, LSD has three main steps. Basically, contours are the parts of images where the gray level rapidly changes. The gradient of the image as magnitude and orientation is employed in LSD for the initial phase. Gradient magnitude gives the power of the changes alongside the corresponding direction and the orientation refers to the calculated magnitudinal direction inside the image.

Gradient of an image consists of the partial derivatives of pixels in direction of $x$ and $y$ as given in Equation 3.2. $I$ refers to the intensity values of an input image. Expansion of the Equation 3.2 is given in Equations 3.3 and 3.4 where the gradients of the pixel at $(x,y)$ location of image $I$ calculated for horizontal and vertical direction separately. $\Delta x$ and $\Delta y$ indicate the distances to be convoluted with the center, respectively. LSD uses a 2×2 mask that refers to the convolution with the 1 pixel away from center.

$$\nabla I = \left( \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right) \tag{3.2}$$

$$g_x(x, y) = \frac{\partial I(x, y)}{\partial x} = \lim_{\Delta x \to 0} \frac{I(x + \Delta x, y) - I(x, y)}{\Delta x} \tag{3.3}$$

$$g_y(x, y) = \frac{\partial I(x, y)}{\partial y} = \lim_{\Delta x \to 0} \frac{I(x, y + \Delta y) - I(x, y)}{\Delta y} \tag{3.4}$$

Gradient magnitude based on the obtained pixel gradients is calculated by using Equation (3.5). Additionally, gradient orientation named as level line angle is deployed to use in LSD algorithm. In this case, Equation (3.6) is adapted. $g_x$ and $g_y$ indicate the

gradient magnitudes of *x* and *y* direction derived from Equation 3.3 and 3.4, respectively.

$$M(x, y) = \sqrt{g_x^2(x, y) + g_y^2(x, y)} \qquad (3.5)$$

$$\theta = tan^{-1}\left[\frac{g_x(x, y)}{-g_y(x, y)}\right] \qquad (3.6)$$

As a last part of the initial phase of LSD, gradient magnitudes are sorted by using a greedy algorithm called pseudo-ordering. The second phase includes the rectangle arrangement procedure. The algorithm expands randomly a selected pixel that is derived from the regions where the gradient magnitudes are of maximum value, to a rectangle shape in gradient direction with its eight-connected pixel neighborhood. Finally, LSD performs a validation process to control the linearity of the detected shape to decide the region as line. Thresholding is performed by multiple criteria, i.e., angle validation, rectangular approximation, number of false alarms computation, and point alignment densities. In our study, usage of LSD is denoted as follow;

$$L_{LSD}^k = LSD(B) \qquad (3.7)$$

where *B* represents the derived background model of original image by Equation (3.1). *k* indicates the number of segmented line. *LSD* symbolizes the technique which sequentially performs the aforementioned processes on background *B*, and segmented lines are recorded as *L* with the corresponding segment number *k*.

LSD was mainly performed for the preliminary detection of lines in our study. According to the abovementioned *LSD* methodology, lines were segmented into same groups until intersecting points of lines, the curvature regions of lines, and immotile spermatozoa effects were seen. Segmented lines were then transferred into clusters using angle information.

### 3.3.3 Angle Clustering of Segmented Lines

Clustering is a way of producing image segmentation without the necessity of a training process [67]. Samples are assigned to different classes based on distances between the sample and cluster centers. Mainly, algorithms aim to minimize the sum of squares within clusters while maximizing the distance of cluster centers. In this study, the K-Means clustering algorithm was utilized because it is easy to implement, fast, and efficient. Slope angles ($\angle^k$) of the segmented lines by LSD are calculated by using

Equations (3.8) and (3.9). First, slopes ($m_k$) of segmented lines are derived from Equation (3.8), then, angles are found by Equation (3.9).

$$y^k = m_k x^k + b_k \tag{3.8}$$

$$\angle^k = \tan^{-1} m_k \, x \frac{180}{\pi} \tag{3.9}$$

Clustering is performed based on the minimal values within the clusters. Equation (3.10) represents the deployed K-Means clustering technique in our study.

$$\arg min_S \sum_{c=1}^{2} \sum_{k \in S_c}^{N} \angle^k - \mu_c^2 \tag{3.10}$$

where $k$ indicates the number of segmented lines ($L_{LSD}^k$) up to total $N$ lines $(l_1, l_2, l_3, \ldots, l_N)$. $c$ is the class number which is two class in our study as *horizontal* and *vertical* alignment. K-Means clustering aims to put $N$ lines into two sets $S = \{S1, S2\}$ taking into account of minimizing the sum of squares. When the algorithm converges a local optimum, the mean and standard deviation of each clusters are calculated by using Equation (3.11) and (3.12), respectively.

$$\mu_c = \frac{1}{N} \sum_{k=1}^{N} \angle_c^k \tag{3.11}$$

$$\sigma_c = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (\angle_c^k - \mu_c)^2} \tag{3.12}$$

$$c = \{1(Vertical), 2(Horizantal)\}$$

According to inner cluster standard deviation, line elimination is maintained by the distances between the angles in cluster and the mean of the cluster until maintaining the following criteria (3.13):

$$0° < \sigma_c < +2° \tag{3.13}$$

Five lines, which are the furthest away from the mean of the cluster, are selected to be removed in each iteration. Then, clustering is performed again over the angles of remaining lines. When the inner cluster standard deviation is converged to Expression 3.13, the line elimination process is finalized. Figures 3.6(a) and 3.6(b) demonstrate K-means clustering on an example image having a 168° (horizontal – equals to 12° in compliment) and 78° (vertical) grid structure. After K-means, the slope angles of each

of the segmented lines by LSD are mostly gathered around two clusters due to the presence of the grid structure.
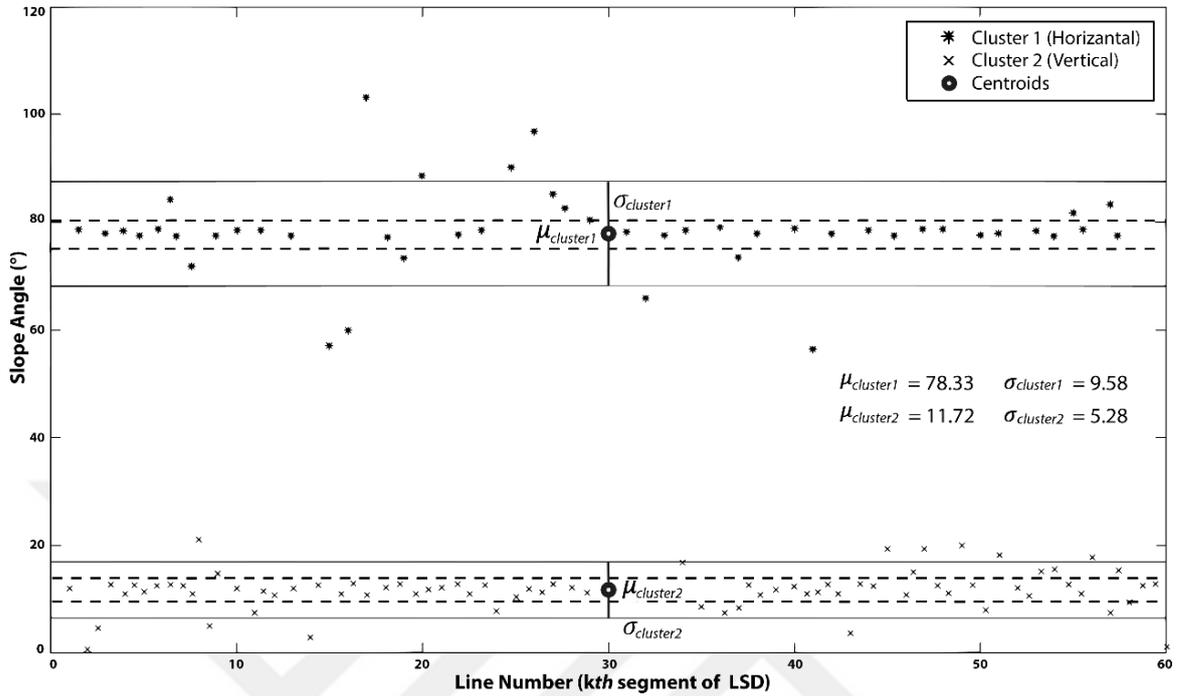


(a)
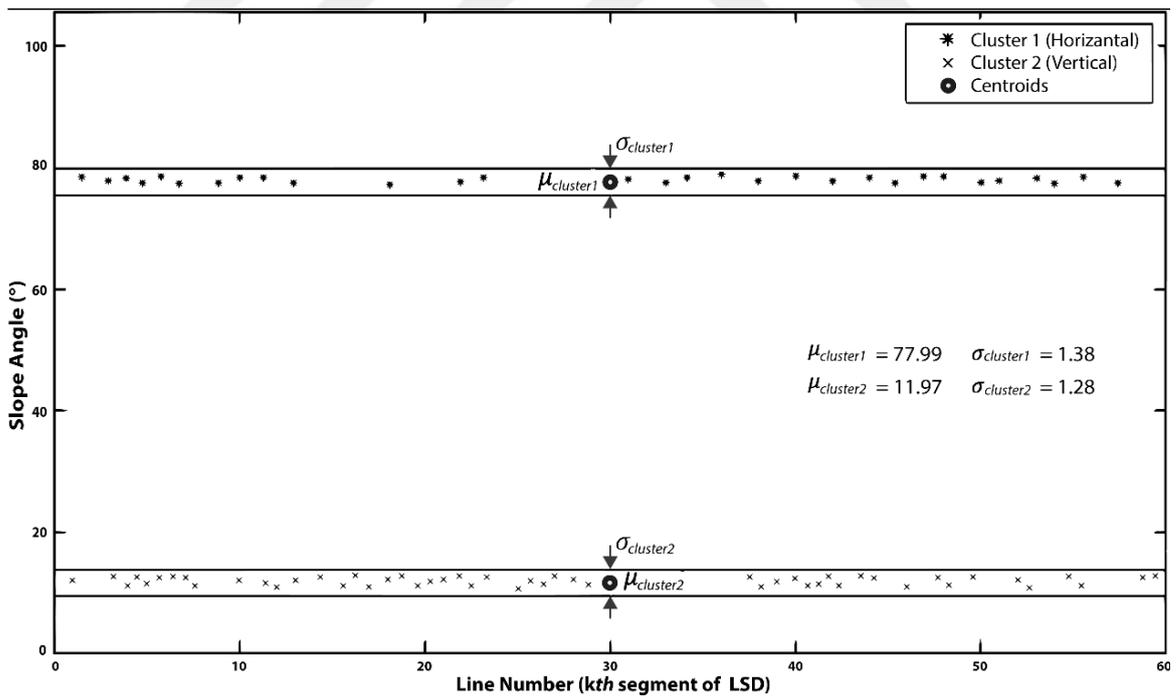


(b)

Figure 3.6  Slope angle clustering of an example image having 12° and 78° grid structure; a) First Iteration, b) Convergent state

*Cluster 1* is demonstrated with a star (*) and a cross (x) indicates *Cluster 2* on the figures. Dotted lines in Figure 3.6 (a) represent the reference zone denoted in Expression (3.13). Straight lines are the ± standard deviation limits around cluster centers. It is calculated as ±9.58 for *Cluster 1* and ±5.28 for *Cluster 2* for this example. In clustering theory, we adopted the K-means technique to align straight lines to the dotted lines level by elimination of lines. The convergence is defined as the alignment of straight lines inside dotted lines. Figure 3.6 (b) shows the convergence state of Figure 3.6 (a) after 4 iterations.

### 3.3.4  Hough Transform

The Hough Transform is a feature extraction technique used in image analysis and computer vision [68]. It is a variation of Radon transform. It is used to detect lines, circles, or other parametric curves by extracting several key features related to gradient changes on an image and performing a voting procedure for a certain class of shapes.

We adopted HT in our study to provide the connection of segmented straight lines after LSD and K-Means clustering processes. Eventually, HT forms curving lines by merging segmented straight lines. Under the HT model, multiple lines, named as virtual lines at different angles passing through to the center point of segmented lines ($L_{LSD}^k$), are formed by using Equations (3.14) and (3.15). Only the remaining LSD lines after the clustering process are enhanced by HT in the direction of the cluster center.

$$y^k = m_k x^k + b_k \qquad k=1, 2, 3, ... N \tag{3.14}$$

$$m_k = tan(a_t^k) \tag{3.15}$$

$$t = \mu_c - 2°, \mu_c - 1.5°, ..., \mu_c + 1.5°, \mu_c + 2° \tag{3.16}$$

where *a* represents the angles of the tested virtual lines inside $\mu_c - 2°$ to $\mu_c + 2°$ with $0.5°$ intervals. Slope (*m*) is derived from Equation 3.15. $x^k$ and $y^k$ refer to the center coordinates of *k*th segmented lines ($L_{LSD}^k$) after elimination process by clustering. $b_k$ is the intersect point of lines. Altogether, nine virtual lines with $0.5°$ intervals for each segmented line were tested. Perpendicular lines to the constituted virtual lines were then created to evaluate the angle alignment state of virtual lines. Distance between the origin and the intersect point of a perpendicular line and virtual line was measured by using Equation (3.17). $d_t^k$ indicates the distance measurement of the *k*th line to origin with respect to *t* degree.

$$d_t^k = (x^k \times \cos a_t^k) + (y^k \times \sin a_t^k) \qquad\qquad (3.17)$$

The distance metric and angle are deployed to transfer the coordinate system into another space, called Hough space. A single point is transferred from (*x, y*) domain to (*a, d*) domain where *a* and *d* represent the angle and the distance. Intersect points of lines in Hough space refer to the direction of the segmented lines. The connection between the segmented lines is maintained by using the derived angle information. To illustrate the usage of HT in the proposed approach, the connection of two segmented lines and their center points ((525,575) for dark line and (575,525) for bright line) is illustrated in Figure 3.7. The interval limitation (*t*) for the tested lines is ignored in this illustration. All the possible angles with 15° intervals are calculated. The axes in the figures represent the coordinates. Virtual lines and perpendicular lines are indicated as solid and dashed lines, respectively.
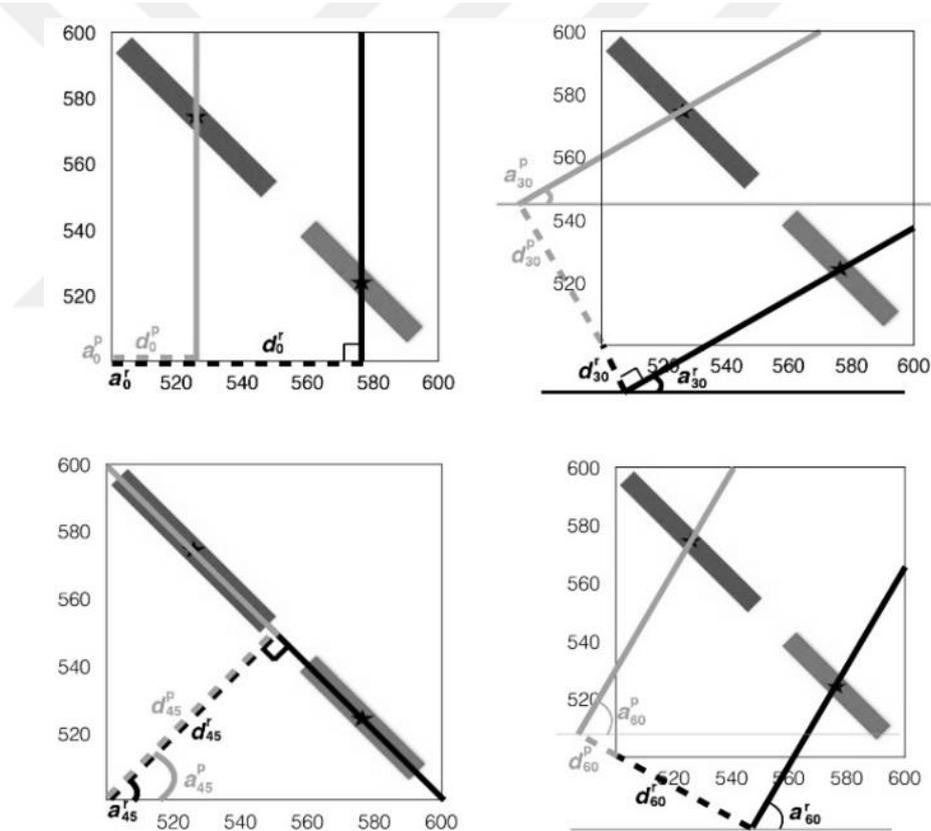
Figure 3.7 Connection of the LSD segmented lines by HT; 0°, 30°, 45°, 60° virtual (solid) and perpendicular (dashed) lines

15° intervals are arranged between the lines from 0° to 180°. Only 4 lines out of 13 angles are graphically drawn. Other results are recorded in Table 3.1 as the distance metric ($d_t^k$), calculated by using Equation (3.17).

Table 3.1 Distance metrics between the origin and intersect point of virtual and perpendicular lines calculated by Equation 3.16

| Segm. Lines (Center Cord.) | 0° | 15° | 30° | **45°** | 60° | 75° | 90° | 105° | 120° | 135° | 150° | 165° | 180° |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Red (525,575) | 525 | 656 | 742 | **778** | 760 | 691 | 575 | 420 | 235 | 35 | -167 | -358 | -525 |
| Green (575,525) | 575 | 691 | 761 | **778** | 742 | 655 | 525 | 358 | 167 | -35 | -236 | -419 | -575 |

According to Table 3.1, lines in Hough space intersect at 45°, which means the alignments of the segmented lines should be connected with a line having a 45° angle. A new line is employed from one center to another with a 45° slope to connect the segmented lines. In this example, segments are in fully linear alignment. However, in most cases, there is curvature between segments. In this case, we calculated the values by using Equation (3.17) and select the angle having the minimum differences.

We adopt HT only on segmented lines in the direction of the mean value of clusters ($\mu_c$) with ±2° intervals. Figure 3.8 indicates the case of direct usage of HT after segmentation of lines by LSD and the effects of the clustering process on the usage of HT. Normally, HT performs to form virtual lines in every angle direction to connect points. As a consequence, corners of segments will have multiple connections. In our proposed method, we restrict angles into the cluster centers with ±2° intervals which provides sharper corners of lines.



(a)                                    (b)
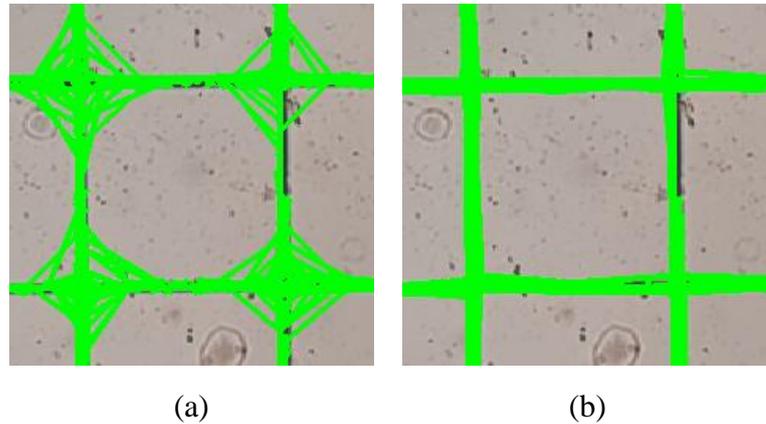
Figure 3.8 Clustering effect on the Line Connection Process: a) Direct HT on LSD segmented lines, b) HT implementation based on clustered angles to LSD segmented lines

Consequently, new lines will be partly in curvature form by the connection of linear lines. In the study, Equation (3.18) is used to reflect the aforementioned HT based processes on $k$th and $m$th *LSD* based segmented lines. $l$ refers to the number of connected line in $L_{HT}^{l}$.

$$L_{HT}^{l} = HT(L_{LSD}^{k}, L_{LSD}^{m})$$ (3.18)

### 3.3.5 Post Processing

In this study, an outline enhancing kernel, the Laplacian of Gaussian (LoG), was applied to the detected grid structure by the proposed approach [69]. The Laplacian is a 2-D isotropic measure of the 2nd spatial derivative of an image. The Laplacian of an image highlights regions of rapid intensity changes, Therefore it is often used for edge detection. The Laplacian is applied to an image that has first been smoothed with something approximating a Gaussian smoothing filter in order to reduce its sensitivity to noise.

Since the input image is represented as a set of discrete pixels, we have to find a discrete convolution kernel that can approximate the second derivatives in the definition of the Laplacian. Employed kernel and the filter effects on two example images are shown in Figure 3.9.
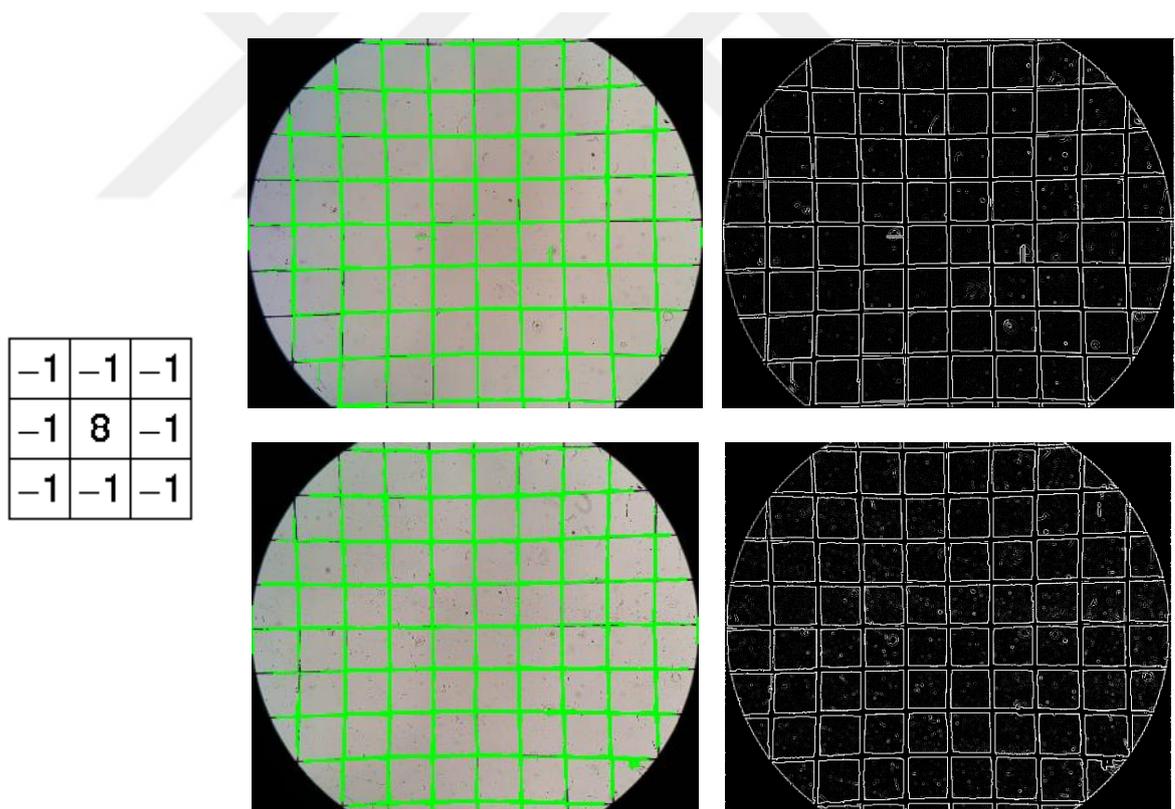
Figure 3.9 Employed Laplacian of Gaussian Kernel and the Effect of Filter on Images

## 3.4 The Evaluation of the Automatic ROI Extraction

The proposed two grid detection based ROI extraction approach was tested on 80 Makler videos obtained from 20 specimens (four videos each). Three Makler chambers were utilized in the sample acquisition step. Each specimen had different semen characteristics. Additionally, the rotation of the Makler chamber and/or microscope range might be different in the acquisition step. In this case, each video is distinctive. Videos were recorded in 1 min. durations and HD resolutions. The recording speed of the camera was 30 Hz/sec. Therefore, each video consists of 1800 frames.

F-measure and accuracy rates are considered as performance metrics for the proposed grid detection approaches in the template-matching-based evaluation. Eighty ground truth images of grids were formed manually. Each video had one ground truth for each stable Makler chamber during the recording process. Similarities between detected and ground truth lines were evaluated using the concept of Sørensen–Dice index over a generated template. Figure 3.10 shows three example Makler images with manually formed GTs as templates.
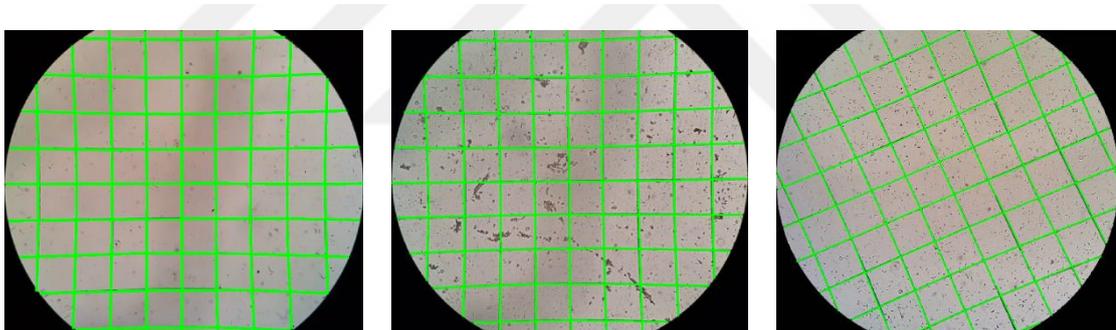


Figure 3.10 Ground Truth examples

In the testing of two approach, f-measure and accuracy were measured using a confusion matrix. In the confusion matrix, True Positive (TP) indicates the total number of pixels on detected lines matched with ground truth lines. True Negative (TN) represents the common number of pixels other than lines in both ground truth and input image. False Positive (FP) symbolizes the number of pixels detected as lines when there were no lines in a ground truth image. False Negative (FN) is the opposite form of FP in that no lines are found using the proposed approach where, in fact, there are lines in a ground truth image. (Table 3.2).

