

Multimodality image registration for visualization in robotic assisted breast biopsy

T. (Trishia) Chemaly

BSc Report

Committee:

Prof.dr.ir. S. Stramigioli
Dr. F.J. Siepel
Dr. S. Rihana
V. Groenhuis, MSc

May 2017

010RAM2017
Robotics and Mechatronics
EE-Math-CS
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands

Acknowledgement

I am using this opportunity to express my gratitude to everyone who supported me throughout my bachelor thesis. Without these individuals, this project wouldn't have been as successful.

It is with greatest honor and appreciation that I thank the University of Twente for granting me this golden, long awaited opportunity to take part in the MURAB project. Thanks to Dr. Françoise Siepel and to my supervisor Vincent Groenhuis for accepting me as a new member in the RaM research group and for guiding me yet providing me with freedom to pursue my part of the project. I portray all my respect and gratitude to Dr. Ferdi van der Heijden for the generous time he offered to enrich me with his expertise in stereo vision and valuable guidance and comments that opened my mind and straightened my work. It was a privilege and once-in-a-lifetime opportunity to work under the supervision of such professionals.

Moreover, I would like to acknowledge the Holy Spirit University of Kaslik, precisely the Faculty of Engineering led by its dean Dr. Barbar Zeghondy for giving me the acceptance to take this step in my Biomedical Engineering studies. I am also grateful to former dean, Mr. Pascal Damien, for his persistent encouragement and acknowledgement of my efforts and capabilities since day one at the university. I wouldn't be here today if it weren't for his guidance.

Furthermore, I would like to express my deepest gratitude to Dr. Sandy Rihana, head of the Biomedical Engineering department at the Holy Spirit University of Kaslik for being a unique mentor to look up to throughout the past five years of my studies. Her generous guidance, expertise, and assistance have paved the way for me choosing this project today. Her consistent instruction from a long distance provided me with more than the support and enlightenment I needed to accomplish this project.

With much respect, I would also like to thank the members of the jury for their valuable time and meticulous remarks.

Last but not least, I would like to thank my family and friends for their consistent moral support, encouragement, and honest advice.

Abstract

Breast cancer, the most commonly diagnosed cancer in women worldwide, is mostly detected through a biopsy where tissue is extracted and chemically examined or pathologist assessed under a microscope. Biopsy robots have been introduced, aiming towards increased precision. The MURAB project aims to revolutionize the way cancer screening and muscle diseases are researched and has the potential to save lives. It intends to create a new paradigm in which, the precision of medical imaging modalities such as MRI and Ultrasound are combined with the precision of robotics in order to target specific areas in the body.

Medical imaging plays a valuable role in targeting malignant tissue accurately and guiding the radiologist during needle insertion in a biopsy. This paper proposes a computer software that can process and combine 3D reconstructed surfaces from different imaging modalities, particularly MRI and camera, showing a visualization of important features during scanning and biopsy in real-time and investigates its feasibility. It deals with the reconstruction, segmentation, transformation, and registration of MRI scans and camera images of the breast, to assist in the MRI/US image fusion part of MURAB. The development of this software promotes exact targeting of (small) lesions, which are visible under MRI thus improving the clinical workflow and contributing to the achievement of MURAB goals. It also aims to combine the detectability of MRI with the physical space of the camera. This study demonstrates that the registration accuracy of the proposed system is acceptable and has potential of clinical application in MURAB.

Table of Contents

Acknowledgement	i
Abstract	ii
List of Figures	v
List of Tables	vii
List of Flow Charts	viii
Acronyms	ix
1 Introduction	1
1.1 Context.....	1
1.2 Overview of the MURAB project.....	1
1.3 Aim of this Project.....	2
1.4 Workflow Overview	2
1.5 Report Organization.....	3
2 Literature Review	4
2.1 Breast Cancer	4
2.2 Breast Biopsy	4
2.3 Biopsy Robots.....	4
2.4 Multi-modal Image Registration.....	5
3 MRI Reconstruction	10
3.1 Markers	10
3.2 Flow Chart	11
3.3 MRI Reconstruction and Enhancement	12
3.4 3D Volume Visualization.....	15
3.5 Markers Localization	17
3.6 Analysis of Results	20
4 Design of the Camera Setup	23
4.1 Methods.....	24
4.2 Materials	24
4.3 Choice of Lenses.....	24
4.4 Distance Range Analysis	25
4.5 Choice of Camera Setup	27
4.6 Monte Carlo Analysis	29
4.7 End effector constraint.....	31
4.8 Conclusion	32
5 Camera Reconstruction	33

5.1	Theory	33
5.2	Implementation	46
5.3	Experimentation and Results	51
6	Registration	60
6.1	Theory and Implementation	60
6.2	Results.....	63
Conclusion and Future Recommendations.....		67
Bibliography		69

List of Figures

Figure 1: MURAB Goals [1]	2
Figure 2: Classification Criteria of Registration Methods	6
Figure 3: Marker type a (left) versus marker type b (right)	10
Figure 4: Breast phantom MRI scanning	11
Figure 5: Original single DICOM Image	12
Figure 6: Bright stretching method [13]	13
Figure 7: Original DICOM Image	14
Figure 8: Enhanced Images. First Approach (Left) - Second Approach (Right)	14
Figure 9: Image obtained after binarization	15
Figure 10: 3D Volume of the breast and markers type a	16
Figure 11: 3D Volume of the breast and markers type b	16
Figure 12: Segmented markers - type a (left) versus type b (right)	17
Figure 13: Image Coordinate System [14]	18
Figure 14: Anatomical coordinate system (left) versus Image coordinate system (right) [14]	18
Figure 15: Affine transformation [14]	19
Figure 16: Top view of camera setup	25
Figure 17: Circle of Confusion	26
Figure 18: Diameter of circle of confusion as a function of working distance	27
Figure 19: Top view of the plot of blur circles as a function of working distance	27
Figure 20: Parallel camera setup (left) versus convergent camera setup (right)	28
Figure 21: Camera positions and verging angle	29
Figure 22: 3D estimated coordinates over 200 Monte Carlo repetitions for different baselines	31
Figure 23: Parallel camera setup (left) versus converging camera setup (right)	32
Figure 24: Pinhole camera model	34
Figure 25: Test image (left) - Radially distorted image (right)	36
Figure 26: Calculating reprojection error [19]	37
Figure 27: Stereo aligned camera set-up [20]	38
Figure 28: Triangulation (1) [20]	41
Figure 29: Triangulation (2) [20]	42
Figure 30: Triangulation(3)	42
Figure 31: Gaussian blur [21]	44
Figure 32: Gaussian kernel for sigma=1 [22]	44
Figure 33: RGB color space (left) versus HSV color space (right)	45
Figure 34: Checkboard calibration images	51
Figure 35: Stereo camera calibration with checkerboard images. (Green circles represent the detected points while red points represent the reprojected points)	51
Figure 36: Graph of reprojection errors as a function of image pairs	52
Figure 37: Stereo camera extrinsics- 1: Camera 1, 2: Camera 2 – The colored planes observed in front of the 2 cameras represent the checkerboard orientations during calibration	52
Figure 38: Rectified image pairs from side view of breast phantom	53
Figure 39: Rectified image pairs from upper view of breast phantom (1)	53
Figure 40: Rectified image pairs from upper view of breast phantom (2)	54
Figure 41: SURF features detected in side view rectified images	54
Figure 42: Disparity map before SURF and range definition (left) versus disparity map after (right)	55
Figure 43: Disparity map before cleaning (left) versus disparity map after (right)	55
Figure 44: Point cloud before cleaning (left) versus point cloud after (right)	56
Figure 45: Left and right original images respectively	56
Figure 46: Left and right undistorted images respectively	57

Figure 47: Left and right gaussian smoothed images respectively	57
Figure 48: Markers localization in left and right camera images respectively	57
Figure 49: Two point clouds reconstructed from two different pairs of images	63
Figure 50: Camera-camera surface registration of 2 point clouds before ICP (left) and after ICP (right) .	63
Figure 51: Camera-camera registration.....	64
Figure 52: MRI-camera registration before ICP	65
Figure 53: MRI-camera matching after ICP	65

List of Tables

Table 1: Centroids Comparison Table	21
Table 2: Error Computation	21
Table 3: Final markers localization results	22
Table 4: Covariance in 3D coordinates estimation as a function of baseline	30
Table 5: 2D pixel coordinates of detected markers in left and right camera images	58
Table 6: 3D mm coordinates of detected markers	59
Table 7: Registration Validation	66

List of Flow Charts

Flow Chart 1: MRI algorithm overview	11
Flow Chart 2: 3D Reconstruction	46
Flow Chart 3: Camera Calibration	47
Flow Chart 4: Disparity Map	48
Flow Chart 5: Markers detection and localization in image	49
Flow Chart 6: Markers 3D localization.....	50
Flow Chart 7: ICP	62

Acronyms

- **ABS** Acrylonitrile Butadiene Styrene
- **DOE** Diffraction Optical Element
- **FOV** Field of View
- **FRI** Fast Research Interface
- **HSV** Hue Saturation Value
- **KCP** KUKA Control Panel
- **KUKA** Keller und Knappich Augsburg
- **LPS** Left Posterior Superior
- **LWR** Light Weight Robot
- **MURAB** MRI and Ultrasound Robotic Assisted Biopsy
- **MSE** Mean Square Error
- **RaM** Robotics and Mechatronics
- **RAS** Right Anterior Superior
- **RCM** Remote Center of Motion
- **RGB** Red Green Blue
- **RMSE** Root Mean Square Error
- **ROS** Robot Operating System
- **SLAM** Simultaneous Localization and Mapping
- **SURF** Speeded Up Robust Features
- **US** Ultrasound
- **ZGT** Ziekenhuisgroep Twente

1 Introduction

1.1 Context

In 2012, about 1.7 million new cases were diagnosed with breast cancer, the most common cancer diagnosed in females around the world. This signifies nearly 25% of all cancers in women and 12% of all new cancer cases [1]. Being the only procedure that can definitely diagnose cancer, a biopsy is a popular imaging-guided, minimally invasive procedure for the examination of suspicious areas.

1.2 Overview of the MURAB project

MURAB stands for MRI and Ultrasound Robotic Assisted Biopsy and is a four-year European project that is part of the Horizon 2000 Framework for Research and Innovation Program. The project includes partners from the University of Verona in Italy, Medical University of Vienna in Austria, Radboud Academic Medical Centre in Nijmegen, ZGT hospital Hengelo and Siemens in the Netherlands, and KUKA Industrial Robotics in Germany with the University of Twente as the leading partner. The aim of the MURAB project is to enhance breast biopsy, specifically addressing issues in MRI-guided biopsy. Instead of multiple MRI imaging repetitions required in the conventional procedure, only one MRI image is taken in the MURAB procedure. The latter can be summarized into the following phases:

- **PHASE 1: Scanning**
 - Autonomous US scanning of the suspicious area by a robot arm
 - Acquisition of elastographic data (US transparent pressure sensor array is used)
- **PHASE 2: Tissue Active SLAM**
 - Lesion localization by MRI – US image merging
 - Elastography model construction (For prediction of lesion movement)
- **PHASE 2: Needle Insertion**
 - Autonomous US needle tracking
 - Autonomous needle navigation
 - Manual needle insertion by the medical professional (He must only penetrate the needle through the skin by applying a constant force in one direction.)

- Virtual stop when target has been reached

For the procedure, a KUKA robotic arm is used.

The goals of the MURAB project can be summarized in Figure 1.

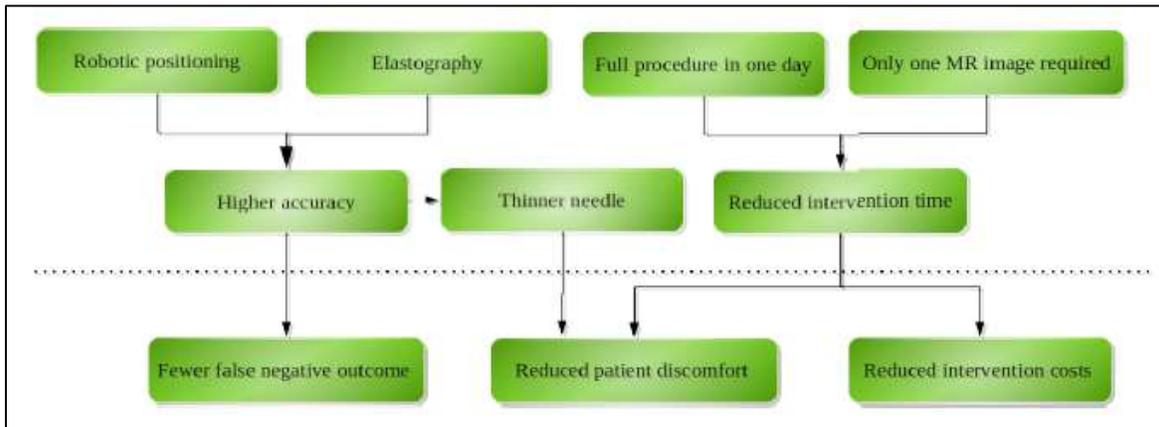


Figure 1: MURAB Goals [1]

1.3 Aim of this Project

The main focus of this final year project is the development of a computer software that can reconstruct the surface of breast phantoms imaged with different modalities, particularly MRI and stereo camera then process and fuse the multiple surfaces. This fusion should allow the visualization and localization of important features, specifically markers. The ultimate goal is to combine MRI lesion detectability with stereo vision localization in the physical space in order to achieve accurate needle placement and tissue extraction in the future. The software deals with the segmentation, registration, and transformation of MRI scans and camera images of the breast, to assist in the MRI/Camera/US image fusion part of MURAB. It should promote exact targeting of (small) lesions and particular features to improve clinical workflow and contribute to MURAB goals.

1.4 Workflow Overview

The implementation of this project comprises four modules:

- 1) **MRI Surface Reconstruction:** The aim of this module is to create a 3D surface from MRI acquired images of a breast phantom and localize fiducial points, markers, placed on the

phantom to later assist in the registration. The acquired images were enhanced then used in reconstruction. The markers were then segmented based on 3D connectivity and area. They were located by the computation of their centroids.

- 2) **Design of the Camera Setup:** Since a major part of this project utilizes a camera setup for stereo vision, a camera setup had to be designed. This module comprises the optics behind the design while taking into consideration the specifications needed for the project.
- 3) **Camera Surface Reconstruction:** The aim of this module is to create a 3D surface from camera acquired images of a breast phantom and localize fiducial points, markers, placed on the phantom to later assist in the registration. Stereo vision techniques were used in the process. The markers were then segmented based on color and area. They were located by the computation of their centroids in image space and triangulated to obtain the physical space 3D coordinates
- 4) **MRI-Camera Registration:** This final module combines the results of MRI and Camera surface reconstruction by aligning the two surfaces. The initial transformation was estimated from the phantom marker locations and then optimized using Iterative Closest Point (ICP).

1.5 Report Organization

This report comprises seven chapters. In the next chapter, an overview of breast biopsy, biopsy robots, and existing image registration techniques used in surgical navigation is provided. In Chapter 3, the theory behind MRI surface reconstruction and marker localization as well as the implementation and results are portrayed. Next, the camera setup design and optical mathematics behind are detailed. Chapter 5 describes the theory behind stereo vision, camera surface reconstruction, and marker localization as well as the implementation and results. The results of Chapter 3 and 5 are combined in Chapter 6 where the MRI-Camera registration is performed and evaluated. Finally, Chapter 7 discusses the results, and groundwork is set for the future.

2 *Literature Review*

2.1 Breast Cancer

Breast cancer, the disease in which malignant tissue forms in the breast, is the most commonly diagnosed cancer in women worldwide and the second major cause of cancer death among women [2]. As for diagnosis, mammography, an X-Ray imaging technique, is currently the first step to be taken. In case a suspicious lesion is observed on the mammogram, an ultrasound (US) guided biopsy is recommended. However, in case the US is not sufficient to accurately localize the lesion, MRI breast biopsy is indicated. Improving the mentioned biopsy techniques will probably allow early detection and render diagnosis more reliable. Thus, it promises to reduce significantly the mortality rate due to breast cancer.

2.2 Breast Biopsy

A biopsy is the only diagnostic procedure that can confirm whether the suspicious area is cancerous or not. During this test, tissue or fluid is removed from the suspicious area for morphology and histology examination [3]. For the collection of tissue, a needle is inserted, either guided by real-time ultrasound imaging or guided by MRI. MRI is required whenever lesions are easier to visualize in MRI images or whenever patient screening indicates high risk. During MRI breast biopsy, the patient is taken out of the MRI machine after imaging and then taken back in after the needle has been placed to check for correct placement. Sometimes, multiple repetitions of this procedure are needed before the needle has been found to have been inserted correctly. This increases patient discomfort, time and cost, and number of false negative diagnoses.

2.3 Biopsy Robots

As expected in the biomedical engineering field, researchers are always looking for improvements to already existing systems and solutions. Screening programs and high end professional training for radiologists have contributed to fast and accurate biopsies. However, technical solutions such as biopsy robots have been also introduced for further improvements. Biopsy robots are well known for high accuracy, dexterity, and repeatability, characteristics that are much needed for a biopsy [4]. Such robots allow automated needle insertion guided by either remote controlled US

scanning and position tracking of the US probe or 3D US volume reconstruction. In such systems, interaction dynamics between the needle force and the breast tissue cause tumor displacement which results in problems such as patient discomfort, clinician fatigue, and compromising the integrity of the tissue specimen. For that, [5] have designed a system for real-time manipulation of the tumor where the US is manipulated and the needle is inserted simultaneously. On one hand, such systems have reported promising accuracies. On the other hand, they can only focus on one part of the biopsy process such as scanning, lesion detection, or needle insertion. They also limit their focus on US guided biopsy, leaving behind MRI biopsy, the more complex intervention. The MURAB project introduces a solution for the aforementioned issues by developing an unconventional biopsy robot that combines both US and MRI. This solution promises to replace the cost and time inefficient MRI breast biopsy.

2.4 Multi-modal Image Registration

Nowadays, medical image registration is valued as an assistant for experts in medical interventions. The importance of this process arises from its ability to help experts in diagnosis, deciding the necessary therapies, localization of the disease, and eventually guiding the intervention. The purpose behind using image fusion is to create a more comprehensive and detailed output image by aligning a target image to a reference image.

Magnetic Resonance Imaging [6]

MRI is a significant invention in the field of medical imaging. It relies on the orientation of protons in a high magnetic field and the manipulation of these by resonant radio-frequency waves. Then, the return of protons to equilibrium is measured. The output images describe the human body thoroughly and contrast soft tissue in a non-invasive, unparalleled manner.

In medical image registration, there are exhaustive applications such as image regeneration, cancer diagnosis, surgical planning, image guided treatment, and others influenced by MRI.

Current Trends in Medical Image Registration [6]

Over time, diverse methods of image fusion were presented. These methods are classified according to the criteria as illustrated in Figure 2.

