SPIM

Thèse de Doctorat

école doctorale **sciences pour l'ingénieur et microtechniques**

UFC UNIVERSITÉ DE FRANCHE-COMTÉ

# Vision and Visual Servoing for Nanomanipulation and Nanocharacterization using Scanning Electron Microscope

Naresh **MARTURI**

SPIM

Thèse de Doctorat

école doctorale **sciences pour l'ingénieur et microtechniques**

UFC UNIVERSITÉ DE FRANCHE-COMTÉ

N° X X X

THÈSE présentée par

**Naresh MARTURI**

pour obtenir le

Grade de Docteur de
l'Université de Franche-Comté

Spécialité : **Automatique**

# Vision and Visual Servoing for Nanomanipulation and Nanocharacterization using Scanning Electron Microscope

Soutenue le Date-19 November 2013 devant le Jury :

| | | |
|---|---|---|
| **Guillaume MOREL** | Rapporteur | Professeur à l'Université Pierre Marie Curie, Paris |
| **Alexandre KRUPA** | Rapporteur | Chargé de recherche, HDR, IRISA / INRIA, Rennes |
| **Michel PAINDAVOINE** | Examinateur | Professeur à l'Université de Bourgogne |
| **Olivier HAEBERLE** | Examinateur | Professeur à l'Université de Haute-Alsace |
| **Michèle ROMBAUT** | Examinateur | Professeur à l'Université Joseph Fourier, Grenoble |
| **Brahim TAMADAZTE** | Examinateur | Chargé de recherche, FEMTO-ST |
| **Nadine LE-FORT PIAT** | Directeur de thèse | Professeur à l'ENSMM |
| **Sounkalo DEMBÉLÉ** | Directeur de thèse | Maître de conférences, HDR, Université de Franche-Comté |

*I cannot teach anybody anything, I can only make them think.*
-Socrates

This work is *dedicated to* my parents, sister, grand mother and my wife.

# Acknowledgements

It would not have been possible to successfully finish this challenging journey of Ph.D. without the support and the help of many people. Foremost, it is a great privilege for me to work in the department of Automatic Control and Micro-Mechatronic systems (AS2M) at FEMTO-ST institute. It provided me with a very good and a friendly work environment, working in where I feel like my three years passed like three days.

My first and the unending gratitude goes to my thesis advisors Prof. Nadine LE-FORT PIAT, ENSMM and Assoc. Prof. Sounkalo DEMBÉLÉ, HDR, Université de Franche-Comté for trusting and offering me this position and also for their patience, motivation and valuable suggestions. Their guidance and the necessary research freedom they have provided helped me a lot in finishing my experiments with the scanning electron microscope and in writing my thesis dissertation. I would like to specially thank Nadine, Emmanuel PIAT and their family for helping me out during the beginning days of my stay in France. I would also like to extend my special thanks to Soun for his continuous support and patience in setting up the experimental system where, the hardware gave us many troubles. Apart from them, I sincerely thank Prof. Ivan KALAYKOV, School of Science and Technology, Örebro, Sweden for encouraging me towards the research in vision-based robotic manipulation.

Next, I would like to sincerely thank my jury members: Prof. Guillaume MOREL, UPMC Paris, Dr. Alexandre KRUPA, CR1 CNRS, HDR, IRISA/INRIA Rennes, Prof. Michèle ROMBAUT, UJF, Prof. Michel PAINDAVOINE, Université de Bourgogne, Prof. Olivier HAEBERLE, Université de Haute-Alsace, Dr. Brahim TAMADAZTE, CR2 CNRS, AS2M, FEMTO-ST institute, for their patience in reading my thesis document and providing me with useful suggestions in extending my work as well as in improving my final report.

My very special thanks go to Dr. Joel AGNUS, Dr. Brahim TAMADAZTE and Mr. David GUIBERT for sharing their experience with me. Joel's help is an unforgettable one. He is my first teacher who taught me the basic details, usage and the operation (especially correcting the astigmatism) of a typical scanning electron microscope. Moreover, he is very instrumental and helped me with all the hardware problems. Brahim's

support is outstanding. He helped me a lot in developing the new visual servoing-based algorithms for autofocusing and nanopositioning using a scanning electron microscope. The discussions I had with Brahim were invaluable. Apart from this, I would also like to thank Brahim for finding my apartment in Crous. Finally, without David's support, the current mechanical set-up inside the microscope's vacuum chamber is almost impossible.

Next, I extend my thanks to Dr. Roland SALUT of MIMENTO for his support in using their scanning electron microscope and also for providing me a complete dataset of images to compare my results on signal-to-noise ratio evaluation. I would like to sincerely thank the administrative staff of our department: Isabelle GABET, Sandrine FRANCHI and Martine AZEMA for their support in completing the never ending paperwork.

Then, I thank my friends and the lab mates Nandish R Calchand, Sergio Lescano, Raj dhara, Ravinder Chutani, Ahmed Mosallam, Zill-e- Hussnain, Lisa Serir, Amelie cot, Margot Billot, Vincent Chalvet, Hector Ramirez and the other department staff for creating and maintaining a friendly and pleasant work atmosphere. Out of all my very special thanks go to Nandish and Sergio who helped me a lot with the French administration process and paperwork. I am also very grateful to all my colleagues who created a friendly atmosphere during the morning coffee which is quite unforgettable and in fact very addictive.

I wish to thank my parents, Vani and Sarma, my sister, Pallavi and my grandmother, Sri lakshmamma. Their love, affection, support and encouragement has always driven me in pursuing whatever I want. Last but not the least, I thank my loving wife and the better half, Sindhu, for being on my side, understanding the work pressure especially during the writing part and encouraging me in all times.

Finally, I thank all my friends in India and Sweden for supporting me all the time.

# Contents

# List of Figures

# List of Tables

# List of Algorithms

# List of Abbreviations

| | |
|---|---|
| **2 1/2 D** | Two and a half dimension. |
| **2D** | Two dimensional. |
| **3D** | Three dimensional. |
| **ACF** | Auto-correlation function. |
| **AFM** | Atomic force microscope. |
| **ARMA** | Autoregressive moving average. |
| **BRIEF** | Binary robust independent elementary features. |
| **BSE** | Back scattered electrons. |
| **CAD** | Computer-aided design. |
| **CCD** | Charge coupled device. |
| **CMOS** | Complementary metal oxide semiconductor. |
| **CNT** | Carbon nanotubes. |
| **CPU** | Central processing unit. |
| **DFT** | Discrete Fourier transform. |
| **DISS** | Digital image scanning system. |
| **DNA** | Deoxyribonucleic acid. |
| **EBID** | Electron beam induced decomposition. |
| **FAST** | Features from accelerated segment test. |
| **FEG** | Field emission gun. |

| | |
|---|---|
| **FIB** | Focused ion beam. |
| **GUI** | Graphical user interface. |
| **IBVS** | Image-based visual servoing. |
| **ICP** | Inductively coupled plasma. |
| **InGaAs** | Gallium indium arsenide. |
| **IP** | Internet protocol. |
| **MLTDEAR** | Mixed Lagrange time delay estimation autoregressive. |
| **MSE** | Mean squared error. |
| **NEMS** | Nano electromechanical systems. |
| **NSOM** | Near-field scanning optical microscope. |
| **ORB** | Orient FAST and rotated BRIEF. |
| **PBVS** | Position-based visual servoing. |
| **PC** | Personal computer. |
| **RAM** | Random access memory. |
| **RANSAC** | Random sample consensus. |
| **ROI** | Region of interest. |
| **SE** | Secondary electron. |
| **SEM** | Scanning electron microscope. |
| **SIFT** | Scale-invariant feature transform. |
| **SNR** | Signal-to-noise ratio. |
| **SPM** | Scanning probe microscope. |
| **STM** | Scanning tunneling microscope. |
| **SURF** | Speeded up robust features. |
| **TCP** | Transmission control protocol. |
| **TEM** | Transmission electron microscope. |
| **USB** | Universal serial bus. |
| **ViSP** | Visual servoing platform. |

Chapter 1

# Context and contributions

Nowadays, nanotechnology is becoming more prominent in the world of sciences. It is expected to play an important role in the coming years with a wide range of applications in the fields of manufacturing, information technology, electronics, and healthcare. Due to the latest advances in nanotechnology, nanorobotics and multi-physics nanocharacterization have gained a significant research interest. They have shown a considerable progress in the recent years, as it became possible to develop novel nanoscale devices and systems with increasing efficiency. The consequence of this growth is an increase in the need for developing reliable and state of the art techniques for nanomanipulation and characterization. Various works that are performed in these areas show a great progress; however, most of them are still isolated operations. This work aims to capitalize on the recent surge in nanorobotics and robust control strategies for manipulation and try to apply them to this problem.

The work performed during this thesis is mostly concentrated or revolved around two themes: vision using SEM towards material characterization and automatic manipulation, and visual servoing for precise nanopositioning. This work (using SEM vision) is a first of its kind that is performed in the department of AS2M (automation and micromechatronics systems). It has been performed in the context of ANR P2N NANOROBUST project that is briefly presented below.

## 1.1 Context

### 1.1.1 NANOROBUST project

The term NANOROBUST is the abbreviated form of the French national project entitled "multi-physics characterization and robotic manipulation of nano-objects under SEM". It was started on November 28, 2011 with a duration of 48 months. The overall work is divided among the four partners of the project: FEMTO-ST (Besançon), ISIR (Paris), IRISA (Rennes) and LPN (Marcoussis). The main objective of this project is to perform fundamental research in the field of nanoscale characterization and manipulation.

The two typical applications that are studied as a part of this project are the in situ characterization of nanostructures for optical NEMS (nanoelectromechanical systems) such as quantum dots, nanowires and nanomembranes, and manipulation of nano-objects under SEM (scanning electron microscope) for observation with TEM (transmission electron microscope). Two types of characterization are considered. The first one aims to determine the physical properties (mechanical, thermal, electrical) of the nanostructures. The initial step in this process is the extraction and placement of these nanostructures on the characterization support. Then the robotic systems can move the sensors along a predefined path for property analysis of the positioned structure. These sensors can be either passive such as a simple AFM (atomic force microscope) tip whose deformation is measured from images or active such as a tuning fork that operates on the frequency drift to evaluate the stiffness of nano-objects.

The second type of characterization is aimed to determine the structural properties of the nanostructures for visual observation with TEM. Analysis of these objects requires precise positioning on a grid that is used inside TEM. A typical example is quantum dots where the nano-objects measure a few nanometers up to a few tens of nanometers in diameter. Standard preparation techniques become unsuitable for observing such nano-objects mostly because of the size and also very time consuming because of their low density. A solution could be by etching the micron sized walls on a semiconductor where the quantum dots will lie on top of these walls. The walls will then be displaced by a controlled nanomanipulation and placed on the TEM observation grid (see figure 1.1).

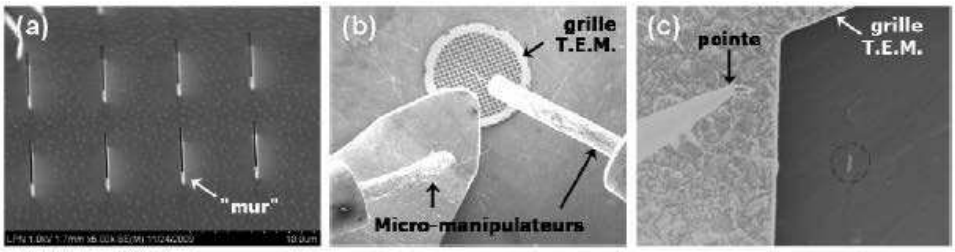



Figure 1.1: Illustration of the problem of handling a nano object on a TEM grid. (a) A series of walls of 2 $\mu m$ obtained by ICP (inductively coupled plasma) etching on epitaxial substrate. (b) Observation TEM grid held by a gripper with two micromanipulators. (c) TEM grid with a ”wall” made (in the circle).

Both tasks are challenging mainly because of the size of nanostructures and requires the development of

- Accurate and reliable handling
- Precise and robust positioning schemes
- Defectless or distortion free SEM imaging

### 1.1.2 Project organization and structure

The overall project can be described in the following steps

- Preparing the nano-objects (preparation)
- Handle, transport and placement of nano-objects (manipulation)
- Performing characterization

Figure 1.2 shows these three stages taking place in situ in the SEM.

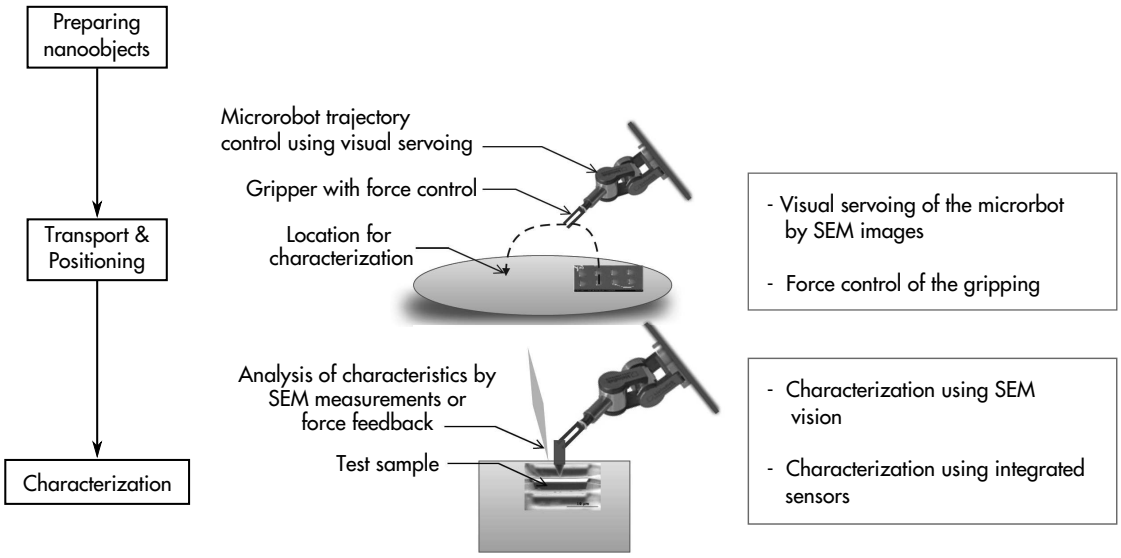



Figure 1.2: Studied process types of the NANOROBUST project.

The overall structure of the project is comprised of six tasks (see figure 1.3) that are briefly presented below.

**Task-1: Coordination**
This task aims for efficient coordination of different stakeholders. All the partners are responsible for this task.

**Task-2: Strategies for micro / nano-handling and characterization**
The main role of this task is to:

- Define gripping strategies that are adapted to the range of objects with different shape and size (e.g. handling / gripping objects like nanoballs or nanowires that may be embedded in a substrate).
- Define techniques for physical characterization of the objects.

ISIR and FEMTO-ST are responsible for this task.

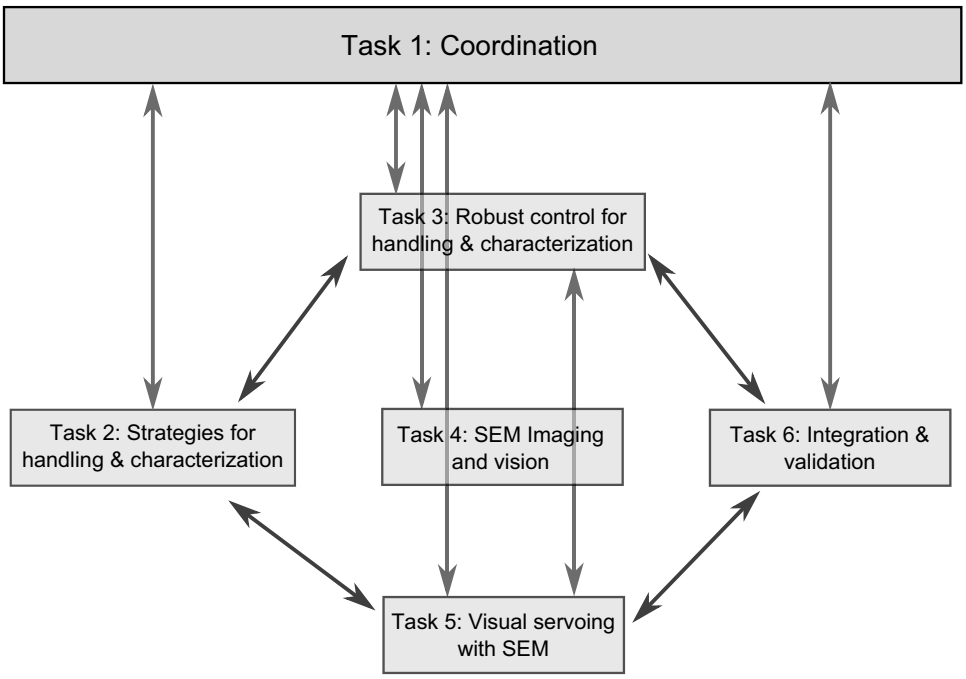



Figure 1.3: NANOROBUST project tasks and coordination.

**Task-3: Robust control for precise handling and characterization**
The main objective of this task is to:

- Study and model the working conditions in a SEM, particularly noise and perturbations.
- Develop robust control laws for the control of micro / nanorobots.
- Increase the accuracy of micro / nanorobots by using refined control methods.

FEMTO-ST and ISIR are responsible for this task.

**Task-4: SEM imaging and vision**
This task aims to perform:

- Evaluation and compensation of noise and drift in real-time.
- Modelling and calibration of distortion and projection.

FEMTO-ST and IRISA are responsible for this task.

**Task-5: Visual servoing with SEM**
This task aims to achieve the manipulation of objects by visual servoing approaches. It includes:

- Definition of visual features from the images.

- Computing the interaction matrices.
- Calculation of the control laws.

IRISA and FEMTO-ST are responsible for this task.

**Task-6: Integration and validation**
This is the final task to integrate all the other tasks and to validate the performance of manipulation techniques in SEM for characterization of nanoscale objects. LPN is responsible for this task.

## 1.2 Thesis outline - author's contributions

The work performed during this thesis is included in the tasks 4 (chapter 3) and 5 (chapters 4 and 5) of the project. The remainder of this document showing the author's contributions towards vision and visual servoing for nanomanipulation and nanocharacterization is organized as follows:

**Chapter 2** presents the existing state of the art in the field of nanorobotics and nanomanipulation. It provides a comprehensive study and a comparison of various microscopic systems available for imaging at nanoscale. Different nanomanipulation systems that are used in other works are analyzed along with their corresponding applications. Apart from this, an overall description of SEM imaging system along with the used hardware set-up is detailed at the end.

**Chapter 3** provides solution for two major problems with SEM imaging: noise and drift. It is a well-known fact that SEM imaging is mostly affected by the non-linearities and instabilities present in the electron column along with the addition of huge amount of noise with high scan rates. This chapter starts with a detailed description towards the noise in SEM imaging. Different techniques and methods that are available for quantifying the level of noise are studied. A simple and real-time method that has been developed for quantifying the noise level and monitoring the image quality is presented. For the problem of drift, an image registration-based method is presented for correcting the drift in real-time. It has been compared with the well-known correlation-based method. Drift compensation is performed at high magnifications ($10k\times - 30k\times$) using a gold on carbon sample.

**Chapter 4** presents the work performed using SEM image focus i.e. autofocus, depth estimation from focus for nanomanipulation and shape reconstruction from focus. It mainly highlights the importance of focus with SEM imaging. While performing an automated vision-based manipulation task, it is important to provide sharp images in order to gain full advantage of the visual measurements as well as to accomplish the task with a good precision. To solve this, a visual servoing-based method to perform efficient autofocusing has been developed. The developed method is robust and fast in

comparison to the traditional search-based techniques. Besides, acquiring inter-object depth information to use it for nanomanipulation is a crucial and difficult task with SEM, since it possess a single imaging sensor. This problem is solved by performing region-based autofocusing. Apart from these, a focus-based method to reconstruct the 3D shape of the objects being viewed is presented. It helps in gaining the knowledge of imaging scene.

**Chapter 5** presents the vision-based automatic nanopositioning under SEM. It starts with an overview of the key topics in nanopositioning and state of the art in visual servoing. The nanopositioning studied in this chapter is to perform characterization of microstructures by probing to measure the structure stiffness and to perform in situ manipulation of micro-nanostructures. Because of the presence of high amount of imaging noise and distortions, it is difficult to rely on any feature tracking-based methods. Considering this problem, two visual servoing-based approaches are presented that do not require any tracking and instead, they use the complete image information. The first one is the intensity-based method that has been implemented using the idea of photometric visual servoing approach. This method uses the overall pixel intensities as visual measurements. The second one is a Fourier domain method. With this method, the 2D motion between the images is considered as visual measurement. For a task of positioning silicon micro-objects using both methods, the control strategies are designed to minimize the positioning error by controlling the platform movement.

**Chapter 6** is dedicated to the software development during the thesis. In fact, this thesis is started from the scratch and it is our primary objective to set-up and program different modules for synchronous operation. Out of all, the most important and primary task performed is the development of software libraries for continuous image acquisition and SEM parameter control. A software user interface application has been developed in order to reduce the operational complexity and process monitoring.

## 1.3 List of author's Publications

### International conferences and workshops

#### Accepted and published

1. **Naresh Marturi**, Brahim Tamadazte, Sounkalo Dembélé, and Nadine Piat, "Visual Servoing Schemes for Automatic Nanopositioning in Scanning Electron Microscope", In *IEEE International Conference on Robotics and Automation (ICRA'14)*,in-press.

2. **Naresh Marturi**, Sounkalo Dembélé, and Nadine Piat, "Depth and Shape Estimation from Focus in Scanning Electron Microscope for Micromanipulation", In *IEEE International Conference on Control, Automation, Robotics and Embedded systems (CARE'13)*, in-press.

3. **Naresh Marturi**, Brahim Tamadazte, Sounkalo Dembélé, and Nadine Piat, "Visual Servoing-Based Approach for Efficient Autofocusing in Scanning Electron Microscope", In *proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'13)*, pp. 2677-2682, 2013.

4. **Naresh Marturi**, Sounkalo Dembélé, and Nadine Piat, "Fast Image Drift Compensation in Scanning Electron Microscope Using Image Registration", In *proceedings of the IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 807-812, 2013.

5. **Naresh Marturi**, Sounkalo Dembélé, and Nadine Piat, "Performance evaluation of scanning electron microscopes using signal-to-noise ratio", In *International Workshop on MicroFactories (IWMF'12)*, Finland, pp. 1-6, 2012.

6. Sounkalo Dembélé, Nadine Piat, **Naresh Marturi**, Brahim Tamadazte, "Gluing free assembly of an advanced 3D structure using visual servoing", In *Micromechanics and Microsystems Europe Workshop*, Germany, 2012.

## International journals

### Accepted

7. Abed Malti, Sounkalo Dembélé, Nadine Piat, Claire Arnoult, **Naresh Marturi**, "Toward Fast Calibration of Global Drift in Scanning Electron Microscopes with Respect to Time and Magnification", In *International Journal of Optomechatronics*, Taylor and Francis, vol. 6, no. 1, pp. 1-16, 2012.

8. **Naresh Marturi**, Sounkalo Dembélé, and Nadine Piat, "Monitoring the Signal to Noise Ratio of a Tungsten Gun Scanning Electron Microscope for Micro-Nanomanipulation", In *Scanning*, in-press.

### In preparation

9. **Naresh Marturi**, Sounkalo Dembélé, and Nadine Piat, "Efficient Depth estimation Technique in Scanning Electron Microscope for Nanomanipulation", In preparation.

10. **Naresh Marturi**, Sounkalo Dembélé, and Nadine Piat, "Accurate Nanopositioning in Scanning Electron Microscope for nanocharacterization", In preparation.

Chapter 2

# Introduction to nanomanipulation and SEM imaging

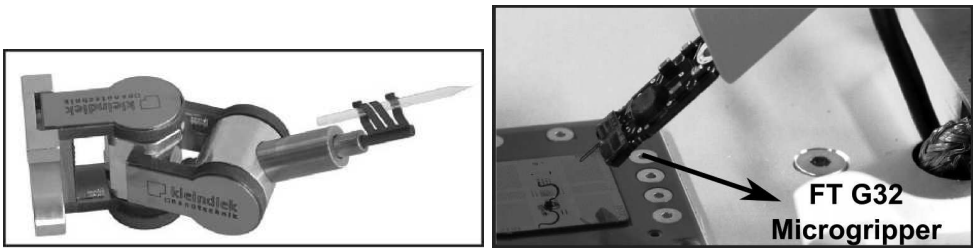
*In this chapter, various background details regarding the nanoscale manipulation are presented. Starting from the basic topics related to the nanorobotics and nanomanipulation, we approach and highlight the role of the imaging sensors in assisting the manipulation of objects at nanoscale. Apart from this, a complete description of a SEM system along with the system hardware set-up used for this thesis are presented.*

## 2.1 Background

With a rapid development in micro-nanoscale technology in the last couple of decades, handling nanometric objects for building complex NEMS-based products has gained a significant research interest. Since the human handling is not a feasible option (almost impossible) at this particular scale, nanorobotics has been emerged as a new field to tackle this critical issue. In general, Nanorobotics deals with the study, design and application of robotic devices close to nanoscale. Specifically, it is concerned with manipulation and assembly of nano-objects using micro or macro devices and programming of robots with overall dimensions at nanoscale [Req03]. It integrates together various disciplines including nanofabrication processes that are used to produce nanoactuators, nanosensors etc., nanorobotic assembly, self-assembly, nanomaterial synthesis, nanocharacterization and nanobiotechnology [DN07].

Being a new field, nanorobotics has been mainly classified into two research areas [Req03, Sit07]. The primary area concerns developing new methods for nanomanipulation and assembly of micro-nanometric objects of interest. Here, the main idea is to design and control large micro-nanorobotic systems that are able to handle nanometric objects with good precision [FE08, BF12]. These approaches are called as *top-down* methods that use well-known processes from semiconductor industry to fabricate small structures with optical lithography. On the other hand, the second area i.e. the *bottom-up* approach is dedicated to manufacture new mobile robotic systems with various capabilities whose overall size is in the range of few micrometers. The bottom-up strategies are generally assembly-based techniques and include techniques such as self-assembly, directed self-assembly etc. For developing these systems, the overall size is a major constraint that makes it challenging to integrate different devices like actuators, sensors etc. altogether. Even though the size miniaturization has been progressed recently using microfabrication techniques, there exist a lot of difficulties in achieving further miniaturizations. Due to this problem, not much work related to nanomanipulation has been performed in this area. In the meanwhile, the former approach has turned out to be one of the possible solutions for handling nano-objects.

In general, nanorobotic manipulation can be defined as the manipulation of nanoscale objects where, contact and surface forces dominate the volume related properties affecting: handling (pick, place, release etc.) and positioning (accuracy, rotation, orientation etc.) operations. So far, considerable research has been performed towards the development of these systems [Sit01]. A basic nanomanipulation system consists of nanomanipulators (figure 2.1(a)) and platform stages for positioning, microscopes for task space perception, handling devices such as microgrippers (see figure 2.1(b)), microprobes for object grasping and different sensors that are integrated with gripper fingers to assist the manipulation process. Apart from these, it also contains a programmed interface for synchronous operation of all the system components. Basically, all these components are mounted on a vibration isolation system to reduce the effect of external environmental disturbances. Most of the time, the manipulation tasks are performed in high vacuum conditions to prevent the sample contamination. The schematic diagram of a basic nanomanipulation system is shown in the figure 2.2.




(a) (b)

Figure 2.1: (a) Kleindiek MM3A robotic micromanipulator [Gmb] (b) FEMTO TOOLS FTG 32 microgripper [Fem11].

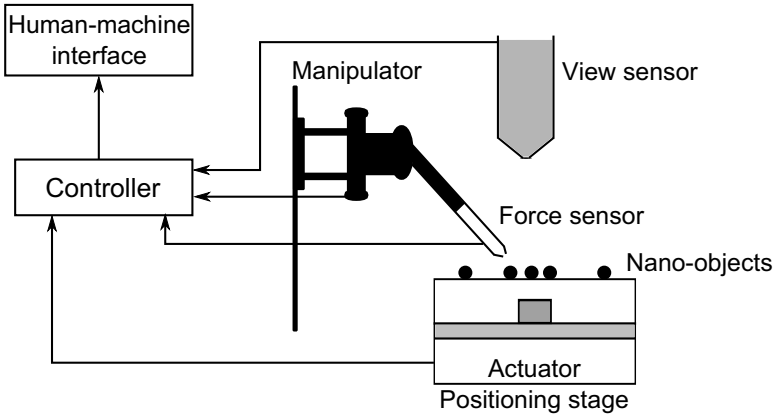



Figure 2.2: Schematic diagram of a basic nanomanipulation system.

### 2.1.1 Nanoscale imaging

In 1878, Ernst Abbe proved that the resolution of the optical microscope is limited by the wavelength of light and the smallest detail that can be resolved is of the order of few micrometers [Fre63]. With the two famous discoveries: the moving electron wave properties by de Broglie in 1924 and the effect of magnetic coil on an electron beam by Busch in 1926-27, many researchers started to work on imaging extremely small objects with electron beam. In 1931, Max Knoll and Ernst Ruska from Technical University of Berlin, Germany successfully built and demonstrated the first electron microscope (figure 2.3).

Apart from electron microscopy that includes TEM and SEM, nanoscale imaging can also be carried out using SPM (scanning probe microscope) technique that includes AFM, STM (scanning tunneling microscope) and NSOM (near-field scanning optical microscope). Except SEM that has been explained in the next section, the rest of the imaging techniques are briefly explained below. A comparison of all these devices is provided in the table 2.1.

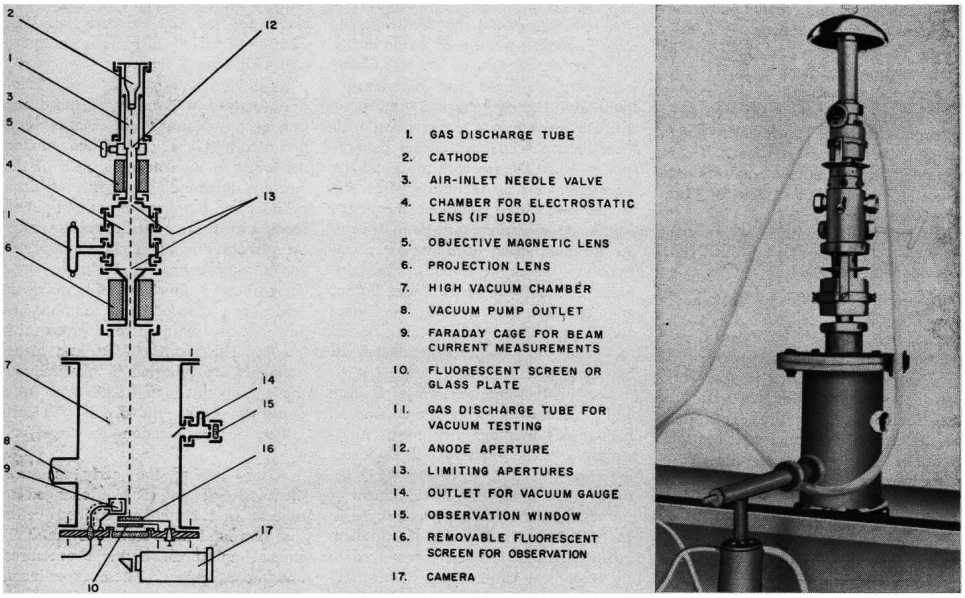



Figure 2.3: The first electron microscope demonstrated by Max Knoll and Ernst Ruska in 1931 [Fre63].

### Transmission electron microscope

Belonging to the family of electron microscopes, a TEM (figure 2.4) uses high energy electrons for imaging. It was built by James Hillier and Albert Prebus at the University of Toronto in 1938. It operates at high voltages ranging from $50-1000\ kV$ and provides a resolution up to $0.1\ nm$. With TEM, a beam of electrons are produced and transmitted through an ultra-thin sample. Later, the unscattered electrons that are transmitted through the sample hit a fluorescent screen at the bottom of the microscope producing the image. By adjusting the voltage of the gun, the speed of the electrons can be modified which inturn modifies the image. Generally, the images produced by TEM are gray scale and can be interpreted as follows: the lighter areas represent the higher number of electrons transmitted and the darker areas represent lower number and dense areas on the sample. To use with TEM, samples need to be sliced thin enough for electrons to transmit.

### Scanning tunneling microscope

STM (figure 2.5) is the first type of SPM that was invented by Binnig and Rohrer in 1982. The STM is comprised of a piezoelectric linear stage and an extremely fine conducting probe that is held close to the sample surface. The probe tip is extremely sharp formed by a single atom. This atomic sized tip is moved across the sample surface in a raster pattern during which the electrons tunnel between the surface and probe generating an electric signal. In order to maintain a constant signal, the probe is raised and lowered during the scanning process. The current used to keep the signal constant

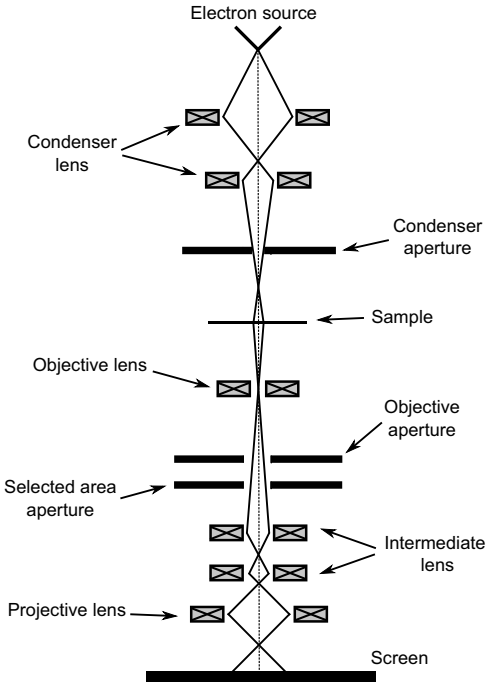



Figure 2.4: Transmission electron microscope schematic diagram.

is used in generating the images. Similar to TEM, a STM is also capable of providing atomic resolution. However, STM can be used only with conducting sample surfaces and provides only surface topographical information.

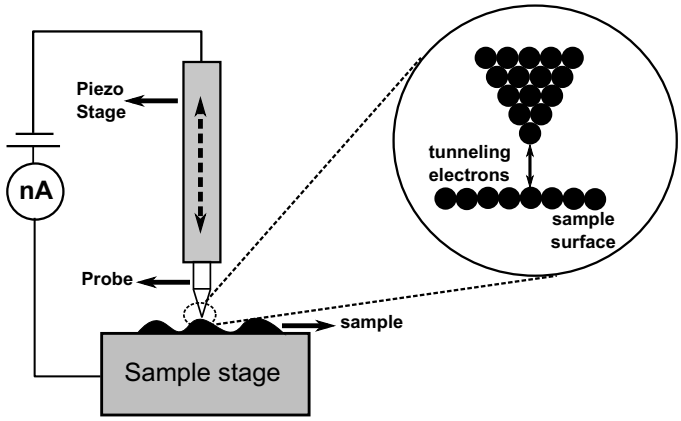



Figure 2.5: Scanning tunneling microscope schematic diagram.

**Atomic force microscope**

The AFM (figure 2.6) was invented by Binnig, Quate, and Gerber in 1986 to overcome the conductivity requirement associated with STM. Similar to STM, an AFM uses a cantilever (made of silicon or silicon nitride) with a sharp tip (radius of few nanometers) to scan the sample surface. When the tip is placed near the sample surface, the forces between the surface and tip deform the cantilever according to Hooke's law [BQG86]. In a typical system, this deflection is monitored by reflecting a laser from the top of the cantilever on to an array of photodiodes. Based on the type of application, different imaging modes (contact, non-contact, lateral force, phase imaging etc.) can be used with AFM. The main advantage of AFM is that it can image any kind of surfaces. The lateral resolution provided by an AFM is 1 $nm$.

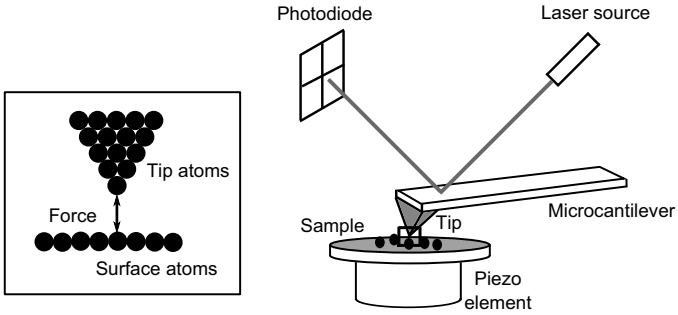



Figure 2.6: Atomic force microscope schematic diagram.

Table 2.1: Comparison of different microscopes used for nanomanipulation.

| **Property** | **STM** | **AFM** | **TEM** | **SEM** |
|---|---|---|---|---|
| Imaging principle | tunneling electrons | inter-atomic forces | electrons | electrons |
| Environment | air / vacuum | all | vacuum | vacuum |
| Dimension | 3D | 3D | 2D | 2D |
| Samples | conductors semi-conductors | All | conductors semi-conductors | conductors semi-conductors |
| Resolution | $> 0.01\ nm$ | $> 0.1\ nm$ | $> 0.01\ nm$ | $> 1\ nm$ |
| Interaction | non-contact | non-contact / contact | non-contact | non-contact |

### 2.1.2 Nanomanipulation systems

The existing nanomanipulation systems can be broadly classified into different types based on the starting point, manipulator-object interaction, utilized nanomanipulator interaction and the technique used for overall system operation as shown in figure 2.7 [Sit01]. The first works on nanomanipulation are known to be performed in 1990 by Eigler and Schweizer [ES90], who managed to position individual Xenon atoms

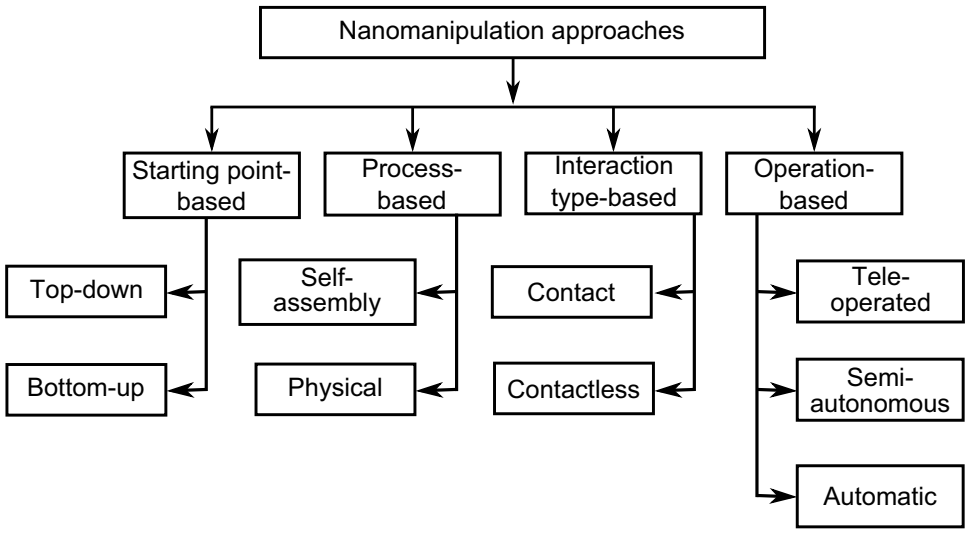



Figure 2.7: Classification of nanomanipulation systems provided by Sitti [Sit01].

in the shape of IBM logo on a single-crystal nickel surface at low temperatures using STM(figure 2.8). Since then, many research groups started working on nanomanipulation using AFM [SRPA95, HKS+98, VSZ04]. The main advantage of using the AFM as a micro-nanomanipulator is its use of single sensor to sense force and 3D topography. Many research groups have showed 2D precise positioning of nanoparticles or CNTs (carbon nanotubes) on surfaces using contact pushing following a *look and move* control scheme. However, the main difficulty comes in imaging and manipulating the object simultaneously in real time. To solve this, many works have used AFM probes with SEM as it is capable of scanning at near real-time [EG94, FSF02, MWJF10]. Apart from using AFM tips, many systems have used piezoelectric actuators for accurate nanopositioning during nanomanipulation as they are capable of providing smaller displacements. These devices are operated by varying voltage across the piezoelectric element. However, these actuators suffer from a major problem of non-linear hysteresis.

For the first time, micro-nanomanipulation using a SEM has been reported and performed by Hatamura and Morishita [HM90] and Sato et al. [SKMH95]. Their manipulation system consists of two micromanipulators to manipulate micro-objects smaller than 100 $\mu m$. In their work, to achieve high positioning accuracy, piezoelectric actuators and parallel plate structures are used. Yu et al. have presented a nanomanipulation system to perform 3D manipulation and characterization of CNTs [YDS+99]. The manipulation is performed in open-loop under slow-scan mode. Fukuda et al. has presented a 16 degrees of freedom nanomanipulation system positioned inside a SEM chamber for the assembly of nanodevices with multiwalled carbon nanotubes [FAD03]. Using this system, they have performed the connections between nanotubes using EBID (electron beam induced decomposition) and mechanochemical bonding. Thereafter, many research groups have reported manipulation of CNTs under SEM [NA03, MHMB04]. In their work, Molhave et al. have used microfabricated, electrostatically actuated tweezers for handling the

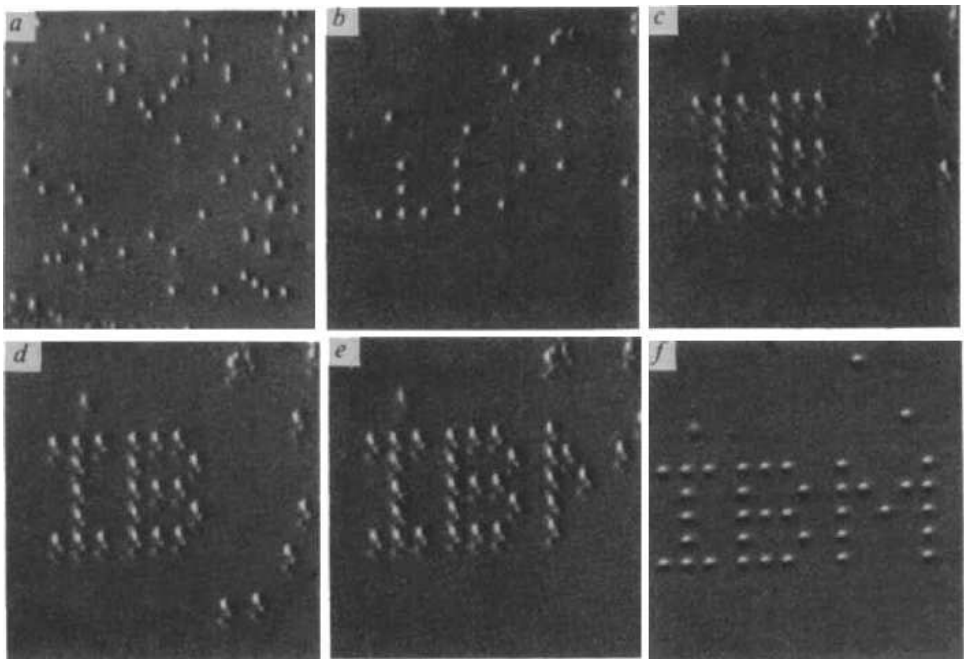



Figure 2.8: A sequence of images taken during the manipulation and arrangement of Xenon atoms [ES90].

nano-objects [MHMB04]. Fatikow et al. have presented an automated nanohandling station inside the SEM chamber [FWH+07]. It consists of two multi degrees of freedom mobile microrobots for positioning the specimen and the end-effector, a CCD (charge coupled device) camera for monitoring platform movement and a touchdown sensor for contact detection and depth estimation. With this setup, they have demonstrated an automated task of handling TEM-lamellae using vision feedback. Considering the problem of using SEM images as feedback information i.e. too noisy with high scan speeds and slow acquisition rate, recently Jasper has presented a new line scan-based position tracking system that is integrated with SEM [Jas09, Jas11]. The new system can determine the position of the objects using few line scans instead of acquiring global image.

### 2.1.3 Applications of nanomanipulation

So far in the literature, nanomanipulation systems are widely used in many fields for various applications. Some of them are listed below.

**Biotechnology:**
The development of new nanomanipulation techniques has given researchers the ability to manipulate single biomolecules and to record mechanical events of biomolecules at the single molecule level [IIY01]. It is also used for controlled manipulation of biological cells,

molecules and tissues, DNA (deoxyribonucleic acid), proteins, adenovirus etc. [GMN+99, GFM+00]. Besides, it is also used for measuring the twisting and bending compliance of DNA [Was03].

**Material science:**
Nanomanipulation plays an important role in building new materials and for analyzing the material properties such as friction, adhesion, electrical, optical etc. As mentioned earlier, many research groups have been working on manipulating and assembling of CNTs. The recent research investigations on CNTs report many suitable properties to use them in nanoscale electronics, mechanics etc. Apart from property analysis, nanomanipulation is also used in preparing samples such as TEM lamella (figure 2.9), nanowires etc.

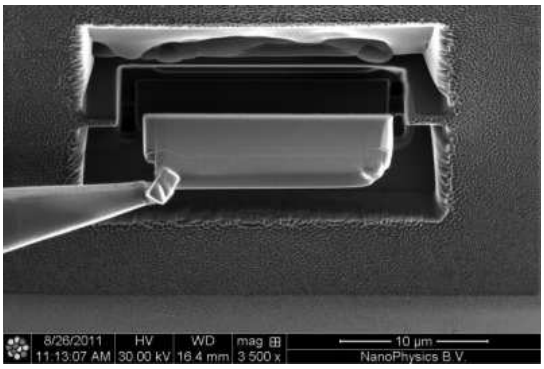



Figure 2.9: TEM sample lift-out using an extraction needle [Nan12].

Apart from the above mentioned applications, nanomanipulation has also its applications in computer technology for data storage [BDD+99, JJJ+09].

## 2.2 SEM Imaging

The SEM is a most widely used instrument of all the available electron microscopes. One of the most important aspects of SEM is its ability in producing the images of three-dimensional specimens at high magnifications and high resolution (better than 1 $nm$). Unlike optical microscopes, a SEM produces images of a sample by raster scanning the sample surface with a focused beam of high energy electrons. In this section, various details regarding the SEM instrumentation and its operating principle are explained.

### 2.2.1 SEM components

The basic architecture along with the major components present inside a SEM electron column is shown in the figure 2.10. All these components perform different roles in generating the electron micrographs. The main components are described below and more details can be found in [GNJ+03].

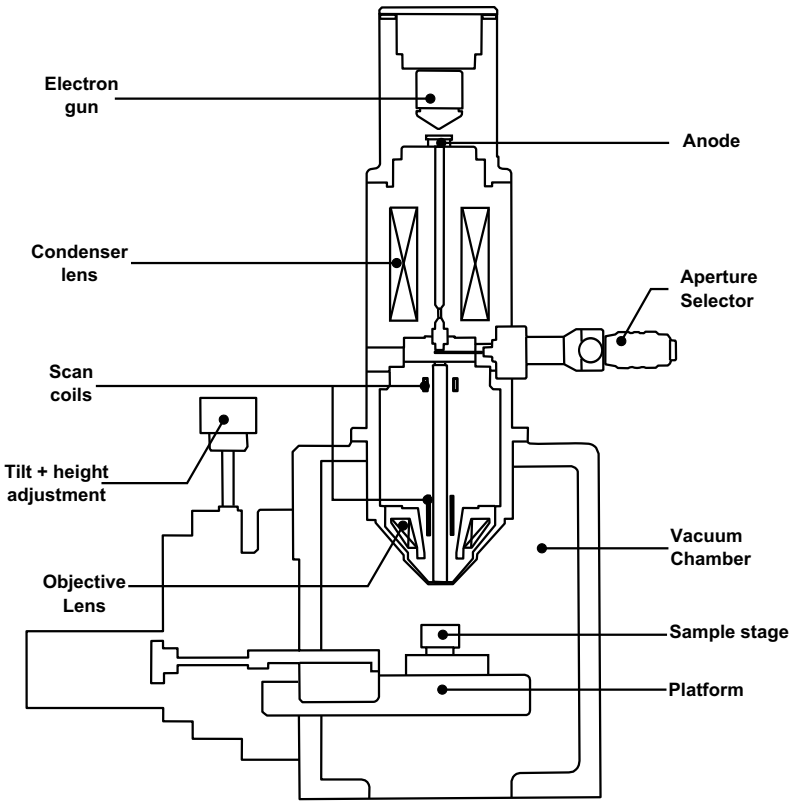



Figure 2.10: Conventional SEM architecture illustrating various components.

**Vacuum system**

A vacuum system is one of the mandatory requirements when working with electron beam. It is mainly responsible to avoid collisions between electrons and the extraneous gas molecules and protecting the filament from oxidation. Without adequate vacuum inside the SEM column, an electron beam can neither be created nor controlled. Typically, most of the available SEMs operate at vacuum of $10^{-4}$ to $10^{-6}$ Torr. In order to achieve this pressure, two types of pumps are used: low vacuum pump or roughing pump and oil diffusion pump. Roughing pump is a type of mechanical pump that provides pressure down to $10^{-3}$ Torr and thereafter high vacuum is achieved by oil diffusion pump.

**The electron gun**

An electron gun (see figure 2.11) is comprised of an electron emitter, a Wehnelt cylinder and an anode. Available electron guns with present day SEMs are conventional electron guns using Tungsten (W) hair-pin, or lanthanum hexaboride (LaB6) tips and field emission guns (FEG) . With the conventional guns, a stable number of electrons are emitted by supplying the necessary filament heating current. On the other hand, FEG (cold emission source) uses the difference in the potential between cathode and anode to excite the electrons. The Wehnelt cylin-

der that encloses the filament acts as a cathode. It contains a small hole in the center through which the generated electrons are centered and exited. Finally, a positively charged anode that is rested under the Wehnelt cap is used to accelerate the electrons. The portion of the electrons that passes through the anode is called as beam current. In general, the acceleration voltage that is supplied to anode varies from less than $1kV$ to $30kV$. A high value for acceleration voltage reduces the beam diameter and increases the resolution. But a too high value leads in charging up and damaging the specimen.