Table 3.2 Confusion Matrix for performance evaluation. The values are number of matched pixels

| | | Pixels in Ground truth | |
|---|---|---|---|
| | | Lines | Other areas |
| Pixels in output image | Detected Lines | True Positive (TP) | False Positive (FP) |
| | Other areas | False Negative (FN) | True Negative (TN) |

Since TN is misleading criteria due to the plenty of pixels outside of lines present in the images, F-measure scores should be calculated. F-measure implicates FP and FN in calculation including precision and recall terms and excluding the impact of TN. In this sense, F-measure scores will be more objective to evaluate in the proposed approach. Formulas of performance metrics, accuracy and F-Measure scores, presented in Equation 3.19 and 3.20.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{3.19}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN} \tag{3.20}$$

$$F - Measure = 2 \times \frac{Precision \, x \, Recall}{Precision + Recall}$$

Performance of the proposed two grid detection techniques were evaluated on 80 images after the post-processing step. Differences of the image pairs (ground truth and proposed second technique result) on two example images are presented in Figure 3.11. The purple color indicates the FN, while the green color represents the FP in the examples.
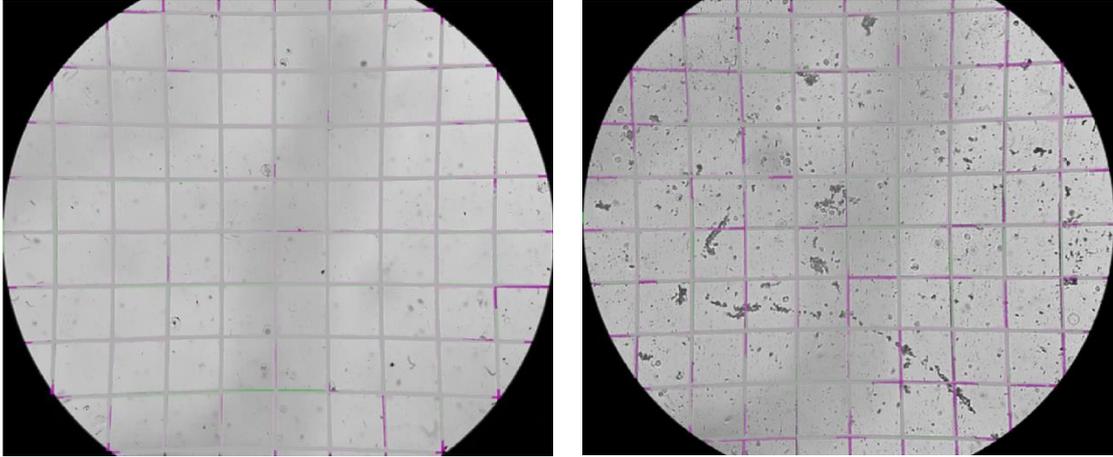
Figure 3.11 Image variation between GT and the result of Proposed Technique for two example image

Results of the template-matching evaluation with regard to performance metrics are listed in Table 3.3. 80 Makler images are divided into 6 subcategories by angle information. Average detection results are registered in the table. The number of tested images for the proposed technique is given in parentheses (Total 80). M1 represents our first approach results and M2 indicates the rates of our proposed second technique. Results prove that the second approach is significantly better than the first one which resulted in a 90.66% accuracy and a 78% F-measure rate on average, but the second combinational approach achieves 95.36% accuracy and 88.58% F-measure rates on averages. However, the accuracy rate is misleading due to too many TN values. It is also cited in literature as a data imbalance problem, or the "Accuracy Paradox". The 95.36% accuracy rate indicates that almost all pixels inside the lines and area between lines are correctly detected. However, it is misleading due to more pixels in area than lines. Hence, F-measure rates give more convincing results. An 88.58% F-measure score is accepted as real success of the presented approach for the grid detection phase in this study.

Table 3.3 Average results of template matching for the ROI detection step

| | Angles | | | | | | | | | | | |
| | 0-30 (14) | | 30-60 (16) | | 60-90 (14) | | 90-120 (15) | | 120-150 (9) | | 150-180 (12) | |
| | $M_1$ | $M_2$ | $M_1$ | $M_2$ | $M_1$ | $M_2$ | $M_1$ | $M_2$ | $M_1$ | $M_2$ | $M_1$ | $M_2$ |
| Acc. | 94 | **97.2** | 87 | **94.4** | 91 | **96.7** | 88 | **91.9** | 91 | **95.4** | 93 | **96.6** |
| F Meas. | 76 | **87.7** | 79 | **88.2** | 75 | **86.8** | 78 | **89.3** | 83 | **91.1** | 77 | **88.4** |

According to the results, we used the second approach for the grid detection process. Then, regions of interests were selected by using the detected grid structure and extracted in the next section.

## 3.5 Region of Interest (ROI) Selection and Extraction

Makler videos might be different in scale and rotation according to the settlement of the video recording device. Additionally, several distortions and/or blurring on the corners of images might occur owing to the camera lens. These challenges require selecting a common area in order to evaluate all images with the same algorithm. To this end, automatic detection of ROI is a necessary part. Therefore, as a final step, ROI are selected by logical queries depending on the detected and post-processed grid structure. Vertical and horizontal lines are individually counted and eliminated according to the following criteria:

$$
\begin{aligned}
S^v_{1...i} &= L^v_{1...i} && if \ i < 6 \\
S^h_{1...j} &= L^h_{1...j} && if \ j < 6
\end{aligned}
\tag{3.21}
$$

$$
\begin{aligned}
k &= floor(i \div 2) + 1 && if \ i \geq 6 \\
l &= floor(j \div 2) + 1 && if \ j \geq 6
\end{aligned}
\tag{3.22}
$$

$$
\begin{aligned}
S^v_{1...6} &= L^v_{k-3...k...k+2} && if \ mod_2(i) = 0 \\
S^h_{1...6} &= L^h_{l-3...l...l+2} && if \ mod_2(j) = 0 \\
S^v_{1...7} &= L^v_{k-3...k...k+3} && if \ mod_2(i) = 1 \\
S^h_{1...7} &= L^h_{l-3...l...l+3} && if \ mod_2(j) = 1
\end{aligned}
\tag{3.23}
$$

where $i$ and $j$ refer to the total number of detected lines as vertical and horizontal after aforementioned techniques, respectively. $S^v_{1...i}$ and $S^h_{1...j}$ denotes the separately selected lines as vertical and horizontal after the elimination by logical expressions. In case of detection of less than 6 lines, each line is automatically accepted as a selected line. In order to form a 6×6 or 5×5 ROI zone for images having a wider range than 6 lines, only the first 3 lines near the center of the image are selected as ROI. In this scenario, lines in the center are named as $k$ for vertical and $l$ for horizontal. Depending on whether it is an odd or even total number of detected lines, different expressions are defined. In case of singularity, the ROI zone is selected as the middle of the image having 6×6 grid, whereas 5×5 grid schema is selected for duality.
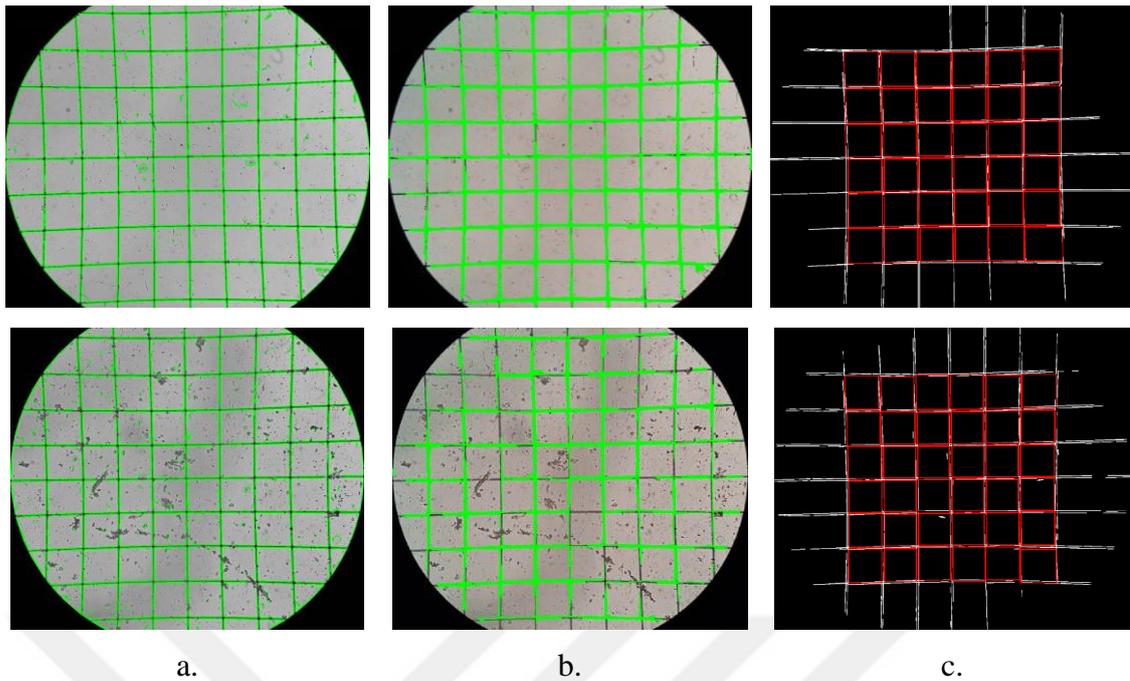
Figure 3.12 Steps of Grid and ROI detection; a)LSD results, b)Enhancement results by HT after clustering, c)Selection of regions of interest after post processing and logical operations

ROI selection approach on two different sample images is demonstrated in Figure 3.12. LSD provides a superficial line detection and segmentation, as shown in Figure 3.12(a). Next to LSD, K-Means clustering is employed over angles, and lines out of the reference limit are eliminated. HT performs over segmented lines by the direction of cluster results. It connects differently segmented lines and highlights each line, as presented in Figure 3.12(b). Post-processing and ROI selection by logical queries are then implemented. ROIs are demonstrated in Figure 3.12(c).

Figure 3.13 indicates the several ROIs of image that will be transferred to semen analysis process. Each square with few or none grid effect will be individually analyzed by Spermatozoa Detection Module after the Video Stabilization process which will be elaborated on in the next chapter.
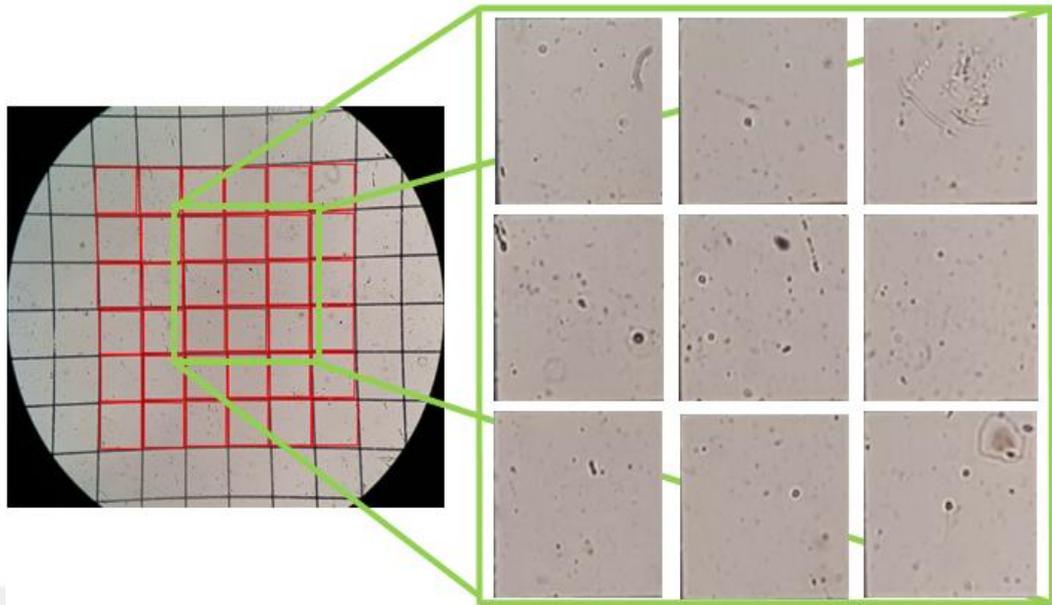
Figure 3.13 Regions of Interest transferred for cell analysis process

# VIDEO STABILIZATION

## 4.1 Introduction

Proposed data acquisition approach can be used in all laboratory without any extra device requirement. In order to provide stable videos, a mobile phone holder as stabilizer was designed and printed out by a 3D printer. Stabilizer provides easy to mount mobile phone to the ocular part; however, it is still not fully stable. Acquired videos of semen samples have vibrations. Therefore, in this thesis, video-microscopy stabilization technique is employed to fix the vibrated frame effects on motile sperm detections. Otherwise, vibration affects the detection process of the spermatozoa. Several software based stabilization studies are reviewed for this step to define the appropriate solution.

Kwon et al. proposed an adapted Kalman Filter for the video stabilization problem [70]. They mainly focused on the jitter effect in real time cameras. So, they extracted the motion vectors to estimate the jitter effects on motion. They emphasized that the technique is not only for camera vibration, but also can be employed in elimination of the effects occurred by other motile objects. Another Kalman based study has been proposed by Tiko and Vahvelian [71]. They formed a practical system constraint with respect to the amount of corrective motion. They performed Kalman based approach between constants and each video frame. Matsushita et al. proposed motion impainting technique to maintain a video stabilization problem [72]. They mainly aimed to stabilize videos in good quality. Normally, stabilized images are in low resolution due to the effect of technique. Spatial and temporal consistency of the video sequence in both static and dynamic image areas were performed in their technique. Additionally, they utilized sharpening idea by using the de-blurring algorithm as post-processing between consequent frames. As a result, their technique provided more quality stabilized videos.

Piva et al. proposed color based segmentation [73]. They extracted two one-dimensional characteristic curves to form an efficient motion estimator. Then, they analogy the consequent extracted curves to predict image displacement. As a validation test, they compared the results of proposed approach with feature based segmentation techniques. Most video stabilization techniques have complexity. The implementation of video stabilization techniques on low resource devices is important as much as its robustness. Auberger and Miro presented a stabilization technique especially for low cost devices [74]. They formed binary motion estimation on some key points in frame to decrease the complexity. Consequently, technique consumed less power than many techniques. Additionally, they adapted the technique to use in not only displacement detection, but also rotation based movement. They evaluated technique in realistic problems. Video stabilization is implemented in many different applications in literature. Liang et al. employed global feature extraction technique of the lane lines and the road vanishing points for the camcorder stabilization mounted on car [75]. They formed a system using priori information of traffic images to reduce the computational and time complexity. Each frame features were combined with one previous frame. According to results, they predicted the stable frame location. Another work made by Oreifej for the moving object detection in turbulence [76]. During object detection process, they faced with non-uniform deformation occurred by turbulence. They extracted background, turbulence and the object as the initial part. Different norms were utilized to eliminate the noises and detect the objects. Similar to [76], Smith et al. focused on the stabilization of hand-held light field video camera [77]. They implemented video camera array to stabilize video from the shaky input in static scenes by employing a space-time optimization between the virtual camera and the camera array. Evaluation has been performed by comparing their algorithm with the state-of-art stabilization software. Liu et al. proposed depth based video stabilization technique in [78]. Feature point tracking idea for video stabilization is fragile in videos having no texture, severe occlusion or camera rotation. In this sense, they employed a depth sensor camera by combining normal camera to avoid noise effects. The fusion of both camera provided much better results but it required more time to process. Battiatto et al. proposed a block matching technique for local extracted motion vectors [79]. They filtered motion vectors by similarity and matching effectiveness to speed up stabilization process. They emphasized the speed of the algorithm as the novel side of technique.

Hu et al. utilized scale invariant features (SIFT) in stabilization problem with the combination of Gaussian kernel filtering and parabolic fitting [80]. They reported that their technique is effective not only in high frequency noise motion, but also minimize the missing area as much as possible. Battiato et al. employed the same feature extraction technique in feature-based motion estimation algorithm idea [81]. They used the features in matching concept to estimate inter-frame motion. Different from direct feature matching idea, Yang et al. used particle filters in stabilization idea [82]. Particle filters were mainly employed to estimate posterior states of nonlinear systems. They roughly extracted scale invariant feature points and then estimate their next important states by particle filters. Shen et al. also utilized scale invariant features and particle filter in their work [83]. Differently, they decreased the feature size by Principle Component Analyses (PCA). Then, they composed block mean squared error metric to use in particle filter. Pinto and Anurenjan utilized another feature extraction technique in their study [84]. They used speed up robust features (SURF) in stabilization for the global motion estimation with Kalman filter as post processing to smooth the result.

In a medical application, Xia et al. used one of the stabilization techniques to clear the Rician noise effect on 3D MR data. They employed forward and inverse variance-stabilizing transformations for the Rician distribution [85]. Another stabilization study on medical imaging made by Aghajani et al [86]. They compensated for global transformation between two consecutive frames. Then, they dedicated a local deformation for stabilization. As a result, they eliminated the motion-based artifacts on iris image set [86]. Sauve et al. utilized stabilization idea on robot-assisted beating heart surgery [87]. They estimated the motion of the heart using two technique: motion estimation and texture tracking. To the end, they efficiently achieved to the estimation of heart beat and provided more stabilized robot-assisted surgery [88].

## 4.2   Feature Matching Based Software Microscope Video Stabilization

Even if a hardware stabilizer designed and mounted on the ocular, there are still vibrations occurred in the videos during the recording. In case of vibrations, motile spermatozoa counting results vary according to utilized counting algorithm. To this end, software based video stabilization is a necessity in our system. In this respect, feature-matching-based video stabilization idea is firstly utilized within the video-microscopy concept. Several feature extraction techniques were tested to sustain more stable videos

in stabilization module. Motile spermatozoa detection algorithm is then adapted to evaluate the efficiency of video stabilization and analyzing part. In case of vibrated frame sequence, the detection of total motile sperms concludes immediate peak values while it is around average values in stable frames.

Makler images have a stable grid texture in the background. Therefore, we employed feature-matching-based video stabilization technique. Speed Up Robust Features (SURF) [90], Binary Robust Invariant Scalable Key-points (BRISK) [91] and Features from Accelerated Segment Test (FAST) [92] were tested in this step to reveal the most convenient one for the implementation in video-microscopy concept, especially for the motile sperm evaluations. We used the idea of priori information image. The vibration is in minimal and non-uniform therefore matching idea is restricted with a metric similar to filtering motion vector. Then, moving pixels were clustered by morphological operations and blob analysis to evaluate the stabilization. The outline of the proposed study is shown in Figure 4.1.
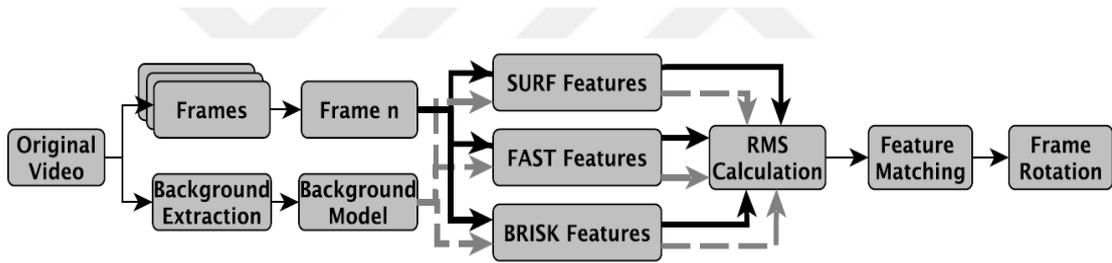


Figure 4.1 The flow chart of the Video Stabilization Methodology

Background models of the videos represent the stable states of pixels along the video sequence. In the Makler images obtained from the ocular part of microscopy as introduced in Figure 2.3, grid will be in the background because it appears in all the images. We used the same background extraction technique explained in Equation 3.1 with different parameters for the stabilization process. Videos are split into sub-videos to obtain more consistent and less noisy background models. Down-sample rate is arranged as 300 frame. As a result, six background models were extracted for the stabilization due to the 1 min. recording time and 30 Hz. sampling specification of video capturing device (1800 frames) with 300 down-sample rate. An example of extracted background model for the usage in stabilization is presented in Figure 4.2.

|     (a)     |     (b)     |
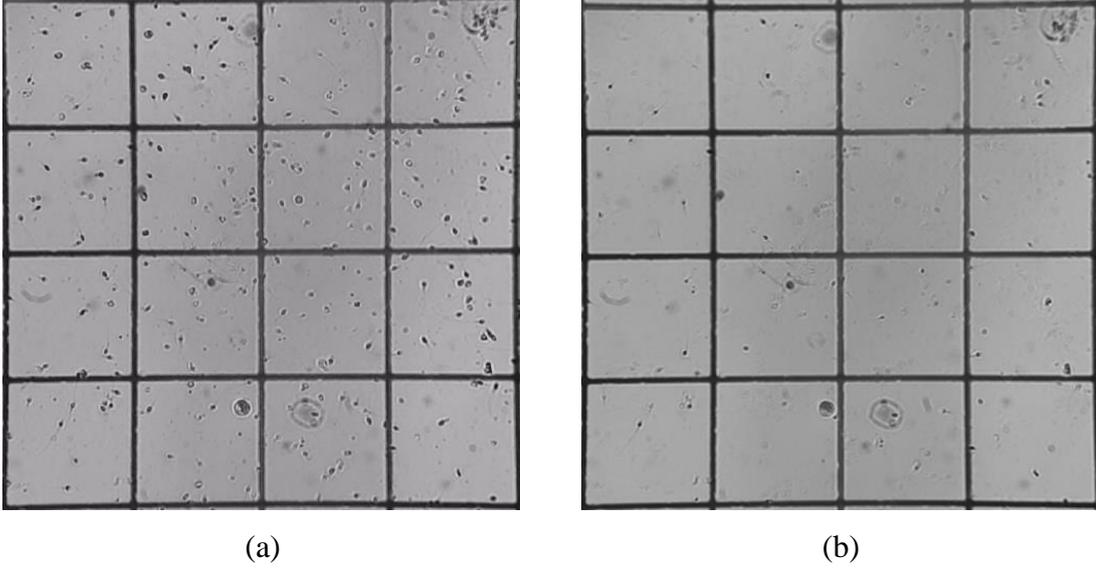
Figure 4.2 a) Original Image, b) Adaptive Mean Filter based Background Extraction

Video stabilization was, then, maintained by re-allocating each frame position ($V(x,y,t)$) in sequence to the nearest location of extracted background model ($B(x,y)^k$). Feature-matching algorithm was employed between frame $t$ and extracted relevant $k_{th}$ background model. Background model was used as the priori information. Relation between $k$ and t is defined in Equation (4.1).

$$(k-1) \times d < t < k \times d \tag{4.1}$$

SURF, BRISK and FAST are the well-known key feature descriptors. In the thesis, we evaluated the usage of each descriptors on video stabilization problem of video-microscopy to reveal the most successive one in practice. Briefly, techniques were performed once over extracted relevant $k$th background model while it was repetitively employed over each frame $t$ of the corresponding video according to Equation 4.1. Root Mean Square Error (*RMSE*) between extracted features of background model and current frame was then calculated by using Equation 4.2.

$$F_{bg}^k(a_m, s_l) = FeatureExtraction(B(x, y)^k)$$

$$F_{fr}^t(b_n, s_l) = FeatureExtraction(V(x, y, t)) \tag{4.2}$$

$$RMSE_{F_{bg}^k, F_{fr}^t} = \sum_{n=1}^{N}\sum_{m=1}^{M}\sum_{l=1}^{L}\sqrt{(F_{bg}^k(a_m, s_l) - F_{fr}^t(b_n, s_l))^2}$$

where $F_{bg}^k$ indicates the extracted features of $k$th background model. $a_m$ represents the $m$th point to be extracted having $s_l$ feature vector. Similar to background feature set,

38

$F_{fr}^t$ shows the feature sets of frame at $t$. $b_n$ is the $n$th detected point for extraction relevant features. $s_l$ will be the same dimension in extraction of the same features.
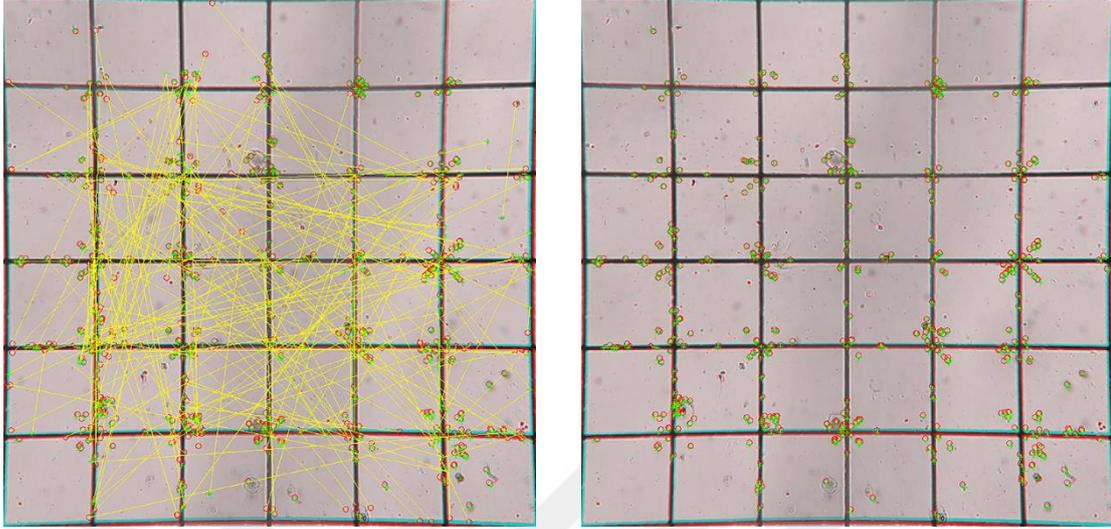


Figure 4.3 Eliminating of similar features at distant coordinates

*RMSE* was then calculated between the features derived from $k$th background and frame at $t$. In this step, a threshold level was constituted to eliminate the effect of feature similarities derived from distant located points. Only the *RMSE* of features within the predefined distance in spatial domain were calculated because of the vibration of camera accepted as minimal level but having high influence in motile sperm detection process. Figure 4.3 indicates the threshold effects on matching of derived features. Features at the nearest 50 pixel distance were used to calculate *RMSE* values.

Next to *RMSE* calculation, frames were rotated according to matching features within defined 50 pixel criteria. In case of fully matched features, rotation process will be neglected. Otherwise, each frame will be rotated to background model location. This provides more stable videos due to the fixed background extracted model.

