Computerized Evaluatution of Microsurgery Skills Training

Payal Jotwani
University of Central Florida

COMPUTERIZED EVALUATION OF MICROSURGERY SKILLS TRAINING

by

PAYAL JOTWANI
Bachelor's in Electronics and Telecommunication, Pune University, 2010

A thesis submitted in partial fulfillment of the requirements
for the degree of Master of Science
in the Department of Electrical Engineering and Computer Science
in the College of Engineering and Computer Sciences
at the University of Central Florida
Orlando, Florida

Summer Term
2015

Major Professor: Hassan Foroosh

ABSTRACT

The style of imparting medical training has evolved, over the years. The traditional methods of teaching and practicing basic surgical skills under apprenticeship model, no longer occupy the first place in modern technically demanding advanced surgical disciplines like neurosurgery. Furthermore, the legal and ethical concerns for patient safety as well as cost-effectiveness have forced neurosurgeons to master the necessary microsurgical techniques to accomplish desired results. This has lead to increased emphasis on assessment of clinical and surgical techniques of the neurosurgeons. However, the subjective assessment of microsurgical techniques like micro-suturing under the apprenticeship model cannot be completely unbiased. A few initiatives using computer-based techniques, have been made to introduce objective evaluation of surgical skills.

This thesis presents a novel approach involving computerized evaluation of different components of micro-suturing techniques, to eliminate the bias of subjective assessment. The work involved acquisition of cine clips of micro-suturing activity on synthetic material. Image processing and computer vision based techniques were then applied to these videos to assess different characteristics of micro-suturing viz. speed, dexterity and effectualness. In parallel subjective grading on these was done by a senior neurosurgeon. Further correlation and comparative study of both the assessments was done to analyze the efficacy of objective and subjective evaluation.

For Mom and Sir.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS/ABBREVIATIONS

ASM	Active Shape Model
CT	Computed Tomography (X-ray)
GMM	Gaussian Mixture Model
JPDAF	Joint Probability Data Association Filter
MAS	Minimal Access Surgery
MHT	Multiple Hypothesis Tracking
MIS	Minimally Invasive Surgery
MRI	Magnetic Resonance Imaging
OR	Operation Room
OSATS	Objective Structured Assessment of Technical Skills
PCA	Principal Component Analysis
RGB	Red Green Blue Color Model
STFT	Short Time Fourier Transform
YUV	Color Model based on Luminance and Chrominance Components

1. INTRODUCTION

The human central nervous system with its intricate neural network, is the most advanced structure known in the universe. It is characterized by complex and intricate anatomy involving the nerves, blood vessels, and deep seated and inaccessible areas [1]. Thus, Neurosurgical procedures pose series of challenges for the neurosurgeons, requiring detailed knowledge of micro-neuro-anatomy and well-versed surgical skills [2]. Owing to the complex anatomy, the margin of error in these surgical procedures is very low. Even a minor surgical error made by the surgeon can lead to irreversible catastrophes like paralysis and death [1,2].

Traditional skills training in neurosurgery is imparted by 'Halstedian apprenticeship method'. In this set up, the trainee is allowed to initially learn by observation and then followed by hands-on experience on human brain (patient) under the supervision of senior neurosurgeon. But, these methods of imparting surgical education and training are constantly being challenged by concerns of patient safety, increased operation room time and cost-ineffectiveness [11]. This suggests the use of alternative methods of instruction for MIS training [5-10].

Moreover, the apprenticeship model subjects the trainees progress to the evaluation by a senior surgeon which cannot be completely unbiased and does not give a component-wise objective feedback indicating distinguished deficit skills. Such a supervised evaluation in a monitored/constrained environment adds to the anxiety of the trainee resulting in poor performance. Teaching, developing, and practicing basic surgical skills via the apprenticeship model of training are no longer considered appropriate in the operating room [12].

The medical fraternity has been confronted with challenges of keeping pace with evolving

technologies and changes in healthcare system [4]. A variety of newer methods to prepare trainees for the operating room have been implemented in recent years, due in large part to the fact that surgical procedures are changing from open surgery to minimal access surgery (MAS) [or minimally invasive surgery (MIS)], a practice requiring new suites of technical skills. Also, iatrogenic errors made by the trainees have drawn increasing attention to the surgical techniques and dexterity of the surgeons, which require elaborate and effectual training [12].

1.1   Background

1.1.1   <u>Neurosurgery Skills Training</u>

In the current state of the art, neurosurgical procedures require high competence of different surgical skills such as endoscopy, microsurgery, high speed drilling and spinal instrumentation [2]. A major proportion of the neurosurgical procedures involve microsurgical techniques, which are complex and require significant level of precision owing to the fine anatomy of the central nervous system.

In neurosurgical procedures, microsurgical skills/techniques are predominantly demonstrated in surgical tasks like dural repair, nerve anastomosis and blood vessel anastomosis, where the tissues being worked on are roughly of the thickness/diameter varying between 0.5 to 3.0 mm [12]. As per the paper published in 1979 [20], the minimum diameter of vessel considered for operation is 0.3 mm in microsurgery. This constraint was found to be guided by limitations of low magnifying optic aid, large suture material, human-hand motor functions like dexterity and degrees of freedom.

The several reasons that contribute to microsurgical procedures being demanding are as follows:

1. The microscope magnifies, and amplifies the small movements into larger displacements when seen from the eye-piece in all dimensions.

2. It limits the visual field.

3. Physiological tremor becomes more prominent/apparent, in-proportion to the length of the surgical instrument.

4. The movements in the direction of the depth are equally altered, which have slower cognitive learning curve as compared to planar movements [3].

Also, long operating hours for the surgeon in the same posture and limited amount of movements can add to his fatigue and affect his performance adversely.

Therefore, in order to regain visual-motor coordination in the new magnified habitat and acquire microsurgical techniques, needs good amount of practice in the environment/field and getting used to, and is not a skill that can be acquired by observation. Also, a continual practice is necessary even after the skill is gained to prevent the loss of dexterity [2,3].

1.1.2 <u>Skills Training and Evaluation: Micro-neurosurgery</u>

Efforts are being made to address the constraints of microsurgery and stretch the limits of operative cases, such as development of finer sutures, higher-magnification microscopes, finer instrument/micro-manipulators and robotic manipulators [3]. However, the key area needing attention still remains surgical skills and dexterity of the neurosurgeons that require elaborate and effectual training as well as continuous practice.

Various institutions in the world have begun microsurgical skills training courses using simulator, synthetic material or training sessions on small laboratory animals to allow the trainees to hone their surgical skills. This training is imparted under variety of schemes, such as several consecutive weeks/months of gradual advancement through the modules or short annual weekend workshops. The non-clinical, surgical skills training methods can be broadly divided into the following four categories:

- Skills Training on Computerized Modules or Virtual Reality Simulators
- Skills Training on Synthetic Modules
- Skills Training on Cadaver Modules

- Skills Training on Animal Modules

Skills training and practice sessions on each of these modules have their own advantages and disadvantages that are still being elucidated in various validation processes [21].

1.1.2.1 Skills Training on Computerized Modules or Virtual Reality Simulators

Simulation based systems have gained popularity over the last 20 years. Simulation may be defined as the imitation of the operation of a real-world process or system over time [22], thus providing an opportunity to learn and experience in a riskless situation. Its broad applicability is exemplified in neuro-endoscopy training. Neuro-endoscopy is a popular surgical choice for the management of intra-ventricular lesions and the treatment of hydrocephalus. Neuro-endoscopic techniques were predominantly demonstrated on cadaver sections as they are the closest one can get preserved structural neuro-anatomy without working with patients [23,24]. However, cadavers do not recreate a realistic semblance for ventriculoscopic work due to lack in the physical tissue properties of tissues such as consistency, color and fluidity. Adding to this high cost of embalming, preservation and storage of cadavers, and also lack of availability in various countries [23,24], have necessitated the adoption of newer methods.

The challenges in development of surgical simulators is to provide a good semblance of the experience of various pathologies, while having standardized training milestones that allow a gradual progression through different modules increasing in the level complexity judged by objective assessments of performance [25].

NeuroTouch (shown in Figure 1.1.2-1), being one of the most popular and advanced virtual reality neurosurgery simulator with haptic feedback, used for imparting skills training and assessment aided by tactile and visual cues. NeuroTouch is a product of collaboration between National Research Council Canada and medical teaching hospitals across Canada. Its primary components are a stereovision system, bimanual haptic tool manipulators, and a high-end computer. The software engine of this simulator runs on 3 processes for computing graphics, haptics, and mechanics. Training tasks are built from MRI scans of patients suffering with different pathologies. The training tasks modeled by this system allow the user to practice surgical skills with 4 different surgical tools. This system has been developed for three different surgical tasks.

- Brain tumor resection: Where the main objective is to execute complete removal of the vascularized tumor tissue without removing non-tumor tissues, using the regular surgical aspirator (suction), the ultrasonic aspirator, and bipolar electro cautery.

- Endoscopic transnasal procedures: Where training scenario involves navigating with an endoscope inside a nasal cavity to find the ostium of the sphenoid sinus while observing the cavity and anatomy through endoscopic view.

- Endoscopic Third Ventriculostomy: Where the user navigates through first and third ventricles to select a target perforation [26-28].

Figure 1.1.2-1 - NeuroTouch: Virtual Reality Neurosurgery Simulator

Another interesting virtual reality neurosurgical simulator is ImmersiveTouch (Fig 1.1.2-2), which facilitates the surgeon in interacting with a 3D computer-generated virtual anatomical model using haptic device. The simulator provides visual, tactile, and audio feedback of a surgical procedure. This system builds the 3D anatomical model cases using the pathologies observed in patient's CT and/or MRI data. The system facilitates the user by allowing him to view relevant fluoroscopic shots if required to give a closer semblance to OR capabilities. This system also has additional 3D monitors so that the trainee can discuss or share different procedural steps with other trainees or seniors. The ImmersiveTouch and MicrovisTouch simulators have been developed for open and percutaneous procedures, and Micro-surgery procedures.

The ImmersiveTouch simulator incorporates:

- A two-handed instrument replicates

- Surgery specific foot pedals

- Fluoroscopy and ultrasound guidance (during procedure steps)

- Endoscopic views (when needed)

- Real time surgical effectiveness and accuracy metrics

- iPad for controls, visualization software and apps [32].

Figure 1.1.2-2 - ImmersiveTouch: Virtual Reality Neurosurgery Simulator

However, the virtual reality simulators are still in the development phase, and require a huge initial investment [29,30]. Also, the computational burden often results in lack of fluidity of real-time interactive simulators due to digital piecewise rendering. The

production of realistic tactile sensory feedback remains unachieved, due to cost requirements. Moreover, the current systems lack the expected semblance to live tissues [31].

1.1.2.2 Skills Training on Synthetic Modules

According to some researchers, synthetic physical simulators are more reliable, effective, and comparatively more cost-efficient as compared to virtual simulators, human cadavers, and animal models [33-38]. Moreover, the synthetic surgical simulators for neurosurgery have better semblance of live tissue as compared to virtual simulators. These designs can also offer repetitive practice sessions, while allowing the user to perform image-guided navigation using CT and MRI data [39,40].

One of the portable synthetic module based training tool is Microvascular Practice Card. It allows repeated practicing of microvascular anastomosis procedure in various situations. This non-animal practice tool helps trainees practice under safe and hygienic conditions that reduces the need of laboratory animals during technical training and can be used in dry labs for practice purposes. This module allows trainees to hone their skills by gradual advancement through tube diameters reducing in diameter from 1.0 mm to tubes with 0.5 and 0.3 mm diameter [43].
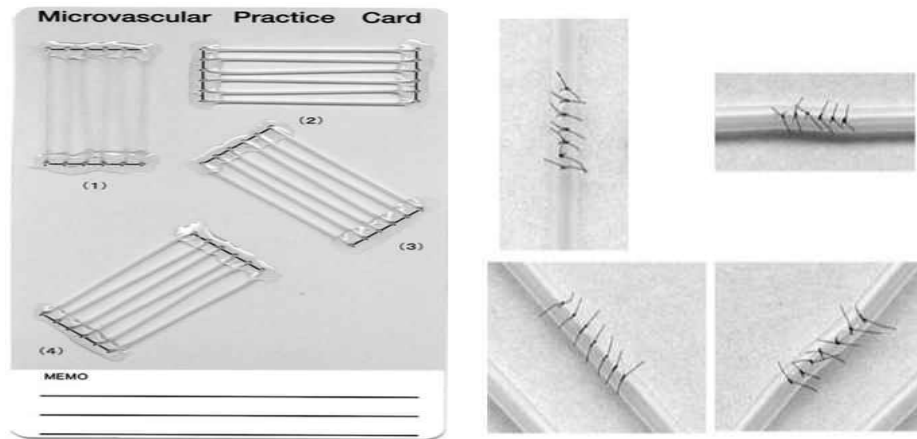
Figure 1.1.2-3 - Microvascular Practice Card

A very important advantage with these modules is that they have much better semblance of OR environment, due to high-fidelity haptic and visual feedback [41]. One of the developments of synthetic surgical simulators involved, a brain silicone replica mimicking normal mechanical properties of a 4-month-old child with hydrocephalus pathology, encased in a replicated skull, and immersed in water. As claimed by the paper, the intraventricular landmarks including the choroid plexus, veins, mammillary bodies, infundibular recess, and basilar artery demonstrating bleeding scenarios, give a much realistic semblance of live tissues. The thinned-out third ventricle floor in this model is replaceable, making this model reusable. Standard neuroendoscopic equipment including irrigation is used [42].

1.1.2.3 Skills Training on Cadaver Modules

As discussed earlier, the cadaver based modules offer the closest structural semblance to the minutiae's of Neuro-anatomy, but due to lack of the physical properties in the tissues such as consistency, color and fluidity, they fail to provide satisfactory experience for surgical skills training. Also, the added cost of embalming, preservation and storage of cadavers, and also lack of availability in various countries [23,24], make them inadequate for surgical skills training. But, they are definitely are a good subject to impart educative lessons and lectures on structural neuro-anatomy using different slices of the brain and spinal cord [1].

1.1.2.4 Skills Training on Animal Modules

Hands-on skills training on small laboratory animals provides trainees the closest semblance to surgical scenario as they are dealing with tissues. The neurosurgical skills training under this module involves nerve and vessel anastomosis procedures performed on femoral artery and sciatic nerve of the anaesthetized albino rats [1]. But, establishment of such laboratories require various ethics committees permissions. Thus, not every institute can establish such a laboratory.

The vitro perfusion module for microvascular anastomosis training is a technique that falls in the category of animal modules and synthetic modules. This module is developed by extracting arteries and veins of 1 mm in diameter from the chicken wings. For the explanation the vessels were cannulated at both ends and mounted on a

platform. These vessels were then supplied with blood obtained from blood banks injected through the proximal catheter, using an automatic perfusor to give a natural semblance of a vascular anastomosis procedure. This is very simple and reliable module facilitating repeated use of low cost materials. Also, vessels explanted from the human placenta or earthworms may be used in a similar manner to develop such modules as supplement for microvascular tissues [44,45]. The discussion above related to the non-clinical surgical skills training methods, is summarized in Table 1.2.

Moreover, the impact of hands-on micro-surgical skills training has been evident in several studies [13,14]. The assessments on synthetic material or anaesthetized small laboratory animal, are mostly conducted by a senior surgeons, thus trainee's performance is still subject to bias and his progress through modules and his evaluation is collectively subjective. OSATS is one popular method that creates a structured way of analyzing the microsurgical task to make evaluation by a senior surgeon more reliable and consistent [15-18]. OSATS stands for Objective Structured Assessment Of Technical Skills; it follows the different parameters to assess the surgical performance shown in Table 1.1, where each parameter is graded on a scale of 5 [19]. One of the drawbacks of this technique is that even though the grading is comparatively more reliable, it is still subject to human errors and undeniably varies with the bias of the evaluator.

Table 1-1 - Global Scale of Objective Structured Assessment of Technical Skills (OSATS)

Respect for tissue	1	2	3	4	5
	Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments		Careful handling of tissue but occasionally caused inadvertent damage		Consistently handled tissues appropriately with minimal tissue damage, no rough handling
Time and motion	1	2	3	4	5
	Many unnecessary moves		Efficient time/motion but some unnecessary moves		Clear economic movement and maximum efficiency
Instrument handling	1	2	3	4	5
	Repeatedly makes tentative or awkward moves with instruments		Competent use of instruments although occasionally appeared still and awkward		Fluid moves with instruments and no awkwardness
Knowledge of instruments	1	2	3	4	5
	Frequently asked for wrong instrument or used inappropriate instrument		Knew names of most and used appropriate instrument		Obviously familiar with the instruments required and their names
Flow of operation	1	2	3	4	5
	Frequently stopped operating and seemed unsure of the next move		Demonstrated some forward planning with reasonable progression of procedure		Obviously planned course of operation with effortless flow from one move to the next
Knowledge of specific procedure	1	2	3	4	5
	Deficient knowledge, needed specific instruction at most steps		Knew all important steps of operation		Demonstrated familiarity with all the aspects of operation
Quality of anastomosis	1	2	3	4	5
	Crossed and tangled suture material, sloppy knots, imprecise and inconsistent suture placement		Minimal tangling, mostly square knots, good suture placement and consistency with some variability		Excellent suture material

Table 1-2 - Non-clinical Skills Training Methods

Skills Training Method	Semblance to Surgical Scenario	Re-Usability/ Repetition	Initial Cost	Maintenance Cost	Patient Specific Cases	Visual and Haptic Feedback	Best suited for
Computerized Modules or Virtual Reality Simulators	Not as good, as most of the products are still in the development phase.	Yes	Very High	Low	Possible	Not good.	Surgical skills practice
Synthetic Modules	Much better than Virtual Modules. But less explored domain.	For Some	Lower than other options	Higher, as repetition requires replacement	Possible but harder.	Quite good	Evaluation and surgical skills training
Cadaver Modules	Good structural semblance, but not tissue properties.	No	High	High cost of embalming and storage.	Possible but tougher to obtain.	Not good.	Education on Neuro-anatomy
Animal Modules	Very good	Once or twice	Moderate	High cost of laboratory maintenance	Not possible	Very good	Surgical skills training and practice

1.1.3   Object Segmentation

Background subtraction is a widely used approach to distinguish and successively detect moving foreground objects from the background. The basic approach involves detecting the moving objects from the difference between the current frame and a reference frame (often called the "background image", or "background model"). One of the basic requirements of this approach to understand the background elements is representation of the scene with certain number of frames with no moving objects. In order to make the subtraction more robust to varying luminance conditions and geometry settings, these must regularly be updated [46].