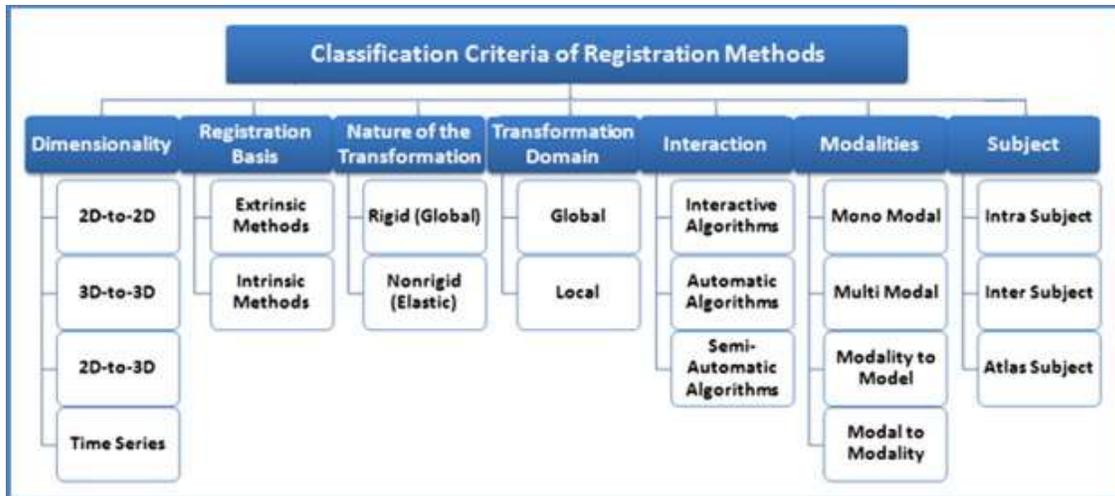


Figure 2: Classification Criteria of Registration Methods

- a) Image dimensionality refers to the number of geometrical dimensions of the image space.
- b) The premise of registration is classified into extrinsic and intrinsic. In extrinsic methods, clearly visible artificial markers are attached to the patient with the necessity of being accurately detectable in all the acquired modalities while intrinsic methods are based on patient provided information, such as voxel image intensities.
- c) The nature of the transformation is divided into two categories: rigid and non-rigid transformations explained in the next section.
- d) The images' coordinate transformation can either be global or local. In the global case, mapping parameters should be valid for the whole image while in the local case, a fraction of the image is transformed, and the mapping parameters are solely valid for a small zone around a selected control point.
- e) Three levels of interaction exist based on user-process relation. In interactive algorithms, the user accomplishes the registration task by feeding the software with the initial transformation parameters. On the other hand, automatic algorithms work independently of any interaction. Finally, in semi-automatic algorithms, the user performs the algorithm initialization through segmenting the data or driving the algorithm to the desired solution.
- f) Four types of registration tasks exist based on the modalities used. In mono-modal tasks, the images of the same medical modality are registered, while in multi-modal tasks, the images involved in the registration process belong to different modalities. There are also modality-to-model and model-to-modality registration tasks where only one image is included while a model

represents the other registration input.

g) Registration methods can also be grouped into intra, inter, and atlas subject registration depending on whether the involved images belong to the same patient, to different patients, or an information database.

Types of Transformations

1) Rigid

A rigid transformation is a linear transformation. It is composed of a translation vector and a rotation matrix and could be defined by Equation 2.1.

$$\vec{p}_2 = \vec{t} + \vec{s}R\vec{p}_1 \quad \text{Equation 2.1}$$

where

$$\vec{p}_1 = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \vec{p}_2 = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \quad \vec{s}_1 = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} \quad \vec{t}_1 = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \quad R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

t(Translation), R(Rotation), s(Scale)

In application, a rigid body registration would allow for only 6 degrees of freedom at most. The 6 degrees of freedom in rigid body registration are translation along the x, y, and z dimensions and rotation around the x, y, and z axes. These preserve the shape and size of the object, but move it around in space.

2) Non rigid

A non-rigid transformation is nonlinear. It could be described as stretching or shrinking. It also allows warping (twisting) of the image.

In application, it causes a distortion of the image by a vertical or horizontal stretch. For example, it allows dilation of the image.

3) Affine

An affine transformation consists of rotation, translation, scaling, and shearing¹.

In application, an affine registration is a linear transformation. It allows for up to 12 degrees of freedom so the scale of the object could be altered or shear could be introduced. It allows us to change position and the geometry of the object. There is no preservation of lengths and angles, but parallel lines are preserved.

2.5 Computer Vision in Surgical Navigation

In MRI and US guided biopsies, the imaging guidance only gives a rough orientation. Consequently, the radiologist estimates the 3D trajectory of the needle insertion on the 2D slices mentally. This is where the importance of surgical navigation is highlighted in order to achieve accurate instrument placement, and it has been widely used in orthopedics and neurosurgery. However, this approach is not yet popular in breast biopsy due to breast movement and deformations.

Physical-to-image space registration is the key technique in surgical navigation. Traditionally, fiducial points or markers are placed on the skin and extracted manually or automatically with detection methods. Nicolau et al [7] used stereo vision to calculate coordinates of markers in physical space. Maier-Hein et al [8] inserted traceable needles with infrared reflective spheres into the abdomen and located them by an optical tracking system. Apparently, detection of fiducials optimizes the procedure of the intervention.

With the development of imaging, markerless registration has also gradually emerged. For example, point cloud registration in open liver surgery has been performed using 3D laser scanning [9]. Nevertheless, such scanners are expensive and cannot acquire surface information in real time.

More recently, depth imaging setups, such as CamCube have become favored. However, their imaging accuracy is lower than that attained with range scanners. Still, the upper hand of these depth cameras is their ability to image in real time and acquire intensity(grayscale) or color (RGB) images simultaneously [10].

¹ Shear: A transformation in which all points along a given line L remain fixed while other points are shifted parallel to L by a distance proportional to their perpendicular distance from L.

A Kinect™ camera based navigation system for percutaneous abdominal puncture [10]

A project has been conducted where the utility of the second generation Kinect has been investigated for guiding needle insertion in abdominal puncture navigation which is an interventional method for the management of abdominal tumors. For Kinect™ surface extraction, the depth image was masked with the abdominal region of interest, and then adjusted and transformed into the 3D world coordinate system based upon a pin-hole camera model. Using Iterative Closest Point (ICP), the abdominal surface extracted from camera images was registered to the surface extracted from CT. ICP based registration would be invalid if the initial alignment is distant from its solution. Manual matching is a common method for obtaining the initial estimate transformation [11]. Another method of feature extraction from multimodality surfaces to establish correspondence was proposed. It deals with the alignment of a CT surface and camera surface with nearly 1 cm accuracy. The registration is then robustly optimized with ICP [12]. However, in [10], the abdominal surface is too smooth, which eliminates the possibility of applying a feature-based matching technique. To estimate a first transformation before ICP, a corresponding point searching algorithm based on a 2D image was proposed. The validation was then based on 48 insertions on a general abdominal phantom and 15 insertions on six beagles, and the corresponding navigation errors were evaluated. The mean target positioning error (TPE) obtained with the phantom experiments was 5.23 mm while the mean TPE obtained with the animal experiments was 6.4 mm. The author concluded that there is hope that the Kinect™ V2-based guidance system be accepted as a clinically approved application since the navigation error of the needle tip does not exceed the limit of 5 mm, thus avoiding destroying too much healthy tissue during the procedure.

3 *MRI Reconstruction*

Reconstructing the 3D surface of the breast from MRI acquired images is a requisite part of this project. It promotes the possibility of accurately localizing malignant tissue in 3D and eventually guiding the robot end effector towards the target where the biopsy needle will be inserted by the radiologist. The shape of the breast is very close to being symmetrical. Hence, localization and eventually registration with other modalities could be misleading in the absence of guiding features. For that, MRI visible markers are placed on the breast phantom.

The main goal of the first phase of this final year project was to reconstruct a 3D volume from acquired MRI DICOM images of a breast phantom and to be able to automatically locate and segment the markers placed on that phantom.

Certain requirements are taken into consideration:

- The accuracy of the surface reconstruction should be at least 1 mm.
- The surface reconstruction should be obtained in anatomical space.

3.1 **Markers**

The implementation was applied to acquisitions obtained from phantoms with two different types of markers glued to the breast phantom with blue Silicone:

- a) Red and purple 3D printed cylindrical shaped markers filled with fish oil
- b) Green 3D printed circular shaped markers filled with Vaseline

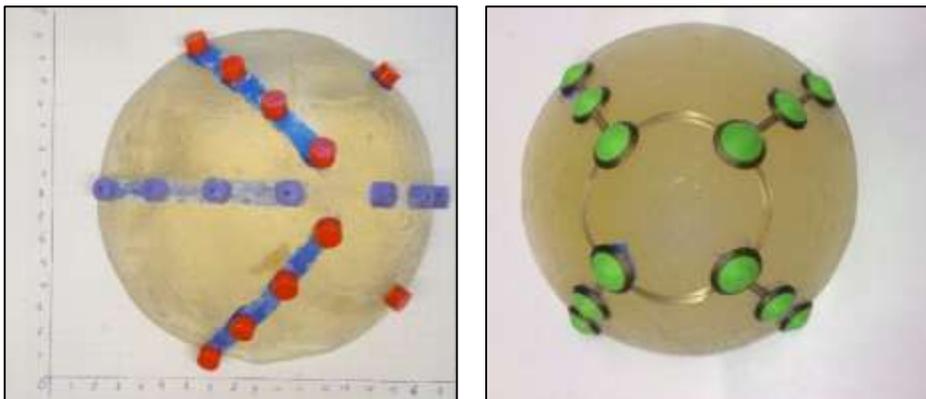


Figure 3: Marker type a (left) versus marker type b (right)²

² In the rest of this document, the left markers will be designated as type a and right markers as type b.

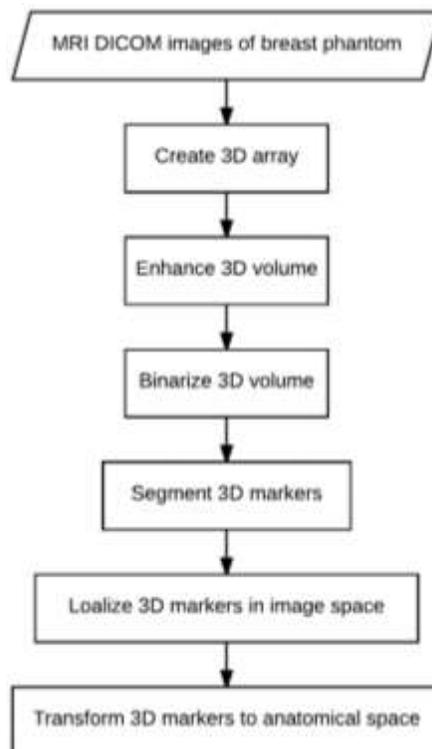


Figure 4: Breast phantom MRI scanning

Fish oil and Vaseline were chosen as fillings for the markers due to their visibility in MRI. Fish oil filled markers have a higher intensity in the MRI images than Vaseline. However, Vaseline is semisolid while fish oil is liquid and leaks out of the markers.

3.2 Flow Chart

In this Chapter, MRI DICOM images will be manipulated in the aim of extracting the 3D MRI coordinates of the markers placed on the breast phantom. Flow Chart 1 gives an overview of the algorithm involved in the process.



Flow Chart 1: MRI algorithm overview

3.3 MRI Reconstruction and Enhancement

3.3.1 Reading DICOM MRI Slices

Since every image is represented by a 2D matrix, the 3D volume could be represented by a 3D array formed by stacking the 2D matrices. At first, an empty 3D array whose size is $512 \times 512 \times N$ (N images, each having the size 512×512) is created. Then, the 3D array is filled with the images.

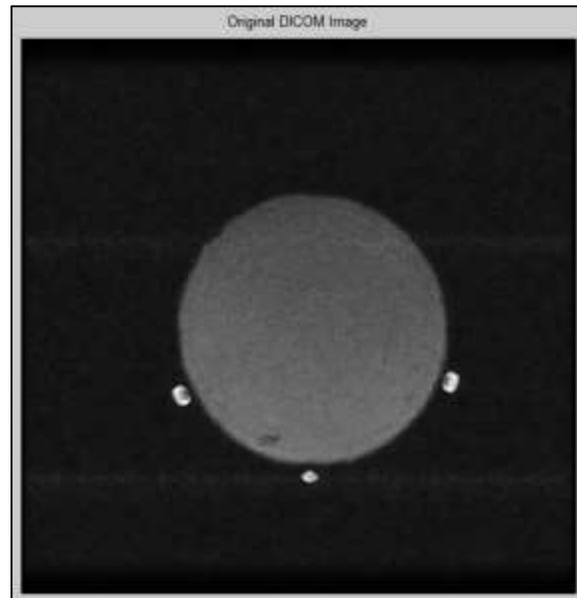


Figure 5: Original single DICOM Image

3.3.2 Image Enhancement

First Approach

A manual approach was considered to remove background noise. The maximum and minimum grayscale values in the image were defined. By manual investigation, any grayscale value less than 500 corresponds to a black pixel. For that, any pixel with a grayscale value less than 500 was replaced with a value of 0 to make it as black as possible and facilitate later manipulation. Then, image contrast was enhanced.

Second approach

In this approach, linear interpolation and bright stretching were used for enhancement of the images forming the 3D array, hence enhancement of the 3D volume.

❖ Linear Interpolation

One of the methods of intensity resolution enhancement is the intensity linear transformation which applies Equation 3.1 to all the image pixels.

$$\rho' = \frac{\rho - \rho_{min}}{\rho_{max} - \rho_{min}} \times \omega_{target} + \rho_{min} \quad \text{Equation 3.1}$$

ρ' : Transformed pixel intensity

ρ : Original pixel intensity

ρ_{min}, ρ_{max} : Minimum and maximum gray values in the original image

ω_{target} : Range of the target intensity space = $2^n - 1$ where n is the number of bits

❖ Bright Stretching

This method is used to enhance the brighter part of the images and is applied using Equation 3.2 .

$$\rho' = \begin{cases} \frac{\rho}{TH} \times SFB & \text{for } \rho < TH \\ \frac{\rho - TH}{255 - TH} \times (255 - SFB) + SFB & \text{for } \rho > TH \end{cases} \quad \text{Equation 3.2}$$

TH : Threshold value (Chosen manually as minimum value in the markers)

SFB : Bright -stretching factor, should be smaller than TH

In the figure below, we can see the stretching and compression methods for the bright-stretching technique. As shown, the pixel values, which are less than the threshold value, will be compressed while the pixel values, which are greater than the threshold value, will be stretched.

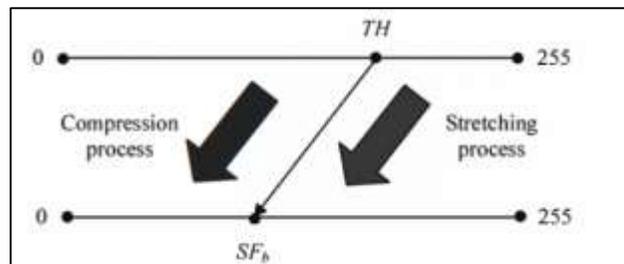


Figure 6: Bright stretching method [13]

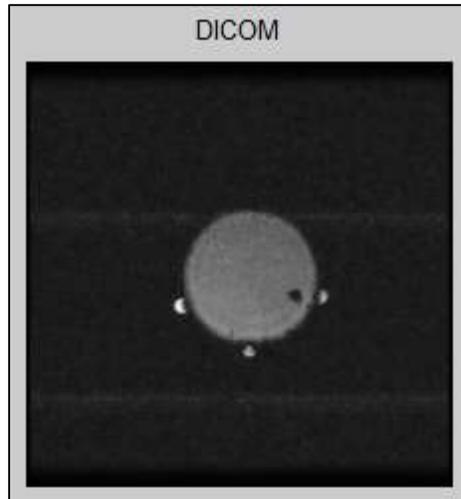


Figure 7: Original DICOM Image

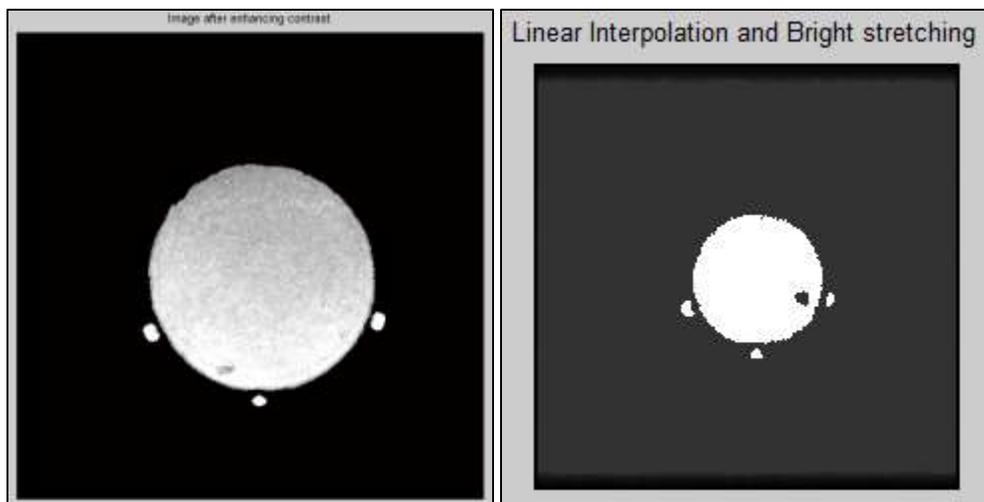


Figure 8: Enhanced Images. First Approach (Left) - Second Approach (Right)

As observed in Figure 8, the first approach modified the original image, removing and adding nonexistent form. For that, the second approach was chosen.

3.3.3 Image Binarization

Otsu's method, invented by Nobuyuki Otsu, is one of various binarization algorithms used to convert a greyscale image to monochrome. This method iterates through all possible threshold values in the image and calculates a measure of spread for the pixel levels each side of the threshold (pixels that either belong to the foreground or background of the image). In this case, the phantom and markers should be in the foreground.

The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum. This method finds the threshold that minimizes the weighted within-class variance

and maximizes the between-class variance

Otsu's method comprises an exhaustive search for the threshold that minimizes the intra-class variance (the variance within the class), defined as a weighted sum of variances of the two classes (Equation 3.3).

$$\sigma_w^2(t) = w_0(t)\sigma_0^2(t) + w_1(t)\sigma_1^2(t) \quad \text{Equation 3.3}$$

Applying Otsu in Matlab using `imbinarize` is done. It detects the most adequate threshold value. Then the binarization is applied.

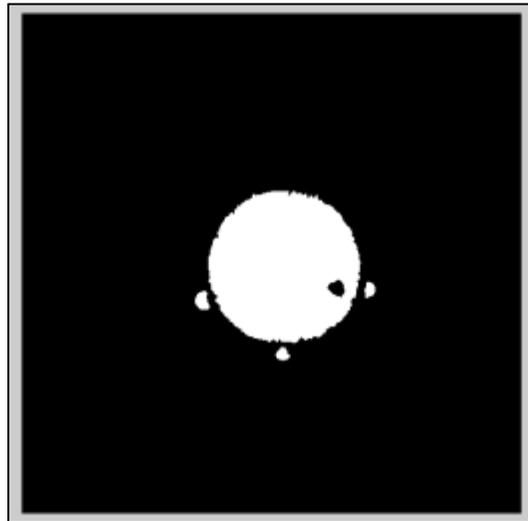


Figure 9: Image obtained after binarization

3.4 3D Volume Visualization

An isosurface is a level set of a continuous function such that this function's domain is 3D space. For the 3D MRI volume visualization, the isosurface method is used. In this method, an object is reconstructed, having its boundaries from points in the 3D array that hold the same value. This is achieved by connecting these points. That value is regarded as a threshold and is called the isovalue. In our case, the white pixels having a binary value of 1 in the 3D array will be used to form the boundaries of the isosurface and visualize the 3D volume.