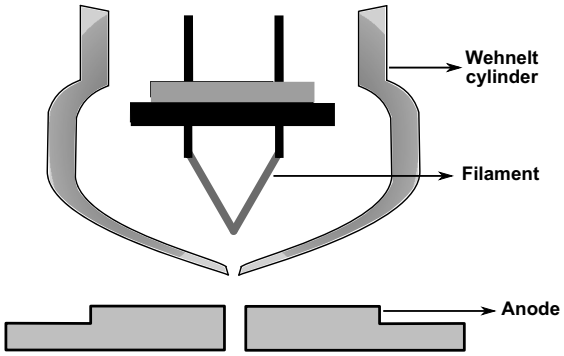



Figure 2.11: Schematic diagram of a SEM electron gun showing various components.

**Electromagnetic lenses**

The two sets of electromagnetic lenses that are available in an electron column are the condenser lenses and the objective lenses. Condenser lenses lying above the aperture strip are mainly responsible for controlling the electron concentration and the diameter of the beam. Objectives lenses that are present under the aperture converges the incoming beam and focus it on the sample surface.

**Objective aperture**

It is a metal rod holding a thin metallic strip (aperture strip) containing small holes (apertures) of different diameters. It is mainly responsible to filter the low energy and non-directional electrons. The aperture size is selected based on the type of the work. For low magnifications a larger aperture can be used and for high resolution imaging a small aperture is required. However, a smaller aperture leads in dark images as the count of the electrons interacting with sample surface is less.

**Scan coils**

The two sets of scan coils that are available in the electron column are used to raster the electron beam in both horizontal and vertical directions on the sample surface. The first pair deflects the beam off the optical axis and the second pair bends the beam back on to the axis at the pivot point of the scan (see figure 2.12). Apart from scan generation, they are also used in controlling the magnification of the instrument.

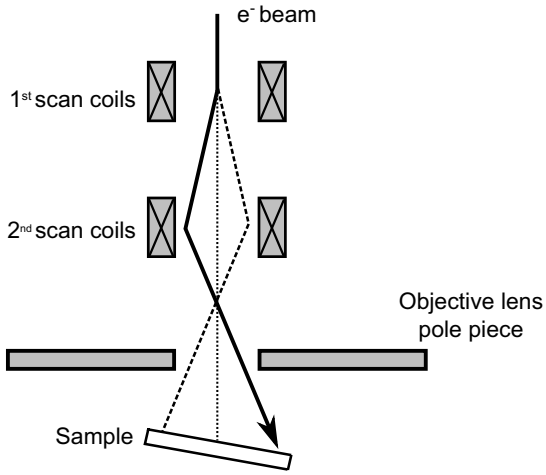



Figure 2.12: Beam deflection by scan coils inside the SEM.

**Electron detectors**

In general, a SEM consists of various detectors each of which is used to collect different types of emitted electrons. The commonly used detectors are SE detector and BSE detector. Two types of SE detectors are available with the modern SEMs. The primary type is the most commonly used Everhart-Thornley detector that is fixed to the side walls of the chamber (see figure 2.13). It works by attracting the emitted secondary electrons by a positive potential applied to a ring around the detector. The second type is in-lens detector that is fixed alongside the electromagnetic lens. BSE detectors are located just below the pole piece of the object lens and collects the emitted back scattered electrons.

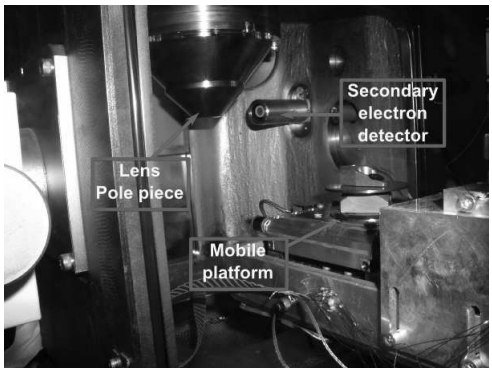



Figure 2.13: SEM chamber showing the secondary electron detector and mobile platform to place samples.

**Positioning platform**

A movable positioning platform is placed in the lower part of the SEM electron

column and is used to load the specimen. It can be controlled using a joystick and contains translational ($X$, $Y$, $Z$), rotational (around $Z$) and tilt (horizontal) movements. It is mainly responsible for better positioning the sample such that it is well exposed to the electron beam.

### 2.2.2 Operation principle

Before dealing with the actual image acquisition process, it is important to know the electron beam interaction with the specimen surface and the emission of various signals that are resulted during this process.

#### Beam - sample interaction

When a focused beam of electrons hits the sample surface, they penetrates into it for some distance (about $1\mu m$) before hitting another particle. These incident electrons are greatly scattered resulting in elastic and inelastic scattering inside the sample forming a region called as *interaction volume*[1]. Elastic scattering results in BSEs (back scattered electrons) where as inelastic scattering produces SEs (secondary electrons), Auger electrons and X-rays. Figure 2.14 shows the interaction of electron beam with a specimen surface and signal emission. The resulted electrons are recorded at their respective detectors. Out of all the emitted signals, SE and BSEs are the most widely used ones and are explained below. More information about SEM interaction volume can be found in [GNJ+03, Rei98].

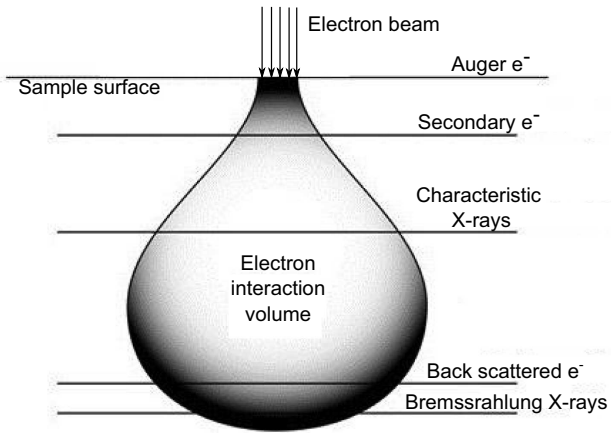
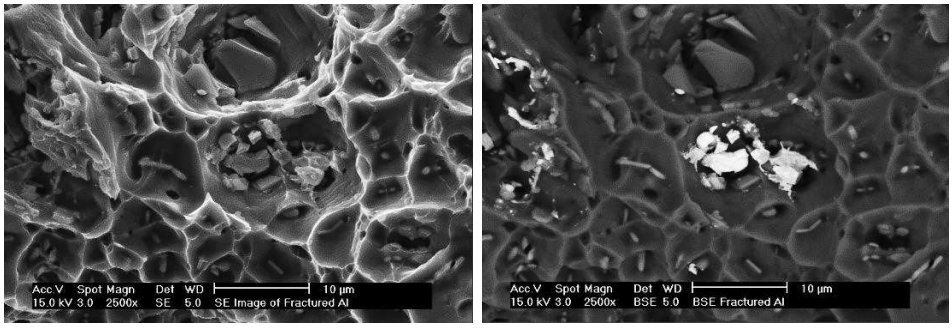



Figure 2.14: Interaction volume and signal emission.

[1]The resulting region over which the incident electrons interact with the sample to produce different signals is termed as interaction volume. Its volume depends on the beam acceleration voltage, specimen topography and the angle of incidence of the primary beam.

**Back scattered electrons:** These are the original beam of electrons that are escaped after the beam-sample interaction. These electrons have approximately the same energy as that of the primary ones and are easy to record. The images formed using the BSEs provide compositional differences, specimen topography, crystal orientation and grain boundaries. Figure 2.15(b) shows the BSE image of a fractured aluminium alloy.

**Secondary electrons:** These electrons are produced when an incident electron excites a weekly bonded outer shell electron (in conductors) or a valence electron (in semi-conductors and insulators). During this process the electrons from the specimen receive kinetic energy from the incident electrons and start moving towards the surface. Mostly, the SEs are emitted by the atoms near the specimen surface. Moreover, the number of SEs is greater than the incident electrons due to the multiple scattering events which subsequently increase the signal level. So with SE mode, the spatial resolution in an image is high. The images formed using the SEs provide surface topography and morphology. Figure 2.15(a) shows the SE image of a fractured aluminium alloy.




(a) (b)

Figure 2.15: (a) SE image and (b) BSE image of a fractured alluminium alloy [Mik04].

### Image acquisition

The image acquisition process starts with the beam generation. Initially, an accelerating voltage is supplied to the electron gun to produce the electrons. These electrons that are redirected by the anode traverse the electron column vertically. The series of electromagnetic lenses and apertures present in the electron column control the beam diameter and focus it on to the sample surface. Finally, to acquire an image, a region on the sample surface has to be scanned by the electron beam in both horizontal and vertical directions with a great speed. This is then performed by the scan coils (one for each direction) by varying the current passing through them as a function of time. As a result of the beam interaction with sample surface, different signals are emitted. The resulted electrons are recorded at their respective detectors and then the gathered information is amplified, digitized and recorded as an image. As the electrons wavelength is much smaller than visible light, no color image can be produced using a SEM. Here

an interesting fact about SEM is, unlike optical systems no lens is directly involved in image acquisition. Figure 2.16 shows the image formation process in SEM.

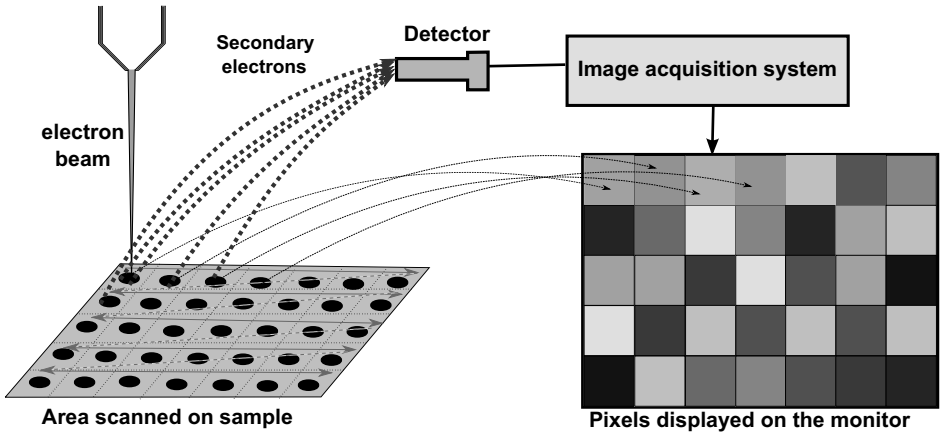



Figure 2.16: Image formation process in a SEM.

In SEM, each pixel is captured individually in a quick succession. The intensity value of a pixel depends mainly on the interaction position of the beam on the sample surface. For a scan length $N$ on the sample surface, the generation of one raster line $R$ is given by equation (2.1).

$$R_n = \sum_{i=1}^{N} I(x_n, y_i) \tag{2.1}$$

where, n = 1 to M lines, $x$ and $y$ are the row and column indices of the interaction point. The total time taken ($t_n$) to produce one raster line is given by equation (2.2).

$$t_n = N t_p n + (n-1) t_{ld} \tag{2.2}$$

where, $t_p$ is the time to produce one pixel given by $t_D + t_d$, $t_D$ is the dwell time[2], $t_d$ is the time delay between the pixels and $t_{ld}$ is the line delay between two raster lines. At $n = M$, $t_n$ becomes the time to produce one single scan image. Besides, the size of a pixel depends on the size of the beam (spot size or probe size) hitting the sample surface. It can be controlled by changing the current passing through the condenser lens. Spot size mostly affects the image resolution. Higher is the spot size, lower is the image resolution. Typically, higher spot sizes are used for low magnification imaging. The magnification in a SEM can be changed by changing the excitation voltage through the scan coils and it is defined as the ratio of the length of the scan on the monitor to the length of the scan on the sample surface.

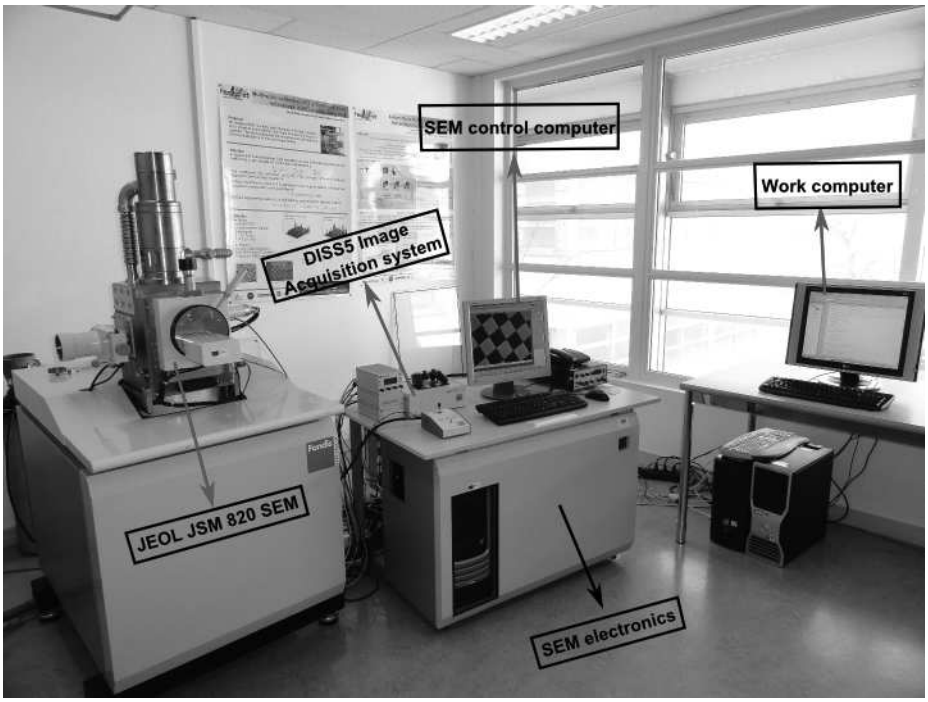



Figure 2.17: Experimental environment showing the hardware set-up.

## 2.3 Experimental hardware set-up

The hardware set-up architecture used for this thesis is shown in the figure 2.17. It consists of a JEOL JSM 820 SEM, an image acquisition system along with a SEM control computer and a work computer. The JEOL JSM 820 is a state-of-the-art tungsten source SEM. Its electron column is equipped with all the elements explained in previous section and its objective aperture strip contains 4 apertures of diameters 100 $\mu m$, 70 $\mu m$, 50 $\mu m$ and 30 $\mu m$ respectively. The accelerating voltage for the SEM varies from 0.3 $kV$ to 30 $kV$ and the achievable resolution at 30 $kV$ is 4.5 $nm$. The magnification of the SEM varies from 10× to 100,000× and the maximum allowable electronic working distance is 50 $mm$. The SEM chamber is equipped with a mobile platform that can be controlled externally using joystick. For this work, a new piezo positioning platform, TRITOR 100 (from piezosystemjena [Pie13]) having 3 degrees of freedom (X, Y and Z) has been mounted on the already available platform (see figure 2.13). It is vacuum compatible and is capable of producing motion up to 100 $\mu m$ on all the three available axes with a resolution of 0.2 $nm$ in open loop. Different axes of this piezo positioning stage are controlled by the NV 40/3 piezo voltage amplifier (from piezosystemjena [Pie13]). It also provides the facility to control the stage axes from remote work computer via serial port interface. Apart from these, a 3 degrees of freedom robot manipulator (Kleindeik MM3A) system is planned to fix inside the SEM vacuum chamber. The end effector

[2]Dwell time is the time taken by the beam to scan one pixel.

of the manipulator will be equipped with a FT G-32 force sensing microgripper (from Femto tools [Fem11]). The maximum opening between the fingers of the gripper is 30 $\mu m$ and their movement can be controlled by varying the actuation voltage. The incoming readings from the force sensors are recorded using a National Instruments data acquisition device (NI 6008 USB). Both the manipulator controller and gripper controllers are interfaced with the work computer.

The SEM control computer is solely responsible for controlling the microscope. It is connected to the image scanning system (DISS5) via USB and SEM electronics. The image scanning system provides an interface that runs from the SEM computer. It enables different functionalities for beam control (acceleration voltage, spot size, focus, magnification etc.) and manual image acquisition with a chosen dwell time. The maximum possible resolution of an image is $16,384 \times 16,384$ pixels and the maximum achievable acquisition speed is 3.1 frames per second using a dwell time of 360 $ns$. A software application to perform automatic image acquisition and control has been designed and is explained in the chapter 6. On the other side, work computer is purely dedicated for performing the control operations using the images acquired from SEM computer.

## 2.4 Conclusion

Now, we have gained sufficient knowledge related to the concepts of nanorobotics, nanomanipulation systems and imaging at nano scale. Since manual control is a difficult and tiresome task at this particular scale, using vision information seems to be a feasible solution. Up on comparing all the available imaging systems, SEM is proved to be more appropriate tool for performing real time manipulation tasks due to its high working range and the capability of providing near real time image feedback with high depth of field, high magnification and high resolution. However, it possess some major limitations and the most important one is the image quality that depends on the scale and speed of imaging and type of the material used. Added to this, time varying distortion (explained in the next chapter) that is usually observed at high magnifications is another critical problem. So, in order to overcome some of these problems associated with SEM to use it for nanomanipulation, we propose some methods and solutions that are detailed in the next chapters.

Chapter 3

# Image analysis and time-varying distortion calibration in SEM

*As a vision sensor, SEM performs a vital role in autonomous micro-nanomanipulation applications. Unlike optical microscopes, a SEM produces images of a sample by raster scanning the sample surface with a focused beam of electrons. When it comes to the sub micrometer range and at high scanning speeds, the images produced by the SEM are noisy and need to be corrected or evaluated before using them. Moreover, SEM image acquisition is mostly affected by the time varying motion of pixel positions in the consecutive images, a phenomenon called drift. In order to perform accurate measurements, it is necessary to compensate this distortion (drift) in advance. In this chapter, we present the methods to evaluate the SEM imaging quality along with the time varying distortion compensation. Various experiments have been performed at different operating conditions and the results are presented.*

## 3.1 Image quality estimation and monitoring

As mentioned in the previous chapter, the images produced by SEM are the result of the electron beam interaction with a sample surface. These images are capable of providing minute details of the scanning surface. However, with fast scanning rates, the images are too noisy and need to be evaluated or corrected before using them. This section provides the information of noise in SEM imaging along with various noise estimation methods and discusses the quality evaluation at different operating conditions.

### 3.1.1 Noise in SEM imaging

In most of the cases, SEM image acquisition is said to be affected by the addition of noise during beam production, its interaction with the specimen surface and also due to the presence of instabilities in the electron column [Rei98]. The primary source of noise present in a SEM image is due to the number of electrons interacted with the specimen surface and the number of emitted electrons recorded at the detector. This type of noise is characterized as shot noise and grow in proportion to the useful signal. It is added due to the fluctuations in the count of emitted primary beam electrons. This type of noise follows Poisson distribution. For SEMs with thermionic guns, this shot noise is more [STP04]. Timischl et al., [TN$^+$12] have described five sources of noise in a SEM: primary emission, secondary emission, scintillator, photocathode, and photomultiplier. They have showed every noise correspond to quantum fluctuations and can be modeled by Poisson statistics. However, in most of the cases the noise added by the detection system is considered to be highly negligible [GNJ$^+$03]. Mulapudi and Joy [SD03] have showed empirically that in a thermionic gun SEM, final image noise follows Gaussian statistics but not Poisson statistics. In order to generate virtual SEM images, Cizmar et al. [CVMP08b] have considered that the final image noise is an addition of a Poisson distribution representing primary emission and a Gaussian distribution representing the others types of noise in the SEM (secondary emission and electronics). Since these previous investigations create ambiguity in selecting the type of noise, in this work, we first study the final image noise. Apart from the above mentioned noises, noise also arises due to the charge-up of specimen surface, alignment variation of electron beam or sample. This type of noise mainly results in gray level fluctuations in the images [Rei98]. Noise is also added due to the external disturbances such as stray magnetic field, mechanical vibrations etc. These types of noises can be seen as image artefacts. Figure 3.1 shows the effect of a short mechanical impulse to the SEM electron column on imaging process.

Practically, noise variation can be determined by acquiring images with varying dwell time. If the dwell time increases, the level of noise decreases as the beam will spend more time at the particular location. However, with the increase in dwell time, the time to capture one frame increases which is not an acceptable constraint for high speed visual tracking applications. Also, if a sample surface is scanned for a long time by electron beam, the surface may get contaminated and unusable [VP05]. The plot shown in figure 3.2 details the relation between acquisition time and dwell time. Figures 3.3(a)

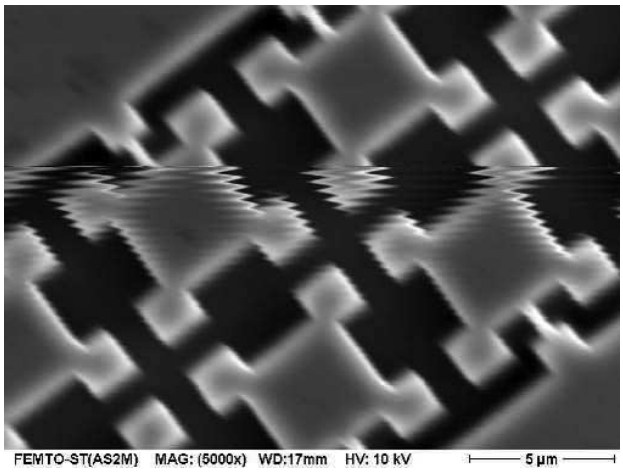



Figure 3.1: Effect of mechanical vibrations on SEM imaging.

and 3.3(b) show two different frames of silicon micro parts captured at 2.9 $\mu s/pixel$ and 92.9 $\mu s/pixel$ dwell times respectively.

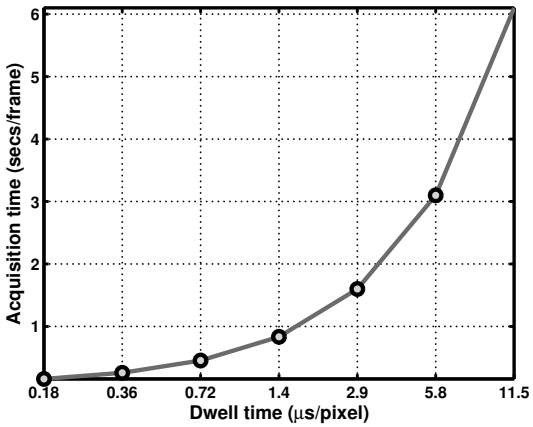
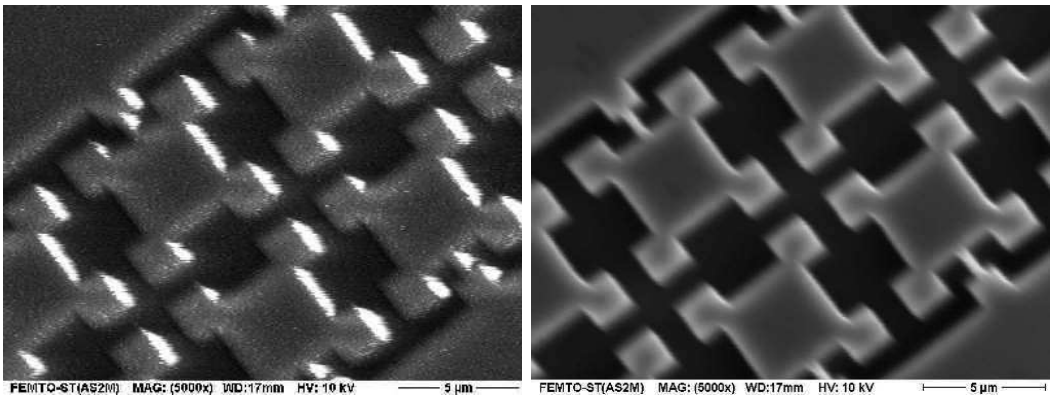



Figure 3.2: Scan time vs. acquisition time with JEOL SEM.

### 3.1.2 Study of final image noise

As mentioned, noise sources are numerous in SEM imaging and each contribute independently towards the final image formation. At the first hand, there are random statistical fluctuations in the primary and secondary electron emissions which produce shot noises independent from one pixel to another and depends on the material being used for imaging. Generally this type of noise satisfies the Poisson statistics:

$$f(\lambda) = \frac{\mu^{\lambda}}{\lambda!} e^{-\mu} \tag{3.1}$$




(a) (b)

Figure 3.3: Images of silicon micro parts acquired at (a) 2.9 $\mu s/pixel$ and (b) 92.9 $\mu s/pixel$.

where, $\lambda$ the number of occurrences of the events and $\mu$ is the mean signal level. However, if the mean number of events is large enough ($\mu$ >10), the Gaussian distribution is a good approximation for the Poisson distribution. On the other hand, we have also noise added by the detection and recording system. Apart from this, noise is also added by the photomultipliers where they emit some signal even in the absence of original signal. This noise is due to the thermal fluctuations in the photocathode. Upon considering all these noises, the final image noise $n(x, y)$ at the pixel $(x, y)$ is a result of superposition of all these partial noises. Under the central limit theorem this final noise can be approximated with Gaussian statistics [Bov10].

In order to find out the distribution followed by the final image noise, we have conducted the initial tests by imaging the microscale calibration rig (gold on silicon) shown in figure 4.9 at two different scan times (180 and 360 ns/pixel). It is performed by computing the noise distribution with in a single image region. Two plain image regions are selected for each image where, one region contains only gold and the other contains only silicon. Figures 3.4(a) to 3.4(d) show the intensity histograms along with the fitted distributions (Gaussian and Poisson) for gold and silicon regions at different scan times.

After analyzing the obtained results, it is clear that any of the two distribution curves (Poisson and Gaussian) do not fit correctly with the real pixel intensity distributions. However, Gaussian distribution seems to provide better fit than Poisson distribution which can be clearly seen with the intensity histograms especially with silicon regions.

### 3.1.3 Noise estimation and SNR quantification methods

To assess the quality of acquired frames, image signal-to-noise ratio (SNR) has been selected as an evaluation parameter in this work. SNR is a commonly used term in the field of signal processing to differentiate or estimate the level of noise present in a

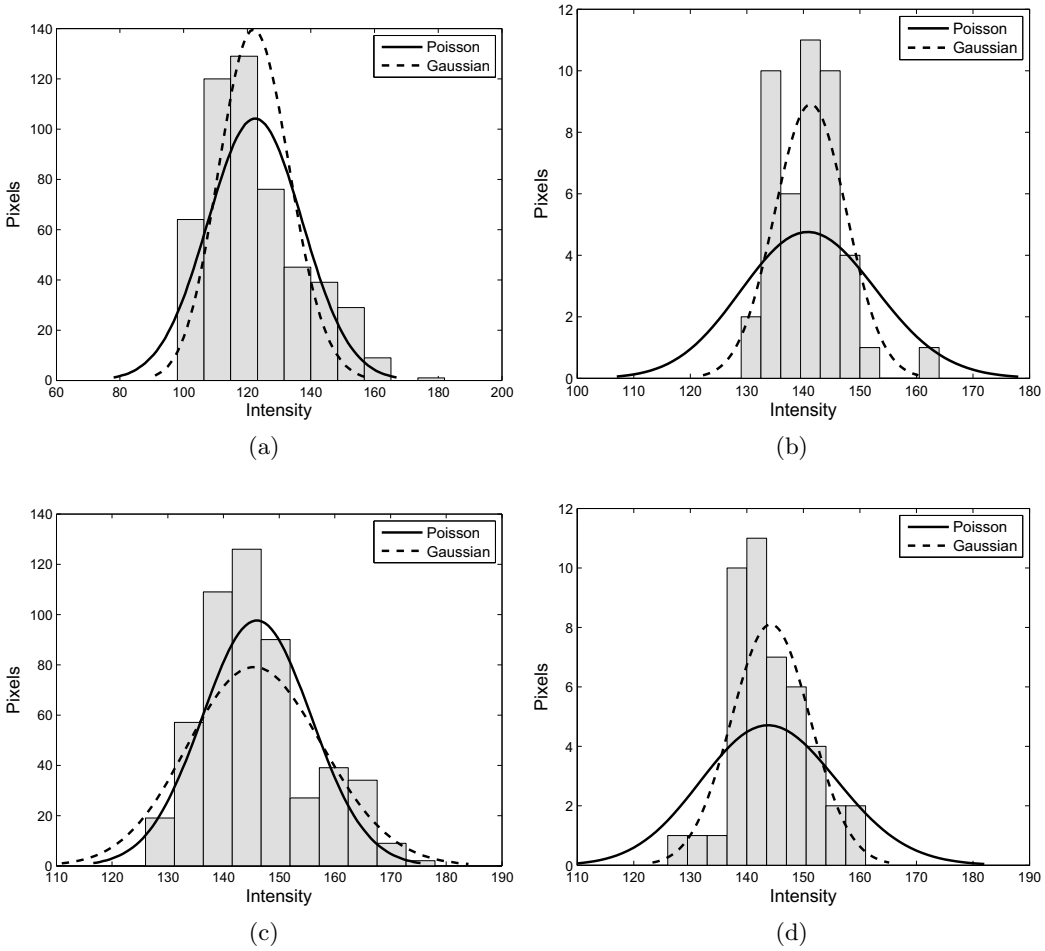



Figure 3.4: Intensity histograms with approximated distributions for (a) and (c) gold regions (b) and (d) silicon regions for the images acquired at 180 and 360 ns/pixel.

recorded signal [BMC04]. In general terms, it is defined as the ratio of desired signal over background noise. Following the industry standards, SNR can be computed as in equation (3.2).

$$SNR \triangleq 10log_{10}\left\{\frac{variance(Signal)}{variance(noise)}\right\} \tag{3.2}$$

And if a signal S is known to be bounded, $S_{min} \leq S \leq S_{max}$, then the SNR can be computed using equation (3.3).

$$SNR \triangleq 20log_{10}\left\{\frac{S_{max} - S_{min}}{std(noise)}\right\} \tag{3.3}$$

In general, for high quality images this ratio is often very high ($> 25\ dB$). Operationally, high SNR values can be achieved by increasing probe current or by decreasing

scanning frequency. In order to compute the SNR using above formulae, it is mainly required to estimate the level of noise added during the final image formation. The other possibility can be by using the auto-correlation technique.

So far in the literature, many research works have proposed various methods using two scan images for the computation of SNR using auto-correlation technique[1] [Era82, LFSK06, SLLT09, SNT11]. These methods require two images that are needed to be perfectly aligned and are obtained by scanning the same specimen area. In addition, they require long processing time and are complex to integrate and use with real time systems. Sim et al. [SCC04] have proposed a mixed Lagrange time delay estimation autoregressive (MLTDEAR) model for estimating the SNR using a single image. Apart from the correlation-based methods, noise in the images can be reduced by performing the frame averaging or by applying filters. Tai et al. [TY08] have proposed a method to estimate the noise in digital images using a first-order gradient-based edge detection algorithm and a Laplacian operator. A method to estimate and minimize the noise by adding external hardware circuits or electronics to the detectors inside SEM is proposed in [JJK11] and the methods for image denoising using the total variation algorithms are explained in [VO96, OSR12]. A thorough evaluation and comparison of various noise estimation methods has been provided by Olsen [Ols93]. In his work, Olsen has used six different techniques that are later categorized into two methods: filter-based methods and block-based methods for estimating the standard deviation of the image noise. He concluded that the filter-based methods, especially the one using median filter provide better results than the block-based methods. Many other researchers have come up with different approaches for noise variance estimation; and most of them are based on the filter-based technique [Imm96, RLU99, PTWZ10]. Different noise reduction techniques along with the SNR computation methods that are most widely used are explained below.

### Frame averaging

Frame averaging is one of the most simple and commonly used noise reduction techniques with digital imaging. Most of the modern days SEMs comes with a hardware unit to perform frame averaging before an image is displayed on a monitor. It produces better quality images by averaging $N$ number of frames from which the noise is averaged among the frames and the final image $f(x, y)$ displayed is an outcome of this process. Theoretically, the process can be described as in equation (3.4).

$$f(x, y) = \frac{1}{N} \sum_{K} f_i(x, y) \tag{3.4}$$

However, the main disadvantage associated with this technique is that the final image is produced after averaging previously acquired frames henceforth increasing the overall image acquisition time.

[1] Auto-correlation technique was first proposed by Frank and Ali [FAA75].

### Image filtering

Apart from averaging the acquired frames, noise can also be reduced by convoluting an image with a spatial filter mask. These methods are software dependent and the overall processing time depends on the size of the applied filter mask. The idea behind these methods is straight forward i.e. good quality images are produced by substituting the intensity value of each pixel by a value determined by the filter mask. These methods are further classified into linear filtering and nonlinear filtering based on the behavior of the applied mask. Linear filters perform noise reduction by averaging the pixels present in the filter neighbourhood. The output response $R$ of a filter $w(k,l)$ of size $P \times Q$ on the image $f(x,y)$ is given by equation (3.5).

$$R = \sum_{k=1}^{P} \sum_{l=1}^{Q} w(k,l) f(x-k,\, y-l) \tag{3.5}$$

However, they also eliminate important image feature details like edges that are vital for object tracking. On the other hand, nonlinear kernels provide noise reduction with considerably less smoothing than linear kernels and preserve important image details. The best known nonlinear filter is the median filter for its ability in filtering the random noise and preserving the image details [GW08, CCR08, PTWZ10].

The median filter was initially introduced in signal processing by Tukey [Tuk74]. Statistically, a median for a given vector of values is the middle value after sorting. In image processing, median filtering is performed by selecting a neighborhood size for the mask and moving this mask through all the pixels present in the image. The filtering is provided within a local kernel neighborhood by sorting the pixels according to their intensity values and by replacing the corresponding pixel with a median value. Figure 3.5 shows the concept of median filtering in image processing.

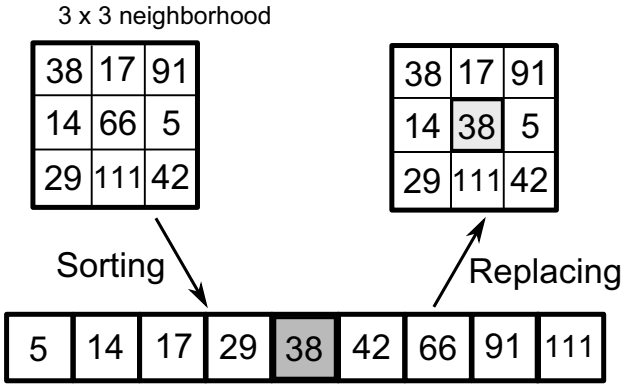



Figure 3.5: Concept of median filtering in image processing.

**SNR estimation using auto-correlation function (ACF)**

As mentioned earlier, one of the traditional ways to estimate the SNR is by using cross-correlation of two microscopic images acquired by scanning the same specimen area. From this, the correlation coefficient $\rho_{12}$ can be computed using equation (3.6).

$$\rho_{12} = \frac{\phi_{12}(0) - \mu_1\mu 2}{\sigma_1\sigma_2} \tag{3.6}$$

where, $\phi_{12}(0)$ is the correlation peak value of the two perfectly aligned images and $\mu_1$, $\mu 2$, $\sigma_1$, $\sigma_2$ are the mean and variance values of the two images, respectively. Then the SNR is formulated as in equation (3.7).

$$SNR = \frac{\rho_{12}}{1 - \rho_{12}} \tag{3.7}$$

Considering the main drawback of acquiring two perfectly aligned images, Thong et al. [TSP01] proposed a method to compute the SNR using single image assuming the noise in the image is additive white noise. The noisy image of gold on carbon sample shown in figure 3.6(b) formed by adding white noise of 20 $dB$ to figure 3.6(a) is used for the demonstration. The ACF values of the noisy image are plotted on a three-dimensional grid and are shown in figure 3.7(a). Now the objective is to estimate the noise-free peak in order to differentiate the noise from noise free image. This can be performed by using polynomial approximation [Smy97]. Other available estimation techniques are piecewise cubic interpolation [KMNF89], spline interpolation [SM98] and autoregressive moving average (ARMA) model [BD91]. Estimated noise and noise-free peaks are shown in figure 3.7(b). Once the noise-free peak is estimated the SNR can be computed using equation (3.8).

$$SNR = \frac{\text{noise free peak} - mean(pixels)^2}{\text{noise peak} - \text{noise free peak}} \tag{3.8}$$

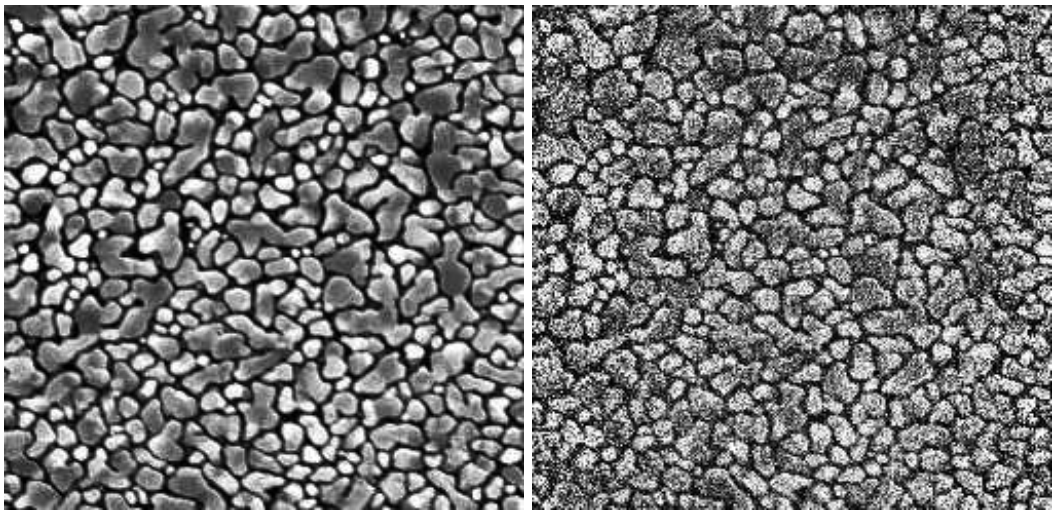
Although this method provides good results, the accuracy of the method is highly dependent on the noise free peak estimation and the nature of the images. Apart from that, the overall processing time is more which makes it difficult to use it with real time applications.

In order to overcome the drawbacks associated with auto-correlation methods, we develop a simple and robust filter-based method using median filter to estimate the amount of noise and to compute the SNR using a single image. The developed method is explained in the next section.

## 3.2 Proposed method of SNR estimation

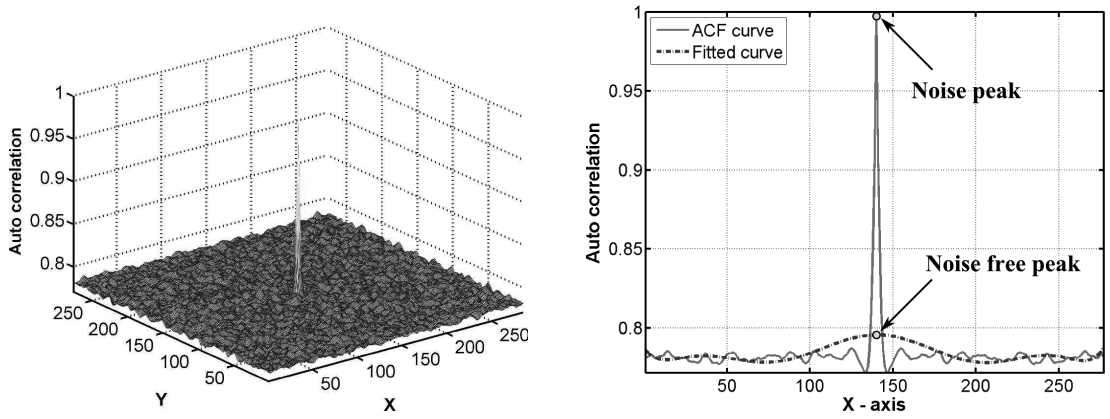
### 3.2.1 Presentation of the method

From the previous study, we assume that the acquired image $g(x, y)$ of silicon microparts sample shown in figure 3.3(a) is formed by the addition of spatially uncorrelated additive



(a) (b)

Figure 3.6: (a) Assumed noise free image of gold on carbon sample (b) Image formed after adding Gaussian noise of $20dB$ to (a).




(a) (b)

Figure 3.7: (a) ACF curve obtained for the image shown in figure 3.6(b) (b) noise peak and the estimated noise-free peak (for demonstration single side of ACF curve is shown).

noise $n(x, y)$ to the original image $f(x, y)$ and is given by equation (3.9).

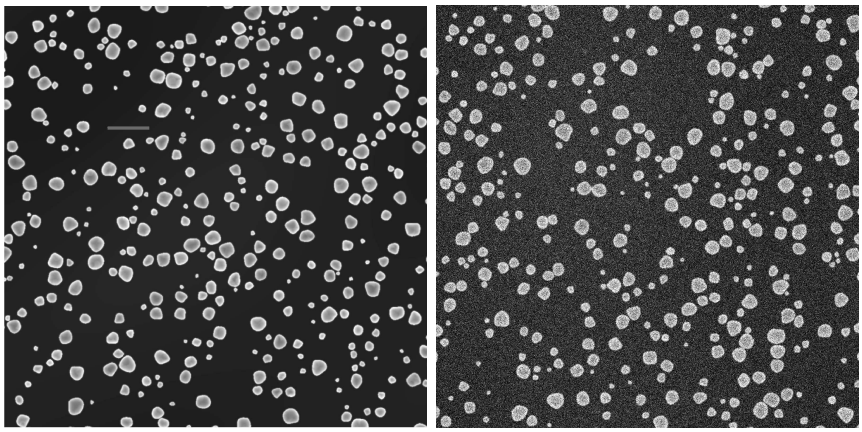
$$g(x, y) = f(x, y) + n(x, y) \tag{3.9}$$

The primary step is to adjust the contrast of the image which can be performed by using histogram equalization. This is optional as most of the present day microscopes come with an integrated module to perform this task during image acquisition. Next, the median filter mask of appropriate size is selected based on the total amount of

processing time and the amount of blurring. The amount of blurring has been computed by calculating the mean squared error (MSE) between the original image and the filtered image $\hat{f}(x, y)$ using equation (3.10).

$$MSE = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} (f(x, y) - \hat{f}(x, y))^2 \tag{3.10}$$

where, M and N are the image dimensions. The artificial noise free and noisy images (Figures 5(A) and 5(B)) of gold on carbon generated by Artimagen library [CVMP08b, CVMP08a] are considered for this purpose. It is a library provided by the National Institute of Standards and Technology (NIST) to produce artificial SEM images with varying noise. The additive Gaussian noise with a standard deviation equal to 0.07 has been added to generate a noisy image. The MSE between original image and noisy image is found to be 110.30. Table 3.1 summarizes the computed MSE values using filters of different size. Figures 3.9(a)-(d) illustrates the blurring details. The plots shown in the figure are the intensity values of the original image and filtered image along the horizontal line (colored in red) shown in the figure 3.8(a). The plots clearly show the effect of blurring.



(a) (b)

Figure 3.8: Artificially generated (A) noise free and (B) noisy images of gold on carbon using Artimagen library. The horizontal line (red) in (A) is the used image profile to analyze the filter response.

In order to accelerate the overall process, a fast sorting algorithm is used in developing the median filter. Table 3.2 shows the total processing time for an image of size $512 \times 512$ pixels using different kernel sizes (implemented in C++).

Now, the goal is to filter $g(x, y)$ by applying the median filter to find a closest estimate $\hat{f}(x, y)$ of the original signal image $f(x, y)$. Thus

$$\hat{f}(x, y) = \text{median}\,\{g(s, t)\} \tag{3.11}$$

Table 3.1: Computed MSE values for various filter sizes.

| Filter Size | MSE |
|---|---|
| $3 \times 3$ | 23.4 |
| $5 \times 5$ | 22.6 |
| $7 \times 7$ | 41.1 |
| $9 \times 9$ | 74.8 |

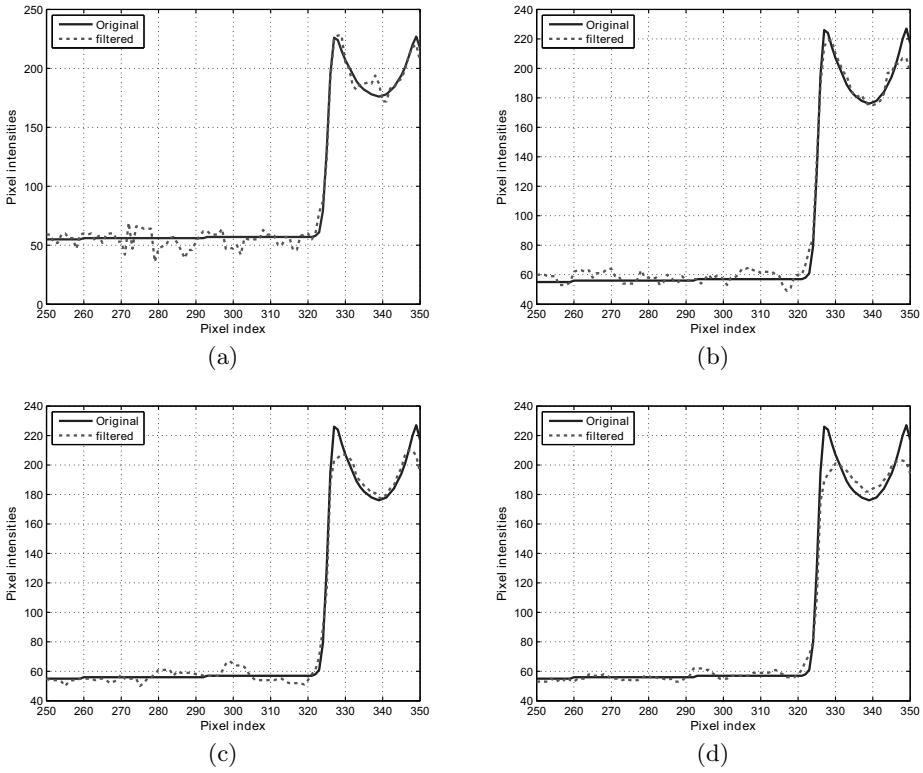



Figure 3.9: Plots of intensities along the red line shown in figure 3.8(a) after filtering with (a) $3 \times 3$ (b) $5 \times 5$ (c) $7 \times 7$ (d) $9 \times 9$.

where $(s, t)$ are the local filter neighborhood pixel coordinates and $(s, t) \in g(x, y)$. For this work a filter of size $5 \times 5$ pixels has been used. Figure 3.10(a) shows the filtered image $\hat{f}(x, y)$ of figure 3.3(a). In-turn this filtered image is subtracted from the original image to produce a noise image $n(x, y)$ given by (3.12) as shown in figure 3.10(b).

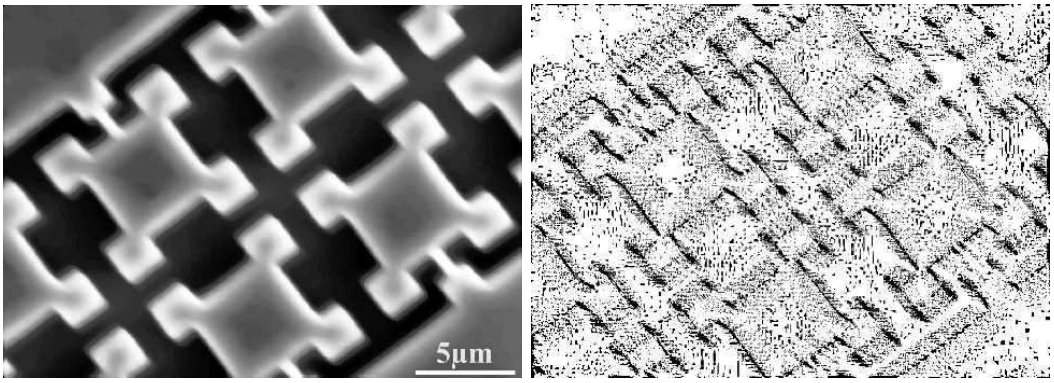
$$n(x, y) = g(x, y) - \hat{f}(x, y) \tag{3.12}$$

Using the estimated noise free image $\hat{f}(x, y)$ and noise image $n(x, y)$ the SNR is computed following the industrial standards as in (3.13).

$$SNR \triangleq 20 \log_{10} \left\{ \frac{std(\hat{f}(x, y))}{std(n(x, y))} \right\} \tag{3.13}$$

Table 3.2: Total processing time for different filter sizes.

| Filter Size | Processing time (ms) |
|---|---|
| $3 \times 3$ | 15.6 |
| $5 \times 5$ | 17.2 |
| $7 \times 7$ | 30.4 |
| $9 \times 9$ | 31.3 |
| $11 \times 11$ | 34.5 |
| $13 \times 13$ | 39.5 |
| $15 \times 15$ | 46.9 |
| $17 \times 17$ | 48.3 |



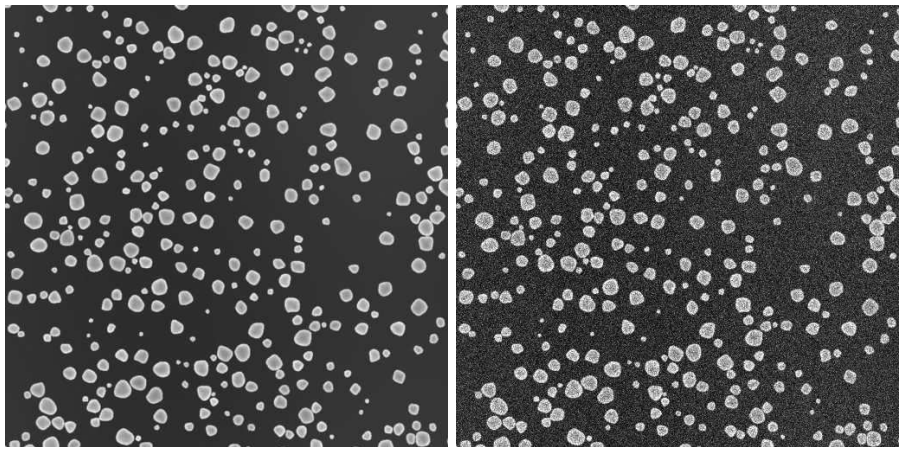

(a) (b)

Figure 3.10: (a) Filtered image of figure 3.3(a) using median filter (b) Noise image formed after subtracting (a) from figure 3.3(a).