Several techniques for performing background subtraction have been proposed, which estimate the background model successfully from temporal sequence of frames.

1.1.3.1   Running Gaussian average

This technique proposes to model the background independently at each pixel (i, j) location by ideally fitting a Gaussian probability density function on the last 'n' pixel values. In order to avoid the overhead of fitting the probability density function from scratch for every new frame time, t, a running (or on-line cumulative) average is computed instead as per equation 1.1.

$$\mu_t = \alpha I_t + (1 - \alpha)\, \mu_{t\text{-}1} \tag{1.1}$$

where,

I_t = pixel's current value

μ_t = previous average

α = empirical weight (stability and quick update)

The standard deviation σ for each pixel can also be computed. This technique has two parameters (μ_t, σ_t) for each pixel, in place of the buffer for last n pixel values. One can classify a particular I_t at time frame t as foreground pixel if it satisfies equation 1.2.

$$| I_t - \mu_t | > k \, \sigma_t \tag{1.2}$$

This, running average computation results in reducing the complexity as well as the memory requirement [47]. Koller et al. [48] modification to this scheme suggested modification in the model given in equation 1.3,

$$\mu_t = M \mu_{t-1} + (1-M)(\alpha I_t + (1 - \alpha) \, \mu_{t-1}) \tag{1.3}$$

where,

      M = 1 if foreground value

         = 0 otherwise

Since, the input values required for this algorithm is only the intensity values of pixels, this technique can accept inputs in form of any of the color models (RGB or YUV) for multiple components or a single channel gray scale images. The update rate for the algorithm can be set, lower the update rate of the background model, the slower the system will be able to adapt to the actual background dynamic [46].

1.1.3.2 Temporal median filter

Temporal median filtering is another well-known technique for background subtraction. The basic fundamental of this technique is to use the median value of the last n frames as the background model [49]. To further increase the stability of the background model, the median should be computed on a special set of values containing the last n, sub-sampled frames and w times the last computed median value. One of the major disadvantages of the median-based approach is that its computation requires a buffer overhead with the recent pixel values. Moreover, the median filter does not accommodate for a rigorous statistical description and does not provide a deviation measure for adapting the subtraction threshold [50].

1.1.3.3 Mixture of Gaussians

This technique uses the knowledge that, over time different background objects are likely to appear at a same (i,j) pixel location, particularly when this is due to a permanent change in the scene's geometry. In such a scenario, above-mentioned background subtraction models will promptly adapt to reflect the value of the current background object. But, sometimes the changes in the background object are not permanent; rather change at a rate faster than the background update. In such cases a single-valued background is not appropriate for the classification purpose. (Example is that of an outdoor scene with trees partially covering a building)

Gaussian mixture model proposes a case of multi-valued background model that is more robust to such variations due to multiple background objects.

This is a probability-based approach of observing a certain pixel value, x, at time t by means of a mixture of Gaussians given in equation 1.4.

$$P(x_t) = \sum_{i=1}^{K} \omega_{i,t} \, \eta \, (x_t - \mu_{i,t} \, \Sigma_{i,t}) \tag{1.4}$$

Where, for each value of the K, Gaussian distributions are to describe only one of the observable background or foreground objects.

When K is set to a value between 3 and 5, gaussians are multi-variate to describe red, green and blue values. If these values are assumed independent, the co-variance matrix, simplifies into a diagonal matrix. In addition, if the standard deviation for the three channels is assumed the same, it further reduces to a simpler problem, $\sigma^2_i I$.

In order to define the discrimination between the foreground and background distributions, first all the distributions are ranked based on the ratio between their peak amplitude, ω_i and the standard deviation, σ_i.

Then the pixels are classified based on the knowledge, that higher and more compact the distribution is, higher is the likelihood of it belonging to the background.

Using this scheme, the first B distributions in the ordered ranking satisfying with a threshold T, are considered as the background (refer equation 1.5).

$$\sum_{i=1}^{B} \omega_i > T \text{ are classified as background pixels} \tag{1.5}$$

Therefore, at each t frame time, two problems are simultaneously solved

- Assigning the new observed value x_t to the best matching distribution
- Estimating the updated model parameters

These concurrent problems can be solved by an expectation maximization (EM) algorithm working on the buffer of the last n frames [51,52].

1.1.4   Object Tracking

The goal of the object-tracking algorithm is to generate the trajectory of an object over time, by locating its position in every video frame, and establishing correspondence between the object instances across frames. There are two different ways to this. First, the possible object regions in every frame can be obtained by means of object detection, and then the tracker can correspond to these individual sets of locations of the objects across frames. And second, the object region and correspondence locations are estimated in an iterative fashion by updating object location and region information obtained from previous frames [53]. Broad classification of tracking techniques can be as seen in the Fig 1-4 [53].
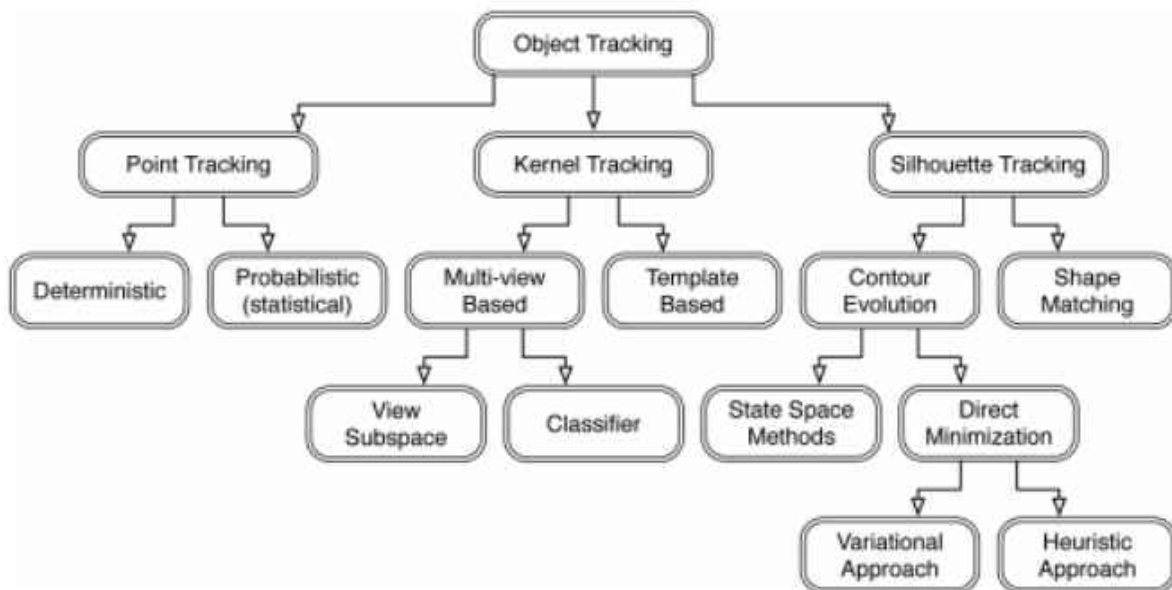
Figure 1-4 - Tracking taxonomy

Point-based Tracking: An object is either represented by a point or by a set of points, such as the centroid or certain interest points. Point correspondence observed between frames using such a technique suffers in the presence of occlusions, misdetections, entries, and exits of objects. This approach can be further classified into two broad categories, namely, deterministic and statistical methods. In the deterministic approach the correspondence problem is constrained by the use of qualitative motion heuristics. While in the probabilistic approach, object measurement is taken explicitly as well as the uncertainties are taken into account to establish correspondence. In general, the point representation is suitable for tracking objects that occupy small regions in an image [54,55]. Deterministic approaches proposed by different papers demonstrate the deployment of distinguished constraint schemes, such proximity, rigidity, maximum velocity, smallest velocity, common motion, temporal coherency, etc. [53,58,59].

Points detected on objects in video sequences are invariably affected by noise. Even the object motions can undergo sudden/random changes. Incorporation of statistics in correspondence methods can improve the accuracy of these tracking problems by taking into account not just the detection but the possible uncertainties. Different approaches that work under such a scheme using state space model with the detections fall under the category of probabilistic point-based tracking techniques. This category of object tracking has had applications in not just tracking but also in the computation of shape-from-motion [60], activity recognition [61], object identification [62], contour tracking [63], etc. One of the popular probabilistic point based tracking technique using single object state estimation, Kalman Filtering is detailed below.

1.1.4.1 Kalman Filter

Kalman filter is a technique that works on the assumption that the state distribution is

Gaussian, to estimate the state of a linear system. This technique can be broadly divided into two components, prediction and correction. The prediction component of the algorithm uses the state model to find an estimation of the position of the object in the current frame is given by the equation 1.6 and 1.7.

$$p\,X^t = DX^{t-1} + W \tag{1.6}$$

$$p\Sigma^t = D\,\Sigma^{t-1}D^T + Q^t \tag{1.7}$$

where,

pX^t = state prediction at time t

$p\Sigma^t$ = covariance prediction at time t

D = state transition matrix, determining the relation between the states at

time t and t-1

Q = covariance factor of white noise

Whereas, the correction component of the algorithm uses the detections in the current frame to update the state of the object as given by equation 1.8, equation 1.9 and equation 1.10.

$$K^t = p\Sigma^t M^T [Mp\Sigma^t M^T + R^t]^{-1} \tag{1.8}$$

$$X^t = pX^t + K^t [Z^t - MpX^t] \tag{1.9}$$

$$\Sigma^t = p\Sigma^t - K^t Mp\Sigma^t \tag{1.10}$$

Where,

$Z^t - \mathbf{Mp}X^t$ = innovation at time t

M = measurement matrix

K = Kalman gain, used for propagation of state models

Over time, various adaptations of Kalman filter algorithm for different applications [80,81].

Primitive geometric shapes/Kernel Tracking: Object shapes can be represented by known regular shapes such as rectangle and ellipse. Motion detection for such object representations can be modeled by translation, affine, or projective homography-based transformations. This technique using primitive geometric shapes, can be used for representing both simple rigid objects and non-rigid objects [53,56].

Silhouette Tracking: This is an efficient technique apt for tracking and recognizing non-rigid objects, using either silhouette or contour representations or both [57]. Objects with complex shapes such as human body with flailing parts(hands, head, and shoulders) cannot be well described by simple geometric shapes. Silhouette- based methods provide a much accurate description for these objects. The goal of a silhouette-based object tracker is to find the object region in each frame by means of an object model generated using the previous frames. These models can be described using color histogram, object edges, object contour or combinations of them [64,65]. Silhouette trackers further divided into two categories, shape matching and contour tracking. Shape matching techniques are based on finding object silhouette in each frame, whereas, contour tracking techniques, work with evolving an initial contour into its new position at each frame [53]. Active Shape Models are one of successful and popularly used technique for object fitting using contour based.

1.1.4.2 Multiple Objects

In a successful multiple object tracking system, we need the algorithm to consistently model the background and all moving objects explicitly tracked. But the motion object tracking models do not take into account the interaction between the moving objects. When tracking multiple objects using Kalman or particle filters, one needs to deterministically associate the different observations in consecutive frames to the particular object and its state. This is a correspondence problem that needs to be addressed before the application of filters. The simplest method to perform correspondence is to use the nearest neighbor approach, except when objects are close to each other. This problem can even lead the filter to fail to converge. There exist several statistical data association techniques to tackle this problem. Joint Probability Data Association Filtering (JPDAF) and Multiple Hypothesis Tracking (MHT) are two widely used statistical data association techniques to solve this problem [66]. MHT works on the conclusion that there is a significant chance of an incorrect correspondence, if tracking correspondence is established using only two frames. Therefore, it tries to establish correspondence decision over detections in several frames, based on several correspondence hypotheses for each object at each time frame [53, 67].

1.1.5 Object Fitting and Modelling

Unlike non-rigid object tracking, quite a bit of work has been done in the field of rigid object tracking and fitting. However, in many problems there is need to model and track the non-rigid object over time. A popular deformable model technique used to solve this type of problem is

Snakes. This technique and its revised variants have gained significant attention and are popularly used for the segmentation of non-rigid 2D and 3D objects [68-74]. But different problems associated with snakes techniques such as initialization sensitivity, lead to unacceptable results. These problems were overcome by a technique called Active Shape Models, by incorporation prior training based on the knowledge of the shapes. Moreover, this technique relies upon a statistical variation possibility is the shapes of a given class of object as presented by the training set [75,76]. A trained ASM model moves towards the convergence of the object shape, given an initial location estimate of the object in a new frame. With each iteration this technique moves each contour point to a better position in the frame and tries to adjust the shape and location of the points in the object contour to fit the model to the object in the new frame. Cootes proposed a method to generate a model of shape and appearance of 2D images objects. Both the papers [77] and [78] have shown the successful results with application of this technique to models with objects exhibiting natural shape, lighting conditions and pose variations.

1.2   Lacunae of Knowledge

I.   Globally there is no validated skills training curriculum incorporated under the doctorate and post-doctorate education and training program for Neurosurgery. A radical change in the training pattern is observed as the need of the hour, as neurosurgery operating rooms (ORs) are now under continuous surveillance of expert, social and legal cameras. And various factors such as work hour restrictions, patient safety concerns, governmental policies, patient loads in ORs, increasing costs and rapidly changing technologies have caused a trend towards minimally invasive surgeries.

Due to the increase in iatrogenic error made by the novice neurosurgeons have drawn increasing attention to surgical skills, necessitating the development of a standardized evaluation and training system for neurosurgery skills training.

II.   There is not much progress made in the field of objective component-wise analysis of neurosurgical skills training and evaluation. Even though, some virtual reality based systems have been developed for certain neurosurgical procedural training, they are in developing phase and lack the representation of the live human surgical field, in terms of semblance of live tissue interaction and feedback both visual and haptic.

1.3   Objective

The objective of this thesis/research work is the development of an algorithm for the evaluation of characteristics micro-neurosurgical skills. As a micro-suturing activity comprehensively encapsulates almost the complete gamut skills required for micro-neurosurgery techniques, the assessment of the skills/performance is done by the application of image segmentation, tracking and object fitting algorithms on the cine clips of the surgical performance of the trainee.

Under the component-wise evaluation scheme we propose, the assessment of microsurgical techniques should be guided by the following characteristics

- Dexterity

- Effectualness

- Speed

- Eye-hand co-ordination

- Instrument-tissue manipulation

For the scope of this Master's thesis, I have limited my work to the evaluation of only the first 3 characteristics.

Such a component-wise evaluation feedback shall help the trainee, in understanding his strengths and weakness', and in turn help improvise their surgical techniques individually and collectively.

1.4   Organization of Thesis

We have comprehensively covered the different state-the-art developments in the field of neurosurgery skills training, and advancements in the field of object tracking and localization. The rest of this thesis is organized into the following chapters. Chapter 3 covers methodology used by our algorithm and results at various intermediate steps. Chapter 4 presents the results detailing the evaluation of the efficacy of different parts of the algorithms deployed. Chapter 5 concludes with reflections on the currently used methods and proposition of future work in the domain.

2. METHODOLOGY

After exploring and studying existing skills training modules and considering the various factors presented in Table 1.1, it was established that the training and evaluation of basic microsurgical skill for neurosurgery should be conducted on synthetic modules under operating microscope for the following reasons:

- Better semblance of human tissue properties: Visual and Haptic

- Use of actual surgical instruments

- Lesser deviance from surgical environment

- Possibility of re-usability of modules

- Reduced cost

- Avoiding potential harm to live tissues patients or small laboratory animals

The research work for this thesis began with acquisition of cine clips of the microsurgical activity performed by the trainee neurosurgeon on synthetic material using silk and monofilament sutures of varied thickness. The synthetic material used in this acquisition varies in thickness 0.4 to 2 mm and gives haptic feedback similar to that of human Dura (protective covering of brain and spinal cord).

2.1   Acquisition Process

This acquisition process is done using a two separate high-end 3-chip cameras installed at the ends of two ocular paths of the operating microscopes, that help us in capturing left and right

frames individually as shown in Figure 2.1-1. The two ocular channels are extracted by the use of an attachment called the beam splitter as shown in Figure 2.1-2.

Figure 2.1-1 - Operating Microscope with a beam splitter, adapter and cameras for the acquisition purpose

Figure 2.1-2 - Beam Splitter for Operating Microscope

The operating microscope used has the capability of manual adjustments to modify the magnification factors. The recordings are done at suitable magnifications for different suture thickness as given by Table 2.1-1. The magnification number to magnification factor conversion is given by the formula.

$$\text{Magnification Factor} = \frac{\text{Focal length of Binocular tube x Magnification of eye pieces x magnification number}}{\text{Focal length of objective}}$$

Where,

Focal length of the tube = 170mm

Magnification of the eye piece = 10 x or 12.5 x

Magnification number = 0.4 x – 2.0 x

Focal length of the objective lens= 85mm for OPMI Pico specifications.

Table 2.1-1 - Recording Parameters for Basic and Intermediate Micro-suturing

Basic and Intermediate Micro-suturing training on Synthetic Material			
Suture	Material	Magnification Number	Magnification Factor
4-0	Silk	0.4	2.83
		0.6	4.25
5-0	Silk	0.4	2.83
		0.6	4.25
7-0	Monofilament	1.0	7.08
		1.6	11.33
9-0	Monofilament	1.6	11.33
		2.5	17.71
10-0	Monofilament	1.6	11.33
		2.5	17.71

2.2 System Design: Dexterity and Speed

Dexterity is characteristic for micro-suturing procedure, which can be measured by checking the tremors of the instrument. The central nervous system is marked by highly intricate anatomy and the thickness of tissues the neurosurgeons operate with are of the diameter as fine as 0.3 mm i.e. the margin of error is as small as 0.3mm. Therefore, the existence of such a tremor in the hand of neurosurgeon can lead to unwanted morbid consequences. The primary factors influencing this characteristic of surgeon's skill are the lack of visual motor coordination in the magnified environment, level of confidence and level of anxiety. This factor can primarily be noticed by observing the instrument movements before the actual surgical manipulation activity begins.