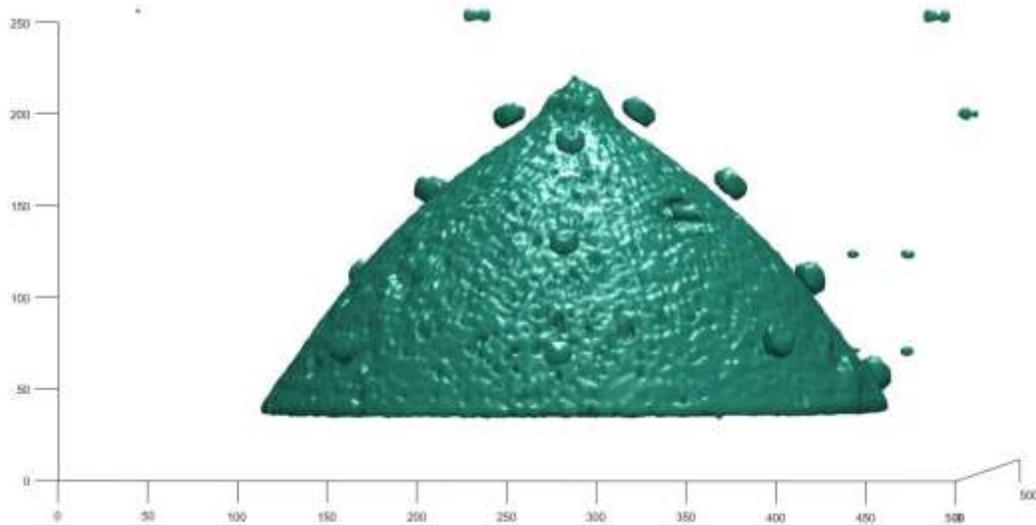


Figure 10: 3D Volume of the breast and markers type a

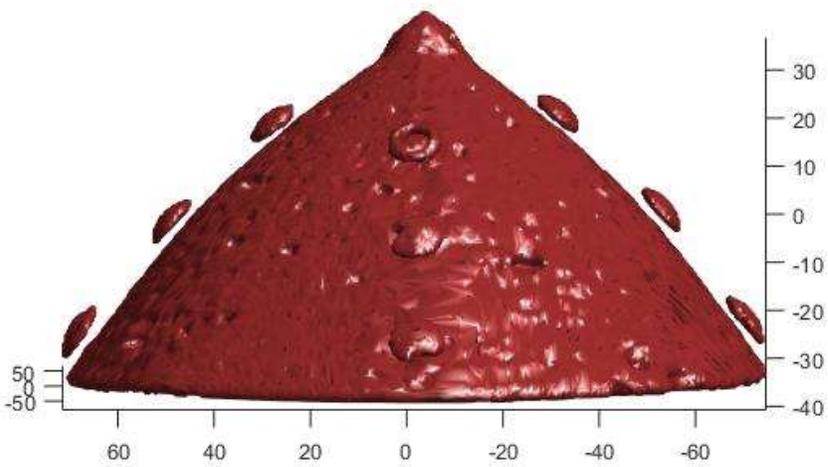


Figure 11: 3D Volume of the breast and markers type b

3.5 Markers Localization

3.5.1 Object Identification

The different objects in the 3D volume were detected based on their connectivity. Connectivity is the way in which voxels or pixels relate to their neighbors. In 3D object detection, 26-connectivity is mainly used where 26-connected pixels are neighbors to every pixel that touches one of their faces, edges, or corners.

3.5.2 Marker Identification and Segmentation

Out of all the blobs identified in our 3D volume, only 12 are markers. To differentiate the markers, area criterion is used knowing that the markers' area is less than that of the phantom but more than any residual noise area.

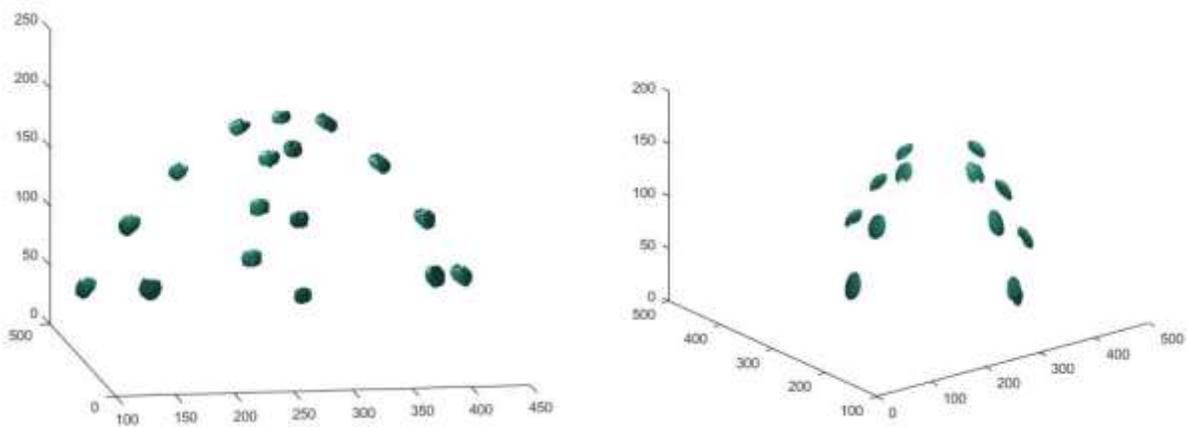


Figure 12: Segmented markers - type a (left) versus type b (right)

3.5.3 Marker Centroid Computation

The markers' centroid coordinates are extracted in the image pixel coordinate frame. These image pixel coordinates must be converted to the MRI coordinate frame. For that, we must first define the two coordinate systems.

- 1) **Image coordinate system:** It describes the way the image was taken with respect to the anatomy. Usually, an array that starts at the upper left corner is obtained. That array is characterized by the origin and spacing seen in Figure 13.

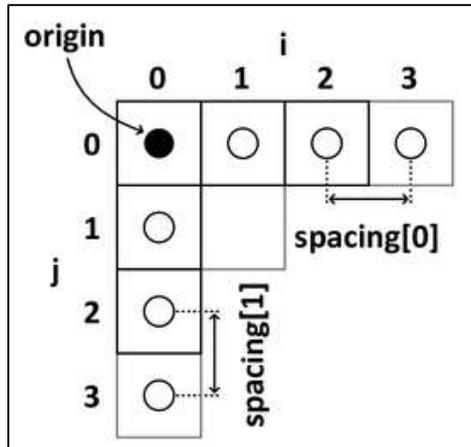


Figure 13: Image Coordinate System [14]

- 2) **Anatomical coordinate system:** It is the most important coordinate system for medical techniques. It consists of the 3 planes: axial, coronal, and sagittal. It has a 3D basis that differs among medical applications.

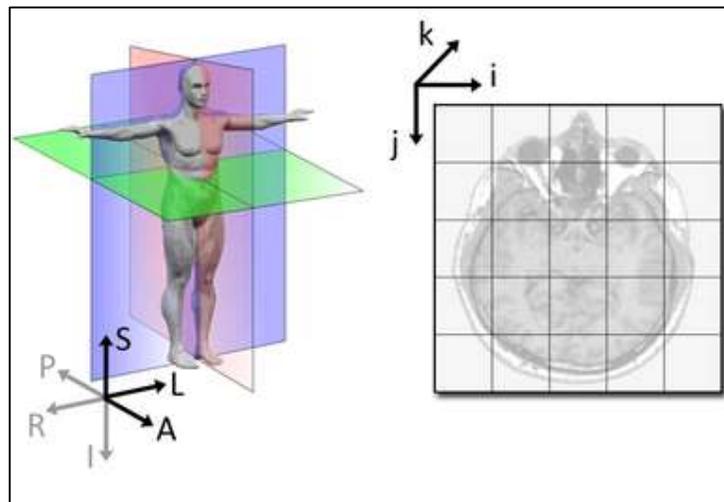


Figure 14: Anatomical coordinate system (left) versus Image coordinate system (right) [14]

RAS Coordinate System

Since the software 3D Slicer will be used to validate marker coordinates, the RAS system was used.

An affine transformation is used to transform the image coordinate system to the anatomical coordinate system:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} \begin{pmatrix} i \\ j \\ k \end{pmatrix} + \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}$$

Figure 15: Affine transformation [14]

A: A matrix that carries all information about space directions and axis scaling. In 3D, A is a 3-by-3 array. To determine its 9 unknowns, a system of 9 equations is solved.

t: Geometric position of the first voxel/pixel

a) Definition of some known points

The geometric position of the first voxel in anatomical space is defined as (**R_origin, S_origin, A_origin**) and the spacing in mm is defined as (**R_spacing, S_spacing, A_spacing**).

(0,0,0) in image space corresponds to (**R_origin, S_origin, A_origin**) in mm in anatomical space
(1,0,0) in image space corresponds to (**R_origin - R_spacing, S_origin, A_origin**) in mm in anatomical space

(0,1,0) in image space corresponds to (**R_origin, S_origin-S_spacing, A_origin**) in mm in anatomical space

(0,0,1) in image space corresponds to (**R_origin, S_origin, A_origin + A_spacing**) in mm in anatomical space

b) Definition of system equations

The system of 9 equations obtained from the points defined in a) is given below.

$$\mathbf{R_origin - R_spacing} = \mathbf{A11*1 + A12*0 + A13*0 + R_origin}$$

$$\mathbf{S_origin} = \mathbf{A21*1 + A22*0 + A23*0 + S_origin}$$

$$\mathbf{A_origin} = \mathbf{A31*1 + A32*0 + A33*0 + A_origin}$$

$$\mathbf{R_origin} = \mathbf{A11*0 + A12*1 + A13*0 + R_origin}$$

$$\mathbf{S_origin - S_spacing} = \mathbf{A21*0 + A22*1 + A23*0 + S_origin}$$

$$\mathbf{A_origin} = \mathbf{A31*0 + A32*1 + A33*0 + A_origin}$$

$$\mathbf{R_origin} = \mathbf{A11*0 + A12*0 + A13*1 + R_origin}$$

$$\mathbf{S_origin} = \mathbf{A21*0} + \mathbf{A22*0} + \mathbf{A23*1} + \mathbf{S_origin}$$

$$\mathbf{A_origin} + \mathbf{A_spacing} = \mathbf{A31*0} + \mathbf{A32*0} + \mathbf{A33*1} + \mathbf{A_origin}$$

c) Definition of transformation matrix

Solving for the equations in b) yields the desired transformation matrix from the image coordinate system (i,j,k) to the anatomical coordinate system (R,A,S).

$$\begin{pmatrix} R \\ S \\ A \end{pmatrix} = \begin{pmatrix} -R_spacing & 0 & 0 \\ 0 & -S_spacing & 0 \\ 0 & 0 & A_spacing \end{pmatrix} + \begin{pmatrix} R_origin \\ S_origin \\ A_origin \end{pmatrix} \quad \text{Equation 03.4}$$

3.6 Analysis of Results

3.6.1 Performance Evaluation

To evaluate the performance of the marker detection algorithm, the software 3D Slicer was used. However, the only way for detecting the marker centroids in 3D Slicer would be in 2D by selecting different image slices and detecting the markers in those slices. The algorithm was applied on 20 different 2D DICOM slices to validate its performance. Table 1 shows (i,j) and (R,S) coordinates of the markers' centroids as obtained from the algorithm and (R,S) coordinates as manually extracted on Slicer.

	Slice Number	Slice Location (mm)	Number of Visible Markers	(i,j)	(R,S)	(R,S) Manually extracted using Slicer
1	104	7.2	3	[148.09,329.76] [266.56,404.48] [396.43,317.28]	[44.26;-30.53] [-4.33;-61.24] [-57.6;-25.40]	[44.8;-30.1] [-3.8;-60.4] [-57.4;-24.8]
2	64	-9.18	4	[146.62,167.93] [265.31,448.81] [270.46,112.21] [390.12,172.28]	[44.86;35.98] [-3.82;-79.46] [-5.93;58.88] [-55.01;34.19]	[45.1;36.8] [-3.5;-78.7] [-5.6;59] [54.6;34.7]
3	67	-7.95	3	[147.06,168.6] [270.50,112.82] [390.48,171.94]	[44.68;35.71] [-5.95;58.63] [-55.16;34.33]	[45.1;36.5] [-5.6;59] [54.6;35.1]
4	68	-7.54	3	[147.45,168.79] [270.62,112.96] [390.33,172.10]	[44.52;35.63] [-5.99;58.58] [-55.1;34.26]	[45.1;36.5] [-5.6;59] [54.6;35.1]

5	150	26.08	3	[185.14,294.87] [267.87,349.79] [357.01,283.69]	[29.06;-16.19] [-4.87;-38.76] [-41.43;-11.6]	[30;-16] [-4.6;-37.8] [-41.2;-11.1]
6	160	30.18	2	[268.76,345.84] [352.63,283.18]	[-5.23;-37.14] [-39.63;-11.39]	[-5.3;-36.4] [-39.1;-10.4]
7	175	36.33	1	[272.16,193.63]	[-6.63;25.42]	[-6.3;25.9]
8	180	38.38	1	[271.35,196.6]	[-6.3;24.2]	[-6;24.9]
9	190	42.48	2	[231.79,255.97] [308.69,249.05]	[9.93;-0.21] [-21.61;2.64]	[10.2;-0.2] [-21.1;3.4]
10	200	46.58	2	[271.56,283.45] [303.85,248.31]	[-6.38;-11.5] [-19.63;2.94]	[-6;-10.7] [-19;3]
11	49	-15.33	2	[114.62,363.56] [431.14,346.42]	[57.99;-44.42] [-71.84;-37.38]	[58.5;-43.8] [-71.5;-36.8]
12	56	-12.46	2	[265.30,452.55] [429.01,345.64]	[-3.81;-81.00] [-70.96;-37.06]	[-3.5;-80.1] [-70.1;-36.5]
13	70	-6.72	3	[147.93,169.29] [270.55,113.53] [389.87,172.56]	[44.33;35.42] [-5.97;58.34] [-54.91;34.08]	[44.1;36.1] [-5.3;58.7] [-54.6;34.7]
14	130	17.8	1	[271.53,149.90]	[-6.37;43.39]	[-6;43.5]
15	155	28.13	3	[184.94,295.59] [268.12,348.18] [354.73,283.49]	[29.15;-16.49] [-4.97;-38.10] [-40.49;-11.52]	[28.2;-15.3] [-4.6;-37.1] [-40.1;-11.1]
16	178	37.56	1	[271.50,195.98]	[-6.36;24.45]	[-6.3;24.9]
17	185	40.43	1	[271.19,198.60]	[-6.23;23.38]	[-6;25.2]
18	195	44.53	3	[232.84,255.70] [271.33,285.31] [305.71,248.93]	[9.51;-0.09] [-6.29;-12.26] [-20.39;2.69]	[9.5;-0.6] [-6;-11.8] [-20.1;3.4]
19	44	-17.38	2	[113.61,364.21] [433.40,347.15]	[58.40;-44.69] [-72.76;-37.68]	[58.2;-44.2] [-72.6;-37.6]
20	57	-12.05	2	[265.28,452.30] [428.90,345.45]	[-3.81;-80.9] [-70.92;-36.98]	[-3.5;-80.1] [-70.4;-36.1]

Table 1: Centroids Comparison Table

The MSE of an estimator **a** with respect to an unknown parameter **b** is defined in Equation 03.5.

Using Table 2, Mean Squared Error (MSE) as well as the RMSE were computed. Table 2 shows the results.

$$\text{MSE (a)} = \text{E [a-b]}^2$$

Equation 03.5

$$\text{RMSE} = \sqrt{\text{MSE}}$$

Equation 03.6

MSE		RMSE	
R	S	R	S
0.194625	0.500225	0.441163	0.707266

Table 2: Error Computation

3.6.2 Final Markers Localization

The 3D printed circular shaped green markers were used for the remaining part of the project due to leakage problems in the first type of markers and shape preference. For that, Table 3 shows the extracted coordinates of the markers type b.

i	j	k	R	A	S
104.12	255.48	55.99	68.23	-29.26	55.99
264.25	109.7	61.88	-3.71	-26.62	61.88
412.81	244.17	61.14	-70.44	-26.95	61.14
260.58	416.79	62.24	-2.06	-26.46	62.24
146.65	253.34	107.17	49.12	-6.27	107.17
260.78	369.15	107.19	-2.15	-6.27	107.19
374.01	245.97	112.2	-53.01	-4.02	112.2
262.9	144.2	119.08	-3.1	-0.93	119.08
259.02	317.08	148.39	-1.35	12.24	148.39
194.3	251.23	152.86	27.71	14.25	152.86
326.27	248.62	157.02	-31.57	16.12	157.02
257.76	180.57	159.68	-0.79	17.31	159.68

Table 3: Final markers localization results

4 Design of the Camera Setup

In the MURAB project, a camera set-up is envisioned in order to measure the geometry of the breasts of patients lying in prone position and localize important features relative to the position of the robot arm. There are some requirements concerning the functionality, performance, and physical dimensions of this device. First of all, the setup will consist of two cameras that are mounted on the end effector of a robot arm. Since the kinematics of the robot arm and the pose of the end effector will be well-known, a single camera can also provide 3D information. However, a second camera will contribute much to the robustness and accuracy of the measurements. Therefore, a dual camera system has been selected. Such a system can reconstruct only parts of the geometry. However, by a scanning motion of the camera set-up using the robot's kinematic degrees of freedom, the surface of the breast can fully be reconstructed except for some possible folds in which no optic line of sight exists.

The skin of a patient often does not have sufficient texture to reliably perform stereo matching. Therefore, a structured light illuminator is needed to provide detailed textures. In the current setup this is accomplished by a laser pattern projection on the skin. If this projector would have been fully geometrically calibrated, it would provide sufficient information to acquire 3D information. Since the calibration of two camera setup is much simpler than the calibration of a camera-projector setup, a setup with two cameras have been selected.

The choice of the laser projector, i.e. the wavelength, the geometrical pattern, and the power are important design considerations. Another possibility for the illumination would be to use a pico-projector. This aspect of the design is not part of the current chapter. We assume the usage of a laser projector with a Diffraction Optical Element (DOE) that can be adjusted later.

The requirements are as follows:

- The accuracy of the surface reconstruction should be at least 1 mm.
- The cameras should be able to capture the field of view of interest defined by the size of the breast with a margin of error (25 cm).
- The specifications should be such that they fit the working space of the robot. This implies that the surface reconstruction should be possible within a range of 40 to 70 cm from the robot arm.

- The physical dimensions should be such that they fit the size of the robot end effector. This implies that the baseline between the two cameras is constricted to a maximum of 15 cm.