### 3.2.2 Precision testing of the developed approach

The precision of the developed method is tested by using a noise free image and by corrupting it with a known amount of noise. Later the SNR values of the corrupted images are computed using the proposed method and are compared with the known values. The noise free image has been generated using Artimagen library.

Figures 3.11(a) and 3.11(b) show the noise free and corrupted images with a white Gaussian noise of 20 $dB$. Table 3.3 summarizes the computed values along with the percent error and MSE computed between the filtered image and the original image. From Tables 3.1 and 3.2, it can be seen that filter size of $5 \times 5$ establishes a good compromise between accuracy and time of SNR computation with the current setup. With another experimental setup, the choice may vary. Upon testing the precision of the developed method in estimating the original SNR, the mean variation of the error is comparatively less for $5 \times 5$ filter (Table 3.3) and the percent error is less than 10 in most of the cases (with a filter size of $5 \times 5$). However, with high amount of noise in the image (especially for SNR $<$ 12 dB) the percent error is more than 10. This is mainly



(a) (b)

Figure 3.11: The artificial noise-free and noisy images of gold on carbon sample produced by Artimagen. Noise image is formed after adding a white Gaussian noise of 20 $dB$.

Table 3.3: Known and computed SNR values using the developed method for artificial gold on carbon image with different filter sizes.

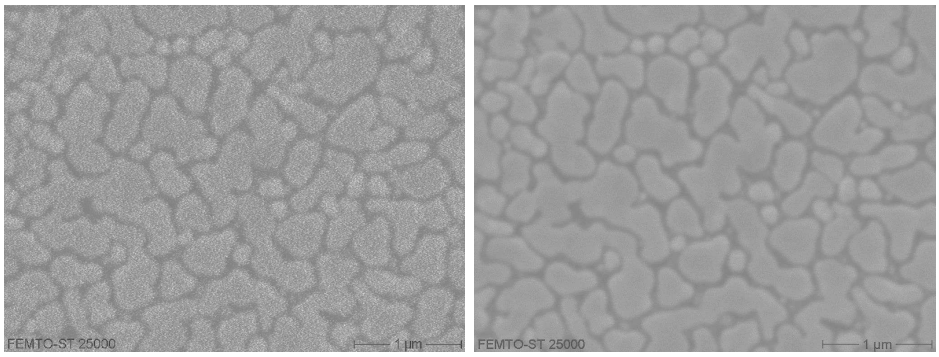
| Known (dB) | $3 \times 3$ | | $5 \times 5$ | | $7 \times 7$ | |
|---|---|---|---|---|---|---|
| | comp | err % | comp | err % | comp | err % |
| 10 | 15.18 | 51.82 | 12.35 | 23.53 | 10.82 | 8.23 |
| 11 | 15.45 | 40.46 | 13.05 | 18.63 | 11.43 | 3.96 |
| 12 | 15.75 | 31.27 | 12.81 | 6.82 | 12.27 | 2.27 |
| 13 | 16.80 | 29.24 | 13.75 | 5.82 | 12.74 | 1.92 |
| 14 | 16.51 | 17.95 | 14.55 | 3.99 | 13.88 | 0.80 |
| 15 | 17.38 | 15.90 | 15.33 | 2.23 | 14.52 | 3.15 |
| 16 | 17.76 | 11.02 | 16.16 | 1.02 | 15.43 | 3.54 |
| 17 | 18.17 | 6.93 | 16.89 | 0.63 | 16.23 | 4.48 |
| 18 | 19.12 | 6.27 | 18.01 | 0.06 | 17.28 | 3.97 |
| 19 | 19.87 | 4.62 | 19.07 | 0.37 | 18.12 | 4.60 |
| 20 | 20.88 | 4.40 | 19.65 | 1.74 | 19.06 | 4.95 |
| 21 | 21.76 | 3.63 | 20.70 | 1.41 | 20.00 | 4.74 |
| 22 | 22.53 | 2.41 | 21.41 | 2.66 | 20.96 | 4.71 |
| 23 | 23.45 | 1.97 | 22.42 | 2.48 | 21.73 | 5.49 |
| 24 | 24.58 | 2.45 | 23.30 | 2.89 | 22.59 | 5.86 |
| 25 | 25.33 | 1.34 | 24.26 | 2.93 | 23.52 | 5.89 |
| 26 | 26.24 | 0.93 | 25.13 | 3.31 | 24.48 | 5.84 |
| 27 | 27.21 | 0.81 | 26.05 | 3.48 | 25.16 | 6.78 |
| 28 | 28.06 | 0.23 | 27.03 | 3.44 | 25.97 | 7.22 |
| 29 | 28.99 | 0.01 | 27.87 | 3.86 | 26.72 | 7.85 |
| 30 | 29.95 | 0.14 | 28.74 | 3.89 | 27.47 | 8.42 |

due to the inability of median filtering in the presence of high amount of noise. In fact, a median filter is able to remove noise in the images only if the noise pixels are less than one half of its neighborhood. This is the reason why a 3x3 filter provides poor error and a $7 \times 7$ filter provides better error than a $5 \times 5$ filter (for some conditions). However, from the performed experiments with our system, it has been found that the SNR level is more than 14 dB in all cases.

## 3.3 Image quality evaluation in SEM at different operating conditions

This section presents the evaluation results of the SEM image quality variation at different operating conditions. For most of the experiments conducted, the images are acquired from JEOL JSM 820 SEM and for all tests, the method explained above has been used for SNR estimation.

### 3.3.1 Quality evaluation with respect to the dwell time




(a) (b)

Figure 3.12: Sample images of gold on carbon obtained at (a) 720 $ns/pixel$ (b) 11500 $ns/pixel$.

Initial tests are performed to evaluate the image quality with a change in the dwell time. It is a well known fact that the SNR increases with increase in dwell time. This experiment has been performed in order to test and support this statement. Nine different SE images of the gold on carbon sample (at $25,000\times$ magnification) and gold on silicon (at $1,000\times$ magnification) are acquired with different dwell times . Figures 3.12(a) and 3.12(b) show two sample images of gold on carbon acquired at different dwell times. The acceleration voltage, aperture size and the working distance used for this test are 10 $kV$, 50 $\mu m$ and 9.3 $mm$ respectively. Figure 3.13 shows the SNR variation with respect to the dwell time.

From the obtained results, it is clear that the image quality increases with increase in the dwell time. This phenomenon can be explained as that the increase in dwell time

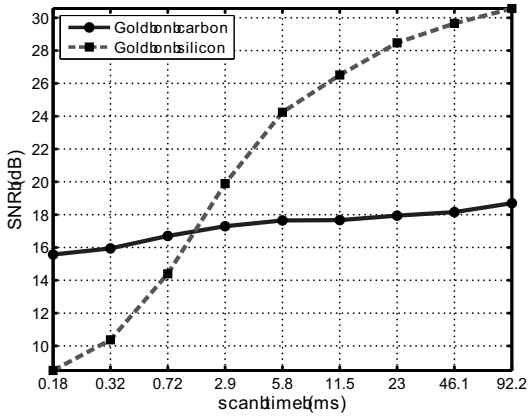



Figure 3.13: SNR evolution with respect to dwell time.

reduces the overall raster rate and increases the amount of time the beam spends at a particular location. Because of this, the electron count recorded at the detector increases subsequently improving the image quality.

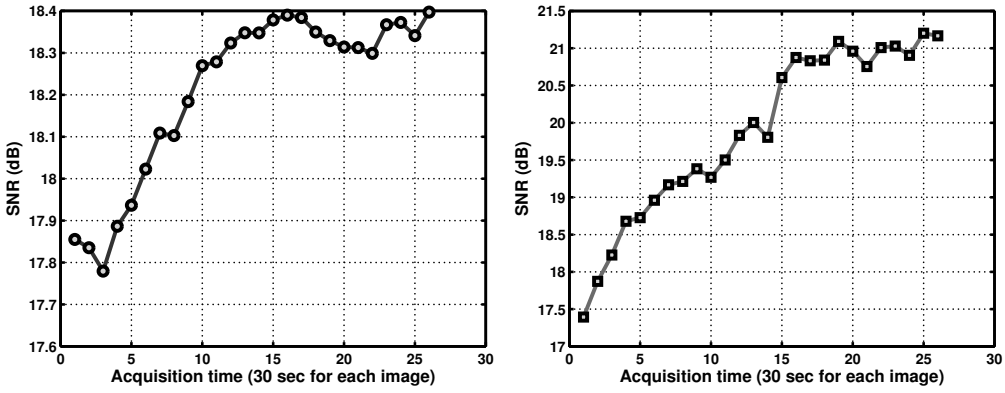
### 3.3.2 Quality evaluation of two different SEMs

Various experiments have been performed to evaluate the image quality of two different SEMs at varying times. Along with the images from our JEOL SEM (tungsten filament), the images acquired from a Carl Zeiss SEM (FEG) have been used for this test. The specimen used for imaging is a standard resolution gold on carbon sample. Selected image size is $512 \times 512$ pixels. These images are acquired using an accelerating voltage of 10 $kV$ for the magnification ranging from $10k\times$ to $100k\times$ with a step of $5k$. For each magnification, 20 images are acquired with a sampling time of 30 seconds for each image. The image quality is then evaluated off-line with respect to the time. The plots shown in figures 3.14(a) and 3.14(b) depict the comparison of SNR evolution of the two instruments with respect to the time at $40,000\times$ magnification and figures 3.15(a) and 3.15(b) show the comparison results at $70,000\times$ magnification.

The obtained results show that the level of noise reduces with increase in time thus, increasing the overall quality of images. This phenomenon is a result of increase in the filament temperature and the count of electrons produced by the electron gun over time. Moreover, it is clear that the level of SNR with FEG SEM is higher than the tungsten filament gun SEM.

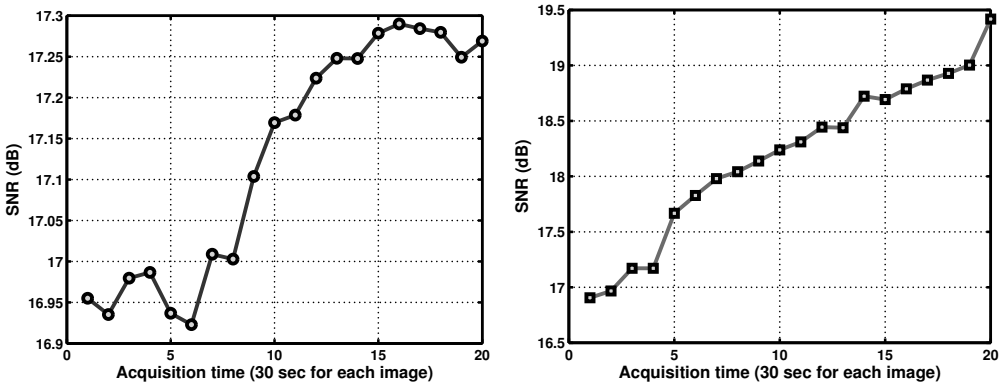
### 3.3.3 Quality evaluation with respect to magnification

This test has been performed to monitor the SNR evolution with a change in the magnification. For this test, two samples are used: gold on carbon and gold on silicon calibration rig. Using gold on carbon sample, 20 images are acquired, one for each magnification ranging from $10k\times$ to $30k\times$ (step = 1000) and with calibration pattern, the images are acquired for the magnification ranging from $250\times$ to $1000\times$ (step = 10). The




(a) (b)

Figure 3.14: SNR vs. acquisition time at $40,000\times$ magnification for (a) tungsten gun SEM (b) FEG SEM.

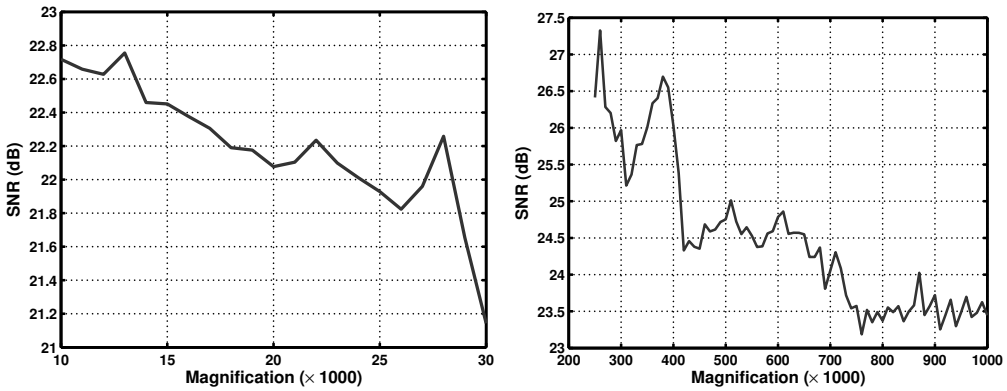



(a) (b)

Figure 3.15: SNR vs. acquisition time at $70,000\times$ magnification for (a) tungsten gun SEM (b) FEG SEM.

used dwell time is 360 *ns/pixel*. Figures 3.16(a) and 3.16(b) show the evolution of the SNR with respect to magnification.

From the obtained results, the evolution of the SNR with respect to magnification is a nonuniform decreasing function for both samples. With an increase in the magnification, since the size of the monitor (image) is fixed, the size of the scan area is narrowed and the surface corresponding to scan a pixel reduces. Therefore, the area of the gold region that actively participates in generating the dominant quantity of original signal reduces due to which the SNR reduces. This phenomenon mostly depends on the material of the sample used for imaging.




(a) (b)

Figure 3.16: Evolution of SNR with respect to magnification for (a) gold on carbon (b) gold on silicon.

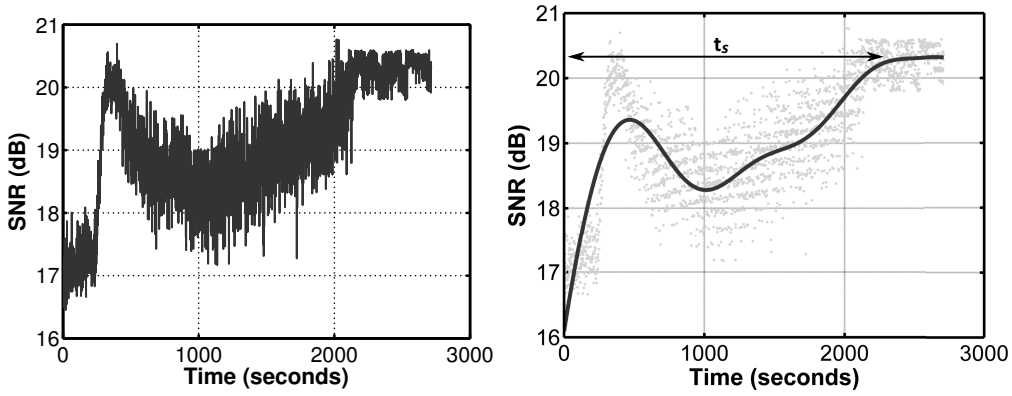
### 3.3.4 Quality evaluation in real time

In this test, the SNR quantification and image quality estimation is performed directly during the time of image acquisition. This mode mainly allows observing the variation in noise and thus image quality with increase in the instrument operational time. For this test, the sample used is a silicon wafer containing the micro-parts of dimensions 6 $\mu m$ × 6 $\mu m$ × 2.5 $\mu m$ (see figure 3.3(b)) and the image size is 512 × 512 pixels.

Initially, an acceleration voltage of 10kV and a magnification suitable for the application (in our case 5000×) are selected using the SEM software interface running in work PC. Both server and image acquisition process are started immediately after supplying the accelerating voltage to the filament and the secondary electron detector is turned on. The total operational time selected is $\sim$ 55 minutes. As the overall operating time is more, the specimen surface is connected to the mounting plate using a copper conductive tape to remove the accumulated charge on the sample surface. Total time taken to acquire, transfer and process a single image is about 400 milliseconds, thanks to fast image acquisition and quick sorting algorithms.

The obtained SNR values are plotted and shown in the figure 3.17(a). The initial output raw data contains multiple peaks that are fitted using polynomial fitting to see the actual response. Fitted data is shown in the figure 3.17(b). After a keen observation of the data, initially the SNR values started to increase for a specific time and then decreases. From then it increases and becomes stable at a point of time. It is the time taken by the filament to reach a specific temperature to produce the stable number of electrons. This time is called as settling time or waiting time, $t_s$ (see figure 3.17(b)).

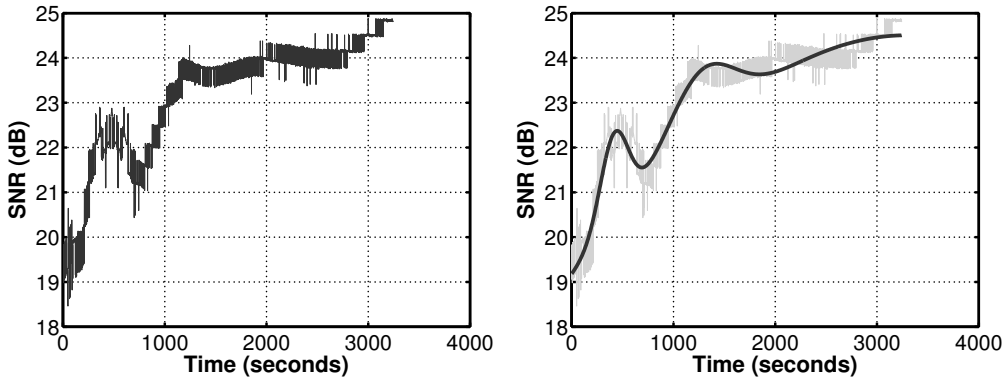
The same experiment is conducted on a different day to observe the consistency of SNR variation. Figures 3.18(a) and 3.18(b) show the raw and fitted plots of the data collected on day 2. From the plots it is clear that the SNR variation is pretty much




(a) (b)

Figure 3.17: (a) Raw data and (b) fitted curve showing the settling time and stable position.

identical to the previously obtained results from day 1 and the only difference is in the length of the peaks.




(a) (b)

Figure 3.18: (a) Raw data and (b) fitted curves obtained on day 2.

### 3.3.5 Quality evaluation with a change in the focus

The evolution of the image quality has been monitored with a change in the focus in this test. More details regarding the focusing process in SEM are explained in the chapter 4. The sample used for this test is a microscale calibration pattern containing chessboard patterns (see figure 4.9) designed and fabricated at FEMTO-ST. The tests are conducted using an aperture size of 50 $\mu m$, 10 $kV$ accelerating voltage at 1000× magnification. For measuring the sharpness, image variance has been used as the sharpness function given

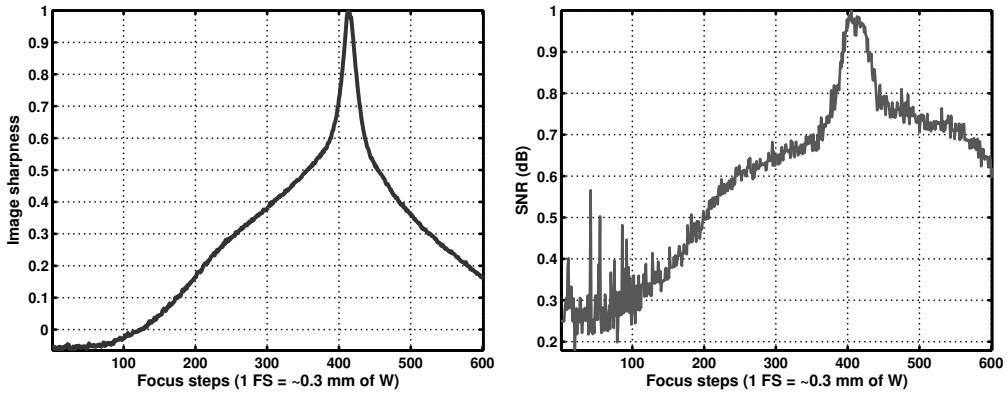
by equation (3.14).

$$\mathbf{S}_{NV} = \frac{1}{MN}\frac{1}{\mu}\sum_{M}\sum_{N}(I(u,v)-\mu)^2 \tag{3.14}$$

where, $M$ and $N$ are the width and height of the image respectively, $\mu$ is the image mean intensity and is given by:

$$\mu = \frac{1}{MN}\sum_{M}\sum_{N} I(u,v) \tag{3.15}$$

Figures 3.19(a) and 3.19(b) show the variation of the sharpness with a change in the electronic working distance. The obtained results show that the SNR level varies simultaneously with image sharpness. Here we can make two conclusions. The primary one is, with a change in the focus step (working distance), the size of beam interacting with the sample surface varies; thus, varying the beam current. At the point of *in-focus*, the current density is high resulting in more signal from the sample improving the SNR. The second one is, with a change in the focus, the level of original signal details present in the image varies. At the point out-of-focus, noise level is predominantly more than that of the signal level, for e.g. at the initial steps of the sharpness curve, the images are completely dark containing no image details other than noise.




(a) (b)

Figure 3.19: (a) Image sharpness (b) SNR variation with respect to the focus steps.

## 3.4 Drift calibration

### 3.4.1 Image Distortions in Scanning Electron Microscopy

As mentioned earlier, unlike optical systems, SEM images are acquired by recording one pixel at a time. Despite of the imaging technique used, SEM images also suffer from different types of aberrations and distortions. The aberrations that are mostly observed

with SEM imaging are: diffraction effects caused by small aperture diameters, spherical aberration due to the geometrical construction of electromagnetic lenses, chromatic aberration due to the electron energy differences, and astigmatism. Apart from the aberrations, SEM image acquisition is known to be affected by two types of distortions, time-varying (drift) and spatially-varying distortions. Bearing these points of distortion in mind, the complete imaging model in a SEM can be functionalized as in equation (3.16).

$$m = \delta_d\{\delta_s(PM)\} \tag{3.16}$$

where, $m$ is the image pixel corresponding to the 3-D scene point $M$, $\delta_d$ is the time-varying distortion, $\delta_s$ is the spatially-varying distortion and $P$ is the projection matrix at a given magnification. According to Cornille [Cor05] and Kratochvil [Kra08], for low magnification, a perspective model of projection can be used and for magnification more than 1000× the model of projection switches to parallel. Hence, in order to recover a 3D point of the scanned scene, the corresponding pixel point has to be corrected initially for drift followed by the spatial distortion. The spatial distortion is constant in time and can be calibrated using many methods [Tsa87, WCH92, Bou99, Zha00]. Kratochvil has mentioned in his work that the spatial distortion is relatively very small across the image (less than 0.5 pixels) and can be easily neglected. However, the drift distortion on the other hand has got much importance and need to be compensated beforehand.

Generally, drift in the SEM imaging can be characterized as the evolution of pixel positions from time to time. It is usually observed with consecutive scans even though there is no change in the device parameters. Most significantly, it affects the images at high magnification. Potential causes of this drift can be mechanical instability of the column or the sample support, thermal expansion and contraction of the microscope components, accumulation of the charges in the column, mechanical disturbances etc. Figure 3.20(a)-(d) shows four images of a gold on carbon sample acquired at different times at 25,000× magnification, without changing any system parameters. By observing the bright spot (pointed inside the circle) in all the images, it is clear that the spot location has been changed from image to image. This is the physical idea of the drift in $x$ and $y$ planes of the sample surface. Even though there exists the drift in $z$ direction, it has a little impact on the image focus.

In the literature, several methods have been proposed to estimate and compensate the drift. According to Cornille [Cor05] and Sutton et al. [SLG+07, SLJ+07], the pixel drift is estimated by the disparity between pairs of points and is fitted by B-splines with respect to time. This model is used to estimate the drift and to remove it in real time. Mokaberi et al. [MR06] has introduced a Kalman filter-based estimator for compensating the drift in AFMs. Many works have used a frequency domain phase correlation method for estimating and compensating the translation produced by the drift in the images [YJB05, YJB08, CVP10, Sne10, MDP+12a]. In his thesis, Cornille has proposed that the drift between pixels or between lines of an image is negligible; instead, the drift between the two images can be considered as a whole.

In general, two approaches are possible for drift compensation. The first one is based on estimating the drift directly from the current image by comparing it with a reference

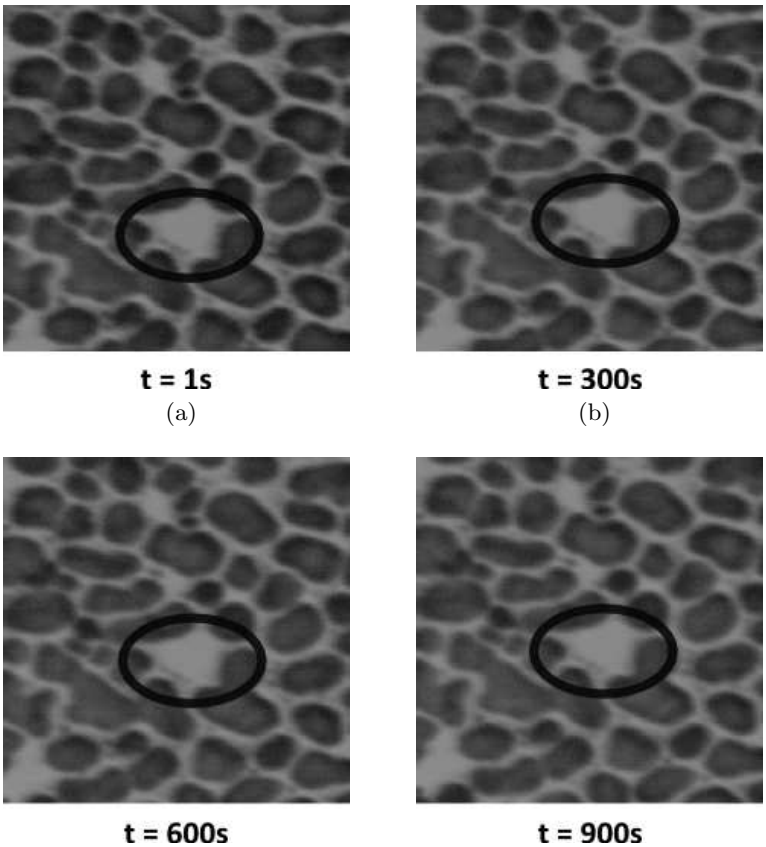



Figure 3.20: Images of gold on carbon sample acquired at different times.

image and correcting the current image in real time [CVP10, MDP+12a]. The second approach is based on developing an empirical model of the drift that is then used to estimate and compensate the drift in real time [Cor05]. This approach is limited by the difficulty in computing a generic model. Due to these limitations, we considered to use the first approach where, the problem of drift compensation has been treated as a registration problem between the images. As majority of the existing drift compensation techniques have been developed using image correlation technique, the achieved results are also compared with the phase correlation method.

### 3.4.2 Image registration and homography

In image processing, image registration is the determination of geometrical transformation to align two different images of a same view. There are many methods available in order to perform image registration. A complete survey about all the available methods is provided by Brown [Bro92], Zitova et al. [ZF03] and Wyawahare et al. [WPA09].

According to them, the most commonly used image registration technique is by using image point features. One such method has been implemented in this work.

Normally for image registration, the images can be acquired by a single sensor at multiple times or by multiple sensors viewing the same scene. For the problem of drift in SEM, as the pixel points evolve with time, it can be viewed as a same scene and the correction can be performed on the current scan by computing the transformation with a reference scan. This transformation can be computed by selecting some points or features that are same in both the images. Each image that has been used for registration will be assigned to a coordinate system that defines the space for that image.

Let $I_0$ and $I_t$ are two SE images of a sample acquired at two different times $t_0$ and $t_1$ respectively. Here, $I_t$ has an unwanted motion (drift) with respect to $I_0$ and need to be compensated. Suppose, $\mathbf{p_1}(x, y) \in I_0$ and $\mathbf{p_2}(\dot{x}, \dot{y}) \in I_t$ are two pixel points, one from each image, the motion between them can be visually reflected by a homography $\mathbf{H}$ between both images and is given by (3.17).

$$\mathbf{p_2} \simeq \mathbf{H}\mathbf{p_1} \tag{3.17}$$

where, $\mathbf{H}$ is a $3 \times 3$ full rank matrix whose elements represent various degrees of freedom. By considering the scaling factor, $\omega$, equation (3.17) is rewritten as in equation (3.18).

$$\omega \begin{bmatrix} \dot{x} \\ \dot{y} \\ 1 \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{3.18}$$

The homography relates the coordinates of two corresponding points through equation (3.19).

$$\dot{x} = \frac{H_{11}x + H_{12}y + H_{13}}{H_{31}x + H_{32}y + 1} \quad \dot{y} = \frac{H_{21}x + H_{22}y + H_{23}}{H_{31}x + H_{32}y + 1} \tag{3.19}$$

Now, obtaining the corrected image $\hat{I}$, corresponds to the registration of $I_t$ with respect to $I_0$. This can be performed in two ways: forward mapping and inverse mapping. Using forward mapping, for each pixel $\mathbf{p_2} \in I_t$, the corrected pixel $\hat{\mathbf{p}} \in \hat{I}$ is obtained directly by copying $\mathbf{p_2}$ using $\mathbf{H}$ as given by (3.20).

$$\hat{\mathbf{p}} = \mathbf{H^{-1}}\mathbf{p_2} \tag{3.20}$$

In this case, some pixels in the corrected image might not get assigned, and would have to be interpolated. This problem could be solved using the second approach, inverse mapping given by equation (3.21).

$$\mathbf{p_2} = \mathbf{H}\hat{\mathbf{p}} \tag{3.21}$$

Using this technique, the position of each corrected pixel $\hat{\mathbf{p}}$ is given and the mapping is performed by sampling the correct pixel $\mathbf{p_2}$ from the current image.

### 3.4.3 Keypoints detection-based drift estimation method

Keypoints detection and matching algorithms form the basis for many computer vision problems like image stitching, data fusion, object recognition etc. The underlying idea with this technique is to extract distinctive features from each image, to match these features for global correspondence and to estimate the motion from the images. The overall task of drift compensation using this method is decomposed into four subtasks that are described below.

#### Keypoint detection

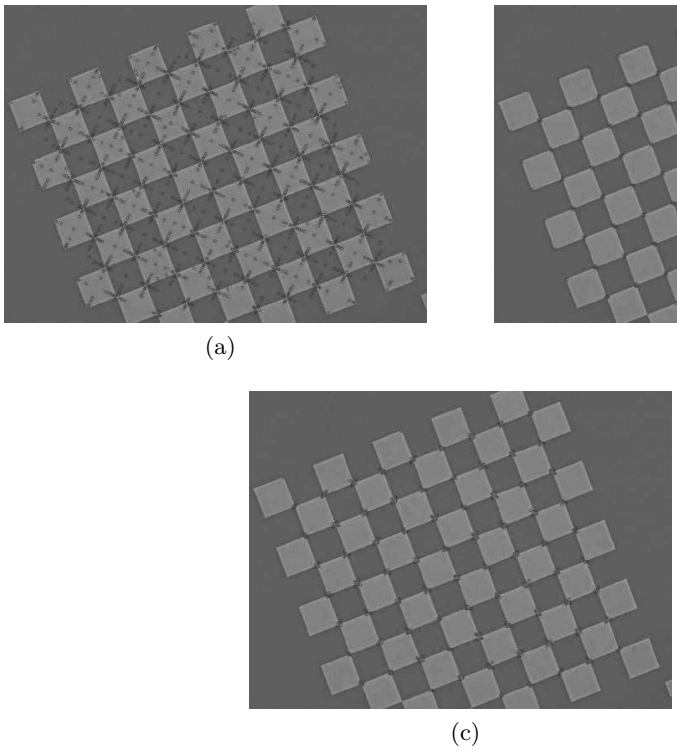
In this step, keypoints are identified from both reference and current images such that the computed points should be same in both images. An interest point detector like Harris detector [HS88], Scale-invariant feature transform (SIFT) [Low04] can be employed for this purpose or alternatively local features like corners, edges etc. can be used. A comprehensive evaluation of the available feature detectors has been provided in [SMB00, MS04, GHT11, MM12]. One main consideration with SEM imaging is, as the brightness and contrast levels are not constant, the detected points should be invariant to the geometry and intensity transformations. For this work of drift estimation, we have tried with three feature point detectors: speeded up robust features (SURF) [BETVG08], features from accelerated segment test (FAST) [RD05] and orient FAST and rotated BRIEF features (ORB) [RRKB11]. Based on the performance evaluation tests, ORB detector has been selected for the final method. Performance evaluation results are discussed in the results section. Figures 3.21(a)-(c) show the detected keypoints using all the three detectors on the images acquired using a microscale calibration pattern at 1500× magnification and figures 3.22(a)-(c) show the detected keypoints for the gold on carbon image shown in figure 3.20(a) .

#### Keypoint description

In this step, all the regions around the detected keypoints from the previous step are converted into compact and stable (invariant) descriptors that can be matched against other descriptors. These descriptors contain information that can be used out of their respective regions. The output of this step is a feature vector that will be used for matching. For this work, two keypoint descriptors, SURF and ORB are used.

**SURF descriptor** This descriptor is used only with the keypoints identified using SURF detector. After identifying the keypoints, the detector assigns an orientation vector to each point. The descriptor is then constructed using this orientation vector and by dividing the region around each keypoint into subregions. For all these subregions, wavelet responses are computed and the sum of the horizontal and vertical responses produces the descriptor vector.

**ORB descriptor** ORB descriptor is an extension technique of binary robust independent elementary features (BRIEF) descriptor. For the detected keypoints, assuming



(a) (b)

(c)

Figure 3.21: Keypoints computed using (a) SURF (b) FAST (c) ORB feature detectors applied on a microscale chessboard pattern.

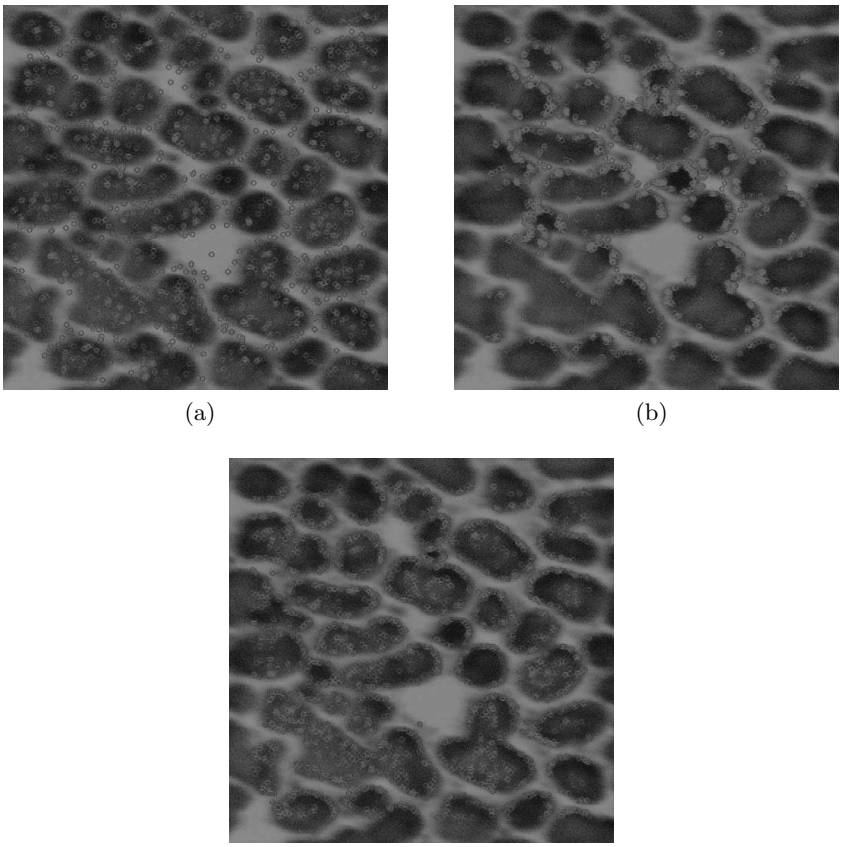
that the corner's intensity has an offset from its center location, an orientation vector is computed from the corner's center to the centroid. This vector is then assigned to the keypoint and used in producing the descriptor.

### Keypoint matching

Once the descriptors are computed, the next step is to extract the correspondences between them. One way is by computing the Euclidean distance **d** between the feature vectors $u_i$ and $v_i$ in the feature space given by (3.22).

$$\mathbf{d} = \sum_i (u_i - v_i)^2 \tag{3.22}$$

Having the Euclidean distance, matching is accomplished by setting a threshold and by returning all the matches less than this value in the images. However, using a large value for the threshold may result in false matches. The other way that has been used in this thesis is to simply match the nearest neighbors (smallest **d**) in the images. This eliminates the problem associated with the previous method. Figure 3.23 shows the matched keypoints using nearest neighbor approach where all matches are shown in figure 3.23(a) and figure 3.23(b) shows the good matches.



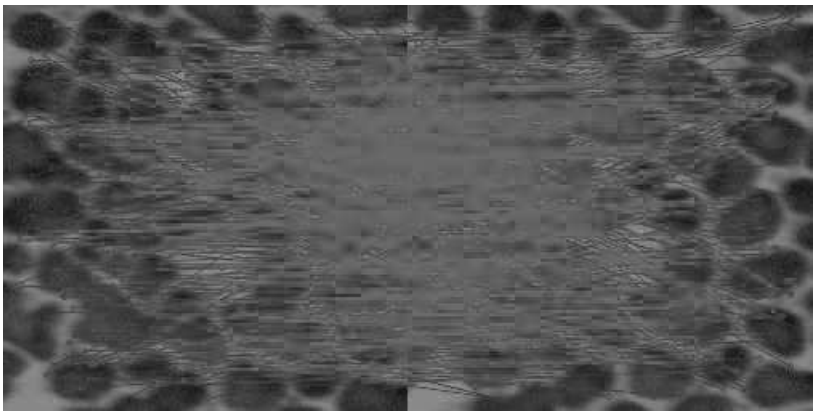
(a) (b)

(c)

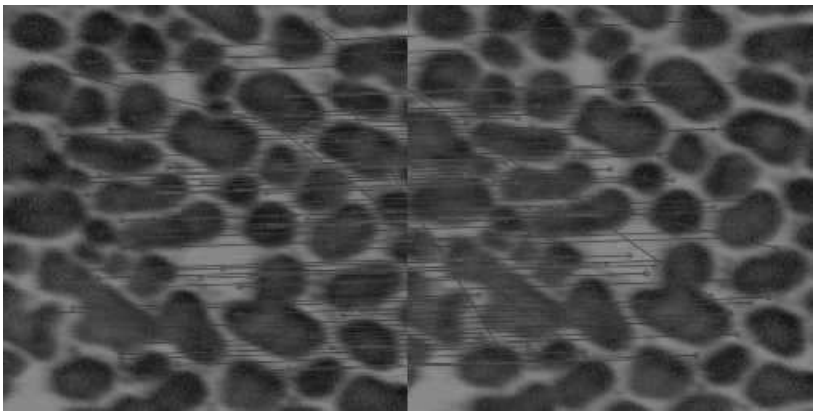
Figure 3.22: Keypoints computed using (a) SURF (b) FAST (c) ORB feature detectors applied on gold on carbon image shown in figure 3.20(a).

### Transformation estimation

The final step is to estimate the drift using the obtained keypoint correspondences. Suppose we get $K$ keypoint correspondent pairs $(x_i, y_i)$ and $(\dot{x_i}, \dot{y_i})$ that are related by an unknown homography $\mathbf{H}$. One way to solve for $\mathbf{H}$ is by using the least squares estimate that is obtained by taking the $2K \times 9$ data matrix $\Psi$ and by finding the eigenvector of $\Psi^T\Psi$ having the least eigenvalue shown in equation (3.23). Since $\mathbf{H}$ is having 8 DOF, atleast 4 points are required to determine the homography.



(a)



(b)

Figure 3.23: Keypoints matching (a) all matches (b) good matches.

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -\dot{x_1}x_1 & -\dot{x_1}y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -\dot{y_1}x_1 & -\dot{y_1}y_1 \\ & & & & \vdots & & & \\ x_K & y_K & 1 & 0 & 0 & 0 & -\dot{x_K}x_K & -\dot{x_K}y_K \\ 0 & 0 & 0 & x_K & y_K & 1 & -\dot{y_K}x_K & -\dot{y_K}y_K \end{bmatrix}_{\Psi} \begin{bmatrix} H_{11} \\ H_{12} \\ H_{13} \\ H_{21} \\ H_{22} \\ H_{23} \\ H_{31} \\ H_{32} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \tag{3.23}$$

Alternatively, the homography can be estimated using an iterative method that has been used in this work, random sample consensus (RANSAC) [HZ00]. The method is briefly summarized in the steps below.

1. Choose four random keypoint correspondences from the available ones and check for co-linearity. If they are collinear, chose another four points.

2. Compute homography that maps the four points exactly to their corresponding points.

3. Find a consensus set for the computed homography by calculating the inliers i.e. find other keypoint correspondences such that their distances from the current model are small.

4. Repeat the step 3 for certain iterations fixed by a threshold and choose the homography (**H**) with largest consensus set (set of inliers).

5. Finally, recompute **H** using all the keypoints in the consensus set.

### 3.4.4 Phase correlation method for drift computation

This is the most common approach for estimating the drift in SEM images. It assumes the drift in images as relative translation. In this work, we use this method for comparing the results obtained from the feature-based methods. The underlying idea behind this method is quite simple and is based on the Fourier shift property. It states that the shift in the coordinate frames of two images in the spatial domain can be realized as the linear phase differences in the Fourier domain. By computing the correlation between the Fourier spectrums, it is possible to estimate the relative shift between the images. The method and the process to derive the shift are explained in detail in chapter 5.

The final outcome after phase correlation is the global drift $(\Delta x, \Delta y)$. Again, the correction is performed using homography. In this case, the matrix **H** only contains the translation information and is given by (3.24).

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \tag{3.24}$$

This method of computing the drift has a remarkable advantage over the traditional cross-correlation techniques; that is its accuracy in finding the peak of correlation at the point of registration. However, this technique can be used only if the motion between two images is a pure translation. In case of rotation, phase correlation produces multiple peaks.

### 3.4.5 Experiments with the system

The experiments with the system are carried out using a standard gold on carbon sample. The acceleration voltage used to produce the beam is $15\,kV$ for and the electronic working distance is set to $6.2\,mm$ in order to keep the sample surface in focus. All tests are performed using SE images with a size of $512 \times 512$ pixels. The $t_D$ is set to $360\,ns$ to get an image acquisition speed of 2.1 frames per second. This speed could be increased up to 3.4 frames per second, but at high scan rates, noise levels in the images are very high.

**Evaluation of the methods**

Initial tests have been performed to evaluate the different methods that are explained above along with the comparison of the keypoint detectors. Gold on carbon sample images acquired after 900 *seconds* from the reference image at $10,000\times$ and $20,000\times$ magnification are used for the evaluation process. Figures 3.24(a) and 3.24(c) show the reference images acquired at $10,000\times$ and $20,000\times$ magnification respectively. The drift compensation has been carried out using both keypoint-based method and phase correlation method. Figures 3.24(b) and 3.24(d) show the corrected images at both magnifications. The images shown are corrected using ORB detector. The black region observed in the borders of the corrected images is due to the compensated drift. Tables 3.4 and 3.5 show the computed homography parameters using all methods at $10,000\times$ and $20,000\times$ magnification respectively. The first four columns ($H_{11}$, $H_{12}$, $H_{21}$, $H_{22}$) provide the rotation information and the last two columns ($H_{13}$, $H_{23}$) provide translation.

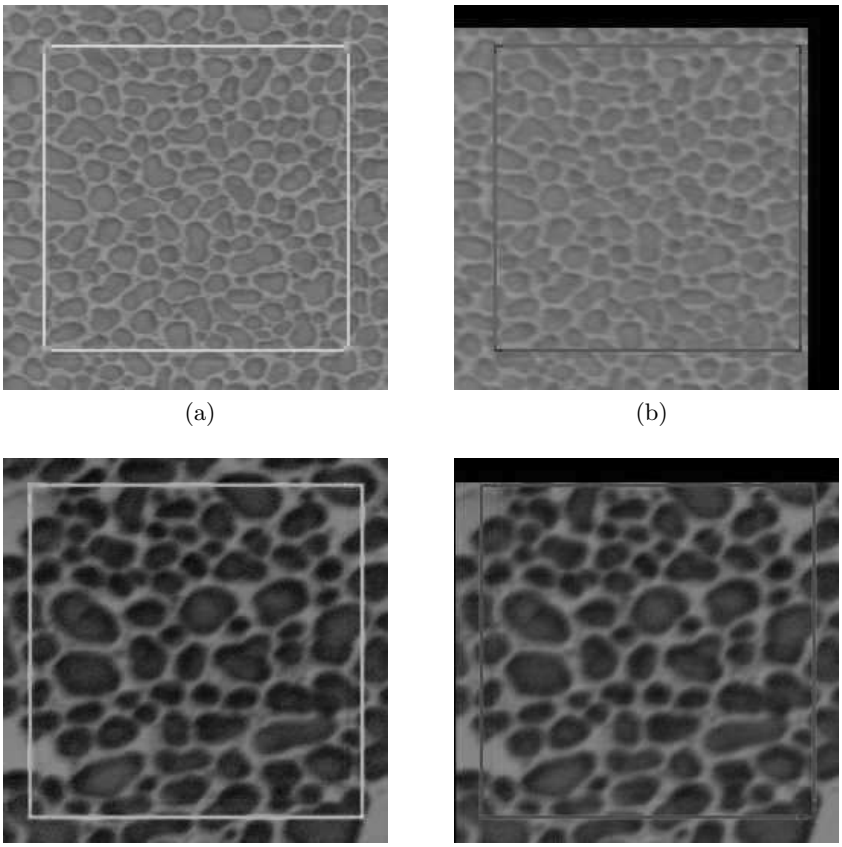
The region of interest (ROI) within the black square shown in every image is used for computing the accuracy of drift correction i.e. to measure how close the corrected image is with the reference image. This is performed by calculating the MSE given by (3.25) between the reference and corrected ROIs. The dimensions of the ROI are updated automatically based on the correction limits and in any case ROI occupies 90% of the corrected region.

$$E_{MSE} = \frac{1}{MN} \sum_{M} \sum_{N} (I_{ROI} - \hat{I}_{ROI})^2 \tag{3.25}$$

where, $I_{ROI}$ and $\hat{I}_{ROI}$ are the regions of interest from reference and current images respectively. The results are summarized in Table 3.6. Figures 3.25(a) and 3.25(b) show the disparity maps computed using the reference and corrected images ROI at $10,000\times$ and $20,000\times$ magnification respectively. These maps provide the relative displacement between the pixels in two images after correction. The total time taken to estimate and correct the drift for one image is shown in Table. 3.7 (implemented in c++). However, the speed is not an important factor with our current system (because of the low frame rate), it has been checked in order to select a specific method that can be used with our new SEM (Carl Zeiss Auriga) having better frame rate in future.

Table 3.4: Homography parameters computed at 10,000× magnification.

| Method | $H_{11}$ | $H_{12}$ | $H_{21}$ | $H_{22}$ | $H_{13}$ | $H_{23}$ |
|---|---|---|---|---|---|---|
| SURF | 0.99 | -0.001 | 0.0009 | 0.99 | -13.57 | 10.13 |
| FAST | 1.03 | -0.005 | 0.006 | 1.01 | -13.968 | 9.73 |
| ORB | 1.003 | -0.003 | 0.0015 | 0.98 | -13.98 | 9.982 |
| Phase corr. | - | - | - | - | -13 | 10 |



(a) (b)

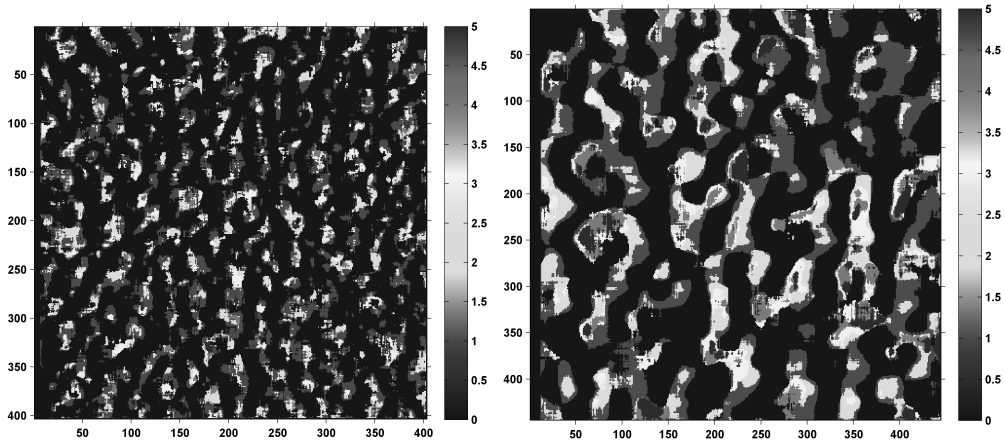
(c) (d)

Figure 3.24: (a) and (b) Reference and corrected images at $10k\times$ magnification. (c) and (d) Reference and corrected images at $20k\times$ magnification. Drift Correction has been performed using ORB detector method. The squares shown in all the images are the automatically computed ROIs. Black region on the borders is the drift displacement after 900s.