The success of the proposed software based stabilization technique was measured by the controlling the vibrated frames. Counting of motile sperms and detection of vibrated frames were both realized based on the pixel changes within predefined duration of frame sequences. Flow chart of stabilization technique evaluation module is presented in Figure 4.4. First, foreground was extracted by frame differencing technique between background and current frame using Equation 2.1. Then, foreground extraction technique is performed for emphasizing the long term motile objects which mostly refers to the motile sperms and eliminates the short vibrations. The details of the

foreground extraction technique will be given in the motile spermatozoa detection section of this thesis.
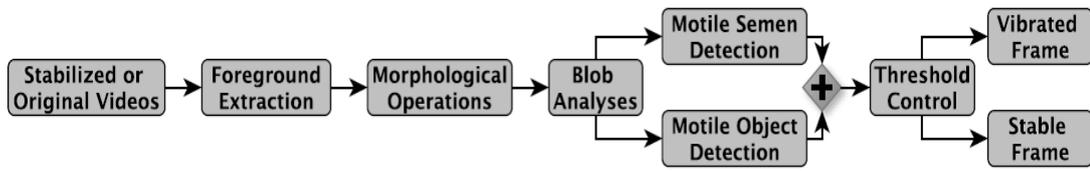


Figure 4.4 The flow chart of the evaluation for the software based video stabilization approach

In stable frames, only the sperms are motile; hence, the number of detected motile object refers to the number of motile sperm. But, vibration distort this generalization. In case of vibration, not only sperms but several grid parts are also registered as sperms due to the utilized segmentation. Even if the segmentation provides to ignore some parts of grids according to defined criteria, vibration still needs to be fixed because of the effect on detection of other sperms. An example of vibration effect on extracted foreground image is demonstrated in Figure 4.5. Each yellow mask indicates the segmentation of moving objects as sperm. White marks refer to motile objects but decided as non-sperms according to blob analysis segmentation. In case of vibration, total number of detected motile objects is increased. Duration of vibration is delineated as vibration interval. High vibration interval results in detection not only motile sperms (yellows), but other immotile objects such as leukocytes and grid parts (whites) in the frame.
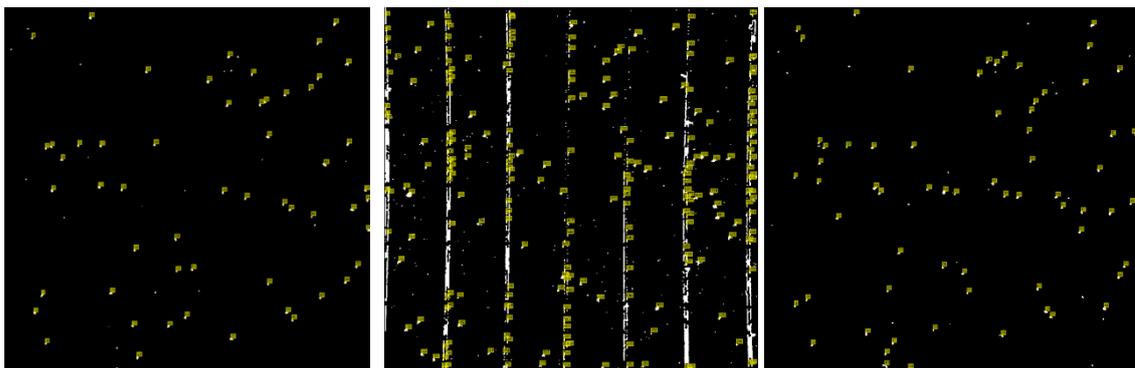


Figure 4.5 Vibration effects on foreground extracted video sequence

Performance evaluation of proposed study was made by controlling the immediate and sharply changes on the counting numbers of detected objects. According to threshold level, frames were delineated as stable or vibrated. In case of vibrations, foreground extraction technique results in remarkable high counting as motile objects due to the

grid structure of Makler material as in Figure 4.5. Otherwise, grid structure remains in background model. Motile object detection module detects only the real motile objects in sample and frame nominated as stable.

Figure 4.6 illustrates the total number of detected objects (*y*) in frames (*x*). Instant peak points drawn with green dotted lines indicate the start of vibrations. Peaks are caused by the incorrectly detection of grid structure as motile objects in vibrations. Red dotted lines indicate the ends of vibrated frames. Red line in vertical alignment represents the calculated mean values of counted motile objects in stable frames. Initially, first 40 frames are reserved for learning interval to constitute initial mean value.
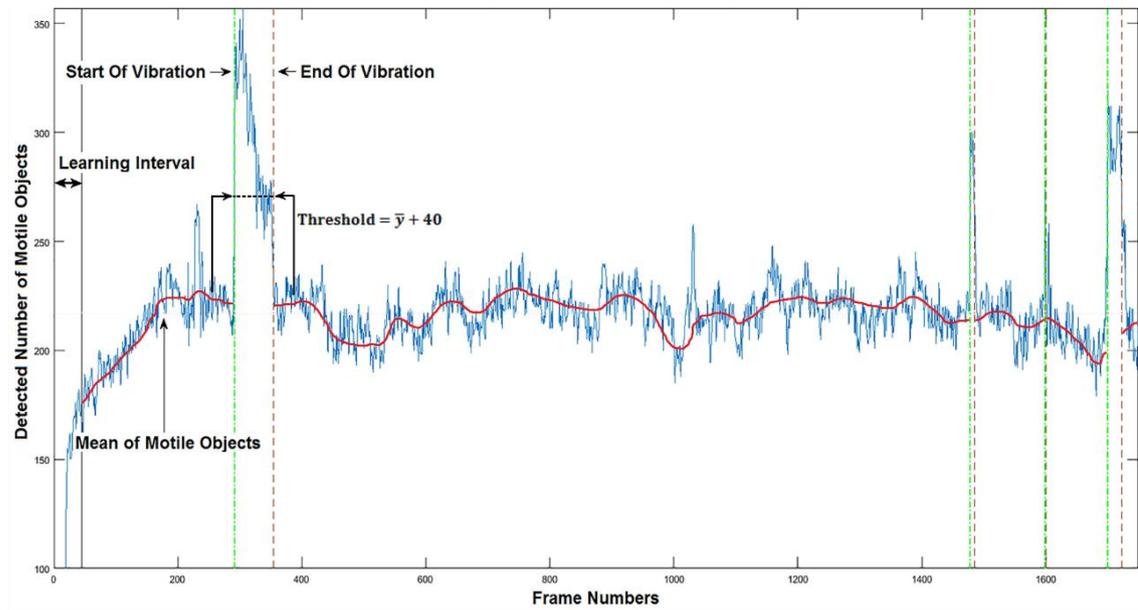


Figure 4.6 Graphically demonstration of vibrated frame detection based on motile object

After learning interval, an adaptive threshold level based on the mean values of total detected motile objects is employed. Threshold level was experimentally determined to be the mean number of motile objects + 40.

$$\bar{y} = \frac{1}{F_l} \sum_{i=1}^{F_l} y_i \tag{4.1}$$

$$\bar{y} = \bar{y} + a \times \frac{y_{F_x} - \bar{y}}{F_l} \tag{4.2}$$

where $y_i$ represents the total detected motile objects at *i*th frame within learning interval. Only the detected motile objects in stable frames were utilized in Equation 4.1 by using constant *a*. Otherwise, vibration control technique based on the criteria results

41

in incorrect detection. Therefore, mean calculation was ignored during the vibration interval. Additionally, in every stable frame, learning frame number ($F_l$) was increased to use updated version in Equation 4.2. Output images for the each main steps in methodology is given on an example image in Figure 4.7.
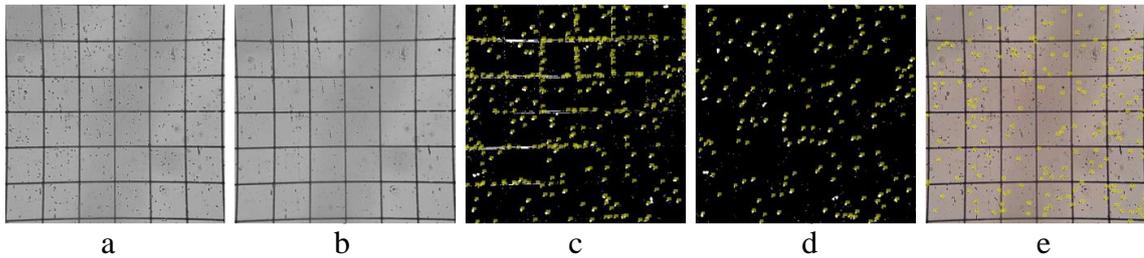


Figure 4.7 Output images of the main steps in methodology: a) Original Image, b) Background Model, c) Foreground Model without Stabilization Process during the vibration, d) SURF based stabilization version of the same image sequence demonstrated in c. e) Counting of segmented sperm cells by using the connected component analysis

## 4.3 The Evaluation of the Video-Microscopy Stabilization Approach

A total of 42 Makler videos of 14 subjects (3 videos per subject) were captured for the evaluation of the proposed software based video stabilization approach. Videos have different vibration intervals that occur during the acquisition step. Additionally, each subject has distinctive number of motile spermatozoa due to the different subject characteristics and disease. Samples were grouped into 3 categories as normospermic, azospermic and oligospermic by an expert in laboratory with visual assessment technique. Motile sperms in 10 square of Makler chamber refer to million in per ml according to Makler instruction manual for visual assessment technique. However, it is impossible to observe all sperms due to the presence of excess number of sperms and limitation of human eye in normospermic case. Therefore, a standard dilution process by using formal saline (SC) was performed. The dilution was adjusted appropriately according to the estimated concentration as 1:1 (sample / dilution) proportion. 7 replicate counts performed on each subjects.

Initially, first module of this study was performed on the original videos with feature extraction techniques. We obtained differently stabilized videos by the matching of corresponding features within the RMSE constraints. Outputs were named with the relevant feature extraction technique as SURF, BRISK and FAST videos. Next, motile

objects and vibrated frame detection module was performed over four type of videos of each sample.

Stabilization success of each descriptor is evaluated by using the detection of immediate peak points of the foreground motile sperm detection. Total vibrated frames of each video sequences are given in Table 4.1. After the detection of extremely high peak of motile pixels, each frame is counted as vibrated frame until the end of vibration, which is the decrements of the detected motile pixels to the level, measured before vibration. Red, purple and cyan represents the lowest vibration detection videos measured during the video 1, 2 and 3 corresponding to each patients, respectively.

According to Table 4.1, best stabilization is obtained by the SURF descriptors. It is measured as minimum vibration on the 37 videos obtained by the SURF feature matching based stabilization technique. FAST results are comparable to the SURF results. Both methods are rotation invariant. Therefore, results are similar. Only the vibrations in 3 videos were measured lower than SURF based stabilization. However, FAST is computationally more efficient than the SURF as illustrated in Table 4.3. BRISK features are worst descriptors in terms of stabilization idea according to presented results. It distorted the entire sequence and increased undesirable vibration effects even if the minimal vibrated frames detected in original raw videos. Main reason behind this problem is the usage of brightness of images, because BRISK calculates different RMSE and matches various points in once. Several graphs belongs to the SURF results in Table 4.1 are demonstrated in Figures 4.8, 4.9, and 4.10 to better observe the effect of SURF descriptor based stabilization on original videos. Red lines are drawn based on the detected motile sperms in original videos while the green lines indicate the stabilized video analysis.

Table 4.1 Total Number of Vibrated Frame Numbers in Sequence

| | Frame | Normospermia | | | | | | | | Oligospermia | | | | Azoospermia | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Countings | Subj1 | Subj2 | Subj3 | Subj4 | Subj5 | Subj6 | Subj7 | Subj8 | Subj1 | Subj2 | Subj3 | Subj4 | Subj1 | Subj2 |
| Original | Vid1 | 53 | 24 | 6 | 28 | 280 | 35 | 132 | 188 | 282 | 88 | 34 | 8 | 48 | 28 |
| | Vid2 | 105 | 50 | 8 | 150 | 47 | 457 | 52 | 114 | 157 | 163 | 60 | 57 | 257 | 57 |
| | Vid3 | 39 | 9 | 26 | 66 | 364 | 188 | 54 | 87 | 105 | 238 | 84 | 16 | 173 | 45 |
| | Minimum | 39 | 9 | 6 | 28 | 47 | 35 | 52 | 87 | 105 | 88 | 34 | 8 | 48 | 28 |
| SURF | Vid1 | 12 | 5 | 1 | 12 | 92 | 4 | 3 | 18 | 12 | 21 | 3 | 18 | 25 | 0 |
| | Vid2 | 150 | 43 | 0 | 9 | 0 | 128 | 2 | 18 | 8 | 4 | 7 | 5 | 88 | 0 |
| | Vid3 | 10 | 7 | 3 | 15 | 16 | 39 | 38 | 12 | 13 | 55 | 11 | 0 | 34 | 0 |
| | Minimum | 10 | 5 | 0 | 9 | 0 | 4 | 2 | 12 | 8 | 4 | 3 | 0 | 25 | 0 |
| BRISK | Vid1 | 63 | 71 | 114 | 107 | 100 | 54 | 106 | 148 | 101 | 77 | 74 | 71 | 174 | 12 |
| | Vid2 | 241 | 232 | 138 | 189 | 8 | 170 | 14 | 89 | 14 | 27 | 389 | 23 | 146 | 6 |
| | Vid3 | 80 | 133 | 396 | 131 | 30 | 127 | 139 | 215 | 24 | 101 | 138 | 1 | 35 | 11 |
| | Minimum | 63 | 71 | 114 | 107 | 8 | 54 | 14 | 89 | 14 | 24 | 74 | 1 | 35 | 6 |
| FAST | Vid1 | 32 | 0 | 34 | 15 | 191 | 19 | 94 | 61 | 82 | 106 | 23 | 182 | 43 | 2 |
| | Vid2 | 436 | 1398 | 35 | 190 | 47 | 124 | 29 | 36 | 68 | 6 | 59 | 54 | 299 | 3 |
| | Vid3 | 432 | 21 | 18 | 65 | 181 | 61 | 72 | 112 | 33 | 40 | 41 | 4 | 40 | 7 |
| | Minimum | 32 | 0 | 18 | 15 | 47 | 19 | 29 | 36 | 33 | 6 | 23 | 4 | 40 | 2 |

44

Figure 4.8 were derived from the sperm counting results of the each frame sequence for Subject 2 and 7, respectively. It can be observed from Figure 4.8 (a) that mostly vibrated frames, nominated as the high peak points of red line, are aligned back to the mean point of sequence. In another normospermic case, Subject 7, resulted in the successful correction of red lines similar to Subject 2. However, descriptor generated vibration effects which are not in the original video as a result of incorrect feature matching. These vibrations are illustrated in Figure 4.8 (b). Despite these adverse effects of SURF descriptor, vibrations were decreased to 5 and 3 frame from 24 and 132 for Subjects 2 and 7, respectively (Table 4.1).



(a) Subject 2                        (b) Subject 7
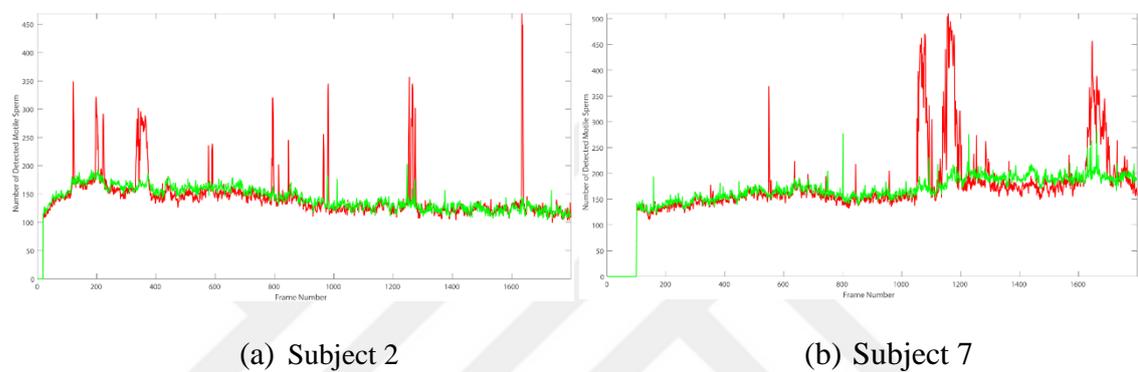
Figure 4.8 Vibration effects on foreground extracted video sequence (Normospermic Case)

Other two subjects are presented in Figure 4.9 as an example of oligospermic case. Several vibrations were stabilized even if it was extremely high as in Figure 4.9 (a), but multiple immediate peaks on the detection of motile sperm occurred as a result of incorrect feature matching. While the final vibration count was determined as 8 and 5 in stabilized SURF videos, they were measured as 157 and 57 in original videos for Subjects 1 and 4, individually.



(a) Subject 1                        (b) Subject 4
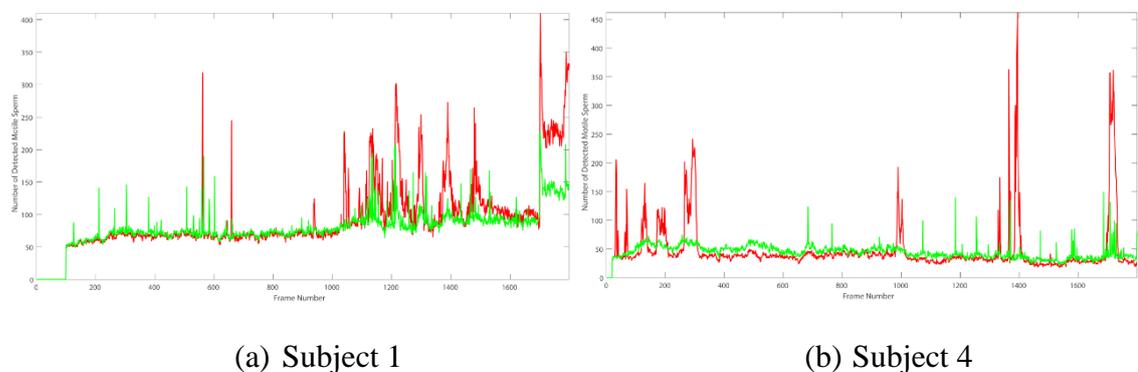
Figure 4.9 Vibration effects on foreground extracted video sequence (Oligospermic Case)

45

Figure 4.10 (a) represents the motile sperm detection process for the Subject 1 in Azospermic case. Azoospermia is the disease as non-sperm presence in semen. However, in vibration, excess number of sperm was measured due to the detection of grid structure as motile sperm. Out of vibrations, presented algorithm resulted in non-zero values. Stabilization by using the SURF features have high impact on Subject 1 as it can be seen on Figure 4.10 (a). 169 frames out of 257 were stabilized by the proposed technique. Motile sperm detection plot of Subject 5 in normospermic case is demonstrated in Figure 4.10 (b) as the most successfully alignment in this study. Vibrations were observed 364 frames (~12 sec.) in original video, but proposed stabilization technique with SURF descriptor reduced the number of vibrated frames to the 16 (~0.4 sec) frame.



(a)  Subject 1 - Azospermic Case          (b) Subject 5 - Normospermic Case
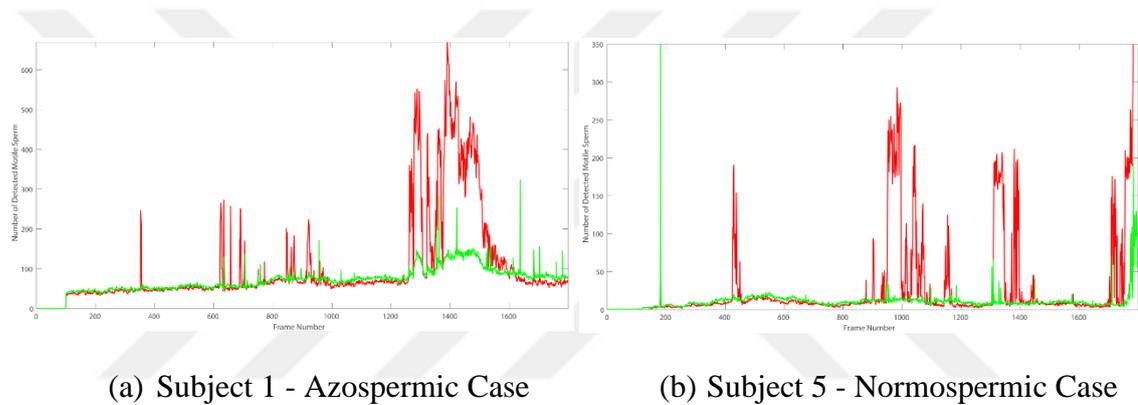
Figure 4.10 Vibration effects on foreground extracted video sequence

In data acquisition step using the ocular part, vibration is unavoidable. Therefore, proposed stabilization technique will improve the sperm detection. In order to see the effects of vibrations, motile sperm counting results in stable frames ($F_s$) and all video sequence were separately measured. Differences of results counted in all sequences and only in stable frames give information about the stability of video. Mean of detected motile objects in stable and all frames are expected to be the same in no vibration case. Otherwise, mean values of all sequences are found to be more than the no vibration case due to the incorrect detection of motile pixels remained in the foreground.

Normally, motile objects should correspond to the motile sperms because the only motile objects in samples are sperms. However, the number of detected motile pixels becomes higher than the stable consequent frames in case of vibration during the sample acquisition step according to employed motile sperm detection technique. Therefore, vibrated frames affect the motile sperm counting analysis. In this respect, the average number of detected motile sperms out of vibrated frames (Stable Frames) was

46

crosschecked to the average value of entire video sequence. Differences are registered in Tables 4.2 and 4.3 as error *e.* In the counting process, the standardization of Makler chamber as in visual assessment technique was employed for 36 square due to the range of presented data acquisition approach. It is better to utilize 36 square instead of 10 square because more field of observation gives more accurate generalization. In visual assessment technique, it is impossible to observe 36 squares at the same time manually, hence, 10 squares are dedicated for human based systems while it is easy to process by computer to analyze wide range. Therefore, counting results within 36 squares were generalized to per ml and then registered in tables.

Convergence between mean values of motile sperm numbers counted in stable and all frames indicates the less vibration status. Otherwise, vibration will cause the more difference in result. In this sense, videos having minimum *e* values should be maintained as the final result in counting process. SURF descriptor based stabilization gave the minimum *e* rate. According to the counting results using BRISK based stabilization, vibration effects in videos got worse than other stabilized and original videos in several subjects.

In order to evaluate the efficiency of proposed computer based approach in terms of only motile sperm counting aspects, another comparison was performed with the visual assessment results. Table 4.4 shows the Visual Assessment (VA) and computer based (CB) counting of SURF stabilized videos results.

It is not expected that the results will be exactly similar between computer based results and visual assessment technique, not only due to the need for software-based developments such as extra segmentation or noise removal filters in computer side, but also due to observer variability. Counting results signify the efficiency of proposed approach in spermiogram tests especially for the normospermic and oligospermic cases since the differences can be ignored when considering the very high number of sperm counts. However, in azospermic cases, system failed due to the segmentation step. It should be measured as 0 sperm numbers, but counting process concluded with 1.5 and 1.3, respectively.

The last metric is the processing times of the descriptors in stabilization module. This metric was evaluated in terms of the applicability of the methods to real-time systems. Descriptor with short processing times will be more effective in establishing a real-time

diagnostic system in order to provide the tests faster. Table 4.5 indicates the time requirement for the entire process of module 1 aforementioned in this study.