This chapter focuses on the physical design of the camera set-up. The design will be optimized for the case in which we have only one pair of images. That is, the scanning process and its impact on the accuracy of the reconstruction will be kept from the discussion.

4.1 Methods

In this project, the 3D surface reconstruction will be achieved by using a dual camera setup and acquiring pairs of images that provide depth information of the object of interest, the breast. The components and design of the camera setup are chosen while taking into consideration all previously defined requirements.

4.2 Materials

The camera setup consists of two cameras, a laser grid, and a 3D printed holder to be fixed on the KUKA arm end effector. The cameras used in this project are two Matrix Vision BlueFox-IGC USB 2.0. The laser grid is a green PICOTRONIC DOE Laser (This grid is out of the scope of this project and will not be detailed any further). The specifications of the cameras are found in [15].

4.3 Choice of Lenses

To optimize a high performance computer vision camera, it has to be coupled with the appropriate lens. In that case, some terms should be defined.

Focal Length: *When the lens is focused at infinity, the distance from the optical center of a camera's lens to that camera's sensor is the focal length. The focal length of a lens is constant and is a key specification of that lens.*

Angle of View: *The angle of view is the angle of subject area that is projected onto the camera's sensor by the lens. We can say that it's the angle over which the sensor can see through the lens. It depends on the lens focal length and the camera sensor size.*

Field of View: *The field of view is a representation of the angle of view. It is a measurement of the subject area instead of the angle.*

To select a lens, three factors are taken into consideration. These are defined as the design constraints and are listed below.

Design Constraints

Field of view $FOV = 250 \text{ mm}$

Working distance $D_{min} = 400 \text{ mm}$

Camera's sensor size $s = 4.8 \text{ mm}$

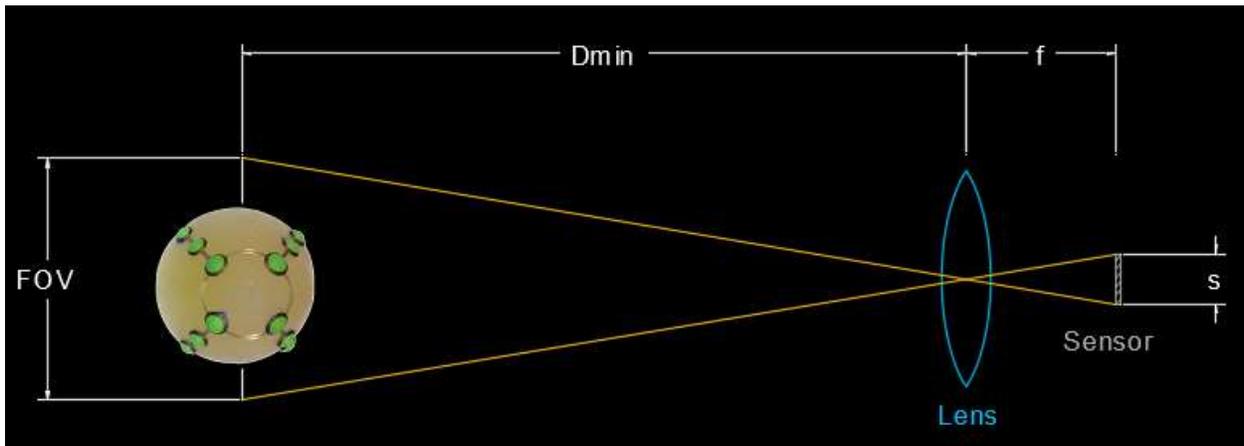


Figure 16: Top view of camera setup

From Figure 16, Equation 4.1 can be defined to choose the appropriate lens for our setup.

$$f = \frac{s}{FOV} \times D_{min} \quad \text{Equation 4.1}$$

When replacing our parameters, we obtain $f = \frac{4.8 \text{ mm}}{250 \text{ mm}} \times 400 \text{ mm} = 8 \text{ mm}$.

4.4 Distance Range Analysis

Once the appropriate lenses for the application have been chosen, a study of the depth of field can be made. In optics, the depth of field is the effective focus range. In other words, it is the range of distance from the camera in which an object can be photographed and still yield a focused image. Thus, defining the depth of field is important in order to define a range of our setup in which all captured images will be sharp. Even if a lens can only be exactly focused at one distance at a time, the decline in sharpness is gradual on either side of the focused distance, in a way that within the DOF, the unsharpness is imperceptible under normal viewing conditions.

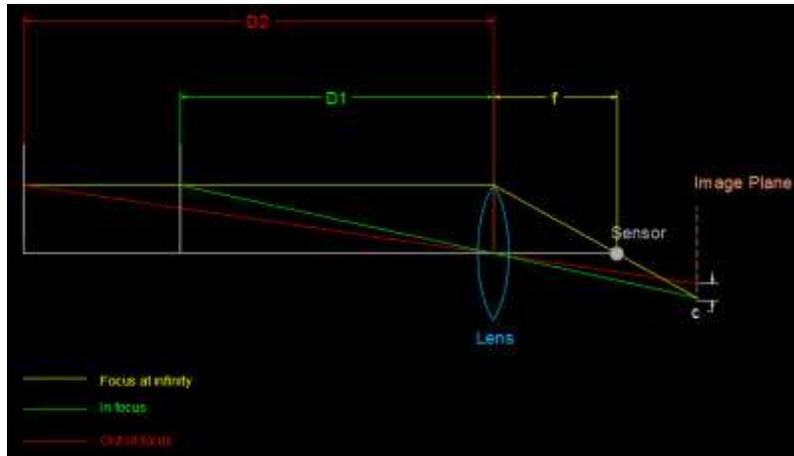


Figure 17: Circle of Confusion

The diameter of the circle of confusion or blur circle can be calculated based on the following equation.

$$c = A \times \frac{|D_1 - D_2|}{D_2} \frac{f}{D_1 - f} \quad \text{Equation 4.2}$$

A : Aperture³ diameter

D₁: In focus working distance

D₂: Current working distance

f : Focal length

A simple MATLAB simulation shows the variation in the blur circle diameter for a distance range of [40 60] cm with $D_1=50$ cm at a fixed aperture diameter $A=1$. The results are shown in Figure 18 and Figure 19.

³ Aperture refers to the opening of a lens's diaphragm through which light passes and is generally written as numbers called f numbers. Controlling the aperture means controlling the area over which light can pass through the camera's lens. As the f number decreases, relative light increases since the aperture diameter increases.

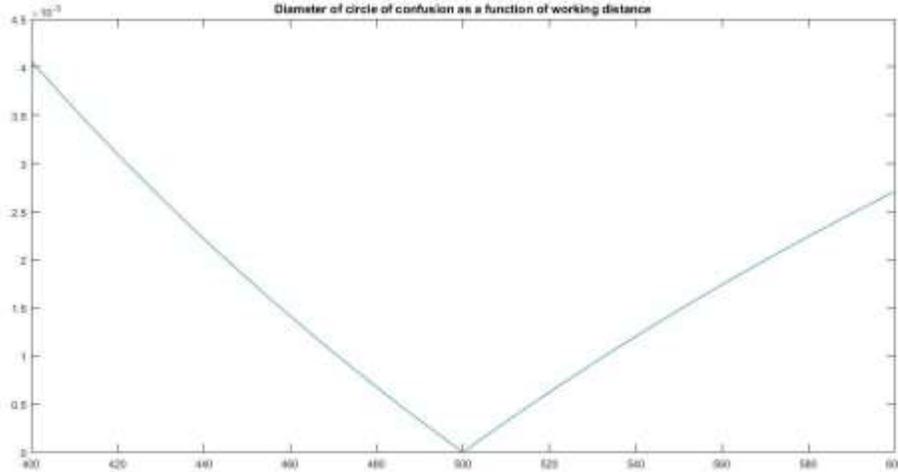


Figure 18: Diameter of circle of confusion as a function of working distance

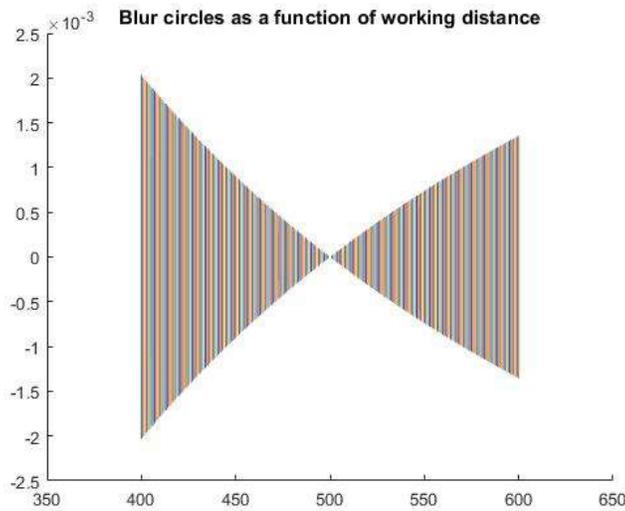


Figure 19: Top view of the plot of blur circles as a function of working distance

The depth of field can be defined as the region where the circle of confusion is less than the resolution of the display.

4.5 Choice of Camera Setup

In the case of this project, a choice has to be made between a parallel and convergent camera setup. The biggest advantage of a parallel stereo camera setup is the maximization of the overall field of view covered by both cameras. However, a parallel stereo camera arrangement has a limited overlap between the fields of views of the two cameras as illustrated in Figure 20 left. On the other hand, a convergent camera setup would be optimal in an indoor application where utility of the camera visual range is maximized and the workspace size is constrained. Such a configuration is illustrated in Figure 20 right. In addition, according to [16], as the verging angle between the stereo

cameras increases, accuracy is improved, reaching its optimal point at a verging angle of 90° .

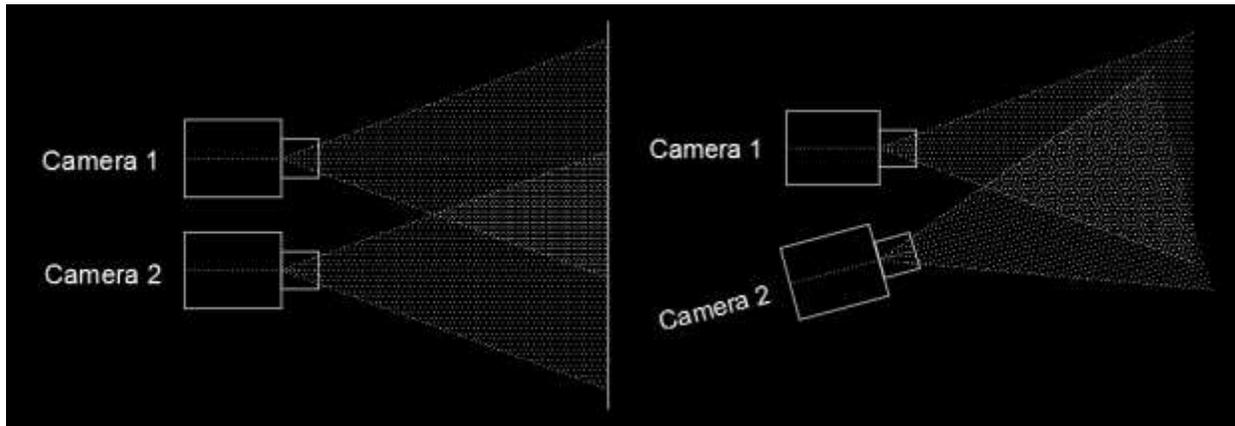


Figure 20: Parallel camera setup (left) versus convergent camera setup (right)

For an optimal setup in this project, the cameras must be verging, meaning non parallel.

4.5.1 Camera Locations

The two cameras must be placed in a convergent setup that allows the acquisition of the FOV of interest at the working distance desired by the application. Hence, the design can be summarized by the following:

Design Parameters

Baseline b

Verging angle θ

Design Constraints

Field of view $FOV = 250 \text{ mm}$

Minimum working distance $D_{min} = 400 \text{ mm}$

Size of the end effector

Design Considerations

Occlusions

Accuracy

The locations of the cameras in our convergent camera system can be defined by the baseline and verging angle.

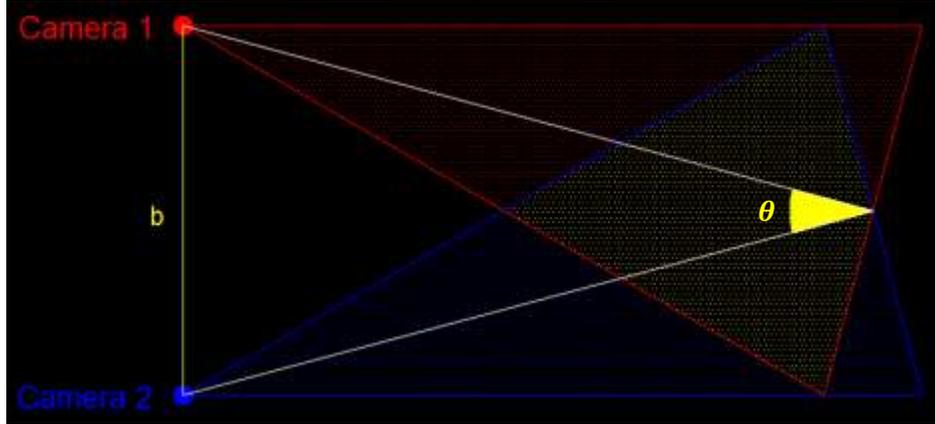


Figure 21: Camera positions and verging angle

From Figure 21, the verging angle θ is function of the baseline b and can be calculated by Equation 4.3.

$$\theta = \arctan \frac{b}{D_{min}} \quad \text{Equation 4.3}$$

This reduces our problem to defining b which should be limited to minimize possible occlusions and maximize accuracy while taking into consideration the size of the end effector to be mounted on the KUKA arm.

4.6 Monte Carlo Analysis

Decision making always involves risk. In that case, it would be a good idea to perform risk analysis. A Monte Carlo simulation allows the investigation of all possible outcomes of certain decisions and assess their corresponding consequences or risk, allowing for better decision making [17]. While deciding on the appropriate baseline for the setup, a Monte Carlo simulation has been performed to test the reliability of a camera setup based on its characteristic baseline b . The steps of this analysis, according to [18], could be summarized by the following.

1. Identify the model of the process or setup to be explored
2. Define the input parameters of the model
3. Create random data
4. Simulate the model for several repetitions and analyze the outputs of the model

Our Monte Carlo analysis is described in this section.

The model

The model whose reliability is to be tested is a converging camera setup with a characteristic baseline b and corresponding verging angle θ . 100 3-dimensional points are considered to be imaged by two cameras (left and right). The image coordinates of these 100 points in the two camera coordinate systems are known. An algorithm is written in order to estimate the corresponding 3D world coordinates of these 100 points based on a homography applied to their known 2D image coordinates.

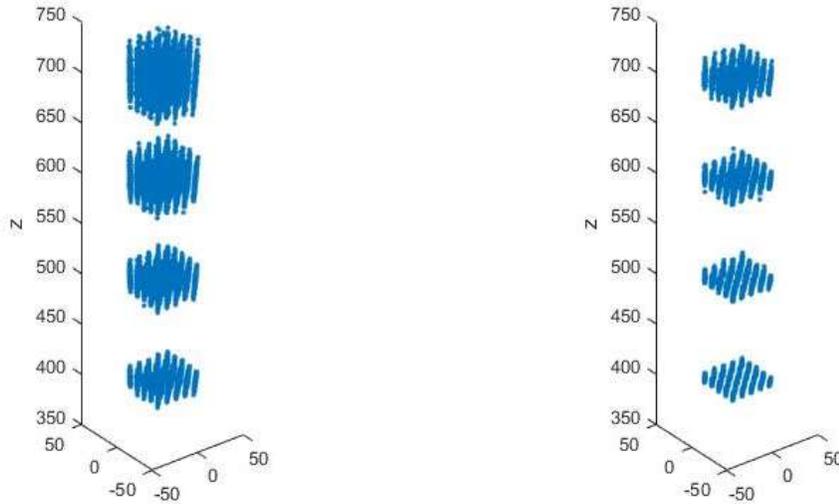
The Simulation

The 3D coordinates are estimated for a number of algorithm repetitions $N_{MC}=200$ where for each repetition, a random noise is added to both images. Finally, the covariance of the resulting estimates obtained over the 200 repetitions is used as a measure of the reliability of the model. The simulation is applied on several models with different characteristic baselines, and the results are compared in Table 4 and Figure 22.

Baseline b (cm)	Covariance in z coordinate
5	11.86
10	2.84
15	1.19
20	0.65
25	0.38

Table 4: Covariance in 3D coordinates estimation as a function of baseline

3D estimated world coordinates for a baseline of 5 cm 3D estimated world coordinates for a baseline of 10 cm



3D estimated world coordinates for a baseline of 15 cm 3D estimated world coordinates for a baseline of 20 cm

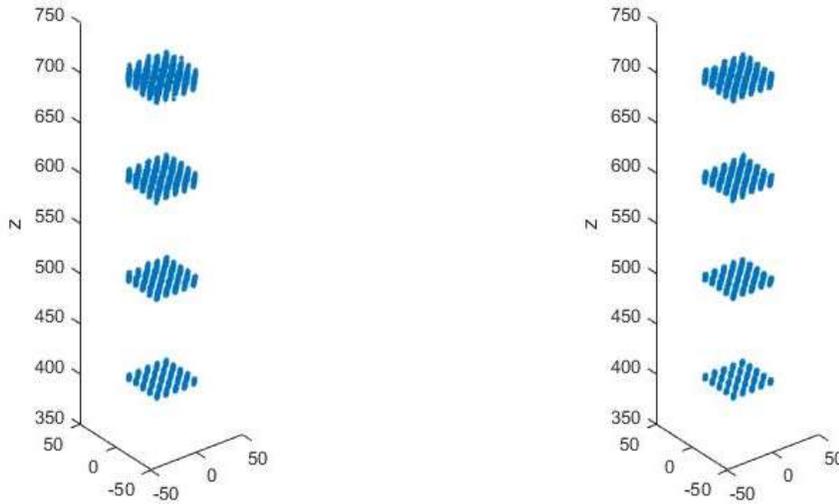


Figure 22: 3D estimated coordinates over 200 Monte Carlo repetitions for different baselines

The results show that a model with a larger baseline is more likely to be more reliable for obtaining 3D coordinates and thus for 3D reconstruction.

4.7 End effector constraint

As previously demonstrated, the larger the baseline, the more accurate the reconstruction results. However, the size of the end effector imposes a limitation to how much we can increase the baseline since the camera holder will be mounted to the latter. The maximum allowable baseline within the size of the end effector is about 15 cm. The corresponding verging angle is calculated

$$\text{as } \theta = \arctan \frac{150}{400} = 20.56^\circ.$$

4.8 Conclusion

Based on the detailed design calculations performed in this chapter, two Matrix Vision BlueFox-IGC USB 2.0 cameras were mounted with 8 mm lenses and placed converged to each other. As a first design, the camera holder was printed with 5.5 cm baseline and verging angle $\theta = \arctan \frac{55}{400} = 7.9^\circ$. The blur effect analyzed in the MATLAB simulation was considered to be acceptable within the working distance of the KUKA arm.