Table 3.5: Homography parameters computed at 20,000× magnification.

| Method | $H_{11}$ | $H_{12}$ | $H_{21}$ | $H_{22}$ | $H_{13}$ | $H_{23}$ |
|---|---|---|---|---|---|---|
| SURF | 1.004 | 0.0009 | 0.001 | 0.99 | 1.003 | 10.54 |
| FAST | 0.99 | 0.004 | 0.002 | 0.99 | 0.73 | 10.66 |
| ORB | 0.97 | 0.001 | 0.002 | 0.94 | 0.98 | 10.49 |
| Phase corr. | - | - | - | - | 1 | 10 |

From the obtained results, ORB keypoint detector-based method shows good speed and accuracy over the other methods in compensating the drift. Even though, SURF-



(a) (b)

Figure 3.25: Disparity maps computed using the reference and corrected image ROIs at (a) $10,000\times$ and (b) $20,000\times$ magnification.

Table 3.6: MSE at 10,000× and 20,000× magnification.

| Method | MSE | |
|---|---|---|
| | 10,000× | 20,000× |
| SURF | **7.1441** | **8.2584** |
| FAST | 7.7011 | 8.9798 |
| ORB | **7.1978** | 8.4198 |
| Phase correlation | 11.0914 | 17.3596 |

Table 3.7: Total time taken for drift compensation.

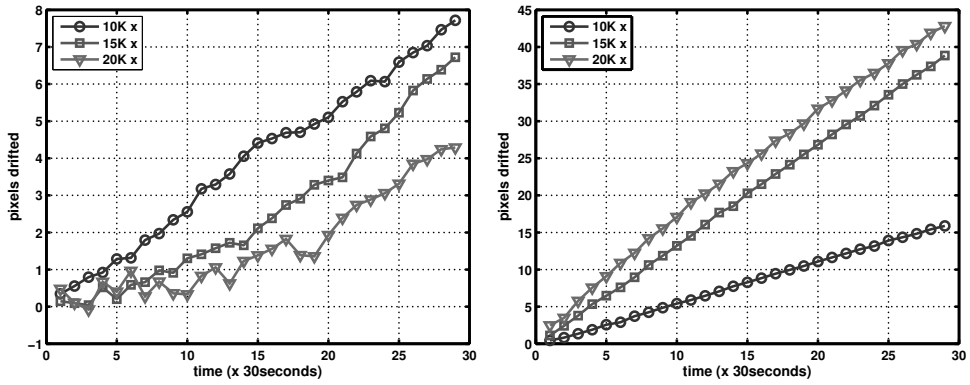
| Method | Time ($ms$) |
|---|---|
| SURF | 871 |
| FAST | 92 |
| ORB | **31** |
| Phase correlation | 78 |

based method also shows good accuracy in correcting the drift; it takes more time for computation. Moreover, all the keypoint-based methods show better performance than the phase correlation method. The major advantage associated with these methods is that the correction can be performed with subpixel accuracy in both rotation and translation. However, these methods are sensitive to the image noise. So the image

quality is monitored continuously during the overall process using the method proposed in section 3.2 [MDP+12b].

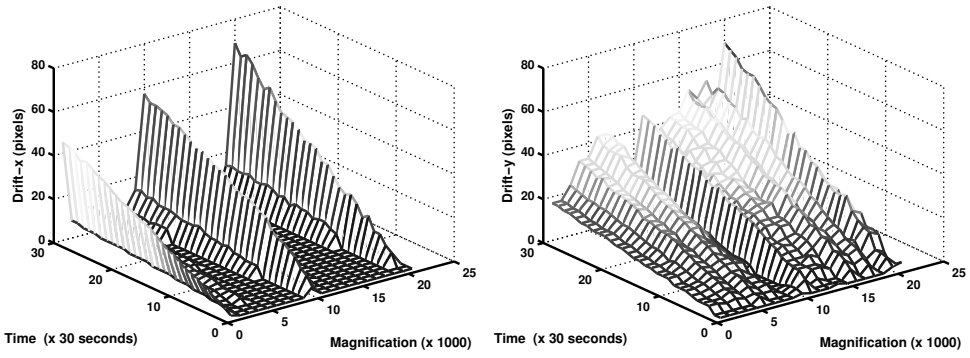
### Drift evolution

Various tests have been performed to evaluate the path followed by the drift at different magnification. For these tests, a set of gold on carbon sample images are acquired for different magnifications ranging from 10,000× to 30,000× with a step change of 1000×. For each magnification, 30 images are acquired at a rate of one image for every 30 *seconds*. ORB detector-based method is used for this test.




(a) (b)

Figure 3.26: Motion of the drift in (a) $x$-axis (b) $y$-axis at different magnifications.




(a) (b)

Figure 3.27: Evolution of drift with respect to time and magnification along (a) $x$-axis (b) $y$-axis.

The plots shown in the figures 3.26(a) and 3.26(b) illustrate the evolution of drift (in pixels) with respect to $x$ and $y$ axes at three different magnifications respectively.

Tables 3.8, 3.9 and 3.10 provides the homography parameters estimated at 3 different magnifications. Plots shown in the figures 3.27(a) and 3.27(b) show the drift evolution with respect to time and magnification.

It is then observed that the drift produced in the images is only a translation in $x$ and $y$ axes and no rotation is involved (from Tables 3.4 and 3.5). The path followed by the drift in $x$ and $y$ axes can be approximated with a linear motion. It is assumed that this drift is a result of the thermal variations inside the chamber and the accumulation of charges on the sample surface over time. However, rotation in the drift can be observed only when there is a change in the focus which is a result of the variations in electromagnetic field produced by the lens[2]. This type of drift can be characterized as the motion in $z$ axis, because the focus can be varied by displacing the sample stage in $z$ direction or by varying the electronic working distance. More details about the focusing in SEM are presented in chapter 4. The presence of the error after global drift compensation (from figure 3.25(b)) clearly demonstrates the existence of pixel level drift. This is mainly occurred due to the charging of the specimen surface in the semi conductive regions. However, for real-time observations, this drift can be neglected as it is not very high.

Table 3.8: Computed homography parameters at $10,000\times$ magnification.

| Time $\times$ 30 $s$ | Rotation | | | | Translation | |
|---|---|---|---|---|---|---|
| | $H_{11}$ | $H_{12}$ | $H_{21}$ | $H_{22}$ | $H_{13}$ | $H_{23}$ |
| 1 | 1.0003 | 0.0006 | 0.0001 | 1.0007 | -1.5048 | 0.8520 |
| 2 | 1.0001 | -0.0000 | 0.0001 | 1.0002 | -2.9327 | 2.0990 |
| 3 | 0.9995 | -0.0002 | -0.0002 | 0.9992 | -4.3075 | 3.2944 |
| 4 | 0.9999 | -0.0004 | -0.0001 | 1.0000 | -5.8901 | 4.3597 |
| 5 | 0.9999 | -0.0000 | 0.0002 | 1.0000 | -7.4547 | 5.4921 |
| 6 | 1.0004 | -0.0001 | 0.0001 | 1.0007 | -8.9595 | 6.4163 |
| 7 | 1.0006 | 0.0001 | 0.0003 | 1.0004 | -10.6892 | 7.5723 |
| 8 | 1.0010 | 0.0002 | 0.0004 | 1.0008 | -12.2627 | 8.5541 |
| 9 | 1.0002 | -0.0004 | 0.0001 | 0.9997 | -13.4724 | 9.8312 |
| 10 | 1.0003 | 0.0003 | -0.0005 | 1.0008 | -15.1336 | 10.8489 |

## 3.5 Conclusion

The SEM is by far an important tool for producing images with nanometer range resolution. Equipped with spectrometers, it is then used as an analytical tool for materials to determine their structural properties (morphology, chemical, crystallography, etc.).

[2]In the presence of electromagnetic field, electrons take helical path. Due to the change in electromagnetic field, the electronic working distance changes simultaneously changing the focus. Thus, when the focus of the device is changed the image rotates.

Table 3.9: Computed homography parameters at $15,000\times$ magnification.

| Time $\times$ 30 $s$ | Rotation | | | | Translation | |
|---|---|---|---|---|---|---|
| | $H_{11}$ | $H_{12}$ | $H_{21}$ | $H_{22}$ | $H_{13}$ | $H_{23}$ |
| 1 | 0.9997 | -0.0003 | 0.0005 | 0.9994 | 0.1454 | 1.1153 |
| 2 | 0.9999 | 0.0002 | -0.0000 | 1.0004 | 0.0948 | 2.4176 |
| 3 | 0.9991 | 0.0005 | -0.0009 | 1.0000 | 0.0419 | 3.7933 |
| 4 | 0.9997 | 0.0006 | -0.0011 | 1.0016 | 0.2418 | 5.1737 |
| 5 | 1.0022 | -0.0001 | 0.0038 | 0.9984 | 0.2009 | 6.4808 |
| 6 | 1.0000 | -0.0003 | 0.0004 | 0.9998 | 0.5882 | 7.6011 |
| 7 | 1.0005 | -0.0001 | -0.0001 | 1.0015 | 0.6385 | 8.9796 |
| 8 | 0.9979 | 0.0002 | -0.0003 | 0.9982 | 0.9805 | 10.6119 |
| 9 | 0.9997 | 0.0007 | -0.0005 | 1.0008 | 0.9119 | 11.8620 |
| 10 | 1.0006 | -0.0000 | -0.0000 | 1.0012 | 1.3045 | 13.1962 |

Table 3.10: Computed homography parameters at $20,000\times$ magnification.

| Time $\times$ 30 $s$ | Rotation | | | | Translation | |
|---|---|---|---|---|---|---|
| | $H_{11}$ | $H_{12}$ | $H_{21}$ | $H_{22}$ | $H_{13}$ | $H_{23}$ |
| 1 | 0.9981 | -0.0005 | -0.0013 | 0.9990 | 0.4803 | 2.4823 |
| 2 | 1.0028 | -0.0001 | 0.0011 | 1.0020 | 0.1037 | 3.4809 |
| 3 | 0.9995 | 0.0027 | -0.0013 | 1.0020 | -0.0783 | 5.7783 |
| 4 | 0.9998 | 0.0002 | -0.0008 | 1.0015 | 0.6695 | 7.5586 |
| 5 | 1.0012 | 0.0005 | 0.0001 | 1.0021 | 0.4050 | 9.0918 |
| 6 | 0.9990 | -0.0008 | 0.0004 | 0.9993 | 1.0130 | 10.8683 |
| 7 | 1.0019 | 0.0013 | 0.0003 | 1.0029 | 0.2744 | 12.2439 |
| 8 | 0.9993 | -0.0002 | -0.0010 | 1.0004 | 0.6748 | 14.1977 |
| 9 | 1.0009 | 0.0013 | -0.0010 | 1.0043 | 0.3033 | 15.5500 |
| 10 | 1.0012 | 0.0010 | 0.0001 | 1.0027 | 0.3315 | 17.0818 |

Today, the trend is to add an ion beam, a gas injector and at least one robot to the microscope column to allow machining, characterization or assembly of structures. In this case the observed scene is dynamic. But electronic imaging is by far known for its sensitivity to the scanning frequency: a high frequency of acquisition leads to low quality images. Apart from this, it is effected by the time varying drift distortion at high magnification. In this chapter, we have demonstrated two methods respectively, for monitoring the image quality and compensating the drift in real time.

The developed method for quality monitoring satisfies a dual constraint: the ability to give an accurate result and the property to run quickly. The first lesson that can be drawn from the obtained evaluation results is that the image SNR level for the FEG

SEM is higher than that of the tungsten gun SEM. This makes sense because the FEG SEM, advanced technology, is more efficient than the filament gun microscope, an older technology. The second lesson, by far the most important, is the instability of the ratio: it varies randomly and greatly depending on time as well as magnification. The analysis of the average change over time revealed two phases. In the first phase, the ratio increases and decreases suddenly: it seems to correspond to the rise in temperature of the filament. The second phase is a slow growing of the ratio that corresponds to a temperature stability of the filament over time. The great merit of this work done allows the quantification of the instability of SEM imaging. The estimator developed has been used in our current works: autofocusing, nano positioning and will be used in our future works of manipulation using visual servoing: machining, assembly and characterization.

The presented drift compensation method is a relatively new (for this problem) and is based on image registration using homography. Unlike the classical correlation-based methods, the implemented method has an ability to correct the rotation, shear and scaling. The obtained results clearly shows that drift in the images corresponds to an inter-frame translation and no rotation is involved. However, the direction and velocity of this phenomenon is not constant and are hard to predict. Apart from the global drift, the pixel level drift that is mostly related to the charging and the used material surface has been observed in the experiments. Even though the developed method is capable of correcting the global drift, a sub-pixel level compensator need to be implemented in order to perform accurate compensation.

Chapter 4

# Autofocusing, depth estimation and shape reconstruction from focus

*Various methods that are developed in relation to the SEM image focus are explained in this chapter. The primary and most important one is the autofocusing technique. In general, fast and reliable autofocusing methods are essential for performing automatic nano-objects positioning tasks using a SEM. So far in the literature, various autofocusing algorithms have been proposed utilizing a sharpness measure to compute the best focus. Most of them are based on iterative search approaches; applying the sharpness function over the total range of focus to find an image in-focus. The developed method in this chapter is a relatively new, fast and direct method of autofocusing that is based on the idea of traditional visual servoing. It performs autofocusing by controlling the focus step using an adaptive gain. It has been validated on the JEOL SEM under different operating conditions.*

*Microhandling of parts with gripper includes a gripping phase and a release phase. In both phases, the detailed knowledge of the inter-object depth (between gripper and the part) and structure of the scene is crucial. As a direct application for the developed autofocusing technique, a method to estimate this inter-object depth has been presented. It computes the depth between micro-objects present in the scene by performing autofocusing separately for each object region windows. These two methods are very important while performing a robotic micromanipulation task. Next, a focus-based 3D reconstruction method has been presented. Both the developments, inter-object depth estimation and 3D shape reconstruction are validated in a robotic handling scenario where the scene contains a microgripper and silicon microstructures.*

## 4.1 Autofocusing in SEM

The automation of nanomanipulation tasks inside a SEM vacuum chamber with nanometric accuracy requires the use of SEM vision feedback in control laws [WSJ+05b]. Such control strategies require efficient image processing algorithms to extract visual features in designing the control laws. Hence, it is essential to use high quality and sharp images to ensure nanometric precision when extracting visual features. To obtain a sharp image and to use the SEM at its maximum potential, it is necessary to apply an efficient, accurate, robust and fast autofocus before performing any nanohandling processes. Besides, autonomous nanomanipulation using SEM vision is a fully unattended operation that requires continuous monitoring of image quality from time to time. Still this task has to be performed by an expert human operator. By implementing an autofocus, this problem can be solved.

Basically in vision systems, focusing is defined as a process of obtaining maximum sharpness in the images by adjusting the objective lens. Most of the present day autofocusing methods are classified into two types: active and passive [BD95]. The active methods are most commonly seen in the modern day digital cameras with CCD (charge coupled device) or CMOS (complementary metal oxide semiconductor) sensor, where the system uses a different measurement system such as infrared and ultrasonic to find the distance between the object and lens i.e. to compute an optimal focal length. Autofocusing is achieved by adjusting the optical system. On the other hand, passive methods do not require any additional devices and they determine the best focus by analyzing the image recorded at the sensor. These methods are categorized into two sub categories where the first one is based on selecting an optimal sharpness function and applying it over a total range of focus steps to find the image with maximum sharpness point. The second one is based on developing a generic imaging model to determine the image defocus information from a set of images. It relates the point spread function of an image with the focal length. The developed autofocusing technique in this work is a kind of passive method that performs focusing directly using the SEM SE images.

Traditional autofocusing methods are mostly concentrated on selecting an optimal sharpness function and applying it over the total range of focus steps to find the image with maximum sharpness (i.e. maximum focus score). Groen et al. [GYL85] and Sun et al. [SDN05] have compared different sharpness functions that are available in the literature. A Fourier transform-based autofocusing method for SEM is presented in [OPT97]. However, Fourier-based method is limited by its complex computations. For SEM, with a high range of focus, applying the sharpness function at each focus step and finding the sharp image is a time-consuming process. To overcome this problem, many works have used the iterative search approaches to find the best focus point [Bat00, RMM09]. However, the electromagnetic lenses present in the SEM show nonlinear hysteresis effects [TSH79, Bat00]. Since the iterative methods are mostly depended upon the search history, the overall accuracy is influenced by this behavior of lenses.

So, in order to provide a dynamic autofocus, a reliable and accurate method has to be developed. In this work, a real-time autofocusing method for SEM has been developed.

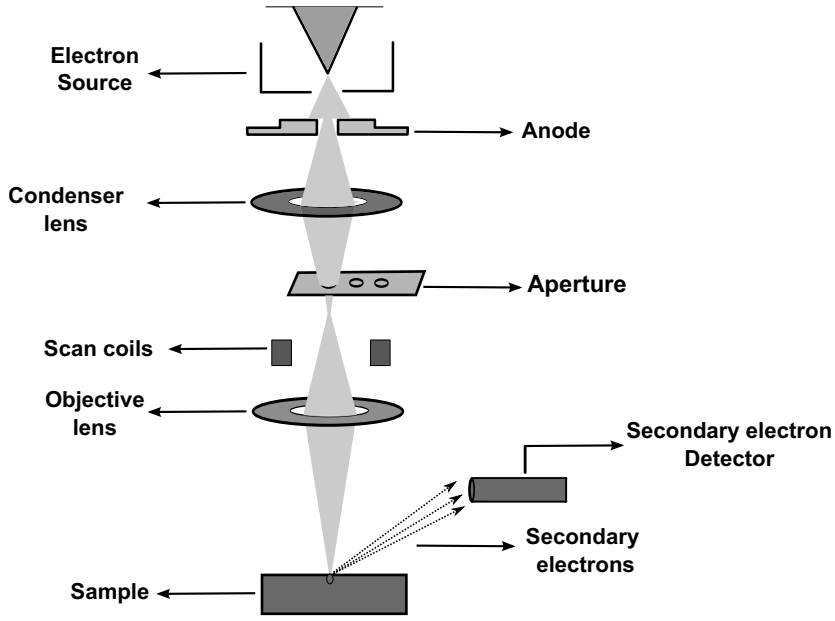



Figure 4.1: SEM probe movement model and operating principle.

It is a kind of passive method that performs focusing directly by controlling the focus step using an adaptive gain that varies with the image sharpness.

Before going into the developed approach, it is necessary to have sufficient knowledge regarding the focusing process in SEM. In the next sections, the focusing process along with the evaluation of different sharpness measures is detailed.

### 4.1.1 Focusing in SEM

The two sets of electromagnetic lenses that contribute towards the focusing in SEM are condenser and objective lenses. The condenser lenses lying above the aperture strip (see figure 4.1) are mainly responsible for controlling the spot size. Initially, these lenses converge the electron beam to a spot. Later, this spot sized beam flares out again and passes through the aperture, where low energy and non-directional electrons are filtered out. The objective lenses that are present below the aperture converge the beam once again and focus it onto the sample surface (figure 4.1). Coarse focusing of the sample surface is performed by adjusting the electronic working distance that is attained by modifying the current passing through the objective lenses. This distance is electronically measured between the focused plane (on the sample surface) and the lower pole piece of the objective lens (electron column) and is not to be confused with the sample stage height that can be controlled externally.

The focusing geometry in a SEM is shown in figure 4.2(a). At a distance $D/2$ on both sides of the focusing plane for a selected magnification, the beam diameter is two times the pixel diameter and produces a blur image. Within the distance $D$ (depth of focus), the image appears to be acceptably in focus. Using the aperture diameter $A$ and

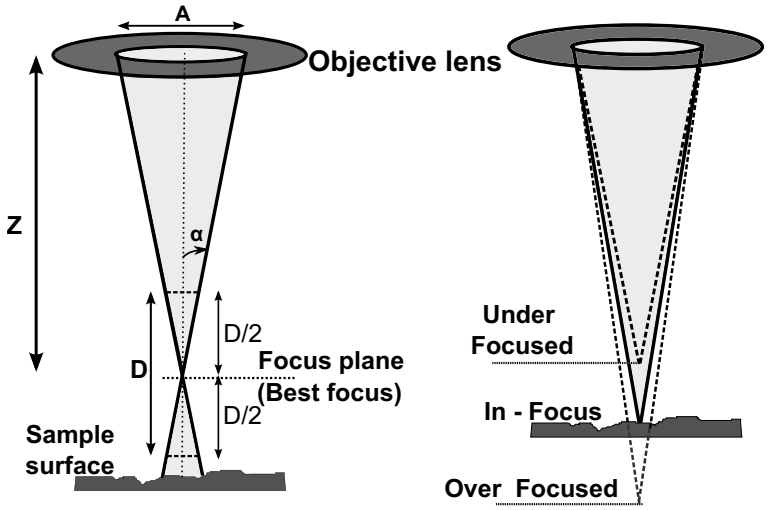



Figure 4.2: (a) Focusing geometry in a conventional SEM. (b) Various focusing scenarios in SEM.

the working distance $Z$, the semi angle of the electron beam can be given by equation (4.1).

$$\alpha = \tan^{-1}\left(\frac{A}{2Z}\right) \simeq \left(\frac{A}{2Z}\right) \text{ at } \alpha < 100\ mrad \tag{4.1}$$

If we consider the resolution on the sample is $\delta_{sample}$, the corresponding resolution on the screen is $\left(\frac{\delta_{screen}}{M}\right)$, where $M$ is the magnification. Depending on the angle $\alpha$ and resolution $\delta$, the depth of focus $D$ is given by (4.2).

$$D = \frac{\delta_{sample}}{\alpha} \simeq \frac{2\delta_{screen}Z}{AM}\ [\mu m] \tag{4.2}$$

So the depth of focus is mainly dependent on the size of aperture selected and electronic working distance. Figure Figure 4.2(b) shows the various focusing scenarios in SEM.

The JEOL SEM used for the experiments is installed with a dynamic focusing module that modifies the current passing through the objective lens and thus the electronic working distance can be preselected. The DISS5 system provides a simple control for the focus by linking the working distance with a series of focus steps (i.e each step modifies the working distance). The relation between focus steps and electronic working distance is shown in Fig. 4.3. Subsequently, the value of working distance (Z) for any given focus step (F) at a fixed sample height can be computed by using the curve equation obtained by approximating it with a polynomial given in equation (4.3). The coefficients of the polynomial are estimated using least squares fit and are summarised in table 4.1.

$$Z = \begin{cases} 50, & \text{if } F \leq 573 \\ 1, & \text{if } F \geq 1218 \\ p_1F^4 + p_2F^3 + p_3F^2 + p_4F + p_5, & \text{otherwise} \end{cases} \tag{4.3}$$

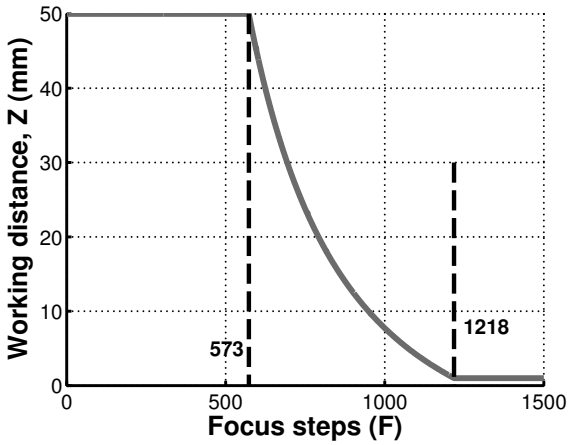



Figure 4.3: Relationship between the focus step and working distance in JEOL SEM used for experiments.

where, $p_1 \dots p_5$ are the polynomial coefficients.

Table 4.1: Computed polynomial coefficients to find the working distance.

| Coefficient | Value |
|---|---|
| p1 | $-7.2 \times 10^{-6}$ |
| p2 | 0.0011 |
| p3 | $-0.074$ |
| p4 | 2.6 |
| p5 | $1.8 \times 10^{3}$ |

### 4.1.2 Evaluation of image sharpness functions

In the context of this work, various sharpness measures that have been used in many focusing algorithms are studied [GYL85, SCN93, SDN05, RMM10]. The principle of accomplishing autofocusing based on the computed sharpness score using a sharpness function is shown in figure 4.4. It is considered that the image with best focus has the maximum value of sharpness score. Different sharpness functions are explained below using the following notations: sharpness function $\mathbf{S}$, image $\mathbf{I}_{M \times N}$ where $M$ and $N$ are the width and height, intensity $I(u, v)$ where $u$ and $v$ are pixel coordinates.

#### Derivative-based sharpness functions

Derivative-based sharpness functions are widely used in many experiments. These methods are based on computing the norm of the gradient between a pixel and its neighbors. The underlying principle behind these methods is that the focused images have high in-

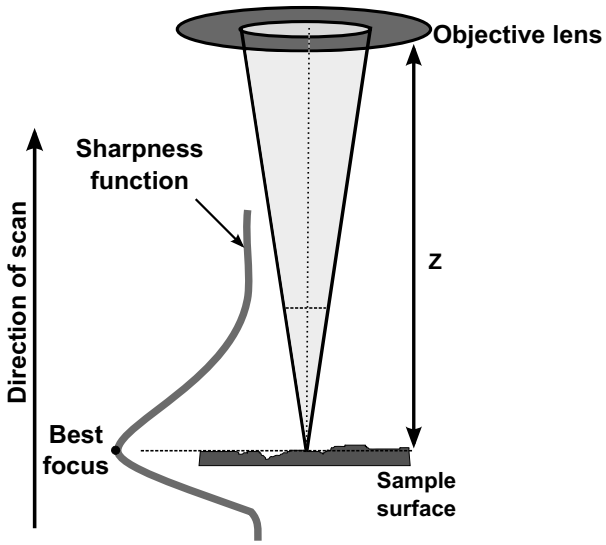



Figure 4.4: The principle of autofocusing in SEM.

tensity differences between the neighboring pixels than in out-of-focus images. However, these methods are sensitive to noise.

**Threshold absolute gradient** [SOdSV+97] This function computes the sum of the absolute values of the first derivative that is greater than a threshold $\theta$.

$$\mathbf{S}_{absgradient} = \sum_{M}\sum_{N} |I(u,v) - I(u+1,v)| \tag{4.4}$$

Where $|I(u,v) - I(u+1,v)| > \theta$.

**Squared gradient** [SOdSV+97] It is similar to the absolute gradient except squaring the differences. So that the gradient values become high and have more influence on the final values.

$$\mathbf{S}_{squaregradient} = \sum_{M}\sum_{N} (I(u,v) - I(u+1,v))^2 \tag{4.5}$$

Where $(I(u,v) - I(u+1,v))^2 > \theta$.

**Modified Laplacian** [NN94] This function convolves the image with a Laplacian edge detection mask given by $L$ to compute the second derivative $d(u,v)$ of the image. $L$ is given by

$$L = \begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix} \tag{4.6}$$

The sharpness function is given by (4.7).

$$\mathbf{S}_L = \sum_M \sum_N |d(u,v)| \tag{4.7}$$

**Brenner [BDH+76]** Brenner's function uses the difference between a pixel value and its second neighboring pixel value.

$$S_{Brenner} = \sum_M \sum_N |I(u,v) - I(u+2,v)| \tag{4.8}$$

**Statistics-based sharpness functions**

Usually, the statistics-based focus measures are less sensitive to image noise and are widely used in many autofocusing algorithms.

**Variance [YOS93]** This function computes the variations in the grey level intensities among the pixels.

$$\mathbf{S}_V = \frac{1}{MN} \sum_M \sum_N (I(u,v) - \mu)^2 \tag{4.9}$$

Where, $\mu$ is the image mean pixel intensity and is given by

$$\mu = \frac{1}{MN} \sum_M \sum_N I(u,v) \tag{4.10}$$

**Normalized variance [YOS93]** This function normalizes the final value by the pixel mean intensity.

$$\mathbf{S}_{NV} = \frac{1}{MN} \frac{1}{\mu} \sum_M \sum_N (I(u,v) - \mu)^2 \tag{4.11}$$

**Correlation [Vol87, Vol88]** Two focusing measures using the correlation technique have been proposed. One is based on the auto-correlation function (ACF) and the other is based on standard deviation function. The ACF based sharpness function is given by

$$\begin{aligned} \mathbf{S}_{Auto-corr} = & \sum_{\substack{M-1 \\ \int \\ u=1}} \sum_{\substack{N \\ \int \\ v=1}} I(u,v) \,.\, I(u+1,v) - \\ & \sum_{\substack{M-2 \\ \int \\ u=1}} \sum_{\substack{N \\ \int \\ v=1}} I(u,v) \,.\, I(u+2,v) \end{aligned} \tag{4.12}$$

The standard deviation-based function is given by (4.13).

$$\mathbf{S}_{std} = \sum_M \sum_N I(u,v) \,.\, I(u+1,v) - M \,.\, N \,.\, \mu^2 \tag{4.13}$$

### Wavelets-based sharpness functions

Various sharpness functions have been proposed using the wavelet transform [YN03].

**Wavelet-1** It applies both high-pass ($H$) and low-pass ($L$) filters to an image and divides it into four sub-images $LL$, $LH$, $HL$ and $HH$. Focus score is computed by summing the absolute values in $HL$, $LH$ and $HH$ regions.

$$\mathbf{S}_{W1} = \sum_{M}\sum_{N} |W_{HL}(u,v)| + |W_{LH}(u,v)| + |W_{HH}(u,v)| \tag{4.14}$$

**Wavelet-2** This function sums the variances in $HL$, $LH$ and $HH$ regions. The mean values in computing the variances for each region are computed using the absolute values.

$$\mathbf{S}_{W2} = \frac{1}{MN}\sum_{M}\sum_{N}(|W_{HL}(u,v)| - \mu_{HL})^2 + (|W_{LH}(u,v)| - \mu_{LH})^2 + (|W_{HH}(u,v)| - \mu_{HH})^2 \tag{4.15}$$

### Histogram based sharpness functions

These functions use histogram $H_i$ (denotes number of pixels with intensity i) of the image.

**Range function [FCC+91]** This function computes the difference between maximum and minimum intensity levels.

$$\mathbf{S}_{range} = \max_i(H_i > 0) - \min_i(H_i > 0) \tag{4.16}$$

**Entropy function [FCC+91]** This function computes the difference between maximum and minimum intensity levels.

$$\mathbf{S}_{entropy} = \sum_i p_i \log_2(p_i) \tag{4.17}$$

Where, $p_i = \frac{H_i}{MN}$ is the relative probability of a pixel with intensity $i$.

From the comparison results provided by Sun et al. [SDN05], six sharpness functions given by equations (4.7), (4.8), (4.11), (4.12), (4.14) and (4.14) lead to optimum results. All these six functions have been evaluated to find a fast and reliable sharpness estimator to use with the autofocusing algorithm. The evaluation results of the different sharpness functions are shown in figure 4.5. After evaluating all the sharpness functions, it is noticed that all the functions have achieved a maximum value at the same focus image with more or less efficiency. Out of all, the normalized variance sharpness function resulted in a curve with less noise and a well-defined peak in less time. Eventhough

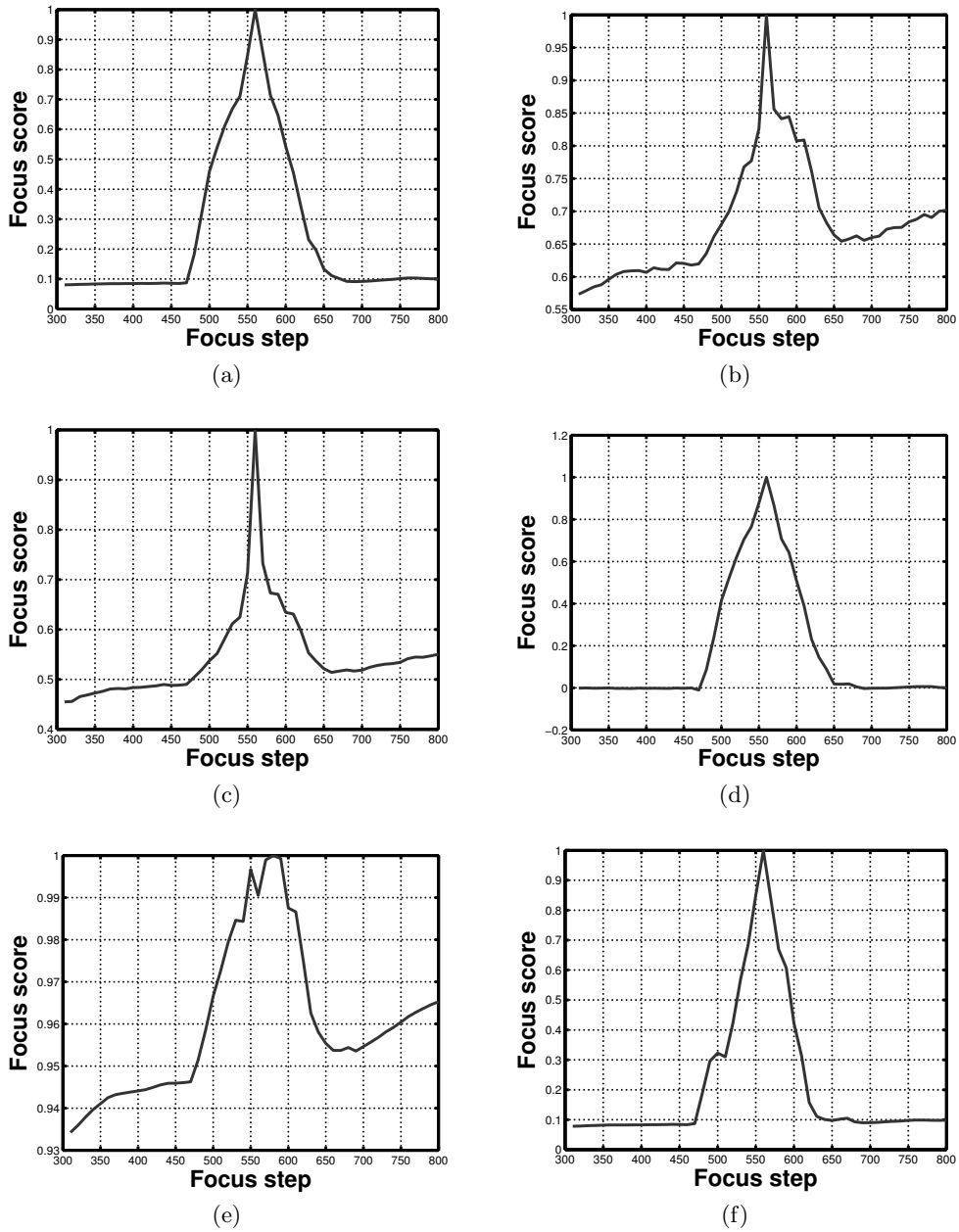



Figure 4.5: Performance of various sharpness functions: (a) Normalized variance (b) Laplacian (c) Brenner (d) Autocorrelation (e) Wavelet-1 (f) Wavelet-2.

Table 4.2: Processing time for various sharpness measures

| **Function** | **Processing time** $(ms)$ |
|---|---|
| Norm. variance | 0.7 |
| Laplacian | 16 |
| Brenner | 15 |
| Auto Correlation | 47 |
| STD correlation | 15 |
| Wavelet-1 | 16 |
| Wavelet-2 | 15 |

the other functions like auto-correlation and wavelet-2 exhibit the similar performance, normalized variance is chosen mainly because of its better performance in different conditions like varying brightness and contrast. Apparently its implementation is simple. Hence, in the design of the visual servoing-based autofocusing, this method has been selected for estimating the sharpness score. The total processing time taken for each function using an image of size $512 \times 512$ is summarized in the table 4.2.

### 4.1.3 Visual servoing-based autofocusing

In this section, the proposed method of autofocusing using the idea of traditional visual servoing is presented. The concept of visual servoing is explained in the chapter 5.

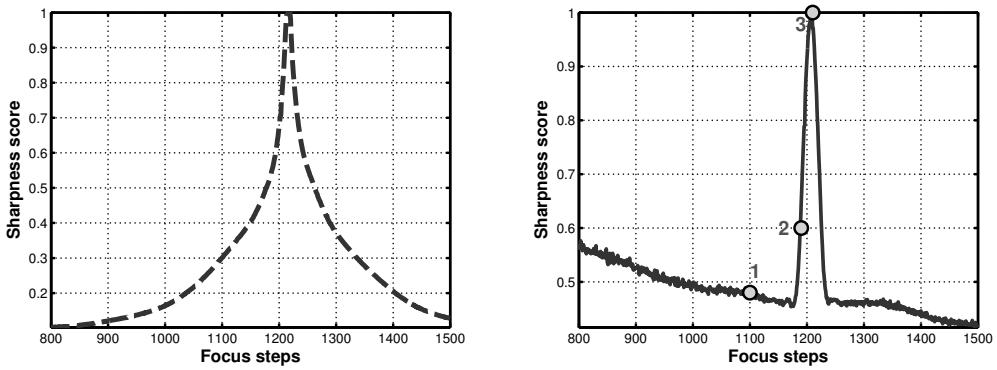
#### Autofocusing using visual servoing

In this work, using the idea of traditional visual servoing, an autofocusing algorithm to control the focus step has been developed. The developed method is based on the following conditions:

- Instead of minimizing the error function, the proposed method maximizes the normalized variance sharpness function given by the equation (4.11).
- Rather than computing the local visual features, the global image information is used in this work. This type of visual servoing is called as direct visual servoing [TLFPM12].
- Most importantly, no reference image or features are used; instead, the method converges to a best focus position i.e. when the sharpness function reaches its maximum value.
- Instead of controlling the camera velocity, the proposed method controls the working distance to obtain, a best focus point.

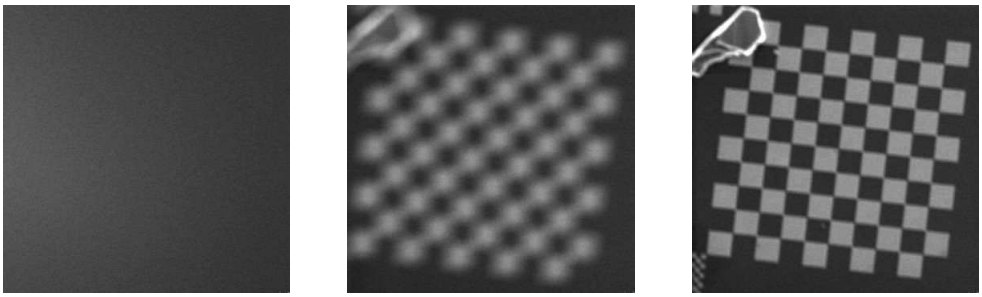
If our goal is to maximize the sharpness function, it is first required to observe the variation of its behavior with respect to the focus steps i.e. with the change in electronic working distance. Figures 4.6(a) and 4.6(b) show this behaviour of the sharpness

function at two different magnifications, 300× and 1000× respectively. It is then observed that, for a specific range of working distance (inside the peak) we get images with some details and the rest are dark or bright images with varying noise. Figures 4.7(a), 4.7(b) and 4.7(c) respectively show the images at points marked by 1, 2 and 3 in figure 4.6(b). From this, the primary goal is to drive the control to reach the focus range using the sharpness information obtained from these plain images. Keeping this in mind, the initial control law has been designed using an adaptive gain that changes with the sharpness score. The corresponding vision-based objective function ($\mathbf{C}$) is given by equation (4.18).




(a) (b)

Figure 4.6: Sharpness score variation with respect to focus steps at (a) 300× (b) 1000×.



(a) (b) (c)

Figure 4.7: Images obtained at the points marked by (a) 1 (b) 2 (c) 3 in figure 4.6(b).

$$\mathbf{C} = \lambda \left( \frac{\bigtriangledown \mathbf{S}}{\| \bigtriangledown \mathbf{S} \|} \right) \tag{4.18}$$

where, $\lambda$ is the adaptive gain that plays vital role in the overall process and is given by equation (4.19).

$$\lambda = \begin{cases} \alpha \left(\frac{\mathbf{S}_0}{\mathbf{S}}\right)^2, & \text{if } \left(\frac{\mathbf{S}_0}{\mathbf{S}}\right) < 1, \\ \alpha, & \text{elsewhere.} \end{cases} \tag{4.19}$$

where, $\alpha$ is a positive integer, $\mathbf{S}_0$ and $\mathbf{S}$ are the sharpness scores computed from first image and the current image. The unit vector $\left(\frac{\bigtriangledown \mathbf{S}}{\|\bigtriangledown \mathbf{S}\|}\right)$ provides the driving direction. Using (4.19), the primary task function $\mathbf{T}_1$ now takes the form given in equation (4.20).

$$\mathbf{T}_1 = F_c + \mathbf{C} \tag{4.20}$$

where, $F_c$ is the current focus step.

The underlying idea is to have a maximum amount of gain at the region far from the focus position so that the next focus step is big enough to reach the focus region. Once it begins to approach the focus position the gain should reduce gradually; so that next focus step will be small or almost equal to unity. At the point of best focus, the unit vector helps in maintaining a constant focus.

Apart from controlling the focus step, it is sometimes required to stop the process at best focus point. Moreover if the focus is set once, it remains unchanged for any magnification. For this purpose, a secondary task has been used along with the primary one. As the primary task can only drive the focus step using sharpness information, it does not have any direct linking with the working distance. In visual servoing, if the primary task does not constrain all degrees of freedom, a secondary task can be added as it has no effect on the primary task [Mar07]. So, for this problem of autofocusing using visual servoing, a secondary task ($\mathbf{T}_2$) has been realized, that links the electronic working distance ($Z$) with the gradient of sharpness function. Later, it has been used as a stopping criterion for the overall process. The secondary task function is given by equation (4.21).

$$\mathbf{T}_2 = \left(\frac{\partial \mathbf{S}}{\partial Z}\right) \mathbf{L}_z \tag{4.21}$$

where, $\left(\frac{\partial \mathbf{S}}{\partial Z}\right)$ is the gradient of the sharpness function with respect to the working distance $Z$ and $\mathbf{L}_z$ is the interaction matrix that links the gradient with the working distance.

Here, assuming the change in the $Z$ is small (especially near best focus point), the gradient has been treated as the difference between consecutive sharpness scores given by equation (4.22). The interaction matrix $\mathbf{L}_z = \frac{-1}{Z}$ where, the $Z$ value at a specific focus step is obtained from equation (4.3).

$$\frac{\partial \mathbf{S}}{\partial Z} = \frac{\mathbf{S}_i - \mathbf{S}_{i-1}}{\triangle Z} \tag{4.22}$$

Using equations (4.20) and (4.21), the final control law to control the focus step with an integrated stopping criterion is given by equation (4.23).

$$F = F_c + \underbrace{\lambda\left(\frac{\bigtriangledown \mathbf{S}}{\parallel \bigtriangledown \mathbf{S} \parallel}\right)}_{\text{main focusing task}} + \overbrace{\left(\frac{\partial \mathbf{S}}{\partial Z}\right)\mathbf{L}_z}^{\text{secondary task}} \tag{4.23}$$

The relationship between the sharpness score and the secondary task function at 1000× magnification is depicted in figure 4.8. The secondary task function thus provides the derivative information of the sharpness score. When the sharpness function reaches it maximum, the secondary task function crosses zero and hence the overall process can be stopped at this point. Because of the presence of noise in the images, the secondary task function have some peaks near zero, to avoid this effect a value of $-0.001$ has been treated as the stopping limit in this work. The total process of autofocusing using visual servoing is summarized in algorithm 1.

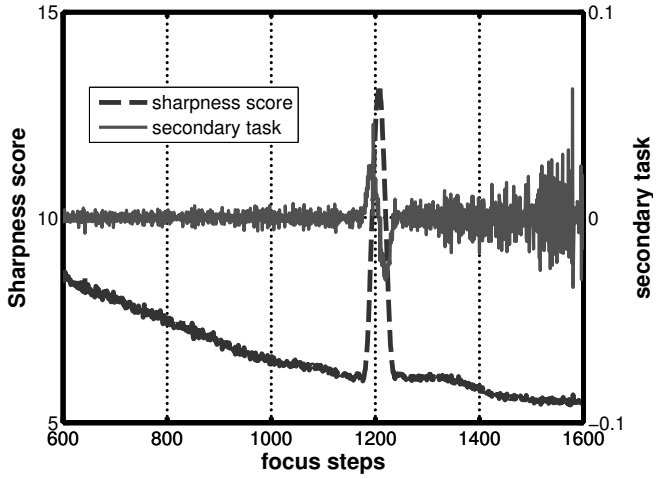



Figure 4.8: Secondary task variation with respect to the focus steps and sharpness score given by equation (4.11).

### 4.1.4 Experimental validations

The proposed method of autofocusing based on visual servoing in SEM has been evaluated at different operating conditions. The experimental setup used for this work has been discussed in the chapter 2. In brief, the SE images acquired by the server are transferred to client over TCP/IP. The control server running from the client computer uses these images and performs sharpness analysis to issue a focus control command to the image server via client. For this work, the image server, the image client and the control server are programmed in C++. An important step before starting the experiment is

**Algorithm 1** Visul servoing-based autofocusing in SEM.

Choose initial value for $\alpha$ ▷ $\alpha$ is a positive value used with gain
Acquire initial image at focus step $F = 10$
Compute $\mathbf{S}_0$ using (4.11) ▷ $\mathbf{S}_0$ is initial sharpness score
$F_{prev} \leftarrow F$
**while** true **do** ▷ Loop until best focus is found
    Set new focus $F_{new}$ ▷ For first iteration $F_{new} = 11$
    Acquire current image
    Compute $\mathbf{S}_{curr}$ using (4.11) ▷ Current sharpness score
    Compute gain $\lambda \leftarrow \alpha \left(\frac{\mathbf{S}_0}{\mathbf{S}_{curr}}\right)^2$
    Compute cost $C \leftarrow \lambda \left(\frac{\nabla \mathbf{S}_{curr}}{\|\nabla \mathbf{S}_{curr}\|}\right)$ ▷ Here, cost is to be maximised
    Compute secondary task $T_2 \leftarrow \left(\frac{\partial \mathbf{S}_{curr}}{\partial Z}\right) L_Z$ ▷ Stop criteria;
    Obtain new focus step $F_{new} \leftarrow F_{prev} + C + T_2$
    **if** $T_2 < -0.001$ **then**
        $bestfocus \leftarrow F_{prev}$
        **return** best focus
    **end if**
    $F_{prev} \leftarrow F_{new}$
**end while**

to tune the astigmatism that has a lot of influence on the imaging process especially at high magnifications. As our SEM does not have an automatic function to control this parameter, a trained professional has performed this task prior to the experiments. For most of the experiments, a microscale calibration rig containing the chessboard patterns that is fabricated at FEMTO-ST (see figure 4.9) is used for imaging. All experiments are performed using images of size $512 \times 512$ pixels and the acceleration voltage used to produce the electron beam is 15 kV.

### Validation of the method

Initial experiments are performed to validate the proposed approach at different magnifications. The tests are conducted using an aperture size of 50 $\mu m$ and $\alpha = 400$ at 300× and 1000× magnifications. The brightness and contrast levels are set to the optimal values. The dwell time is set to 360 $ns$ to achieve an acquisition speed of 2.1 frames per second.

The autofocusing process starts from an initial focus step ($F$) of 10. Figures 4.10(a) and 4.10(b) show the variation of sharpness scores computed using equation (4.11) and the corresponding gain values for each iteration at 300× and 1000× respectively. The secondary task function variation for both magnifications is shown in the figures 4.11(a) and 4.11(b). From the obtained results, it can be noticed that the focusing process stopped precisely at the maximum of sharpness function i.e. at this point the secondary

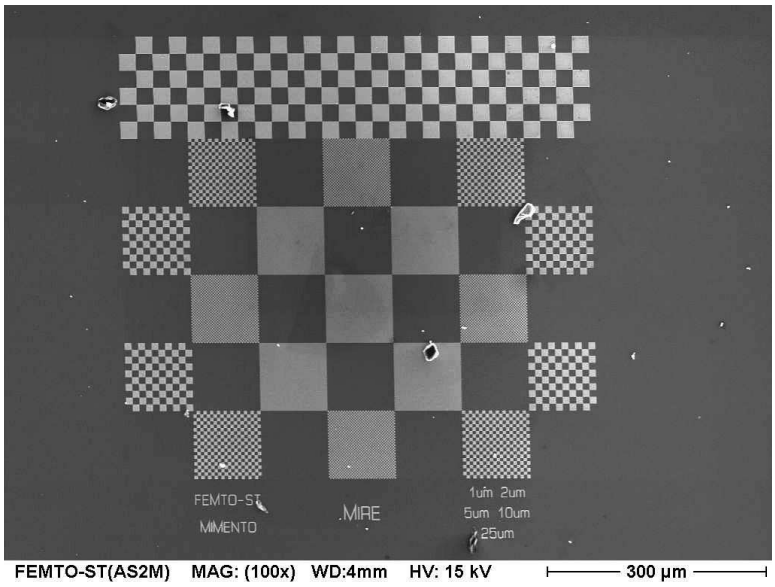
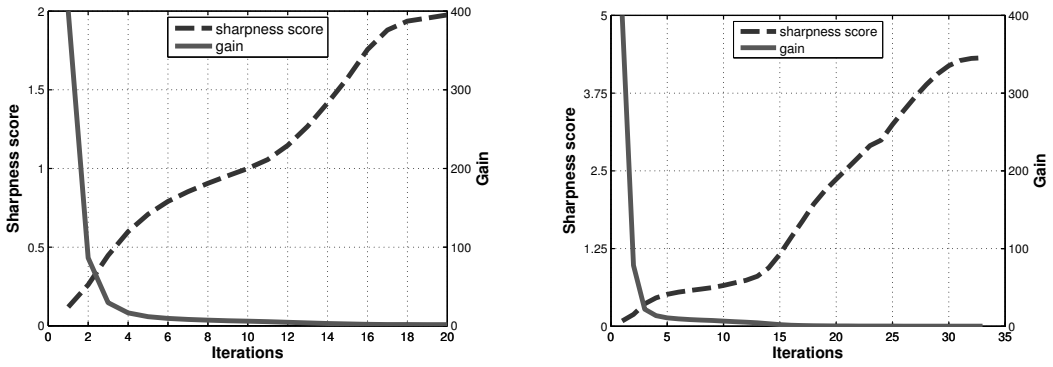



Figure 4.9: Microscale calibration rig with chessboard patterns used for imaging.

task function reached −0.001. Figure 4.12 shows some of the images acquired during the focusing process at 300× magnification. Figure 4.12(c) shows the final image *in-focus*.