Table 4.2 Total Detected Motile Sperms for Normospermic Cases

| Motile Sperms (millions/ml) | | Subj1 | | | Subj2 | | | Subj3 | | | Subj4 | | | Subj5 | | | Subj6 | | | Subj7 | | | Subj8 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $y_{Fv+Fs}$ | $y_{Fs}$ | e | $y_{Fv+Fs}$ | $y_{Fs}$ | e | $y_{Fv+Fs}$ | $y_{Fs}$ | e | $y_{Fv+Fs}$ | $y_{Fs}$ | e | $y_{Fv+Fs}$ | $y_{Fs}$ | e | $y_{Fv+Fs}$ | $y_{Fs}$ | e | $y_{Fv+Fs}$ | $y_{Fs}$ | e | $y_{Fv+Fs}$ | $y_{Fs}$ | e |
| Original | Vid1 | 63.6 | 62.8 | 0.8 | 47.2 | 46.9 | 0.3 | 59.4 | 59.4 | 0 | 49.4 | 48.6 | 0.8 | 35 | 18.9 | 16.1 | 50.3 | 49.4 | 0.9 | 53.1 | 51.4 | 1.7 | 55.6 | 54.2 | 1.4 |
| | Vid2 | 66.7 | 65 | **1.7** | 59.2 | 58.1 | 1.1 | 56.7 | 56.7 | 0 | 48.6 | 43.9 | 4.7 | 22.8 | 21.4 | 1.4 | 79.2 | 54.4 | 24.8 | 53.3 | 51.9 | 1.4 | 56.1 | 52.8 | 3.3 |
| | Vid3 | 71.4 | 70.8 | 0.6 | 53.1 | 53.1 | 0 | 54.7 | 54.4 | 0.3 | 51.1 | 48.3 | 2.8 | 55 | 32.8 | 22.2 | 52.2 | 45 | 7.2 | 53.3 | 51.9 | 1.4 | 60.6 | 58.9 | 1.7 |
| | Average | 67.2 | 66.2 | 1 | 53.2 | 52.7 | 0.5 | 56.9 | 56.8 | 0.1 | 49.7 | 46.9 | 2.8 | 37.6 | 24.4 | 13.2 | 60.6 | 49.6 | 11 | 53.2 | 51.7 | 1.5 | 57.4 | 55.3 | 2.1 |
| SURF | Vid1 | 69.2 | 69.2 | **0** | 49.7 | 49.7 | **0** | 59.4 | 59.4 | **0** | 51.1 | 50.8 | **0.3** | 27.2 | 25.6 | **1.6** | 65.6 | 65.6 | **0** | 54.4 | 54.4 | **0** | 54.4 | 54.4 | **0** |
| | Vid2 | 71.4 | 69.2 | 2.2 | 63.1 | 62.5 | **0.6** | 56.9 | 56.9 | **0** | 46.1 | 46.1 | **0** | 25.8 | 25.8 | **0** | 72.8 | 69.7 | 3.1 | 57.8 | 57.8 | **0** | 55.8 | 55.6 | **0.2** |
| | Vid3 | 71.6 | 71.6 | **0** | 56.9 | 56.9 | **0** | 56.1 | 56.1 | **0** | 52.8 | 50.3 | 2.5 | 25.6 | 25.6 | **0** | 62.8 | 62.8 | **0** | 51.4 | 51.4 | **0** | 61.7 | 61.7 | **0** |
| | Average | 70.7 | 70 | **0.7** | 56.6 | 56.4 | **0.2** | 57.5 | 57.5 | **0** | 50 | 49.1 | **0.9** | 26.2 | 25.7 | **0.5** | 67.1 | 66.0 | **1** | 54.5 | 54.5 | **0** | 57.3 | 57.2 | **0.1** |
| BRISK | Vid1 | 65.8 | 65 | 0.8 | 48.3 | 47.2 | 1.1 | 48.6 | 46.1 | 2.5 | 51.7 | 49.7 | 2 | 26.7 | 24.7 | 2 | 65.6 | 64.7 | 0.9 | 56.7 | 55 | 1.7 | 54.4 | 52.2 | 2.2 |
| | Vid2 | 65.8 | 60 | 5.8 | 52.8 | 46.4 | 6.4 | 50.3 | 46.9 | 3.4 | 47.8 | 44.2 | 3.6 | 25.8 | 25.8 | 0 | 71.1 | 67.5 | 3.6 | 57.5 | 57.5 | 0 | 54.4 | 53.1 | 1.3 |
| | Vid3 | 73.6 | 72.8 | 0.8 | 56.4 | 54.2 | 2.2 | 47.8 | 31.1 | 16.7 | 51.4 | 45.3 | 6.1 | 25.3 | 25 | 0.3 | 60.8 | 58.3 | 2.5 | 50.3 | 46.1 | 4.2 | 63.3 | 59.2 | 4.1 |
| | Average | 68.4 | 65.9 | 2.5 | 52.5 | 49.3 | 3.2 | 48.9 | 41.4 | 7.5 | 50.3 | 46.4 | 3.9 | 25.9 | 25.2 | 0.8 | 65.8 | 63.5 | 2.3 | 54.8 | 52.9 | 2 | 57.4 | 54.8 | 2.5 |
| FAST | Vid1 | 67.5 | 66.9 | 0.4 | 47.2 | 47.2 | 0 | 58.6 | 58.3 | 0.3 | 49.4 | 49.2 | 0.2 | 28.9 | 23.3 | 5.6 | 62.5 | 62.5 | 0 | 53.9 | 52.2 | 1.7 | 54.4 | 53.6 | 0.8 |
| | Vid2 | 62.2 | 60 | 2.2 | 52.5 | 43.1 | 9.4 | 59.2 | 58.9 | 0.3 | 48.3 | 43.1 | 5.2 | 23.3 | 22.8 | 0.5 | 70.8 | 67.8 | **3** | 55.3 | 55 | 0.3 | 54.4 | 53.9 | 0.5 |
| | Vid3 | 72.2 | 60.3 | 11.9 | 55.3 | 54.7 | 0.6 | 51.7 | 51.7 | 0 | 50.6 | 48.9 | **1.7** | 29.7 | 25.3 | 4.4 | 60 | 58.9 | 1.1 | 51.7 | 50.8 | 0.9 | 62.8 | 59.4 | 3.4 |
| | Average | 67.3 | 62.4 | 4.8 | 51.7 | 48.3 | 3.3 | 56.5 | 56.3 | 0.2 | 49.4 | 47.1 | 2.4 | 27.3 | 23.8 | 3.5 | 64.4 | 63.1 | 1.4 | 53.6 | 52.7 | 1 | 57.2 | 55.6 | 1.6 |

$y_{Fv+Fs}$ = In all frame set (vibrated + stable)    $y_{Fs}$ = Only in stable frames    e = error

Table 4.3 Total Detected Motile Sperms for Oligospermic and Azospermic Cases

| Motile Sperms (millions/ml) | | Oligospermic | | | | | | | | | | | | | | Azoospermic | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Subj1 | | | Subj2 | | | Subj3 | | | Subj4 | | | Subj1 | | | Subj2 | | |
| | | $y_{F_v+F_s}$ | $y_{F_s}$ | e | $y_{F_v+F_s}$ | $y_{F_s}$ | e | $y_{F_v+F_s}$ | $y_{F_s}$ | e | $y_{F_v+F_s}$ | $y_{F_s}$ | e | $y_{F_v+F_s}$ | $y_{F_s}$ | e | $y_{F_v+F_s}$ | $y_{F_s}$ | e |
| Original | Vid1 | 26.4 | 15.6 | 10.8 | 5.3 | 2.8 | 2.5 | 3.6 | 3.1 | 0.6 | 16.9 | 16.9 | 0.0 | 1.7 | 0.6 | 1.1 | 1.7 | 1.1 | 0.6 |
| | Vid2 | 19.4 | 15.8 | 3.6 | 6.4 | 2.5 | 3.9 | 3.9 | 2.5 | 1.4 | 19.4 | 18.6 | 0.8 | 7.2 | 1.1 | 6.1 | 2.5 | 1.4 | 1.1 |
| | Vid3 | 16.9 | 14.2 | 2.8 | 3.9 | 3.1 | 0.8 | 4.7 | 2.8 | 1.9 | 16.7 | 16.4 | 0.3 | 9.4 | 0.6 | 8.9 | 2.2 | 1.1 | 1.1 |
| | Average | 20.9 | 15.2 | 5.7 | 5.2 | 2.8 | 2.4 | 4.1 | 2.8 | 1.3 | 17.7 | 17.3 | 0.4 | 6.1 | 0.7 | 5.4 | 2.1 | 1.2 | 0.9 |
| SURF | Vid1 | 18.9 | 18.6 | **0.3** | 2.2 | 1.9 | **0.3** | 3.3 | 3.3 | **0.0** | 18.9 | 18.9 | **0.0** | 0.8 | 0.6 | **0.3** | 0.8 | 0.8 | **0.0** |
| | Vid2 | 16.7 | 16.1 | **0.6** | 2.5 | 2.5 | **0.0** | 2.8 | 2.8 | **0.0** | 18.3 | 18.3 | **0.0** | 2.5 | 1.1 | **1.4** | 1.4 | 1.4 | **0.0** |
| | Vid3 | 16.4 | 16.4 | **0.0** | 3.1 | 3.1 | **0.0** | 3.3 | 3.3 | **0.0** | 18.9 | 18.9 | **0.0** | 1.1 | 0.8 | **0.3** | 1.7 | 1.7 | **0.0** |
| | Average | 17.3 | 17.0 | **0.3** | 2.6 | 2.5 | **0.1** | 3.1 | 3.1 | **0.0** | 18.7 | 18.7 | **0.0** | 1.5 | 0.8 | **0.6** | 1.3 | 1.3 | **0.0** |
| BRISK | Vid1 | 20.6 | 18.3 | 2.2 | 3.3 | 1.9 | 1.4 | 4.4 | 3.6 | 0.8 | 19.2 | 18.1 | 1.1 | 4.7 | 1.4 | 3.3 | 1.1 | 1.1 | 0.0 |
| | Vid2 | 17.2 | 15.3 | 1.9 | 2.8 | 2.5 | 0.3 | 15 | 5.8 | 9.2 | 18.3 | 18.1 | 0.3 | 3.9 | 1.1 | 2.8 | 1.4 | 1.4 | 0.0 |
| | Vid3 | 16.4 | 16.1 | 0.3 | 4.2 | 2.5 | 1.7 | 5.8 | 3.1 | 2.8 | 20.6 | 20.6 | 0.0 | 1.1 | 0.8 | 0.3 | 1.9 | 1.7 | 0.2 |
| | Average | 18.1 | 16.6 | 1.5 | 3.4 | 2.3 | 1.1 | 8.4 | 4.2 | 4.3 | 19.4 | 18.9 | 0.5 | 3.2 | 1.1 | 2.1 | 1.5 | 1.4 | 0.1 |
| FAST | Vid1 | 20.3 | 18.3 | 1,9 | 5 | 1.9 | 3.1 | 3.6 | 3.3 | 0.3 | 19.2 | 15.8 | 3.3 | 1.1 | 0.6 | 0.6 | 0.8 | 0.8 | 0.0 |
| | Vid2 | 22.5 | 14.2 | 8.3 | 2.2 | 2.2 | 0.0 | 4.2 | 2.5 | 1.7 | 16.7 | 16.4 | 0.3 | 8.1 | 0.8 | 7.2 | 1.4 | 1.4 | 0.0 |
| | Vid3 | 16.4 | 15.6 | 0.8 | 3.3 | 2.2 | 1.1 | 3.6 | 3.1 | 0.6 | 18.9 | 18.9 | 0.0 | 1.4 | 0.8 | 0.6 | 1.9 | 1.7 | 0.2 |
| | Average | 19.7 | 16.0 | 3.7 | 3.5 | 2.1 | 1.4 | 3.8 | 3.0 | 0.8 | 18.2 | 17.0 | 1.2 | 3.5 | 0.7 | 2.8 | 1.4 | 1.3 | 0.1 |

$y_{F_v+F_s}$ = In all frame set (vibrated + stable)    $y_{F_s}$ = Only in stable frames    e = error

Table 4.4 Counting results by an expert with visual assessment technique

| million/ml | Normospermic | | | | | | | | | | | | | | | | | | | | | | Oligospermic | | | | | | | | Azoospermic | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Subj1 | | Subj2 | | Subj3 | | Subj4 | | Subj5 | | Subj6 | | Subj7 | | Subj8 | | Subj1 | | Subj2 | | Subj3 | | Subj4 | | Subj1 | | Subj2 | |
| | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB | VA | CB |
| Video1 | 65 | 69.2 | 52 | 49.7 | 56 | 59.4 | 47 | 51.1 | 24 | 27.2 | 63 | 65.6 | 52 | 54.4 | 47 | 54.4 | 14 | 18.9 | 3 | 2.2 | 3 | 3.3 | 18 | 18.9 | 0 | 0.8 | 0 | 0.8 |
| Video2 | 67 | 71.4 | 59 | 63.1 | 52 | 56.9 | 41 | 46.1 | 21 | 25.8 | 62 | 72.8 | 47 | 57.8 | 51 | 55.8 | 15 | 16.7 | 3 | 2.5 | 4 | 2.8 | 16 | 18.3 | 0 | 2.5 | 0 | 1.4 |
| Video3 | 68 | 71.6 | 62 | 56.9 | 51 | 56.1 | 48 | 52.8 | 24 | 25.6 | 60 | 62.8 | 52 | 51.4 | 64 | 61.7 | 17 | 16.4 | 3 | 3.1 | 3 | 3.3 | 15 | 18.9 | 0 | 1.1 | 0 | 1.7 |
| Average | 66.6 | 70.7 | 57.7 | 56.6 | 53 | 57.5 | 45.3 | 50 | 23 | 26.2 | 61.7 | 67.1 | 50.3 | 54.5 | 54 | 57.3 | 15.3 | 17.3 | 3 | 2.6 | 3.3 | 3.1 | 18.7 | 19.4 | 0 | 1.5 | 0 | 1.3 |
| Diff | 4.1 | | 1.1 | | 4.5 | | 4.7 | | 3.2 | | 5.4 | | 4.2 | | 3.3 | | 2 | | 0.4 | | 0.2 | | 0.7 | | 1.5 | | 1.3 | |

VA = Visual Assessment          CB = Computer Based (Proposed Approach)          Diff = Difference between methods

Table 4.5 Processing Times for the Video Stabilization Module

| ms. | Normospermic | Oligospermic | Azoospermic |
|---|---|---|---|
| SURF | 5524 | 4785 | 3448 |
| BRISK | 4023 | 3669 | 2704 |
| FAST | 2254 | 1957 | 782 |

# SPERMATOZOA DETECTION AND COUNTING

## 5.1 Introduction

The most important part of this thesis is the determination of the total number of sperm in a given sample - in other words, sperm concentration. Infertility is highly related with the sperm numbers, hence, counting process is crucial. Infertility zone is described in the manual published by WHO as the presence of the immotile detected sperm number as more than 70% of all sperms [3]. In this sense, correctly identified of motile sperm number mostly reveals the possible infertility disease.

The detection of sperms is the first phase of the counting process. Olalla et al. used Otsu's thresholding and wavelet transform for sperm detection in their study [14]. As last step, they performed SVM classifier to segment the acrosomes of sperm as intact or damaged. In another studies published by Alegre et al. [15, 16], segmentation was performed using the same method but classification was realized based on different feature extraction techniques. Discrete Wavelet Transform (DWT) was employed to extract Haralick Features [15] and Contour based features [16]. Features were classified by different methods to analyze the acrosome states.

In order to segment exact sperm size instead of only acrosome classification, Khachane et al. proposed a fuzzy rule based classification [17]. They segmented sperm into regions as head, mid-piece and tail. Mainly, spatial features such as major and minor axis, areas, and perimeters of sperm regions were used in classification. Authors defined different logical expressions in terms of fuzzy meaning. Another study, Liu et al. [18] mostly focuses on independently tracking of sperm head and tail. This study reveals the correlation between tail beating amplitude and head motility. Authors subtracted consecutive two frames to extract motile spermatozoon, and then used Kalman Filter to

track. More detailed study of sperm head detection was published by Chang et al [19]. They presented an improved two-stage framework for not only acrosome but also nucleus detection of sperm. They performed a combination of K-Means clustering and histogram analysis. Additionally, they utilized different color spaces into clustering. Anbumozhi employed fuzzy logic into the neural network classification for brain tumors [93]. Eltoukhy et al. implemented watershed segmentation to extract ROI for breast cancer [94]. Then, they performed wavelet transforms to extract features. SVM classification, regression trees and rule classifiers were used in classification step. Mageshwari et al. used morphological operations to segment HIV cells [95]. Another cell segmentation was performed by Al-Dulaimi et al. adapting "active contours" in the segmentation [96]. Bijar et al. segmented the sperm cell into acrosome, nucleus, and mid-piece parts [97]. They used a Bayesian classifier to utilize the adaptive mixtures method and Markov random field model. In another study, sperm-based segmentation was performed by Ilhan and Elbir [98]. They tested descriptor effects on the determination of sperm cell obtained from the embedded microscope camera.

Microscopic images are mostly noisy images due to optical effects. Hence, several preprocesses should be applied. Abbiramy et al. [20] used a sequentially combination of Laplace and Median Filter, Li et al. [21] adopted 2D gauss filter, and Karabiber et al. [22] performed a low pass filter over images in their sperm counting approaches. A different study based on fluorescence dying technique was published in [26]. Authors simply performed morphological operations such as dilation and erosion to erase irrelevant objects. Baâzaoui et al. [99] enhanced the images by entropy based fuzzy region growing techniques to minimize optical effects and provide semi-automated segmentation.

In the thesis, spermatozoa detection was performed over the extracted and stabilized ROIs in two categories; Motile and Immotile Spermatozoa regarding to the abovementioned studies.

## 5.2 Immotile Spermatozoa Detection

Detection of immotile objects was performed over the background models which have been extracted in ROI extraction step. Mean of the pixels alongside the defined frame set will reflect of stable objects. Equation 3.1 is performed to obtain the background model. 10 seconds of videos (300 frames) was arranged as down sampling interval to

skip 300 frames after each processed frame. Basically, down sampling provides to ignore the motile sperm by detection of altered pixels within defined time interval. In other words, observing constant pixels of images over down sampling interval gives the detection of the steady object. Each ROI has three background models *(k)* in immotile spermatozoon detection due to the arrangement of sub-sample size (900 frame). Examples of background extraction of two squares of ROI are shown in Figure 5.1. Immotile spermatozoa and other immotile objects remained in background model while motile spermatozoa are removed.

Sample Image 1    Background 1    Sample Image 2    Background 2
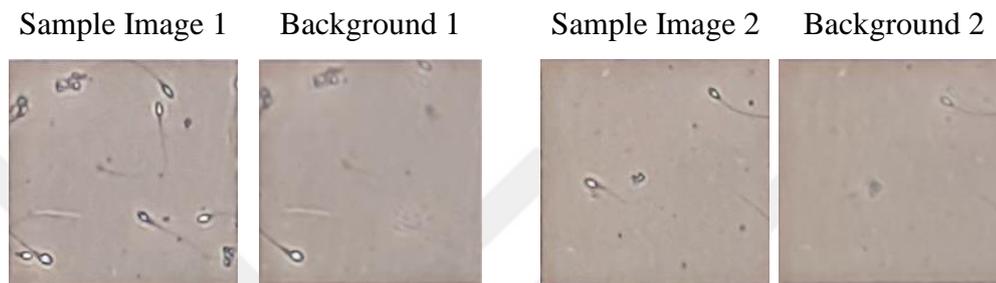


Figure 5.1 Adaptive Mean Filter based Background Extractions

Not only immotile spermatozoa but also stable objects such as blood cells, debris or halos remain on the background image. Therefore, segmentation of the objects is required to select only immotile spermatozoa from the stable objects. We tested two technique for the segmentation of the spermatozoa.

### 5.2.1 Fuzzy C-Means Based Segmentation

Clustering is a segmentation process. In short, it utilizes different groups by adjusting the similarities of objects as maximum inside the same group (cluster) compared to others. In regular clustering, each object can be assigned to one specific single group. However, in Fuzzy C-Means (FCM) clustering, each object can be represented by different clusters via similarity measures [100]. Therefore, it is also referred to as a soft clustering technique.

Fuzzy logic-based segmentation, FCM, was initially tested for the immotile spermatozoa segmentation. Spatial features of segmented objects are extracted with blob analysis. Object clustering as sperm/non-sperm is performed with the feature elimination of objects by comparing already defined parameters of manually selected spermatozoa.

FCM is similar to the regular K-means clustering technique, but different in that it involves the membership values of each object to class with a fuzzifier level, which is the determination of fuzziness. Equation 5.1 indicates the $k^{th}$ centroid calculation of FCM.

$$c_k = \frac{\sum_{s=1}^{S} w_k(s)^m s}{\sum_{s=1}^{S} w_k(s)^m} \tag{5.1}$$

where $S$ refers to total number of samples. $k$ and $s$ represents the selected cluster and sample. $w_k(s)$ is the degree of being in the $k^{th}$ cluster for sample $s$. $m$ is the fuzzifier level. Next to the calculation of the centroid for $k^{th}$ cluster ($c_k$), updates are performed until the convergence by using Equation 5.2.

$$\underset{c}{argmin} \sum_{i=1}^{S} \sum_{j=1}^{K} w_{ij}^m \left\| s_i - c_j \right\|^2 \tag{5.2}$$

$K$ indicates the total number of clusters defined in the first iteration. FCM aims to minimize the centroids by controlling Equation 5.2. $w_{ij}^m$ indicates the degree of element $s_i$ belonging to cluster $c_j$. It is calculated by Equation 5.3.

$$w_{ij} = \frac{1}{\sum_{k=1}^{K} \left( \frac{\left\| s_i - c_j \right\|}{\left\| s_i - c_k \right\|} \right)^{\frac{2}{m-1}}} \tag{5.3}$$

After the segmentation of extracted ROI by FCM, pixel-based spatial analysis, by another name blob analysis, was performed over segmented parts to extract spatial-based features. Objects were classified as sperm according to previously defined parameters of sperm shapes. Extracted features are demonstrated in Figure 5.2.



$$35 < S \leq 60$$

$$1.7 < \frac{Axis_{major}}{Axis_{minor}} \leq 2.2$$
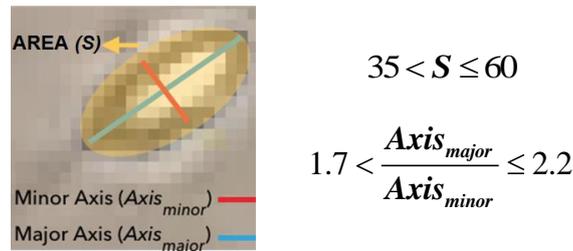
Figure 5.2 Extracted Features by Blob Analysis

Figure 5.3 demonstrates two examples of the segmentation steps of extracted ROI. The threshold levels of each feature for the determination of sperm cells were assigned as

the average values of 100 manually selected sperms' features. As a result, the features of the sperm cells should be in the range presented in Figure 5.2. Clinical concentration results will be given in the spermatozoa counting part of this chapter.
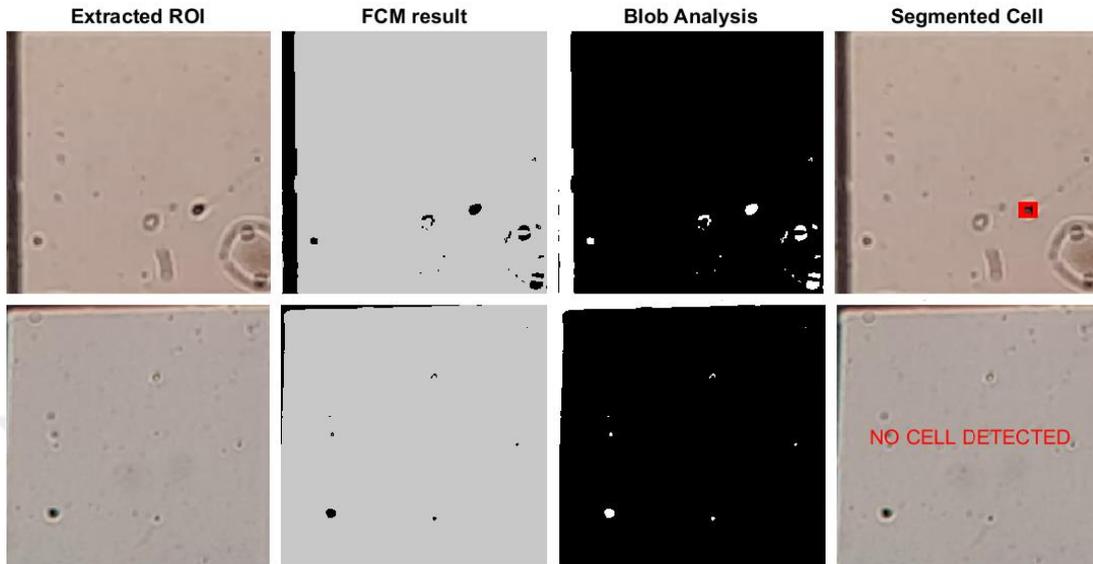


Figure 5.3 Segmentation Steps of Sperm Cells

### 5.2.2 Active Contour with Dual Thresholding

Gray level histograms of the images play important role in the segmentation. According to the histogram chart of the extracted ROI of the Makler images, a normal distribution can be observed. Thresholding is an essential part of any image retrieval applications [101]. The key idea behind the method is replacing each pixel with black pixel if the intensity value is less and/or more than predefined threshold. Dual thresholding over normal distribution was applied to signify immotile objects (immotile spermatozoa, white blood cells, leucocyte, debris etc.) remained in background model. In terms of dual thresholding, Expression 5.4 is utilized over Gaussian curve of Histogram values illustrated in Figure 5.4.
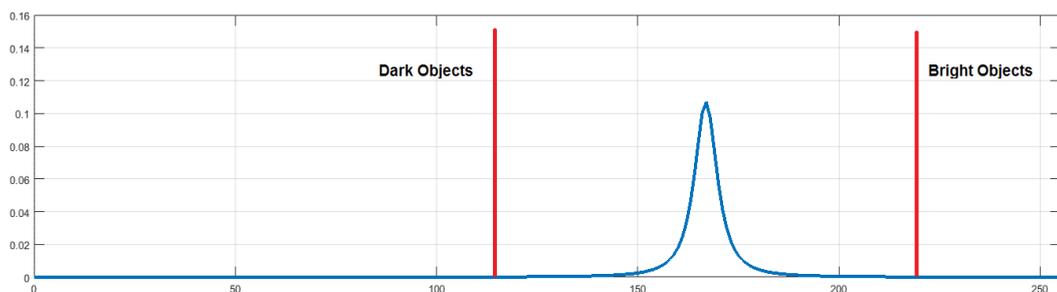


Figure 5.4 Dual Thresholding process over Gaussian curve of Histogram Values

$$\begin{cases} T_{ij} = 1 & T_{ij} < S_1 & \textit{Dark Objects} \\[2mm] T_{ij} = 0 & S_1 \geq T_{ij} \geq S_2 & \textit{Background} \\[2mm] T_{ij} = 1 & T_{ij} > S_2 & \textit{Bright Objects} \end{cases} \tag{5.4}$$

$T_{ij}$ represents the gray level pixel value of the image at $(i,j)$ location. The dual thresholding levels were calculated by Gaussian parameters denoted as $S_1$ and $S_2$. Mathematically the threshold are calculated using Equation 5.5.

$$S_1 = \mu\big(H(T)\big) + \big(3 \times \sigma\big(H(T)\big)\big)$$
$$S_2 = \mu\big(H(T)\big) - \big(3 \times \sigma\big(H(T)\big)\big) \tag{5.5}$$

where $H(T)$ indicates the histogram of image $T$. According to Gaussian theorem, if the distribution is normal, mean and standard deviation can be calculated over the distribution of histogram by using Equation (5.6).

$$\mu = \frac{1}{N}\sum_{i=1}^{N} h_i$$
$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(h_i - \mu)^2} \tag{5.6}$$

where $h_i$ denotes $i$th histogram level and $N$ is the total number. 8 bit histogram was used in this study. Hence, $N$ is equal to 255 histogram level. Following to thresholding process, several morphological operations were performed to emphasize the segmentation of objects. Morphological methods rely on the mathematical and logical operations. In order to give brief information of methods, erosion performs a simple 'logic AND' operation and dilation carries out 'logic OR' process with a mask structure around the segmented object. Example outputs of abovementioned segmentation and post processing steps on extracted background model are illustrated in Figure 5.5.

| Background 1 | Threshold | Background 2 | Threshold |
|---|---|---|---|
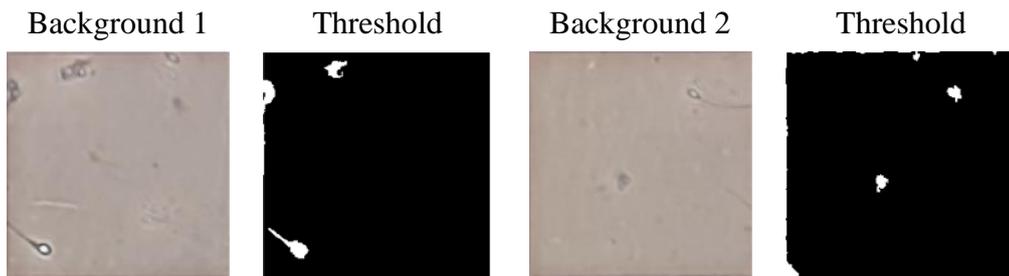


Figure 5.5 Dual thresholding segmentation results

Afterwards, active contours as a diffusion method in spatial domain was performed to strengthen spermatozoa with their edges and tails as fully detectable form [102]. Active contour was used as delineating the outline of objects. In some specific cases, using features derived by only histogram-based segmentation is not effective method for elimination of objects. Tails of spermatozoa might be missed out due to similar intensities with background. This similarities may result in elimination of spermatozoa in next stages because of the segmentation of incorrectly derived features. Active contours were principally employed in this study to fix sperm detection problem by detecting the contours of complete structure. It stands on the energy minimizing between internal and external forces. In theory, deformable spline is employed on object. External forces pull the spline towards object contours and internal forces resist deformation. Strictly changing points (edges) with high value ends up the algorithm. Active contour (also known as snake equation) is presented in Expression 5.7. Details of internal and external forces are formulized in Equation 5.8 and 5.9, respectively.

$$\int_{0}^{1} E_{snake}(v(c))d_c = \int_{0}^{1} E_{internal}\left(v\left(c\right)\right) + E_{external}(v(c))d_c \tag{5.7}$$

$$E_{internal} = E_{cont} + E_{smth} \tag{5.8}$$

$$E_{external} = (w_{line} \times E_{line}) + (w_{edge} \times E_{edge}) \tag{5.9}$$

where $v(c)$ represents the contour defined by a set of n points. Contour energy nominated as $E_{snake}$ consists of separately calculated internal ($E_{internal}$) and external ($E_{external}$) energies of objects. Internal energies are defined as the sum of the continuity ($E_{cont}$) and smoothness ($E_{smooth}$). It is mainly used to control snake deformation on contours. External energy is a combination of the feature energies of images such as edges and lines. $w_{line}$ and $w_{edge}$ are the weights of the features. $E_{internal}$ can be expanded as in Equation 5.10.

$$E_{cont} = \frac{1}{2} \times (\alpha(c) \times |v(c)|^2)$$

$$\tag{5.10}$$

$$E_{smth} = \frac{1}{2} \times (\beta(c) \times |v(c)|^2)$$

58

$\alpha$ and $\beta$ are user-defined weights that control the internal energy function's sensitivity to the amount of stretch in the snake and the amount of curvature in the snake, respectively, and thereby control the number of constraints on the shape of the snake. In practice, a large weight $\alpha(c)$ for the continuity term penalizes changes in distances between points in the contour. A large weight $\beta(c)$ for the smoothness term penalizes oscillations in the contour and will cause the contour to act as a thin plate.