The problem of analyzing two characteristics of the micro-suturing activity, viz. Speed and Dexterity, can be done by the blocks of algorithmic units shown in Figure 2.2-1.
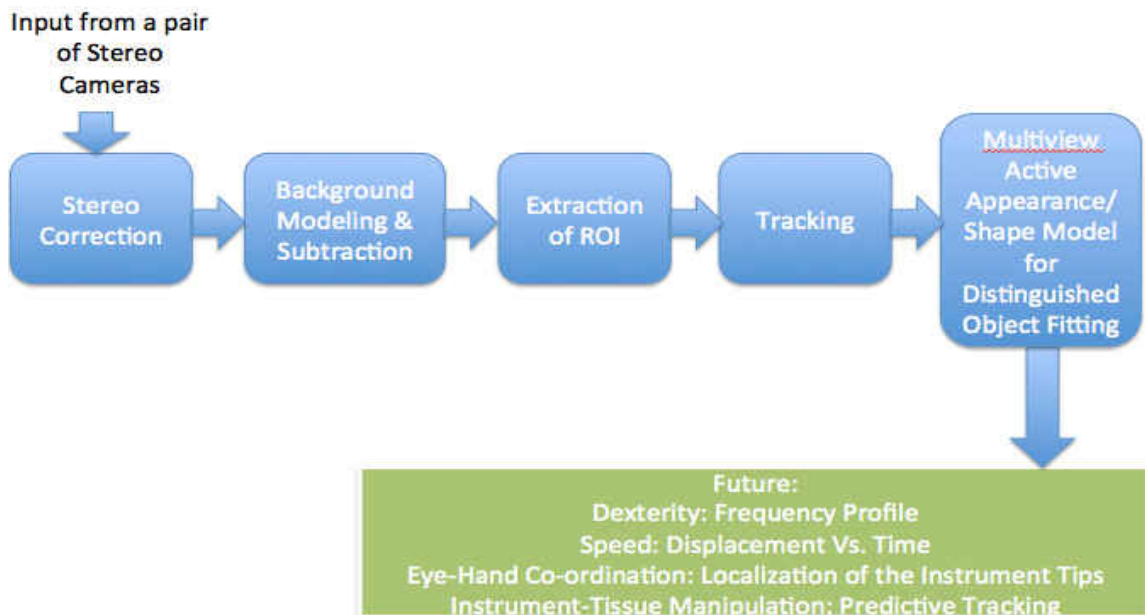
Figure 1.1.5-1 - Basic Block Diagram for Speed and Dexterity Computation

Lets go through all these blocks and their respective algorithms one by one.

2.2.1 **Stereo Correction Block**

After the acquisition procedure of the left eye and right eye views was obtained. It was observed that there existed minor translation and rotation based distortion between the two views. Also this system had inherent parallax i.e. there is a difference in the apparent position of the object or points, when observed from two different view-points (refer Figure). Thus, this problem was identified to fall under the domain of plane-induced parallax. Such a problem can be solved by, finding homography between the two views, and then warping one of the image obtained by one view to the other. This is guided by equation 2.1.
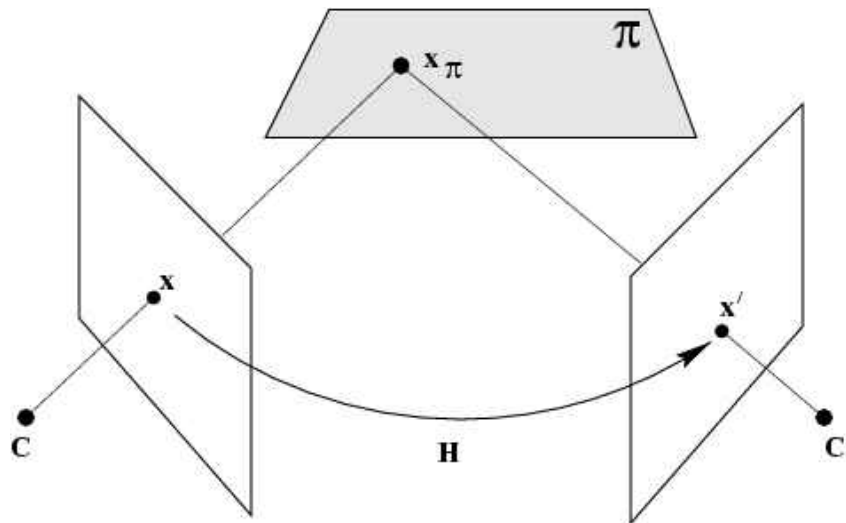
Figure 2.2.1-1 - Plane Induced Parallax

x' = Hx   (2.1)

where,

  x=apparent position of the point as observed from the left camera

  x'=apparent position of the point as observed from the right camera

  H=Homography between the two views

The following algorithm details the steps deployed to perform the stereo correction.

Algorithm

Step 1: Read the stereoscopic images of a planar surface used for microsurgical activity as shown in Figure 2.2.1-2
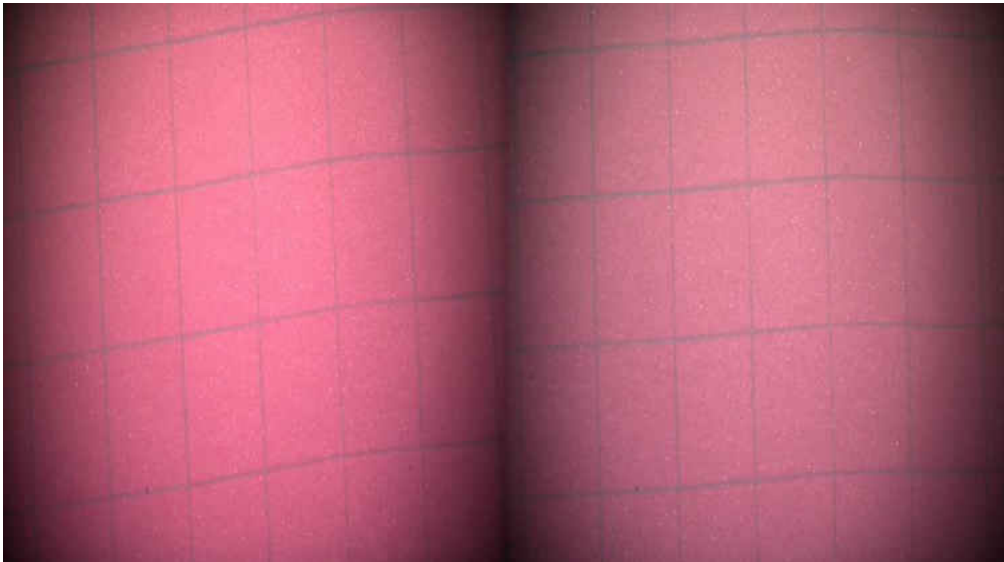
Figure 2.2.1-2 - Stereoscopic frame

Step 2: Separate the stereoscopic frame to read the left and right frames separately as shown in Figure 2.2.1-3 and 2.2.1.4 and find the SIFT features on both the frames respectively, as shown in Figure 2.2.1-5 and 2.2.1.6

Figure 2.2.1-3 - Left Image
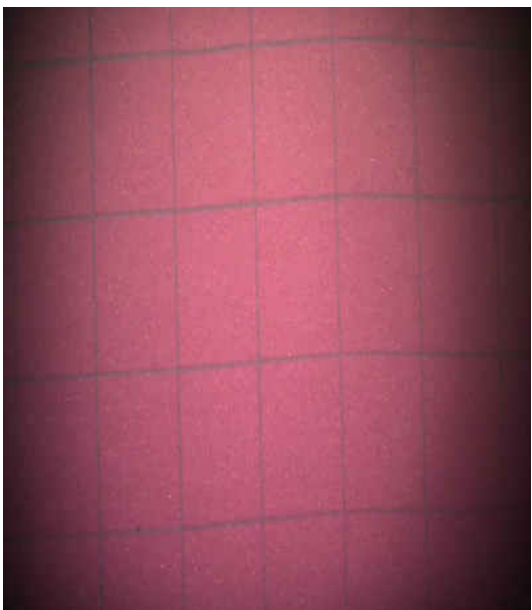
Figure 2.2.1-4 - Right Image
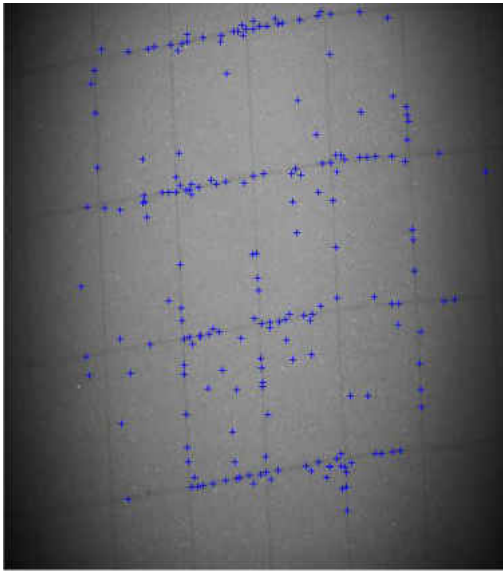
Figure 2.2.1-5 - SIFT features on Left Image
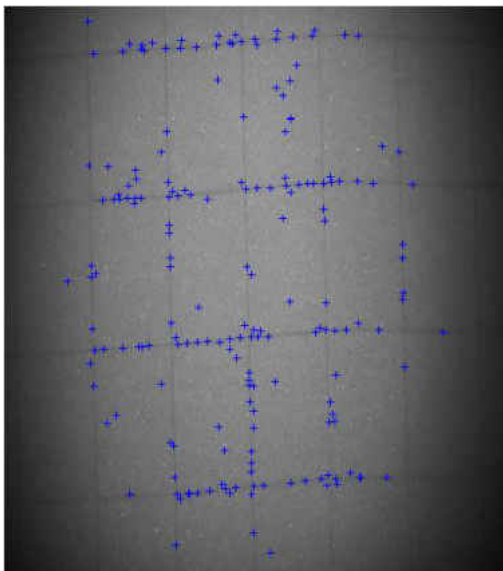
Figure 2.2.1-6 - SIFT features on Right Image

Step 3: Align the SIFT feature points in order to understand the correspondences

Step 4: Apply Ransac to Fit a 2D homography using the overestimated system on the

aligned SIFT feature points

Step 5: Compute the homography matrix for the left image with respect to its

right image using equation 2.1.

Step 6: Rectify/Warp the both the images and their weight mask by using the

homography found in step 5, as shown in Figure 2.2.1-7, 2.2.1-8, 2.2.1-9 and

2.2.1-10
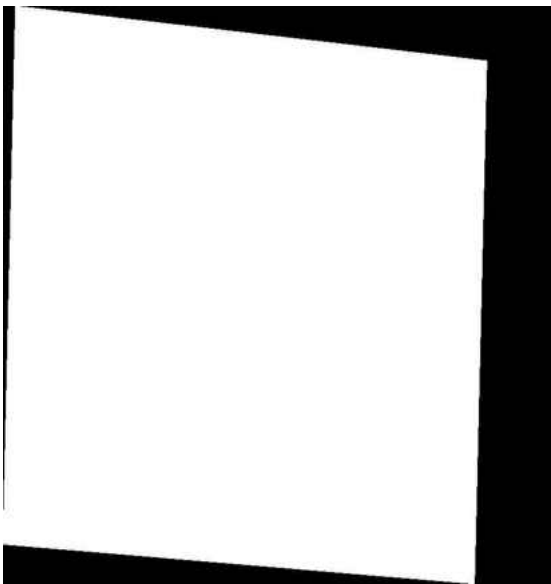
Figure 2.2.1-7 - Warped Mask: Left
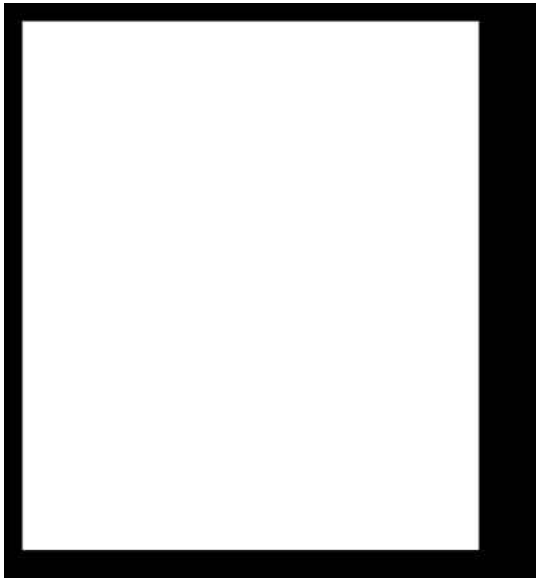
Figure 2.2.1-8 - Warped Mask: Right
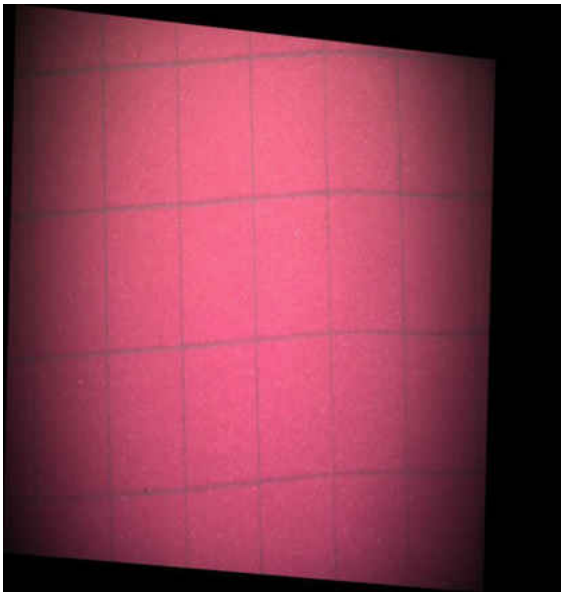
Figure 2.2.1-9 - Warped Left Image
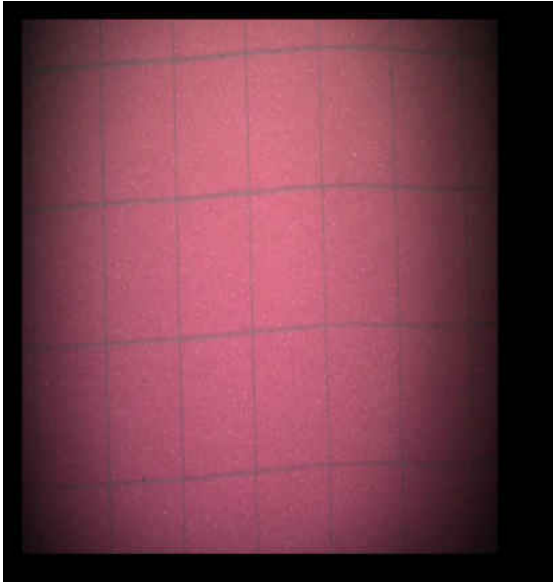
Figure 2.2.1-10 - Warped Right Image

Step 7: Overlay and crop these images to observe the warping correction in Figure 2.2.1-11
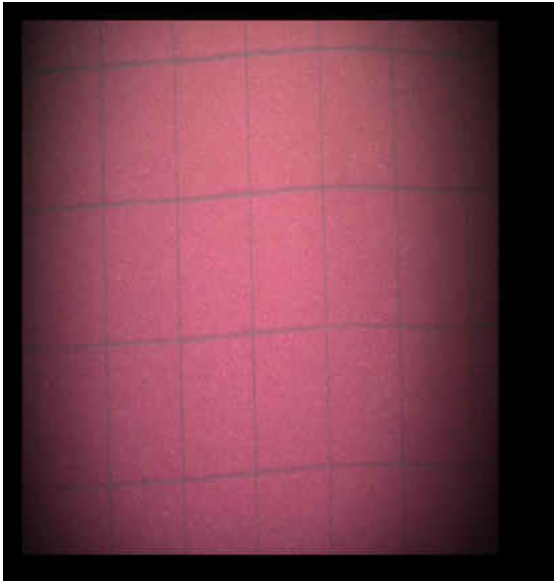
Figure 2.2.1-11 - Superimposed and cropped Images

demonstrating the correction

2.2.2 Background Subtraction and Modeling Block

Background subtraction and modeling is a method used for segmenting out and clean detection of the moving objects. The main purpose of this method is to get a clear extraction of the region of interest. If one has a model of how the background pixels behave, the "subtraction" process is very simple. But, in our problem due to high illumination of the region, the situation is much more challenging as the pixel values for the background keep varying just due to the minutest change in illumination.