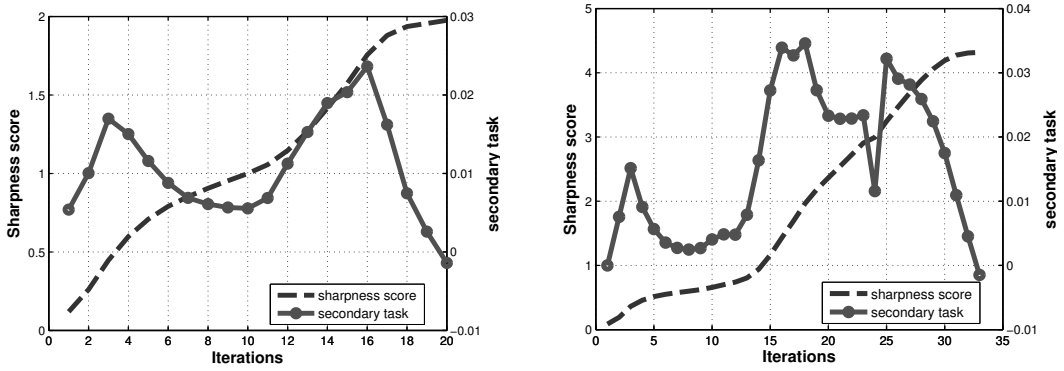



(a) (b)

Figure 4.10: Sharpness score and corresponding gain values at (a) 300× (b) 1000× magnifications respectively.

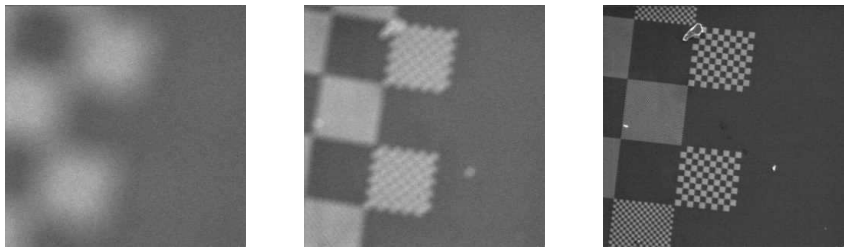
### Validation of the developed method for speed and accuracy

The performance of the proposed autofocusing method is evaluated by comparing it with the traditional search-based algorithms at different experimental conditions. An iterative search approach has been employed for this purpose. The search-based method




(a) (b)

Figure 4.11: Secondary task variation with respect to the sharpness score at (a) 300× (b) 1000× magnifications respectively.



(a) (b) (c)

Figure 4.12: (a) and (b) Images acquired during the focusing process (c) Final focused image.

is implemented over the total range of focus steps using three iterations and is explained in the following steps:

1. For each iteration, divide the total range of focus steps into four sub-regions.
2. For each sub-region, search for the global maxima by scanning with a fixed step size selected for the iteration. The step sizes used for each iteration are 50, 20 and 1 respectively.
3. The maximum value found after the final iteration is the best focus point.

The accuracy obtained by this method has been treated as the best accuracy, since it scans each step for the best focus in its final iteration. Hence, the focus step returned by this method is used as the reference step for computing the step error. However, the hysteresis effects produced by the lens are not taken into account while computing the step error.

The proposed method is compared to the iterative search-based approach in terms of total time taken and accuracy of focusing at different magnifications. Here, the total

time taken also includes the time taken to acquire and transfer the images. The first four tests are performed using $\alpha = 300$. First test is performed using optimal values of brightness, contrast and scan speed (1.4 frames per second). With this configuration, the amount of noise in the images is less. Table 4.3 summarizes the obtained results. When compared with the search-based approach, the average step error is found to be 2 focus steps.

Table 4.3: Time taken and accuracy at optimal conditions.

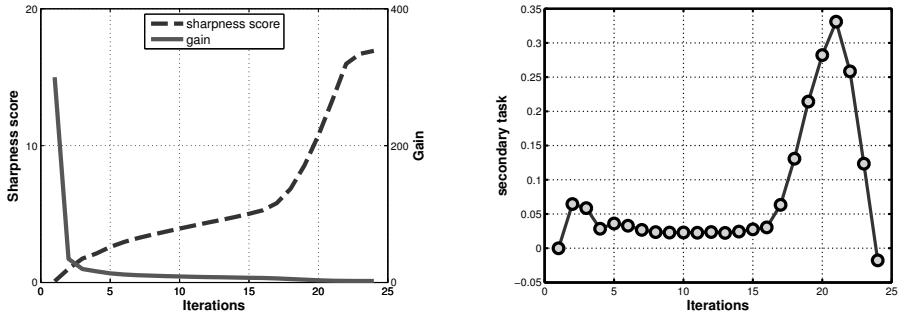
| Mag (×) | Time ($s$) | | Accuracy (Focus steps) (Focus step $\simeq 0.3$ $mm$) | | |
|---|---|---|---|---|---|
| | proposed | search | proposed | search | step error |
| 300 | 42.703 | 117.97 | 1222 | 1219 | 3 |
| 500 | 48.5 | 109.62 | 1220 | 1218 | 2 |
| 800 | 43.46 | 112.03 | 1221 | 1219 | 2 |
| 1000 | 48.42 | 109.4 | 1222 | 1219 | 3 |

Second test is performed with an increased brightness and with a scan speed used for the first test. With an increase in brightness, the textures present in a defocused image appear to be interfering with each other making it difficult to obtain a focused image. The obtained results are summarized in table 4.4 and the average step error is 3 focus steps. Figures 4.16(a) - 4.16(c) show the images obtained during this test at 1000× magnification. Figure 4.13(a) shows the variation of sharpness score and corresponding gain values with each iteration. The secondary task variation is shown in figure 4.13(b).

Table 4.4: Time taken and accuracy at changed brightness.

| Mag (×) | Time ($s$) | | Accuracy (Focus steps) (Focus step $\simeq 0.3$ $mm$) | | |
|---|---|---|---|---|---|
| | proposed | search | proposed | search | step error |
| 300 | 42.06 | 110.64 | 1217 | 1213 | 4 |
| 500 | 49.48 | 118.06 | 1218 | 1213 | 5 |
| 800 | 48.29 | 111.6 | 1215 | 1212 | 3 |
| 1000 | 51.37 | 111.74 | 1216 | 1213 | 3 |

Third test is performed with an increase in the scan speed. This change in the scan speed increases the amount of noise in images. The dwell time used for this test is 180 $ns$ with which 3.2 frames per second can be acquired. Table 4.5 shows the results obtained with this test and the computed average step error is 3 focus steps. Figures 4.16(d) - 4.16(f) show the images obtained during this test at 800× magnification. Figure 4.14(a) shows the variation of sharpness score and corresponding gain values with each iteration at 800× magnification. The secondary task variation is shown in figure 4.14(b).

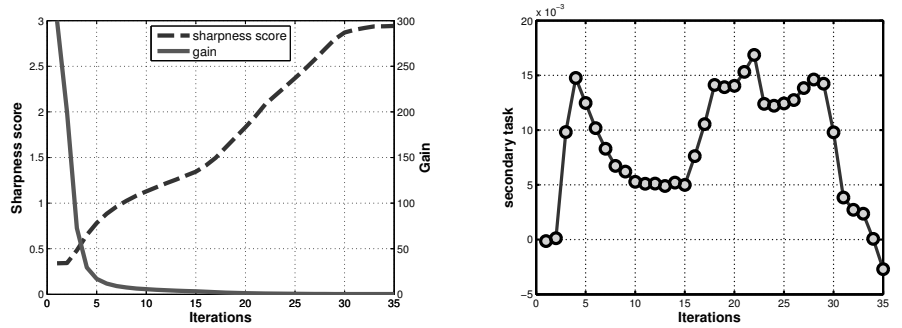



(a) (b)

Figure 4.13: (a) Sharpness score and corresponding gain values with increased brightness test (b) Secondary task variation.

Table 4.5: Time taken and accuracy with change in the dwell time.

| Mag (×) | Time ($s$) | | Accuracy (Focus steps) (Focus step $\simeq 0.3$ $mm$) | | |
|---|---|---|---|---|---|
| | proposed | search | proposed | search | step error |
| 300 | 17.39 | 64.28 | 1224 | 1219 | 5 |
| 500 | 16.23 | 62.15 | 1223 | 1219 | 4 |
| 800 | 16.61 | 62.84 | 1222 | 1219 | 3 |
| 1000 | 14.37 | 65.09 | 1222 | 1218 | 4 |




(a) (b)

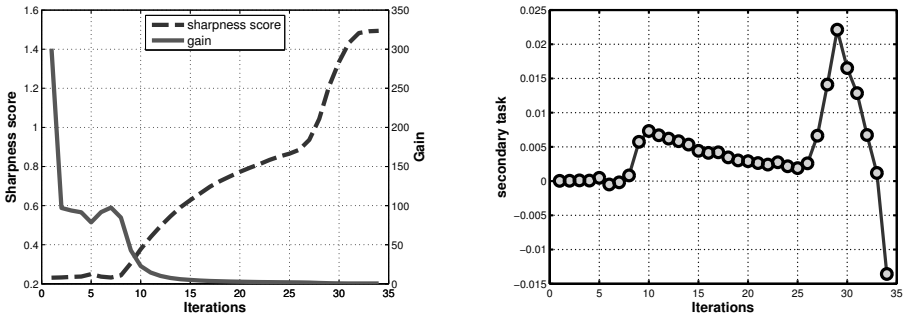
Figure 4.14: (a) Sharpness score and corresponding gain values with increased scan speed (b) Secondary task variation.

Fourth experiment is performed using a silicon dioxide sample containing the microparts of dimensions 6 $\mu m \times 6$ $\mu m \times 2$ $\mu m$. The main reason behind this test is that the surface of this sample is not a good conductor and contaminated. Due to this,

charges are accumulated on the sample surface resulting in the images with grey level fluctuations. Table 4.6 summarizes the obtained results. The average step error computed in this case is 2 focus steps. Figures 4.16(g) - 4.16(i) show the images obtained during this test at 500× magnification. Figure 4.15(a) shows the variation of sharpness score and corresponding gain values with each iteration at 500× magnification. The secondary task variation is shown in figure 4.15(b).

Table 4.6: Time taken and accuracy with silicon dioxide sample at optimal conditions.

| Mag (×) | Time ($s$) | | Accuracy (Focus steps) (Focus step $\simeq 0.3\ mm$) | | |
|---|---|---|---|---|---|
| | proposed | search | proposed | search | step error |
| 300 | 39.1 | 116.82 | 1473 | 1471 | 2 |
| 500 | 36.45 | 99.07 | 1468 | 1471 | -3 |
| 800 | 41.61 | 108.65 | 1472 | 1471 | 1 |
| 1000 | 34.07 | 101.4 | 1476 | 1473 | 3 |




(a) (b)

Figure 4.15: (a) Sharpness score and corresponding gain values with silicon dioxide sample containing microparts (b) Secondary task variation.

Final experiments are performed to examine different speeds that can be attained by the proposed method. The overall speed of the autofocusing process can be changed by changing the $\alpha$ value (see equation (4.23)) or by changing the dwell time. In this case, speeds are tested by changing the $\alpha$ and the obtained results are summarized in table 4.7.

## 4.2 Depth estimation from focus

In this section, a method for estimating the *inter-object* depth using image focus information is presented. The developed method is a direct application of the previously explained visual servoing-based autofocusing technique.

Table 4.7: Time taken (*seconds*) by the proposed method with various $\alpha$ values.

| Mag (×) | $\alpha = 100$ | $\alpha = 200$ | $\alpha = 300$ | $\alpha = 400$ |
|---|---|---|---|---|
| 300 | 96.01 | 69.7 | 42.91 | 19.7 |
| 1000 | 98.3 | 70.17 | 48.4 | 21.4 |

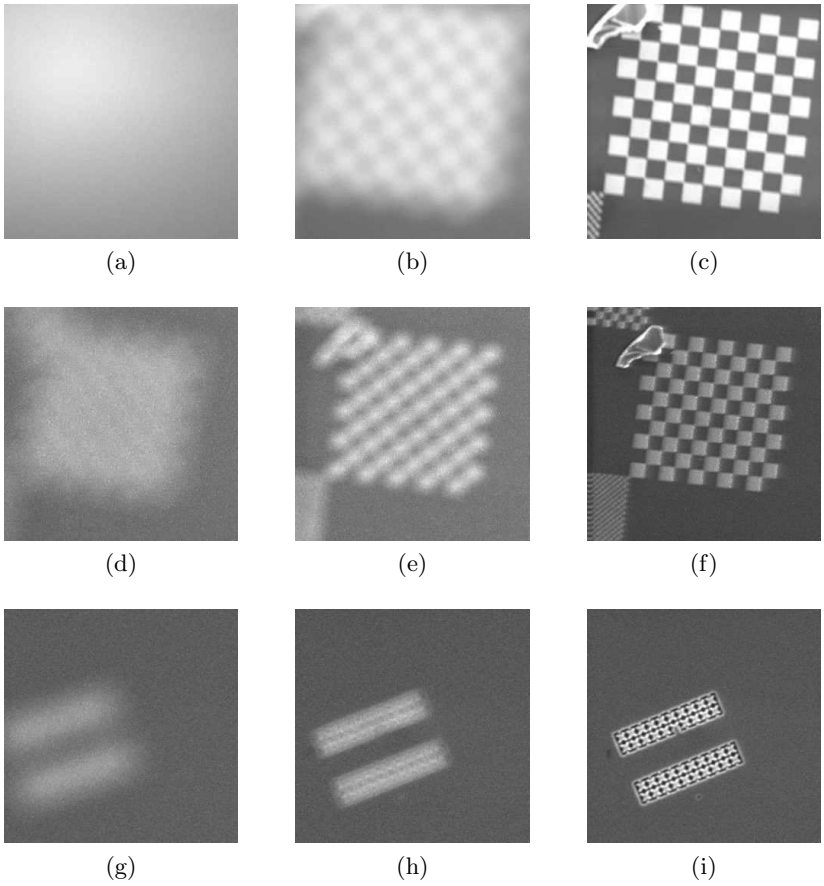


Figure 4.16: Images acquired during the focusing process with (a)-(c) increased brightness at 1000× magnificatin (d)-(f) increased scan time at 800× magnification (g)-(i) silicon dioxide sample containing microparts at 500× magnification.

### 4.2.1 Introduction

Since the early stages of research using SEM, it has been a great interest for many researchers to extract the depth information from the images to use it in micromanipulation and assembly applications. Even though SEM images are capable of providing surface topography information, they are purely two dimensional. The pixel intensity values do not contain any information regarding the height. In order to tackle this

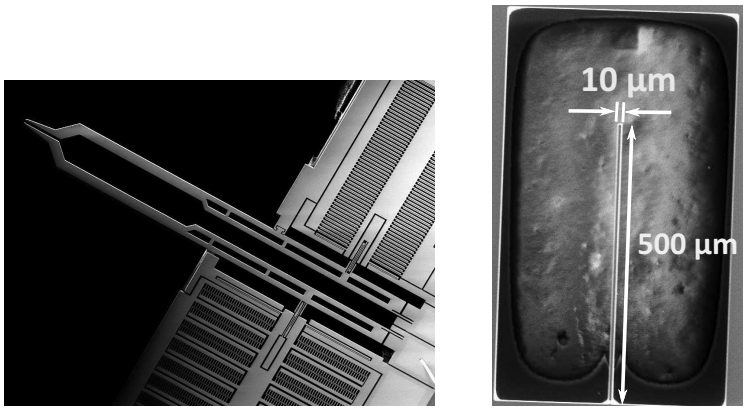
problem, many research works have tried to apply stereoscopic imaging-based technique that is commonly seen in the optical systems [JJ95, PLB06, RMM+12]. As the system possesses only a single imaging sensor, most of them tried to acquire a stereo pair of images by concentrically tilting the sample with a specific angle. The main difficulty lies in determining the accurate tilt angle and finding the correspondence between the acquired noisy SEM images. Moreover, it requires a precise calibration. However, tilting a sample is not a feasible option when performing an autonomous task. To overcome this problem, Jahnisch and Fatikow have developed a special hardware system for beam deflection in order to observe the sample with different angles [JF07]. However, the system is limited by its small angle of deflection. Another solution can be by using a FIB system along with SEM [TF11], but it increases the overall cost of the system and damages the sample. Eichhorn et al. [EFW+08] have used a touchdown sensor to detect the contact between two micro-objects and simultaneously computing the depth in nanometres. Schmoeckel et al. [SWK01] have used a laser triangulation method for depth detection using an additional optical microscope that is fixed inside the SEM vacuum chamber. A photometric stereo-based approach using multiple back scattered electron detectors has been used by Chen et al. in their work [CMK12].

Apart from the stereo imaging-based methods that are relatively complex, the depth information can be extracted simply by using the focus information in the images. This concept of estimating the depth is well-known in the field of optical microscopy and is commonly termed as *depth from focus* [Sub88, GS90]. Most importantly this technique avoids two main problems: correspondence and occlusion that are commonly seen in the case of stereo imaging-based methods. The underlying idea behind these techniques is that they use a sequence of images obtained by moving the microscope along its optical axis with a predefined finite number of steps. Depth is then computed by applying a focus measure on the images to find the camera position at which the object in frame appears to be in focus. One such technique has been used in this work.

In this work, instead of acquiring multiple images at fixed step, the depth information is directly estimated by autofocusing the regions containing the objects. This process is explained using an experimental scenario.

### 4.2.2 Experimental scenario

In this work, the inter-object depth estimation has been performed using a FT G32 force sensing microgripper from FEMTO TOOLS [Fem11] (see figure 4.17(a)) and silicon microparts of dimensions $10\mu m \times 500\mu m \times 20\mu m$ (see figure 4.17(b)). The gripper is fixed to a support inside the SEM vacuum chamber and the microparts are placed on the positioning table underneath the gripper. Normalized variance (given by equation (4.11)) has been used as the focus measure. In this case, the sharpness curve contains two peaks as show in the figure 4.18.




(a) (b)

Figure 4.17: The objects used for the experiments (a) FEMTO TOOLS FT G32 micro-gripper (b) silicon micropart.

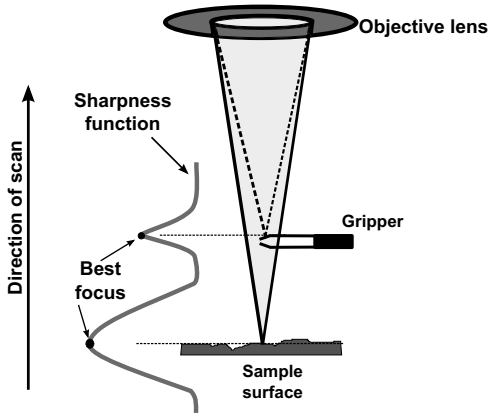



Figure 4.18: Principle of focusing and sharpness score variation with the presence of gripper.

### 4.2.3 Inter-object depth estimation

As an initial approach, the sharpness function has been applied over the total range of focus steps (i.e. by changing the working distance) to find the individual object depth. The resulted curve is shown in the figure 4.19. As expected, two local maxima have been found indicating the height of the gripper and the microparts with respect to the lens pole piece. Although this approach determined the absolute height, it is a time consuming process to scan the entire region of focus.

In order to overcome this problem, in this work, we first segment the image into different regions of interest (ROI). For each ROI, the best focus position is computed using the previously explained visual servoing-based autofocusing technique. By computing the working distance that is associated with the obtained best focus step using equation

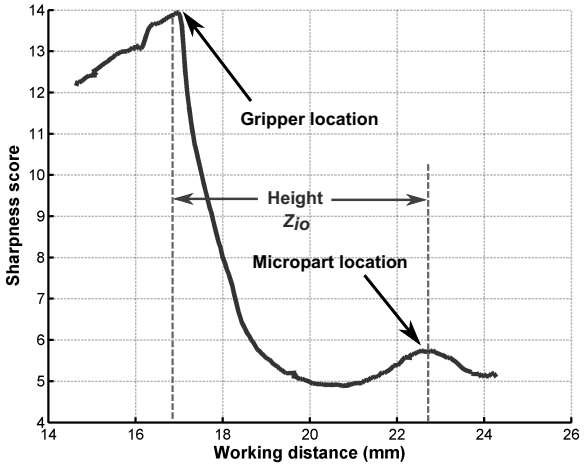


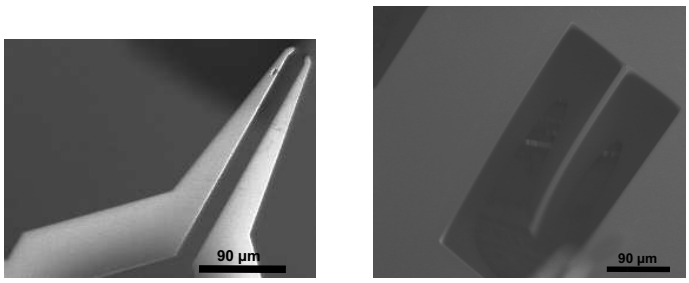

Figure 4.19: Sharpness function indicating the gripper and micropart locations.

(4.3) provides the height of the object. Finally, the inter-object depth $Z_{io}$ is obtained by equation (4.24).

$$Z_{io} = \mid Z_{part} - Z_{gripper} \mid \tag{4.24}$$

where, $Z_{part}$ and $Z_{gripper}$ are the obtained heights of micropart and gripper regions respectively.

For this task, the ROI is determined on the basis of thresholding the pixel intensity values. The value of this threshold has been selected manually. Optionally, the ROI can also be determined by using object recognition or tracking algorithms [WK03, YC09, Sie11]. Figures 4.20(a) and 4.20(b) show the best focused images of each ROI. The plots shown in the figure 4.21 show the results of autofocusing process applied on different regions. Figures 4.21(a) and 4.21(b) show the heights obtained for gripper and micropart regions respectively when they are far from each other and the computed $Z_{io}$ is 5.632 $mm$. Figures 4.21(c) and 4.21(d) show the heights obtained for gripper and micropart regions respectively when they are near and the computed $Z_{io}$ is 1.23 $mm$.




(a) (b)

Figure 4.20: Focused ROI of (a) FT G32 microgripper (b) silicon microparts.

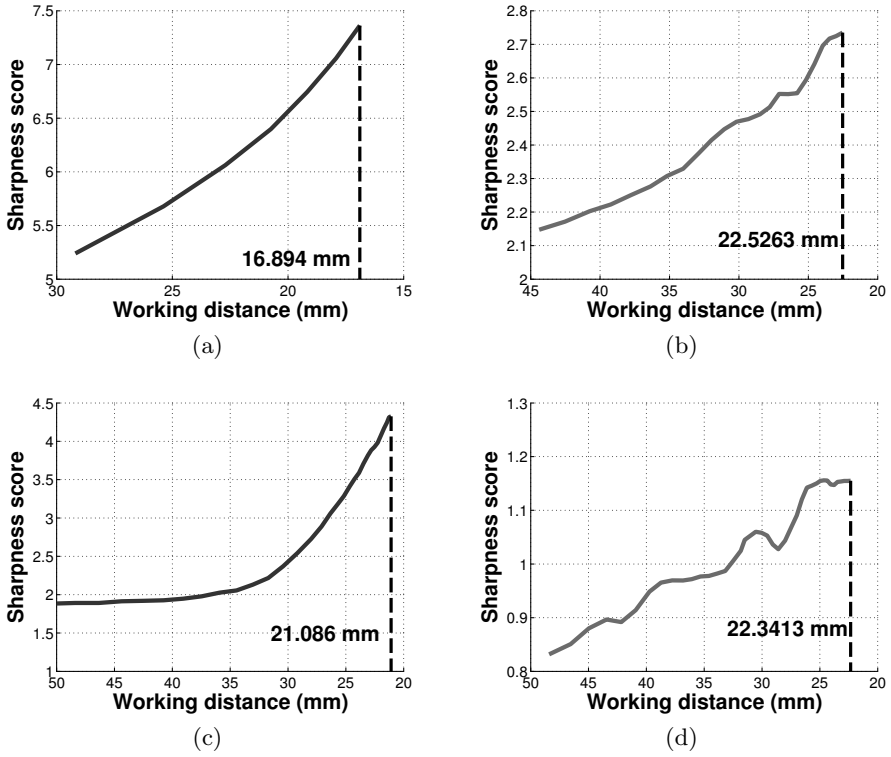



Figure 4.21: Results of autofocusing at different regions. (a) Gripper (b) micropart heights when they are far from each other. (c) Gripper (d) micropart heights when they are near.

## 4.3 Shape reconstruction from focus

In this section, a method to reconstruct the 3D shape of micro-objects present in the scene is detailed. Unlike inter-object depth estimation, the sharpness of each pixel present in the ROI has been maximized in its local neighborhood. The size of the neighborhood is pre-selected. This type of approach is commonly termed as *shape from focus* [NN94]. The underlying idea behind this technique is: it uses a sequence of images obtained by moving the microscope along its optical axis with a predefined finite number of steps. Depth of a pixel point is then computed by applying a focus measure on the acquired images to find the camera position at which the object in frame appears to be in focus [NN94, NNŠ03b, NNŠ03a, Wed04, FTDLFP08]. These type of methods mostly rely on the advantage of weak depth of focus in optical systems. In the case of SEM, in spite of having a high depth of focus, it is still possible to use the focus-based approach by simple modifications in the system imaging parameters. So in this section, the details regarding various system parameters that are influencing the depth of field are presented first. Also for this case, normalized variance has been used as the sharpness measure.

### 4.3.1 Reducing the depth of focus

As mentioned earlier in section 4.1, from equation (4.2), the depth of focus ($D$) mainly depends on the aperture diameter and the working distance. So in order to reduce $D$, we have three possible options: increasing the aperture diameter ($A$), increasing the magnification ($M$), or reducing the electronic working distance ($Z$). By increasing $A$ the semi angle of the beam increases and simultaneously $D$ decreases. This option can be used only with the SEMs having multiple exchangeable apertures. Figure 4.22 shows the relation between aperture diameter and depth of focus. The other option is to increase M. But increase in magnification limits the field of view making it difficult to track all the available parts. The relation between the magnification and depth of focus is shown in figure 4.23. Apart from these two options, depth of field can be reduced by decreasing the electronic working distance i.e. to move the sample (focal plane) close towards the objective lens pole piece. However, this is not a feasible choice, as increase in sample height may damage the lens. The minimum acceptable working distance is about $5 - 8\,mm$.

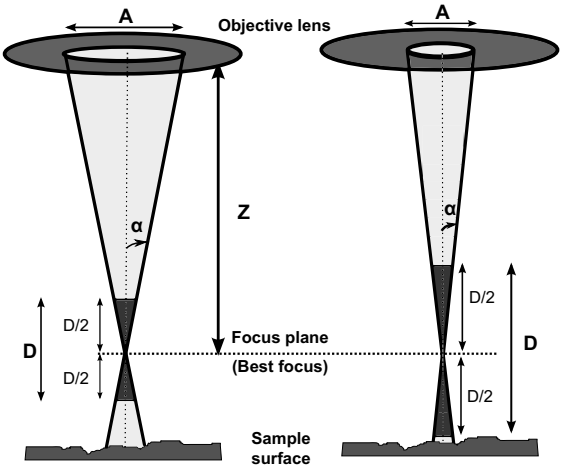



Figure 4.22: Relationship between the aperture diameter and depth of focus in SEM.

In this work, the option of varying the aperture diameter has been selected. The aperture strip present in the JEOL SEM column contains four apertures of diameters $100\,\mu m$, $70\,\mu m$, $50\,\mu m$ and $30\,\mu m$ respectively. By observing the amount of blur, an aperture with 70 $\mu m$ has been selected for this work. However, high aperture sizes may result in low resolution. The magnification used is $300\times$ and the image size is $512 \times 512$ pixels. The measured electronic working distance with the focal plane lying on gripper surface is 23.7 $mm$. With these parameters the computed $D$ is 293 $\mu m$. So to achieve good resolution in reconstruction it is necessary to take images with a distance step greater than 293.

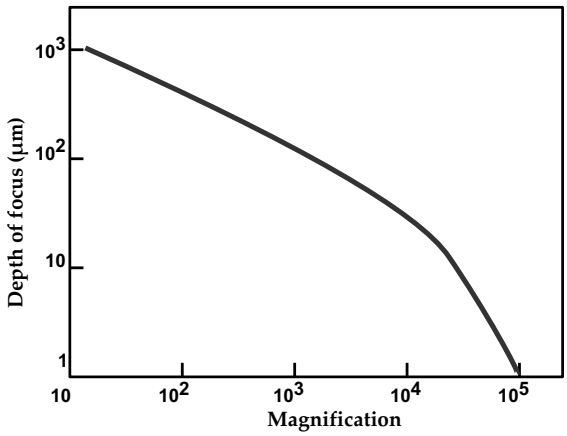
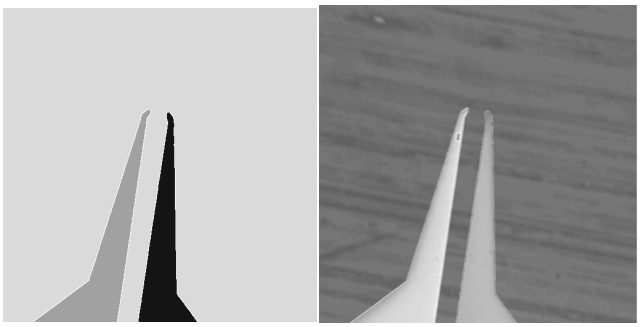



Figure 4.23: Relationship between the magnification and depth of focus in SEM.

### 4.3.2 Experimental shape reconstruction

In this process, the first step is to perform autofocusing on the global scene in order to acquire a sharp image to compute the ROIs. Unlike for depth estimation where a global ROI has been used for each part, multiple ROIs are determined to increase the accuracy of reconstruction. For this task, the ROIs are computed by applying the watershed transformation [GW08]. Figure 4.24 shows the computed ROI demonstrating the segmented gripper fingers from platform surface. Next, multiple images are acquired by varying the electronic working distance. Figure 4.25 shows two images out of a series of 6 images acquired with a varying working distance of 300 $\mu m$. The acquired images are indexed with their corresponding focus step. Then, for each pixel point $(u, v)$ in the ROI for each image, normalized variance has been computed in a local neighborhood of size $n \times n$ using equation (4.25). In this work, a neighborhood of size $5 \times 5$ has been used.



(a) (b)

Figure 4.24: Segmented regions of the (a) gripper fingers and (b) platform surface formed after applying watershed transformation.

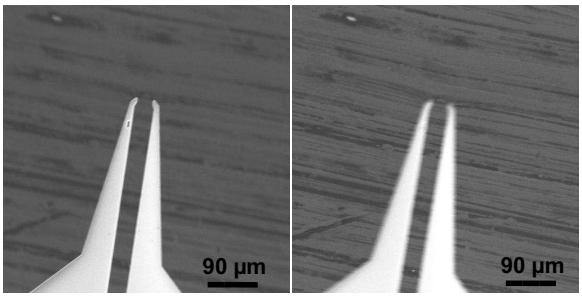



Figure 4.25: Images acquired for reconstruction with varying working distance.

$$\mathbf{S}_{local}(K) = \frac{1}{n^2}\frac{1}{\bar{\mu}} \sum_{u=u-\left(\frac{n-1}{2}\right)}^{u+\left(\frac{n+1}{2}\right)} \sum_{v=v-\left(\frac{n-1}{2}\right)}^{v+\left(\frac{n+1}{2}\right)} (I(u,v) - \bar{\mu})^2 \tag{4.25}$$

where, $K = 1 \ldots k$ number of images, $\bar{\mu}$ is the local neighborhood pixels mean, $I$ is the grey level intensity of the pixel.

Now, the image in which the $\mathbf{S}_{local}$ is maximum for the current point is determined. The height of the point is then computed using the index (focus step) of the particular image from equation (4.3). As the entire process of reconstruction is depended on the sharpness information, the scene to be reconstructed should contain good textures. Figure 4.26 shows the initially obtained depth map.

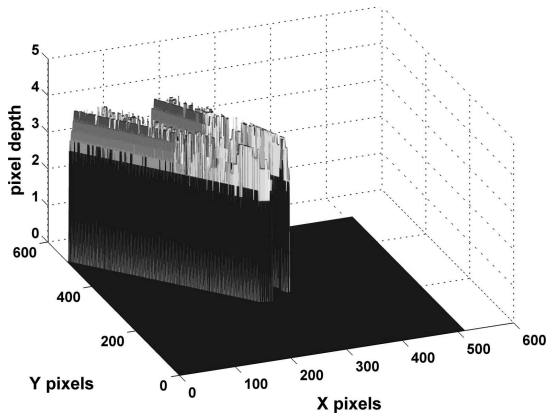



Figure 4.26: Initially estimated depth map for the regions containing gripper fingers.

From the initial results, it has been observed that some pixel points lead to erroneous heights. This is due to the quality of images and used step between the working distances. These heights are processed by local thresholding for each region. The threshold value is chosen to be the index value possessed by maximum number of pixels in the specific region. An alternative solution is by using a parable fit for sample of variances as described by Niederoest et al. [NNŠ03b]. The obtained depth map is approximated using the `gridfit` program provided by D'Errico in Matlab® [D'E05]. A sample matlab code snippet using `gridfit` and visualizing the final output is shown in listing 4.1. The total

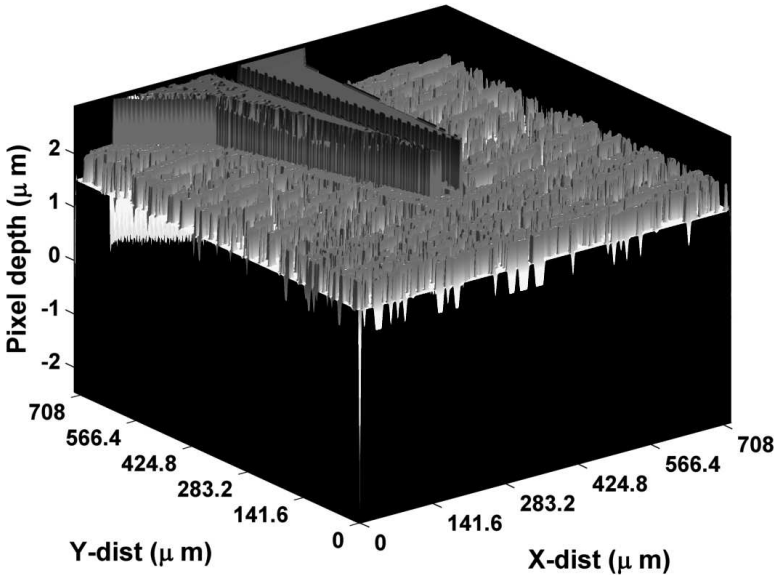
process of shape reconstruction from focus is summarized in algorithm 2. Figures 4.27(a) and 4.27(b) show the reconstructed images formed after surface approximation. Figures 4.28(a) and 4.28(b) show the images formed after overlaying the original texture.

```
x = X;  % x data vector
y=  Y;  % y data vector
z = Z;  % z data vector
%  x and y vectors defining the nodes in the grid
gx=0:1:512;
gy=0:1:512;
% Option 'tilesie' is for simplifying larger grid
[zg,xg,yg] = gridfit(x,y,z,gx,gy,'tilesize',4,'overlap',0.5);
figure;
surf(xg,yg,zg);
shading interp;
camlight right
lighting phong
```

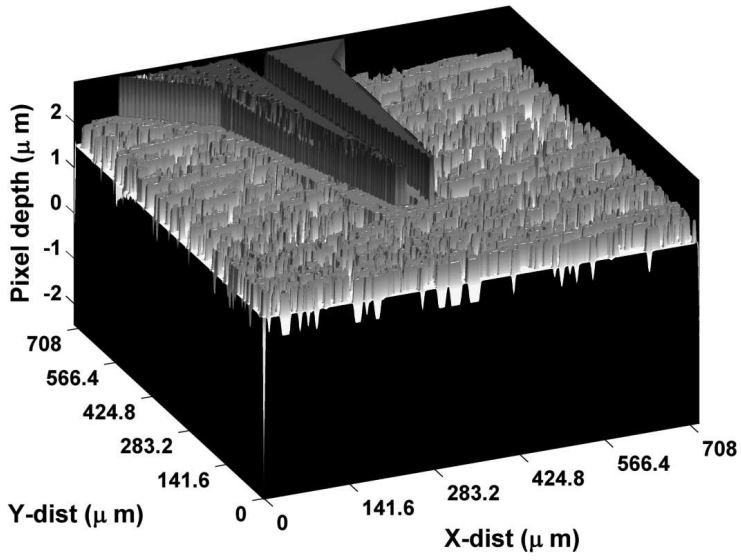
Listing 4.1: Matlab code snippet demonstarting the usage of `gridfit`

**Algorithm 2** Shape reconstruction from focus in SEM.

Perform autofocus on global scene and get working distance $Z$
Compute multiple ROIs
**for** $Z = Z - 1.46 : 0.29 : Z + 1.46$ **do**
    Acquire image and index it with the corresponding focus step
    Save the image in stack
**end for**
Select the neighborhood size, $n$
**for** ROI = 1: total number of ROI **do**
    **for** $K$=1: total number of images **do**
        **for** $u$=1: image width **do**
            **for** $v$=1:image height **do**
                Compute $\mathbf{S}_{local}$ of each pixel $(u, v)$ in ROI using (4.25)
                Get the image index with maximum $\mathbf{S}_{local}$
                Assign $Z$ using (4.3)
            **end for**
        **end for**
    **end for**
**end for**
Get a reconstructed image adding all ROI pixel depths
Perform median filtering, thresholding and surface interpolation
Get final reconstructed image




(a)

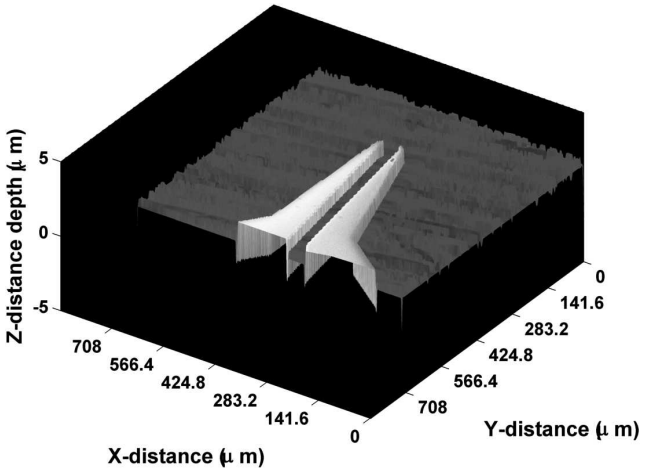



(b)

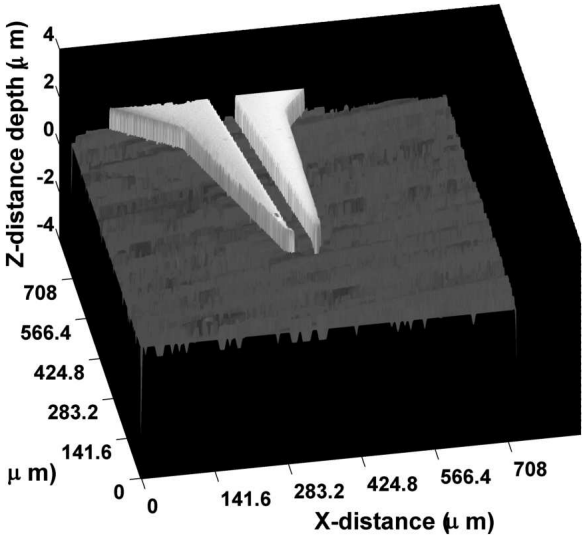
Figure 4.27: Reconstructed images formed after surface approximation.

The developed method is also validated using a gold on carbon specimen. Figure 4.29 shows the sample stub for JEOL SEM and the region pointed by red rectangle is used for reconstruction.

Figures 4.30(a) and 4.30(b) show the initial depth map and the image formed after surface interpolation respectively. Figure 4.31 shows the texture overlaid image.

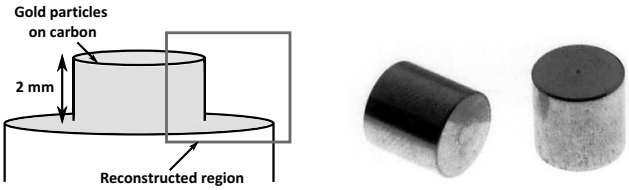



(a)



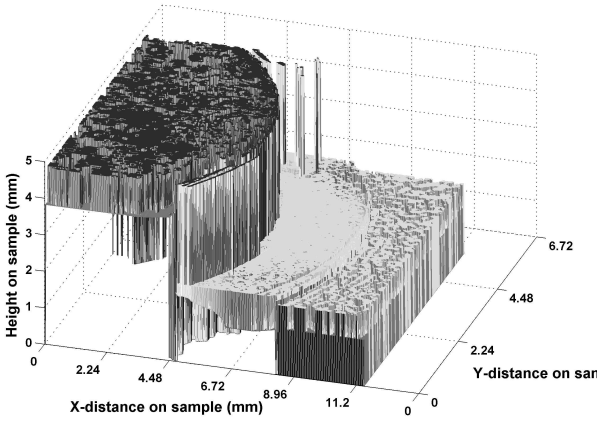

(b)

Figure 4.28: Reconstructed images formed after overlaying the original texture.

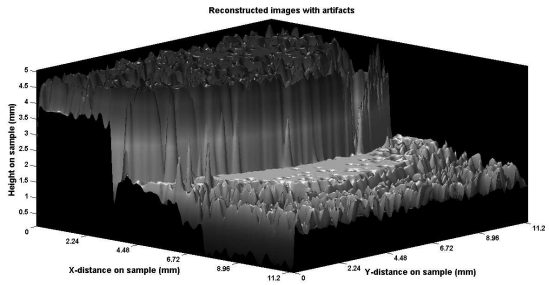



(a) (b)

Figure 4.29: Gold on carbon sample stub for JEOL SEM.




(a)




(b)

Figure 4.30: (a) Initially obtained depth map for gold on carbon sample. Reconstructed images formed after (b)surface approximation.

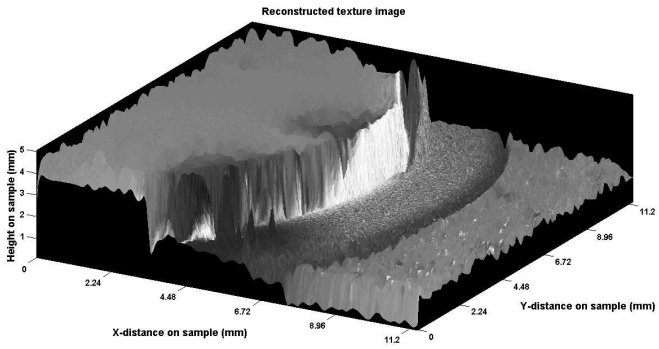



Figure 4.31: Reconstructed gold on carbon stub image formed after overlaying the original texture.

## 4.4 Conclusion

Throughout this chapter, different methods that are developed around the image focus information are presented. At the first hand, a new, direct and robust visual servoing-based autofocusing method for SEM has been presented. For this purpose, different sharpness measures are evaluated and from the obtained evaluation results normalized variance has been selected as the sharpness measure in designing the control law. Unlike the traditional methods, the developed method performs autofocusing directly by controlling the size of the focus step using an adaptive gain. Based on the sharpness score, the value of the gain is adjusted such that a best focus position is ensured. The secondary task function used in the control law precisely stops the overall process, when a best focus position is reached. This method has been integrated with the instrument's software control that has been specially developed at FEMTO-ST and validated using the JEOL SEM at different operating conditions. The operating conditions include varying brightness, scan speed and sample. The obtained experimental results at these conditions prove the robustness of developed approach. Moreover, it has an ability to skip the local maxima that affect the focusing process. The obtained experimental results from the tests performed in comparison with the existing search-based approach prove the fastness and precision of the developed method in finding the best focus.

Inter-object depth estimation is a crucial and challenging task when performing a robotic micromanipulation task in SEM. In this work, a method to estimate the inter-object depth using image focus has been presented. It is a direct application of the previously presented autofocusing technique and estimates the inter-object depth by finding the point of maximum sharpness separately for each object present in the scene. The results obtained with JEOL SEM for a scene containing microgripper and a silicon microstructure are convincing: fast and accurate. Finally, a method to reconstruct the 3D shape of micro-objects present in the scene is presented. It is performed by acquiring multiple images by varying working distance and by estimating the depth of each pixel point in its local neighborhood. It has been validated on the scenes containing microgripper and gold on carbon stub and the reconstructed results are presented. The accuracy of reconstruction has not been tested due to the limitations associated with the current system.

The future works are devoted to perform an automatic micro-nanomanipulation task using the developed autofocusing and depth estimation techniques. Before performing an automatic task, it is also planned to validate all the developed approaches under a more advanced FEG SEM that also includes the improvement of the approaches in terms of accuracy.

Chapter 5

# Automatic nanopositioning in SEM

*This chapter discusses vision-based automatic nanopositioning in SEM . Two visual servoing approaches that are developed to control the 2 degrees of freedom piezo-positioning platform are presented. The first approach uses the total pixel intensities present in an image for designing the control law. In this case, the positioning error and the platform control are directly linked with the intensity variation. The second approach is a frequency domain method that uses Fourier transform to compute the relative motion between images. In this case, the control law is designed to minimize the error i.e. the 2D motion between current and desired frames by controlling the platform movement. Both the methods are validated at different experimental conditions for a task of positioning silicon microparts using piezo-positioning platform.*

## 5.1 Overview of micro-nanopositioning

Over the last decade, nanorobotics has gained significant attention due to the advances in nanoscience and nanotechnology. It has a wide variety of applications in industry, scientific and biomedical domains. It integrates various disciplines that includes nanofabrication and nanomanipulation. Nanofabrication is used to generate different components that are used to produce nanoactuators, nanosensors etc. Nanomanipulation on the other hand, is used to manipulate different micro-nanoobjects like CNTs (carbon nanotubes) [FAD03, FNLE09], InGaAs nano springs [HHB+09], TEM samples [LBK12] and silicon nanowires [ZXQ+09] in order to perform dynamic analysis and characterization of their structural, mechanical, electrical or optical properties [Bon93]. Moreover, manipulation of nanometric objects also benefit in developing and building complex NEMS [AMH+03]. The consequence of this strong demand is the emergence of an active research area concerning the development of assembly and handling technologies on the micro-nanoscale.

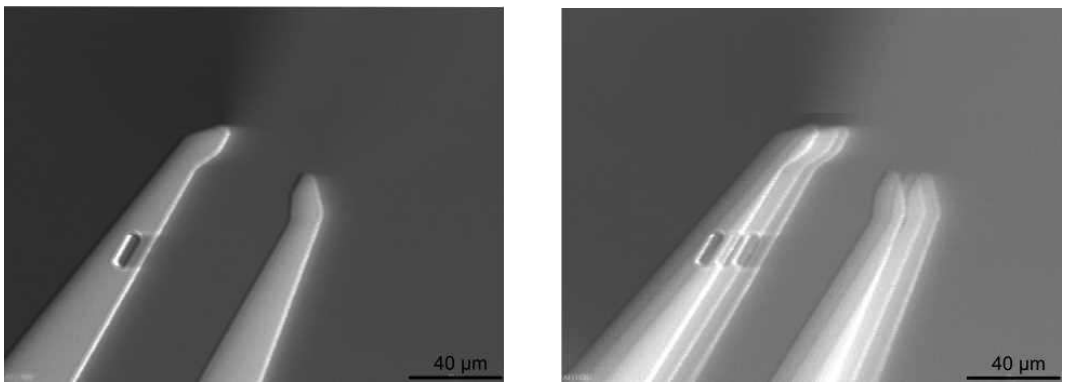
Until now, great progress has already been realized in the development of microassembly stations [YGN01,WP04,PVKN06,ACC+13], microgrippers [BBN+06,AHGP09], precise manipulation systems [PCL+04] as well as the robust control strategies [TDF+08, Tam09,TLFPD10]. In order to perform nanomanipulation, the basic tasks include identification, positioning and handling of the objects. Out of all, positioning is more challenging at this particular scale mainly due to the lack of accurate feedback information. As mentioned earlier in chapter 2, this problem can be resolved by considering microscopic imaging systems such as SEM, TEM or AFM as the underlying sensors to control the behavior of robotic devices during the process. Generally optical microscopes are employed for this task. However, they are limited by their small depth of field and resolution and are not suitable at this particular scale. With its ability of producing images with high resolution at high magnifications in real-time, a SEM is always favored as an imaging tool for automatic nanomanipulation applications. Even though, TEM and AFM provide images with high resolution, they are limited by their working range.

So far, most of the nanopositioning tasks using a SEM are performed by skilled operators who use a joystick device to control the positioning platform movement [WSJ05a]. However, due to the presence of uncertainties and instabilities in the SEM electron column, it is always a difficult and tiresome task to perform. So, it is necessary to develop a system that can automatically handle the positioning problems. When we think of an automatic positioning system, the first thing that can be used is the microscope image information. This process of controlling a robotic device using vision information is a well-known technique called as *visual servoing* [HHC96]. The concept of visual servoing is explained in detail in the next section. In this way, the positioning errors can be minimized and the task can be well repeated.