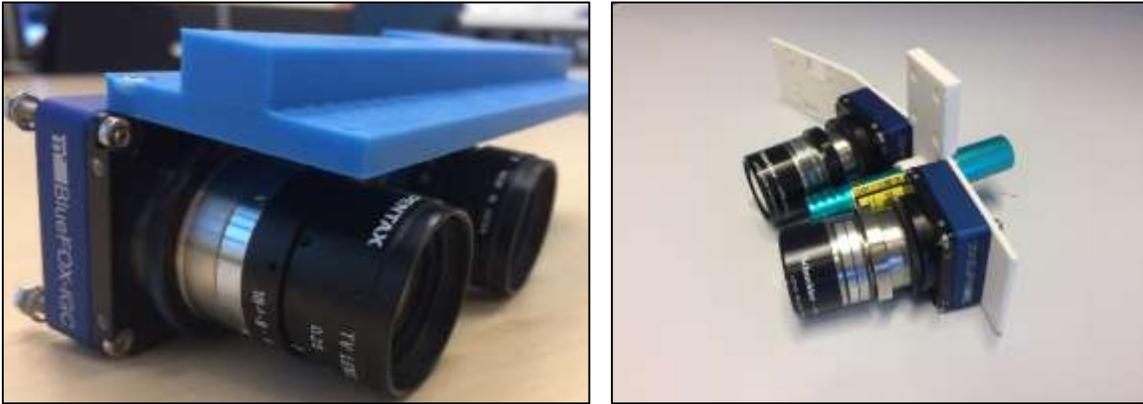


Figure 23: Parallel camera setup (left) versus converging camera setup (right)

5 Camera Reconstruction

Once the camera setup is ready, the next step would be using the setup to obtain the surface reconstruction of the breast. The case involves stereo vision which is nowadays a very popular approach in robotics applications where object detection, obstacle detection, or depth estimation is needed. In this project, the goal is to obtain the dimensional structure of the breast and eventually locate the depth relative to the robot end effector as well as accurately locate the markers needed for the registration.

In computer stereo vision, two cameras acquire images of the same scene, but they are separated by a distance – exactly like our eyes. An algorithm compares the images while placing them on top of each other in the aim of finding the features that match. The shift between the images is called the disparity and is used to calculate objects' distance from the camera setup.

To be able to match the images and eventually calculate the disparity and depth, the position of one camera must be accurately defined with respect to the second. For that, camera calibration is needed as the first step of the whole process.

5.1 Theory

5.1.1 Single Camera Calibration

Problem Statement

Camera calibration is very important in 3D computer vision because it allows the extraction of metric data from 2D images. It is the estimation of the intrinsic and extrinsic camera parameters using images with reference objects. These parameters define the transformation between an object in 3D space and its 2D image obtained by a camera.

Two types of parameters should be defined:

- Extrinsic parameters defining the orientation(rotation) and location(translation) of the camera
- Intrinsic parameters defining the characteristics of the camera

Notation

The pinhole camera model (Figure 24) will be used.

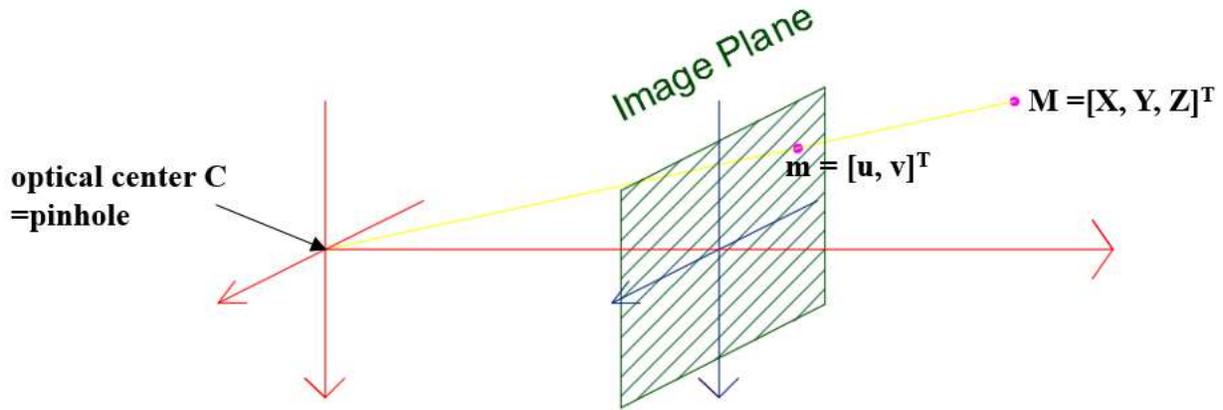


Figure 24: Pinhole camera model

$M = [X, Y, Z]^T$ is a 3D point in world coordinates and $m = [u, v]^T$ is its corresponding image 2D point, defined as the intersection between the optical ray from M passing through the optical center C and the image plane. Consequently, $M' = [X, Y, Z, 1]^T$ and $m' = [u, v, 1]^T$ would be the corresponding augmented vectors of M and m respectively.

The 3D point M is associated to its image projection m by Equation 5.1.

$$sm' = K[R \ t]M' = PM' \tag{Equation 5.1}$$

$$\text{with } K = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & 0 & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

s : Arbitrary scale factor

K : Intrinsic matrix (Camera calibration matrix)

$[R \ t]$: Extrinsic parameters (Rotation and translation relating the world coordinate system to the camera coordinate system)

(u_0, v_0) : Coordinates of the principal point

(α, β) : Scale factors in the image's u and v axes

γ : Parameter describing the skew of the two image axes

$P = K[R \ t]$: Camera projection matrix (Holding both intrinsic and extrinsic parameters)

Finally, there are 6 extrinsic parameters including the rotation matrix 3 degrees of freedom and the translation's 3 parameters while there are 5 intrinsic parameters $(\alpha, \beta, \gamma, u_0, v_0)$.

Calculating the Intrinsic and Extrinsic Parameters

Solving for the Camera Projection Matrix

At first, we recall that $sm' = PM'$ (Equation 5.1). We assume that we have the knowledge of a number n of reference points M' in the model and their corresponding image points m' . This method will allow us to find the camera projection matrix P and consequently the intrinsic and extrinsic parameters.

For each pair (M', m') , we have:

$$sm' = PM'$$
$$s \times \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

For each reference point M' , solving for s using the third equation leaves us with 2 equations. Having 12 unknowns, at least 6 reference points are needed to solve for P (At least 6x2 equations). Next, our parameters could be estimated.

Solving for the Parameters

$$P = K[R \ t] = [KR \ Kt]$$

$$\text{Let } A = KR = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{bmatrix}$$

$$\text{Let } B = AA^T = KR(KR)^T = KRR^T K^T = KK^T$$

Knowing P , we can calculate A as well as B and extract the intrinsic matrix K .

Next, the extrinsic parameters are obtained from

$$R = K^{-1}A$$
$$t = K^{-1} \times \begin{bmatrix} P_{14} \\ P_{24} \\ P_{34} \end{bmatrix}$$

Checker Board Calibration Technique

The checker board technique is an example of a 2D plane model calibration technique. It is mostly used for its good accuracy and easy implementation. The squares in the checkerboard are detected and their corners are used as reference points for the calibration. The number of squares in the X direction differs than that in the Y direction so that the axes could be differentiated. In this

calibration technique, the checkerboard plane is assumed to have $Z=0$. Hence, each point is represented as $M=[X,Y]^T$ and is related to m , its corresponding image point, by a homography H such that:

$$sm' = HM'$$

$$H = K[r_1 \ r_2 \ t]$$

Considerations

a) Radial Distortion

Lens Distortion is a geometrical transform characterized by Equation 5.2 and Equation 5.3.

$$u = x(1 + k(x^2 + y^2)) \quad \text{Equation 5.2}$$

$$v = y(1 + k(x^2 + y^2)) \quad \text{Equation 5.3}$$

(x, y) : Coordinates in the source domain (Undistorted image points)

(u, v) : Coordinates in the destination domain (Distorted image points)

Radial distortion is a simple model of lens distortion, and its distorting terms depend only on a radius $r = \sqrt{x^2 + y^2}$ where x and y are the undistorted camera coordinates. It occurs due to lens imperfections and could be compensated during camera calibration by estimating distortion parameters. Its effect could be then neutralized through further processing.

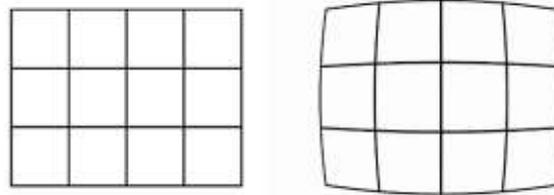


Figure 25: Test image (left) - Radially distorted image (right)

b) Reprojection Error

The reprojection error is a geometric error corresponding to the image distance in pixels between a projected point and a measured one. It could be used in qualifying camera calibration accuracy

by quantifying how closely an estimate of a 3D point recreates the point's true projection. Mathematically speaking, if \mathbf{M} is the 3D point, \mathbf{m} its corresponding true image point, and \mathbf{n} its corresponding reprojected point, then the reprojection error would be $\mathbf{d}(\mathbf{m},\mathbf{n})$, the Euclidean distance between \mathbf{m} and \mathbf{n} .

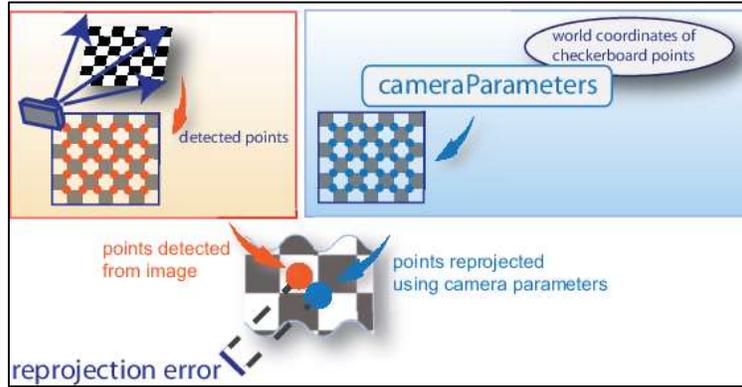


Figure 26: Calculating reprojection error [19]

When calibrating using the checkerboard method, the reprojection error is calculated as the difference between the actual detected pattern keypoint, and a corresponding world point projected into the same image. In single camera calibration, an average reprojection error is acceptable if it is less than one pixel.

5.1.2 Stereo Camera Calibration

In stereo photography, the same 3D object is imaged by two camera systems. Each camera has its own coordinate system: CCS_1 and CCS_2 . These two systems are related by a transformation composed of a translation and a rotation. The translation is the baseline between the two cameras (Figure 27). It is expressed in CCS_1 and denoted by 1t_2 . The orientation of CCS_2 relative to CCS_1 is defined by the rotation matrix 1R_2 . Hence, points represented in CCS_2 are represented in CCS_1 .

$${}^1M = {}^1R_2 {}^2M + {}^1t_2 \quad \text{Equation 5.4}$$

In stereo camera calibration, the intrinsic matrices are those of the individual cameras and the extrinsic parameters are defined by rotation and translation of one camera with the respect to the other.

Stereo Rectification

Stereo rectification allows to create a set-up of two identical cameras that are exactly aligned. This set-up allows for easier processing and straightforward epipolar geometry. In practice, the cameras are never perfectly aligned. However, stereo rectification allows to virtually rotate the cameras and facilitate epipolar geometry.

Before going into the details of rectification, a fully aligned stereo set-up must be described. The camera orientations are equal ${}^2R_1 = {}^1R_2 = I$, and the baseline is oriented horizontally ${}^1t_2 = [t_x \ 0 \ 0]^T$. The horizontal directions of both image planes are parallel to the baseline. The two optical axes (z-axes) are parallel to each other and orthogonal to the base line.

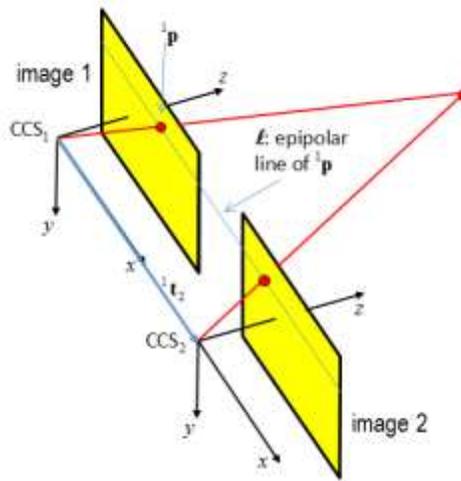


Figure 27: Stereo aligned camera set-up [20]

Hence, there are two requirements for rectifying two images:

- 1) Make the virtual calibration matrices equal
- 2) Rotate the two cameras such that they are oriented in the same direction

The direction of the baseline 1t_2 constrains the orientation of the two cameras. The horizontal axes (rows of pixels) of the cameras should be aligned with the base line, but the vertical axes (columns of pixels), can be chosen freely because the cameras can be rotated along the base line and still maintain the requirements for a full alignment. The third direction, the optical axis, is fixed to be orthogonal to the new image planes.

The first image is aligned to the image by applying a rotation matrix ${}^aR_1 = \begin{bmatrix} u_x^T \\ u_y^T \\ u_z^T \end{bmatrix}$ to CCS_1 .

Since the x-axis of the new coordinate system should be aligned with the baseline 1t_2 ,

$$u_x = \frac{{}^1t_2}{\|{}^1t_2\|}$$

$$u_y^T {}^1t_2 = 0$$

$$u_z^T {}^1t_2 = 0$$

A reasonable choice of the new y axis would be such that the orthogonal direction of the optical axis is preserved as much as possible. This is obtained by:

$$u_y = u_x \times \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$u_z = u_x \times u_y$$

Finally, the alignment of the first camera is accomplished by the homography:

$${}^aH_1 = K_1 {}^aR_1 K_1^{-1}$$

The second image is rectified by undoing its rotation relative to camera 1, applying the same alignment to the baseline as that developed for camera 1, and correcting for the difference in calibration matrix.

Finally, the alignment of the second camera is accomplished by the homography:

$${}^aH_2 = K_1 {}^aR_1 {}^1R_2 K_2^{-1}$$

Stereo Matching and Disparity Map

Stereo matching constitutes finding the corresponding points of all of an image's pixels in another image. It could be defined as:

Given a pixel ${}^1\mathbf{m}=[\mathbf{u} \ \mathbf{v}]^T$ in the first image, determine the corresponding pixel ${}^2\mathbf{m}=[\mathbf{u}' \ \mathbf{v}']^T$ in the second image. Do this for all pixels in the first image. Keep in mind the constraint that ${}^2\mathbf{m}$ should be on the same epipolar line associated with ${}^1\mathbf{m}$.

Disparity is the spatial shift of the same point in the two images. Once stereo matching has been applied, the disparity for each pixel ${}^1\mathbf{m}$ and its corresponding ${}^2\mathbf{m}$ can be calculated, and the so-called ***disparity map*** is introduced. This map decodes for each pixel the depth.

When working with two rectified images, all epipolar lines are horizontal so that each ${}^1\mathbf{m}$ and its corresponding ${}^2\mathbf{m}$ are on the same row of the image, i.e. $\mathbf{v}'=\mathbf{v}$.

If (\mathbf{u},\mathbf{v}) and $(\mathbf{u}',\mathbf{v}')=(\mathbf{u}',\mathbf{v})$ are corresponding points, then the disparity could be calculated as the pixel shift: $\mathbf{D}=\mathbf{u}-\mathbf{u}'$

One of the problems encountered during experimentation was that point clouds obtained from the first trials were too sparse in data points. Such point clouds were reconstructed from disparity maps with too little reliable points. In order to obtain a valid reconstruction, it is important to have a reliable disparity map. Otherwise, the 3D data points wouldn't be sufficient to create a dimensional reconstruction. And since the disparity map is obtained from the matched features in the two images, an adequate stereo matching method must be chosen. MATLAB uses Semi-Global Block Matching algorithms where the sum of absolute differences (SAD) of each block of pixels in the image is compared. For this project, a Speeded Up Robust Features (SURF) detector has been used in order to boost the stereo matching and define a disparity range from the features detected. One of the problems encountered during experimentation was that point clouds obtained from the first trials were too sparse in data points. Such point clouds were reconstructed from disparity maps with too little reliable points. In order to obtain a valid reconstruction, it is important to have a reliable disparity map. Otherwise, the 3D data points wouldn't be sufficient to create a dimensional reconstruction. And since the disparity map is obtained from the matched features in the two images, an adequate stereo matching method must be chosen. MATLAB uses Semi-Global Block Matching algorithms where the sum of absolute differences (SAD) of each block of pixels in the image is compared. For this project, a Speeded Up Robust Features (SURF) detector has been used in order to boost the stereo matching and define a disparity range from the features detected.

In computer vision, SURF is defined as a local feature detector and descriptor, partly inspired by the scale-invariant feature transform (SIFT) descriptor. SURF consists of detection, descriptor, and matching detailed next.

1) Detection

SURF first filters each of the pair of images, I , using square shaped filters defined by Equation 5.5.

$$S(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j) \quad \text{Equation 5.5}$$

It then finds points of interest in each image based on the Hessian matrix which is a square matrix with second order partial derivatives of a function or field as elements. The points of interest are chosen where the determinant of the Hessian matrix is maximal since the determinant measures the local change around a certain point. Given a point $p = (x, y)$ in an image I , the Hessian matrix $H(p, \sigma)$ at point p and scale σ , is:

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{yx}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix}$$

Equation 5.6

where $L_{xx}(p, \sigma)$, etc is the convolution of the second order derivative of the filter with the image at the point x .

2) Descriptor

A descriptor provides a unique and robust description of an image feature by describing certain traits such as shape, texture, the intensity distribution of the pixels within the neighborhood of the point of interest, etc. In SURF, a descriptor is obtained for every point of interest identified in the Detection step.

First, a reproducible orientation is selected and fixed based on information from a circular region around the point of interest. Then a square region associated to the previous orientation is constructed, and the SURF descriptor is extracted from it.

3) Matching

Matching pairs of features in the two images whose disparity is to be found are deduced by comparing their descriptors.

Triangulation for 3D Point Clouds [20]

Triangulation is the principle of estimating depth from images taken from two different points of view. A triangle is formed between the two focal points of the two cameras and the 3D point of interest as depicted in Figure 28.

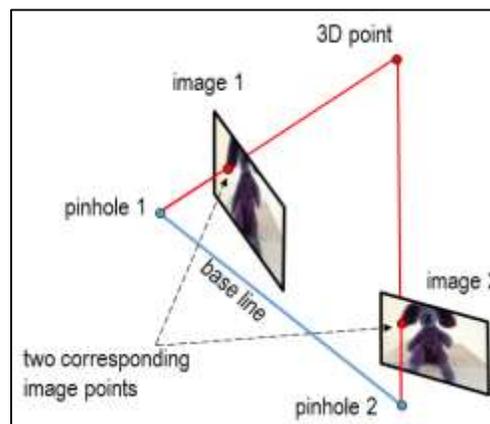


Figure 28: Triangulation (1) [20]

Knowledge of the length of the base line (the line between the two focal points), and the two angles formed by the base line and the two rays from the 3D point, suffices to calculate the depth of the 3D point as depicted in Figure 29.