Such a problem, can be best modeled by Adaptive Gaussian Mixture Model. Gaussian Mixture Model is a probabilistic model that assumes all the data points are generated from the mixture of finite number of Gaussian distributions with unknown parameters [82]. The principle behind the working of the GMM has already been covered in the introduction section. This technique does not have high memory requirement as the parameters are updated in each frame. The basic flow diagram for GMM is as shown in Figure 2.2.2-1, constituting of three parts: background initialization using the first 'n' frames as per our problem modeling, background modeling and updating with changes observed in new frames, and foreground pixel classification and extraction with the generation of a foreground mask [51,79].
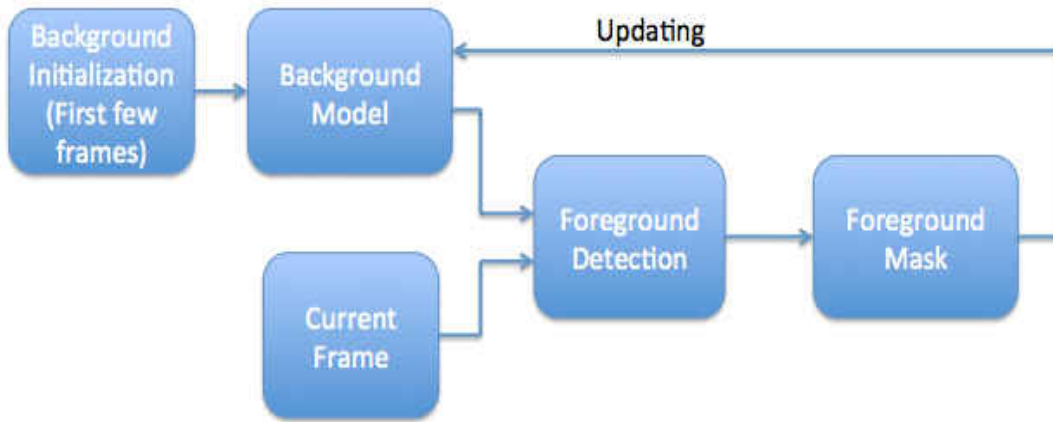
Figure 2.2.2-1 - GMM Block Diagram

After obtaining the output of a binary mask of the object as the output of GMM based foreground detection, in order to have a complete silhouette of the object we have explored some basic morphological operations like opening and closing with a defined kernel size found appropriate for the problem (narrowed down on a certain size purely based on experimentation). The input image and respective masks obtained by adaptive Gaussian mixture model and application of morphological operations are as shown in Figure 2.2.2-2 and 2.2.2-3 respectively.
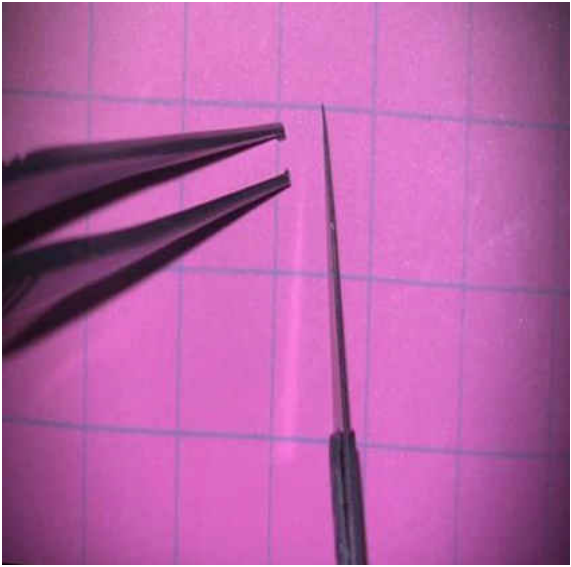
Figure 2.2.2-2 - Input Frame
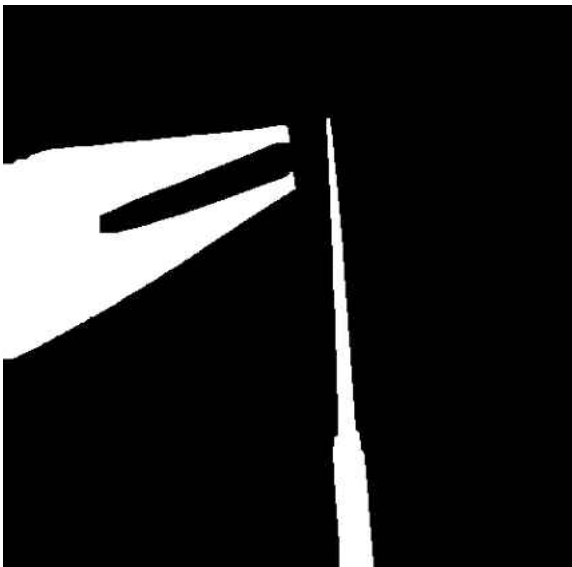
Figure 2.2.2-3 - Frame Mask

2.2.3 Multiple Object Tracking

This algorithm is based on the groundwork of clean object segmentations obtained in section 2.2.2, followed by application of blob analysis to segregate the different moving objects. This algorithm maintains individual tracks for each object. Once the tracks are initialized for a certain object, detections of the same object in the current frame are assigned to the respective tracks based solely on the probability associated with the motion of the object. This motion for each track is estimated using individual Kalman filter objects for each track. Algorithm deployed for multiple instrument tracking used in this project is detailed below [83].

Algorithm

Step 1: Read Video and for each frame do step 2 to step 22

Step 2: Read prevDetections from track's structure

Step 3: Read current frame from video object

Step 4: Look for the foreground pixels based on the background model if

frameNum>NumFramesUsedForBackgroundModeling

Step 5: Initialize or update the background model using this frame

Step 6: If any prevDetections not empty do steps 7 to 12

Step 7: Determine collision or overlap in objects if

numberOfCurrentDetections < prevDetections and unexpected increase in size

Step 8: If the condition in step 6 is true do steps 9 else jump to step 10

Step 9: Delete the currentDetections

Step 10: Determine if an objects is breaking into two parts

prevDetections < numberOfCurrentDetections and unexpected decrease in size

Step 11: If the condition in step 10 is true do steps 12 else jump to step 13

Step 12: Delete the respective track data from the structure as it was storing incorrect

tracks

Step 13: If tracks not empty predict new locations

Step 14: if tracks not empty assign current detections to respective tracks

Step 15: if any tracks were assigned in step 14 update them

Step 16: If certain track was not assigned, add predicted value to the detection

Step 17: if the last update to the track is a predicted location do steps 18 to 19

Step 18: if the last predicted location is not within the image or at an incorrect location do

step 19 else jump to step 20

Step 19: Remove the predicted location from the track

Step 20: Delete tracks whose invisibility/life ratio is > threshold

Step 21: If any detection is unassigned, create a new track

Step 22: Display the reliable tracks>minVisibility

Step 23: Stop

Throughout the running of the above algorithm, we have accumulated each and every object tracks of different objects in the video frames in a structure called AllTracks as shown in Figure 2.2.3-1.

Figure 2.2.3-1 - AllTracks structure

This structure is then refined to obtain a copy of only shortlisted tracks in EvalTracks, satisfying

the criteria of minimum life and visibility count as shown in Figure 2.2.3-2.

Figure 2.2.3-2 - EvalTracks structure

The assessment of Speed and Dexterity based on the basis of centroid locations is then conducted

on EvalTracks using the technique mentioned in section 2.2.4.

2.2.4   <u>Short Time Fourier Transform</u>

Short time fourier transform is a popular technique for frequency domain analysis in smaller time windows or local sections of signal as it varies over time [80].

We have deployed this technique to understand the different frequency components present in individually x and y component of the displacement of the object. The narrowing down of the time frame where higher frequency components are observed, helps this software in prompting the trainee where he went wrong, also this avoids over-averaging as opposed to traditional Fourier transform. The algorithmic steps for this technique are detailed below.

<u>**Algorithm**</u>

Step 1: For each track in EvalTracks shown in Figure 2.2.3-2 do step 2 to step 7

Step 2: Compute displacement signal using centroid locations for both x and y components individually

Step 3: Generate a Hamming Window of the defined window size

Step 4: For each temporal section in the displacement signal of x and y components do steps 5 to 6

Step 5: Perform Windowing using the equation 2.2

$$xW=x(1:wLen).*HammingWindow \tag{2.2}$$

Step 6: Compute Discrete Fourier Transform using equation 2.3

$$X(k) = \sum_{n=1}^{N} x(n) e^{\frac{2\pi(n-1)(k-1)}{N}} \tag{2.3}$$

Step 7: Using the X(k) obtained in step 6 for different temporal sections fill the bins of the STFT Matrix for the respective displacement signal, as shown in Figure 2.2.4-1

Step 8: Stop

Application of this algorithm to a given displacement signal yields STFT bin matrix as seen in Figure 2.2.4-1 and Figure 2.2.4-2. Figure 2.2.4-1 illustrates, the impact of the presence of tremor or high frequency component in the beginning of the signal, on the STFT matrix.
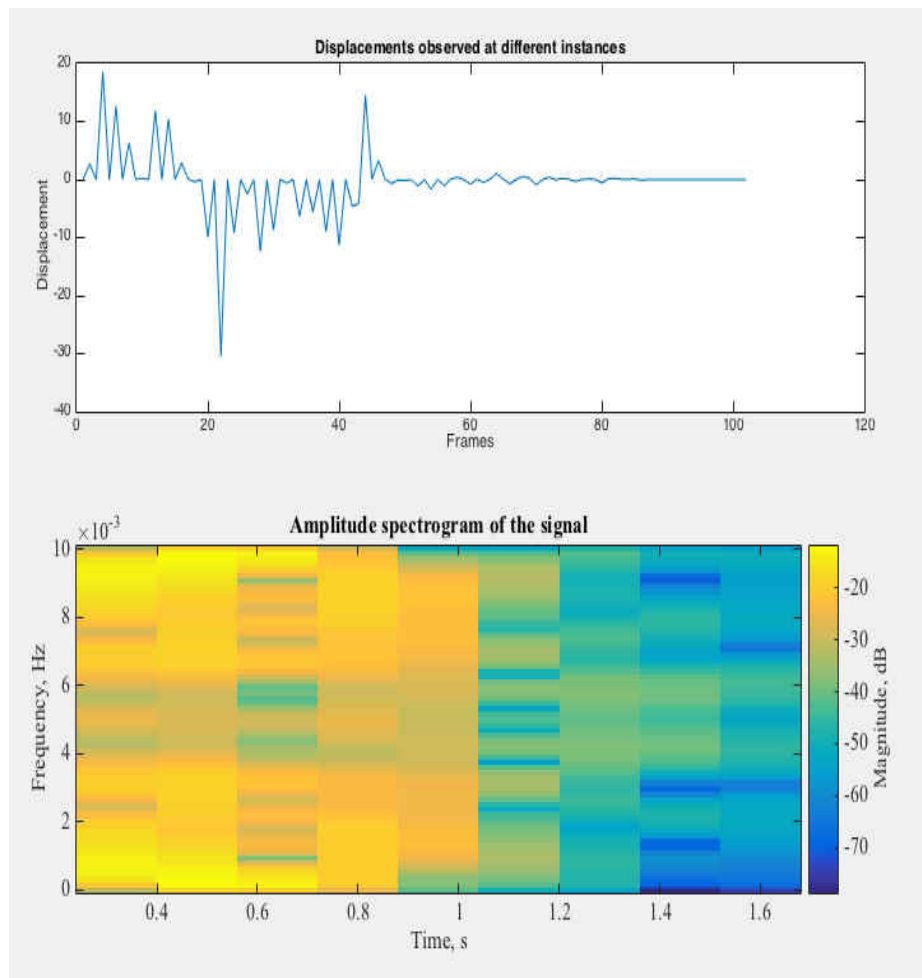
Figure 2.2.4-1 - Results STFT: (Top) Input Displacement with more tremor initially (Bottom) Corresponding STFT matrix plot frequency vs. time
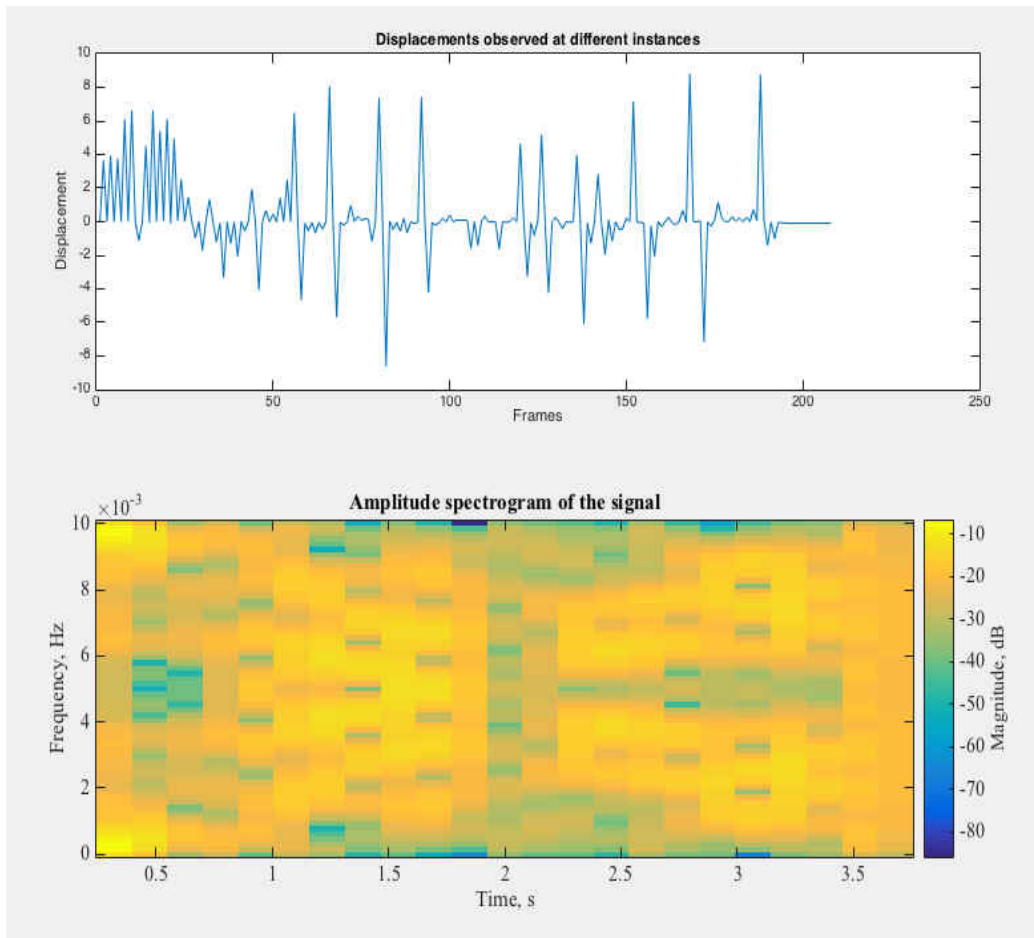
Figure 2.2.4-2 - Results STFT: (Top) Input Displacement with more tremor initially (Bottom) Corresponding STFT matrix plot frequency vs. time

2.2.5   <u>Assessment: Dexterity and Speed</u>

Once such STFT matrices are obtained for x and y components of each signal, we compute the presence of high frequency components above a certain magnitude threshold, to give a score for dexterity out of 4. The algorithm used to compute this score is as detailed below.

<u>Algorithm</u>

Step 1: For each track in EvalTracks observed in a TestVideo shown in Figure 2.2.3-2 do step 2

to step 5

Step 2: Compute STFT matrices for each displacement signal: x and y components

Individually, do steps 3

Step 3: Find the score each signal using its STFT matrix and equation 2.4

Score = Weighted sum of Frequency Components>ThresholdSpecified $\quad\quad$ (2.4)

Step 4: Find the average of the of the score obtained for x and y components

Step 5: Normalize this score using the total number of frames and duration of the respective track

And store them in DexterityScore(EvalTracks(n))

Step 6: Compute FinalDexterity Score using equation 2.5

FinalDexterityScores=$\frac{1}{N}\sum_{n=1}^{N}$ DexterityScore(EvalTracks(n)) $\quad\quad$ (2.5)

Step 7: Stop

2.2.6   Assessment: Speed

In parallel, compute the score for speed for each individual track in EvalTracks. This computation is guided by the algorithm given below.

Algorithm

Step 1: For each track in EvalTracks observed in a TestVideo shown in Figure 2.2.3-2 do step 2

to step 5

Step 2: Compute displacement signal: x and y components individually, do steps 3

Step 3: Find the individual speeds

Step 4: Find the average of the of the two speeds obtained for x and y components

Step 5: Normalize this speed using the total number of frames and duration of the respective

track and store them in Speed(EvalTracks(n))

Step 6: Compute FinalSpeed using Speed scores of each track in EvalTrack equation 2.6

$$FinalSpeed = \frac{1}{N}\sum_{n=1}^{N} Speed(EvalTracks(n)) \qquad (2.6)$$

Step 7: Assign FinalSpeedScore based on the difference between ideal speed and final speed of

observed for the TestVideo, guided by the equation 2.7

$$FianlSpeedScore = F(|IdealSpeed - FinalSpeed|) \qquad (2.7)$$

Both these scores are then converted into a grade using the Table 2.2.4-1.

Table 2.2.6-2 - Score to Grade Conversion for Dexterity and Speed

Score	Grade
4	A
3	B
2	C
1	D
0	E

2.2.7 **Frame Segregation**

Also all the frame numbers in EvalTracks mark the presence of object in track-able form, these frame numbers can then be segregated in a set named FrameNumsInEvalTracks. We can determine the set of frames with collisions or without any objects, in set named FrameNumsWithProbableColission by using the formula given in equation 2.8 below,

FrameNumsWithProbableColission = AllFrameNums – FrameNumsInEvalTracks   (2.8)

Once, these frames are determined we know that, these can used for further evaluation of characteristics like Eye-hand Co-ordination and Instrument-Tissue Manipulation.

This is a vital result as it enables us in not just extracting the frames where dexterity and speed

need to be evaluated, but it also helps in narrowing down to the frames where future algorithm on eye-hand tracking and instrument-tissue manipulation evaluations can be applied as shown in Figure 2.2.7-1.
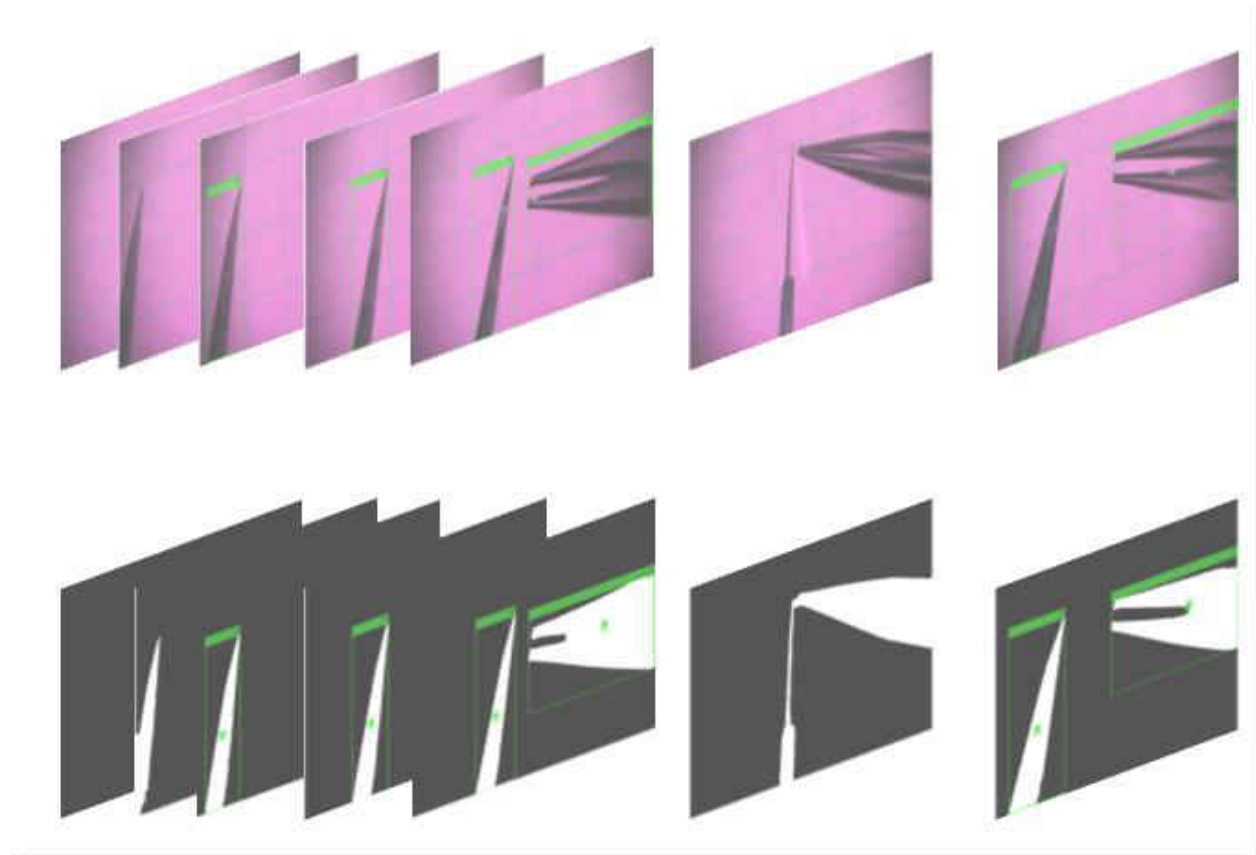
Figure 2.2.7-1 - Showing tracking of the object in different frames, except when the objects are too close: where other characteristics like eye-hand co-ordination and instrument-tissue manipulation need to evaluated

2.2.8   Object Fitting: Active Shape Model

Active Shape Models is a distinguished object fitting technique. It has applications in the area of non-rigid object segmentation, recognition and tracking. This technique has been popularly been used in face detection algorithms and biomedical imaging.