In general, most of the visual servoing control strategies are based on minimizing the error between current features and reference features observed in the images [CH06]. These features are local geometric contents (like edges, corners etc.) of an object or specially used fiducial markers and are extracted from visual tracking. This tracking process plays a vital role in the development of visual servoing strategies. Many re-

cent works have used cross-correlation-based template matching and active contours for tracking micro-objects in semi-automated tasks [SF06, FWH+07]. The active contours approach uses energy minimizing splines to track an object's shape in the image. Even though, both the methods show convincing results with the noisy SEM images, they are more or less computationally complex. Apart from these 2D tracking approaches, recent works have used 3D model-based tracking. Kratochvil [Kra08, KDN09] and Tamadazte et al. [TMDLFP10] have proposed different methods for tracking the micro-objects using 3D CAD (computer-aided design) models [YN05]. To use these methods, a previously defined or known model is required. As full frame acquisition in SEM is a time taking process, Kratochvil in his work has used small ROI scans for tracking. These ROIs are selected based on an analysis of the model.

However, with SEM as an imaging tool, acquiring images to use them with visual servoing is always a challenging task. This is mainly due to the addition of huge amount of noise during image acquisition process at higher scan rates. Moreover, the pixels are acquired one at a time slowing down the overall acquisition rate. Apart from this, when the objects are in motion, the images contain multiple occurrences of the object due to the sequential raster scanning of the surface (see figure 5.1). This phenomenon mainly raises the difficulty in applying any visual tracking algorithms in real time. In order to tackle this problem, we opted to develop a tracker-less nanopositioning in this work. Many recent developments have shown that the tracking process for visual servoing can be completely omitted and replaced by the global image information like pixel grey level intensities also called as *photometric visual servoing* [TLFPM10, CM11], image entropies [DM09] or image gradient [MC10] etc. Apart from these methods, tracking can also be replaced by using a non-geometric image measurements as explained by Benhimane and Malis [BM07]. For this case, parameters obtained from the 2D motion model are used.




(a) (b)

Figure 5.1: Images of the microgripper acquired when (a) stable (b) moving.

In this work, by considering the direct approach of using the global image information, two methods have been implemented for automatic positioning of the silicon

micro-objects by controlling the motion of piezo-positioning platform. The first one uses the global pixel intensity values and the second one is a frequency domain technique based on 2D motion estimation. Both the developed methods are evaluated at different experimental conditions such as varying scan speed and magnifications. Before explaining the methods, the details regarding the concept of visual servoing are presented in the next section.

## 5.2 Basic visual servoing methods

As mentioned earlier, visual servoing[1] is a closed-loop control mechanism to control a robotic device using vision feedback information. Since the last couple of decades, it has been studied extensively and used in many applications starting from simple pick and place tasks to the present day's complex manipulations. The initially developed systems are based on a simple *look and move* configuration (see figure 5.2). In the terms of manipulation, these systems work in two phases: tracking and control. The former phase is used to continuously track and update the position information of the environment or object being viewed and the later phase is used to control the robot position based on the information provided by tracking phase. These two phases are repeated continuously until the overall task is accomplished.

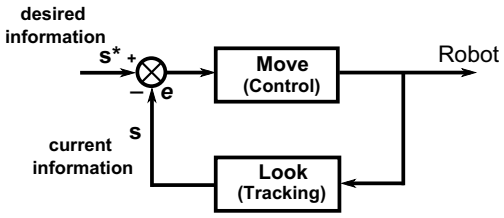



Figure 5.2: Classical look and move visual servoing.

In general, the underlying idea behind this technique is to compute the relationship between image and robot spaces to estimate and control the robot joint velocities. The image information used to generate a control can be either a two dimensional (in image coordinate system) or a three dimensional pose (in the robot base or camera coordinate systems). Based on this information, the visual servoing strategies are mainly classified into the following three types:

- Position-based visual servoing
- Image-based visual servoing
- 2 1/2 D visual servoing

All these types are explained below.

[1]The term *visual servoing* has been first introduced in the literature by Hill and Park in 1979 [HP79].

### 5.2.1 Basic idea of visual servoing

According to Chaumette and Hutchinson [CH06], the visual servoing task can be seen as a task function and its main aim is to minimize the error $\mathbf{e}(t)$ given by equation (5.1).

$$\mathbf{e}(t) = \mathbf{s}(\mathbf{m}(t)) - \mathbf{s}^* \tag{5.1}$$

where, $\mathbf{s}$ and $\mathbf{s}^*$ being the current and reference features, $\mathbf{m}(t)$ is a vector containing the set of measurements. Now to design a control scheme, a relation between the time variations of $\mathbf{s}$ i.e. $\dot{\mathbf{s}}$ and camera instantaneous velocities $\mathbf{v}$ is required. Here $\mathbf{v} = (v, \omega)$ where, $v$ and $\omega$ are respectively the linear and angular velocities of the camera frame. Now the relationship between $\dot{\mathbf{s}}$ and $\mathbf{v}$ is given by

$$\dot{\mathbf{s}} = \mathbf{L_s}\mathbf{v} \tag{5.2}$$

where $\mathbf{L_s}$ is the *image Jacobian* or the *interaction matrix* that links both parameters. Now, using equations (5.1) and (5.2) the time variation of error is

$$\dot{\mathbf{e}} = \mathbf{L_s}\mathbf{v} \tag{5.3}$$

The final control law that ensure an exponential decrease of the error is given by

$$\mathbf{v} = -\lambda \mathbf{L_s}^{\dagger}\mathbf{e} \tag{5.4}$$

where, $\mathbf{L_s}^{\dagger}$ is *Moore-Penrose pseudo inverse* of $\mathbf{L_s}$ and $\lambda$ is the positive gain.

Normally with real time systems, it is impossible to compute a perfect $\mathbf{L_s}$. So in this case, an estimate must be realized and the final control law now takes the form given by equation (5.5).

$$\mathbf{v} = -\lambda \widehat{\mathbf{L_s}^{\dagger}}\mathbf{e} \tag{5.5}$$

where, $\widehat{\mathbf{L_s}}$ is an estimate for $\mathbf{L_s}$.

### 5.2.2 Position-based visual servoing (PBVS)

Usually, PBVS Position-based visual servoing is referred to as 3D visual servoing since the image measurements are used in computing the 3D pose of an object being tracked with respect to the camera or world coordinate systems [WSN87, TMCG02]. More specifically, in the context of a rigid body positioning task, the considered current and desired features are expressed as 3D poses (see figure 5.3). The Cartesian error to be minimized is the difference between these poses. The main advantage associated with this method is that the control is performed directly in the Cartesian space.

Since the pose of an object is estimated in the camera frame, the corresponding transformations between the camera and the robot frames have to be known before in order to generate the control. For computing the pose of an object using image measurements, it requires the camera intrinsic parameters[2] along with the 3D model

[2]Intrinsic parameters of a camera can be obtained from camera calibration.

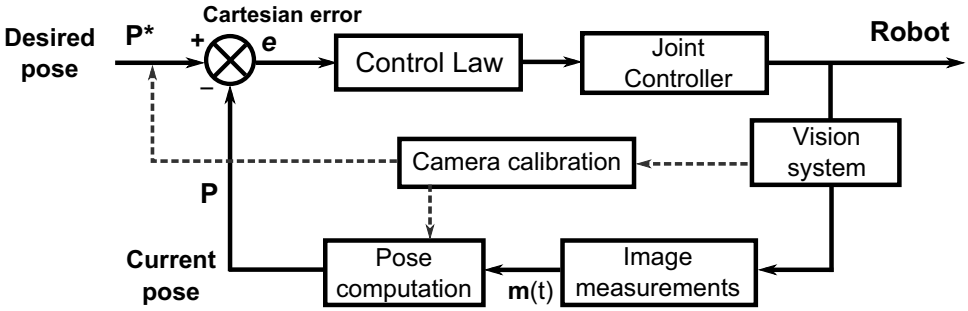



Figure 5.3: Block diagram depicting PBVS.

of the object. So, for the task of accurate positioning, it mainly requires a precise calibration. Furthermore, the image processing algorithms to compute the object pose have to be more sophisticated. Adding to this, as the control is not generated in the images, the possibility of missing the target in camera's field of view is high. Due to these reasons, many researchers have opted to use 2D visual servoing over PBVS.

### 5.2.3 Image-based visual servoing (IBVS)

IBVS is usually referred to as 2D visual servoing, since it uses the 2D visual features that are extracted locally from the images for error minimization [WSN87,FM89]. These features can be any local geometric contents (like edges, corners etc.) of an object or specially used fiducial markers and are extracted from visual tracking. The reference features $\mathbf{s}^*$ are obtained by teaching approach where the robot is moved to a desired position first and the features at this location are recorded. Inturn, these features are compared with the current features to generate an error value. For each new image, the error values are updates simultaneously.The control block diagram for IBVS is shown in figure 5.4.

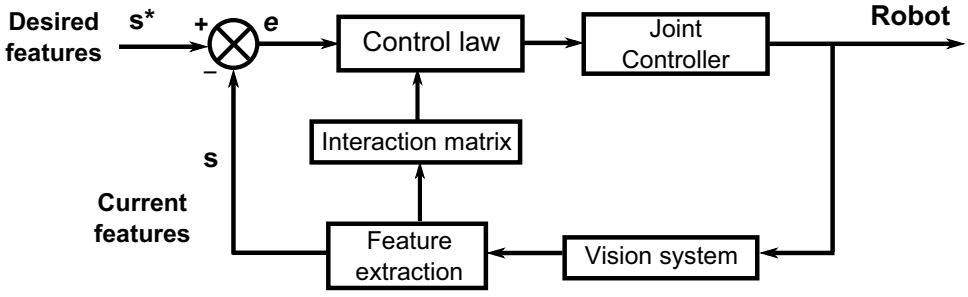



Figure 5.4: Block diagram depicting IBVS.

For IBVS, the image measurements $\mathbf{m}$ are usually the pixel coordinates of a set of points. This method is mainly involved in the computation of interaction matrix that provides a relationship between the scene and camera frames. For a 2D point $\mathbf{p}(x, y)$ in the image frame that is a projection of the 3D point $\mathbf{P}(X, Y, Z)$ in the camera frame,

the interaction matrix $\mathbf{L_p}$ associated with $\mathbf{p}$ is given by

$$\mathbf{L_p}(x, y, Z) = \begin{bmatrix} \frac{-1}{Z} & 0 & \frac{x}{Z} & xy & -(1+x^2) & y \\ 0 & \frac{-1}{Z} & \frac{y}{Z} & (1+y^2) & -xy & -x \end{bmatrix} \tag{5.6}$$

In $\mathbf{L_p}$, $Z$ is the depth of the point relative to camera frame and need to approximated when a single camera is used. Instead of approximation, the depth values can be computed using the method proposed by Papanikolopoulous and Khosla for adaptive depth estimation [PK93]. However, variable depths may lead to the local minima and unstable behavior of the robot [Cha98].

Now, for an eye-to-hand configuration, the control law to control the robot using joint velocities $\mathbf{q}$ is given by equation (5.7) [DK10].

$$\dot{\mathbf{q}} = -\lambda \mathbf{J_s}^{\dagger} \mathbf{e} \tag{5.7}$$

where, $\mathbf{J_s}$ is the features Jacobian given by

$$\mathbf{J_s} = -\mathbf{L_s}\, {}^{c}\mathbf{V}_{\mathcal{F}}\, {}^{\mathcal{F}}\mathbf{J}_n \tag{5.8}$$

where, ${}^{c}\mathbf{V}_{\mathcal{F}}$ is the transformation matrix from camera frame $R_c$ to the robot frame $R_{\mathcal{F}}$. This matrix can be estimated using the camera calibration methods [Tsa87]. ${}^{\mathcal{F}}\mathbf{J}_n$ is the robot Jacobian matrix expressed in the robot reference frame $R_{\mathcal{F}}$.

Despite of requiring the camera calibration, IBVS is known to be robust to the robot and camera calibration errors. The only difference when using a precise calibration and coarse calibration is that with precise calibration convergence rate is faster.

### 5.2.4 2 1/2 D Visual servoing

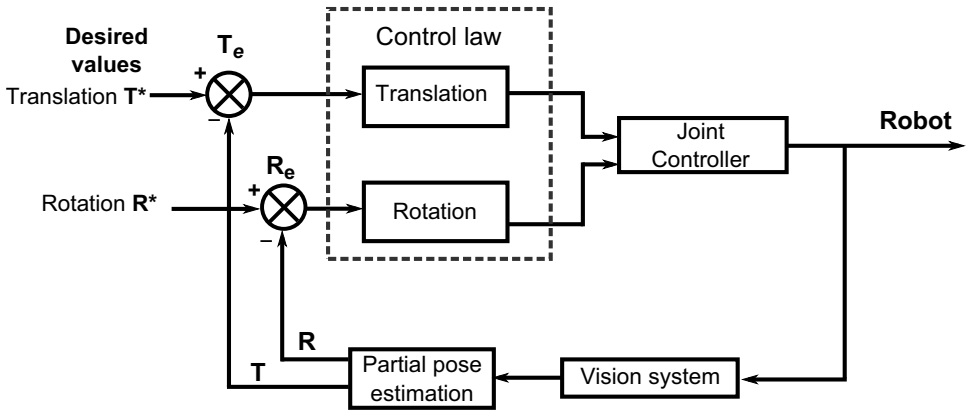



Figure 5.5: Block diagram depicting 2 1/2 D visual servoing [MCB99].

This new concept of 2 1/2 D visual servoing also called as hybrid visual servoing has been first introduced by Malis et al. [MCB99]. The block diagram of this schema

as provided by Malis et al. is shown in the figure 5.5. This method is actually a combination of both the two methods explained above and possess some good advantages. In comparison to PBVS, it does not require any 3D model. This method is based on estimating the rotation and translation of the camera using current and desired images. Since the rotation and translation is controlled by two independent control loops, it ensures the convergence of the control law in the entire task space if the camera intrinsic parameters are known.

## 5.3 Task description and geometrical modelling of positioning stage

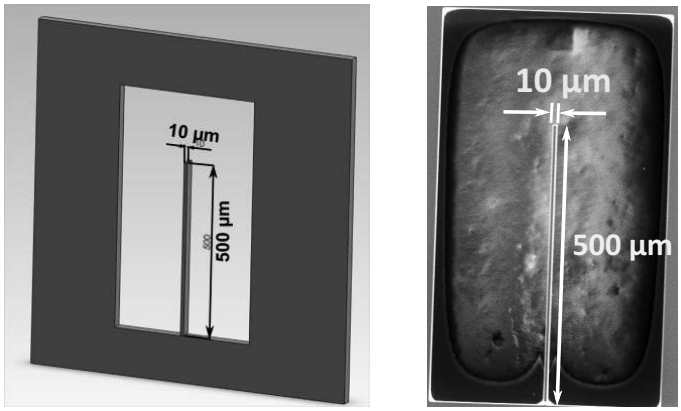
### 5.3.1 Experimental positioning task

The nanopositioning studied in this work is for two objectives. The primary one is to perform characterization of microstuctures by probing to measure the structure stiffness. The second one is to perform in situ manipulation of micro-nanostructures by handling for lift-out from their support. For this study, the silicon microparts (see figures 5.6(a) and 5.6(b)) of dimensions 10 $\mu m \times 100\ \mu m \times 20\ \mu m$ are considered. The parts are placed on the piezo-positioning stage and the main objective is to position the parts automatically in the desired location by controlling the platform's movement. For demonstration, the desired location has been selected by an operator using the specially designed GUI. Once the desired location is selected, the platform moves back to its initial position $(-19\,V, -19\,V)$ from where the servoing starts. Different experimental conditions that include varying scan speed and magnifications are tested for both methods. For all experiments, the SEM secondary electron images of size $512 \times 512$ are used for developing the vision-based control. The acceleration voltage used to generate the beam is 15 $kV$. The control laws for both the methods are programmed in C++, where the intensity-based method uses ViSP (visual servoing platform) [MSC05, Pro13] library and Fourier-based method uses FTTW3 library [FJ05, BR10].

### 5.3.2 Voltage-displacement model of piezo positioning platform

Recall the hardware setup explained in chapter 2, a 3 degrees of freedom $(X, Y, Z)$ open loop piezo positioning platform (TRITOR 100 from Piezosystem Jena GmbH) has been used for this work. The positioning platform is mounted inside SEM vacuum chamber and is controlled using a 3 channel piezo controller NV 40/3. The maximum possible motion on all axes is up to 100 $\mu m$ with a resolution of 0.2 $nm$.

In general, the displacement provided by the positioning platform is a result of the voltage supplied to the piezo actuator. As the control laws explained in the next section can only provide the displacements, it is necessary to compute the relationship between displacement and input voltage. Apart from that, it is a well-known fact that the piezoelectric materials exhibit strong nonlinear hysteresis. So, it is also required to compensate this effect beforehand. In order to accomplish this, the experiments are performed by increasing the voltage from initial value $(-19\ V)$ to maximum value $(110\ V)$




(a) (b)

Figure 5.6: Silicon microparts used for the experiments.

with a step change of 1 $V$ and then decreasing back to the initial value. The respective displacements are measured using a laser interferometer. The tests are conducted for each axis ($x$ and $y$) separately and the obtained hysteresis curves are shown in figure 5.7(a) and 5.7(b). Note that, even though $z$ can be controlled up to 100 $\mu m$ it is not used in this work since this movement is less than the depth of field at the used magnification and cannot be traced. Also, since the variation in $z - axis$ varies image focus, a method to control the focus is already explained in previous chapter. The ascending and descending curves for each axis are approximated by fourth order polynomial given by (5.9), (5.10), (5.11) and (5.12) respectively. The coefficients of the polynomial are estimated by least squares fit. The estimated coefficients for this work are summarized in tables 5.1 and 5.2. Finally the platform control is performed using the displacement vector $\mathbf{d}_{(x,y)} = [dx\ dy]^\top$ to supply voltages $[V_x\ V_y]^\top$ for the piezo-actuator (figure 5.8).

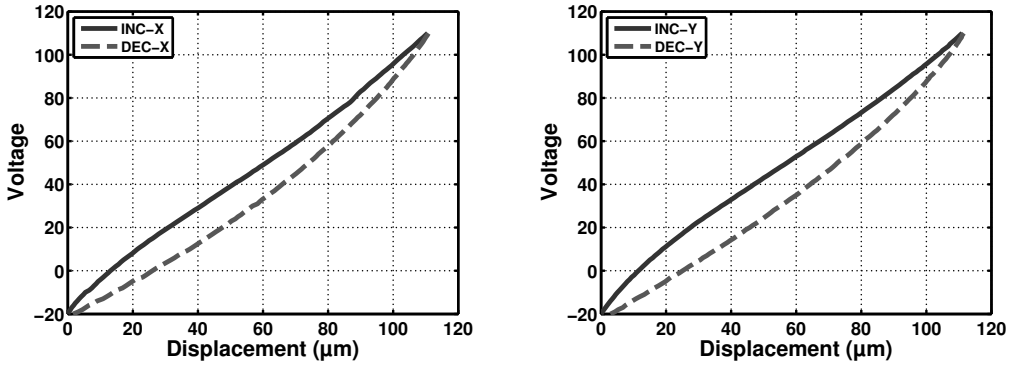
$$V_{x-inc} = a_{1i}d_x^4 + a_{2i}d_x^3 + a_{3i}d_x^2 + a_{4i}d_x + a_{5i} \text{ if } \dot{x} > 0 \tag{5.9}$$

$$V_{x-dec} = a_{1d}d_x^4 + a_{2d}d_x^3 + a_{3d}d_x^2 + a_{4d}d_x + a_{5d} \text{ if } \dot{x} < 0 \tag{5.10}$$

$$V_{y-inc} = b_{1i}d_y^4 + b_{2i}d_y^3 + b_{3i}d_y^2 + b_{4i}d_y + b_{5i} \text{ if } \dot{y} > 0 \tag{5.11}$$

$$V_{y-dec} = b_{1d}d_y^4 + b_{2d}d_y^3 + b_{3d}d_y^2 + b_{4d}d_y + b_{5d} \text{ if } \dot{y} > 0 \tag{5.12}$$

where, $V_x, V_y$ and $d_x, d_y$ are the input voltages and displacements for $x$ and $y$ axes respectively, $a_{ki}, a_{kd}, b_{ki}$ and $b_{kd}$ are polynomial coefficients for increasing and decreasing curves, $\dot{x}$ and $\dot{y}$ are the change in displacements.




(a) (b)

Figure 5.7: Hysteresis curves for (a) $x$-axis (b) $y$-axis.

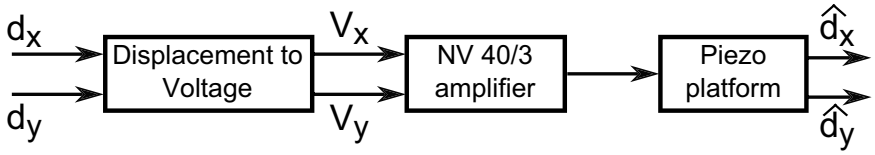



Figure 5.8: Voltage-displacement control model of the piezo-positioning platform.

Table 5.1: Computed polynomial coefficients for $x$-axis.

| k | $a_{ki}$ (Inc-$x$) | $a_{kd}$ (Dec-$x$) |
|---|---|---|
| 1 | -5.3343 $\times 10^{-7}$ | 2.6118$\times 10^{-7}$ |
| 2 | 0.00017528 | -2.2086$\times 10^{-5}$ |
| 3 | -0.017184 | 0.0032187 |
| 4 | 1.6457 | 0.73526 |
| 5 | -19.018 | -21.086 |

Table 5.2: Computed polynomial coefficients for $y$-axis.

| k | $b_{ki}$ (Inc-$y$) | $b_{kd}$ (Dec-$y$) |
|---|---|---|
| 1 | -5.6939$\times 10^{-7}$ | 4.9101$\times 10^{-7}$ |
| 2 | 0.00018878 | -7.3072$\times 10^{-5}$ |
| 3 | -0.020198 | 0.0056593 |
| 4 | 1.8527 | 0.77264 |
| 5 | -19.171 | -22.073 |

## 5.4 Nanopositioning using intensity-based visual servoing

In this section, a control law to perform automatic nanopositioning in SEM using the pixel intensity values is explained. First the derivation of control law is explained followed by the experimental validations.

### 5.4.1 Intensity-based visual servoing

This method is based on considering the intensity values of all the pixels present in an image as visual features i.e. $\mathbf{s} = \mathbf{I}$. By considering all the pixels in an image $f(x, y)$ of size $M \times N$, the visual feature is given by

$$\mathbf{s} = \mathbf{I} = \left(I_{(1,1)}, I_{(1,2)} ... I_{(M,N)}\right)^\top \tag{5.13}$$

where, $I_{(u,v)}$ is the intensity of a pixel at location $(u, v)$ and $\mathbf{I}$ is a column vector of size $M \times N$. In this case, the error is

$$\mathbf{e} = \mathbf{I} - \mathbf{I}^* \tag{5.14}$$

If we consider the problem of error minimization as an optimization problem, the primary goal will be to minimize the cost $C$ given by equation (5.15).

$$C = \mathbf{e}^\top \mathbf{e} = \left(\mathbf{I} - \mathbf{I}^*\right)^\top \left(\mathbf{I} - \mathbf{I}^*\right) \tag{5.15}$$

The goal Image $\mathbf{I}^*$ is defined by teaching approach where the platform is moved to a desired position and the image is acquired. Once, the target is defined, the main goal of the controller is to regulate the cost function from an unknown position. When $C$ is minimum, the current position corresponds to the desired position. To visually reflect this cost function, a set of images are acquired by moving the platform around the target position. The cost is computed using these images offline and is shown in the figure 5.9. From the figure, the cost becomes minimum at the desired position; however, the convergence is not smooth. The main reason is that the intensity variations in SEM imaging are not constant.

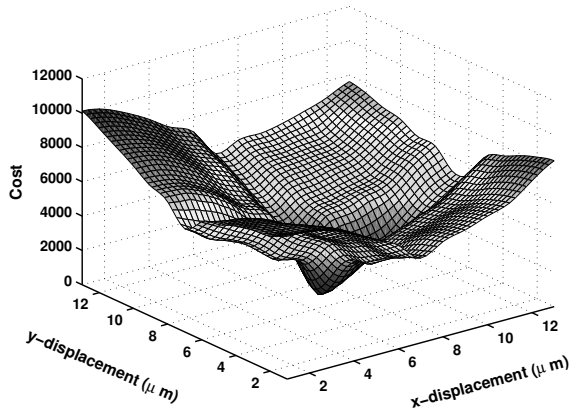



Figure 5.9: Visual representation of the cost function for intensity-based visual servoing.

Now to control the motion of the positioning stage, an estimation of the interaction matrix is required to link the temporal variation of pixel intensities with the camera instantaneous velocities. As shown by Marchand [Mar07], it can be derived by considering the *optical flow constraint equation* (OFCE) provided by Horn and Schunck [HS81].

According to OFCE, the intensity of a pixel $I(x, y, t)$ in an image is same before and after a small displacement $(dx, dy)$ for a time interval $dt$. We then have

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \tag{5.16}$$

The first order Taylor expansion of (5.16) gives (5.17).

$$\frac{\partial I}{\partial x}dx + \frac{\partial I}{\partial y}dy + \frac{\partial I}{\partial t}dt = 0 \tag{5.17}$$

On dividing (5.17) by $dt$, we get:

$$\nabla I_x \dot{x} + \nabla I_y \dot{y} + \dot{I} = 0 \tag{5.18}$$

where, $\nabla I_x = \frac{\partial I}{\partial x}$ and $\nabla I_y = \frac{\partial I}{\partial y}$.

Now, if we consider the interaction matrix related to a 2D point $m(x, y)$ (projection of 3D point $M(X, Y, Z)$) in the image which is given by equation (5.19).

$$\mathbf{L}_{(x,y)} = \begin{bmatrix} -\frac{1}{Z} & 0 & \frac{x}{Z} & xy & -(1+x^2) & y \\ 0 & -\frac{1}{Z} & \frac{y}{Z} & (1+y^2) & xy & -x \end{bmatrix} \tag{5.19}$$

Using $\mathbf{L}_{(x,y)}$, equation (5.2) can be decomposed into (5.20).

$$\dot{x} = \mathbf{L}_x \mathbf{v} \text{ and } \dot{y} = \mathbf{L}_y \mathbf{v} \tag{5.20}$$

By substituting (5.20) in (5.18), we get

$$\dot{\mathbf{I}} = -\left[\nabla \mathbf{I}_x \mathbf{L}_x + \nabla \mathbf{I}_y \mathbf{L}_y\right] \mathbf{v} \tag{5.21}$$

From equation (5.21), the interaction matrix $\mathbf{L}_{\mathbf{I}(x,y)}$ is

$$\mathbf{L}_{\mathbf{I}(x,y)} = -\left[\nabla \mathbf{I}_x \mathbf{L}_x + \nabla \mathbf{I}_y \mathbf{L}_y\right] \tag{5.22}$$

Here, $\mathbf{L}_{\mathbf{I}(x,y)}$ is the interaction matrix of size $1 \times 6$ for one pixel. For entire image, it can be rewritten as

$$\mathbf{L}_{\mathbf{I}} = \begin{bmatrix} \mathbf{L}_{\mathbf{I}(1,1)} \\ . \\ . \\ . \\ \mathbf{L}_{\mathbf{I}(M,N)} \end{bmatrix} \tag{5.23}$$

. Using $\mathbf{L}_{\mathbf{I}}$, the time variation of pixel intensities is given by

$$\dot{\mathbf{I}} = \mathbf{L}_{\mathbf{I}} \mathbf{v} \tag{5.24}$$

The deduced interaction matrix allows us to control up to 6 degrees of freedom. However, the used positioning platform can be controlled only 2 degrees of freedom ($x$ and $y$).

Now, using equations (5.22) and (5.14), the control law given by equation (5.5) can be used. However, it has been shown that, a more feasible solution can be by using a control

law derived of the form Levenberg-Maquardt optimization technique [Mal04, CMC08]. The control law derived of this form can be highly non-linear and improve the robustness [CM11]. The final control law is then given by equation (5.25).

$$\mathbf{v} = -\lambda {}^{c}\mathbf{V}_{p} \left(\mathbf{H} + \mu diag(\mathbf{H})\right)^{-1} \mathbf{L_I}^{\top} \mathbf{e} \tag{5.25}$$

where, $\lambda$ and $\mu$ are positive gains, $\mathbf{H} = \mathbf{L_I}^{\top}\mathbf{L_I}$ is the Hessian matrix and ${}^{c}\mathbf{V}_{p}$ is the transformation matrix from camera frame $\mathcal{R}_c$ to platform frame $\mathcal{R}_p$ and can be obtained from sensor calibration.

As mentioned earlier, the platform can be controlled only with the voltages and a voltage-displacement model has been derived. So the velocities computed from control law are now converted to displacements $\mathbf{d}_{(x,y)}$ of the platform using equation (5.26).

$$\mathbf{d}_{(x,y)} = \mathbf{v}_{avg} t \tag{5.26}$$

where, $\mathbf{v}_{avg} = \frac{\mathbf{v}_0 + \mathbf{v}_{cur}}{2}$ is the average velocity, $\mathbf{v}_0$ is initial velocity, $\mathbf{v}_{cur}$ is the current velocity and $t$ is the sampling time taken.

For each iteration, the displacement vector is updated as given by equation (5.27) and the corresponding voltages computed from equations (5.9), (5.10), (5.11) and (5.12) are used to control the platform.

$$\mathbf{d}_{new} = \mathbf{d}_{cur} + \mathbf{d}_{prev} \tag{5.27}$$

where, $\mathbf{d}_{new}, \mathbf{d}_{cur}$ and $\mathbf{d}_{prev}$ are the updated, current and previous displacements respectively. The block diagram depicting the overall control is shown in figure 5.10.

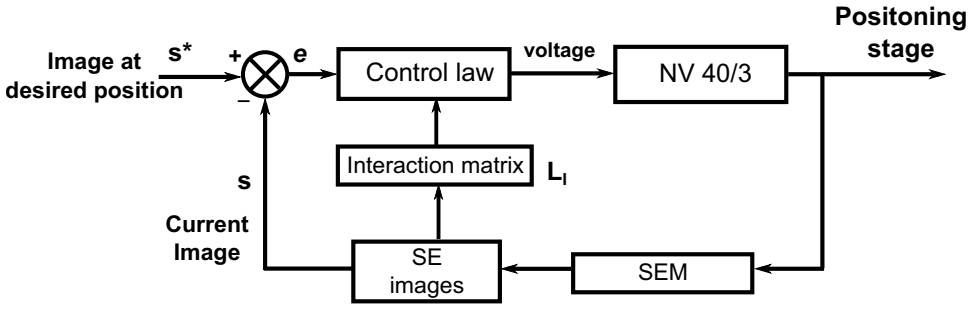



Figure 5.10: Block diagram depicting intensity-based visual servoing.

### 5.4.2 Nanopositioning at optimal scan speed

The initial experiments are performed to position the silicon microparts with a normal scan speed of 720 nanoseconds per pixel providing a frame rate of 2.2 frames per second. The magnification is fixed to 300 ×. For this test, the user selected voltages are $50\,V$ for $x$ channel and $60\,V$ for channel. Figures 5.11(a) and 5.11(b) show respectively the images acquired at desired and initial positions. Figures 5.11(c) to 5.11(f) show error $(\mathbf{I} - \mathbf{I}^{*})$ at different locations. The final error at the end of the positioning task is

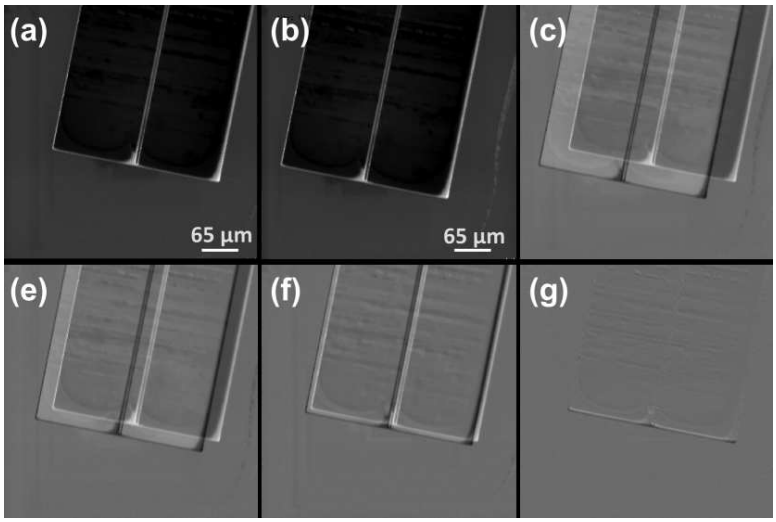



Figure 5.11: Series of images depicting the intensity-based visual servoing at normal scan speed. (a) Represent user selected desired position. (b) Initial image in the process. (c) - (f) Errors at different positions. (g) Final error.

shown in figure 5.11(g). Fig. 5.12(a) and Fig. 5.12 show the displacement and voltage variations with each iteration during the positioning process. Figure 5.12(c) shows the cost variation with each iteration.

### 5.4.3 Nanopositioning at high scan speed

Second tests are performed to position the microparts using the images acquired with an increased raster scanning speed. Normally in SEM imaging, high scanning rates during image acquisition leads to the increased noise levels in images. More details about noise in SEM imaging noise are explained in chapter 3. This test has been performed to check the methods efficiency in reaching the desired position at noisy conditions. For this test a scan speed of 360 nanoseconds (maximum allowed) per pixel that provides a frame rate of 3.1 frames per second has been used. Initially selected voltages for desired location are 50 $V$ and 60 $V$ for $x$ and $y$ channels respectively. Figures 5.13(a) and 5.13(b) show images at desired position and initial location respectively. Figures 5.13(c) and 5.13(d) show the error at first and final iterations during the positioning. Figures 5.14(a), 5.14(b) and 5.14(c) show the variation of displacements, voltages and cost during the process.

### 5.4.4 Nanopositioning at increased magnification

This experiments are performed to position the micropats at high magnification. This task has been performed at a magnification of 800×. Simultaneously, the method is also validated with increased scan speed at the selected magnification. The selected scan time is 360 nanoseconds per pixel. The voltages selected for desired position are 30 $V$

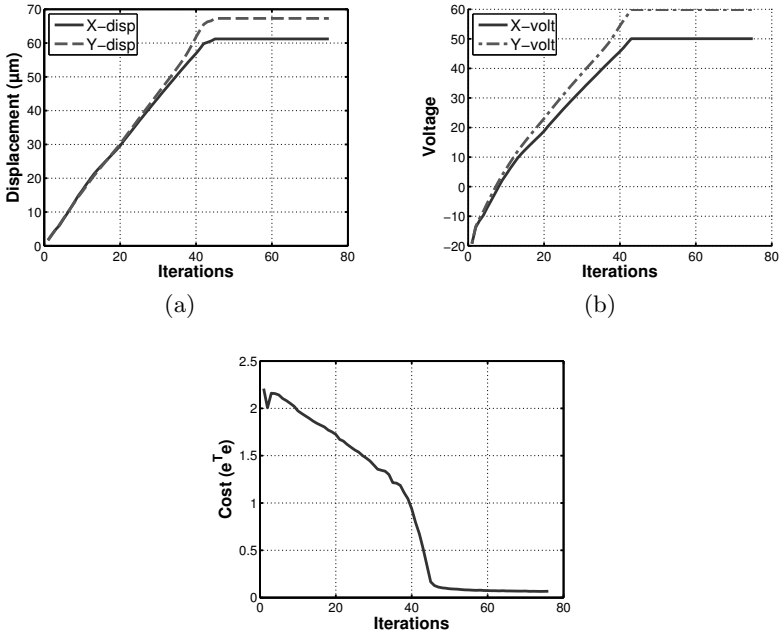



Figure 5.12: (a) Displacement (b) voltage (c) cost variations in $x$ and $y$ axes of the positioning stage with each iteration using optimal scan speed.

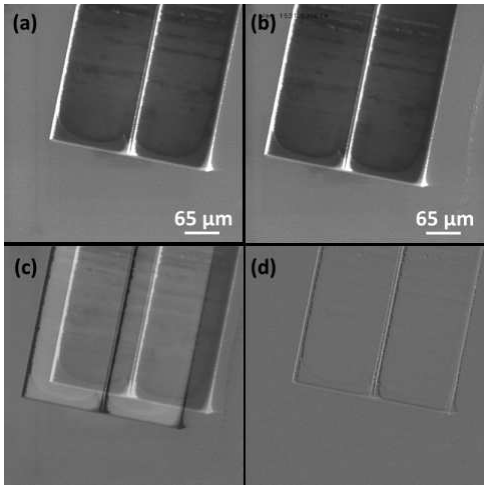



Figure 5.13: (a) Image acquired at desired location. (b) initial image (c) error at initial position (d) error at final position using high scan speed.

and 60 $V$ respectively for $x$ and $y$ axes. Figures 5.15(a) and 5.15(b) show the images at desired position and initial position respectively. Figure 5.15(c) shows the error at initial position and figure 5.15(d) shows the error at final position during the positioning

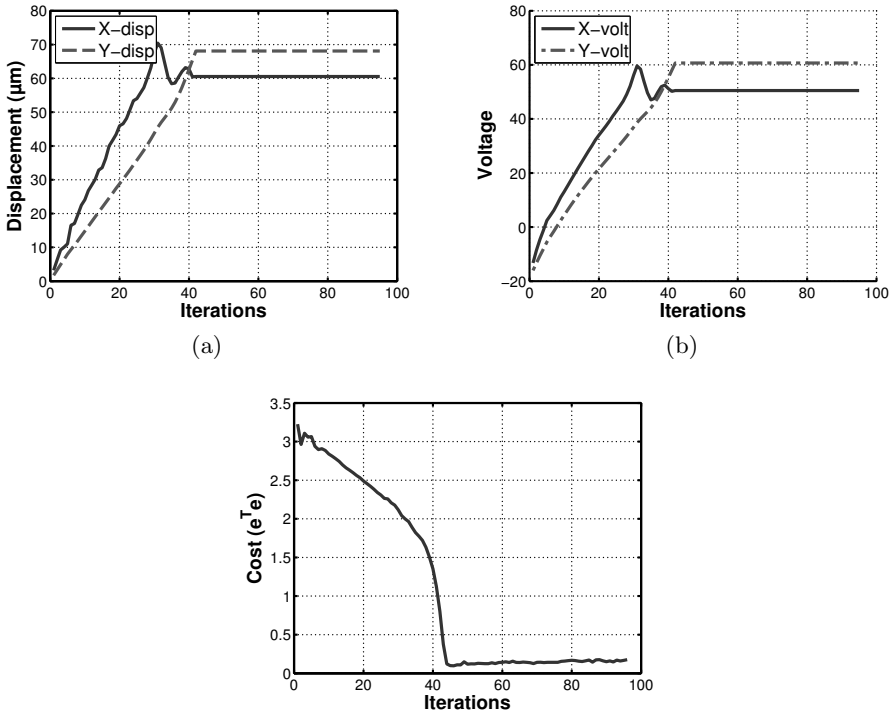



Figure 5.14: (a) Displacement (b) Voltage (c) cost variations during the positioning task using intensity-based visual servoing at high scan speed.

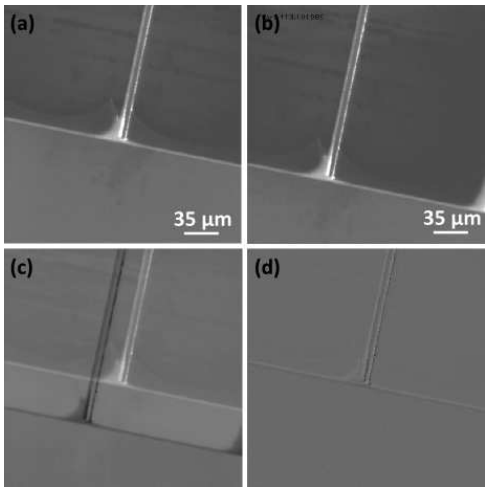



Figure 5.15: Images acquired at (a) desired location (b) initial position. Error images at (c) initial position (d) final position during the Fourier-based visual servoing process at high magnification (800×).

task. The displacement and voltage variation plots are shown in Figures 5.16(a) and 5.16(b) respectively. Figure 5.16(c) shows the cost variation with each iteration at the selected magnification.

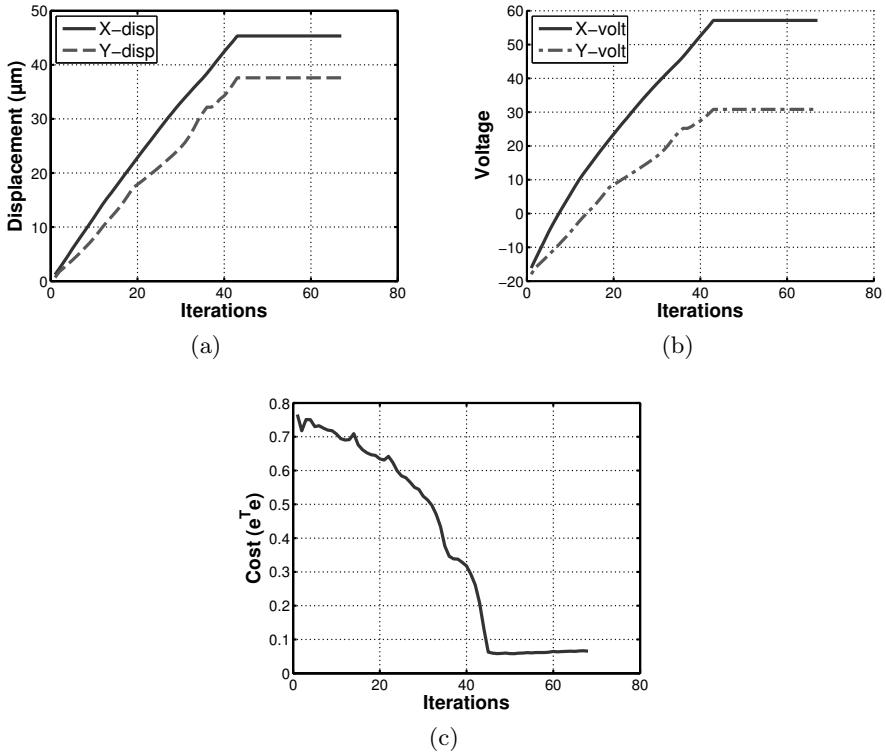



Figure 5.16: (a) Displacement (b) Velocity variations during the positioning task using intensity-based visual servoing at high magnification (800×).

### 5.4.5 Nanopositioning at unstable conditions

The final experiments are conducted to perform the positioning task at unstable conditions i.e. at varying brightness and contrast and to check the method's efficiency. For this test, once the desired position is selected and the positioning has started, the contrast and brightness are varied manually. The selected magnification is 300× and scan speed is 720 nanoseconds per pixel. Eventually, the nanopositioning task failed during this test since it requires constant or less variations in the intensity. Figures 5.17(a) and 5.17(b) show the reference and initial images. Figures 5.17(c) and 5.17(d) show the images where the contrast and brightness values are changed during this test. The displacement, voltage and error variations are shown in figures 5.18(a), 5.18(b) and 5.18(c) respectively. The zero voltage during the test indicates that the voltage is out of the range.

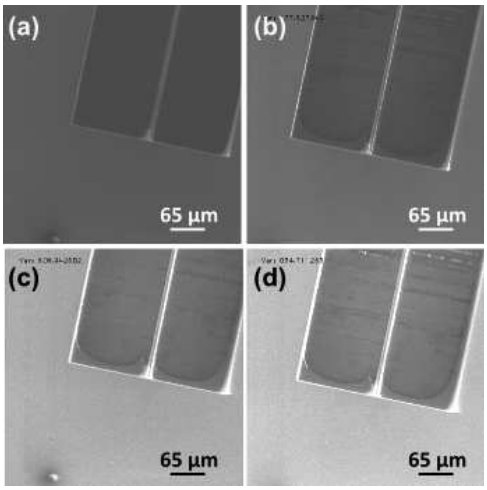
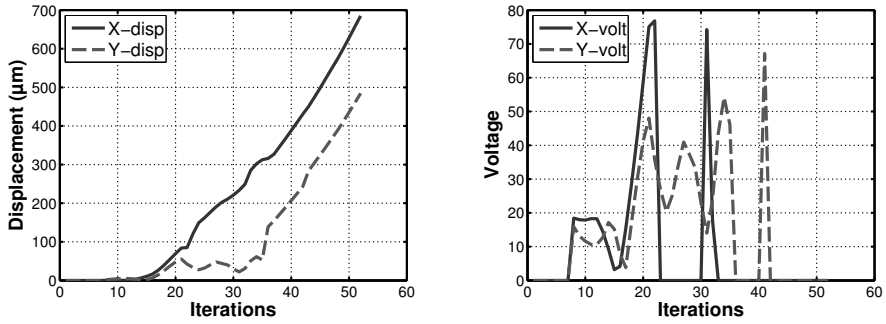



Figure 5.17: Series of images acquired during nanopositioning at unstable conditions using intensity-based method. (a) Desired image. (b) - (d) Images acquired during the process.




(a) (b)

Figure 5.18: (a) Displacement (b) Velocity variations during the positioning task using intensity-based visual servoing at unstable conditions.

## 5.5 Nanopositioning using Fourier-based visual servoing

In this section, another method to perform the nanopositioning task has been discussed. This method is based on estimating the motion between the current image and reference image in Fourier domain. It uses the idea of homography-based visual servoing [BM06]. The main reason to select the Fourier domain is that with SEM imaging, the brightness and contrast are not constant and it is known that Fourier-based motion estimation is robust to intensity variations and image noise. With this method, the control strategy is quite simple and does not require any direct estimation of the interaction matrix. In this section, first different methods to estimate the motion in Fourier domain are explained. Later the control strategy along with the positioning experiments are presented.

### 5.5.1 Fourier-based motion estimation

In image processing, Fourier transformation is an important tool that is used in wide variety of applications such as spectral analysis, image filtering, image compression and image registration [RC96,Pau98,SJ06,SSKS09]. Normally for image data that is discrete, the main basis to perform spectral analysis is by using discrete Fourier transform (DFT) that is a sampled version of the Fourier transformation. It works by decomposing an image in the spatial domain into sine and cosine components in the frequency domain. Each point in the Fourier image represent the corresponding frequency of the points in the spatial domain.

In general, the DFT produce an image with complex numbers that can displayed as two images: magnitude image and phase image. Figures 5.19(b), 5.19(e), 5.19(h), 5.19(c), 5.19(f) and 5.19(i) show the centrally shifted log-magnitude spectrums and phase spectrums of the images shown in figures 5.19(a), 5.19(d) and 5.19(g). It can be seen from the images that the magnitude spectrum is insensitive to image translation, but it rotates by the same angle of a rotated image. The image in figure 5.19(d) is a translated version of figure 5.19(a) and their magnitude spectrum is same, but their phase angles are different. Figure 5.19(f) shows this. In general, most of the information is present in the magnitude spectrum and phase angle images contain little intuitive information. Based on this concept, the motion estimation is performed independently and is explained below.

#### Translation estimation by phase correlation

Let us suppose that $f(x, y)$ and $g(x, y)$ are two images acquired at desired and current locations respectively and are translated by an amount of $(\delta_x, \delta_y)$ in $x$ and $y$ axes. The Fourier transforms $\mathcal{F}(u, v)$ and $\mathcal{G}(u, v)$ for both images are given by (5.28) and 5.29).

$$\mathcal{F}(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi\left\{\frac{ux}{M}+\frac{vy}{N}\right\}} \tag{5.28}$$

$$\mathcal{G}(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} g(x,y)e^{-j2\pi\left\{\frac{ux}{M}+\frac{vy}{N}\right\}} \tag{5.29}$$

where, $M$ and $N$ are image dimensions and $x$ and $y$ are the spatial pixel coordinates.

Now, according to the Fourier shift property

$$\mathcal{G}(u,v) = \mathcal{F}(u,v)e^{-j2\pi\left\{\frac{u}{M}\delta_x+\frac{v}{N}\delta_y\right\}} \tag{5.30}$$

The normalized cross power spectrum $\widehat{C}(u, v)$ is given by (5.31).

$$\widehat{C}(u,v) = \frac{\mathcal{F}(u,v)\overline{\mathcal{G}(u,v)}}{|\,\mathcal{F}(u,v)\overline{\mathcal{G}(u,v)}\,|} = e^{j2\pi\left\{\frac{u}{M}\delta_x+\frac{v}{N}\delta_y\right\}} \tag{5.31}$$

where, $\overline{\mathcal{G}(u,v)}$ is the complex conjugate of $\mathcal{G}(u, v)$.