Since the active counter method requires predefined shape information of desired contour, or seed points, the method is not fully automated method on finding contours in images. It depends on other mechanisms such as interaction with a user to set up a mask, some higher-level image understanding process, or information from image data. We used the segmentation result obtained by histogram-based thresholding as mask (seed point) for deformation. Energy of snake was calculated by minimizing the differences of internal and external force energies by expending or collapsing the mask.

Each counter of segmented objects was estimated by using snake active contour algorithm. After active contour calculation, another segmentation was utilized to detect specifically spermatozoa within all segmented stable objects similar to FCM clustering approach. Blob analyses were implemented to get several feature sets in this respect. Area, Eccentricity, Perimeter and Circularity of the detected contours were extracted as features. Mathematically demonstration of the features are given in Expression 5.11.

$$Area_k = \sum_{x=1}^{H}\sum_{y=1}^{W}(S_{x,y})_k$$

$$Perimeter_k \approx \pi\left[3\times(a_k+b_k)-\sqrt{(3a_k+b_k)\times(a_k+3b_k)}\right]$$

$$Eccentricity_k = \sqrt{1-\frac{b_k^2}{a_k^2}}$$

$$Circularity_k = \frac{Perimeter_k^2}{4\times\pi\times Area_k}$$

(5.11)

$N$ indicates the total number of segmented objects with the contours in the image. $(S_{x,y})_k$ represents the pixel value at $(x, y)$ location of $k_{th}$ segmented object. Value of the pixels inside of detected object are equal to 1 because of the thresholding process. Hence, sum of the pixel gives the area of the objects in image. $a_k$ and $b_k$ refers to major

and minor axes of the objects, respectively. Major axis explains the maximum distance and minor axis indicates the minimum distance between the contours. Perimeter was calculated by using Ramanujan approximation formula [103].

The area and perimeter features might be misleading parameters. Therefore, eccentricities and circularities were also calculated to strengthen the segmentation results. Eccentricities of objects tent to be 0 if the objects are in a circular structure. However, spermatozoa are mostly elliptical. In this case, eccentricity of the spermatozoon should be between 0 and 1. Another parameter, circularity, is taken into account as well. Average of the manually selected 100 spermatozoa features with $\pm\sigma$ interval were considered as reference values for each derived feature sets of objects. Reference standards are listed in Table 5.1.

Table 5.1 Features and Reference Values

| Feature Name | Criteria | $\pm\sigma$ |
|---|---|---|
| Area | 83 | 29 |
| Perimeter | 46 | 22 |
| Eccentricity | 0.6 | 0.1 |
| Circularity | 2.37 | 1.1 |

Segmentation was performed by the correlation between reference and object values. In some cases only the size is not distinctive because of the plenty of debris and blood cells with the similar size with spermatozoa. Therefore, size information is strengthened with other features that are more identical and mostly the same for all spermatozoa shapes. Figure 5.6 shows the segmentation and marking only the immotile spermatozoa in the images.
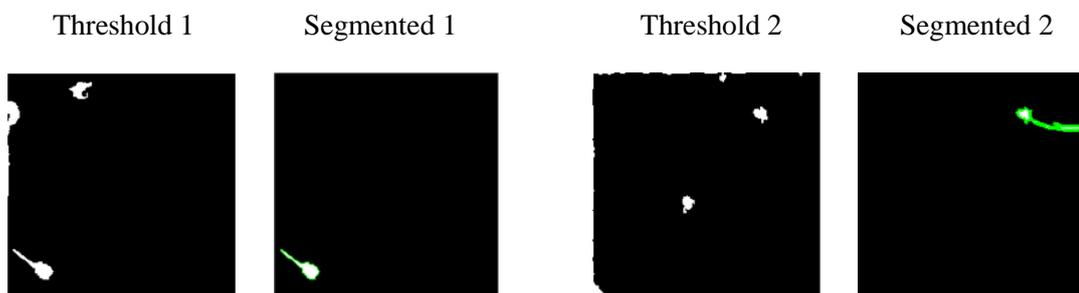


Figure 5.6 Segmentation results based on derived features by blob analyses over active contours

The results of this approach for the immotile spermatozoa detection will be given together with the motile spermatozoa detection results in the spermatozoa counting section.

## 5.3    Motile Spermatozoa Detection

In semen specimen, motile spermatozoa are the only moving objects. So, foreground detection process is enough to detect motile spermatozoa. Additionally, the performance of the video stabilization module is measured in this respect. Frame differencing was utilized between the adaptive mean background model and the original video sequence to extract the foreground model [104]. Method was performed over 90 images (3 seconds). Formula of algorithm is shown in Equation 5.12.

$$P[D(t)] = P[V(t)] - P[B]$$

$$F(x, y) = \sum_{i=1}^{f} \left[ \left( 1 - D(x, y, t - i) \right) \times \frac{1}{f} \right] \tag{5.12}$$

where $B$ indicates the extracted background from Equation 2.1 and $V(t)$ is the frame at time $t$. Subtraction of frames with background is realized over densities represented here by P. $D(t)$ is subtracted image at time $t$ between background model and video frame. After all, $F(x, y)$ is the foreground image over $f$ frame which refers to 90 frames in this study. Totally 10 foreground models were extracted due to the duration of video sub-samples (900 frames). Example results of foreground extraction is illustrated in Figure 5.7.

| Sample 1 | Foreground 1 | | Sample 2 | Foreground 2 |
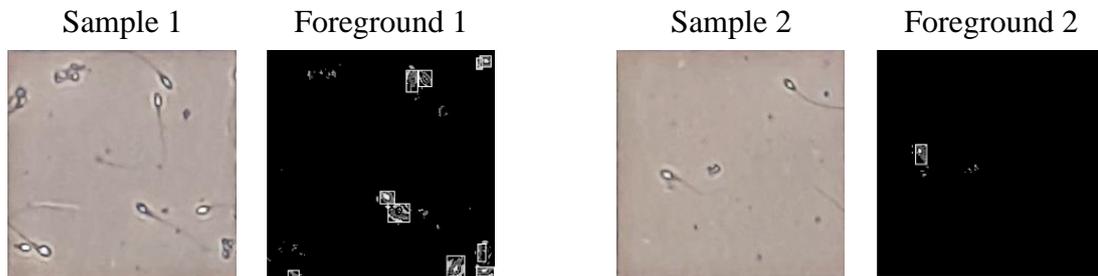


Figure 5.7 Original images and Foreground models of two example images

## 5.4    Spermatozoa Counting

Motile and immotile spermatozoon detections were performed over foreground and background models of videos derived from 90 and 300 frames, respectively. Totally, 10 foreground and 3 background models were extracted over each sub-sample (900

frames). In this way, detection process was repeated every 3 seconds for motile and 10 seconds for immotile spermatozoon. Counting process in computer side was maintained with the similar idea of using Makler chamber in VA technique. Normally, number of sperm counted in 10 squares refers to the sperm concentration as millions in per ml for given sample. In this study, ROI were increased from 10 to 36 squares. Detections of motile and immotile spermatozoon inside the each square were individually counted. Additionally, given main sample was split into 30 seconds sub-samples. Each partition was analyzed separately to verify and ensure the results. Entire sperm detections as motile and immotile spermatozoa of sub-samples were, then, calculated by sum of detection in each square. Counting process was repeated as the total number of arranged sub-samples. Final result of the given sample was accepted as the average results of each counting result of sub-samples. Mathematically expressions of counting process are presented as in follows:

$$a_k = \frac{1}{10} \times \frac{10}{36} \times \sum_{f=1}^{10} \sum_{w=1}^{36} m_{f,w}$$

$$d_k = \frac{1}{3} \times \frac{10}{36} \times \sum_{b=1}^{3} \sum_{w=1}^{36} n_{b,w}$$

$$A_{patient\_no}^{sample\_no} = \frac{1}{n} \sum_{k=1}^{n} a_k \qquad (5.13)$$

$$D_{patient\_no}^{sample\_no} = \frac{1}{n} \sum_{k=1}^{n} d_k$$

$$n = \frac{t}{30}$$

In the expressions, $a_k$ and $d_k$ denote motile and immotile spermatozoa numbers detected inside of sub-samples. $n$ is the total number of sub-sample partition derived from $t$ seconds main sample. $A$ and $D$ refer to cumulative counted numbers of $a$ and $d$, respectively. Cumulative results are calculated by averaging results of sub-samples. $m_{f,w}$ and $n_{b,w}$ indicate the detected motile and immotile spermatozoa in foreground and background models. $w$ refer to extracted ROI square number, $f$ represents the foreground model number, and $b$ indicates the background model number.

In the clinical tests, we followed the instruction reported by Björndahl et al. [11]. Detection technique was tested on semen specimen which were collected from refrained males from any sexual activity (no ejaculation) for at least two days, but not more than

7 days. Patients have no missed early ejaculate fractions. Samples were requested from subjects as ejaculating into provided sterile sample cup by masturbation in the morning between 9 – 11 am. Samples were kept in 37° C for 30 min. after ejaculation process. Then, liquefaction, volume and viscosity analysis have been determined visually by the experts. Next to the semen analysis, 5 μl. volume of semen or sperm suspension was loaded into the Makler chamber according to manual. First, an expert manually counted the sperm cells and registered under the VA in the tables. Manual counting procedure was done in duplicate and compared; counts were repeated on new samples when the difference between duplicates exceeded the acceptance limits.  Then, the videos of Makler chambers under the 20x Microscope magnification were recorded by the smartphone based data acquisition approach illustrated in Figure 2.3. Videos were transferred to a computer over a local network. CSCTAS analyses were then initialized. Olympus BX50 microscope was used in both VA and CSCTAS analysis.

The evaluation of the FCM clustering-based sperm counting analysis was made by comparing the results with the visual assessment technique. This part was tested as a preliminary clinical research, hence, only 5 subjects were selected for immotile sperm-counting analysis. Each extracted ROI was individually analyzed by FCM-based segmentation. Then, the results of 36 squares total were collected. Normally, counting sperm in 10 squares of the entire area of Makler indicates the millions per ml in the visual assessment technique. Similarly, we performed the idea over 36 squares to generalize the counting result as millions per ml. The results of proposed counting approach (FCM) and visual assessment (VA) is presented in Table 5.2.

Table 5.2 Immotile spermatozoa results of FCM based segmentation and Visual Assessment technique

| million / ml. | Subject 1 | | | Subject 2 | | | Subject 3 | | | Subject 4 | | | Subject 5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $V_1$ | $V_2$ | $V_3$ | $V_1$ | $V_2$ | $V_3$ | $V_1$ | $V_2$ | $V_3$ | $V_1$ | $V_2$ | $V_3$ | $V_1$ | $V_2$ | $V_3$ |
| FCM | 3.4 | 3.8 | 4.2 | 10.7 | 12.2 | 12.1 | 29.7 | 31.8 | 27.3 | 42.8 | 35.2 | 36.2 | 72.1 | 81.9 | 80.4 |
| VA | 0 | 0 | 0 | 7 | 9 | 7 | 24 | 27 | 22 | 34 | 29 | 31 | 40+ | 40+ | 40+ |
| *error* | 3.4 | 3.8 | 4.2 | 3.7 | 3.2 | 5.1 | 5.3 | 4.8 | 5.3 | 8.8 | 6.2 | 5.2 | - | - | - |

Sperm counting results indicate that the proposed approach should be improved. Subject 1 is the case for no sperm present in semen. However, a blob analysis-based segmentation resulted in incorrect sperm detection. Additionally, more sperm than in the VA technique was counted in all cases. *Error* is indicated in the differences. In the

case of more sperm in semen, the proposed approach concludes more *error*. Therefore, different feature extraction techniques such as wavelets or individual sperm analysis as motile or immotile should be utilized in the counting part to decrease the *error* rate. On the other hand, in of over sperm cases such as Subject 5, it is impossible to count manually. Therefore, it is noted as "*40+*" to indicate more than 40 million per ml. Our proposed counting method successfully reported the counting results in such cases. In our second approach to this problem, we implemented active contour with a modified Otsu thresholding; dual thresholding for increasing the success of the spermatozoa detection.

The second technique for the immotile and motile detection was tested on 32 semen videos obtained at four different times from 8 subjects. Patients were classified into normal and abnormal classes according to the reference semen parameters published by the WHO in VA technique. Experts separately counted the motile and immotile sperm cells within 10 squares to generalize the result as million per ml. Therefore, manually counted numbers inside 10 squares directly recorded into table as sperm concentration in millions per ml according to manual of Makler. Addition to VA, one of the basic CASA system, SQA-Vision, was also utilized in the study. It is an automated sperm quality analyzer, which provides sperm concentrations, motility percentage as progressive and non-progressive movement, and roughly morphology assessment. However, recording or observing the sample in computer side is not possible. It does not give full view of the semen in details. It is an integrated and compact system which cannot analyze the samples having the concentration less than 5 million/ml. However, system was employed in the tests owing to compare the Makler counting results of both techniques as computerized and manual to one of CASA system. According to WHO manual published in 2010, normal sperm concentration and active sperm ratio is accepted as at least 15 million/ml and 40% of concentration, respectively. Results and the comparison between VA, CASA and CSCTAS will be interpret in this scope.

In the tables, $A_p^s$ and $D_p^s$ denote motile and immotile spermatozoa numbers for specimen *s* of subject *p*, respectively. Numbers written in normal font represent the generalized sperm concentration in millions per ml. while italic ones indicate the detected total sperm cells by developed counting software. Additionally, the percentage of motile spermatozoa over total sperm numbers is given in tables under "%" column to compare

the detection of motility rate between methods. All the comparative results in standard deviation and the average rates will be given in the last table of this chapter.

Table 5.3 indicates the results for the azoospermia case. Azoospermia is defined as the absence of the sperm cells in ejaculated semen. Each sample of subject 1 is evaluated correctly by VA technique owing to observed none sperm in 10 squares. However, proposed CSCTAS has partly misdetections in immotile spermatozoa evaluation, which is indicated in italic numbers. When the results are generalized from 36 squares to ml by using Equation 5.13, these misdetections can be negligible. There is no misdetection on motile spermatozoa counting. On the other hand, SQA CASA system concludes with "No Operation" because it is not functional on evaluation of samples less than 5 million/ml sperm concentrations.

Table 5.3 Test results of Subject 1 (Azoospermia)

| | $A_1^1$ | $D_1^1$ | % | $A_1^2$ | $D_1^2$ | % | $A_1^3$ | $D_1^3$ | % | $A_1^4$ | $D_1^4$ | % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| CASA | No Operation | | | | | | | | | | | |
| CSCTAS | 0 | 0.2 *(1)* | 0 | 0 | 0.5 *(2)* | 0 | 0 | 0.2 *(1)* | 0 | 0 | 0.2 *(1)* | 0 |

The results for the Oligospermia case are given in Table 5.4. Oligospermia explains the cases having less sperm concentration than normal level which is defined as 15 million/ml in the WHO manual. Subject 2 has an oligospermia diagnoses according to concentration evaluated between 3 and 5 million/ml. presented in Table 5.4. Again, CASA system reports "No Operation" because of the same reason as in azoospermia case. VA and CSCTAS results are similar, but CSCTAS results are more reliable and effective owing to 36-square analyses instead of only 10 as in VA.

Table 5.4 Test results of Subject 2 (Oligospermia and Asthenospermia)

| | $A_2^1$ | $D_2^1$ | % | $A_2^2$ | $D_2^2$ | % | $A_2^3$ | $D_2^3$ | % | $A_2^4$ | $D_2^4$ | % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VA | 0 | 4 | 0 | 1 | 3 | 25 | 0 | 2 | 0 | 0 | 4 | 0 |
| CASA | No Operation | | | | | | | | | | | |
| CSCTAS | 0.3 *(1)* | 2.8 *(10)* | 9 | 0.3 *(1)* | 2.5 *(9)* | 10 | 0.3 *(1)* | 3.6 *(13)* | 7 | 0.3 *(1)* | 3.1 *(11)* | 8 |

The manual of Makler suggests that the multiple test should be performed to obtain more stable results. The main purpose of this is to expand the analyzing area which provides more trustful and objective results. It can be seen on motility rates clearly. VA results are inconsistent over motility rates due to only one or none motile sperm

observation within 10 squares. When the range of field is expanded in analysis, sperm counting over wide range provides more objective and consistent results. Motility rate is evaluated between 8% and 10% by CSCTAS whereas VA is reported as 0% to 25%. Another diagnoses, asthenospermia, can be achieved by the aid of CSCTAS for this subject. Asthenospermia is identified as the low motility rates. On the other hand, VA is inefficient to perform such analyses for this subject.

Table 5.5 Test results of other Subjects

| | $A_3^1$ | $D_3^1$ | % | $A_3^2$ | $D_3^2$ | % | $A_3^3$ | $D_3^3$ | % | $A_3^4$ | $D_3^4$ | % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VA | 6 | 8 | 42 | 5 | 8 | 38 | 6 | 9 | 40 | 4 | 8 | 33 |
| CASA | 14.1 | | 36 | 13.7 | | 34 | 14.2 | | 35 | 13.8 | | 36 |
| CSCTAS | 5.2 *(19)* | 7.5 *(27)* | 41 | 4.4 *(16)* | 7.7 *(28)* | 36 | 5.2 *(19)* | 7.7 *(28)* | 40 | 3.88 *(14)* | 7.2 *(26)* | 35 |
| | $A_4^1$ | $D_4^1$ | % | $A_4^2$ | $D_4^2$ | % | $A_4^3$ | $D_4^3$ | % | $A_4^4$ | $D_4^4$ | % |
| VA | 10 | 12 | 45 | 10 | 11 | 47 | 11 | 11 | 50 | 10 | 11 | 47 |
| CASA | 22.2 | | 45 | 21.9 | | 47 | 21.4 | | 49 | 22.4 | | 46 |
| CSCTAS | 8 *(29)* | 10.6 *(38)* | 44 | 8.9 *(32)* | 10 *(36)* | 47 | 9.4 *(34)* | 8.9 *(32)* | 51 | 9.7 *(35)* | 10.2 *(37)* | 48 |
| | $A_5^1$ | $D_5^1$ | % | $A_5^2$ | $D_5^2$ | % | $A_5^3$ | $D_5^3$ | % | $A_5^4$ | $D_5^4$ | % |
| VA | 2 | 28 | 7 | 3 | 23 | 11 | 2 | 21 | 9 | 2 | 23 | 8 |
| CASA | 28.3 | | 6 | 27.2 | | 8 | 27.5 | | 8 | 26.6 | | 8 |
| CSCTAS | 2.5 *(9)* | 27.2 *(98)* | 8 | 3.3 *(12)* | 25.2 *(91)* | 11 | 1.6 *(6)* | 20.8 *(75)* | 8 | 1.9 *(7)* | 21.4 *(77)* | 8 |
| | $A_6^1$ | $D_6^1$ | % | $A_6^2$ | $D_6^2$ | % | $A_6^3$ | $D_6^3$ | % | $A_6^4$ | $D_6^4$ | % |
| VA | 21 | 10 | 68 | 23 | 11 | 68 | 21 | 13 | 62 | 22 | 13 | 63 |
| CASA | 34.1 | | 65 | 34.4 | | 67 | 33.8 | | 66 | 32.9 | | 65 |
| CSCTAS | 21.1 *(76)* | 9.7 *(35)* | 68 | 22.5 *(81)* | 11.1 *(40)* | 67 | 20 *(72)* | 11.1 *(40)* | 64 | 21.9 *(79)* | 10.8 *(39)* | 67 |
| | $A_7^1$ | $D_7^1$ | % | $A_7^2$ | $D_7^2$ | % | $A_7^3$ | $D_7^3$ | % | $A_7^4$ | $D_7^4$ | % |
| VA | 21 | 24 | 47 | 20 | 22 | 47 | 19 | 23 | 45 | 22 | 25 | 47 |
| CASA | 44.1 | | 48 | 43.7 | | 46 | 44.8 | | 46 | 45.1 | | 49 |
| CSCTAS | 20.2 *(73)* | 21.6 *(78)* | 48 | 19.1 *(69)* | 21.9 *(79)* | 47 | 18.3 *(66)* | 23.8 *(86)* | 44 | 20.8 *(75)* | 23.3 *(84)* | 47 |
| | $A_8^1$ | $D_8^1$ | % | $A_8^2$ | $D_8^2$ | % | $A_8^3$ | $D_8^3$ | % | $A_8^4$ | $D_8^4$ | % |
| VA | 23 | 25 | 48 | 24 | 27 | 47 | 24 | 25 | 49 | 25 | 25 | 50 |
| CASA | 50.8 | | 48 | 51.1 | | 47 | 50.3 | | 48 | 50.7 | | 49 |
| CSCTAS | 22.5 *(81)* | 23.8 *(86)* | 48 | 23.8 *(86)* | 25.2 *(91)* | 48 | 22.7 *(82)* | 23.9 *(86)* | 48 | 23.6 *(85)* | 24.7 *(89)* | 49 |

Analyzing of other 6 subjects are given in Table 5.5. SQA CASA does not provide individual numerical results for immotile and motile sperm concentration. System gives the motile spermatozoa ratio over total number. The results recorded in tables as the concentration and the motile ratio. According to Table 5.5, motility rates of VA and CSCTAS are more similar. Because the developed sperm counting approach is basically inspired from VA technique. Only the difference is about the field of range in

calculation. Sperm concentration results between VA and CSCTAS are slightly different, but developed counting software offer more objective and consistent results. According to obtained concentration results, CSCTAS technique is effective and can be used instead of VA and CASA.

According to experiments, the analysis of immotile spermatozoa by CSCTAS causes several misdetections due to presence of noise and undesirable particles in the wide range field of view. Results of subject 1 indicate this problem. However, generalization of the results by using Makler standardization with 36 square minimizes the problem and provides better understanding over diagnoses. Since the number of motile spermatozoa detected by VA and CSCTAS are closer, foreground extraction method are perfectly adapted to the proposed approach.

The average and standard deviation of 4 analysis for each subjects are presented in Table 5.6. Individual detected sperm numbers by all techniques are averaged to present better comparison between the techniques. Results also indicate the consistency of the technique on practical use. Normally, similar results are supposed to obtain within sub-samples of certain main sample. Standard deviation represents the probabilities of getting similar results of samples that given by same subject in different times. Minimal values of standard deviation refer to more consistent technique.

Table 5.6 Average results and Standard Deviations of 4 samples for each subjects

| | Motile Spermatozoa | | | | | | Immotile Spermatozoa | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | VA | | CASA | | CSCTAS | | VA | | CASA | | CSCTAS | |
| | Avg. | $\sigma$ | Avg. | $\sigma$ | Avg. | $\sigma$ | Avg. | $\sigma$ | Avg. | $\sigma$ | Avg. | $\sigma$ |
| Subject 1 | 0 | 0 | | | 0 | 0 | 0 | 0 | | | 0.28 | 0.15 |
| Subject. 2 | 0.25 | 0.5 | | | 0.3 | 0 | 3.25 | 0.96 | | | 3 | 0.47 |
| Subject. 3 | 5.25 | 0.96 | 5.1 | 0.17 | 4.65 | 0.68 | 8.25 | 0.5 | 9 | 0.16 | 7.53 | 0.24 |
| Subject 4 | 10.25 | 0.5 | 10 | 0.21 | 9 | 0.74 | 11.25 | 0.5 | 12.2 | 0.59 | 9.93 | 0.73 |
| Subject 5 | 2.25 | 0.5 | 1.7 | 0.24 | 2.33 | 0.75 | 23.75 | 2.99 | 26.6 | 0.90 | 23.65 | 3.07 |
| Subject 6 | 21.75 | 0.96 | 22.2 | 0.66 | 21.4 | 1.10 | 11.75 | 1.5 | 12.1 | 0.32 | 10.68 | 0.67 |
| Subject 7 | 20.5 | 1.29 | 21.1 | 0.85 | 19.6 | 1.12 | 23.5 | 1.29 | 23 | 0.57 | 22.65 | 1.07 |
| Subject 8 | 24 | 0.82 | 24.4 | 0.36 | 23.15 | 0.65 | 25.5 | 1 | 26.4 | 0.51 | 24.4 | 0.67 |

SQA CASA is the most consistent technique due to the obtained lower standard deviation than other techniques. But the difference between the CASA and CSCTAS is less than VA. Maximum difference is recorded as 2.1 in immotile spermatozoa concentration of Subject 5 which refers to 2.1 million/ml. differences. This difference

with respect to 25 million sperm concentration is insignificant. Difference between VA and CSCTAS is less than CASA because of their similarity. It is precise that the more sample testing will reduce CSCTASs deviation when compared to VA due to 36 square analyses of semen.