We have used this technique to locate and determine the object in the microsurgical activity video sequence. The algorithm deployed for the same is as detailed below.

Algorithm

Step 1: Annotated Training Data with major landmark points and interspersed with

additional points on each of the image of a particular class of objects is provided to the

algorithm, as shown in Figure 2.2.8-1.
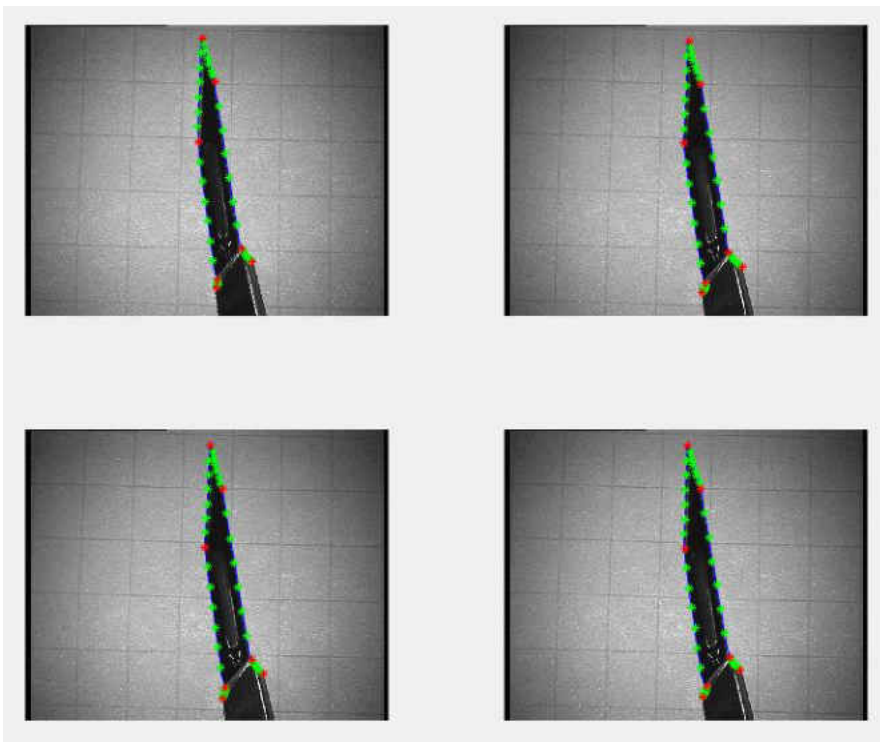
Figure 2.2.8-1 - Annotated training images: Knife

Step 2: The above training data is used to first align the different training shapes into a common

co-ordinate frame using a popular approach called Procruste's Analysis [75,77]

Step 3: This realigned data is then stored in a structure named TrainingData,

with following attributes,

> Vertices for Landmark and interspersed points,

> Lines describing the connectivity of these points

> and the Images that are now in common coordinate frame

Step 4: Empty elements from the structure TrainingData are removed

Step 5: Active Shape model is developed in this step by modeling the variations between

> TrainingData set contours, using PCA model

Step 6: The results of step 5 are stored in a structure named ShapeData,

> with following attributes, Eigen values, Eigen vectors, x-mean, x and lines.

Step 7: The variations in the shapes of the training data for the instrument can be as seen

> in Figure 2.2.8-2
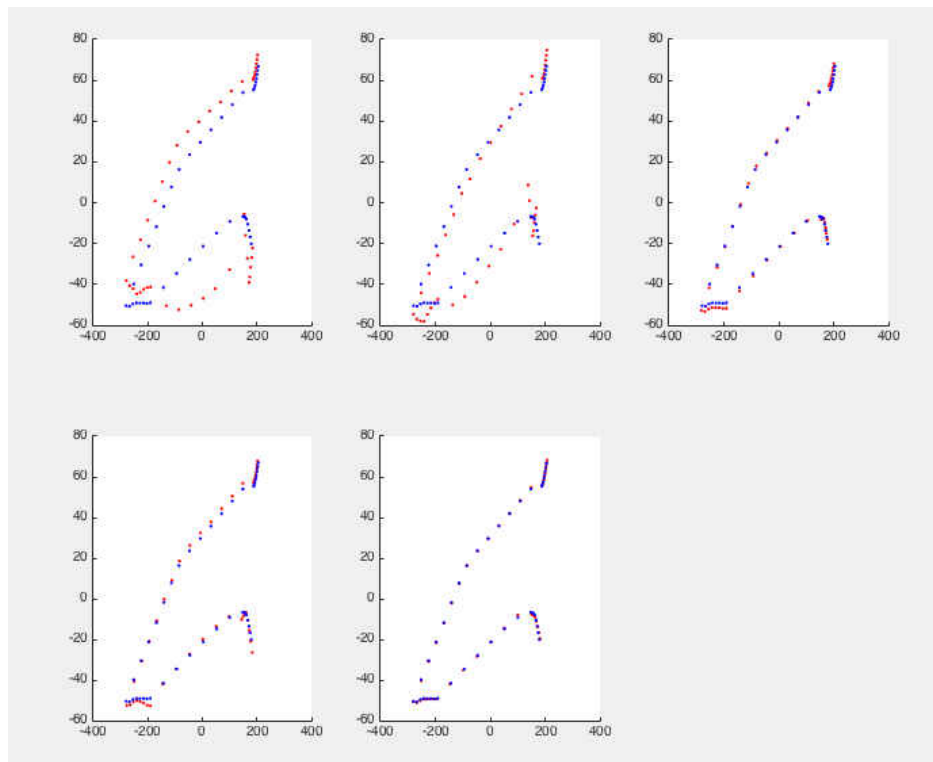
Figure 2.2.8-2 - Variations of different shape as compared to the mean shape

Step 8: In this step an Active appearance model from the given training images is built, which samples a intensity pixel profile/line perpendicular to each contour points. This is then used to build correlation matrices for each landmark (which is used later to converge to the object contour edge).

Step 9: Test image with either a manual initialized location or automatically initialized location are provided. The algorithm computes similar intensity profile lines along the contour points as done in step 8 for the training images and uses the Active appearance model and Active shape model to observe comparative variations to converge on the approximated contour points in the test frame.

This step converges by reducing the covariance offset, as the initialized contour points move towards the actual contour edge.

Step 10: Once the iterations converge, the two silhouettes are compared: Original Mask and Generated Mask. Values of Precision and Recall are obtained by equations 2.9 and 2.10.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive+False Positive}} \qquad (2.9)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive+False Negative}} \qquad (2.10)$$

Step 11: These values of Precision and Recall and then projected on PCA subspace for classification purposes

2.2.9 Evaluation of Effectualness

End effectualness of the micro-suturing activity performed by the trainee neurosurgeon, is one of the key assessment characteristic of his performance. This algorithm evaluates the end result of the micro-suturing activity, on the basis of the following components.

- Bite Thickness is given by A + B, as shown in Figure 2.2.9-2

  Thickness of Bites = A+B   (2.11)

- Approximations is given by C + D, as shown in Figure 2.2.9-2   (2.12)

- Angle of knot with respect to the cut as shown in Figure 2.2.9-1

- Inter-suture distance as shown in Figure 2.2.9-1
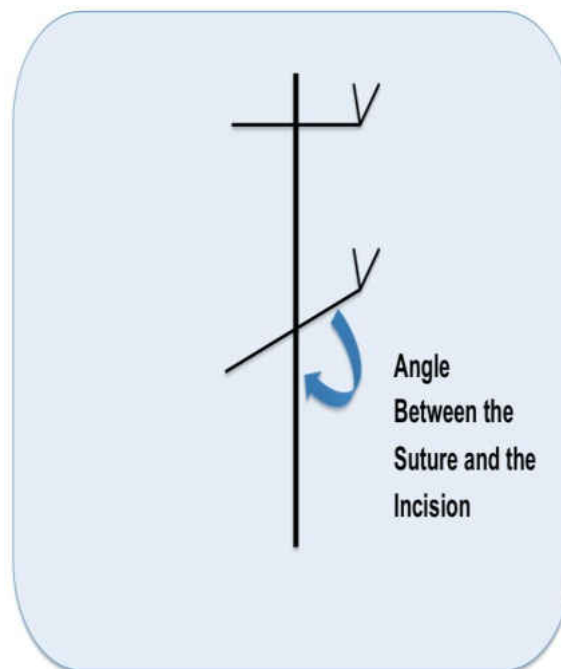
Figure 2.2.9-1 - Angle of the knot with that of the cut and inter-suture distance
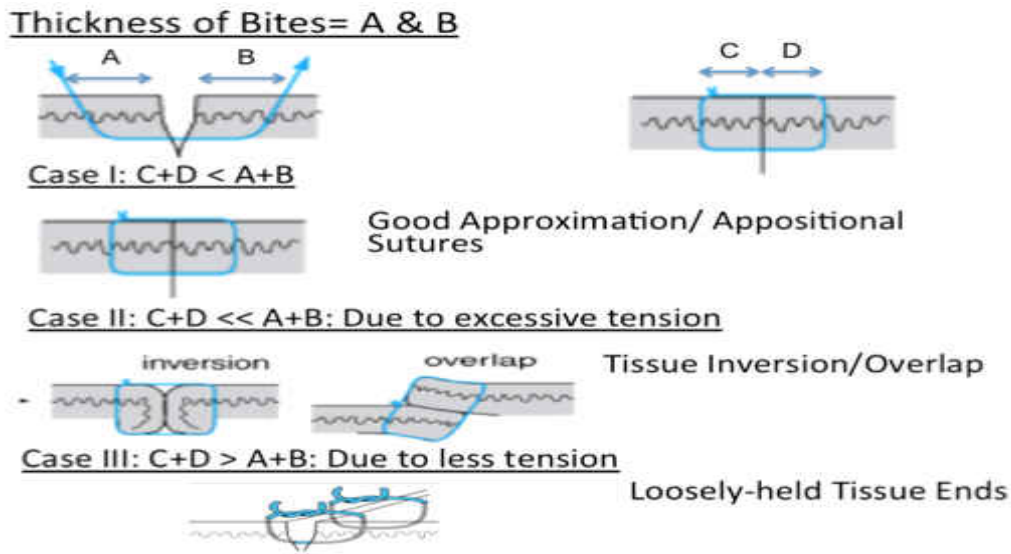
Figure 2.2.9-2 - Bite Thickness and Approximations

This algorithm is applied on top of the basic Gaussian mixture model and foreground or region of interest extraction procedures discussed in section 2.2.2. First, the acquisition of the frames with the cut, bites and final knots is done. These frames are then sent to the Background subtraction model (section 2.2.2) for the extraction of region of interest. Once, this is achieved the respective frames are analyzed to measure different characteristics as per the steps detailed below.

<u>Algorithm</u>

Step 1: The cut made, is analyzed using hough transform based line and curve fitting

Techniques as shown in Figure 2.2.9-3. The result of the same is as observed by the

annotated cut mark in Figure 2.2.9-4.
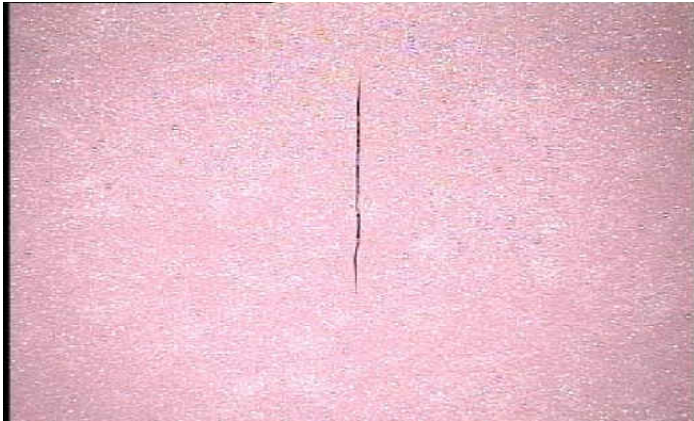
Figure 2.2.9-3 - Input cut Image
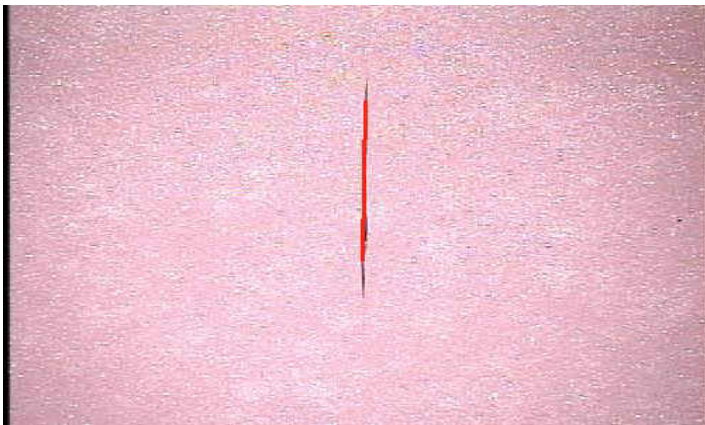
Figure 2.2.9-4 - Annotated Cut Mark: with an observed angle of 89.13 degree

Step 2: Thickness of the bites taken on either sides of the cut are analyzed in the input frame, Figure 2.2.9-5. The identification of the each bite points is done by the deployment of corner detecting algorithm on top of multi-resolution approach. Once these bite points are obtained their distance from cut measured and summed to get the Thickness of bites, as given by equation 2.11.
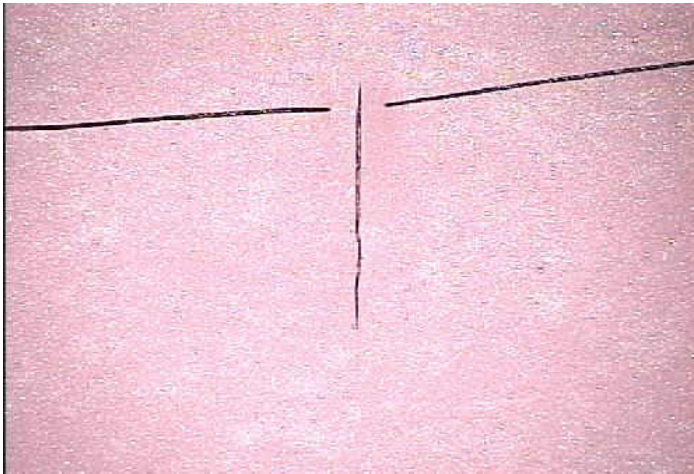
Figure 2.2.9-5 - Frame input to measure the bites
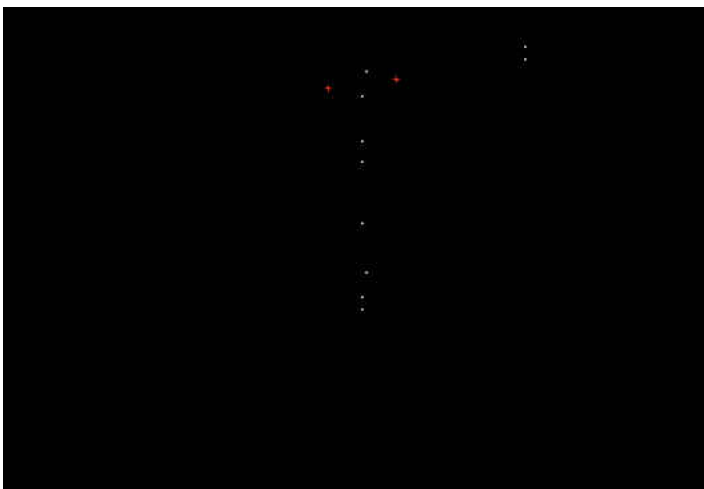
Figure 2.2.9-6 - Bites as analyzed by multi-resolution approach

The Bite Thickness observed in this particular sample were,

A=24, B=39

Thickness of Bites = A+B = 63

Step 3: Knot Extraction and analysis is done by the deployment of Hough transform based line fitting techniques, on Figure 2.2.9-7 to record the orientation of the knot as well as the length as shown in Figure 2.2.9-8 and Figure 2.2.9-9.
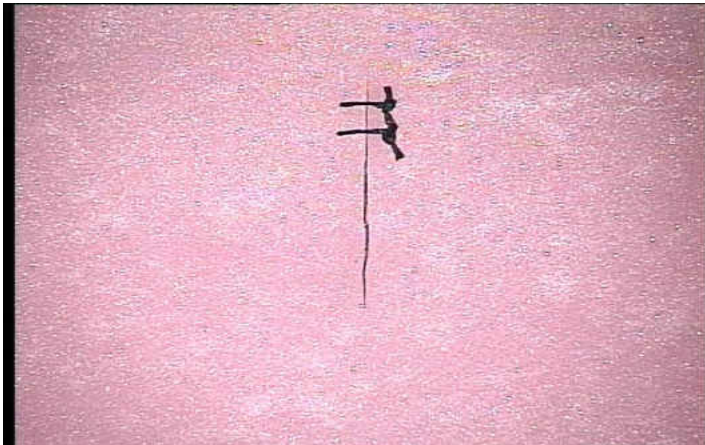
Figure 2.2.9-7 - Input Image with knots
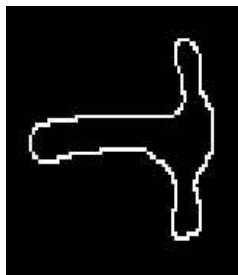
Figure 2.2.9-8 - Extracted knot contour

Figure 2.2.9-9 - Knot Silhouette with annotated line signifying the orientation of the knot

Measured parameters in the above example are,

Orientation of the knot= 6 degrees

Length of the knot = C+D= 59

The knot length and knot orientation recorded in step 3 are the compared with the orientation of the cut and bite thickness recorded in steps 1 and 2 respectively.