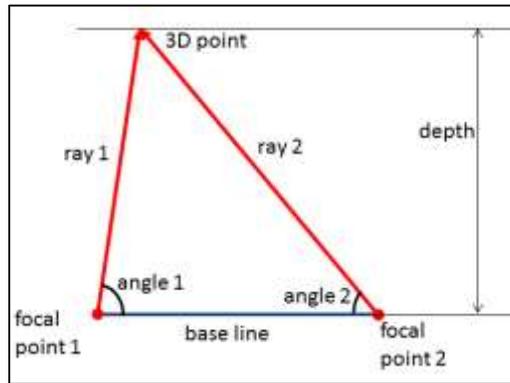


Figure 29: Triangulation (2) [20]

Triangulation helps in obtaining 3D point clouds by encoding the depth. Knowing the disparity D at a given pixel position (u, v) , the goal is to find the 3D position $M = [X \ Y \ Z]^T$ of the 3D point. This point is to be expressed in CCS_1 coordinates.

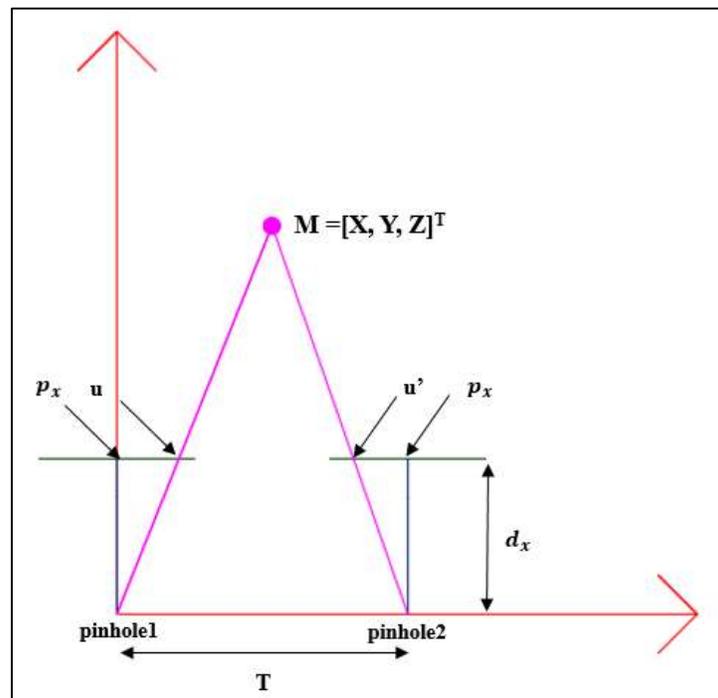


Figure 30: Triangulation(3)

Figure 30 allows us to compute the following equations for ray 1 and 2 respectively:

$$Z = \frac{d_x}{u-p_x} X \quad \text{Equation 5.7}$$

$$Z = \frac{d_x}{u'-p_x} (X-T) \quad \text{Equation 5.8}$$

Solving for Z yields Equation 5.9.

$$Z = \frac{d_x T}{u-u'} = \frac{d_x T}{D} \quad \text{Equation 5.9}$$

X and Y are obtained by applying perspective projection equations which yield:

$$X = \frac{u - p_x}{d_x} Z \quad \text{Equation 5.10}$$

$$Y = \frac{v - p_y}{d_y} Z \quad \text{Equation 5.11}$$

Eventually, applying equations to the whole disparity map yields a 3D point cloud defined by:

$$X_{cloud}(u, v) = [X(u, v) \ Y(u, v) \ Z(u, v)]^T$$

5.1.3. Marker Localization

Similarly to the final step in MRI reconstruction, the breast phantom markers must be localized for use in registration at the end of the project. What characterizes the markers used is their shape and color. However, the shape of the markers as viewed in the images may differ largely between two acquisitions depending on the position of the cameras relative to the breast. For that, color remains the only trait by which the markers can be identified in the stereo reconstruction process. Once the markers are localized in 2D in a pair of corresponding images, they are triangulated to obtain their 3D coordinates.

Gaussian Filter

In Gaussian smoothing, a 2-D convolution operator is used to blur images and remove detail and noise. We can say it is similar to the mean filter but it uses a different kernel that represents the

shape of a Gaussian hump. Applying a Gaussian blur has the effect of reducing the image's high-frequency components. It is thus a low pass filter and is generally used when we want nearest neighborhood pixels to have the most influence on the output.

In 2-D, a Gaussian is represented by

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \text{Equation 5.12}$$

(x, y) : Image pixel

σ : Gaussian distribution standard deviation, determines extent of smoothing



Figure 31: Gaussian blur [21]

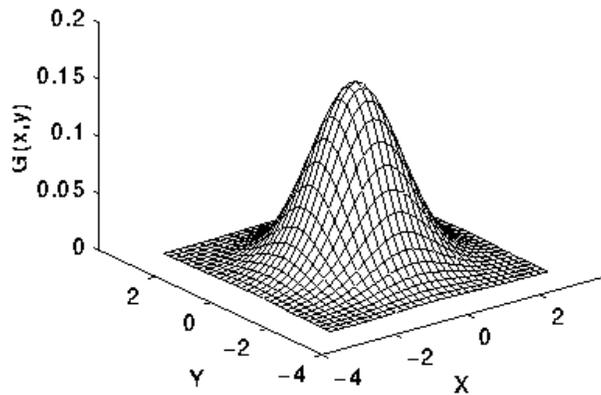


Figure 32: Gaussian kernel for sigma=1 [22]

RGB versus HSV

Red Green Blue (RGB) and Hue Saturation Value (HSV) are both color representations of an image. On one hand, RGB codes each pixel content of red, green, and blue from 0 to $2^{\text{number of bits}}$ (black to full intensity of the corresponding color). On the other hand, in HSV, hue defines color, saturation defines the intensity of the color, and value defines the brightness of the color. It is useful in computer vision because it differentiates in color and intensity, it is close to human perception, and it is invariant to illumination transforms. For example, different lighting in a room won't affect results. For a more flexible marker detection technique, a mask can be created based on the HSV characteristics of the markers and applied to the image in the aim of isolating and segmenting the markers.

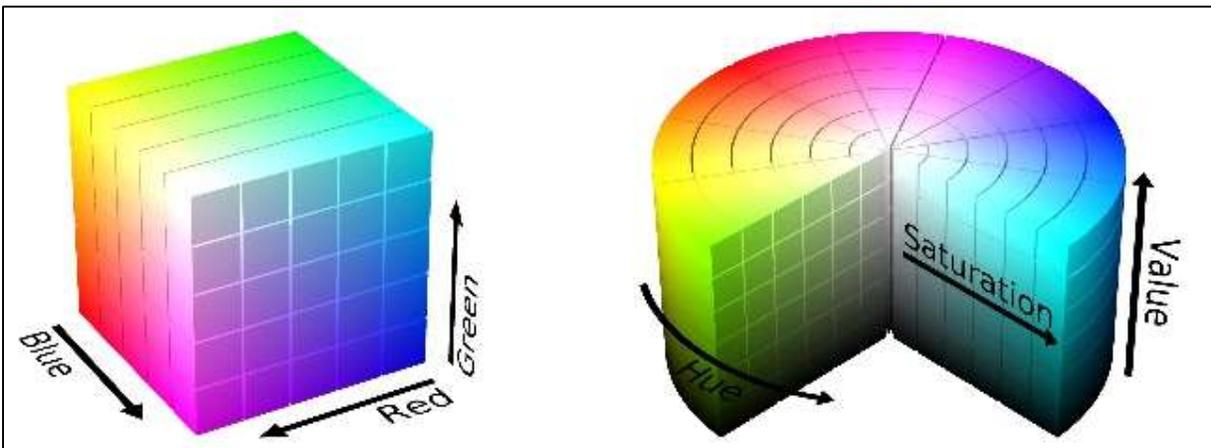
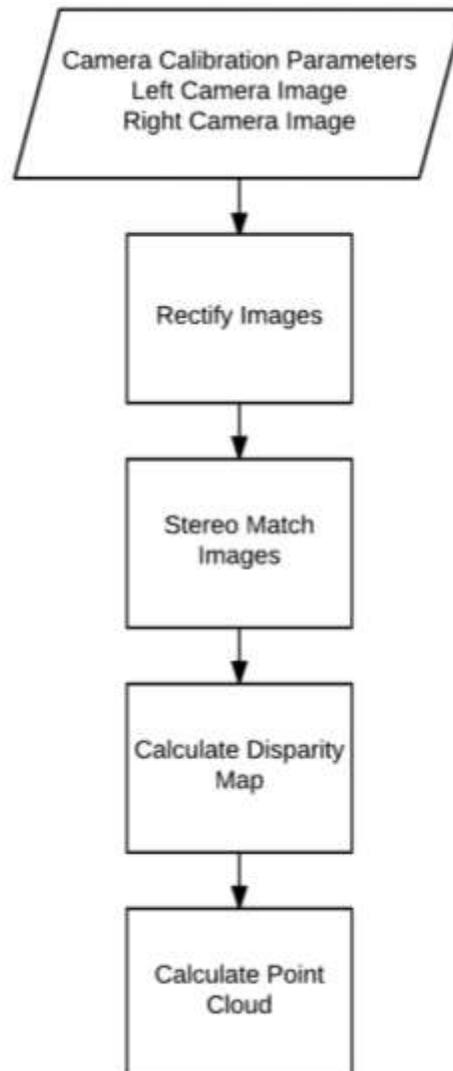


Figure 33: RGB color space (left) versus HSV color space (right)

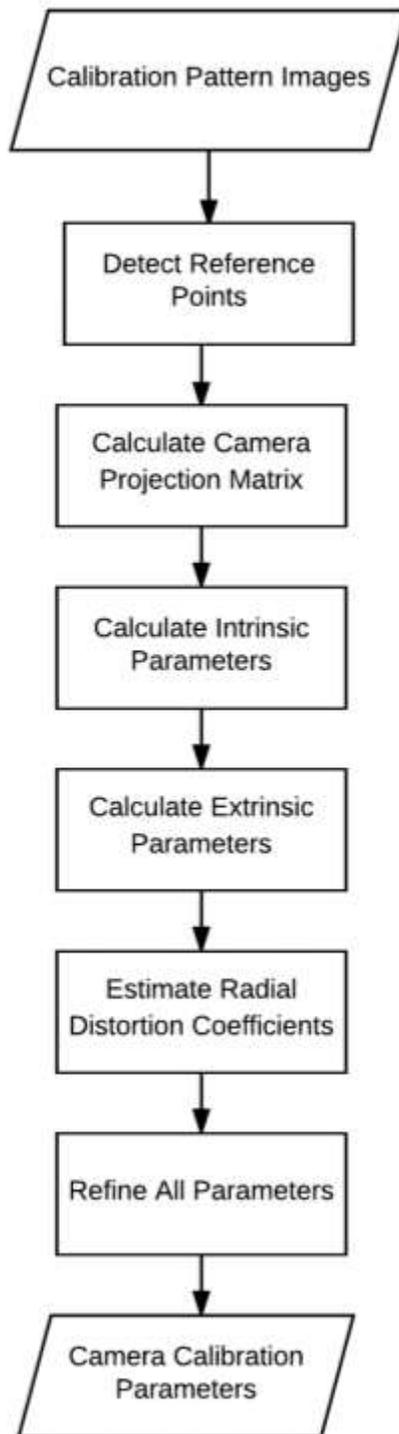
5.2 Implementation

The implementation of this chapter was done in MATLAB using the MATLAB Computer Vision System Toolbox. The algorithm can be summed up by Flow Chart 2, Flow Chart 3, Flow Chart 4, Flow Chart 5, and Flow Chart 6.



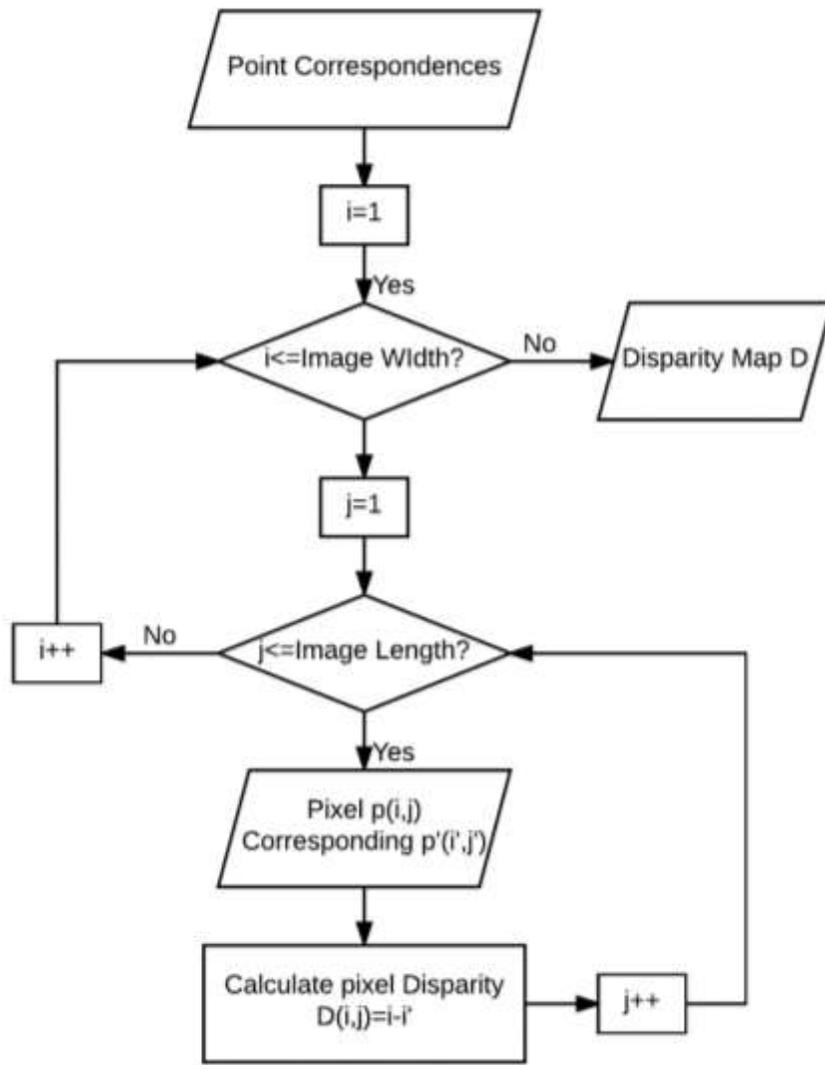
Flow Chart 2: 3D Reconstruction

Flow Chart 2 sums up the 3D reconstruction process in computer vision. Once the dual cameras are calibrated, the camera calibration parameters can be used to rectify the pair of images acquired. The rectified images can be then placed on top of each other and matched for common features. The disparity map is then obtained based on these features, and finally, the point cloud or surface reconstruction is estimated from the disparity map.



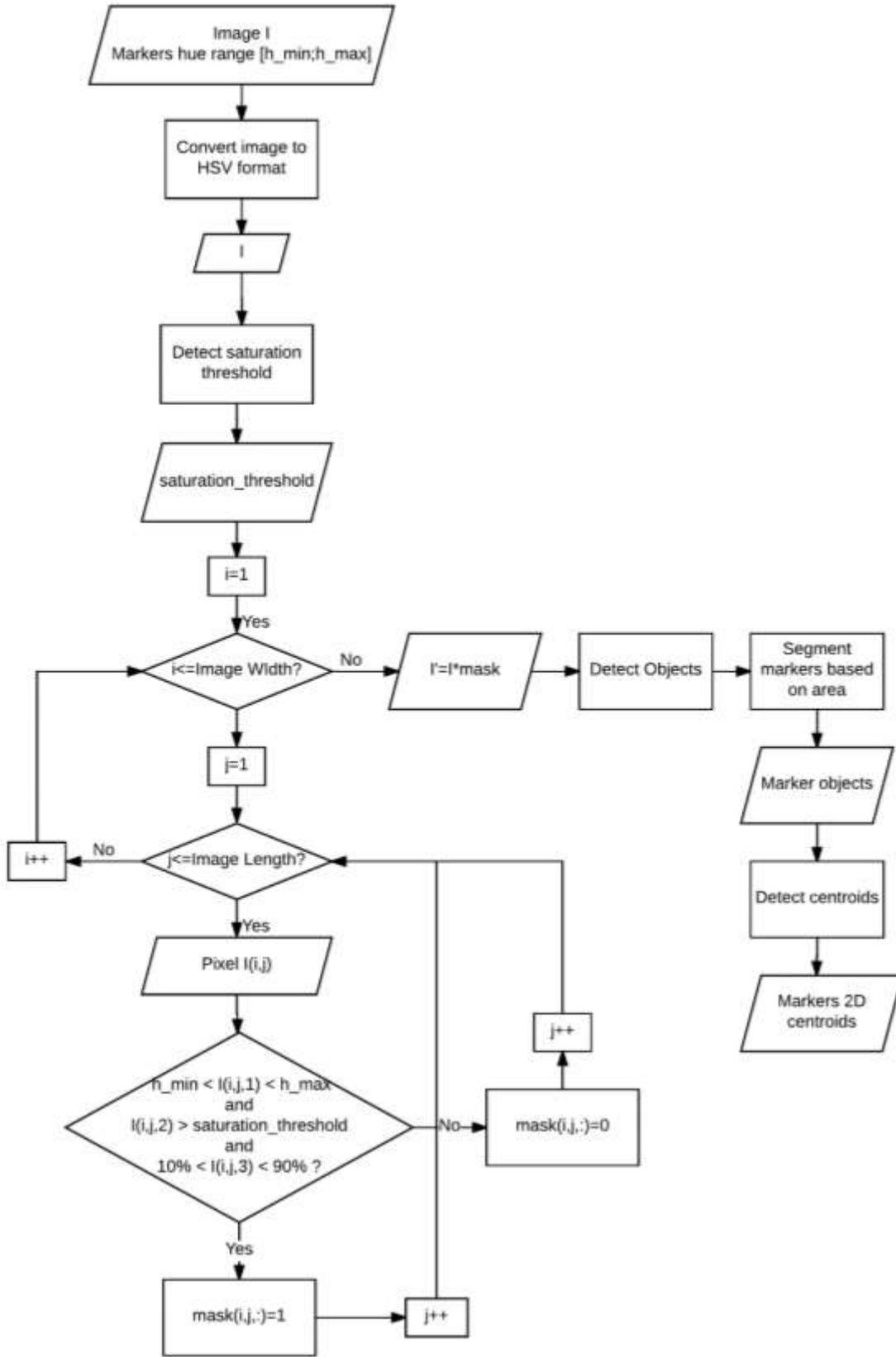
Flow Chart 3: Camera Calibration

Flow Chart 3 sums up the camera calibration steps that were explained in detail in the section 5.1 of this chapter. The output of the flow chart is the camera calibration parameters which will be used as inputs to Flow Chart 2.