Results of sperm concentration analysis is individually demonstrated by Bland-Altman plot as motile and immotile spermatozoa concentration in Figures 5.8 and 5.9. Bland-Altman plot, or difference plot, is an informative demonstration on comparing two techniques [105]. The main idea is to demonstrate the differences between techniques against the averages of techniques. Solid horizontal line indicates the mean difference, and the dashed horizontal lines represents the limits of agreement, which are defined as the mean difference plus and minus 1.96 times the standard deviation of the differences.
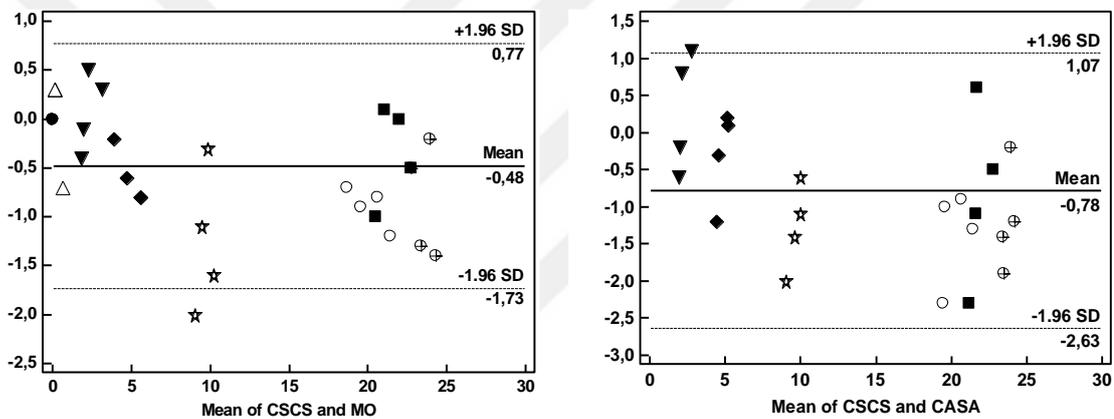


Figure 5.8 Bland-Altman plot analysis of motile spermatozoa concentration

Figure 5.8 indicates the scatter plot of two counting approach of motile spermatozoa concentration in terms of the Bland-Altman plot. According to the agreement level, graph proves that the developed CSCTAS is considered to be in agreement and may be used interchangeably with VA and CASA techniques. Only one sample analysis in each technique remained out of agreement level with slight difference.

Figure 5.9 displays Bland-Altman plot for the immotile spermatozoa concentration analysis. Similar to motile spermatozoa concentration scatter plot, developed CSCTAS can be utilized interchangeably with VA and CASA due to the obtained results remaining inside of defined agreement level.
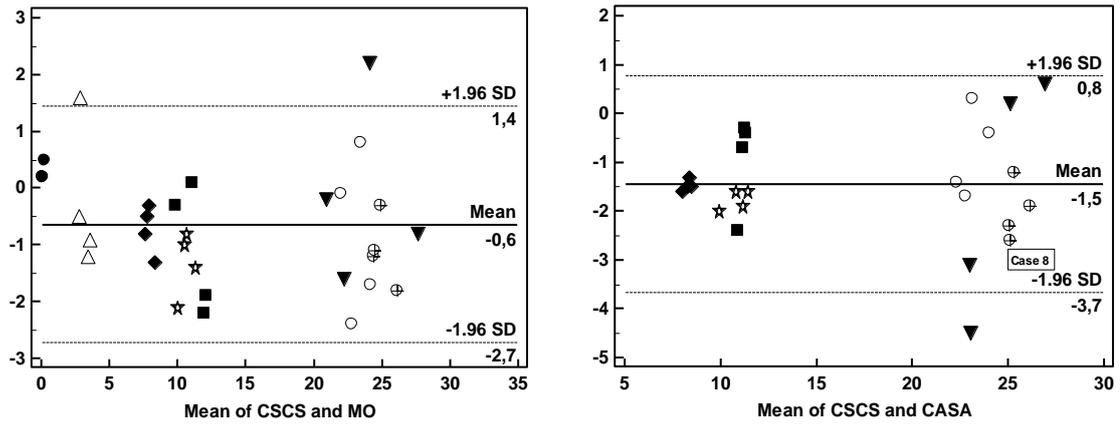
Figure 5.9 Bland-Altman plot analysis of immotile spermatozoa concentration

# SPERMATOZOA TRACKING AND TRAJECTORY ANALYSIS

## 6.1    Introduction

Sperm movement characteristic plays an important role in male fertility and motion information is the key parameter for analyzing the sperm movement. Motion is often represented by the trajectories after tracking process. In this respect, tracking is the first step for the motility analysis.

Sperm tracking is a challenging matter, because sperm have same size and shape, move fast, there is an uncertainty in their motions. Therefore, most studies have been performed to reduce tracking mistakes, especially when sperms collide with each other. Robust multi-target tracking algorithms, developed originally for radar applications and video processing, have addressed similar challenges successfully in other domains [106]. Over the years, interest in applying such algorithms to track viruses, bacteria, stem cells, sub-cellular organelles and other biological particles has increased [107]. We reviewed several tracking studies applied on especially for the sperm tracking problem.

Firstly, Katz and Davis introduced the automatic sperm tracking in the mid-1980s [108]. In the proposed system, a user-selected gray level threshold is applied to all video frames to identify sperm pixels, and the centroids of the manually segmented blobs are then accepted as sperm positions. A circular validation region centered at each position in one frame is used to select a path measurement from the next frame in order to track the sperm. User selects the radius according to the shape of spermatozoa. This method works well for tracking a small number of well-separated targets in the absence of clutter, but its effectiveness rapidly degrades if targets become closely spaced and their overlapping conditions contain multiple conflicting measurements. In these cases,

CASA systems often exclude the affected tracks from analysis or continue the track by selecting the nearest-neighbor measurements.

Beresford-Smith and Van Helden tested radar tracking algorithms in sperm tracking problem by modifying the probabilistic data association filter (PDAF) to track a single sperm [109]. Shi et al. and Liu et al. applied the recent ad-hoc methods for tracking single sperm through occlusion problem [110, 18]. Tomlinson et al. explained a CASA system that is capable of tracking multi-targets [111]. They can track multiple sperm and classify their motility using 1-sec video clips. However, the collisions or over long durations of sperm motility problems was not addressed in the paper. Su et al. developed a lens-free holographic imaging sensor to track the sperms in the 3D swimming paths during $10 - 20$ seconds in highly diluted sample preparations [112]. Their method requires a significant computational cost (greater than 2 hours of post-processing). In [113], Berezansky et al. used the mean shift and optical flow techniques for sperm detection and tracking, respectively. But, the execution time was excessive and the method cannot detect immotile sperm. Sorensen et al. studied multi-sperm tracking using both Kalman and particle filters [114]. Bar-Shalom et al. applied joint probabilistic data association filter (JPDAF) for the sperm measurement-to-track association conflicts occurring during real and apparent cell-to-cell collisions [115]. They tested the approach on the air traffic control systems. This approach uses independent Kalman filters to estimate the position and velocity of each sperm tracked. The recent work of Ristic et al. [116] summarizes these methods along with suggestions for consistent assessment of tracking algorithm performance.

Moving objects should directly indicate the motile spermatozoa because the only motile objects are spermatozoa in the semen viscosity. In the tracking part of this study, we used the detection points obtained by the foreground detection technique (frame differencing) which refers to the motile objects in the semen sample as the initialization point for the tracking techniques. We have tested two tracking approaches on the detected sperm locations. First, Mean Shift tracking was adopted [117]. Then more advanced and estimation based technique, Kalman Filter, was utilized due to the inefficient Mean Shift tracking results.

## 6.2    Tracking with Mean Shift

Mean Shift tracking is a feature searching technique in the spatial domain within a certain range [117]. After the defining initial points of objects, technique extracts several features such as color, intensity, edge, histogram etc. of the initial points. Then, in the next frame, extracted features are sought within the certain range defined by the users. This part named as Mode-Seeking. Tracking is concluded by the probability density function (PDF) which is the likelihood over object locations. The point which gives the higher likelihood is connected for the tracking. The Mean Shift tracking idea for the sperm tracking problem is demonstrated in Figure 6.1.
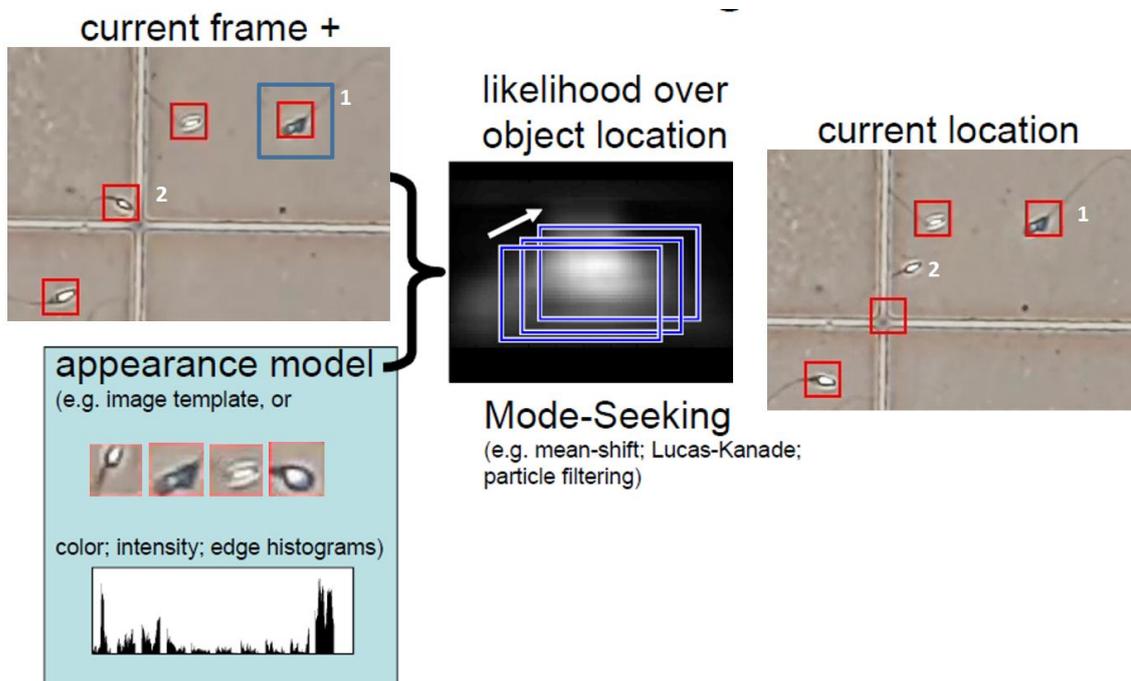


Figure 6.1 Mean Shift tracking demonstration on sperm tracking problem

Mean Shift tracking is a robust technique when the features are distinctive from the background and objects are identical. But, in sperm tracking problem, it is not an efficient technique due to the similar appearance models in the frame such as multiple spermatozoa and the grid structure. In case of sperm or grid occlusion for specific sperm tracking, technique confuses the derived features of the appearance model in the mode-seeking step. An example of successful and unsuccessful tracking can be observed in Figure 6.1. In the figure, the sperm marked as 1 is successfully tracked by seeking the extracted features in the next frame, but tracking the number 2 failed due to the similar extracted features with the grid structure.

## 6.3 Tracking with Recursive Kalman Filter

As a result of unsatisfactory performance of visual evaluation for the Mean Shift tracking technique, we applied the Kalman approach suggested in [118]. Recursive $\alpha$ $\beta$ Kalman Filter (RKF) is employed over the sequential image sets. RKF is based on Bayesian approach which employs the current and previous states and measurements of the objects and estimates the objects new states in the next frame [124].

$$P(x \mid m) = \frac{P(x) \times P(m \mid x)}{P(m)}$$
(6.1)
$$P(x_t \mid x_{t-1}, m_t)$$

Equation 6.1 indicates the Bayesian theory in which P(x) is the prior probability and $P(m/x)$ is the measurement at the state $x$. $P(x/m)$ explains the estimated new state of the object according to the obtained measurements and previous states. Recursive form of regular Bayesian equation by changing $P(x)$ with $P(x/m)$ for each frame and calculating new $P(x/m)$ due to the information obtained from former frame constitutes the RKF which can be formulated as in Equation 6.2. Velocity of the spermatozoa is accepted as the measurement ($m$) variable in the equations. The parameters used in the Equations are given in Equation 6.3.

$$x_{t+1} = x_t + v_t^x t + \frac{1}{2} a_t t^2$$
$$y_{t+1} = y_t + v_t^y t + \frac{1}{2} a_t t^2$$
(6.2)

$$v_{t+1}^x = v_t^x + at$$
$$v_{t+1}^y = v_t^y + at$$
(6.3)

where $x_{t+1}$ and $y_{t+1}$ indicate the next state estimation in 2D coordinate plane. Images are in 2D form, hence, [$x$ $y$] and [$v^x$ $v^y$] are calculated individually for the horizontally and vertically state and velocity, respectively. The estimation of velocity for the estimated state of object is nominated as the measurement ($m$) variable in Kalman Filter Equation 6.3 and obtained by the velocity and acceleration at the current time $t$. $v_{t+1}^x$ and $v_{t+1}^y$ indicate the estimation of velocity for the estimated state [$x, y$].

In the spermatozoa tracking problem, we utilized the $\alpha$ $\beta$ Kalman Filter for the estimation of the next state. $\alpha$ and $\beta$ are the coefficients for the calculation of state

and velocity estimation. Additionally, *E* parameter is defined as *error* parameter. $\alpha$ $\beta$ Kalman Filter can be formulated as in Equation 6.4. The last 10 sequential images were analyzed for the estimation of next frame and the quantization was selected as $0.1t$. State is indicated by $\bar{x}_t$. Former state information is multiplied by $\alpha$ coefficient. New state is estimated by the summing of current measurement which provides the movement of the object's state and nominates as the velocity and its acceptable error limit.

$$\bar{x}_t = \alpha \times x_{t-1} + \beta \times m_t + \mathrm{E}_x \tag{6.4}$$

Equation 6.5 is the expanded matrix form of given Equation 6.2 and 6.3 in the Kalman Filters.

$$\begin{bmatrix} x_{t+1} \\ y_{t+1} \\ v^x_{t+1} \\ v^y_{t+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & t & 0 \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_t \\ y_t \\ v^x_t \\ v^y_t \end{bmatrix} + \begin{bmatrix} \dfrac{t^2}{2} \\ \dfrac{t^2}{2} \\ t \\ t \end{bmatrix} a \tag{6.5}$$

$\alpha$, $\beta$, and $E_x$ is the Gaussian parameters in Equation 6.4. They can be written as in expressions 6.6, 6.7 and 6.8 according to Equation 6.5.

$$\alpha = \begin{bmatrix} 1 & 0 & t & 0 \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.6}$$

$$\beta = \begin{bmatrix} \dfrac{t^2}{2} \\ \dfrac{t^2}{2} \\ t \\ t \end{bmatrix} \tag{6.7}$$

$$E_x = \begin{bmatrix} \dfrac{t^4}{4} & 0 & \dfrac{t^3}{2} & 0 \\ 0 & \dfrac{t^4}{4} & 0 & \dfrac{t^3}{2} \\ \dfrac{t^3}{3} & 0 & t^2 & 0 \\ 0 & \dfrac{t^3}{3} & 0 & t^2 \end{bmatrix} \qquad (6.8)$$

The variation in the velocity is also used in the updating process of the states. In this respect, Equation 6.9 is utilized to update the velocity parameter.

$$\overline{z_t} = C\overline{x_t} + E_z$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ v_t^x \\ v_t^y \end{bmatrix} \qquad (6.9)$$

$$E_z = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

where $\overline{z_t}$ indicates the updates in the velocity, in other words, sensor prediction during the tracking. Updating is performed by the variation in the estimated state $\overline{x_t}$. Therefore, $C$ constant is defined only for the coordinates [x, y]. Additionally, a specified error parameter for the possible sensor errors is set to standard deviation of the Gaussian curve and nominated as $E_z$.

$$x_{est} = \overline{x_t} + K(z_t - \overline{z_t}) \qquad (6.10)$$

Equation 6.10 indicates the last estimated state including the update effects in the velocities to the pre-estimated state. K is the Kalman gain parameter which aggravate the effects of changes in velocity. Equation 6.10 provides more accurate state estimation and $x_{est}$ is the final location to be used for the tracking in the study.

The detections of spermatozoa and the estimated locations by Kalman Filter should be assigned for the track identification and updating process. Euclidian distance between the detected and estimated location of each spermatozoa is utilized for the cost matrix.

Local minimum results in wrong assignments due to the presence of plenty of spermatozoa. This is an assignment problem, which is illustrated in Figure 6.2.
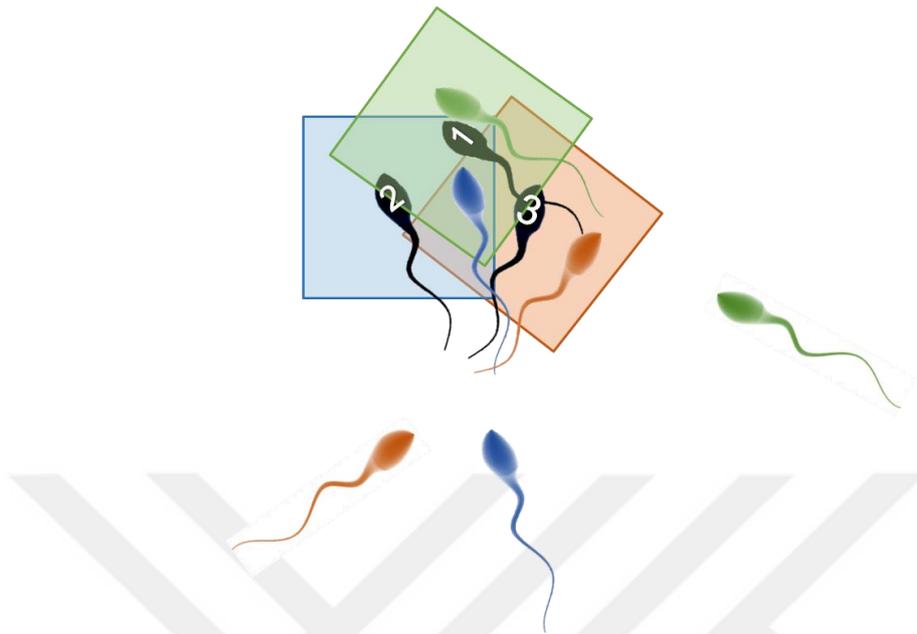


Figure 6.2 Assignment Problem of the detection and the estimation location of the spermatozoa

Black spermatozoa refer to the estimated locations at time $t+1$ of the orange, blue and green spermatozoon at time $t$. Updating the parameters of Kalman Filter requires the verification of the estimations with the detections of the spermatozoa at time $t+1$. A matching for the verification is necessary within a certain range. The squares around the estimations are used for the matching range. However, squares can involve multiple detection. Without an assignment algorithm, finding the optimal solution takes excessive times. Additionally, minimum distance matching results in inappropriate matching which is misleading for the updating process.

Assign the detections to estimated track positions were performed by Hungarian Algorithm which is utilized for the optimization of the distance vector to provide less effected trajectories from the immediate motilities. The Hungarian method is an algorithm to find an optimal assignment for a given cost (distance) matrix. In this algorithm, distance vector between the detection and the estimated track position will be assigned by using the global minimum instead of local minimums. Steps of the algorithm are as follows;

Step 1. Subtract the smallest entry in each row from all the entries of its row.

Step 2. Subtract the smallest entry in each column from all the entries of its column.

Step 3. Draw lines through appropriate rows and columns so that all the zero entries of the cost matrix are covered and the minimum number of such lines is used.

Step 4. Test for Optimality: If the minimum number of covering lines equals to the total number of state, an optimal assignment of zeros is possible and optimization is finished. (ii) If the minimum number of covering lines is less than the total states, an optimal assignment of zeros is not possible yet. In that case, proceed to Step 5.

Step 5. Determine the smallest entry not covered by any line. Subtract this entry from each uncovered row, and then add it to each covered column. Return to Step 3.

| Distance Matrix (Cost Matrix) | | Estimated Locations | | |
|---|---|---|---|---|
| | | 1 | 2 | 3 |
| Detections | Blue | 250 | 400 | 350 |
| | Orange | 400 | 600 | 350 |
| | Green | 200 | 400 | 250 |

Minimum Distance Matching

$Blue \rightarrow 1$
$Orange \rightarrow 3$
$Green \rightarrow 2$
Cost = 1000

| Distance Matrix (Cost Matrix) | | Estimated Locations | | |
|---|---|---|---|---|
| | | 1 | 2 | 3 |
| Detections | Blue | 250 | 400 | 350 |
| | Orange | 400 | 600 | 350 |
| | Green | 200 | 400 | 250 |

| Distance Matrix (Cost Matrix) | | Estimated Locations | | |
|---|---|---|---|---|
| | | 1 | 2 | 3 |
| Detections | Blue | 250 | 400 | 350 |
| | Orange | 400 | 600 | 350 |
| | Green | 200 | 400 | 250 |

Hungarian Algorithm

$Blue \rightarrow 2$
$Orange \rightarrow 3$
$Green \rightarrow 1$
Cost = 950

Figure 6.3 Calculation of the cost matrix for assignment problem

Numerical example of the illustrated three spermatozoa assignment problem in Figure 6.2 is given in Figure 6.3. Minimum distance assignment searches for the local minimums which might be the misleading assignment. In this example, the incorrect

assignment of matching the local minimums can be observed in the Figure 6.3. According to the minimum distance matching results, the estimation of the green and blue spermatozoa are updated by the detection of the blue and green spermatozoa, respectively. It is a misleading assignment, which affects the all assignments. Contrary to these results, Hungarian assignment results in the correct matching. Detection of the blue and green spermatozoa are assigned to the correct estimated tracks to update the parameters.

## 6.4    Feature Extraction and Trajectory Classification

Tracking was performed over the RKF based estimated states with their matching by Hungarian algorithm inside of the predefined ROI. Long term tracking might be misleading due to the possible long duration of blinking or noise appearing and immediate changes in the direction and/or velocity of the motilities. Therefore, trajectories of each 90 frame (3 sec.) were partly analyzed. The analyzing schema over an example extracted trajectory is shown in Figure 6.4.
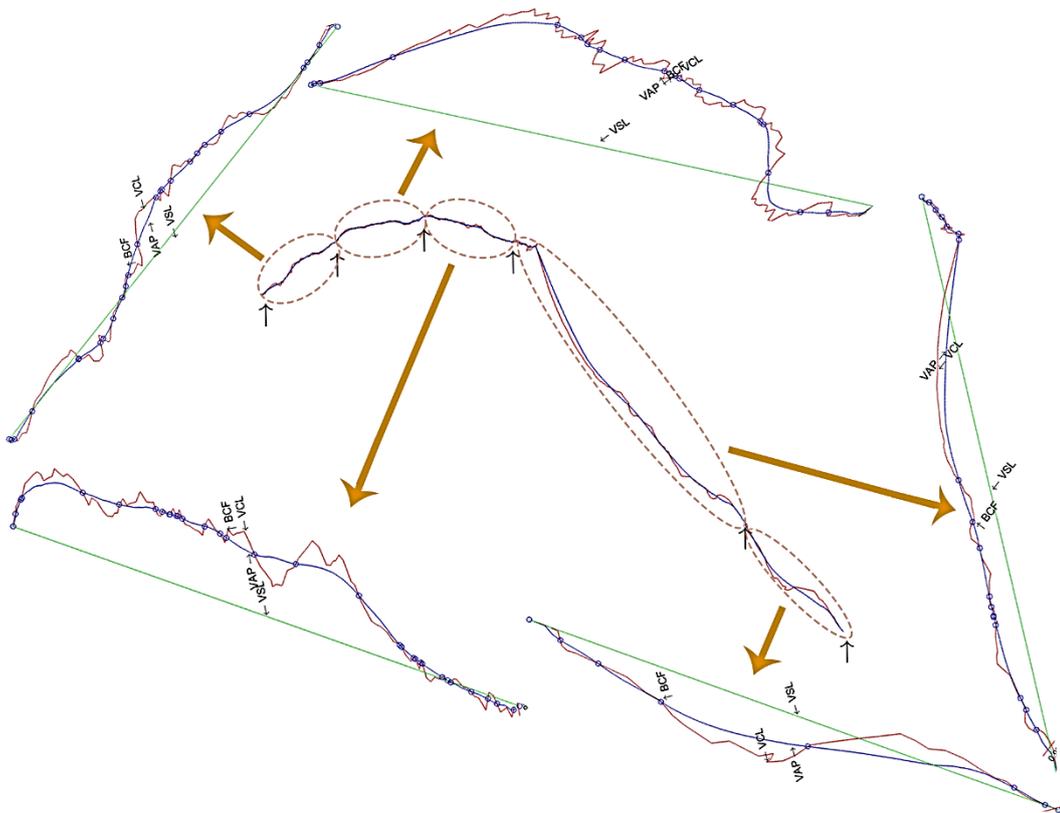


Figure 6.4 Trajectory splitting and analyzing

WHO has defined the several motility parameters of the spermatozoa [3]. Trajectory parameters are demonstrated in Figure 6.5. Several metrics are calculated from the

78

demonstrated VAP (Average Path Velocity), VCL (Curve Linear Velocity) and VSL (Straight Line Velocity) parameters.
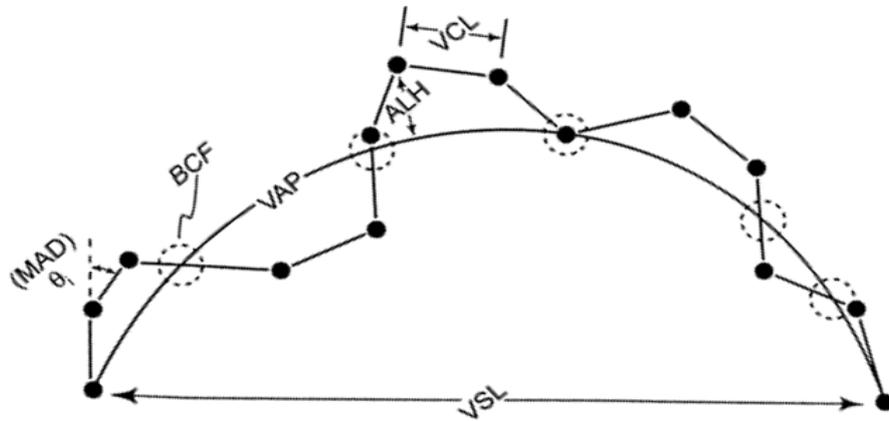


Figure 6.5 Evaluation Parameters in Visual Assessment Technique defined by WHO

The features that will be used in clustering is as follows;

- Curve Linear Velocity (VCL) - Velocity parameter calculated from the total path

- Average Path Velocity (VAP) - Velocity parameter calculated from the smoothed path

- Straight Line Velocity (VSL) - Velocity parameter calculated from the direct path between start and end points of the motility

- Amplitude of Lateral Head displacement (ALH) - Magnitude of lateral displacement of a sperm head about its average path

- Beat Cross Velocity (BCF) - The frequency of the intersection between the spermatozoa head piece and the trajectory

- Straightness (STR) - A measure of the linear convergence of the spermatozoa to the smooth path calculated by VSL/VAP

- Linearity (LIN) - A measure of the linear convergence of the spermatozoa to the curve linear path calculated by VSL/VCL

- Wobble (WOB) - A measure of oscillation of the actual path about the average path calculated by VAP/VCL

- Mean Angular Displacement (MAD) - Immediate turning angle of the sperm head along its curvilinear trajectory

WHO defined 4 type sperm motilities in the latest laboratory manual as follows;

Grade A - Fast progressive motility

Grade B - Progressive motility

Grade C - Non-progressive motility. Vibrations or inconsistent movements.