In this particular example, the Total Bite Thickness > Length of the Knot. Thus, categorizing this sample as good approximation.

Angle the angle of the knot with respect to that of the cut for this example was measured to be 96.8 degrees.

3. RESULTS

To understand the efficacy of our algorithms, this chapter is broadly classified in the following three sections to individually understand the different components like Tracking, Dexterity and Speed, Object Fitting and Effectualness.

3.1   Tracking Algorithm

The multiple object-tracking algorithm deployed to track and localize the instruments is built on basic Kalman Filter approach, followed by maintaining individual tracks for each of the objects observed.

Tracking algorithm was run on video samples with single and multiple instruments involving different interaction between the foreground objects as well as interaction between the foreground and background objects.

As the results of the localization of the instruments are stored in the form of list of centroids observed in each frame for different instrument tracks, to understand the efficacy of the tracking algorithm, it was studied by recording the observations on the WEIZMANN Dataset). This helped is in understanding the accuracy of the localization of objects by this algorithm. (http://www.wisdom.weizmann.ac.il/~vision/SpaceTimeActions.html)

The results generated on these can be studied in the following steps.

Step 1: Comparison of the centroids generated by the algorithm to the ground truth. The results on one of the samples in the Dataset can be seen in Figure 3.1-1. Figure 3.1-2 (a), (b) and (c) illustrate with plots he difference in the location of the centroids as compared to ground truth for the sample shown in Figure 3.1-1, in direction of motion, in direction

perpendicular to the direction of motion and overall difference in position of the centroid

as compared to the Ground Truth respectively.

Figure 2.2.9-1 - Tracks: Centroid locations measured by the Algorithm are shown in blue and ground truth shown in red
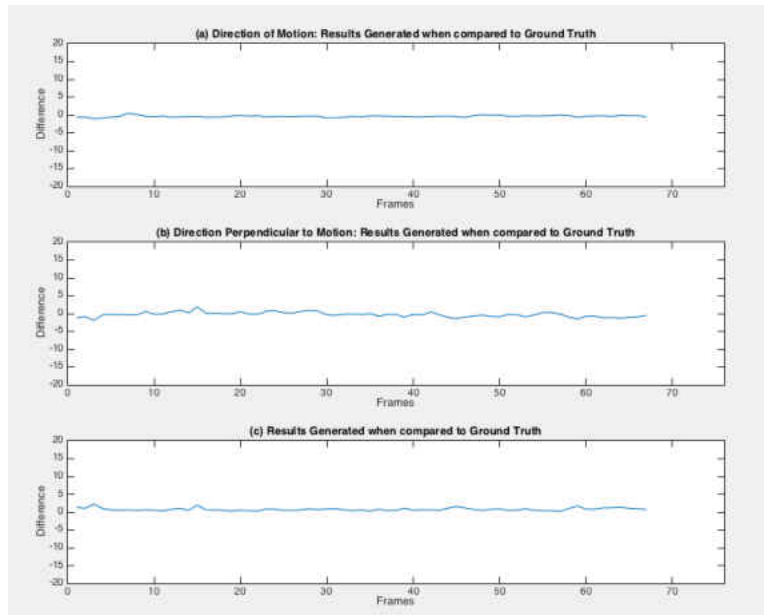
Figure 2.2.9-2 - The difference in the location of the Centroids as compared to ground truth for the sample shown above is (a) direction of motion (b) direction perpendicular to the direction of motion (c) Overall difference in position of the centroid as compared to the Ground Truth

This difference in the position of the centroids was tabulated and the difference in the values was analyzed by the application of threshold T (=1 pixel).

The accuracy percentage was then studied by the following equation 3.1,

$$\text{Accuracy Percentage} = 1 - \frac{\text{Number frames with Difference in Position>T}}{\text{Total Number of Frames}} \text{ x } 100 \quad (3.1)$$

Accuracy percentage for different samples in the WEIZMANN data set, were then recorded as shown in Table 3.1.

Table 3.1-1 - Accuracy Percentage for WEIZMANN Dataset

Sample Name	Accuracy Percentage
Eli Walk	100
Daria Walk	96.15
Lena Walk	98.68
Lyova Walk	95.74
Denis Walk	78.69

Average accuracy was then computed to be 93.85 %.

3.2 Dexterity and Speed

In order to understand the effectualness of the dexterity and speed evaluation done by system developed as compared to that done by senior neurosurgeons. Therefore, the data of scores was obtained by a blinded evaluation done by the senior surgeon and the score provided by our algorithm i.e. objective and subjective analysis (as shown in Table 3.2-1 and Figure 3.2-1).

Table 3.2-2 - Percentage Match for Dexterity and Speed

Factors	Percentage Match
Dexterity	45.00%
Speed	60.00%
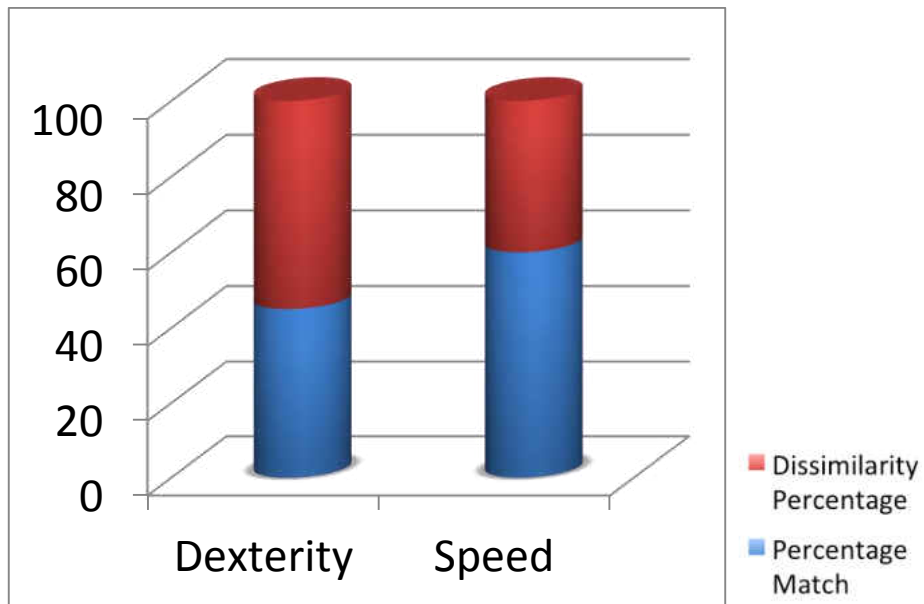
Figure 2.2.9-1 - Similarity vs. dissimilarity percentages of the results of Dexterity and Speed

It was observed that even though computerized analysis cannot completely surpass the training under apprenticeship model, it is required to remove the ambiguity of biasedness of scoring as well as to intimate trainees about the specific target areas (that need improvisation).

3.3   Object Fitting

As discussed in section 2.2.8, active shape model based technique is deployed for object fitting and recognition in this system. This particular technique is used in our system for classification purpose of the instruments. Thus, we have studied the performance of this algorithm in order to maximize the difference between the different classes of objects, by varying the following parameters.

Case I: Contour Points between the given Landmark Points

The contour points and landmark points are annotated by green stars and red stars as shown in Figure 2.2.8-1.

The results were studied by using a sample set of 10 images each of the positive and negative sample. But for the convenience in understanding, they have been demonstrated for one positive and negative example by varying the number of contour points between landmarks. The Figure 3.3-1 illustrates the difference in measurement for precision and recall between the positive and negative samples at different set number of contour points interpolated between the landmark points.
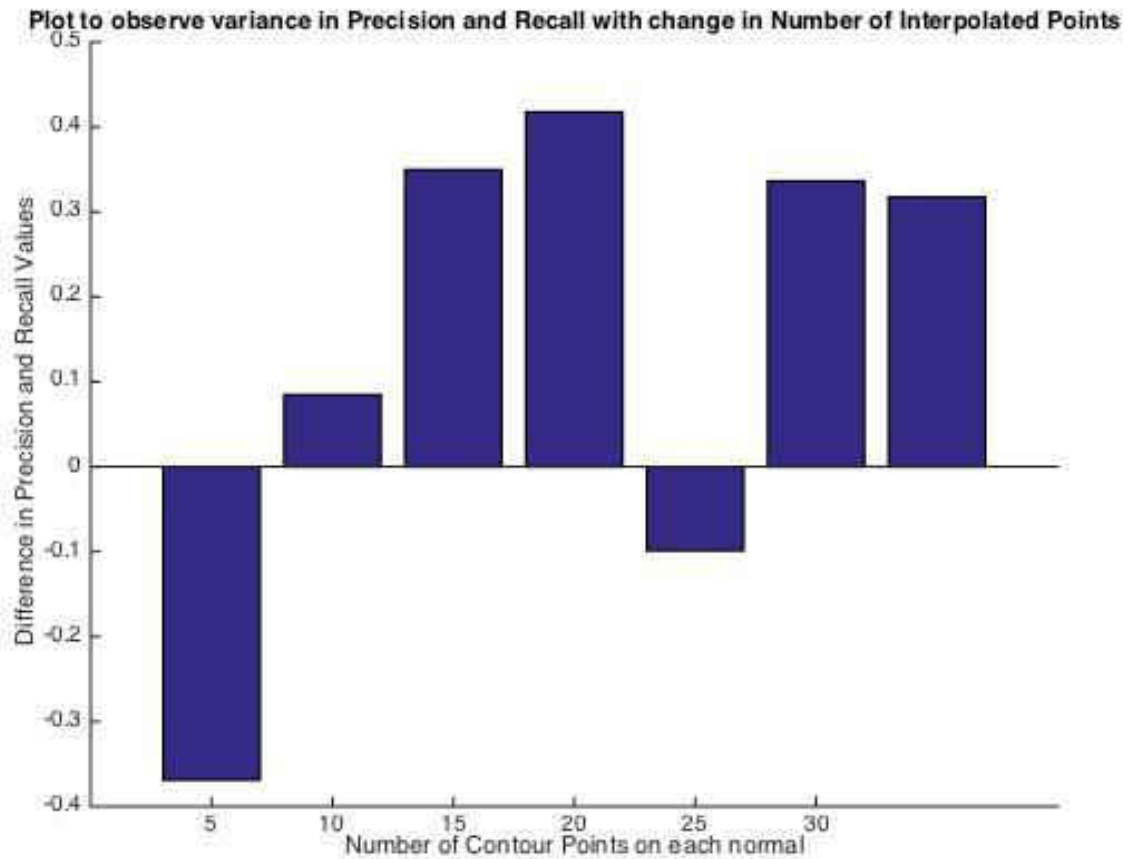
Figure 2.2.9-1 - Difference in measurement for precision and recall between the positive and negative samples at different set number of contour points interpolated between the landmark points

As observed in Figure 3.3-1 the difference in the two sets of examples i.e. positive and negative was maximized when the number of contour points between the landmark points was set to 20.

Case II: Length of Landmark Intensity Profile

The landmark intensity profile length is the length in direction of the normal at each contour point (annotated by green stars and red stars as shown in Figure 2.2.8-1).

This length guides the modeling of active appearance sub-unit discussed in section 2.2.8. The

impact of varying length of landmark profiles on the classification problem was done using a sample set of 10 images each of the positive and negative sample. But for the convenience in understanding, they have been demonstrated for one positive and negative example by varying the number of contour points between landmarks. The Figure 3.3-2 illustrates the difference in measurement for precision and recall between the positive and negative samples at different set landmark profile lengths.

Figure 2.2.9-2 - Difference in measurement for precision and recall between the positive and negative samples at different set landmark profile lengths

As observed in Figure 3.3-2 the difference in the two sets of examples i.e. positive and negative

was maximized when the landmark intensity profile length was set to 6 pixels.

Case III: Search Length for Contour Points

The search length for contour points guides the search model while converging or fitting the contour on test image (as discussed in section 2.2.8). This is the search length in the direction of the normal on each side of the contour. The impact of varying search length of contour points on the classification problem, was studied using a sample set of 10 images each of the positive and negative sample. But for the convenience in understanding, they have been demonstrated for one positive and negative example by varying the number of contour points between landmarks. The Figure 3.3-3 illustrates the difference in measurement for precision and recall between the positive and negative samples at different search lengths of contour points.

Figure 2.2.9-3 - Difference in measurement for precision and recall between the positive and negative samples at different search lengths of contour points
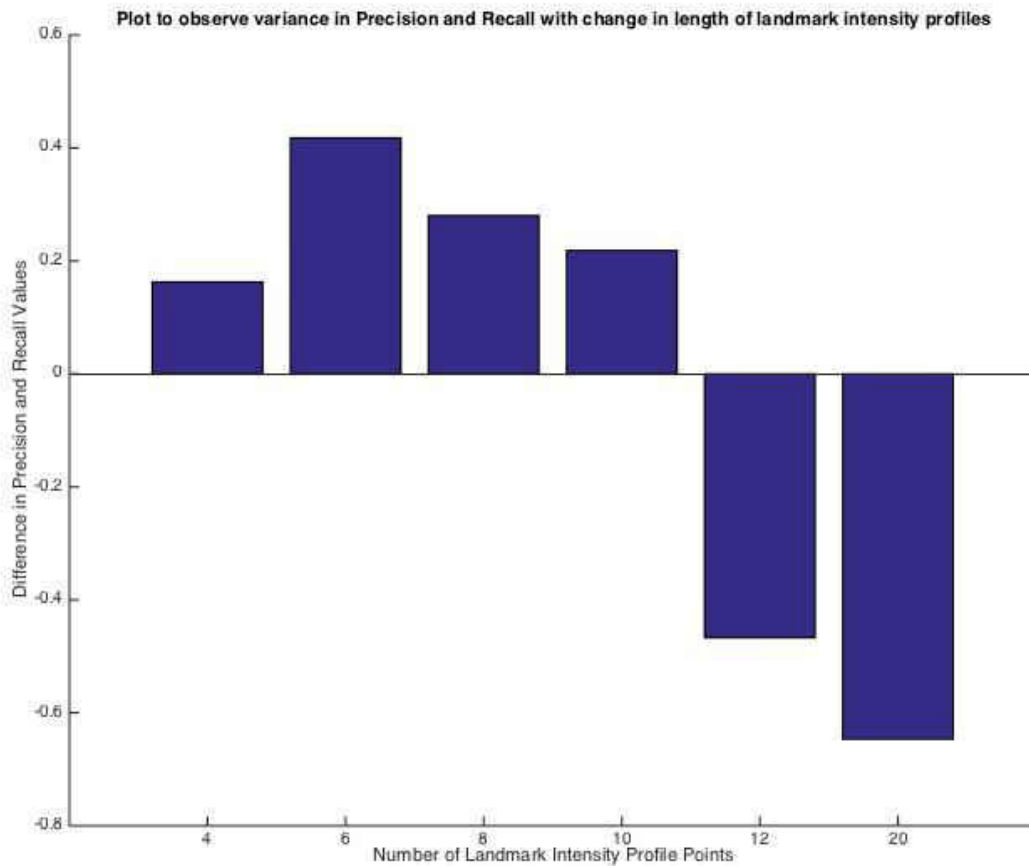
As observed in Figure 3.3-3 the difference in the two sets of examples i.e. positive and negative was maximum when the search lengths at contour points was set to 8 pixels.

Case IV: Scale

The search model in algorithm discussed in section 2.2.8 is modeled for a multi-resolution approach based object fitting technique. The parameter scale aids in defining the number of the scales the search model of the algorithm should work on while converging the object contour points on the given image. The impact of varying scale parameter on the classification problem,

was studied using a sample set of 10 images each of the positive and negative sample. But for the convenience in understanding, they have been demonstrated for one positive and negative example by varying the number of contour points between landmarks. The Figure 3.3-4 illustrates the difference in measurement for precision and recall between the positive and negative samples at different values of scale parameter.
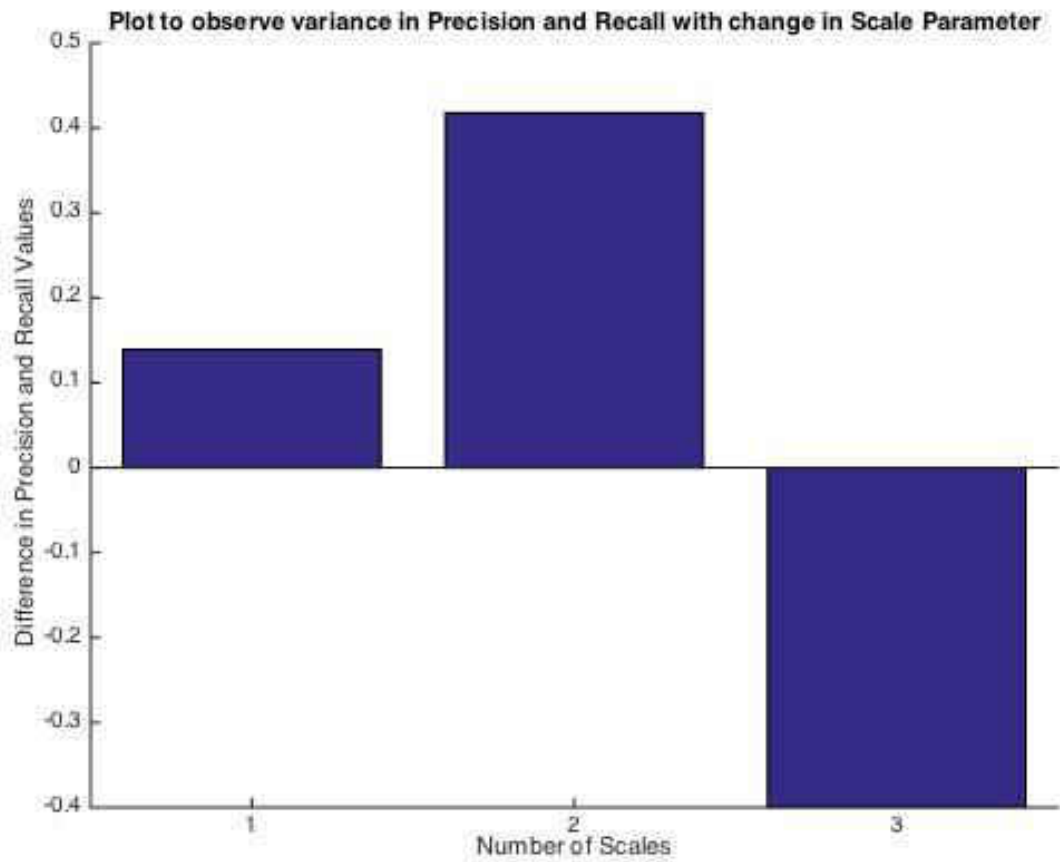
Figure 2.2.9-4 - Difference in measurement for precision and recall between the positive and negative samples at different values of scale parameter