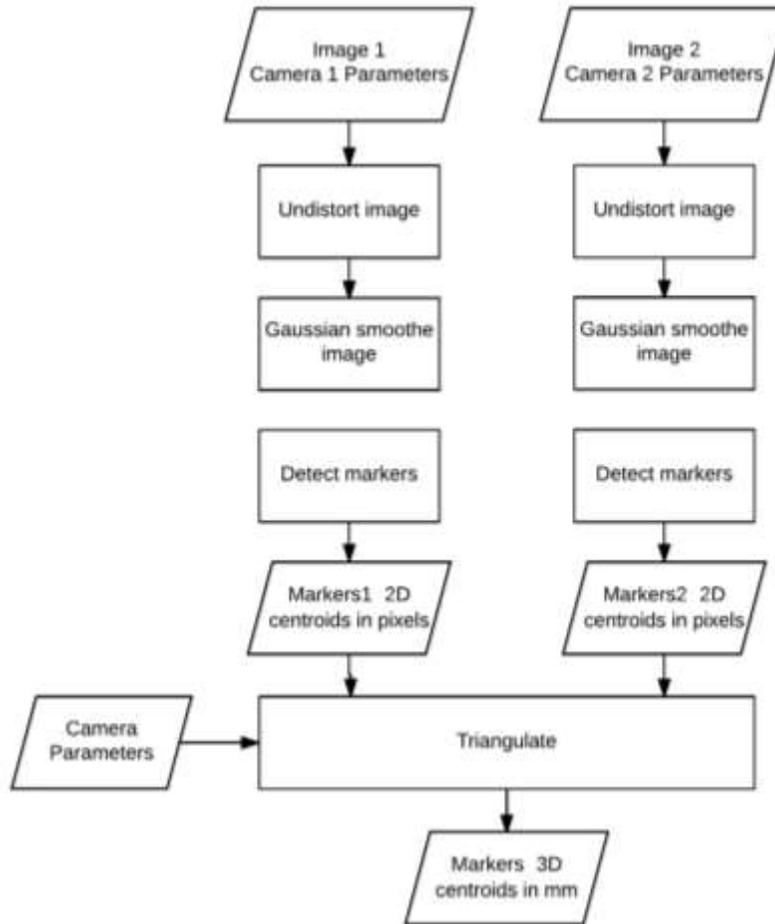


Flow Chart 4: Disparity Map

Flow Chart 4 depicts how the disparity map is obtained by going through all the pixel correspondences in the two images and obtaining the disparity of each matching pair.



Flow Chart 5: Markers detection and localization in image



Flow Chart 6: Markers 3D localization

Flow Chart 5 sums up how the markers are located in each of the two corresponding image pairs and localized in 2D. Each image is first filtered with Gaussian smoothing. A mask is then created based on hue, saturation, and value parameters. The hue range is characteristic of the specific green color of the markers with a margin of error. It could be detected from any HSV model or from the hue layer of the HSV format image. The saturation threshold is detected automatically in MATLAB based on the Otsu method defined previously in Chapter 3. The value range is defined as [10%; 90%] in order to eliminate the darkest and brightest values. The mask is later multiplied to the image. Finally, the markers are segmented based on their area, and their centroids are obtained in pixels.

Flow Chart 6 depicts how the 2D pixel coordinates extracted from the pair of images are used to obtain the 3D mm coordinates of markers.

5.3 Experimentation and Results

5.3.1 Camera Calibration

The stereo camera was calibrated using the MATLAB Stereo Camera Calibrator App and a checkerboard pattern with a square of 10 mm dimension. Every time the focus or aperture in the camera was adjusted, a new calibration session had to be done. For each session, more than 20 pairs of checkerboard images were acquired with the setup for calibration (Figure 34). The stereo parameters were then extracted while taking into consideration the correction for distortion. The calibration was generally always acceptable, yielding a reprojection error of less than 0.5 pixels. The final and best calibration session gave a reprojection error of 0.14 pixels (Figure 36). Knowing that the pixel dimension for the specifications of the camera used is $6\mu\text{m}$, this implies that the reprojection error is about $0.14 \times 6 = 0.84 \mu\text{m}$.

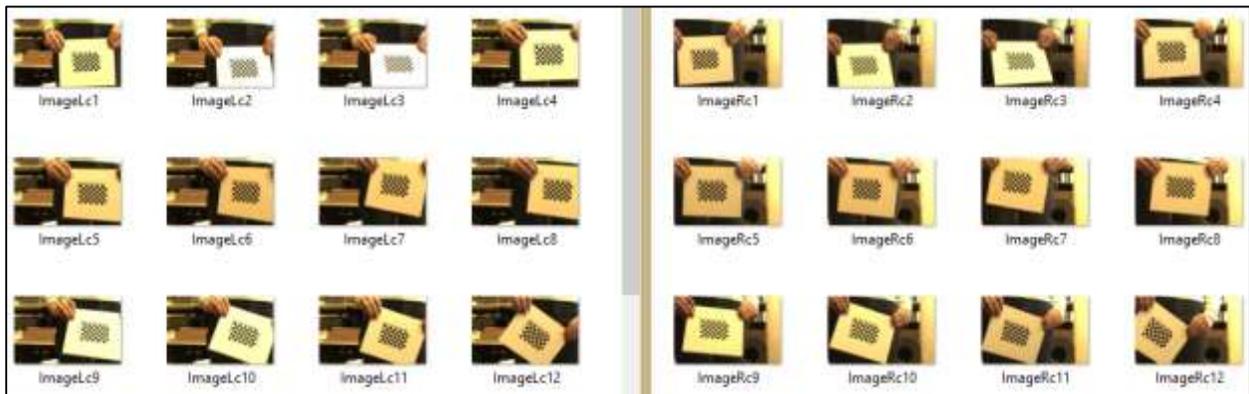


Figure 34: Checkerboard calibration images

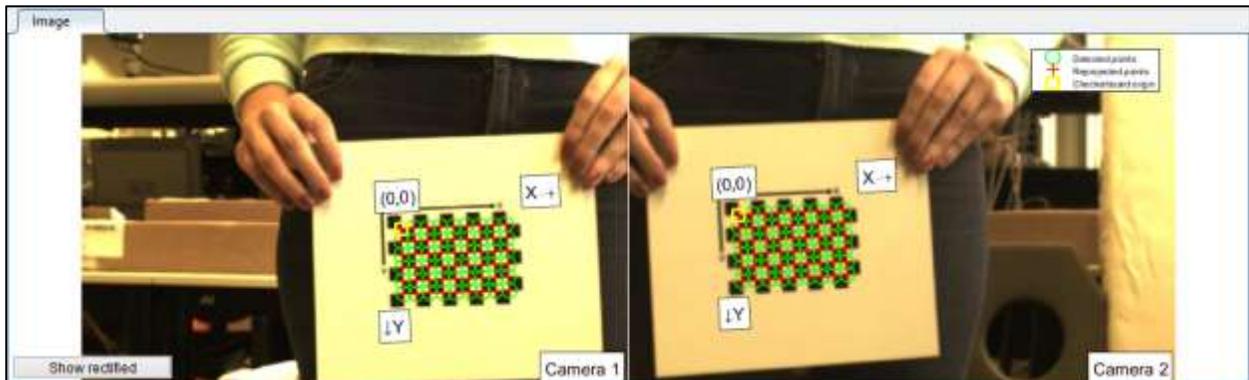


Figure 35: Stereo camera calibration with checkerboard images. (Green circles represent the detected points while red points represent the reprojected points)

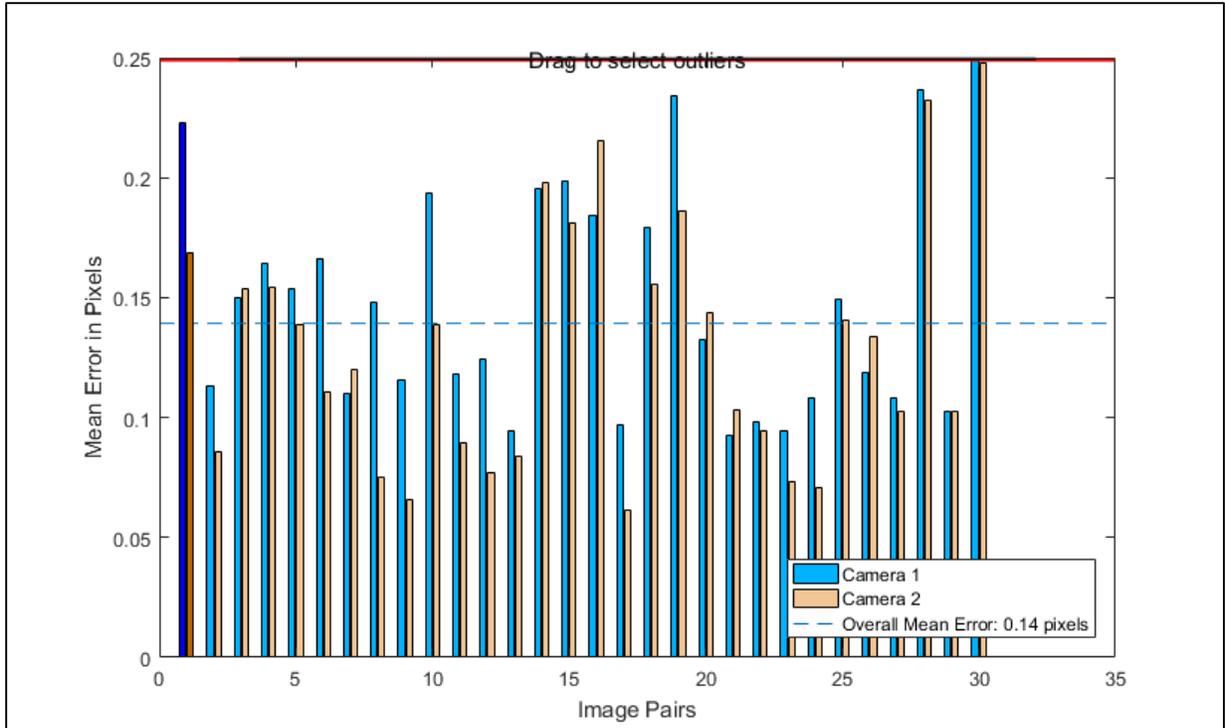


Figure 36: Graph of reprojection errors as a function of image pairs

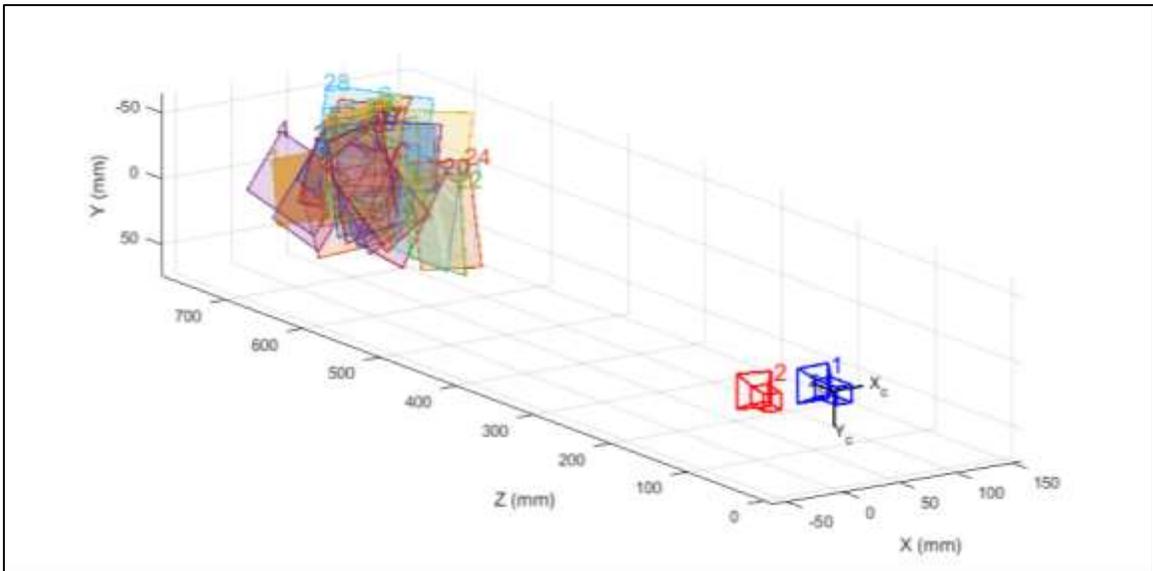


Figure 37: Stereo camera extrinsics- 1: Camera 1, 2: Camera 2 – The colored planes observed in front of the 2 cameras represent the checkerboard orientations during calibration

5.3.2 Stereo Rectification

Multiple image pairs were acquired by the stereo camera from different angles, and corresponding image pairs were rectified using the camera calibration parameters obtained.

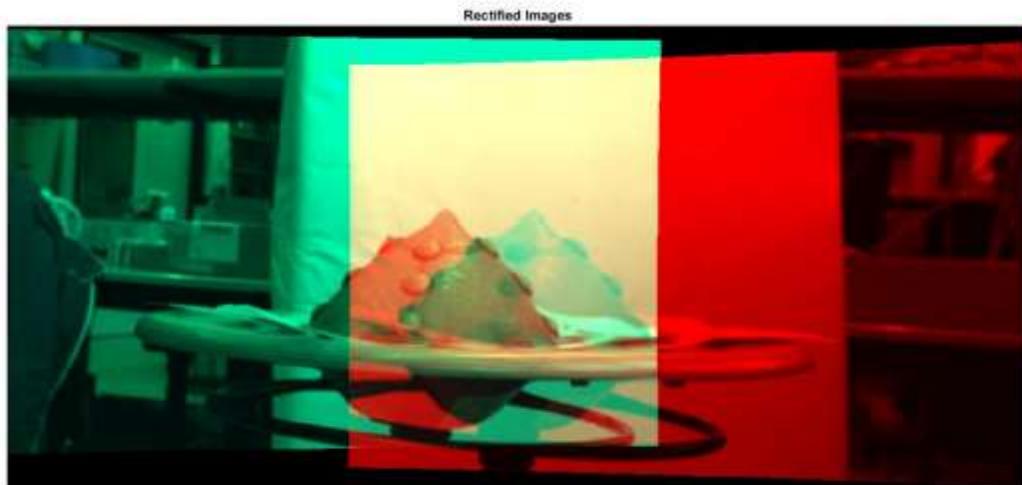


Figure 38: Rectified image pairs from side view of breast phantom

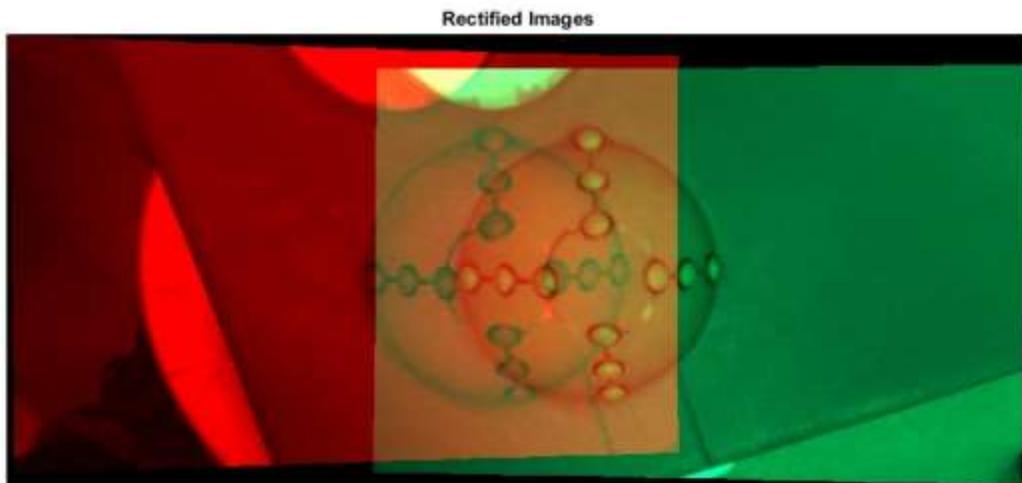


Figure 39: Rectified image pairs from upper view of breast phantom (1)

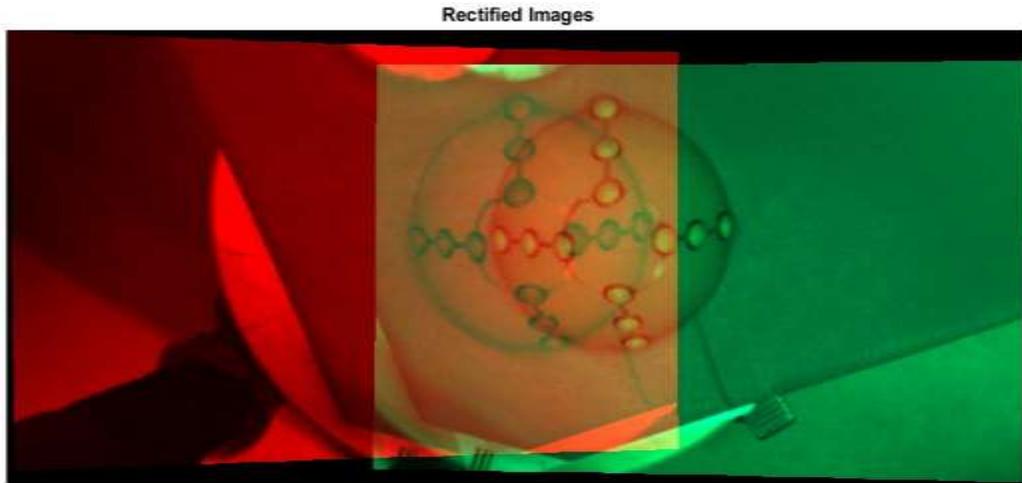


Figure 40: Rectified image pairs from upper view of breast phantom (2)

5.3.3 Stereo Matching and Disparity Map

Experimentation proves that using SURF features and range definition improves the disparity map as observed in Figure 42.