Grade D - Stable and immotile.

Motilities have been classified according to the reference values for specific features defined by the Witkowski et al. [39]. The reference values of features for the each motility types are explained as follows;

Grade A = VAP > 25 μm/s and LIN > 75%

Grade B = VAP < 25 μm/s and LIN > 75%

Grade C = LIN < 75%

Grade D = VAP < 5 μm/s

## 6.5 The Evaluation of the Spermatozoa Tracking Approach

It is important to first group motion trajectories by clustering homogeneous trajectories into the same clusters before further modeling trajectory distributions and learning motion patterns in order to handle different motion patterns more effectively and efficiently. As an initial work, we verified our derived features of the trajectories in the determination of the motility analysis with the reference values [39] by using hierarchical clustering [119]. We merged trajectories derived from 32 videos of 8 subjects (4 videos of each subject including 3 healthy, 3 low motile, 2 non-motile case) and then clustered into 4 clusters referring to the each grades. Totally 89438 trajectories were included in the clustering of the features. The purpose of this part is to verify the system in terms of feature extraction process by comparing each cluster centers to the reference values explained in [39]. Dendrogram is given in Figure 6.6.
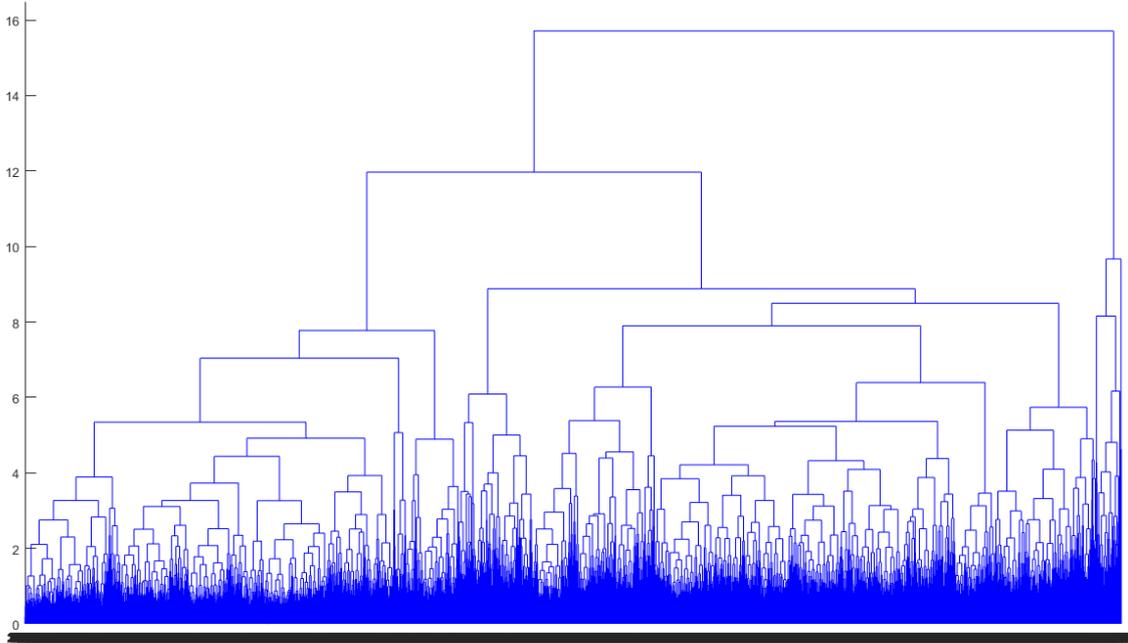
Figure 6.6 Dendrogram of the hierarchical clustering for 89438 trajectories

Cluster centers and the reference values are compared in Table 6.1. System is verified according to the similar cluster centers of the trajectory features obtained by the proposed system with the reference values explained in [39].

Table 6.1 Cluster Centers and the Referance Values

|  | VCL (μm/s) | VSL (μm/s) | VAP (μm/s) | STR (%) | LIN (%) | WOB (%) | BCF (#) |
|---|---|---|---|---|---|---|---|
| Cluster 1 (Grade A) | 35.3 | 29.2 | **32.7** | 89.2 | **82.7** | 92.6 | 11 |
| Reference Grade A | - | - | **>25** | - | **>75** | - | - |
| Cluster 2 (Grade B) | 21.8 | 17.4 | **19.4** | 89.3 | **79.5** | 88.6 | 14 |
| Reference Grade B | - | - | **<25** | - | **>75** | - | - |
| Cluster 3 (Grade C) | 11.8 | 4.9 | 8.5 | 57.2 | **40.8** | 67.9 | 17 |
| Reference Grade C | - | - | - | - | **<75** | - | - |
| Cluster 4 (Grade D) | 3.1 | 0.8 | **1.4** | 42.3 | 17.9 | 39.1 | 35 |
| Reference Grade D | - | - | **<5** | - | - | - | - |

The distribution of the three clusters of the motilities as Progressive (A+B), Non-Progressive (C) and Immotile (D) over the extracted features are shown in Figure 6.7. Plots also indicate that the reference values for the specific features such as LIN and VAP play important role on clustering. Additionally, extra features such as STR and VSL which can be employed in the determination of the progressive and immotile

motility, respectively. Plots also emphasized that the WOB feature is not useful for the clustering or classification process of the motilities.
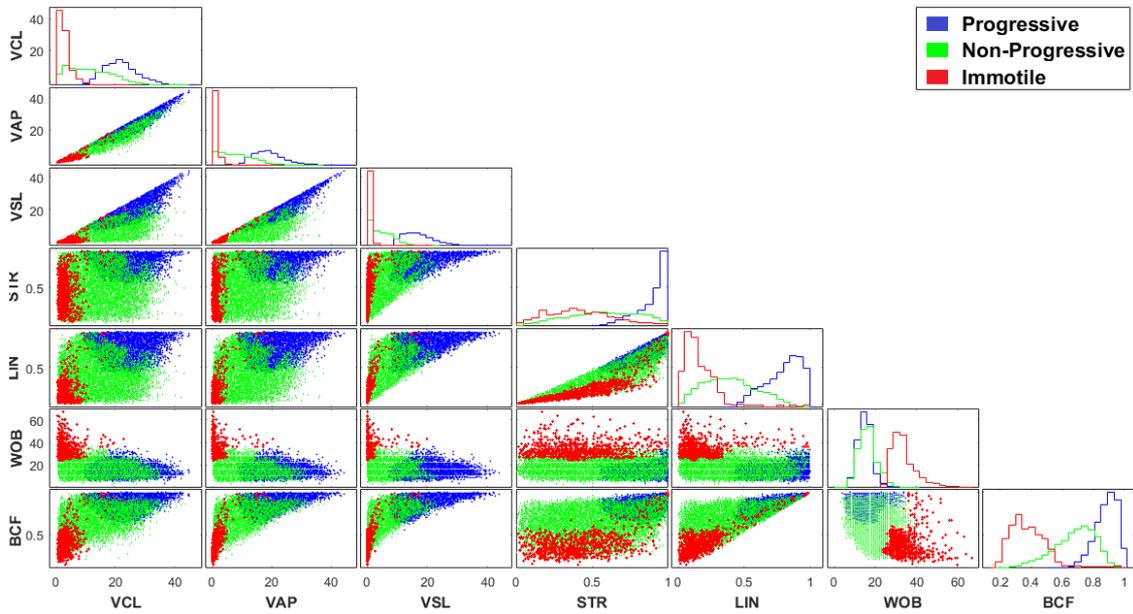


Figure 6.7 The distribution of the clusters over the extracted features

An example of the clustering results of each motilities over one specific patient is given in Figure 6.8. Grade A has rapid and fast straight characteristic as indicated in Figure 6.8 (a). Grade B is in linear format but the speed is lower than Grade A. Figure 6.8 (b) shows the Grade B trajectories recorded in the same duration with Grade A. The lengths of trajectories are shorter than Grade A which indicates the slower motilities than Grade A. In Figure 6.8 (c), inconsistent, elliptical movements can be observed as Grade C. Lastly, Grade D motilities are given in Figure 6.8 (d).

Grade A is the extremely fast motilities which cannot be observed frequently in the laboratories. Therefore, many CASA systems report the fast or rapid motility analysis in the normal progressive motility. In our system, we report separately. Hence, in the test results, our result were compared by one of the basic CASA systems by summing the Grade A and B in the progressive motility analysis.
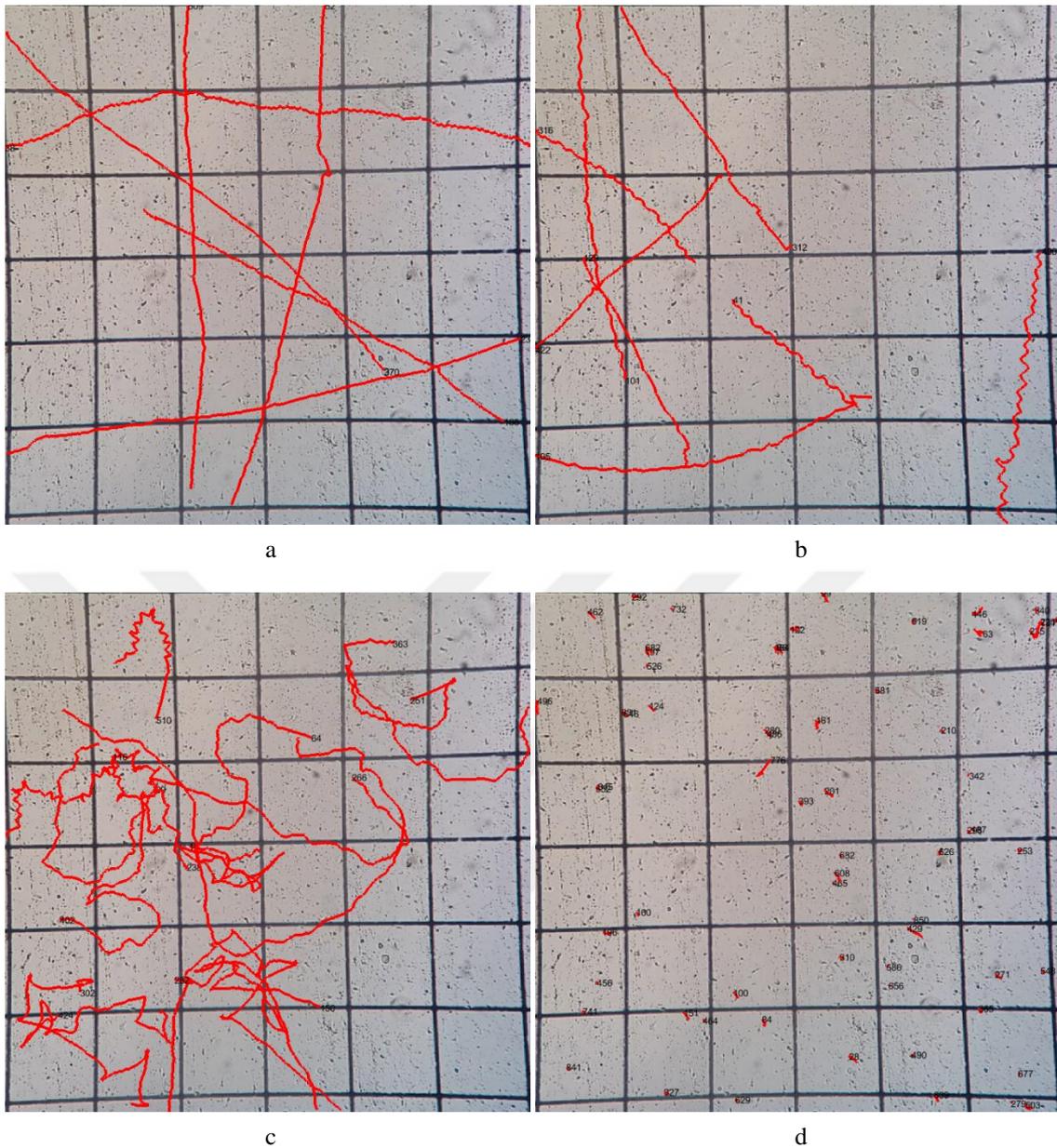
Figure 6.8 Motility Feature clustering over a subject a) Grade A – Fast progressive, b)
Grade B – Progressive, c) Grade C – Non-progressive, d) Grade D – Stable

After the verification of the proposed tracking technique, we evaluated the tracking
analysis in the diagnosis of 6 subjects by using the reference values as the threshold
values. The same semen samples were also tested in one of the basic CASA systems;
SQA-Vision. It is utilized in analysis and the results are accepted as the ground truth.
The comparative results are given in Table 6.2.

Table 6.2 Motility Analysis Results

| | | Total Spermatozoa | Motile Spermatozoa | Type of Motility (%) | | | |
|---|---|---|---|---|---|---|---|
| | | | | A | B | C | D |
| Subject 1 | V1 | 478 | 238 | 0 | 21.2 | 21.7 | 57 |
| | V2 | 523 | 221 | 0 | 17.0 | 20.8 | 62 |
| | V3 | 569 | 225 | 0 | 21.2 | 18.1 | 60.5 |
| | **Mean** | **523** | **228** | **0** | **19.9** | **20.2** | **59.8** |
| | **SQA-V** | **507** | **202** | **26** | | **8** | **66** |
| Subject 2 | V1 | 405 | 197 | 0.06 | 28.1 | 17.1 | 54.5 |
| | V2 | 438 | 209 | 0.19 | 32 | 17.4 | 50.2 |
| | V3 | 464 | 219 | 0.06 | 34.1 | 18.9 | 46.8 |
| | **Mean** | **435** | **208** | **0.11** | **31.4** | **17.8** | **50.5** |
| | **SQA-V** | **406** | **180** | **34** | | **11** | **55** |
| Subject 3 | V1 | 212 | 115 | 2.01 | 51.3 | 9.2 | 37.5 |
| | V2 | 245 | 122 | 1.57 | 48.7 | 11.01 | 38.7 |
| | V3 | 262 | 142 | 1.82 | 47.4 | 10.6 | 40.2 |
| | **Mean** | **240** | **126** | **1.8** | **49.1** | **10.3** | **38.8** |
| | **SQA-V** | **260** | **135** | **52** | | **5** | **43** |
| Subject 4 | V1 | 377 | 180 | 0 | 10.4 | 27.8 | 61.8 |
| | V2 | 321 | 174 | 0 | 8.7 | 34.3 | 57 |
| | V3 | 352 | 197 | 0 | 11.1 | 30.8 | 58.1 |
| | **Mean** | **350** | **184** | **0** | **10.1** | **30.9** | **58.9** |
| | **SQA-V** | **320** | **188** | **9** | | **30** | **61** |
| Subject 5 | V1 | 499 | 155 | 4.8 | 21.6 | 14.7 | 58.9 |
| | V2 | 524 | 172 | 6.1 | 19.9 | 12.4 | 61.6 |
| | V3 | 578 | 169 | 5.7 | 22.1 | 11.9 | 60.4 |
| | **Mean** | **533** | **165** | **5.5** | **21.2** | **13** | **60.3** |
| | **SQA-V** | **532** | **162** | **24** | | **7** | **69** |
| Subject 6 | V1 | 399 | 175 | 0.2 | 32.2 | 22.2 | 45.4 |
| | V2 | 404 | 192 | 0 | 34.7 | 18.2 | 47.1 |
| | V3 | 417 | 174 | 0 | 31.9 | 24,2 | 43.9 |
| | **Mean** | **406** | **180** | **0.06** | **32.9** | **21.5** | **45.4** |
| | **SQA-V** | **378** | **180** | **40** | | **8** | **52** |

SQA-Vision system reports three types of motilities as Progressive (Grade A + B), Non-Progressive (Grade C), and Immotile (Grade D). According to the results, system outputs are similar to the SQA-Vision reports in the detection of progressive and immotile movement types. Only the analysis for the high concentration samples is

different due to the challenges for detection process. But, SQA-Vision is also not a preferable CASA system in the laboratories because of the inefficiency in the low and/or high concentration samples. Therefore, results obtained in those cases can be neglected and should be compared with more advanced systems.

# CHAPTER 7

## CONCLUSION AND DISCUSSION

Spermiogram tests are currently carried out by Visual Assessment (VA) and Computer Assisted Sperm Analysis (CASA) systems in laboratories. VA is easy to implement and inexpensive, but highly dependent on the user experiences and expertise, which called as observer variability problem. Additionally, results might be misleading such as the diagnoses of asthenospermia in oligospermia cases because of the human factor. In order to eliminate the disadvantages of VA, CASA systems have been utilized. CASA systems have advantages over VA, but they are more complicated and expensive. Also the systems have dependencies and limitations. In this study, we developed a technique combining computerized systems and manual evaluations. We proposed a new computerized approach in which data acquisition is performed as in visual assessment, but analysis is done by a computer.

Today, Makler chambers are the most commonly used counting chambers in the visual assessment analysis. Therefore, computer-based systems for analyzing the Makler images will be an innovation. In this respect, videos of the samples over the Makler counting chamber were recorded from the ocular part of the microscopy by a mounted smartphone in the developed system. Then, videos transferred to the server to be analyzed in computer side software which is named as Computerized Sperm Counting and Tracking Analyzing System (CSCTAS). First phase of this software is the ROI detection and the extraction module. Hough transform (HT) with the clustering idea is initially tested for this aim. However, the results indicated that the technique should be improved. Then we proposed a combinational line detection approach. It was tested on 80 videos and the overall performance was evaluated with 95.36% and 88.58% accuracy and F-measure scores, respectively. Eventually, ROIs in Makler images were correctly extracted from images.

In our approach, we mainly aim to form a cheap and easily accessible sperm analyzing framework. Therefore, a holder accepting all mobile phones is designed and employed for mounting the phone on the ocular part of microscope. However, camera cannot be fully fixed on ocular part, hence, small vibrations can occur during the recording. Therefore, video stabilization technique is applied to prevent the counting results from the vibration effects that occur in the microscope. The main purpose is to provide more accurate and consistent results. Otherwise, employed motile spermatozoa detection algorithm results in more sperms in vibrated frames. We developed feature matching based video stabilization technique to decrease the vibrations on the videos. Three feature extraction techniques (SURF, FAST and BRISK) were employed to extract features of each frames. Then, each frame was rotated according to the match results of the features with the extracted background to stabilize the video sequence. In terms of counting vibrated frames, SURF is the most efficient algorithm with the minimum vibrations detected after the rotation process. On the other hand, FAST is the fastest algorithm in terms of processing times, while the SURF requires lots of time and system resources. BRISK increases the vibration of original video in several cases, hence BRISK is not suitable to use in stabilization algorithms. FAST is a moderate technique according to all performance metrics. However, counting of motile sperms is crucially important process. Small variations in result can have high affects in counting due to the generalization rule of Makler Chamber. Therefore, SURF is selected as the employed feature extraction technique in stabilization module. Each recorded video in the laboratories were initially stabilized by the SURF feature matching technique. Then the detection module for motile and immotile spermatozoa is performed over the stabilized videos.

Fuzzy c-means (FCM) and blob analysis-based segmentation were tested on each extracted and stabilized sub-images (a total of 36 sub-images from one frame) for the sperm-counting. Results were compared with the visual assessment technique. According to the visual assessment technique, results indicate that the proposed counting approach should be improved. Therefore, we utilized a hybrid method including active contours and dual thresholding in the spermatozoa detection step in the proposed CSCTAS.

32 video samples belonging to 8 subjects were tested for the spermatozoa detection by the proposed hybrid method. Four samples were requested from each subject in

different times to verify the results. Data organization was automatically performed by forming 30 seconds length sub-samples from the given main samples in the server. More extensive, objective and comprehensive counting analyses of the sperms are performed by the developed CSCTAS. Number of sperm cells were counted as immotile and motile spermatozoa according to the motilities inside the individual regions of interest. After applying the standardization calculation procedure of Makler, concentration results obtained by the VA and one of the basic CASA system, SQA, were compared with our approach (CSCTAS).

CASA is the most consistent analyzing technique according to the minimal standard deviation in the concentration results of identical subjects. CSCTAS employs the automatization of VA technique. Therefore, CSCTAS results are similar to the VA technique. However, CSCTAS is more consistent due to the less standard deviation in sample concentration analysis of the same subject. Additionally, Bland-Altman plot emphasized that the CSCTAS is in agreement limit with VA and CASA result which indicates that the proposed approach can be employed interchangeably with VA and CASA.

In the tracking module, α β Kalman Filter was used. Hungarian algorithm is utilized in the object assignments. Trajectories were split into 3 seconds length (90 frame) to avoid wrong assignments due to the immediate changes in motilities of the spermatozoa. Seven features from the trajectories were then extracted. As an initial work, we verified our system extracted features with the reference values of the each type of the motilities. In this respect, we merged all trajectories obtained from 32 videos of 8 subjects (4 videos of each subject including 3 healthy, 3 low motile, 2 non-motile cases). Totally 89438 trajectories were clustered into 4 clusters referring to the fast/rapid progressive, progressive, non-progressive, and stable. We checked the cluster centers of each feature with the reference values. Because of our cluster centers and the reference values are similar, our approach can be assumed as verified. Next, we tested our tracking approach on 18 videos of 6 subjects. Additionally, we analyzed the same semen sample in one of the CASA like system, SQA-Vision. According to the comparative results, our system is efficient in the determination of the fast and normal progressive motilities. However, system should improve in terms of the non-progressive motility analysis. According to the reference parameters for the type of the motility, non-progressive motility should be classified by only one feature; LIN. However, in our case, we realized that the extra

features should be appropriately included in the classification. Clustering results of the each features also indicates that the one feature for the clustering is not enough. In this respect, we will propose new parameter definitions for the motility analysis in future studies.

CSCTAS reported more detailed results in several cases and surpassed CASA in terms of being portability, cost and modularity. CSCTAS stands out with its hardware independency, implementation simplicity on any kind of phase-contrast microscopy, requirement of less parameters, and implementation and running cost. Software is not a hardware dependent solution as it is in the CASA. It only requires an extra camera apparatus. According to the obtained spermatozoa counting and tracking results, proposed system can be utilized instead of visual assessment technique and offers a computerized solution to spermatozoa analyzing much cheaper than the CASA systems and more reliable and observer invariable than the VA.

In further studies, we will define new parameters for determination of the motility types and also add morphological analysis capability to the system for the determination of the spermatozoa shapes.

# REFERENCES

[1]     Keel, B. A. and Webster, B. W, (1990). "Handbook of the laboratory diagnosis and treatment of infertility", CRC Press.

[2]     Roser, M. "Fertility". Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/fertility/. 27 Jan 2017.

[3]     World Health Organization (2010). WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction, Cambridge University Press, 5th Edition, Cambridge, England.

[4]     Carlsen, E., Giwercman, A., Keiding, N. and Skakkebæk, N. E., (1992). "Evidence for decreasing quality of semen during past 50 years", British Medical Journal, 305(6854):609-613.

[5]     Swan, S. H., Elkin, E. P. and Fenster, L., (2000). "The question of declining sperm density revisited: an analysis of 101 studies published 1934-1996", Environmental health perspectives, 108(10):961.

[6]     Grace, M. C., (2014). "Semen Assessment", Urologic Clinics of North America, 41(1):163-167.

[7]     Makler, A., (1980). "The improved ten-micrometer chamber for rapid sperm count and motility evaluation", Fertility and Sterility, 33(2):160-166.

[8]     Smith, J. T. and Mayer, D. T., (1955). "Evaluation of sperm concentration by the hemocytometers method: Comparison of four counting fluids", Fertility and Sterility, 6(3):271–275.

[9]     Granger, G. A. and Kolb, W. P., (1968). "lmphocyte in vitro cytotoxicity: Mechanisms of immune and non-immune small lymphocyte mediated target cell destruction", The Journal of Immunology, 101(1) 111–120.

[10]    Mahmoud, A. M., Depoorter, B., Piens, N. and Comhaire, F. H., (1997). "The performance of 10 different methods for the estimation of sperm concentration", Fertility and sterility, 68(2):340–345.

[11]    Björndahl, L., Mortimer, D., Barratt, C. L., Castilla, J. A., Menkveld, R., Kvist, U. and Haugen, T. B., (2010). "A practical guide to basic laboratory andrology". Cambridge University Press, 1st Edition, Cambridge, England.

[12]    Amann, R. P. and Waberski, D., (2014). "Computer-assisted sperm analysis (CASA): capabilities and potential developments", Theriogenology, 81(1) 5–17.

[13]    Rittscher, J., Machiraju, R. and Wong, S. T., (2008). "Microscopic image analysis for life science applications", Artech House, Norwood, MA.

[14]    Olalla, O. G., Alegre, E., Robles, L. F., Malm, P. and Bengtsson, E., (2015). "Acrosome integrity assessment of boar spermatozoa images using an early fusion of texture and contour descriptors", Computer Methods and Programs in Biomedicine, 120(1):49-64.

[15]    Alegre, E., Castro, V. G., Rodríguez, R. A. and García-Ordás, M. T., (2012). "Texture and moments-based classification of the acrosome integrity of boar spermatozoa images", Computer Methods and Programs in Biomedicine, 108(2):873-881.

[16]    Alegre, E., Biehl, M., Petkov, N. and Sanchez, L., (2013). "Assessment of acrosome state in boar spermatozoa heads using n-contours descriptor and RLVQ", Computer Methods and Programs in Biomedicine, 111(3):525-536.