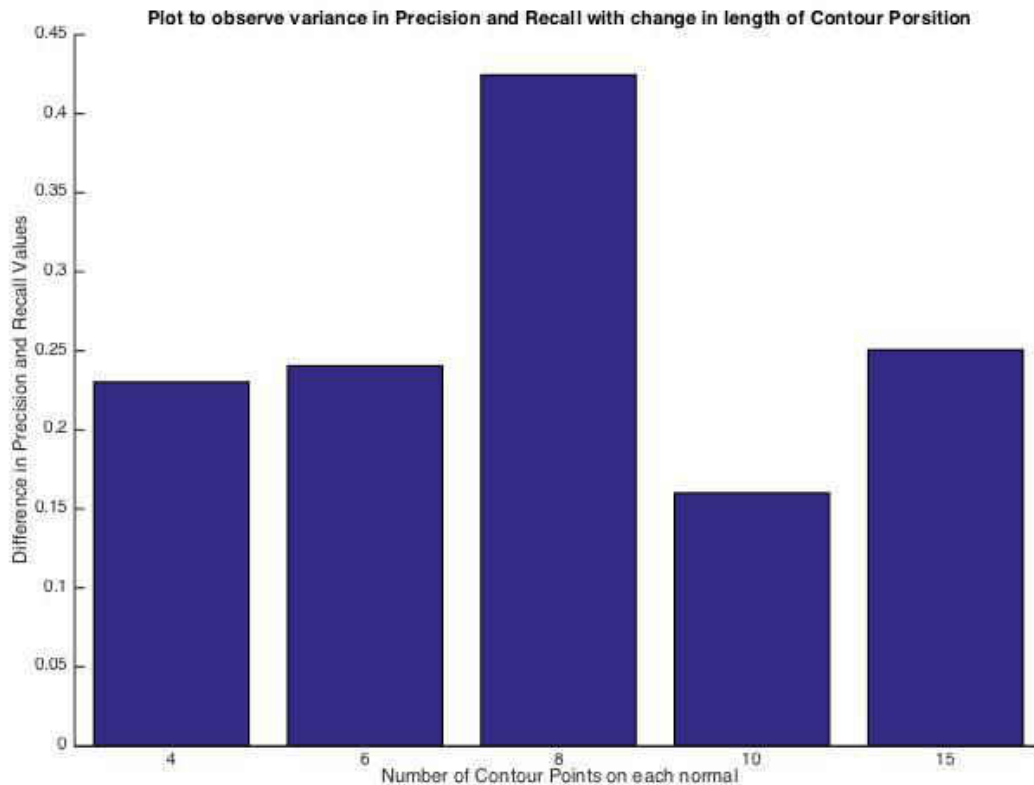As observed in Figure 3.3-4 the difference in the two sets of examples i.e. positive and negative

was maximum when the parameter scale was set to 2.

Case V: Iterations

The search model in algorithm discussed in section 2.2.8 converges for the purpose of object fitting problem on the given test image in defined number of iterations. The impact of varying iterations parameter on the classification problem, was studied using a sample set of 10 images each of the positive and negative sample. But for the convenience in understanding, they have been demonstrated for one positive and negative example by varying the number of permitted iterations. It was observed that once all the other parameters are appropriately set there is almost no variation observed in the difference in the two sets of examples (i.e. positive and negative).

Once all the parameter were set in accordance to maximize the difference between two classes of object / instruments, the system was run for 63 samples each of positive and negative samples to demonstrate the segregation of data on Precision vs. Recall plot as shown in Figure 3.3-5.
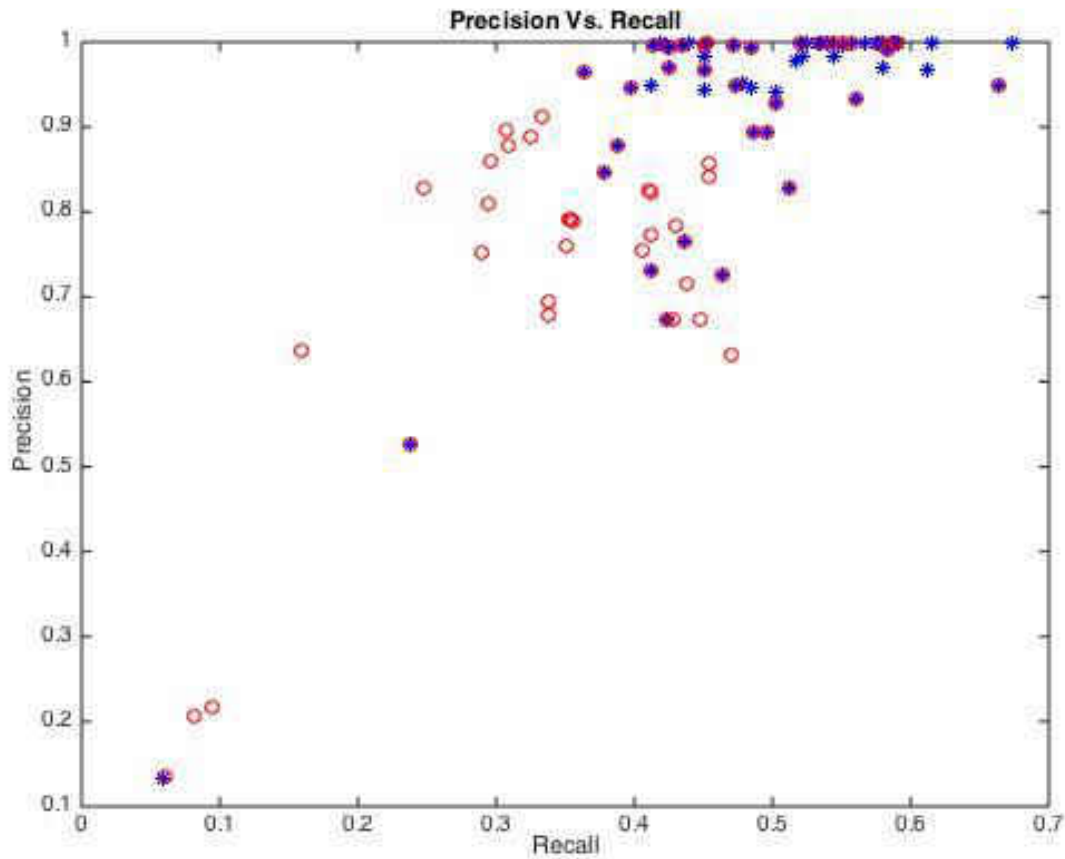
Figure 2.2.9-5 - Precision Vs. Recall for Positive and Negative Samples

3.4   Object Classification

Further, for classification purposes PCA is deployed on the precision and recall values of two classes observed in Figure 3.3-5 and the resultant classifier is as shown in Figure 3.4-1.

In order to estimate the potential of the algorithm in object classification, precision and recall values were computed using the TP, TN, FP and FN values obtained in Figure 3.4-1.
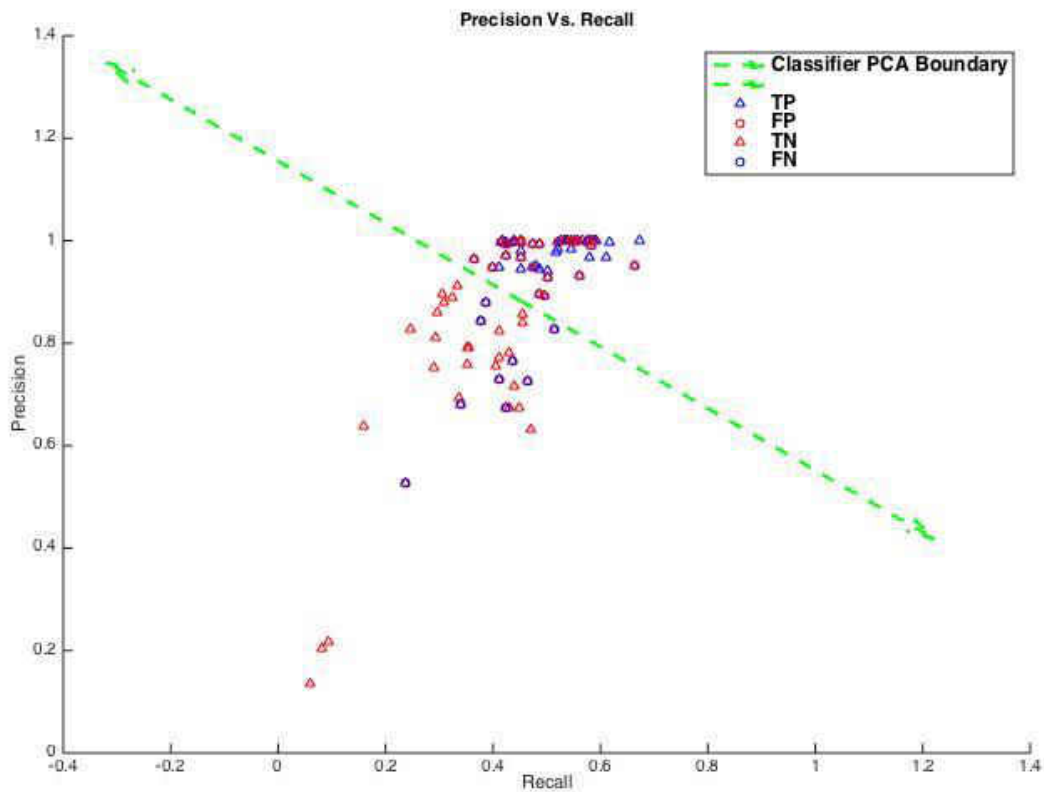
Figure 2.2.9-1 - Precision Vs. Recall Value based Object Classifier

The values of precision and recall for object classification using the current algorithm were observed as:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} = 0.675$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} = 0.857$$

3.5   Effectualness

Data was obtained after comparing evaluation results obtained from the software and by the examiner i.e. objective and subjective analysis (as shown in Table 3.5-13, Figure 3.5-1 and 3.5-

2).

Figure 3.5-1 details the similarity vs. dissimilarity percentages of the results of effectualness of 4-0 and 5-0 sutures, whereas Figure 3.5-2 demonstrates the similarity and dissimilarity percentage of 9-0 and 10-0 sutures. It was observed that even though computerized analysis cannot completely surpass the training under apprenticeship model, it is required to remove the ambiguity of biasedness as well as to intimate trainees about the specific target areas (that need improvisation).

Table 3.5-3 - Percentage Match between Objective and Subjective Assessment: Effectualness

Factors	Percentage Match Thick Sutures (4-0, 5-0)	Percentage Match Fine Sutures (7-0, 9-0, 10-0)
Bites	38.46%	62.50%
Approximations	46.15%	50.00%
Angle of knot w.r.t. cut	38.46%	37.50%
Inter-Suture Distance	38.46%	62.50%
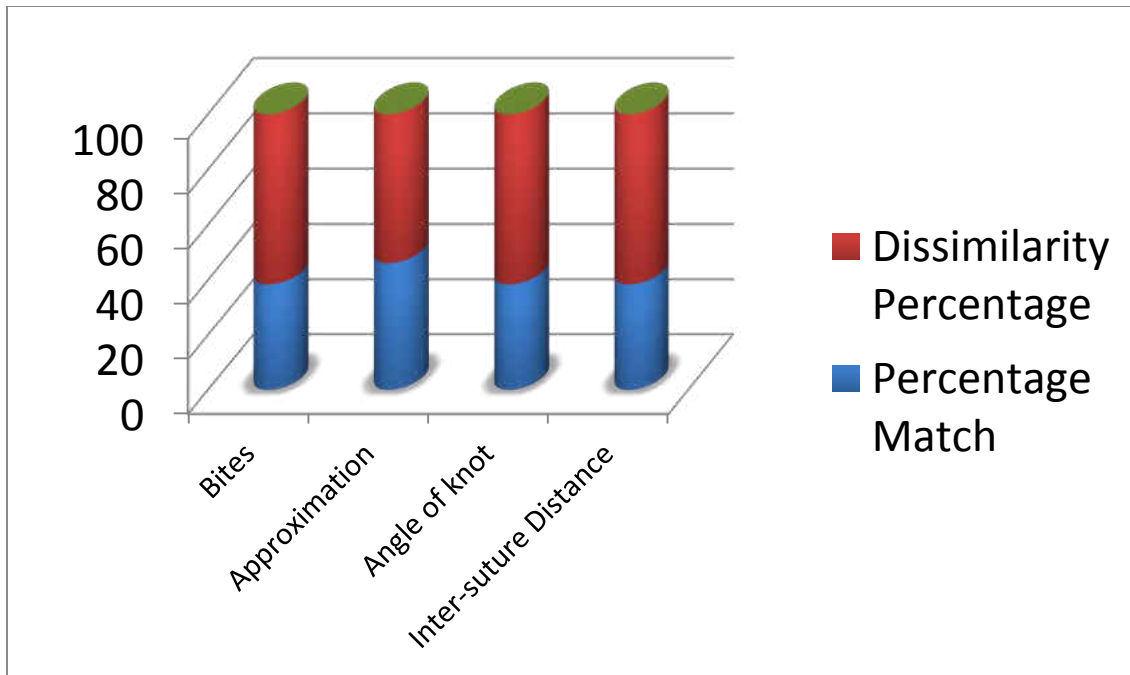
Figure 2.2.9-1 - Similarity vs. dissimilarity percentages of the results of effectualness of 4-0 and 5-0 sutures
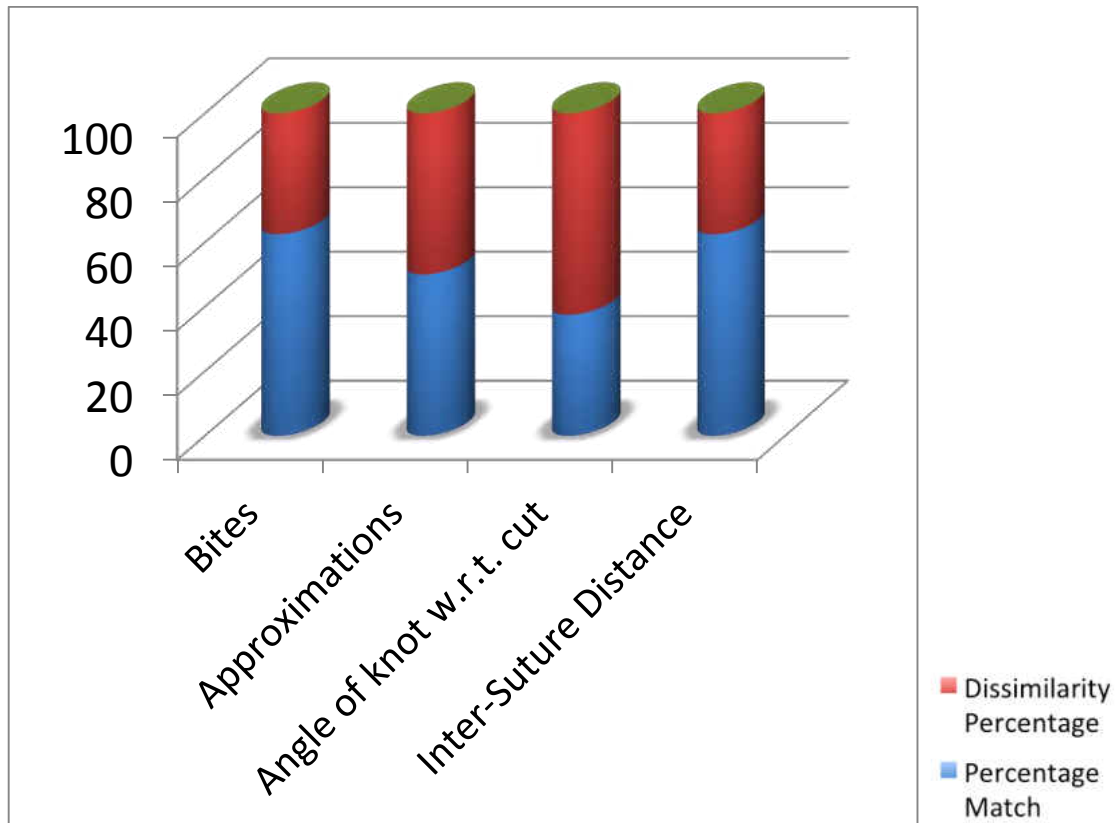
Figure 2.2.9-2 - Similarity vs. dissimilarity percentages of the results of effectualness of 9-0 and 10-0 sutures

It was realized that such a computerized evaluation and assessment system can provide an unbiased, component-wise, individual and collective assessment of micro-suturing techniques under supervised yet unsupervised environment. Thus, this system can be used to validate as well as supplement the evaluation done by an examiner, while helping the trainees understand their flaws and improving their micro-suturing skills in a more focused manner.

4. CONCLUSION

4.1   Conclusion and Final Thoughts

In this thesis we have explored different methods to segment, detect and track the location of the instruments and analyze their movements as well as end results of microsurgical activity. The key contribution of this work is that different microsurgical instruments can be identified and tracked consistently for the frames where dexterity and speed need to be evaluated. Also a byproduct of this research was the identification of set of frames or section of the video clip where other characteristics i.e. techniques like eye-hand co-ordination and instrument-tissue manipulation need to be evaluated.

A detailed presentation and analysis of the problem being targeted and various developments in this field were explored for their efficacy in providing surgical skills training. It was argued that even though techniques like active shape models seem promising in the area of non-rigid object localization and recognition, they take a couple of seconds to execute for each frame. More techniques for on the go, tracking learning and detection might be suitable for a real-time system. It was realized that such a computerized evaluation and assessment system can provide an unbiased, component-wise, individual and collective assessment of micro-suturing techniques under supervised yet unsupervised environment. Thus, this system can be used to validate as well as supplement the evaluation done by an examiner, while helping the trainees understand their flaws and improving their micro-suturing skills in a more focused manner.