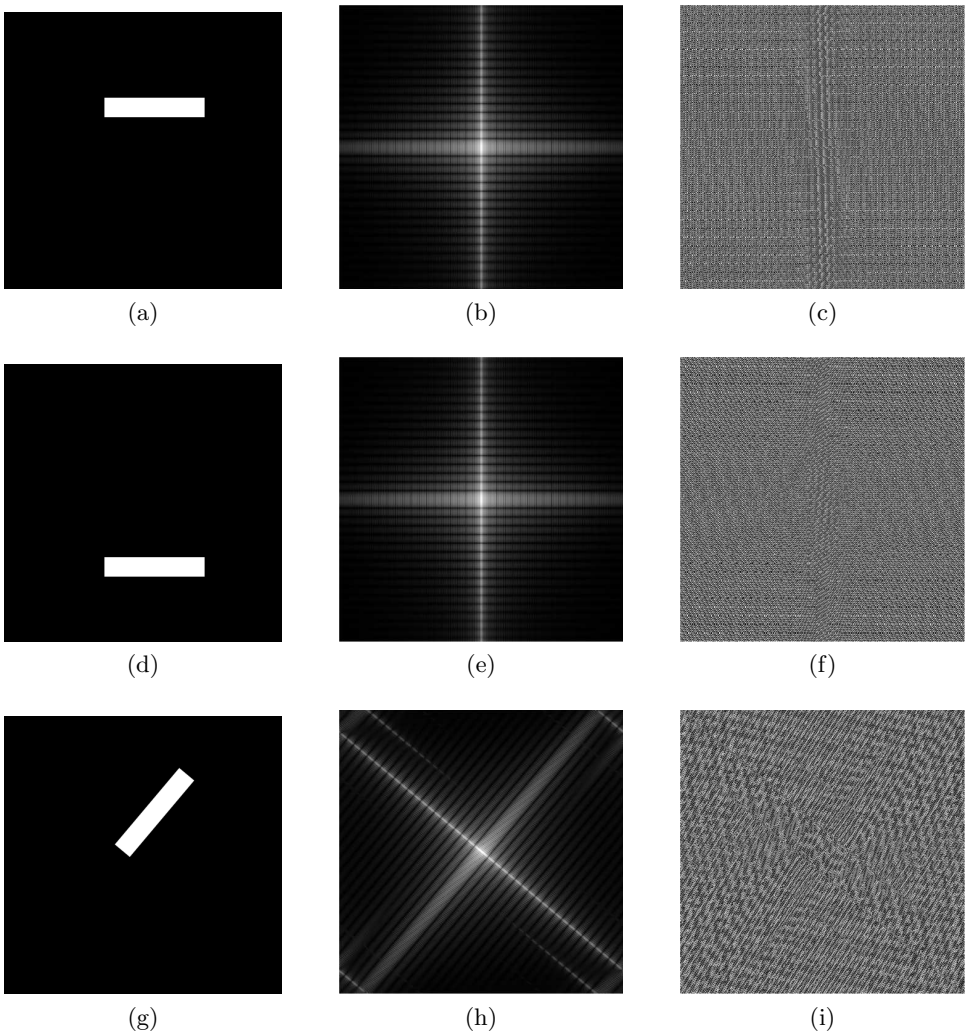


Figure 5.19: (a)(d)(g) Images in spatial domain (b)(e)(h) respective magnitude spectrums (c)(f)(i) respective phase spectrums.

The cross power spectrum is normalized in order to compensate the intensity variations. Now, (5.31) can be solved for overall translation $(\delta_x, \delta_y)$ in two ways. The first one is to solve directly in the Fourier domain using a coordinate system containing two frequency axes and one phase difference axis. The slopes produced by $\frac{u}{M}\delta_x + \frac{v}{N}\delta_y = 0$ provides $(\delta_x, \delta_y)$. However, it is computationally complex. The other method that is used in this work is to find the inverse Fourier transform of (5.31) that results in a Dirac delta function given by (5.32).

$$\mathcal{D}(\delta_x, \delta_y) = \mathcal{F}\mathcal{T}^{-1}(\widehat{C}(u, v)) \tag{5.32}$$

Finally, the 2D translation is computed by the maximum of (5.32).

$$(\tau_x, \tau_y) = argmax\,\{\mathcal{D}(\delta_x, \delta_y)\} \tag{5.33}$$

where, $\tau_x$ and $\tau_y$ are the estimated translations in $x$ and $y$ directions respectively.

**Roto-translation estimation in Fourier domain**

This section provides the information regarding the computation of rotatory motion between images using the Fourier transformation. However, as the used positioning stage does not features any rotation, it is not tested in the real time. The rotation between two images can be extracted by converting the Fourier spectrum to log-polar transformation. If $f_2(x, y)$ is rotated (by $\theta_0$) and translated (by $(\delta_x, \delta_y)$) version of $f_1(x, y)$, then they are related by

$$f_2(x, y) = f_1(x\cos\theta_0 + y\sin\theta_0 + \delta_x,\ -x\sin\theta_0 + y\cos\theta_0 + \delta_y) \tag{5.34}$$

Using Fourier translation and rotation properties

$$\mathcal{F}_2(u, v) = e^{-j2\pi[u\delta_x + v\delta_y]}\mathcal{F}_1(u\cos\theta_0 + v\sin\theta_0, -u\sin\theta_0 + v\cos\theta_0) \tag{5.35}$$

where, $\mathcal{F}_1$ and $\mathcal{F}_2$ are the Fourier transforms of $f_1$ and $f_2$ respectively.

Let $\xi_1$ and $\xi_2$ be the magnitude spectrums of $\mathcal{F}_1$ and $\mathcal{F}_2$ respectively, then

$$\xi_2(u, v) = \xi_1(u\cos\theta_0 + v\sin\theta_0, -u\sin\theta_0 + v\cos\theta_0) \tag{5.36}$$

Now, to compute the rotation, the magnitude spectrums are transformed to log-polar form, then

$$\xi_2(\rho, \theta_0) = \xi_1(\rho, \theta - \theta_0) \tag{5.37}$$

Where, $\rho$ and $\theta$ are the radius and angle in polar coordinates. From the equation (5.37), the following properties are true:

1. Rotation in image rotates the magnitude spectrum by the same angle.
2. Rotation in spatial domain can now be seen as pure translational displacement in polar representation.

Now, by computing the global correlation between $\xi_2(\rho, \theta_0)$ and $\xi_1(\rho, \theta - \theta_0)$ provides a peak in the correlation surface, that corresponds to an estimate $\hat{\theta}$ to the angle of rotation.

**Control scheme**

For positioning the platform, a control scheme has been designed considering the 2D translation and rotation motions computed from the previous steps as the visual features. By using such features, the resulting control can allow us to decouple the control of rotation and translation (in our case only translation is used). The current features are

$$\mathbf{s}(t) = [\mathbf{s}_v, \mathbf{s}_\omega]^\top \tag{5.38}$$

where, $\mathbf{s}_v = [\tau_x\ \tau_y\ 0]^\top$ are the translations in $x$, $y$ and $z$, and $\mathbf{s}_\omega = (\mathbf{u}\theta)$ are the rotation features with $\mathbf{u}$ being the axis and $\theta$ being the angle of rotation. The translation in $z$ is considered to be zero, since it is not used in this work. However, it is considered only to analytically derive a control law for 6 degrees of freedom. The final objective is to drive the platform to the desired location i.e. $\mathbf{s}^* = [\mathbf{0}_{1\times 6}]^\top$. The task function $\mathbf{e}$ to be regulated is then given by

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_v \\ \mathbf{e}_\omega \end{bmatrix} = \mathbf{s} - \mathbf{s}^* = [\tau_x\ \tau_y\ 0\ \mathbf{u}\theta]^\top \tag{5.39}$$

In this case, the shape of the task function that has been computed offline by considering only 2D translation is shown in the figure 5.20. By comparing the shape of cost functions given by both methods, it is clear that the Fourier-based method shows better convergence. It is mainly due to its robustness to intensity variations.

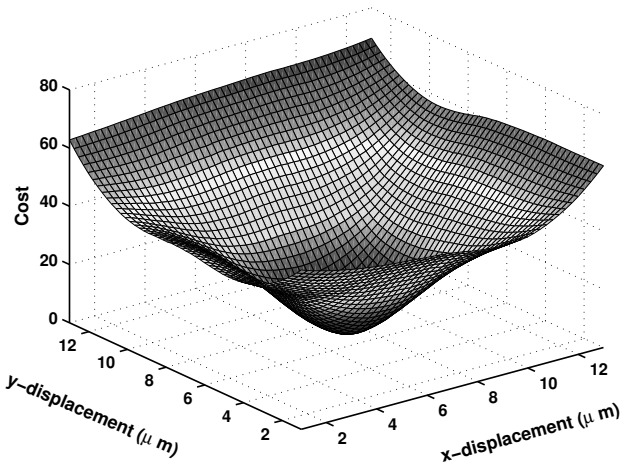



Figure 5.20: Cost function for Fourier-based visual servoing.

Now, the relationship between time variation of visual features $\dot{\mathbf{s}}$ and the camera instantaneous velocity $\mathbf{v} = [\mathbf{v}_v\ \mathbf{v}_\omega]^\top$ is given by

$$\dot{\mathbf{s}} = \begin{bmatrix} \dot{\mathbf{s}}_v \\ \dot{\mathbf{s}}_\omega \end{bmatrix} = \begin{bmatrix} \mathbf{L}_v & \mathbf{0} \\ \mathbf{0} & \mathbf{L}_\omega \end{bmatrix}_{6\times 6} \begin{bmatrix} \mathbf{v}_v \\ \mathbf{v}_\omega \end{bmatrix}_{6\times 1} \tag{5.40}$$

where, $\mathbf{L}_v$ and $\mathbf{L}_\omega$ are the $3\times 3$ interaction matrices to link linear and angular velocities. Here, $\mathbf{L}_v$ is a $3\times 3$ identity matrix and $\mathbf{L}_w$ is a $3\times 3$ matrix given by equation (5.41) [BM07].

$$\mathbf{L}_w = \mathbf{I} - \frac{\sin\theta}{2}[\mathbf{u}]_\times - \sin^2\left(\frac{\theta}{2}\right)(2\mathbf{I} + [\mathbf{u}]_\times^2) \tag{5.41}$$

By considering the exponential convergence of the error i.e. $\dot{\mathbf{e}} = -\lambda \mathbf{e}$, where $\lambda$ is a positive gain value, the control law is given by

$$\begin{bmatrix} \mathbf{v} \\ \omega \end{bmatrix} = \begin{bmatrix} -\lambda_v \mathbf{L}_v & \mathbf{0}_3 \\ \mathbf{0}_3 & -\lambda_\omega \mathbf{L}_\omega \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{e}_v \\ \mathbf{e}_\omega \end{bmatrix} \tag{5.42}$$

where, $\lambda_v$ and $\lambda_\omega$ are positive translational and rotational gains.

In this work, the positioning platform can be controlled only with 2 degrees of freedom (translation in $x$ and $y$), so the final control used is

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = -\lambda_v {}^c\mathbf{V}_p \mathbf{L}_v{}^{-1} \begin{bmatrix} \tau_x \\ \tau_y \end{bmatrix} \tag{5.43}$$

where, $v_x$ and $v_y$ are the platform velocities in $x$ and $y$ directions, respectively. Finally, the displacement vector $\mathbf{d}_{(x,y)}$ is computed using equation (5.26). The control block diagram for nanopositioning using Fourier-based visual servoing is shown in figure the 5.21.

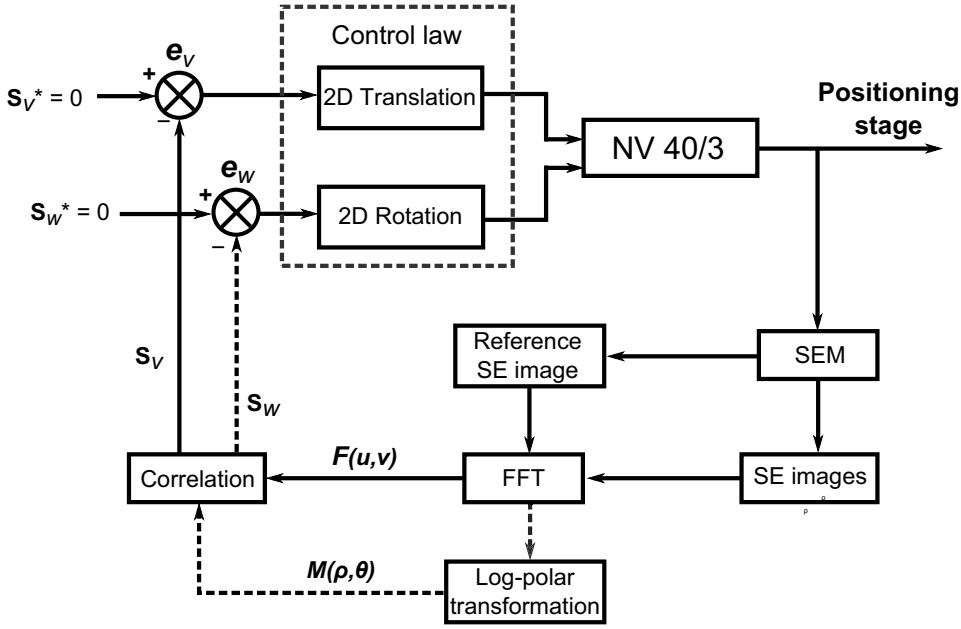



Figure 5.21: Block diagram depicting Fourier-based visual servoing.

Similar to the intensity-based visual servoing, for each iteration, the displacements are updated as given by equation (5.27) and the corresponding voltages computed from equations (5.9), (5.10), (5.11) and (5.12) are used to control the platform.

### 5.5.2 Experimental validation and positioning at optimal scan rates

The initial experiments are performed using this method to perform the positioning task at optimal scan rates. The selected scan speed is 720 nanoseconds per pixel. Similar to

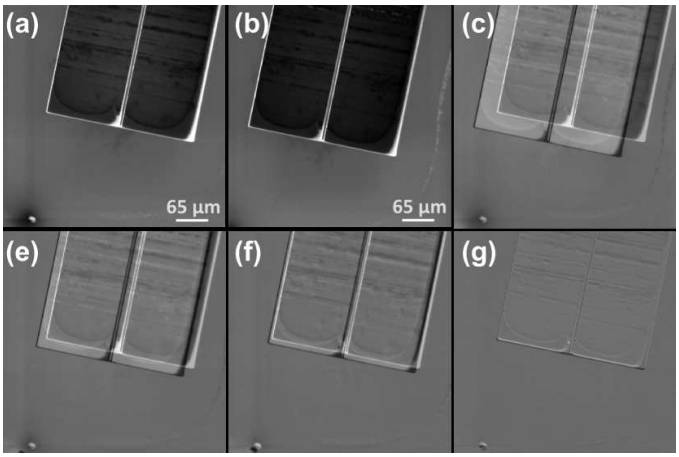



Figure 5.22: Series of images showing the positioning task using Fourier-based visual servoing at normal scan speed. (a) Represent user selected desired position. (b) Initial image in the process. (c) - (f) Errors at different positions. (g) Final error.

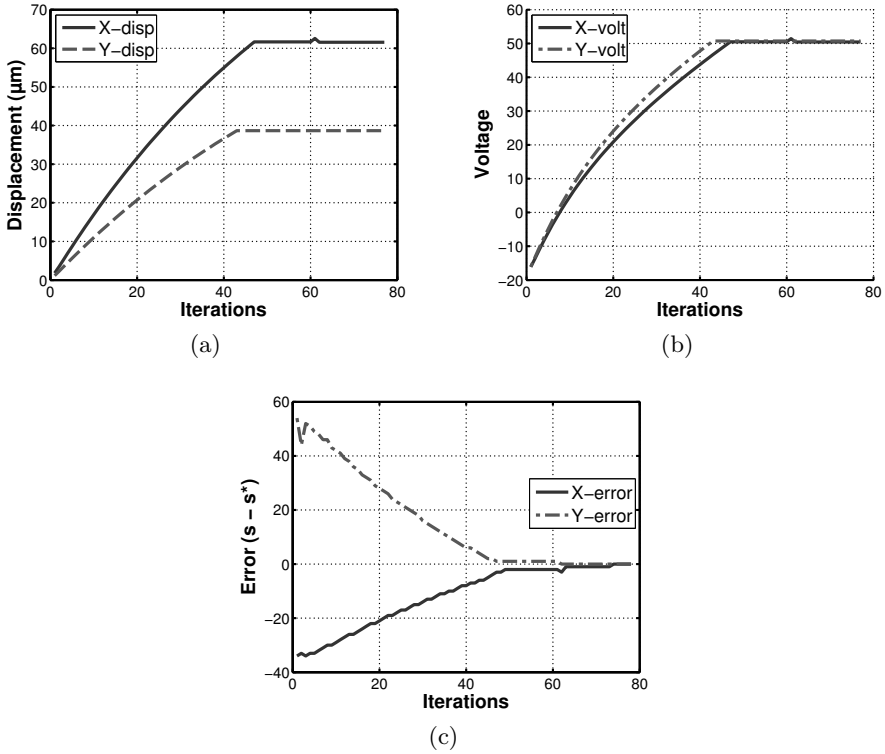



Figure 5.23: (a) Displacement (b) voltage (c) error variations in $x$ and $y$ axes of the positioning stage with each iteration using optimal scan rates.

the intensity-based method, a desired location is selected by the operator by changing the platform voltages using GUI. For this task, a voltage of 50 $V$ has been selected for both the channels and the image acquired at this desired location is shown in figure 5.22(a). Figures 5.22(b) to 5.22(f) show the image acquired at initial position and the errors during the process respectively. Figure 5.22(g) shows the finally achieved position. In this case, subtracted images are used to visually depict the error. Figures 5.23(a), 5.23(b) and 5.23(c) show the displacement, voltage and error variations during the process.

As a comparison of both the methods, it can be seen that both the methods are succeeded in reaching the desired position; however, Fourier-based method shows better accuracy.

### 5.5.3 Experimental validation and positioning at increased scan rate

Second tests are carried out to perform the positioning task using the explained method with an increased raster scanning speed. The selected scan speed for this test is 360 nanoseconds per pixel and the desired location voltages are selected to be 40 $V$ and 50 $V$ respectively for $x$ and $y$ channels. Figure 5.24(a) shows the image acquired at desired location and figure 5.24(b) shows the image at initial location. Figures 5.24(c) and 5.24(d) show the error images at initial and final location during the positioning task.

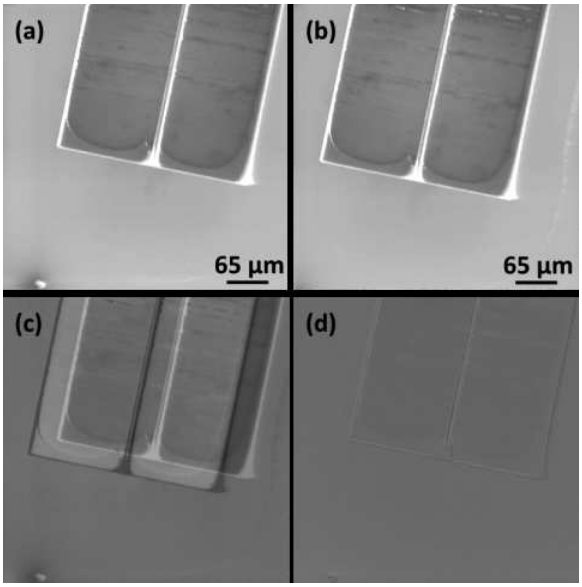



Figure 5.24: Images acquired at (a) desired location (b) initial position. Subtracted images (error) at (c) initial position (d) final position during the Fourier-based visual servoing process with high scan speed.

Figures 5.25(a), 5.25(b) and 5.25(c) show the displacement, voltage and error variation during the process. Despite of having high amount of noise in the images, the

positioning task has been accomplished successfully. However, similar to the previous test, method-2 shows better performance.

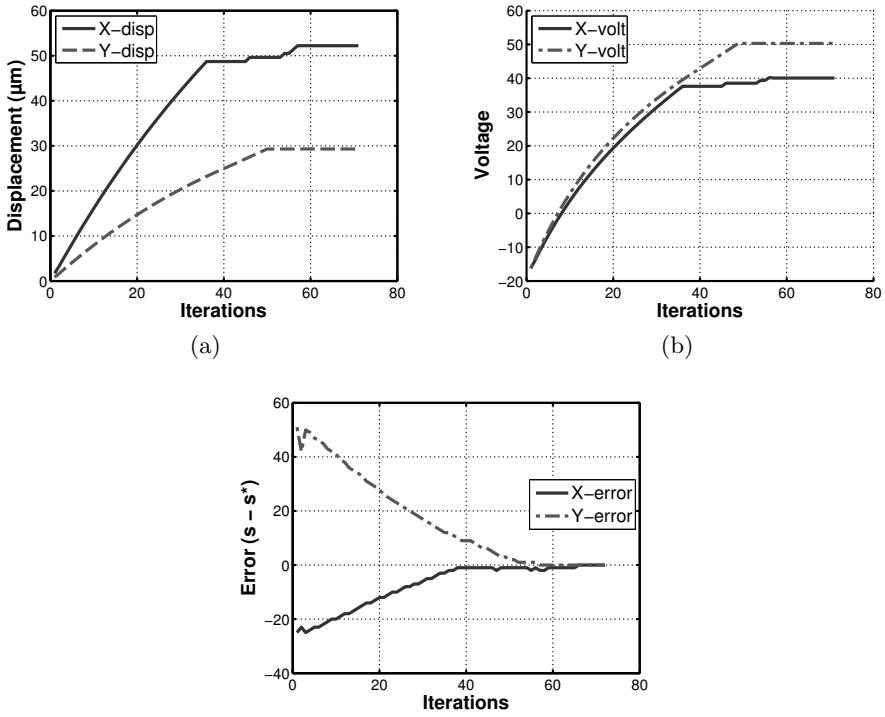



Figure 5.25: (a) Displacement (b) Velocity and (c) error variations during the positioning task using Fourier-based visual servoing at high scan speed.

### 5.5.4 Experimental validation and positioning at high magnification

The final tests are performed to carry out the positioning task at increased magnification. The selected magnification is 800×. Apart from validating at high magnification, the scan speed (360 nanoseconds per pixel) is also set to high. The selected desired location voltages are 40 $V$ and 50 $V$ for $x$ and $y$ channels respectively. Figures 5.26(a) and 5.26(b) show the images acquired at desired location and initial location respectively. Figures 5.26(c) and 5.26(d) show the errors at initial and final positions respectively.

Figures 5.27(a), 5.27(b) and 5.27(c) show the displacement, voltage and error variation during the process. From the obtained results, despite of increased magnification and having high amount of noise, the positioning task has been accomplished successfully.

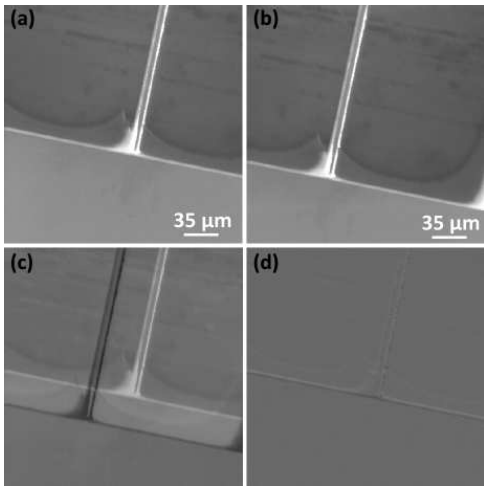



Figure 5.26: Images acquired at (a) desired location (b) initial position. Subtracted images (error) at (c) initial position (d) final position during the Fourier-based visual servoing process at high magnification (800×).

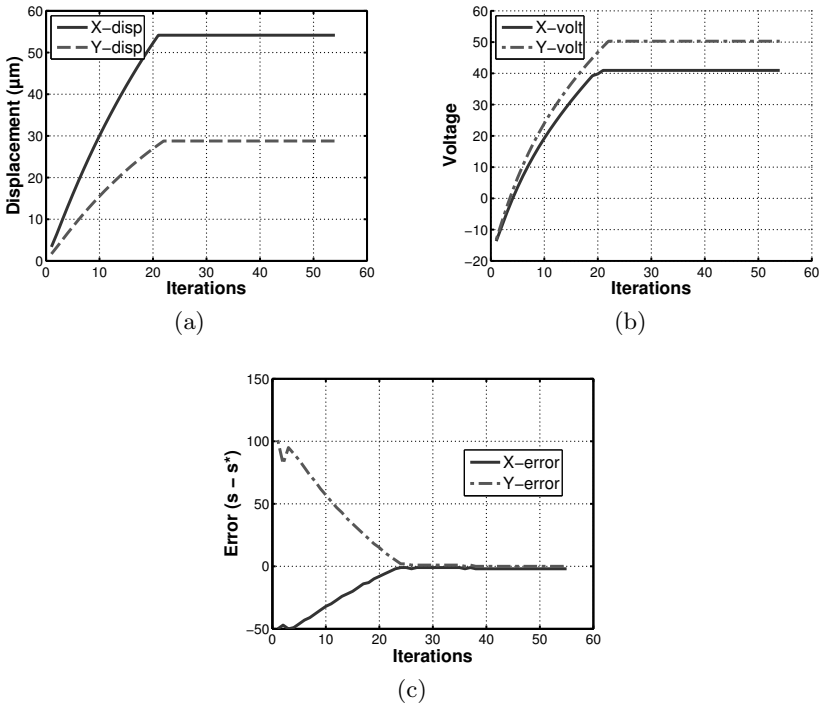



Figure 5.27: (a) Displacement (b) Velocity and (c) error variations during the positioning task using Fourier-based visual servoing at increased magnification (800×).

### 5.5.5 Experimental validation and positioning at unstable conditions

Similar to the intensity-based method, the positioning has been tested with this method at unstable conditions. The voltages at desired location are $40\,V$ and $50\,V$. Figures 5.28(a) and 5.28(b) show the reference and initial image captured during this process respectively. Figures 5.28(d) - 5.28(f) show the error. The displacement, voltage and error variations are shown in figures 5.29(a), 5.29(b) and 5.29(c) respectively. From the obtained results it is clear that, despite of having instability in the imaging conditions, the positioning task has been successfully accomplished. It is mainly due to the robustness of Fourier-based method to image noise and intensity variations.

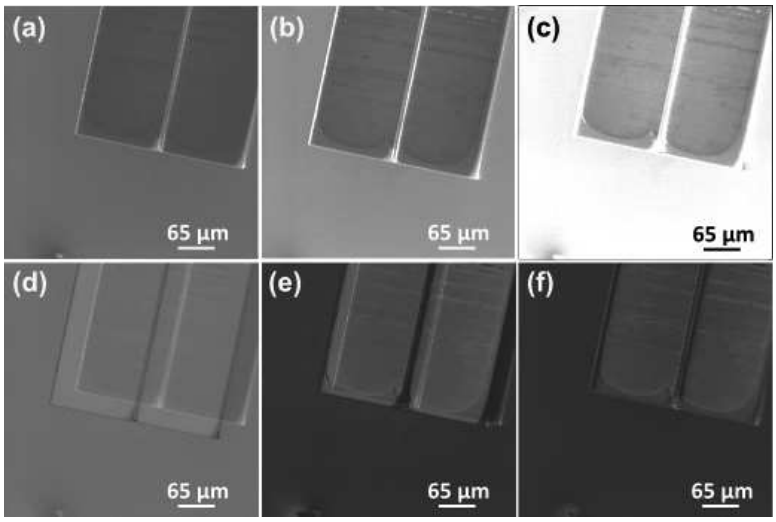



Figure 5.28: Series of images during nanopositioning at unstable conditions using Fourier-based method. (a) Desired image. (b)-(c) Images acquired during the process with varying brightness and contrast. (d) - (e) Errors at different positions. (f) Final error.

## 5.6 Accuracy of positioning and discussion

Since the existing system (SEM) does not allow using any external displacement measuring devices like laser interferometers, in this work, the positioning accuracy has been measured directly from the images. Eventhough, the estimated accuracy is not reliable (due to the present of noise); it has been computed to demonstrate the efficiency of the methods in performing a nanopositioning task. Besides, it depends on the magnification used (from (5.44)) and increases with increase in the magnification. The accuracy is estimated using the final error image by multiplying the number of error pixels with pixel dimension on the sample. The pixel dimension $P$ on the sample is computed using (5.44).

$$P = \frac{D}{G} \quad [\mu m] \tag{5.44}$$

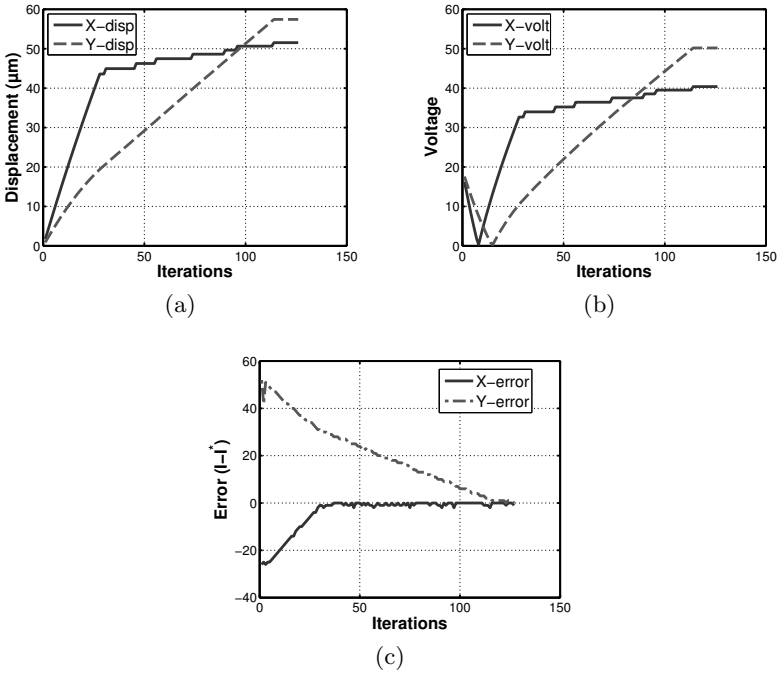



Figure 5.29: (a) Displacement (b) Velocity variations during the positioning task using Fourier-based visual servoing at unstable conditions.

where, $D$ is the pixel dimension on the screen (constant) and $G$ is the magnification. With our system, the computed $D$ value is 212.3 for a screen size of $512 \times 512$ pixels. The accuracies computed with different tests demonstrated above are summarized in table 5.3.

Table 5.3: Estimated positioning accuracy ($\mu m$) achieved by both methods.

| Condition | Intensity method | | Fourier method | |
|---|---|---|---|---|
| | x ($\mu m$) | y ($\mu m$) | x ($\mu m$) | y ($\mu m$) |
| Optimal speed | 0.707 | 1.414 | 0.707 | 0.707 |
| High scan rate | 1.061 | 2.123 | 0.707 | 0.707 |
| High magnification | 0.398 | 0.5307 | 0.265 | 0.265 |

From the obtained results, it can be seen that both the methods accomplished the positioning task successfully in almost all cases. However, Fourier-based method shows better accuracy in all conditions. Moreover, it performed well in the case of unstable and noisy conditions (variation in brightness and contrast) where the intensity-based method failed in reaching the desired position.

## 5.7 Conclusion

An automatic nanopositioning task of silicon microstructures using SEM has been presented in this chapter. It has been accomplished using two approaches of visual servoing. The first method is a photometric approach where all the gray level intensities of an image are used as visual features. The error variation and platform movement are directly linked with the intensity variation. The second method is based on estimating the 2D motion between images. In this chapter, the motion estimation has been performed using the image frequency information in Fourier domain. Since SEM image acquisition is mostly affected by the additive noise during image acquisition, Fourier-based method provides good robustness and is mostly unaffected by the intensity variation. The main advantage of using these kind of methods is that they do not require any complex image processing and tracking algorithms in designing the control law. Both the methods are validated at different experimental conditions. Even though both methods succeeded in accomplishing the overall task, Fourier-based method shows better behaviour and accuracy.

The future work will concentrate on performing the positioning task using both approaches on another SEM and will try to improve the positioning control in terms of accuracy. Apart from this, it will also attempt to control the positioning of a 3 degrees of freedom microactuator (MM3A Kleindiek) containing a microgripper to perform a complete manipulation task.

Chapter 6

# Software development

*This chapter presents various software modules that have been developed to integrate all the vision and visual servoing methods explained in the previous chapters. The developed software also includes an application GUI that has been used to control the SEM remotely as well as to monitor the running process.*

## 6.1 Software architecture

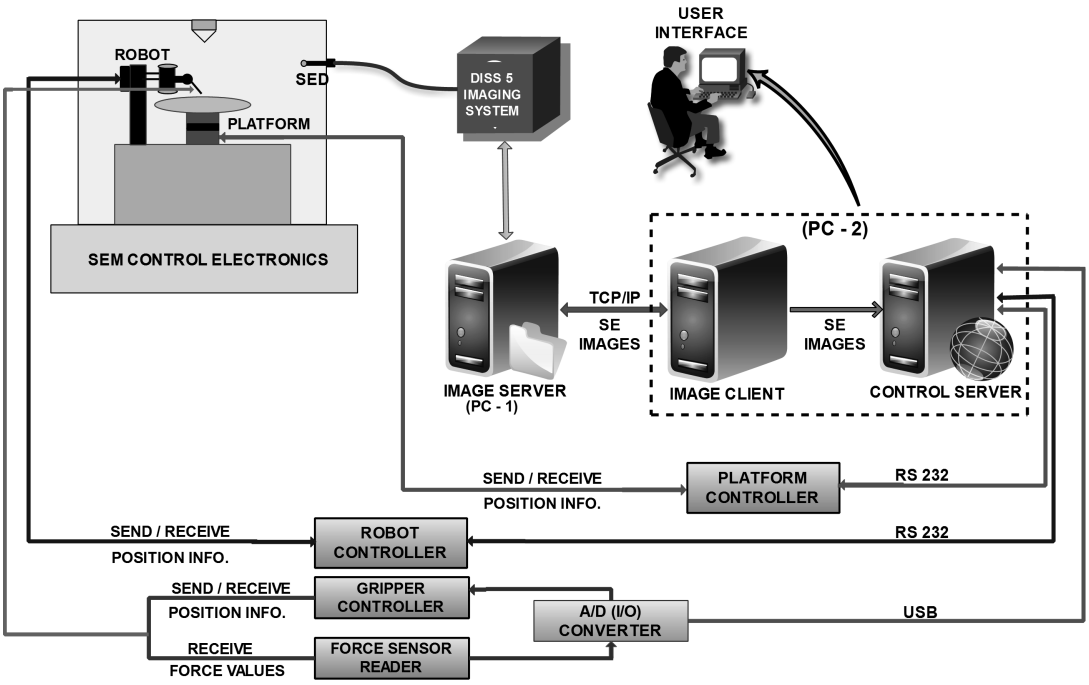



Figure 6.1: Software architecture.

The developed software architecture for synchronous operation of the various hardware components (presented in chapter 2) is shown in the figure 6.1. The communication between the SEM control computer (Intel Pentium 4, CPU 2.24 GHz and 512 MB of RAM) and the work computer (Intel Core 2 Duo, CPU 3.16 GHz, and 3.25 GB of RAM) is accomplished by implementing a cross-platform client-server model using TCP/IP . Both client and server are implemented in C++ and tested with windows XP. The primary and the vital module is the image server program that runs continuously from the SEM computer. It is mainly responsible for passing the scan control commands to the image acquisition system (DISS5) and to receive the incoming scan data. Inturn, this data is saved as an image. Later, the saved image is transferred to the image client running on the work computer. The server connects and communicates with the DISS5 when a client is connected. The communication between client and server is demonstrated in the figure 6.2. The other module that is running on the work computer is control server. As seen in the figure 6.1, the control server communicates with robot and platform controllers via serial port connections and with gripper controller via USB. It uses the image data received by client to compute the control sequence for different devices. All the modules are developed as individual programs in C++. Besides, the image client and control server are wrapped under a common program whose functionalities can be accessed using a special purpose GUI (APROS3) that has been developed during the thesis. More details about the GUI are presented in the next section.

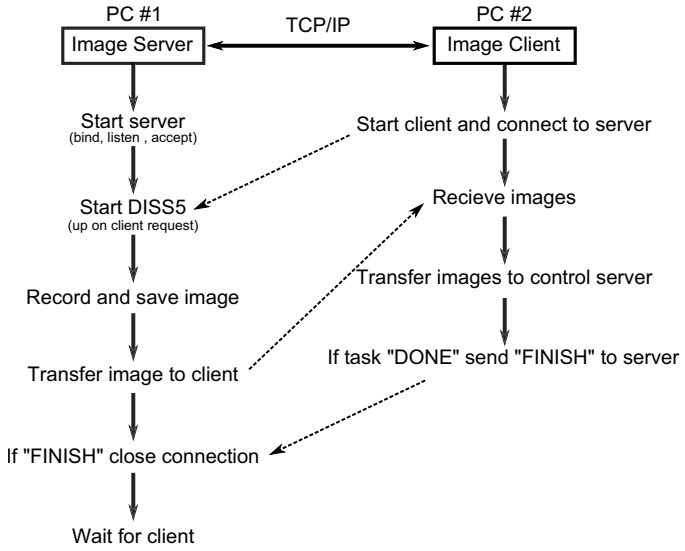



Figure 6.2: Client-server communication model.

## 6.2 Scan control and image acquisition

As mentioned earlier, DISS5 scan controller from Point Electronics GmbH [Pie13] has been used with the JEOL SEM. It allows the SEM to be controlled from an external computer system. During this thesis, in order to integrate and use the SEM as an imaging sensor, a software library (`diss5main`) has been developed in C++ to access the scan controller. The developed library possesses the following functionalities: open device for communication over USB, set scan parameters, read the incoming buffer data, process the data and close device communication. The communication with the device is established with the initial scan and is closed only when the process is finished (i.e. with a client request). It writes all the required scan parameters (explained below) to the device with the initial scan and only scan area dimensions are updated with each scan. This feature allows to increase or decrease the image size during the process (for speeding up the scan).

At any magnification the DISS5 is capable of providing a maximum scan area of $16384 \times 16384$ pixels. For a typical application, the image scan area is determined by selecting or setting the paramters XOffset (left offset), YOffset (upper offset), XRep (width-1), YRep (height-1), XStep (horizontal step) and YStep (vertical step) (see figure 6.3). For a desired image size of $W \times H$, these parameters can be computed using the sample code shown in the listing 6.1. The digital magnification can be modified by changing the step sizes.

Some of the other scan commands are:

- `LineRepeat` specifies the number of times a raster line has to be scanned.
- `FrameRepeat` specifies the number of times a frame has to be scanned.
- `LineStartDelay` is the time (ms) to wait before starting scanning a single line.

- `SampleRepeat` determines the oversampling for analog channels ($2^n - 1,\ n = 0, 1 \cdots 15$). This parameter helps in increasing or decreasing the scan time.

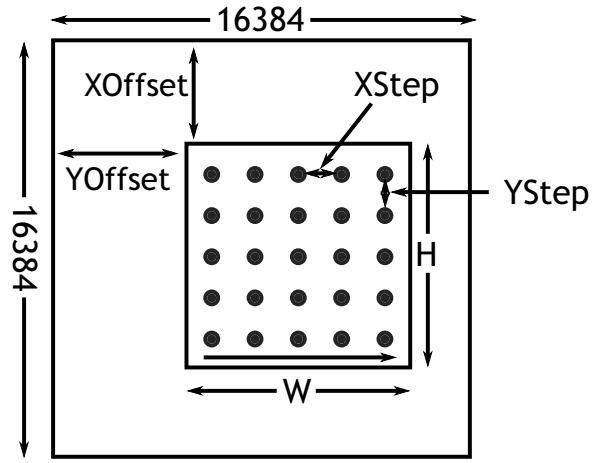



Figure 6.3: The scanning area selection using DISS5 scan controller.

```
 1 XRep = W-1
 2 YRep = H-1
 3 XStep = 16384 / W
 4 YStep = 16384 / H
 5 if (XStep < YStep)
 6        YStep = XStep
 7 else
 8        XStep = YStep
 9 end if
10 XOffset = (16384 - W * XStep) / 2
11 YOffset = (16384 - H * YStep) / 2
```

Listing 6.1: Pseudocode for computing the scan parameters.

Apart from communicating with the scan controller, a wrapper library (`semRemote`) containing the functions to access/modify the device parameters such as coarse focus, magnification, working distance, spot size and scan rotation etc. has been developed during this thesis (in C++). It communicates with the `SEMServer` program (provided by Point Electronics GmbH) running from SEM control computer (PC-1) to gain remote access to these parameters.

## 6.3 APROS3 software GUI

During this thesis, a software application GUI, APROS3 (Application PROgram for SEM Station Setup) has been developed to integrate all the developed software modules as well as to monitor the undergoing process. It mainly reduces the complexity of operating different system components like positioning stage, robot etc. The interface is designed in such a way that a non-expert user can operate the SEM and system devices without much effort. APROS3 is designed and built using Qt (for user interface), Open CV (for image processing), ViSP (for visual servoing) and FFTW3 (for Fourier domain operations) libraries. It is programmed in C++ with cross-platform compatibility and currently, it has been tested under Windows XP. The overall software is programmed in such a way that each developed module contains individual class interface. This feature helps in the future development and integration of the different modules without modifying the existing main interface.

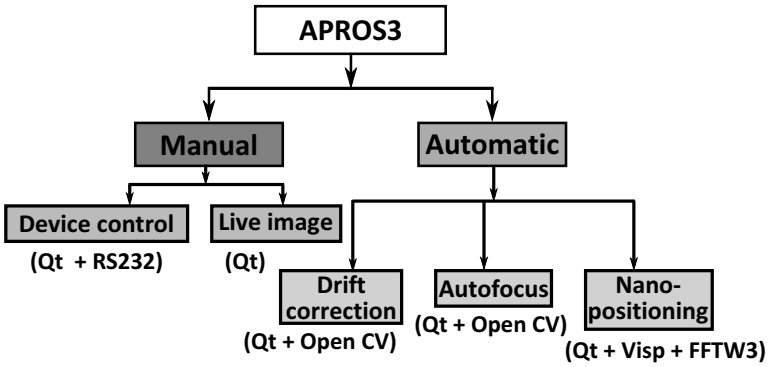



Figure 6.4: Architecture diagram for APROS3 software.

The total software is divided into two core components: manual control and automatic control (figure 6.4). At the first hand, manual control enables the user to access and operate the existing system devices. The main interface window containing the manual control options is shown in the figure 6.5. Besides device operation, it also includes a simple interface for live imaging that contains different functionalities such as image acquisition, modifying basic SEM parameters etc. The live imaging window is shown in figure 6.6.

On the other hand, automatic control component includes the automated tasks (drift compensation, autofocusing etc.) developed during this thesis. Due to the time constraint, some of the tasks are not integrated into the software and will be added in the future. Every module in this component represents a specific task and provides an option for the user to save the task output as formatted data. The developed interfaces for drift compensation and autofocusing are shown in the figures 6.7 and 6.8. The drift compensation module can be seen as live imaging module with added compensation functionality. It provides an option for the user to choose between different methods. The autofocusing module is a one time operation that exits when the task is finished.

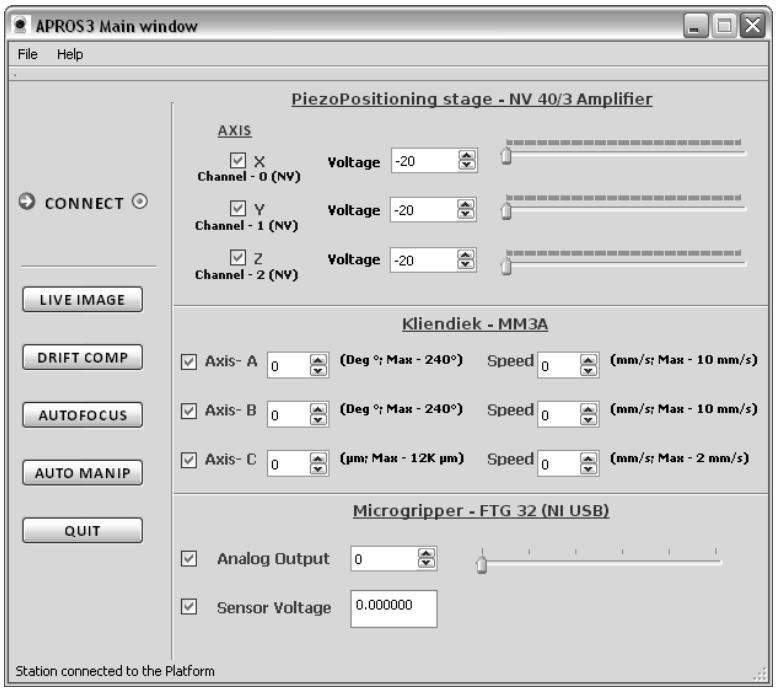



Figure 6.5: The software main window containing device control widgets.

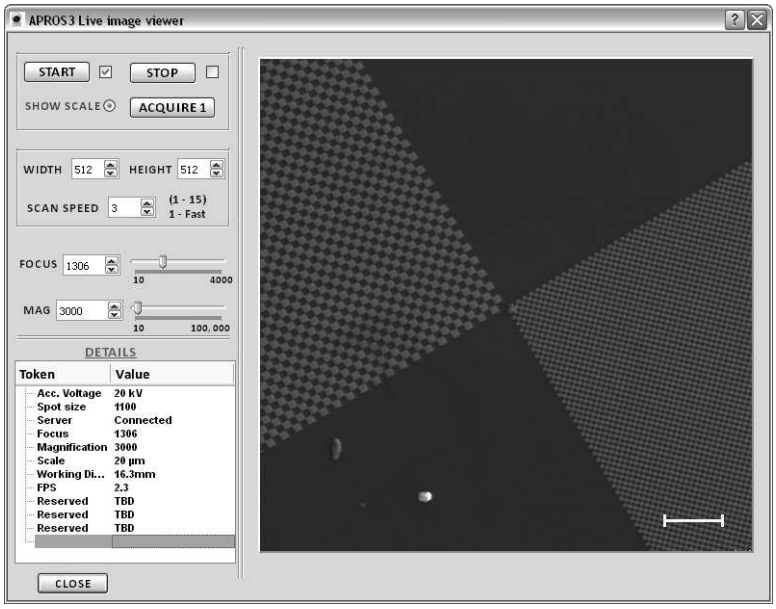



Figure 6.6: The live image viewing window of the software.

It provides continuous visual monitoring of the task. The real-time plotting has been accomplished using `QCustomplot` add-on library for Qt.

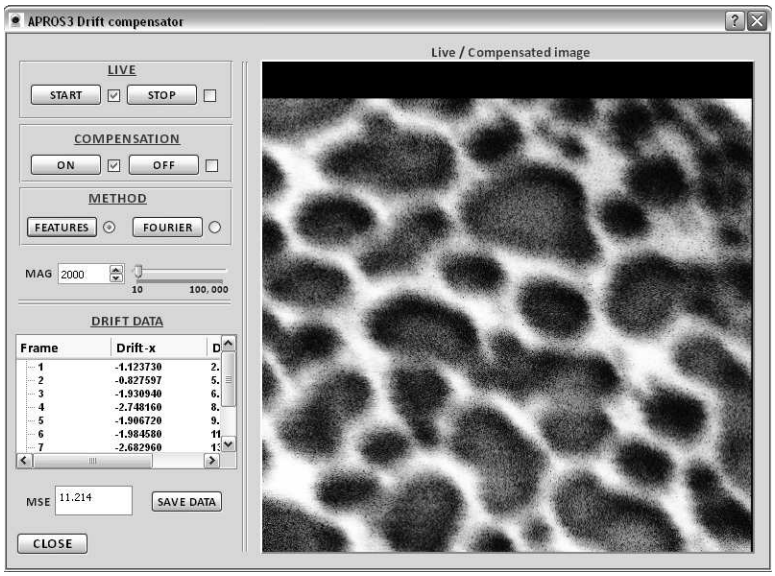



Figure 6.7: The interface window for performing real-time drift compensation.

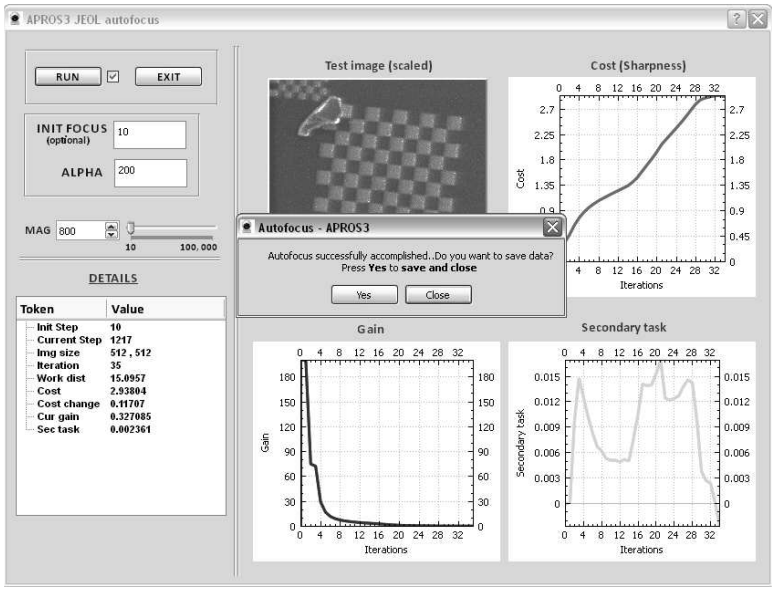



Figure 6.8: The interface window for autofocusing.

## 6.4 Conclusion

This is the first of all the tasks performed during this thesis. The developed software architecture enables all the system devices to communicate synchronously with each other. All the modules in the current architecture are programmed afresh and are developed like individual classes that can use the functionalities of other classes. This

type of design provides a flexibility in developing or modifying individual modules. As a final task, a software user interface application is developed to provide an easy operation of JEOL SEM and other hardware components. All the developed methods (during this thesis) are integrated with this software.

Chapter 7

# Summary and future perspectives

*This very last chapter provides a general summary of all the work performed in the framework of this thesis. Here, we present the main contributions regarding each developed task along with some points of view to continue the work in future.*

## 7.1 Summary

Being a new field, micro-nanorobotics has many challenges to solve. The primary and most important problem it is facing currently is the lack of effective processes for building and characterizing nanostructures. In order to solve this and to perform efficient nanomanipulation, reliable techniques and tools are indispensable. Various works that are performed in these areas are still isolated and requires the intervention of an expert human operator in accomplishing the overall task. In this thesis we show that, this problem can be tackled effectively up to an extent using the microscopic vision information. The overall work done is divided into two parts: vision and visual servoing. They are summarized below.

### 7.1.1 Part-1: Imaging with SEM

From the analysis of various microscopic devices, it has been found that, SEM suits better for performing real time manipulation tasks due to its high working range and the capability of providing near real time image feedback. The primary advantage associated with SEM imaging is its capability in producing images with high resolution (better than 1 nm) and high magnification. Nowadays, the trend is to add an ion beam, a gas injector and at least one robot to the microscope column to allow machining, characterization or assembly of structures. In this case, the observed scene is dynamic. But electronic imaging is by far known for its sensitivity to the scanning frequency i.e. a high frequency of acquisition leads to low quality images. From the study we found that, many sources contribute to the addition of this noise. Out of all, the most important

one is the noise due to the number of electrons interacted with the specimen surface and the number of emitted electrons recorded at the detector. In this case, it is necessary to analyze the image quality beforehand.