[17]    Khachane, M. Y., Manza, R. R. and Ramteke, R. J., (2015). "Fuzzy rule based classification of human spermatozoa", International Conference on Electrical, Electronics, Signals, Communication and Optimization, 24-25 January 2015, Visakhapatnam, India.

[18]    Liu, J., Leung, C., Lu, Z. and Sun, Y., (2013). "Quantitative Analysis of Locomotive Behavior of Human Sperm Head and Tail", IEEE Transactions on Biomedical Engineering, 60(2):390-396.

[19]    Chang, V., Saavedra, J. M., Castañeda, V., Sarabia, L., Hitschfeld, N. and Härtel, S., (2014). "Gold-standard and improved framework for sperm head segmentation", Computer Methods and Programs in Biomedicine, 117(2):225-237.

[20]    Abbiramy, V. S., Shanthi, V. and Allidurai, C., (2010). "Spermatozoa detection, counting and tracking in video streams to detect asthenozoospermia", International Conference on Signal and Image Processing (ICSIP), 15-17 December 2010, Chennai, India.

[21]    Li, Q., Chen, X., Zhang, H., Yin, L., Chen, S., Wang, T. and Zhang, R., (2012). "Automatic human spermatozoa detection in microscopic video streams based on OpenCV", International Conference on Biomedical Engineering and Informatics (BMEI), 16-18 October 2012, Chongqing, China.

[22]    Karabiber, F., Pabuccuoglu, S. and Arik, S., (2009). "The Fast Moving Cell Counting System using Analog CNN and DSP Microprocessors", International Symposium on Innovations Intelligent Systems and Applications (INISTA), 29 June – 1 July 2009, Trabzon, Turkey.

[23]    Ghasemian, F., Mirroshandel, S. A., Azad, S. M, Azarnia, M. and Zahiri, Z., (2015). "An efficient method for automatic morphological abnormality detection from human sperm images", Computer Methods and Programs in Biomedicine, 122(3):409-420.

[24]    Möller, M., Burger, M., Dieterich, P. and Schwab, A., (2014). "A framework for automated cell tracking in phase contrast microscopic videos based on normal velocities", Journal of Visual Communication and Image Representation, 25(2):396-409.

[25] Oku, H., Ishikawa, M., Ogawa, N., Shiba, K. and Yoshida, M., (2008). "How to track spermatozoa using high-speed visual feedback", 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS) 21-24 August 2008, Vancouver, Canada.

[26] Savkay, O. L. and Yalcin, M. E., (2012). "Analysis of sperm motility with CNN architecture", 13th International Workshop on Cellular Nanoscale Networks and Their Applications (CNNA), 29-31 August 2012, Turin, Italy.

[27] Boyers, S. P., Davis, R. O. and Katz, D. F., (1989). "Automated semen analysis", Year Book Medical Publishers.

[28] Gillan, L., Evans, G. and Maxwell, W. M. C., (2005). "Flow cytometric evaluation of sperm parameters in relation to fertility potential", Theriogenology, 63(2):445-457.

[29] SQA Automated Sperm Quality Analyzers - Accurate and Rapid Assessments of Male Fertility: https://mes-global.com, 01 Jan 2018.

[30] SCA® CASA System: http://micropticsl.com, 01 Jan 2017.

[31] Rodzen, C., Clinical Sperm Analysis Products: https://hamiltonthorne.com, 01 Jan 2017.

[32] Lu, J. C., Huang, Y. F. and Lü, N. Q., (2014). "Computer-aided sperm analysis: past, present and future", Andrologia, 46(4):329-338.

[33] Sherins, R. J., (1991). "Clinical use and misuse of automated semen analysis", Annals of the New York Academy of Sciences, 637(1):424-435.

[34] Mortimer, S. T., (1997). "A critical review of the physiological importance and analysis of sperm movement in mammals", Human Reproduction Update, 3(5):403-439.

[35] Davis R. O. and Katz, D. F., (1992). "Standardization and comparability of CASA instruments," Journal of Andrology, 13(1):81–86.

[36] Boryshpolets, S., Kowalski, R. K., Dietrich, G. J., Dzyuba, B. and Ciereszko, A., (2013). "Different computer-assisted sperm analysis (CASA) systems highly influence sperm motility parameters", Theriogenology, 80(7):758-765.

[37] Kathiravan, P., Kalatharan, J., Karthikeya, G., Rengarajan, K. and Kadirvel, G., (2011). "Objective Sperm Motion Analysis to Assess Dairy Bull Fertility Using Computer-Aided System – A Review", Reproduction in Domestic Animals, 46(1): 165-172.

[38] Feitsma, H., Broekhuijse, M. L. W. J. and Gadella, B. M., (2011). "Do CASA systems satisfy consumers demands? A critical analysis", Reproduction in Domestic Animals, 46(s2):49-51.

[39] Witkowski, L., (2013). "A computer system for a human semen quality assessment", Biocybernetics and Biomedical Engineering, 33(3):179-186.

[40] Wilson-Leedy, J. G. and Ingermann, R. L., (2007). "Development of a novel CASA system based on open source software for characterization of zebrafish sperm motility parameters", Theriogenology, 67(3):661-672.

[41]     Shi, L. Z., Nascimento, J. M., Chandsawangbhuwana, C., Botvinick, E. L. and Berns, M. W., (2008). "An automatic system to study sperm motility and energetics", Biomedical microdevices, 10(4):573-583.

[42]     Masood, S., Sharif, M., Masood, A., Yasmin, M. and Raza, M., (2015). "A Survey on Medical Image Segmentation", Current Medical Imaging Reviews, 11(1):3-14.

[43]     McLean, G., Prescott, B. and Kotturi, D., (1993). "Hierarchical clustering for automated line detection", Communications, Computers and Signal Processing IEEE Pacific Rim Conference, 19-21 May 1993, Victoria BC, Canada.

[44]     Schuster, G. M. and Katsaggelos, A. K., (2003). "Robust line detection using a weighted MSE estimator", Proceedings of the International Conference on Image Processing (ICIP), 14-17 September 2003, Barcelona, Spain.

[45]     Herout, A., Dubská, M. and Havel, J., (2013). "Review of hough transform for line detection", Real-Time Detection of Lines and Grids, Springer-Verlag, London, U.K., 3-16.

[46]     Saha, T. B., Tchiotsop, D., Tchinda, R. and Kenne, G., (2015). "Automated Extraction of the Intestinal Parasite in the Microscopic Images Using Active Contours and the Hough Transform", Current Medical Imaging Reviews, 11(4): 233-246.

[47]     Li, W. C. and Tsai, D. M., (2011), "Defect Inspection in Low-Contrast LCD Images Using Hough Transform-Based Nonstationary Line Detection", IEEE IEEE Transactions on Industrial Informatics, 7(1): 136-147.

[48]     Hayashi, N., Tomizawa, T., Suehiro, T. and Kudoh, S., (2012). "A robust white line detection technique for double circular operator", IEEE International Conference on Mechatronics and Automation (ICMA), 5-8 August 2012, Chengdu, China.

[49]     Ahmad, T., Bebis, G., Nicolescu, M., Nefian, A. and Fong, T., (2015). "An Edge-Less Approach to Horizon Line Detection", IEEE 14th International Conference on Machine Learning and Applications (ICMLA), 9-11 December 2015, Miami, USA.

[50]     Lie, W., Lin, T. C. I., Lin, T., Hung, K., (2005). "A robust dynamic programming algorithm to extract skyline in images for navigation", Pattern Recognition Letter, 26(2):221-230.

[51]     Chen L, Wang L and Xiong J., (2010). "Straight-Line Intersection Point Detection Based on Curve Fitting of Hough Parameter Space", International Conference on e-Education, e-Business, e-Management and e-Learning, 22-24 January 2010, Sanya, China.

[52]     Lee, J., Han, J. and Whang, K., (2007). "Trajectory clustering: a partition-and-group framework", Proceedings of the 2007 ACM SIGMOD international conference on Management of data, 11-14 June 2007, New York, NY, USA.

[53]     Piciarelli, C., Foresti, G. L., (2006). "On-line trajectory clustering for anomalous events detection", Pattern Recognition Letter, 27(15):1835-1842.

[54]     Wu, S., Weinstein, S. P., Conant, E. F., Schnall, M. D. and Kontos, D., (2013). "Automated chest wall line detection for whole-breast segmentation in sagittal breast MR images", Medical physics, 40(4).

[55]    Uyen, T. V. N., Alauddin, B., Laurence, A. F. P. and Kotagiri, R. "An effective retinal blood vessel segmentation method using multi-scale line detection", Pattern Recognition, 46(3):703-715.

[56]    Dave, R. N. and Bhaswan, K., (1992). "Adaptive fuzzy c-shells clustering and detection of ellipses", IEEE Transaction Neural Network, 3(5):643-662.

[57]    Rajesh, N.D., (1989). "Use of the Adaptive Fuzzy Clustering Algorithm to Detect Lines in Digital Images", Symposium on Visual Communications, Image Processing and Intelligent Robotics Systems International Society for Optics and Photonics, 01 November 1989, Philadelphia, PA, USA.

[58]    Behrenbruch, C. P., Petroudi, S., Bond, S., Declerck, J. D., Leong, F. J. and Brady, J. M., (2004). "Image filtering techniques for medical image post-processing: an overview", British Journal of Radiology, 77:126-132.

[59]    Karuppathal, R. and Palanisamy, V., (2014). "An Automotive Approach for Brain Tumor Segmentation Based on Gaussian distribution and Level Set Method", Current Medical Imaging Reviews, 10(4):290-296.

[60]    Meenakshi, R. and Anandhakumar, P., (2015). "A Hybrid Brain Tumor Classification and Detection Mechanism Using Knn and Hmm", Current Medical Imaging Reviews, 11(2):70-76.

[61]    Matsopoulos, K. G., Economopoulos, L. T., (2013). "A Review of Five Automatic Point Correspondence Methods for Application on Medical Images", Current Medical Imaging Reviews, 9(4):263-282.

[62]    Erdem, C. E, Tekalp, A. M, Sankur, B., (2001). "Metrics for performance evaluation of video object segmentation and tracking without ground-truth", Proceedings of the International Conference on Image Processing, 7-10 October 2001, Thessaloniki, Greece.

[63]    İlhan, H. O. and Karabiber, F., (2016). "Computer Based Grid Detection on Makler Images", International Journal of Signal Processing Systems, 4:323-327.

[64]    İlhan, H. O. and Karabiber, F., (2017). "Automatic Detection of Regions of Interest in Makler Images by Combinational Approach and Sperms Analysis by Fuzzy C-Means", Current Medical Imaging Reviews, 13:90-98.

[65]    Von-Gioi, R. G., Jakubowicz, J., Morel, J. M. and Randall, G., (2010). "LSD: A fast line segment detector with a false detection control", IEEE transactions on pattern analysis and machine intelligence, 32(4):722-732.

[66]    Stauffer, C. and Grimson, W. E. L., (1999). "Adaptive background mixture models for real-time tracking", Computer Vision and Pattern Recognition, 2:246-252.

[67]    Hartigan, J. A. and Wong, M. A., (1979). "Algorithm AS 136: A k-means clustering algorithm", Journal of the Royal Statistical Society. Series C (Applied Statistics), 28(1):100-108.

[68]    Illingworth, J. and Kittler, J., (1988). "A survey of the Hough transform", Computer vision, graphics, and image processing, 44(1):87-116.

[69]    Jain, R., Kasturi, R. and Schunck, B. G., (1995). Machine vision, New York: McGraw-Hill.

[70] Kwon, O., Shin, J. and Paik, J., (2005). "Video stabilization using Kalman filter and phase correlation matching", In International Conference Image Analysis and Recognition (pp. 141-148), Springer, Berlin, Heidelberg.

[71] Tico, M. and Vehvilainen, M., (2005). "Constraint motion filtering for video stabilization", IEEE International Conference in Image Processing, ICIP 2005, 11-14 September 2005, Genova, Italy.

[72] Matsushita, Y., Ofek, E., Tang, X. and Shum, H. Y., (2005). "Full-frame video stabilization", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), 20-26 June 2005, San Diego, CA, USA.

[73] Piva, S., Zara, M., Gera, G. and Regazzoni, C. S., (2003). "Color-based video stabilization for real-time on-board object detection on high-speed trains", IEEE Conference on Advanced Video and Signal Based Surveillance, 21-22 July 2003, Miami, FL, USA.

[74] Auberger, S. and Miro, C., (2005). "Digital video stabilization architecture for low cost devices", 4th International Symposium on Image and Signal Processing and Analysis (ISPA 2005), 15-17 September 2005, Zagreb, Croatia.

[75] Liang, Y. M., Tyan, H. R., Chang, S. L., Liao, H. Y. and Chen, S. W., (2004). "Video stabilization for a camcorder mounted on a moving vehicle", IEEE Transactions on Vehicular Technology, 53(6):1636-1648.

[76] Oreifej, O., Li, X. and Shah, M., (2013). "Simultaneous video stabilization and moving object detection in turbulence", IEEE transactions on pattern analysis and machine intelligence, 35(2):450-462.

[77] Smith, B. M., Zhang, L., Jin, H. and Agarwala, A. (2009). "Light field video stabilization", IEEE 12th International Conference on Computer Vision, 29 September – 2 October 2009, Kyoto, Japan.

[78] Liu, S., Wang, Y., Yuan, L., Bu, J., Tan, P. and Sun, J. (2012). "Video stabilization with a depth camera", IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), 16-21 June, Providence, RI, USA.

[79] Battiato, S., Puglisi, G. and Bruna, A. R. (2008). "A robust video stabilization system by adaptive motion vectors filtering", IEEE International Conference on Multimedia and Expo, 23-26 June, Hannover, Germany.

[80] Hu, R., Shi, R., Shen, I. F. and Chen, W. (2007). "Video stabilization using scale-invariant features", 11th International Conference in Information Visualization (IV'07), 4-6 July, Zurich, Switzerland.

[81] Battiato, S., Gallo, G., Puglisi, G. and Scellato, S. (2007). "SIFT features tracking for video stabilization", 14th International Conference on Image Analysis and Processing (ICIAP 2007), 10-14 September 2007, Modena, Italy.

[82] Yang, J., Schonfeld, D., Chen, C. and Mohamed, M. (2006). "Online video stabilization based on particle filters", IEEE International Conference on Image Processing, 8-11 October 2006, Atlanta, GA, USA.

[83] Shen, Y., Guturu, P., Damarla, T., Buckles, B. P. and Namuduri, K. R., (2009). "Video stabilization using principal component analysis and scale invariant feature transform in particle filter framework", IEEE Transactions on Consumer Electronics, 55(3).

[84]    Pinto, B. and Anurenjan, P. R., (2011). "Video stabilization using speeded up robust features", 2011 International Conference on Communications and Signal Processing, 10-12 February 2011, Kerala, India.

[85]    Xia, Y., Gao, Q., Cheng, N., Lu, Y., Zhang, D. and Ye, Q., (2017). "Denoising 3-D magnitude magnetic resonance images based on weighted nuclear norm minimization", Biomedical Signal Processing and Control, 34:183-194.

[86]    Aghajani, K., Yousefpour, R., Shirpour, M. and Manzuri, M. T., (2016). "Intensity based image registration by minimizing the complexity of weighted subtraction under illumination changes", Biomedical Signal Processing and Control, 25:35-45.

[87]    Sauvée, M., Noce, A., Poignet, P., Triboulet, J. and Dombre, E., (2007). "Three-dimensional heart motion estimation using endoscopic monocular vision system: From artificial landmarks to texture analysis", Biomedical Signal Processing and Control, 2(3):199-207.

[88]    Piccardi, M., (2004). "Background subtraction techniques: a review", IEEE International Conference on Systems, Man and Cybernetics. 10-13 October 2004, The Hague, Netherlands.

[90]    Bay, H., Tuytelaars, T. and Van Gool, L., 2006. "Surf: Speeded up robust features", European conference on computer vision, Springer, 404-417, Berlin, Heidelberg.

[91]    Zhangh, L. and Liu, Y., (2015). "Improved binary robust invariant scalable key-points algorithm fusing depth information", Journal of Computer Applications, 35(8):2285-2290.

[92]    Viswanathan, D. G., (2009). Features from accelerated segment test (fast): https://wenku. baidu.com/view/af9743d725c52cc58ad6beac.html, 01.01.2018.

[93]    Anbumozhi, S., (2016). "Performance Analysis of Brain Tumor Detection based on Fuzzy Logic and Neural Network Classifier", Current Medical Imaging Reviews, 12(4):304-312.

[94]    Eltoukhy, M. M., Farag, A. K. and Abdelwahed, N. M. A., (2017). "Fusion of Wavelet and Morphological Features for Breast Cancer Diagnosis in Ultrasound Images", Current Medical Imaging Reviews, 12(4):290-297.

[95]    Mageshwari, V., Aroquiaraj, I. L. and Dharani T., (2016). "A realistic medical image segmentation of abnormal HIV in blood cells using morphological operations", International Conference on Inventive Computation Technologies (ICICT), 2016, Coimbatore, India.

[96]    AL-Dulaimi, K., Banks, J., Tomeo-Reyes, I. and Chandran, V., (2016). "Automatic segmentation of HEp-2 cell Fluorescence microscope images using level set method via geometric active contours", 23rd International Conference on Pattern Recognition (ICPR), 2016, Cancun, Mexico.

[97]    Bijar, A., Mikaeili, M., Benavent, A. P. and Khayati, R., (2012). "Segmentation of sperm's Acrosome, nucleus and mid-piece in microscopic images of stained human semen smear", 8th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP), 2012, Poznan, Poland.

[98]    Ilhan, H. O. and Elbir, A., (2016). "The evaluation of detectors and descriptors on determination of semen cell", International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), 2016, Sinaia, Romania.

[99]    Baâzaoui, A., Barhoumi, W., Ahmed, A. and Zagrouba, E., (2017). "Semi-Automated Segmentation of Single and Multiple Tumors in Liver CT Images Using Entropy-Based Fuzzy Region Growing", Inovation and Research in Biomedical Engineering, 38(2):98-108.

[100]   Bezdek, J. C., Ehrlich, R. and Full, W., (1984). "FCM: The fuzzy c-means clustering algorithm", Computers & Geosciences, 10(2-3):191-203.

[101]   Freire-Gormaly, M., Ellis, J. S., Bazylak, A. and MacLean, H. L., (2015). "Comparing thresholding techniques for quantifying the dual porosity of Indiana Limestone and Pink Dolomite", Microporous and Mesoporous Materials, 207:84-89.

[102]   Chan, T. F. and Vese, L. A., (2001). "Active contours without edges", IEEE Transactions on image processing, 10(2):266-277.

[103]   Chudnovsky, D. V. and Chudnovsky, G. V., (2004). "Approximations and complex multiplication according to Ramanujan", In Pi: A Source Book, Springer, New York, NY.

[104]   Cheung, S. C. S. and Kamath, C., (2005). "Robust background subtraction with foreground validation for urban traffic video", EURASIP Journal on Advances in Signal Processing, 2005: 726261

[105]   Dewitte, K., Fierens, C., Stöckl, D. and Thienpont, L. M., (2002). "Application of the Bland–Altman plot for interpretation of method-comparison studies: a critical investigation of its practice", Clinical chemistry, 48(5):799-801.

[106]   Blackman, S. and Popoli, R., (1999). "Design and Analysis of Modern Tracking Systems", Artech House Radar Library.

[107]   Chenouard, N., Smal, I., De Chaumont, F., Maška, M., Sbalzarini, I. F., Gong, Y. and Cohen, A. R., (2014). "Objective comparison of particle tracking methods", Nature methods, 11(3):281.

[108]   Katz, D. F., Davis, R. O., Delandmeter, B. A. and Overstreet, J. W., (1985). "Realtime analysis of sperm motion using automatic video image digitization," Computer Methods and Programs in Biomedicine, 21(3):173-182.

[109]   Beresford-Smith B. and Helden D. V., (1994). "Applications of radar tracking algorithms to motion analysis in biomedical images", International Conference on Image Processing, 1994.

[110]   Shi, L. Z., Nascimento, J., Berns, M. W. and Botvinick, E. L., (2006). "Computer-based tracking of single sperm", Journal of biomedical optics, 11(5):054009.

[111]   Tomlinson, M. J., Pooley, K., Simpson, T., Newton, T., Hopkisson, J., Jayaprakasan, K. and Pridmore, T., (2010). "Validation of a novel computer-assisted sperm analysis (CASA) system using multitarget-tracking algorithms", Fertility and sterility, 93(6):1911-1920.

[112] Su, T. W., Xue, L. and Ozcan, A., (2012). "High-throughput lensfree 3D tracking of human sperms reveals rare statistics of helical trajectories", Proceedings of the National Academy of Sciences, 109(40):16018-16022.

[113] Berezansky, M., Greenspan, H., Cohen-Or, D. and Eitan, O., (2007). "Segmentation and tracking of human sperm cells using spatio-temporal representation and clustering", In Medical Imaging: Image Processing International Society for Optics and Photonics, 6512:65122M.

[114] Sørensen, L., Østergaard, J., Johansen, P. and Bruijne, M., (2008). "Multi-object tracking of human spermatozoa", In Medical Imaging: Image Processing, International Society for Optics and Photonics, 6914:69142C.

[115] Bar-Shalom, Y., Daum, F. and Huang, J., (2009). "The probabilistic data association filter," IEEE Control Systems Magazine, 82–100.

[116] Ristic, B., Vo, B. N., Clark, D. and Vo, B. T., (2011). "A metric for performance evaluation of multi-target tracking algorithms", IEEE Transactions on Signal Processing, 59(7):3452-3457.

[117] Comaniciu, D. and Meer, P., (2002). "Mean shift: A robust approach toward feature space analysis", IEEE Transactions on pattern analysis and machine intelligence, 24(5):603-619.

[118] Ristic, B., Arulampalam, S. and Gordon, N., (2003). "Beyond the Kalman filter: Particle filters for tracking applications", Artech house.

[119] Johnson, S. C., (1967). "Hierarchical clustering schemes", Psychometrika, 32(3):241-254.

# CURRICULUM VITAE

## PERSONAL INFORMATION

**Name Surname**           : Hamza Osman İLHAN

**Date of birth and place** : 05.06.1987 - Nallıhan

**Foreign Languages**      : English

**E-mail**                 : hamza.ilhan@gmail.com

## EDUCATION

| Degree | Department | University | Date of Graduation |
|--------|-----------|-----------|--------------------|
| Master | Computer Engineering | Yalova University | 2012 |
| Undergraduate | Computer Education | Marmara University | 2010 |

## WORK EXPERIENCE

| Year | Corporation/Institute | Enrollment |
|------|----------------------|-----------|
| 2012 - Present | Yıldız Technical University | Research Assistant |
| 2011 - 2012 | Middle East Technical Univ. | Research Assistant |

## PUBLISHMENTS

### Papers

1. Ilhan, H.O., Aydin, N. "A novel data acquisition and analyzing approach to spermiogram tests", Biomedical Signal Processing and Control, 41, 129-139, 2018

2. Ilhan H.O., Bilgin G. "Sleep stage classification via ensemble and conventional machine learning methods using single channel EEG signals", International Journal of Intelligent Systems and Applications in Engineering, 5(4):174-84, 2017

3. İlhan H.O., Karabiber F., "Automatic Detection of Regions of Interest in Makler Images by Combinational Approach and Sperms Analysis by Fuzzy C-Means", Current Medical Imaging Reviews, vol.13, pp.1-1, 2017

4. İlhan H.O., Karabiber F., "Computer Based Grid Detection on Makler Images", International Journal of Signal Processing Systems, no.4, pp.323-327, 2016

5. İlhan H.O., Amasyali M.F., "Active Learning as a Way of Increasing Accuracy", International Journal of Computer Theory and Engineering, vol.6, no.1, pp.460-465, 2014

6. İlhan H.O., Amasyali M.F., "Comparing Informative Sample Selection Strategies in Classification Ensembles", International Journal of Machine Learning and Computing, vol.4, pp.79-84, 2014

7. Esenkanova J., İlhan H.O., Yavuz S., "Pre-Mapping System with Single Laser Sensor Based on Gmapping algorithm", International Journal of Electrical Energy, vol.1, pp.97-101, 2013

### Conference Papers

1. Çelik E., İlhan H.O., Elbir A., "Detection and Estimation of Down Syndrome Genes by Machine Learning Techniques", Sinyal işleme ve iletişim uygulamaları kurultayı, İSTANBUL, TÜRKIYE, 15-18 Mayıs 2017, pp.1-1

2. Elbir A., İlhan H.O., "Churn Analizinde Sınıflama ve Kümeleme Tekniklerinin Uygulanması", Akıllı Sistemlerde Yenilikler ve Uygulamaları Konferansı, ANTALYA, TÜRKIYE, 5-7 Ekim 2017, ss.1-5

3. İlhan H.O., Aydin N., "Bilgisayar Destekli Hareketli Sperm Sayımı", Tıp Teknolojileri Kongresi'17, TRABZON, TÜRKIYE, 12-14 Ekim 2017, pp.1-5

4. İlhan H.O., Aydin N., "The Contribution of DSP Integration to ARM Cores in SBCs for the Video Decoding Process", International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017), Chennai, HINDISTAN, 21-22 Eylül 2017, pp.1-5

5. İlhan H.O., Avcı C., "Sleep Stage Classification by Ensemble Learning Methods with Active Sample Selection Techniques", International Artificial Intelligence and Data Processing Symposium'17, MALATYA, TÜRKIYE, 16-17 Eylül 2017, pp.1-5

6. İlhan H.O., Elbir A., "The evaluation of detectors and descriptors on determination of semen cell", International Symposium on INnovations in Intelligent SysTems and Applications, Sinaia, ROMANYA, 2-5 Ağustos 2016, pp.1-4

7. İlhan H.O., Çelik E., "The Mesothelioma Disease Diagnosis with Artificial Intelligence Methods", 10th IEEE International Conference on Application of Information and Communication Technologies, Bakü, AZERBAYCAN, 12-14 Ekim 2016, pp.837-840

8. Delibaşoğlu İ., Kösesoy İ., İlhan H.O., Çetin Ö., "A Comparative Study of Different IHS Transformation Models with a Tradeoff Parameter", Seoul International Conference on Applied Science and Engineering, , Kasım 2014

9. İlhan H.O., Elçiçek S., Buldu A., Savaş K., "A web based controlled greenhouse system and monitoring the system's data", Marmara University, , Nisan 2010


**Projects**

1. BAP 2016-04-01-DOP01

2.