4.2   Future Work

Current state-of-the-art approaches to skills training in neurosurgery are being made, which highlight involvement of graphics and haptic devices, but they tend to create an environment far from the one in the OR in terms of unnaturalness in response of the visual feedback as well as with the dimensions/feel of the haptic devices. Our current technique definitely lacks closeness to the pathological and anatomical semblance to the live surgery on human brain and spinal cord. But, it gives the trainees an opportunity to work with actual instruments under the operating microscope, and in turn help them hone their surgical skills in similar environment.

Work in the field of activity recognition can be explored, for different steps in the micro-suturing procedure, to aid in the identification of the intent. This problem will fall under the category of complex activity recognition guided by features or cues like instruments present, presence of suture material, movements of the instrument, interaction with the tissue, etc. This can help the algorithm have better judgments of the surgical technique, and in turn produce more reliable evaluations.

Once the development of the computer vision-based surgical evaluation system is achieved with a good level of efficiency, the area of Mixed-reality can be explored to give a closer semblance to surgical scenario and activity area. This can also be the first step towards the development of Robotic Micro-neurosurgery.

4.3   Conference Presentations

- Payal Jotwani, Ashish Suri and Hassan Foroosh et al. "Paradigm Shift From Apprenticeship Training Model To Computerized Evaluation: Effectualness of Micro-Suturing In

Neurosurgery Skills Training." Proceedings of 30th Annual Meeting AANS (American Association of Neurological Surgeons) / CNS Section on Disorders of the Spine and Peripheral Nerves, 2014.

- Payal Jotwani, Ashish Suri and Hassan Foroosh et al. "Stereoscopy-based Computerized Evaluation of Microneurosurgery Skills." Proceedings of 2014 CNS Annual Meet (Congress of Neurological Surgeons Society)

LIST OF REFERENCES

1. Suri, A. "Neurosurgery Education and Training School Website: Cadaver Anatomy Tutorial". http://www.aiimsnets.org/cadaver.asp

2. Suri, A. et al. "Neurosurgery Skills Training - Past, Present and Future: Hands-On Skills Training Modules to Virtual Reality Simulation". Journal of Neurological Surgery Part B: Skull Base, Issue S 01, 2014.

3. Phillipe, A. et al. "Current Method of Training in Microsurgery. Telemicrosurgery: Robot Assisted Surgery" Section 7.2, pp-60.

4. Jotwani, P. et al. "Free-access open-source e-learning in comprehensive neurosurgery skills training".

5. Cameron, J. et al. "Williams Stewart Halsted: Our Surgical Heritage". Annals of Surgery, Vol. 225 (5): pp 445–458, 1997.

6. Kavic, M. S. "Competency and the six core competencies". [editorial] JSLS, Vol. 6: pp 95- 97, 2002.

7. Sherman, I. et al. "Personal recollections of Walter E. Dandy and his Brain Team". Journal of Neurosurgery, Vol. 105: pp 487, 2006.

8. Sachdeva AK. "Acquisition and maintenance of surgical competence". Semin Vasc Surg., Vol. 15: pp 182-190, 2002.

9. Barnes, R. W., Lang, N. P., Whiteside, M. F. "Halstedian technique revisited. Innovations in teaching surgical skills", Ann Surg. Vol. 210(1): pp118-21, 1989.

10. Rutkow, I. M. William Stewart Halsted and the Germanic influence on education and training programs in surgery". Surg. Gynecol. Obstet., Vol. 147: pp 602–606, 1978.

11. Balasundaram, I., Aggarwal, R., Darzi, L. A. "Development of a training curriculum for microsurgery". Br J Oral MaxillofacSurg, Vol. 48: pp 598–606, 2010.

12. Jotwani, P. et al. "Paradigm Shift from apprenticeship training model to computerized valuation: Effectualness of micro-suturing in neurosurgery skills training". Proceedings of 30th Annual Meeting of the AANS/CNS Section on Disorders of the Spine and Peripheral Nerves. Abstarct No. 403.

13. Atkins, J. L., Kalu, P. U., Lannon, D. A. et al "Training in microsurgical skills: does course- based learning deliver"? Microsurgery, Vol. 25: pp 481–485, 2005.

14. Chan, W. Y., Matteucci, P., Southern, S. J., "Validation of microsurgical models in microsur- gery training and competence: A review". Microsurgery Vol. 27: pp 494–499, 2007.

15. Nielsen, P. E., Foglia, L. M., Mandel, L. S. et al "Objective structured assessment of technical skills for episiotomy repair", Am. J. Obstet. Gynecol. Vol. 189: pp1257–1260, 2003.

16. Siddiqui, N. Y., Stepp, K. J., Lasch, S. J. et al "Objective structured assessment of technical skills for repair of fourth-degree perineal lacerations". Am. J. Obstet. Gynecol., Vol. 199: pp 671–676, 2008.

17. VanBlaricom, A. L., Goff, B. A., Chinn, M. et al "A new curriculum for hysteroscopy training as demonstrated by an objective structured assessment of technical skills (OSATS)". Am. J. Obstet. Gynecol., Vol. 193: pp 1856–1865, 2005.

18. Swift, S. E., Carter, J. F. "Institution and validation of an observed structured assessment of technical skills (OSATS) for obstetrics and gynecology residents and faculty". Am. J. Obstet. Gynecol., Vol. 195: pp 617–621, 621–613, 2006.

19. Martin, J. A., Regehr, G., Reznick, R. et al "Objective structured assessment of technical skill (OSATS) for surgical residents". Br. J. Surg. Vol. 84: pp 273–278, 1997.

20. Acland, R. D. "Microvascular anastomosis: a device for holding stay sutures and a new vascular clamp". Surgery, Vol. 75: pp 185–187, 1974.

21. Giselle, C. et al. "The role of simulation in neurosurgery". Childs Nerv. Syst. Vol. 30: pp 1997–2000 DOI 10.1007/s00381-014-2548-7, 2014.

22. Banks, J., Carson, J., Nelson, B., and Nicol, D. "Discrete-Event System Simulation". Prentice Hall. pp 3. ISBN 0-13-088702-1, 2001.

23. McLachlan, J. C., Bligh, J., Bradley, P. and Searle, J. "Teaching anatomy without cadavers". Med. Educ. Vol. 38: pp 418–424, 2004.

24. Resch, K. D. M. and Perneczky, A. "Use of plastined crania in neuroendoscopy". J Int. Soc. Plastination, Vol. 6: pp 15–16, 1992.

25. Zymberg, S. T., Vaz-GuimaraesFilho, F. and Lyra, M. "Neuroendoscopic training: presentation of a new real simulator". Minim. Invasive Neurosurg. Vol. 53: pp 44–46, 2010.

26. Nicholas Gélinas-Phaneuf et al. "Assessing performance in brain tumor resection using a novel virtual reality simulator". Int. J. CARS DOI 10.1007/s11548-013-0905-8.

27. Rosseau, G. et al. "The development of a virtual simulator for training neurosurgeons to perform and perfect endoscopic endonasaltranssphenoidal surgery." Neurosurgery. Vol. 73(1): pp 85-93. doi: 10.1227/NEU.0000000000000112, 2013.

28. GmaanAlZhrani et al. "Proficiency Performance Benchmarks for Removal of Simulated Brain Tumors Using a Virtual Reality Simulator NeuroTouch". Journal of Surgical Education Vol. 16: pp iii48, 2014.

29. De Paolis, L. T., De Mauro, A., Raczkowsky, J. and Aloisio, G. "Virtual model of the human brain for neurosurgical simulation". Stud. Health Technol. Inform. Vol. 150: pp 811–815, 2009.

30. Gallagher, A. G., Cates, C. U. "Virtual reality training for the operating room and cardiac catheterisation laboratory". Lancet Vol. 364: pp 1538–1540, 2004.

31. Malone, H. R., et al. "Simulation in neurosurgery: a review of computer-based simulation environments and their surgical applications". Neurosurgery Vol. 67(4): pp 1105-16. doi: 10.1227/NEU.0b013e3181ee46d0, 2010.

32. Immersive Touch Website: http://www.immersivetouch.com/index.html

33. Botden, S. M., Buzink, S. N., Schijven, M. P., Jakimowicz, J. J. "Augmented versus virtual reality laparoscopic simulation: what is the difference? A comparison of the ProMIS augmented reality laparoscopic simulator versus LapSim virtual reality laparoscopic simulator". World J. Surg. Vo. 31: pp 764–772, 2007.

34. Condino, S., Carbone, M., Ferrari, V., Faggioni, L., Peri, A., Ferrari, M. et al "How to build patient-specific synthetic abdominal anatomies. An innovative approach from physical toward hybrid surgical simu- lators". Int. J. Med. Robot Vol. 7: pp 202–213, 2011.

35. Debes, A. J., Aggarwal, R., Balasundaram, I., Jacobsen, M. B. "Construction of an evidence-based, graduated training curriculum for D-box, a webcam-based laparoscopic basic skills trainer box". Am. J. Surg. Vol. 203: pp 768–775, 2012.

36. Halic, T., Kockara, S., Bayrak, C. and Rowe, R. "Mixed reality simulation of rasping procedure in artificial cervical disc replacement (ACDR) surgery". BMC Bioinformatics Vol. 11(6): S11, 2010.

37. Jain, M., Tantia, O., Khanna, S., Sen, B. and Sasmal, P. K. "Hernia endotrainer: results of training on self-designed hernia trainer box". J. Laparoendosc. Adv. Surg. Tech. A Vol. 19: pp 535–540, 2009.

38. Khine, M., Leung, E., Morran, C. and Muthukumarasamy, G. "Homemade laparoscopic simulators for surgical trainees". Clin. Teach. Vol. 8: pp118–121, 2011.

39. Coelho, G., Kondageski, C., Vaz-GuimaraesFilho, F., Ramina, R., Hunhevicz, S. C., Daga, F., Lyra, M. R., Cavalheiro, S. and Zymberg, S. T. "Frameless image-guided neuroendoscopy training in real simulators". Minim. Invasive Neurosurg. Vol 54: pp 115–118, 2011.

40. Zymberg, S. T., Vaz-GuimaraesFilho, F. and Lyra, M., "Neuroendoscopic training: presentation of a new real simulator". Minim. Invasive Neurosurg. Vol. 53: pp 44–46, 2010.

41. Seymour, N. E., Gallagher, A. G., Roman, S., O'Brien, M. K., Bansal, V. K., Andersen, D. K. and Satava, R. M. "Virtual reality training improves operating room performance: results of a randomized, double-blinded study". Ann. Surg. Vol. 236(4): pp 458–63, discussion 463–4, 2002.

42. Gerben, E. Breimer. "Design and evaluation of a new synthetic brain simulator for endoscopic third ventriculostomy". Journal of Neurosurgery: Pediatrics. Vol. 15(1): pp 82-88, 2015.

43. Matsumura N., et al. "A newly designed training tool for microvascular anastomosis techniques: Microvascular Practice Card. Surgical Neurology Vol. 71: pp 616–620, 2009.

44. Krishnan K. G., et al. "Simple and viable in vitro perfusion model for training microvascular anastomoses". Microsurgery Vol. 24(4): pp 335-8, 2004.

45. Mantovani, G., Fukushima, W. Y., Baik, Cho A., Aita, M. A. and Mazzetti, M. V. "Use of earthworms for microsurgery training". J. Reconstr. Microsurg. Vol. 25: pp 275–278, 2009.

46. Piccardi, M., "Background subtraction techniques: a review". IEEE International Conference on Systems, Man and Cybernetics 2004

47. Wren, C., Azarhayejani, A., Darrell, T. and Pentland, A.P. "Pfinder: real-time tracking of the human body," IEEE Trans. on Patfern Anal. and Machine Infell., Vol. 19(7): pp 780-785, 1997.

48. Koller, D., Weber, J., Huang, T., Ma, J. 14 Ogasawara, G., Rao, B., and Russell, S., "Towards Robust Automatic Traffic Scene Analysis in Real-time", Proc. ICPR Vol. 94: pp 126-131, 1994.

49. Lo, B.P.L. and Velastin, S.A. "Automatic congestion detection system for underground platforms," Proc. ISIMP Vol. 2001: pp 158-161, 2001.

50. Cucchiara, R., Grana, C., Piccardi, M., and Prati, A. "Detecting moving objects, ghosts, and shadows in video streams," ZEEE Tram on Paftem Anal. and Machine Infell, Vol. 25(10): pp 1337-1442, 2003.

51. Stauffer, C. and Grimson, W.E.L., "Adaptive background mixture models for real-time tracking," Proc. IEEE CVPR Vol. 1999: pp 24 & 252, 1999.

52. Wayne Power, P. and Schoonees, J. A., "Understanding background mixture models for foreground segmentation", Proc. of IVCNZ, pp 267-271, Nov. 2002.

53. AlperYilmaz et al. "Object Tracking: A Survey." ACM Computing Surveys, Vol. 38(4), Article 13, Publication date: December 2006.

54. Veenman, C., Reinders, M., and Backer, E.. Resolving motion correspondence for densely moving points. IEEE Trans. Patt. Analy. Mach. Intell. 23(1): pp 54–72, 2001.

55. Serby, D., Koller-Meier, S., and Gool, L. V. "Probabilistic object tracking using multiple features. In IEEE International Conference of Pattern Recognition (ICPR)". pp 184–187, 2004.

56. Comaniciu, D., Ramesh, V., and Meer, P. Kernel-based object tracking. IEEE Trans. Patt. Analy. Mach. Intell. Vol. 25: pp 564–575, 2003.

57. Yilmaz, A., Li, X., and Shah, M. "Contour based object tracking with occlusion handling in video acquired using mobile cameras". IEEE Trans. Patt. Analy. Mach. Intell. Vol. 26(11). pp 1531–1536, 2004.

58. Sethi, I. and Jain, R. "Finding trajectories of feature points in a monocular image sequence". IEEE Trans. Patt. Analy. Mach. Intell. Vol. 9(1). pp 56–73, 1987.

59. Salari, V. and Sethi, I. K. "Feature point correspondence in the presence of occlusion". IEEE Trans. Patt. Analy. Mach. Intell. Vol. 12(1). pp 87–91, 1990.

60. Matthies, L., Szeliski, R., and Kanade, T. "Kalman filter-based algorithms for estimating depth from image sequences". Int. J. Comput. Vision Vol. 3(3): pp 209–238,1989.

61. Vaswani, N., Roychowdhury, A., and Chellappa, R. "Activity recognition using the dynamics of the configuration ofinteracting objects". In IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp 633–640, 2003.

62. Zhou, S., Chellapa, R., and Moghadam, B. "Adaptive visual tracking and recognition using particle filters". In Proceedings IEEE International Conference on Multimedia and Expo (ICME). pp 349–352, 2003.

63. Isard, M. and Blake, A. Condensation - conditional density propagation for visual tracking. Int. J. Comput. Vision Vol. 29(1). pp 5–28, 1998.

64. Bouaynaya, N., Qu, W., and Schonfeld, D., "An online motion-based particle filter for head tracking applications," Proceedings of the IEEE International Conference on

Acoustics, Speech, and Signal Processing, vol. 2, Philadelphia, PA, USA, pp 225–228, March 2005.

65. Bouaynaya, N. and Schonfeld, D., "On the optimality of motion-based particle filtering," IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 7, pp. 1068–1072, July 2009.

66. Cox, I. J. "A review of statistical data association techniques for motion correspondence". Int. J. Com- put. Vision 10(1): pp 53–66, 1993.

67. Reid, D. B. "An algorithm for tracking multiple targets". IEEE Trans. Autom. Control Vol. 24(6): pp 843–854, 1979.

68. Singh, A., Goldgof, D., Terzopoulos, D., "Deformable Models in Medical Image Analysis". IEEE Computer Society, ISBN: 0818685212.

69. Kass, M., Witkin, A., Terzopoulos, D., "Snakes: Active Contour Models". International Journal of Computer Vision, Vol. 1(4): pp 321-331, 1988.

70. Lobregt, S., Viergever, M., "A discrete dynamic contour model". IEEE Transactions on Medical Imaging, 14(1): pp 12 -24, 1995.

71. McInerney, T., Terzopoulos, D., "Topologically adaptable snakes". Proceedings of the Fifth International Conference on Computer Vision, pp.840 -845, 1995

72. Cohen, L., "On active contour models and balloons". CVGIP: Image understanding, Vol. 53(2): pp 211-218, 1991.

73. Grzeszczuk, R., Levin, D., "Brownian strings": Segmenting images withstochastically deformable contours. IEEE PAMI, Vol. 19(10): pp 1100 -1114, 1997.

74. Herlin, I., Nguyen, C., Graffigne, C., "A deformable region model using stochastic proces ses applied to echocardiographic images". Proceedings of Computer Visionand   Pattern Recognition, pp 534 -539, 1992.

75. Cootes, T., Taylor, C., Cooper, D., Graham, J., "Active Shape Models - Their trainingand application". Computer Vision and Image Understanding, Vol. 61(1): pp 38-59, 1995.

76. GhassanHamarneh et al."Deformable Spatio-Temporal ShapeModels: Extending ASM to 2D+Time". British Machine Vision Conference, BMVC 2001, Vol. 1, pp. 13-22, Manchester, UK. pp 10-13, September 2001.

77. Cootes, T.F., Hill, A., Haslam, J., Taylor, C.J. "The use of active shape models for locating structure in medical images". Image and Vision Computing Vol. 12(6): pp 355–366, 1994.

78. Hamarneh, G., Abu-Gharbieg, R., Gustavsson, T.: Review Active Shape Models - Part I: Modeling Shape and Gray Level Variations. In: SSAB, 1998.

79. Kaewtrakulpong,P. and Bowden, R. "An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection", In Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01, Video Based Surveillance Systems: Computer Vision and Distributed Processing, September 2001.

80. Broida, T. And Chellappa, R. "Estimation of object motion parameters from noisy images". IEEE Trans. Patt. Analy. Mach. Intell. Vol. 8(1): pp 90–99, 1986.

81. Beymer, D. and Konolige, K. "Real-time tracking of multiple people using continuous detection". In IEEE International Conference on Computer Vision (ICCV) Frame-Rate Workshop, 1999.

82. Wang, R. et al. "Trends and Applications in Knowledge Discovery and Data Mining." PAKDD 2013 International Workshops: DMApps, DANTH, QIMIE, BDM, CDA, CloudSD, Gold Coast, QLD, Australia. pp 601-605, April 14-17, 2013.

83. Mathworks. "Motion-based multiple Object Tracking." http://www.mathworks.com/help/vision/examples/motion-based-multiple-object-tracking.html