In order to accomplish this, we have developed a method to quantify the level of image SNR. It is based on image non-linear filtering and can be used in real time for quality monitoring and noise level quantification. Initially, it has been used to compare the imaging performances of two SEMs: one with a tungsten gun and the other with a FEG. From the conducted experiments, we found that the FEG SEM provides good quality images which is quite obvious mainly due to the nature of the beam production. Later, it has been used to evaluate the image quality at different conditions. The evolution of the SNR with respect to magnification is found to be a non-uniform decreasing function. This result seems to be surprising at first but becomes logical after analysis. Indeed, increasing the magnification results in a decrease of the surface corresponding to a pixel and thus a reduction in the number of electrons emitted. When we evaluated the image quality with respect to the focus, it provided a maximum quality with the image at best focus. This is mainly due to the concentration of electrons that is more at this point and produces the maximum number of secondary electrons. When we evaluated the image quality in real time, the SNR level increased rapidly to an extent (for a time of about 500 $s$) and then decreased (for a time of about 750 $s$). Thereafter, it increased irregularly up to stabilization point for a time of the order of 3000 $s$. This time corresponds to the time required by the filament to attain a specific heat to produce a stable number of primary electrons. As a result, it is recommended to wait at least 50 minutes from the ignition of JSM 820 SEM before starting an experiment.

The second problem that is considered in the vision module is to compensate the time varying distortion also called as image drift in real time. This phenomenon is mainly observed at high magnifications and can be characterized as the evolution of pixel positions from time to time. In order to solve this issue, we have implemented a new image registration-based global drift compensation method utilizing the idea of homography estimation from keypoints detection and matching. Three keypoint detectors are tried for this purpose. Out of all, ORB detector showed the better speed and accuracy in estimating and compensating the drift with an average MSE of 7.5. We have also compared the developed method with the traditional correlation-based method in order to find a better solution. When performed the experiments with varying magnifications ranging from $10,000\times$ to $30,000\times$ (with a step of $1000\times$), it has been found that the drift produced in the images is only a translation in $x$ and $y$ axes and no rotation is involved. The path followed can be approximated by linear motion. It is also observed that the velocity of drift increases with time as the number of pixels evolving is more at higher times. It is assumed that this drift might be caused due to two reasons: 1) thermal variations inside the chamber 2) effect of charging and the type of sample surface. However, the most straightforward reason is still unknown.

### 7.1.2 Part-2: Visual servoing under SEM

As an initial step towards the visual servoing using SEM, we concentrated on the problem of computing the inter-object depth. Genrally, in order to perform an autonomous manipulation task using SEM, the knowledge of inter-object depth is crucial. With SEM as an imaging sensor, it is highly challenging to find this depth information. In this thesis we show that, this problem can be overcome by using the image focus. For this purpose, we have developed a visual servoing-based autofocusing technique that performs the focusing by controlling the size of the focus step using an adaptive gain. When compared this method with the traditional search-based approaches, it has shown a great speed and precision in reaching the best focus. Moreover, it has been tested and validated at different experimental conditions including high noise, high brightness, high magnification, and with a different sample. An added advantage of this method is that it is not effected by the hysteresis shown by the SEM electromagnetic lenses and possess an ability to skip the local maxima that affect the focusing process. Using this method, we compute the inter-object depth simply by focusing each object independently in their respective ROIs. From the obtained best focus step, the related electronic working distance is computed and is used in estimating the inter-object depth. Note that, as the maximum time taken to perform autofocusing is about 19 $s$ (for a frame rate of 2.1 frames per second) for an image size of $512 \times 512$ pixels, it can be more fast for smaller regions while estimating the depth and can be used in real time manipulations. Apart from inter-object depth, we compute the individual pixel point depth from the focus in order to reconstruct the object structure. In this case, the depth is estimated by acquiring multiple frames with varying working distance and by searching for the image in which the pixel sharpness is maximum. Even though, the method suffers from the high depth of field property of SEM, it has provided acceptable results.

Another important work performed during this thesis is the development of visual servoing schemes for nanopositioning. For the first time, two visual servoing methods that are completely free from image processing are used with SEM. In general, to perform characterization of micro-nanostructures by probing to measure the structure stiffness and to perform in situ nanomanipulation, a reliable and efficient nanopositioning is required. At this small scale, this task is heavily challenging due to inadequate feedback information. Due to this reason many works have relied on the internal position sensors which is not a feasible solution at nanoscale. One solution could be by using SEM image information as feedback. However, the primary difficulty comes in using these images as they are heavily noisy (since we use fast acquisition). Due to this, it is mostly difficult to apply any tracking methods. So, in order to overcome this problem, we have implemented two positioning schemes that does not require any tracking and uses the total image information for feedback. The first one uses overall pixel intensities of an image as visual measurements and the second one is a frequency domain method that uses Fourier transform to compute the relative motion between images. For both the methods, respective control laws are designed to minimize the positioning error. After performing the positioning at different conditions that include, varying scan speed and

magnification, we conclude that frequency domain approach seems to perform better than the intensity-based approach. This is quite obvious, since the brightness and contrast are never be constant with SEM imaging and frequency domain method is robust to these variations. However, the intensity-based method also succeeded in almost all tests and can be used for positioning with FEG SEMs since they provide good SNR and less intensity variations.

## 7.2 Future perspectives

Due to the system limitations and components maintenance, some of the tasks are left over to be completed in future. Along with these, some other issues mentioned below can also be taken into account.

**Sub-pixel drift compensation:** The drift compensation explained in **chapter 3** is capable of correcting the global image drift. However, from the performed experiments, we found the existence of some uncorrected pixels after compensation (see figure 7.1). From the analysis, we came to know that these pixels come from the conducting regions on the sample and it shows that the electron beam drifts more towards the conducting regions resulting in a pixel level drift. So to perform accurate measurements from the images, it is necessary to implement sub-pixel drift compensation.

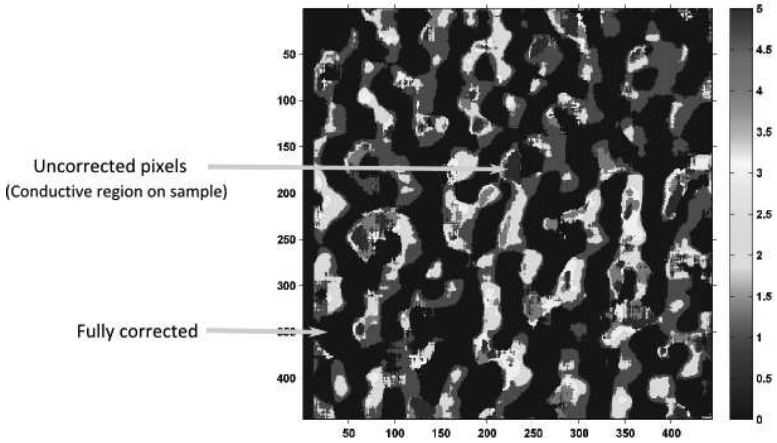



Figure 7.1: Disparity map showing the corrected and uncorrected pixels.

**Fine depth estimation:** The inter-object depth detection explained in **chapter 4** can only provide coarse depth, since it uses focus information. This is acceptable for some applications where the objects are large enough. However, to perform manipulations with good precision, it is recommended to find a way to estimate the fine depth. As an example, we can refer to the work performed by Eichhorn et al. [EFW+08] where they have used a *touch down sensor* to detect the contact between two surfaces and to estimate the fine depth.

**Positioning control of microrobotic actuator:** Due to the system space limitations, the microrobot positioning is not performed during this thesis. It is planned to perform this task in the near future on an other SEM with bigger vacuum chamber. Both the methods explained in **chapter 5** will be considered for this purpose. Apart from that, the future work also includes to improve the positioning control in terms of accuracy.

**Micropart handling using vision and force information:** One of our main goals for the future is to perform a complete manipulation task that requires precise handling of the micro / nano-objects. Example applications could be probing a nanowire (figure 7.2(a)), handling microspheres (figure 7.2(b)) and electrical characterization of nanowires (figure 7.2(c)). In order to perform this, it is required to integrate the force information (coming from gripper fingers) along with vision information. To accomplish this task, a dual mode control strategy has to be designed that can switch between force and visual control.

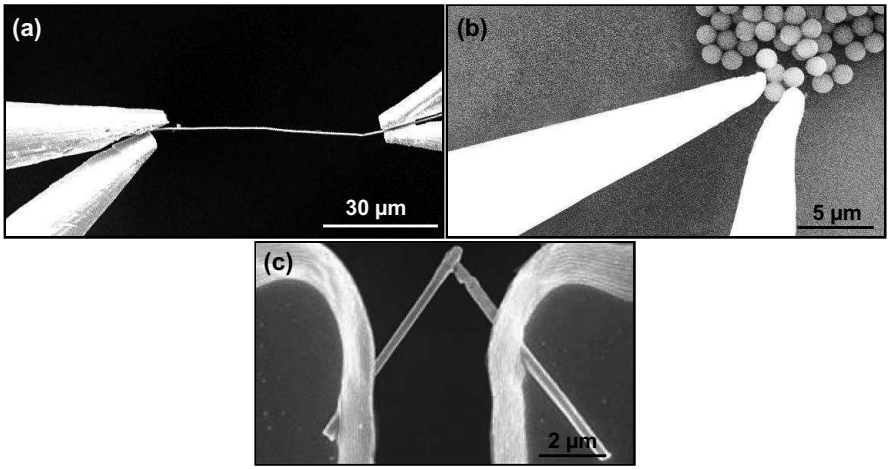



Figure 7.2: (a) Nanowire probing (b) handling microspheres (c) electrical characterization of nanowires. Images courtesy: (a) and (b) Kleindiek Nanotechnik GmbH (c) Gu et al. [GYGG05].

**3D reconstruction of standing acoustic waves for surface characterization:** 3D Visualization of the standing acoustic waves on the crystal surfaces is very useful for the surface characterization and observing defects (figure 7.3). While performing the 3D reconstruction of these waves, it is necessary to take into account the spatial distortion. Even though the spatial distortion is not high with SEM imaging, it is recommended to develop a method to compensate this effect. Moreover, the focus-based 3D surface reconstruction method explained in chapter 4 need to be tuned i.e. to modify the depth of focus according to the size of the waves.

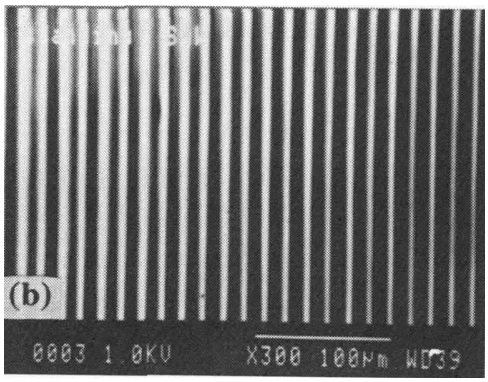



Figure 7.3: Standing acoustic waves observed on the surface of LiNbO3 [RBT94].

**Software testing and migration to new SEM:** Even though, the developed software application (APROS3) is running well in most of the cases, it has not tested thoroughly. Also due to the time constraint, the nanopositioning module is not included in the final interface. So for the future work, it is highly recommended to perform an overall unit testing and to integrate the positioning module. Once this task is finished, it is planned to transfer all the developments made along with the software application to another SEM (Carl Zeiss Auriga 60 with a FEG) in our department.

# Bibliography


[ACC+13] Joël Agnus, Nicolas Chaillet, Cédric Clévy, Sounkalo Dembélé, Michaël Gauthier, Yassine Haddab, G Laurent, Philippe Lutz, Nadine Piat, Kanty Rabenorosoa, et al. Robotic microassembly and micromanipulation at femto-st. *Journal of Micro-Bio Robotics*, pages 1–16, 2013.

[AHGP09] Joël Agnus, David Hériban, Michaël Gauthier, and Valérie Pétrini. Silicon end-effectors for microgripping tasks. *Precision engineering*, 33(4):542–548, 2009.

[AMH+03] Kanna Aoki, Hideki T Miyazaki, Hideki Hirayama, Kyoji Inoshita, Toshihiko Baba, Kazuaki Sakoda, Norio Shinya, and Yoshinobu Aoyagi. Microassembly of semiconductor three-dimensional photonic crystals. *Nature materials*, 2(2):117–121, 2003.

[Bat00] Christopher F. Batten. Autofocusing and astigmatism correction in the scanning electron microscope. Master's thesis, University of Cambridge, 2000.

[BBN+06] Felix Beyeler, Dominik J Bell, Bradley J Nelson, Yu Sun, A Neild, S Oberti, and J Dual. Design of a micro-gripper and an ultrasonic manipulator for handling micron sized objects. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 772–777. IEEE, 2006.

[BD91] Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods*. Springer-Verlag, New York, 1991.

[BD95] J Baina and J Dublet. Automatic focus and iris control for video cameras. In *Fifth International Conference on Image Processing and its Applications*, pages 232–235. IET, 1995.

[BDD+99] G Binnig, M Despont, U Drechsler, W Haberle, M Lutwyche, P Vettiger, HJ Mamin, BW Chui, and Th W Kenny. Ultrahigh-density atomic force

microscopy data storage with erase capability. *Applied Physics Letters*, 74(9):1329–1331, 1999.

[BDH+76] John F Brenner, Brock S Dew, Brian J Horton, Thomas King, Peter W Neurath, and William D Selles. An automated microscope for cytologic research a preliminary evaluation. *Journal of Histochemistry & Cytochemistry*, 24(1):100–111, 1976.

[BETVG08] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Computer vision and image understanding*, 110(3):346–359, 2008.

[BF12] Malte Bartenwerfer and Sergej Fatikow. Nanorobot-based handling and transfer of individual silicon nanowires. *International Journal of Intelligent Mechatronics and Robotics (IJIMR)*, 2(2):34–46, 2012.

[BM06] Selim Benhimane and Ezio Malis. Homography-based 2d visual servoing. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 2397–2402. IEEE, 2006.

[BM07] Selim Benhimane and Ezio Malis. Homography-based 2d visual tracking and servoing. *The International Journal of Robotics Research*, 26(7):661–676, 2007.

[BMC04] Tamal Bose, François G Meyer, and Mei-Qin Chen. *Digital signal and image processing*. J. Wiley, 2004.

[Bon93] Dawn A Bonnell. *Scanning tunneling microscopy and spectroscopy: theory, techniques, and applications*. VCH New York, 1993.

[Bou99] Jean-Yves Bouguet. *Visual methods for three-dimensional modeling*. PhD thesis, Citeseer, 1999.

[Bov10] Alan C Bovik. *Handbook of image and video processing*. Access Online via Elsevier, 2010.

[BQG86] G Binnig, CF Quate, and Ch Gerber. Atomic force microscope. *Physical review letters*, 56(9):930, 1986.

[BR10] John C Bowman and Malcolm Roberts. Fftw++: A fast fourier transform c++ header class for the fftw3 library, 2010. `http://http://www.fftw.org/`.

[Bro92] Lisa Gottesfeld Brown. A survey of image registration techniques. *ACM computing surveys (CSUR)*, 24(4):325–376, 1992.

[CCR08] James C Church, Yixin Chen, and Stephen V Rice. A spatial median filter for noise removal in digital images. In *IEEE Southeastcon*, pages 618–623. IEEE Southeastcon, 2008.

[CH06] François Chaumette and Seth Hutchinson. Visual servo control. i. basic approaches. *Robotics & Automation Magazine, IEEE*, 13(4):82–90, 2006.

[Cha98] Francois Chaumette. Potential problems of stability and convergence in image-based and position-based visual servoing. In *The confluence of vision and control*, pages 66–78. Springer, 1998.

[CM11] Christophe Collewet and Eric Marchand. Photometric visual servoing. *IEEE Transactions on Robotics*, 27(4):828–834, 2011.

[CMC08] Christophe Collewet, Eric Marchand, and François Chaumette. Visual servoing set free from image processing. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 81–86. IEEE, 2008.

[CMK12] Deshan Chen, Atsushi Miyamoto, and Shun'ichi Kaneko. Robust surface reconstruction in sem with two bse detectors. In *Mechatronics (MECATRONICS), 9th France-Japan & 7th Europe-Asia Congress on and Research and Education in Mechatronics (REM)*, pages 64–70. IEEE, 2012.

[Cor05] Nicolas Cornille. *Accurate 3D shape and displacement measurement using a scanning electron microscope*. PhD thesis, University of South Carolina, 2005.

[CVMP08a] Petr Cizmar, András E Vladár, Bin Ming, and Michael T Postek. Artificial sem images for testing resolution-measurement methods. *Microscopy and Microanalysis*, 14(S2):910–911, 2008.

[CVMP08b] Petr Cizmar, András E Vladár, Bin Ming, and Michael T Postek. Simulated sem images for resolution measurement. *Scanning*, 30(5):381–391, 2008.

[CVP10] Petr Cizmar, András E Vladár, and Michael T Postek. Real-time scanning charged-particle microscope image composition with correction of drift. *Microscopy and Microanalysis*, 17(2):302, 2010.

[D'E05] John R D'Errico. Surface fitting using gridfit. MATLAB Central File Exchange, 2005.

[DK10] Etienne Dombre and Wisama Khalil. *Modeling, Performance Analysis and Control of Robot Manipulators*. Wiley Online Library, 2010.

[DM09] Amaury Dame and Eric Marchand. Entropy-based visual servoing. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 707–713. IEEE, 2009.

[DN07] Lixin Dong and Bradley J Nelson. Tutorial-robotics in the small part ii: nanorobotics. *Robotics & Automation Magazine, IEEE*, 14(3):111–121, 2007.

[EFW+08] Volkmar Eichhorn, Sergej Fatikow, Thomas Wich, Christian Dahmen, Torsten Sievers, Karin Nordström Andersen, Kenneth Carlson, and Peter Bøggild. Depth-detection methods for microgripper based cnt manipulation in a scanning electron microscope. *Journal of Micro-Nano Mechatronics*, 4(1-2):27–36, 2008.

[EG94] AV Ermakov and EL Garfunkel. A novel afm/stm/sem system. *Review of scientific instruments*, 65(9):2853–2854, 1994.

[Era82] SJ Erasmus. Reduction of noise in tv rate electron microscope images by digital filtering. *Journal of Microscopy*, 127(1):29–37, 1982.

[ES90] Donald M Eigler and Erhard K Schweizer. Positioning single atoms with a scanning tunnelling microscope. *Nature*, 344(6266):524–526, 1990.

[FAA75] J Frank and L Al-Ali. Signal-to-noise ratio of electron micrographs obtained by cross correlation. *Nature*, 256:376–379, 1975.

[FAD03] Toshio Fukuda, Fumihito Arai, and Lixin Dong. Assembly of nanodevices with carbon nanotubes through nanorobotic manipulations. *Proceedings of the IEEE*, 91(11):1803–1818, 2003.

[FCC+91] Lawrence Firestone, Kitty Cook, Kevin Culp, Neil Talsania, and Kendall Preston. Comparison of autofocus methods for automated microscopy. *Cytometry*, 12(3):195–206, 1991.

[FE08] Sergej Fatikow and Volkmar Eichhorn. Nanohandling automation: trends and current developments. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 222(7):1353–1369, 2008.

[Fem11] FemtoTools. `http://www.femtotools.com`, 2011.

[FJ05] Matteo Frigo and Steven G Johnson. The design and implementation of fftw3. *Proceedings of the IEEE*, 93(2):216–231, 2005.

[FM89] John T Feddema and Owen R Mitchell. Vision-guided servoing with feature-based trajectory generation [for robots]. *IEEE Transactions on Robotics and Automation*, 5(5):691–700, 1989.

[FNLE09] Toshio Fukuda, Masahiro Nakajima, Pou Liu, and Haitham ElShimy. Nanofabrication, nanoinstrumentation and nanoassembly by nanorobotic manipulation. *The International Journal of Robotics Research*, 28(4):537–547, 2009.

[Fre63] Martin M Freundlich. Origin of the electron microscope. *Science*, 142(3589):185–188, 1963.

[FSF02] Stephan Fahlbusch, Alexsandr Shirinov, and Sergej Fatikow. Afm-based micro force sensor and haptic interface for a nanohandling robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 2, pages 1772–1777. IEEE, 2002.

[FTDLFP08] Guillaume Fortier, Brahim Tamadazte, Sounkalo Dembélé, and Nadine Le Fort-Piat. Estimating the 3d orientation of a microgripper by processing the focus data from the images delivered by a videomicroscope. In *International Workshop on Robotic and Sensors Environments (ROSE)*, pages 58–63. IEEE, 2008.

[FWH+07] Sergej Fatikow, Thomas Wich, Helge Hülsen, Torsten Sievers, and Marco Jähnisch. Microrobot system for automatic nanohandling inside a scanning electron microscope. *IEEE/ASME Transactions on Mechatronics*, 12(3):244–252, 2007.

[GFM+00] Martin Guthold, Michael R Falvo, Garrett W Matthews, Scott Paulson, Sean Washburn, Dorothy A Erie, Richard Superfine, Frederick P Brooks Jr, and Taylor II Russell M. Controlled manipulation of molecular samples with the nanomanipulator. *IEEE/ASME Transactions on Mechatronics*, 5(2):189–198, 2000.

[GHT11] Steffen Gauglitz, Tobias Höllerer, and Matthew Turk. Evaluation of interest point detectors and feature descriptors for visual tracking. *International journal of computer vision*, 94(3):335–360, 2011.

[Gmb] Kleindiek Nanotechnik GmbH. Micromanipulator for electron microscopy. `http://www.nanotechnik.com/mm3a-em.html`.

[GMN+99] M. Guthold, G. Matthews, A. Negishi, R. M. Taylor, D. Erie, F. P. Brooks, and R. Superfine. Quantitative manipulation of dna and viruses with the nanomanipulator scanning force microscope. *Surface and Interface Analysis*, 27(5-6):437–443, 1999.

[GNJ+03] J Goldstein, DE Newbury, DC Joy, CE Lyman, P Echlin, E Lifshin, L Sawyer, and JR Michael. *Scanning electron microscopy and X-ray microanalysis*, volume 1. Springer Us, 2003.

[GS90] Bernd Girod and Stephen Scherock. Depth from defocus of structured light. In *Advances in Intelligent Robotics Systems Conference*, pages 209–215. International Society for Optics and Photonics, 1990.

[GW08] Rafael C Gonzalez and RE Woods. *Digital image processing (International ed.)*. Pearson Prentice Hall, Upper Saddle River, New Jersey, 2008.

[GYGG05] Zhiyong Gu, Hongke Ye, David H Gracias, and David Gracias. The bonding of nanowire assemblies using adhesive and solder. *JOM*, 57(12):60–64, 2005.

[GYL85] Frans CA Groen, Ian T Young, and Guido Ligthart. A comparison of different focus functions for use in autofocus algorithms. *Cytometry*, 6(2):81–91, 1985.

[HHB+09] Gilgueng Hwang, Hideki Hashimoto, Dominik J Bell, Lixin Dong, Bradley J Nelson, and Silke Schön. Piezoresistive ingaas/gaas nanosprings with metal connectors. *Nano letters*, 9(2):554–561, 2009.

[HHC96] Seth Hutchinson, Gregory D Hager, and Peter I Corke. A tutorial on visual servo control. *IEEE Transactions on Robotics and Automation*, 12(5):651–670, 1996.

[HKS+98] TL Hansen, A Kühle, AH Sørensen, J Bohr, and PE Lindelof. A technique for positioning nanoparticles using an atomic force microscope. *Nanotechnology*, 9(4):337, 1998.

[HM90] Yotaro Hatamura and Hiroshi Morishita. Direct coupling system between nanometer world and human world. In *Micro Electro Mechanical Systems, 1990. Proceedings, An Investigation of Micro Structures, Sensors, Actuators, Machines and Robots. IEEE*, pages 203–208. IEEE, 1990.

[HP79] J Hill and W Park. Real time control of a robot with a mobile camera. In *Proceedings of the 9th International Symposium on Industrial Robots*, pages 233–246, 1979.

[HS81] Berthold KP Horn and Brian G Schunck. Determining optical flow. *Artificial intelligence*, 17(1):185–203, 1981.

[HS88] Chris Harris and Mike Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, page 50. Manchester, UK, 1988.

[HZ00] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*, volume 2. Cambridge Univ Press, 2000.

[IIY01] Yoshiharu Ishii, Akihiko Ishijima, and Toshio Yanagida. Single molecule nanomanipulation of biomolecules. *TRENDS in Biotechnology*, 19(6):211–216, 2001.

[Imm96] John Immerkaer. Fast noise variance estimation. *Computer Vision and Image Understanding*, 64(2):300 – 302, 1996.

[Jas09] Daniel Jasper. High-speed position tracking for nanohandling inside scanning electron microscopes. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 508–513. IEEE, 2009.

[Jas11] Daniel Jasper. *SEM-based motion control for automated robotic nanohandling*. PhD thesis, 2011.

[JF07] Marco Jähnisch and Sergej Fatikow. 3-d vision feedback for nanohandling monitoring in a scanning electron microscope. *International Journal of Optomechatronics*, 1(1):4–26, 2007.

[JJ95] Drahomira Janová and Jiří Jan. Robust surface reconstruction from stereo sem images. In *Computer Analysis of Images and Patterns*, pages 900–905. Springer, 1995.

[JJJ+09] Ara Jo, Wonchul Joo, Won-Hyeog Jin, Hyojin Nam, and Jin Kon Kim. Ultrahigh-density phase-change data storage without the use of heating. *Nature Nanotechnology*, 4(11):727–731, 2009.

[JJK11] KO Jung, W Joo, and DH Kim. Scanning electron microscope image enhancement using spread spectrum through dither signal imposition. *Journal of Electron Microscopy*, 60(6):367–373, 2011.

[KDN09] Bradley E Kratochvil, Lixin Dong, and Bradley J Nelson. Real-time rigid-body visual tracking in a scanning electron microscope. *The International Journal of Robotics Research*, 28(4):498–511, 2009.

[KMNF89] D Kahaner, CB Moler, S Nash, and GE Forsythe. *Numerical methods and software*. Prentice-Hall Englewood Cliffs, NJ, 1989.

[Kra08] Bradley E Kratochvil. *Visual tracking for nanorobotic manipulation and 3d reconstruction in an electron microscope*. PhD thesis, Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 17953, 2008.

[LBK12] Lorenz Lechner, Johannes Biskupek, and Ute Kaiser. Improved focused ion beam target preparation of (s)tem specimen-a method for obtaining ultrathin lamellae. *Microscopy and Microanalysis*, 18(02):379–384, 2012.

[LFSK06] C Liu, WT Freeman, R Szeliski, and SB Kang. Noise estimation from a single image. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 901–908, 2006.

[Low04] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[Mal04] Ezio Malis. Improving vision-based control using efficient second-order minimization techniques. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pages 1843–1848. IEEE, 2004.

[Mar07] Eric Marchand. Control camera and light source positions using image gradient information. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 417–422. IEEE, 2007.

[MC10] E. Marchand and C. Collewet. Using image gradient as a visual feature for visual servoing. In *EEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5687–5692, 2010.

[MCB99] Ezio Malis, Francois Chaumette, and Sylvie Boudet. 21/2d visual servoing. *IEEE Transactions on Robotics and Automation*, 15(2):238–250, 1999.

[MDP+12a] Abed C Malti, Sounkalo Dembélé, Nadine Piat, Claire Arnoult, and Naresh Marturi. Toward fast calibration of global drift in scanning electron microscopes with respect to time and magnification. *International Journal of Optomechatronics*, 6(1):1–16, 2012.

[MDP+12b] Naresh Marturi, Sounkalo Dembélé, Nadine Piat, et al. Performance evaluation of scanning electron microscopes using signal-to-noise ratio. In *International Workshop on MicroFactories.*, pages 1–6, 2012.

[MHMB04] Kristian Molhave, Torben Mikael Hansen, Dorte Norgaard Madsen, and Peter Boggild. Towards pick-and-place assembly of nanostructures. *Journal of Nanoscience and Nanotechnology*, 4(3):279–282, 2004.

[Mik04] Meier Mike. Introduction to materials science for engineers cd rom. `http://www.pearson-studium.de/books/3827370597/cd01/index.htm`, 2004.

[MM12] Ondrej Miksik and Krystian Mikolajczyk. Evaluation of local detectors and descriptors for fast feature matching. In *21st International Conference on Pattern Recognition (ICPR)*, pages 2681–2684. IEEE, 2012.

[MR06] Babak Mokaberi and Aristides AG Requicha. Drift compensation for automatic nanomanipulation with scanning probe microscopes. *IEEE Transactions on Automation Science and Engineering*, 3(3):199–207, 2006.

[MS04] Krystian Mikolajczyk and Cordelia Schmid. Scale & affine invariant interest point detectors. *International journal of computer vision*, 60(1):63–86, 2004.

[MSC05] Éric Marchand, Fabien Spindler, and François Chaumette. Visp for visual servoing: a generic software platform with a wide class of robot control skills. *Robotics & Automation Magazine, IEEE*, 12(4):40–52, 2005.

[MWJF10] U Mick, M Weigel-Jech, and S Fatikow. Robotic workstation for afm-based nanomanipulation inside an sem. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 854–859. IEEE, 2010.

[NA03] Yoshikazu Nakayama and Seiji Akita. Nanoengineering of carbon nanotubes for nanotools. *New Journal of Physics*, 5(1):128, 2003.

[Nan12] NanoPhysics. Fib tem sample preparation. `http://www.nanophysics.nl/Services/tem-sample-preparation.html`, 2012.

[NN94] Shree K Nayar and Yasuo Nakagawa. Shape from focus. *IEEE Transactions on Pattern analysis and machine intelligence*, 16(8):824–831, 1994.

[NNŠ03a] Markus Niederöst, Jana Niederöst, and Jiří Ščučka. *Automatic 3D reconstruction and visualization of microscopic objects from a monoscopic multifocus image sequence*. ETH, Swiss Federal Institute of Technology, Institute of Geodesy and Photogrammetry, 2003.

[NNŠ03b] Markus Niederöst, Jana Niederöst, and Jiří Ščučka. *Shape from focus: fully automated 3D reconstruction and visualization of microscopic objects*. Citeseer, 2003.

[Ols93] SI Olsen. Estimation of noise in images: An evaluation. *CVGIP: Graphical Models and Image Processing*, 55(4):319 – 323, 1993.

[OPT97] KH Ong, JCH Phang, and JTL Thong. A robust focusing and astigmatism correction method for the scanning electron microscope. *Scanning*, 19(8):553–563, 1997.

[OSR12] N Ouarti, B Sauvet, and S Régnier. High quality real-time video with scanning electron microscope using total variation algorithm on a graphics processing unit. *International Journal of Optomechatronics*, 6(2):163–178, 2012.

[Pau98] Bourke Paul. Image filtering in the frequency domain. Paul Bourke personal page., 1998. `http://paulbourke.net/miscellaneous/imagefilter/`.

[PCL+04] L-M Peng, Q Chen, XL Liang, S Gao, JY Wang, S Kleindiek, and SW Tai. Performing probe experiments in the sem. *Micron*, 35(6):495–502, 2004.

[Pie13] Piezosystemjena. Piezo actuators & nano stages, positioning equipment. `http://www.piezosystem.com/`, 2013.

[PK93] Nikolaos P Papanikolopoulos and Pradeep K Khosla. Adaptive robotic visual tracking: Theory and experiments. *IEEE Transactions on Automatic Control*, 38(3):429–445, 1993.

[PLB06] Ezequiel Ponz, Juan Luis Ladaga, and Rita Dominga Bonetto. Measuring surface topography with scanning electron microscopy. i. ezeimage: a program to obtain 3d surface data. *Microscopy and Microanalysis*, 12(2):170–177, 2006.

[Pro13] Lagadic Project. Visp website, 2013. `http://www.irisa.fr/lagadic/visp/visp.html`.

[PTWZ10] Z Pei, Q Tong, L Wang, and J Zhang. A median filter method for image noise variance estimation. In *Second International Conference on Information Technology and Computer Science (ITCS)*, pages 13–16, 2010.

[PVKN06] Martin Probst, Karl Vollmers, Bradley E Kratochvil, and Bradley J Nelson. Design of an advanced microassembly system for the automated assembly of bio-microrobots. In *Fifth International Workshop on Microfactories*, pages 25–27, 2006.

[RBT94] DV Roshchupkin, M Brunel, and L Tucoulou. Visualization of surface acoustic waves by scanning electron microscopy. *Le Journal de Physique IV*, 4(C5):C5–1229, 1994.

[RC96] B Srinivasa Reddy and Biswanath N Chatterji. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 5(8):1266–1271, 1996.

[RD05] Edward Rosten and Tom Drummond. Fusing points and lines for high performance tracking. In *Tenth IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 1508–1515. IEEE, 2005.

[Rei98] L Reimer. *Scanning Electron Microscopy: Physics of Image Formation and Microanalysis*, volume 11. Springer, 2nd edition, 1998.

[Req03] Aristides AG Requicha. Nanorobots, nems, and nanoassembly. *Proceedings of the IEEE*, 91(11):1922–1933, 2003.

[RLU99] K Rank, M Lendl, and R Unbehauen. Estimation of image noise variance. *IEE Proceedings - Vision, Image and Signal Processing*, 146(2):80–84, 1999.

[RMM09] M Rudnaya, RMM Mattheij, and JML Maubach. Iterative autofocus algorithms for scanning electron microscopy. *Microscopy and Microanalysis*, 15(S2):1108–1109, 2009.

[RMM10] M Rudnaya, RMM Mattheij, and JML Maubach. Evaluating sharpness functions for automated scanning electron microscopy. *Journal of microscopy*, 240(1):38–49, 2010.

[RMM+12] S Roy, J Meunier, AM Marian, F Vidal, I Brunette, and S Costantino. Automatic 3d reconstruction of quasi-planar stereo scanning electron microscopy (sem) images*. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 4361–4364. IEEE, 2012.

[RRKB11] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: an efficient alternative to sift or surf. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2564–2571. IEEE, 2011.

[SCC04] KS Sim, HT Chuah, and Z. Cheng. Robust image signal-to-noise ratio estimation using mixed lagrange time delay estimation autoregressive model. *Scanning*, 26(6):287–295, 2004.

[SCN93] M Subbarao, TS Choi, and A Nikzad. Focusing techniques. *Optical Engineering*, 32(11):2824–2836, 1993.

[SD03] Mulapudi Satya and Joy David. Is sem noise gaussian. *Microscopy and Microanalysis*, 9(S02):982–983, 2003.

[SDN05] Yu Sun, Stefan Duthaler, and Bradley J Nelson. Autofocusing algorithm selection in computer microscopy. In *EEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 70–76. IEEE, 2005.

[SF06] Torsten Sievers and Sergej Fatikow. Real-time object tracking for the robot-based nanohandling in a scanning electron microscope. *Journal of Micromechatronics*, 3(3/4):267, 2006.

[Sie11] T Sievers. Global sensor feedback for automatic nanohandling inside a scanning electron microscope. In *Intelligent Production Machines and Systems-2nd I* PROMS Virtual International Conference 3-14 July 2006*, page 289. Access Online via Elsevier, 2011.

[Sit01] Metin Sitti. Survey of nanomanipulation systems. In *1st IEEE Conference on Nanotechnology (IEEE-NANO)*, pages 75–80. IEEE, 2001.

[Sit07] Metin Sitti. Microscale and nanoscale robotics systems [grand challenges of robotics]. *Robotics & Automation Magazine, IEEE*, 14(1):53–60, 2007.

[SJ06] KK Sharma and SD Joshi. Image registration using fractional fourier transform. In *IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, pages 470–473, 2006.

[SKMH95] Tomomasa Sato, Takayuki Kameya, Hideki Miyazaki, and Yotaro Hatamura. Hand-eye system in nano manipulation world. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 1, pages 59–66. IEEE, 1995.

[SLG+07] MA Sutton, N Li, D Garcia, N Cornille, JJ Orteu, SR McNeill, HW Schreier, X Li, and AP Reynolds. Scanning electron microscopy for quantitative small and large deformation measurements part ii: experimental validation for magnifications from 200 to 10,000. *Experimental Mechanics*, 47(6):789–804, 2007.

[SLJ+07] MA Sutton, N Li, DC Joy, AP Reynolds, and X Li. Scanning electron microscopy for quantitative small and large deformation measurements part i: Sem imaging at magnifications from 200 to 10,000. *Experimental Mechanics*, 47(6):775–787, 2007.

[SLLT09] KS Sim, JK Lee, MA Lai, and CP Tso. Gaussian-taylor signal-to-noise ratio estimation for scanning electron microscope images. *Journal of Microscopy*, 236(1):18–34, 2009.

[SM98] McKinley Sky and Levine Megan. Cubic spline interpolation. *College of the Redwoods*, 1998.

[SMB00] Cordelia Schmid, Roger Mohr, and Christian Bauckhage. Evaluation of interest point detectors. *International Journal of computer vision*, 37(2):151–172, 2000.

[Smy97] GK Smyth. Polynomial approximation. *Encyclopedia of Biostatistics*, 1997.

[Sne10] Michael T Snella. Drift correction for scanning-electron microscopy. Master's thesis, Massachusetts Institute of Technology, 2010.

[SNT11] KS Sim, ME Nia, and CP Tso. Image noise cross-correlation for signal-to-noise ratio estimation in scanning electron microscope images. *Scanning*, 33(2):82–93, 2011.

[SOdSV+97] Andrés Santos, C Ortiz de Solorzano, Juan José Vaquero, JM Pena, Norberto Malpica, and F Del Pozo. Evaluation of autofocus functions in molecular cytogenetic analysis. *Journal of Microscopy*, 188(3):264–272, 1997.

[SRPA95] DM Schaefer, R Reifenberger, A Patil, and RP Andres. Fabrication of two-dimensional arrays of nanometer-size clusters with the atomic force microscope. *Applied Physics Letters*, 66(8):1012–1014, 1995.

[SSKS09] K Singh, N Singh, P Kaur, and R Saxena. Image compression by using fractional transforms. In *International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom)*, pages 411–413, 2009.

[STP04] KS Sim, JTL Thong, and JCH Phang. Effect of shot noise and secondary emission noise in scanning electron microscope images. *Scanning*, 26(1):36–40, 2004.

[Sub88] Murali Subbarao. Parallel depth recovery by changing camera parameters. In *ICCV*, pages 149–155, 1988.

[SWK01] Ferdinand Schmoeckel, Heinz Worn, and Matthias Kiefer. The scanning electron microscope as sensor system for mobile microrobots. In *8th IEEE International Conference on Emerging Technologies and Factory Automation*, volume 2, pages 599–602. IEEE, 2001.

[Tam09] Brahim Tamadazte. *Vision 3D et Commande par asservissements visuels pour la micromanipulation et le micro-assemblage de MEMS. Application à l'automatisation d'une station de micro-assemblage.* PhD thesis, Université de Franche-Comté, 2009.

[TDF+08] B. Tamadazte, S. Dembélé, G. Fortier, L. Fort-Piat, et al. Automatic micromanipulation using multiscale visual servoing. In *IEEE International Conference on Automation Science and Engineering (CASE)*, pages 977–982. IEEE, 2008.

[TF11] Robert Tunnell and Sergej Fatikow. 3d position detection with an fib-sem dual beam system. In *Proceedings of the 10th WSEAS international conference on communications, electrical & computer engineering, and 9th WSEAS international conference on Applied electromagnetics, wireless and optical communications*, pages 128–133. World Scientific and Engineering Academy and Society (WSEAS), 2011.

[TLFPD10] Brahim Tamadazte, Nadine Le Fort-Piat, and Sounkalo Dembélé. Robotic micromanipulation and microassembly using monoview and multiscale visual servoing. *IEEE/ASME Transactions on Mechatronics*, (99):1–11, 2010.

[TLFPM10] Brahim Tamadazte, Nadine Le Fort-Piat, and Eric Marchand. A direct visual servoing scheme for automatic nanopositioning. *IEEE/ASME Transactions on Mechatronics*, (99):1–9, 2010.

[TLFPM12] Brahim Tamadazte, Nadine Le-Fort Piat, and Eric Marchand. A direct visual servoing scheme for automatic nanopositioning. *IEEE/ASME Transactions on Mechatronics*, 17(4):728–736, 2012.

[TMCG02] Benoit Thuilot, Philippe Martinet, Lionel Cordesses, and Jean Gallice. Position based visual servoing: keeping the object in the field of vision. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pages 1624–1629. IEEE, 2002.

[TMDLFP10] Brahim Tamadazte, Eric Marchand, Sounkalo Dembélé, and Nadine Le Fort-Piat. Cad model-based tracking and 3d visual-based control for mems microassembly. *International Journal of Robotics Research*, 29(11):1416–1434, 2010.

[TN+12] F Timischl, S Nemoto, et al. A statistical model of signal–noise in scanning electron microscopy. *Scanning*, 34(3):137–144, 2012.

[Tsa87] Roger Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *Robotics and Automation, IEEE Journal of*, 3(4):323–344, 1987.

[TSH79] WJ Tee, KCA Smith, and DM Holburn. An automatic focusing and stigmating system for the sem. *Journal of Physics E: Scientific Instruments*, 12(1):35, 1979.

[TSP01] J Thong, KS Sim, and J Phang. Single-image signal-to-noise ratio estimation. *Scanning*, 23:328–336, 2001.

[Tuk74] JW Tukey. Nonlinear (nonsuperposable) methods for smoothing data. *Congressional Record (EASCON)*, 74:673–681, 1974.

[TY08] Shen-Chuan Tai and Shih-Ming Yang. A fast method for image noise estimation using laplacian operator and adaptive edge detection. In *3rd International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pages 1077–1081. IEEE, 2008.

[VO96] CR Vogel and ME Oman. Iterative methods for total variation denoising. *SIAM Journal on Scientific Computing*, 17(1):227–238, 1996.

[Vol87] D Vollath. Automatic focusing by correlative methods. *Journal of Microscopy*, 147(3):279–288, 1987.

[Vol88] D Vollath. The influence of the scene parameters and of noise on the behaviour of automatic focusing algorithms. *Journal of Microscopy*, 151(2):133–146, 1988.

[VP05] A Vladar and M Postek. Electron beam-induced sample contamination in the sem. *Microscopy and microanalysis*, 11:764, 2005.

[VSZ04] W Vogl, M Sitti, and MF Zah. Nanomanipulation with 3d visual and force feedback using atomic force microscopes. In *4th IEEE Conference on Nanotechnology*, pages 349–351, 2004.

[Was03] Masao Washizu. Dna manipulation in electrostatic fields. In *7th international Conference on Miniaturized Chemical and Biochemical Analysts Systems*, pages 869–873, 2003.

[WCH92] Juyang Weng, Paul Cohen, and Marc Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–980, 1992.

[Wed04] Jan Wedekind. Focus set based reconstruction of micro-objects. In *IEEE International Conference on Mechatronics & Robotics (MechRob'04)*, pages 754–756, 2004.

[WK03] SK Weeratunga and C Kamath. An investigation of implicit active contours for scientific image segmentation. In *Visual Communications and Image Processing Conference*, 2003.

[WP04] M Weck and C Peschke. Equipment technology for flexible and automated micro-assembly. *Microsystem technologies*, 10(3):241–246, 2004.

[WPA09] Medha V Wyawahare, Pradeep M Patil, and Hemant K Abhyankar. Image registration techniques: an overview. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2(3):11–28, 2009.

[WSJ05a] Nathan A Weir, Dannelle P Sierra, and James F Jones. A review of research in the field of nanorobotics. *Sandia Report*, 2005.

[WSJ+05b] Thomas Wich, Torsten Sievers, Marco Jähnisch, Helge Hülsen, and Sergej Fatikow. Nanohandling automation within a scanning electron microscope. In *Industrial Electronics, 2005. ISIE 2005. Proceedings of the IEEE International Symposium on*, volume 3, pages 1073–1078. IEEE, 2005.

[WSN87] Lee Weiss, Arthur C Sanderson, and Charles P Neuman. Dynamic sensor-based control of robots with visual feedback. *IEEE Journal of Robotics and Automation*, 3(5):404–417, 1987.

[YC09] Huei-Fang Yang and Yoonsuck Choe. Cell tracking and segmentation in electron microscopy images using graph cuts. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 306–309. IEEE, 2009.

[YDS+99] MinFeng Yu, Mark J Dyer, George D Skidmore, Henry W Rohrs, XueKun Lu, Kevin D Ausman, James R Von Ehr, and Rodney S Ruoff. Three-dimensional manipulation of carbon nanotubes under a scanning electron microscope. *Nanotechnology*, 10(3):244, 1999.

[YGN01] Ge Yang, James A Gaines, and Bradley J Nelson. A flexible experimental workcell for efficient and reliable wafer-level 3d micro-assembly. In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 1, pages 133–138. IEEE, 2001.

[YJB05] Q Yang, S Jagannathan, and EW Bohannan. Block phase correlation-based automatic drift compensation for atomic force microscopes. In *5th IEEE Conference on Nanotechnology*, pages 370–373. IEEE, 2005.

[YJB08] Q Yang, S Jagannathan, and EW Bohannan. Automatic drift compensation using phase correlation method for nanomanipulation. *IEEE Transactions on Nanotechnology*, 7(2):209–216, 2008.

[YN03] Ge Yang and Bradley J Nelson. Wavelet-based autofocusing and unsupervised segmentation of microscopic images. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 3, pages 2143–2148. IEEE, 2003.

[YN05] Kemal Berk Yesin and Bradley J Nelson. A cad model based tracking system for visually guided microassembly. *Robotica*, 23(4):409–418, 2005.

[YOS93] TTE Yeo, SH Ong, and R Sinniah. Autofocusing for tissue microscopy. *Image and vision computing*, 11(10):629–639, 1993.

[ZF03] Barbara Zitova and Jan Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.

[Zha00] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.

[ZXQ+09] Yong Zhu, Feng Xu, Qingquan Qin, Wayne Y Fung, and Wei Lu. Mechanical properties of vapor- liquid- solid synthesized silicon nanowires. *Nano letters*, 9(11):3934–3939, 2009.

# Index

**Résumé :**

Avec les dernières avancées en matière de nanotechnologies, il est devenu possible de concevoir, avec une grande efficacité, de nouveaux dispositifs et systèmes nanométriques. Il en résulte la nécessité de développer des méthodes de pointe fiables pour la nanomanipulation et la nanocaractérisation. La détection directe par l'homme n'étant pas une option envisageable à cette échelle, les tâches sont habituellement effectuées par un opérateur humain expert à l'aide de microscope électronique à balayage équipé de dispositifs micro-nanorobotiques. Toutefois, en raison de l'absence de méthodes efficaces, ces tâches sont toujours difficiles et souvent fastidieuses à réaliser. Grâce à ce travail, nous montrons que ce problème peut être résolu efficacement jusqu'à une certaine mesure en utilisant les informations extraites des images. Le travail porte sur l'utilisation des images électroniques pour développer des méthodes automatiques fiables permettant d'effectuer des tâches de nanomanipulation et nanocaractérisation précises et efficaces. En premier lieu, puisque l'imagerie électronique à balayage est affectée par les instabilités de la colonne électronique, des méthodes fonctionnant en temps réel pour surveiller la qualité des images et compenser leur distorsion dynamique ont été développées. Ensuite des lois d'asservissement visuel ont été développées pour résoudre deux problèmes. La mise au point automatique utilisant l'asservissement visuel, développée, assure une netteté constante tout au long des processus. Elle a permis d'estimer la profondeur inter-objet, habituellement très difficile à calculer dans un microscope électronique à balayage. Deux schémas d'asservissement visuel ont été développés pour le problème du nanopositionnement dans un microscope électronique. Ils sont fondés sur l'utilisation directe des intensités des pixels et l'information spectrale, respectivement. Les précisions obtenues par les deux méthodes dans différentes conditions expérimentales ont été satisfaisantes. Le travail réalisé ouvre la voie à la réalisation d'applications précises et fiables telles que l'analyse topographique, le sondage de nanostructures ou l'extraction d'échantillons pour microscope électronique en transmission.

**Mots-clés :** Microscopie électronique, nanomanipulation, nanopositionnement, asservissement visuel

**Abstract:**

With the latest advances in nanotechnology, it became possible to design novel nanoscale devices and systems with increasing efficiency. The consequence of this fact is an increase in the need for developing reliable and cutting edge processes for nanomanipulation and nanocharacterization. Since the human direct sensing is not a feasible option at this particular scale, the tasks are usually performed by an expert human operator using a scanning electron microscope (SEM) equipped with micro-nanorobotic devices. However, due to the lack of effective processes, these tasks are always challenging and often tiresome to perform. Through this work we show that, this problem can be tackled effectively up to an extent using the microscopic vision information. It is concerned about using the SEM vision to develop reliable automated methods in order to perform accurate and efficient nanomanipulation and nanocharacterization. Since, SEM imaging is affected by the non-linearities and instabilities present in the electron column, real time methods to monitor the imaging quality and to compensate the time varying distortion were developed. Later, these images were used in the development of visual servoing control laws. The developed visual servoing-based autofocusing method ensures a constant focus throughout the process and was used for estimating the inter-object depth that is highly challenging to compute using a SEM. Two visual servoing schemes were developed to perform accurate nanopositioning using a nanorobotic station positioned inside SEM. They are based on the direct use of global pixel intensities and Fourier spectral information respectively. The positioning accuracies achieved by both the methods at different experimental conditions were satisfactory. The achieved results facilitate in developing accurate and reliable applications such as topographic analysis, nanoprobing and sample lift-out using SEM.



**Keywords:** Electron microscopy, nanomanipulation, nanopositioning, visual servoing

■ École doctorale SPIM - 16 route de Gray F - 25030 Besançon cedex

■ tél. +33 (0)3 81 66 66 02 ■ ed-spim@univ-fcomte.fr ■ www.ed-spim.univ-fcomte.fr