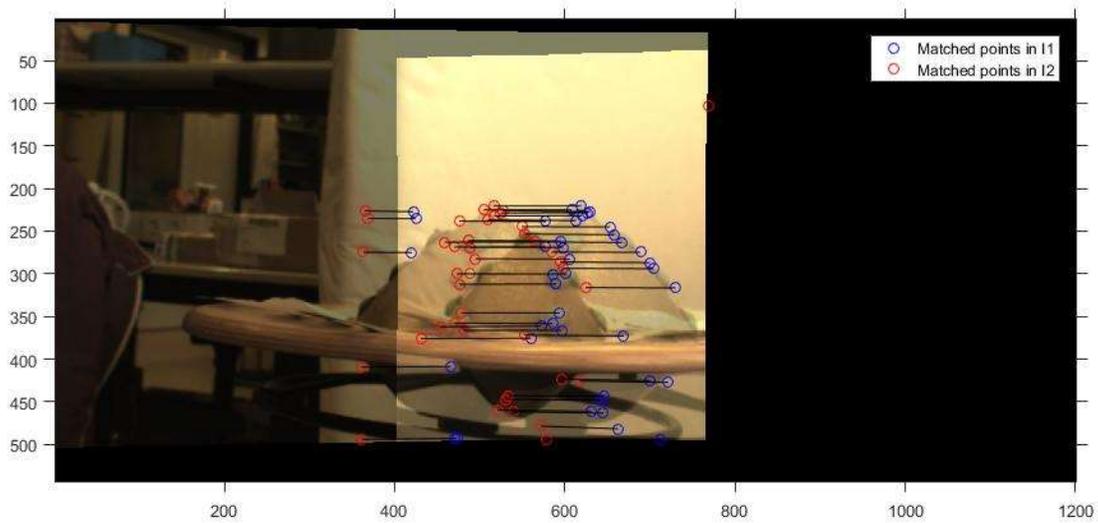


Figure 41: SURF features detected in side view rectified images

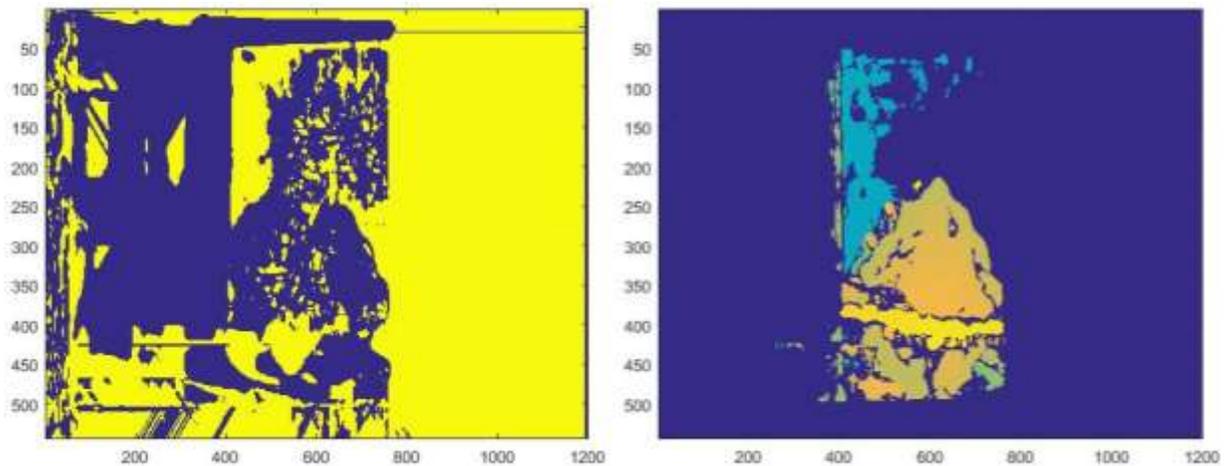


Figure 42: Disparity map before SURF and range definition (left) versus disparity map after (right)

5.3.4 Surface Reconstruction

Having the disparity map, the breast phantom can now be reconstructed by obtaining the X,Y, and Z coordinates of all points in the disparity map. Since our region of interest is the breast phantom and not the whole scene, the disparity map can be further improved by cleaning out the parts of the disparity map corresponding to the environment. Once the 3D coordinates of the points are obtained, the disparity map could be cleaned of all points lying outside the 40-70 cm working space of the KUKA arm. All connected pixels forming an area less than that of the breast can also be considered as noise and removed by setting their disparity to -1. The surface is then reconstructed again.

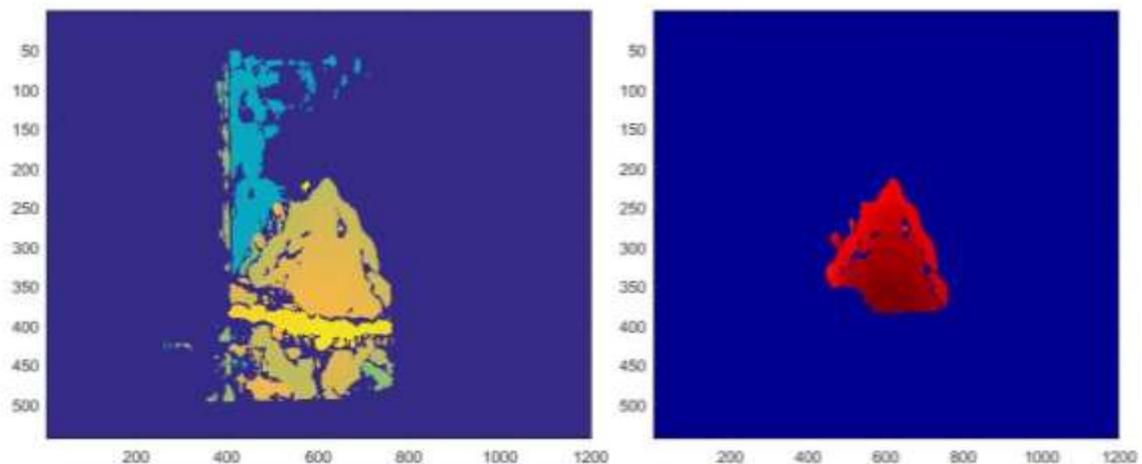


Figure 43: Disparity map before cleaning (left) versus disparity map after (right)

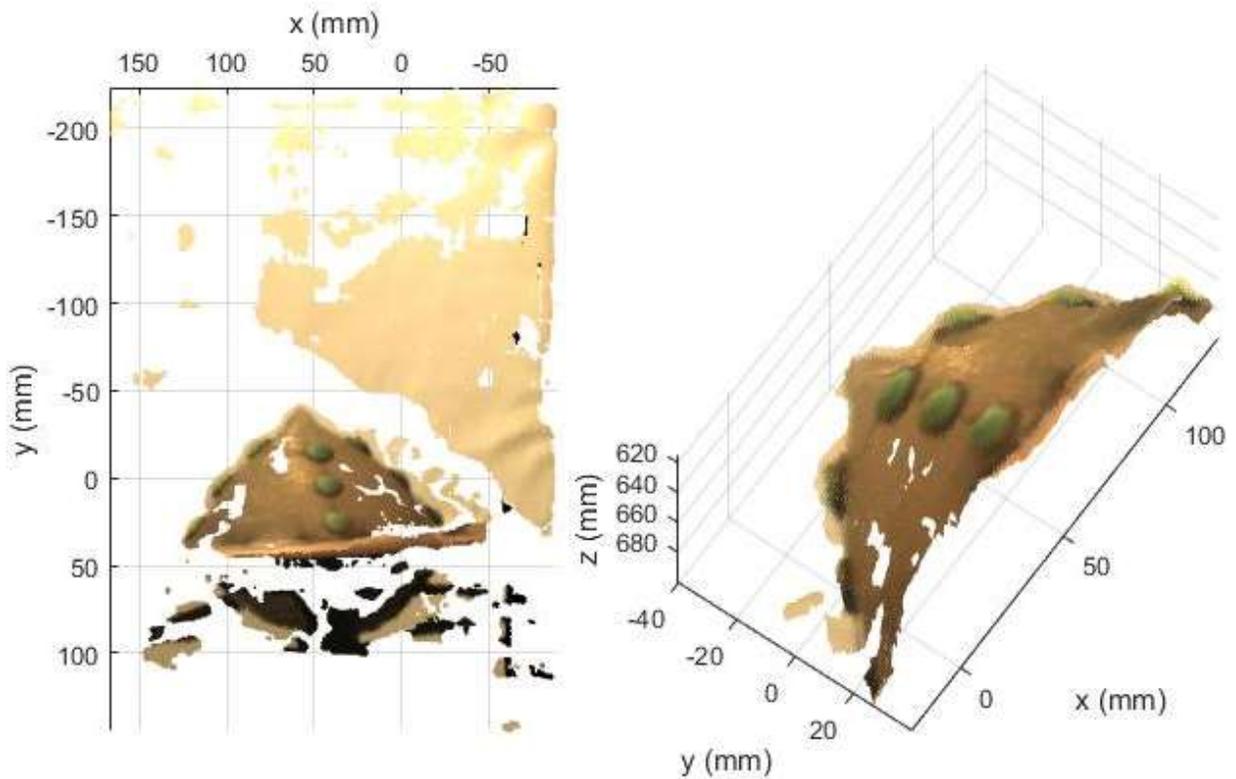


Figure 44: Point cloud before cleaning (left) versus point cloud after (right)

5.3.5 Markers Localization

The markers localization algorithm is implemented in MATLAB. A pair of images is acquired using the camera setup, and the markers are detected in both images. Their centroids are localized in pixel units then triangulated to find the corresponding 3D coordinates in mm.

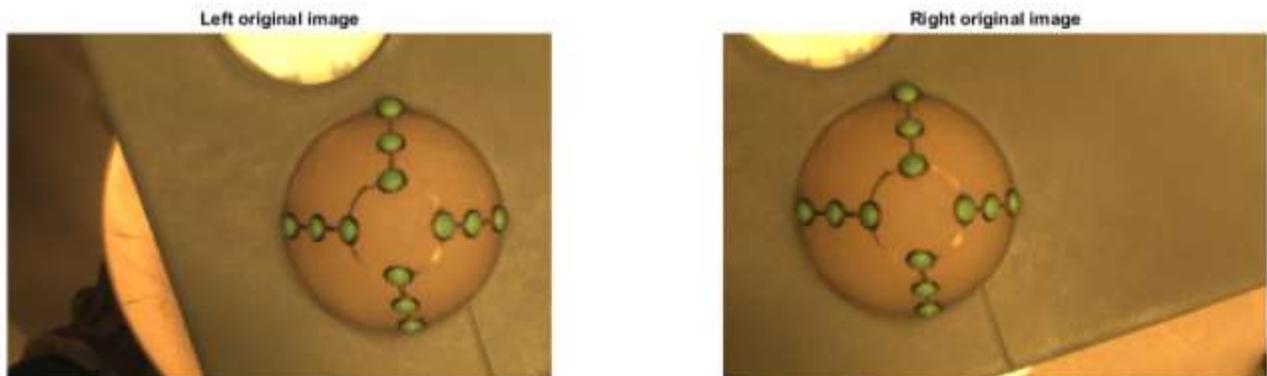


Figure 45: Left and right original images respectively

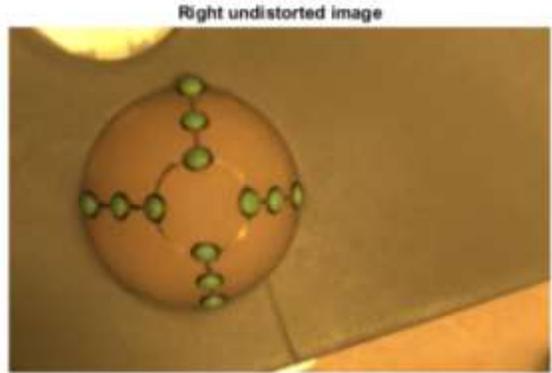
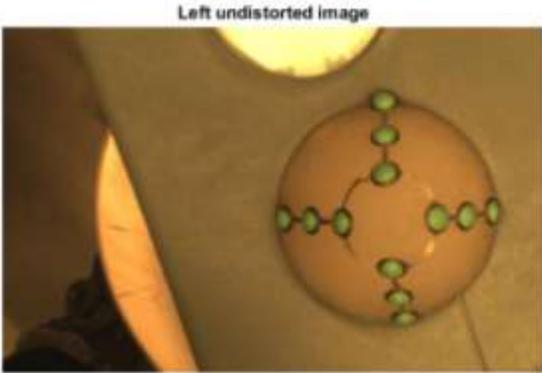


Figure 46: Left and right undistorted images respectively

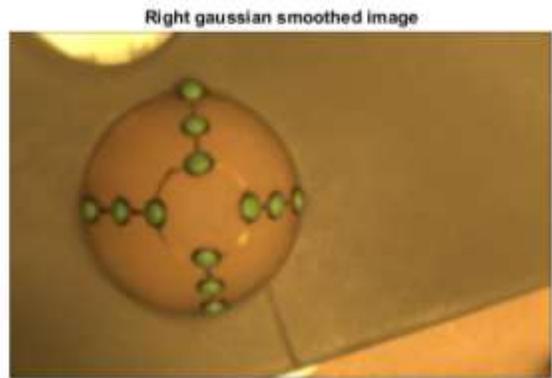
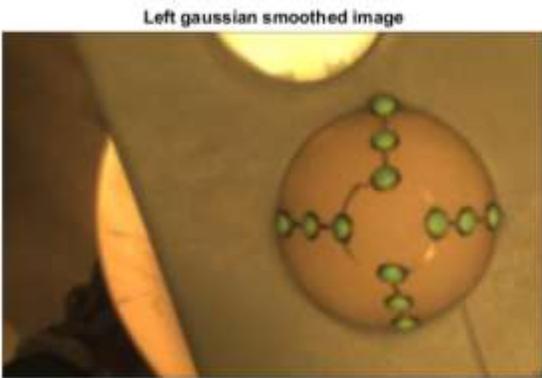


Figure 47: Left and right gaussian smoothed images respectively

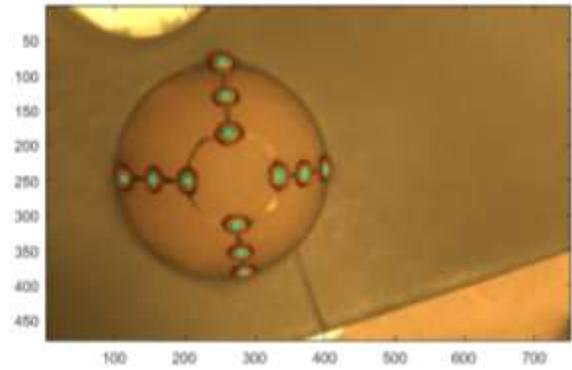
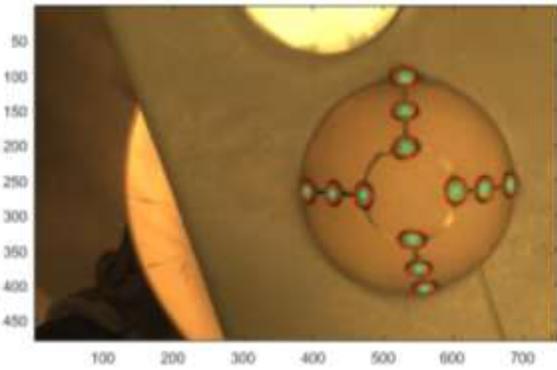


Figure 48: Markers localization in left and right camera images respectively

2D Centroid Coordinates

The results of the 2D and markers localization are stored in Table 5 and Table 6 respectively.

	<i>Left Image</i>		<i>Right Image</i>	
	<i>X (px)</i>	<i>Y (px)</i>	<i>X (px)</i>	<i>Y (px)</i>
<i>Marker 1</i>	<i>389.27</i>	<i>266.35</i>	<i>111.16</i>	<i>247.56</i>
<i>Marker 2</i>	<i>426.09</i>	<i>268.79</i>	<i>153.33</i>	<i>249.30</i>
<i>Marker 3</i>	<i>470.25</i>	<i>271.63</i>	<i>201.86</i>	<i>251.19</i>
<i>Marker 4</i>	<i>529.59</i>	<i>102.52</i>	<i>253.88</i>	<i>81.8</i>
<i>Marker 5</i>	<i>530.10</i>	<i>150.00</i>	<i>257.81</i>	<i>129.53</i>
<i>Marker 6</i>	<i>530.6</i>	<i>202.13</i>	<i>262.04</i>	<i>181.84</i>
<i>Marker 7</i>	<i>540.26</i>	<i>336.12</i>	<i>273.49</i>	<i>314.68</i>
<i>Marker 8</i>	<i>550.22</i>	<i>377.46</i>	<i>279.38</i>	<i>354.94</i>
<i>Marker 9</i>	<i>558.18</i>	<i>405.09</i>	<i>284.5</i>	<i>381.75</i>
<i>Marker 10</i>	<i>604.53</i>	<i>265.83</i>	<i>335.48</i>	<i>243.88</i>
<i>Marker 11</i>	<i>645.13</i>	<i>262.99</i>	<i>371.04</i>	<i>240.91</i>
<i>Marker 12</i>	<i>681.32</i>	<i>256.70</i>	<i>401.37</i>	<i>234.17</i>

Table 5: 2D pixel coordinates of detected markers in left and right camera images

	X(mm)	Y(mm)	Z(mm)
Marker 1	6.93	-0.22	692.03
Marker 2	24.03	0.88	665.48
Marker 3	43.34	2.06	643.01
Marker 4	72.68	-77.63	665.1
Marker 5	71.28	-53.9	649.82
Marker 6	69.8	-29.08	633.94
Marker 7	74.25	31.12	634.11
Marker 8	81.77	51.5	658.38
Marker 9	87.8	66.08	676.09
Marker 10	101.99	-0.56	627.42
Marker 11	122.89	-1.82	642.31
Marker 12	143.53	-4.9	661.52

Table 6: 3D mm coordinates of detected markers

6 Registration

The closure of this project is attained at the point where it all comes together, the registration. Image registration algorithms combine different images by aligning target images to a reference image. In our application, 3D registration will be applied in the aim of fusing MRI and camera reconstructed surfaces. The two sets of surface measurements are transformed into a common coordinate system, and the output is a fused model. The goal is to contribute to the SLAM phase of MURAB multiple imaging modalities are combined in order to accurately locate the lesion in the patient's breast.

Knowing that the common detectable aspects of the MRI and camera reconstructions include the markers and the surface of the breast phantom, it would be intuitive to focus on feature based registration. The markers are considered as extrinsic landmarks or features of the volume (fiducial points). The camera reconstructed surface is chosen as the reference volume since it depicts the relative position of the breast to the camera, and hence the robot. A geometrical transformation should be obtained to map the target surface (MRI surface) to the reference surface. Finally, we obtain a point by point correspondence between the reference and target volumes.

During the process, the major requirement kept in mind is to limit the error to less than 1 mm.

6.1 Theory and Implementation

6.1.1 Registration Approach

Since both MRI and camera surface reconstruction are obtained in mm, no scaling is needed. The shape of the reconstructions should be preserved as well. For that, rigid transformations will be used in the registration.

A common approach for 3D surface registration is using the iterative closest point (ICP) algorithm which aims to find the transformation between a point cloud and another point cloud by minimizing the square errors between the corresponding entities. It then assumes closest points correspond to each other and computes the best transform. For that, the ICP is only applicable in the case where a good first estimate of transformation is available. It could be then used to optimize

the process.

If we know correct correspondences, we can find a good first estimate of translation and rotation. As discussed previously, the landmarks or fiducial points that could be used to obtain the first transformation are the centroids of the markers placed on the breast phantom.

Procrustes Transformation

Having localized the centroids of the green markers placed on the phantom in both MRI and camera setups, the first transformation between the two surfaces can be estimated by transforming the MRI markers to the Camera markers using Procrustes. A Procrustes transformation is a geometric transformation that only involves translation, rotation, uniform scaling, or a combination of these transformations. Hence, it may change the size, but not the shape of a geometric object. In our case, scaling is not needed since both surfaces are in mm. Being a rigid transformation, the Procrustes equations are defined by Equation 2.1. The rotation and translation in Procrustes are defined such that the two objects will be superimposed and their shapes will be compared. The objective is to obtain a similar placement and size, by minimizing a measure of shape difference called the Procrustes distance between the objects.

Translation

The target surface is transformed to the reference one so that the mean of all the object's points (centroid) lies at the origin.

$$(x, y, z) \rightarrow (x - \bar{x}, y - \bar{y}, z - \bar{z})$$

Rotation

The optimum rotation is represented by a 3-by-3 rotation matrix R whose optimum value is obtained by singular value decomposition.

Shape Comparison

The Procrustes distance is used when comparing the shape of the superimposed surfaces. It can be defined by

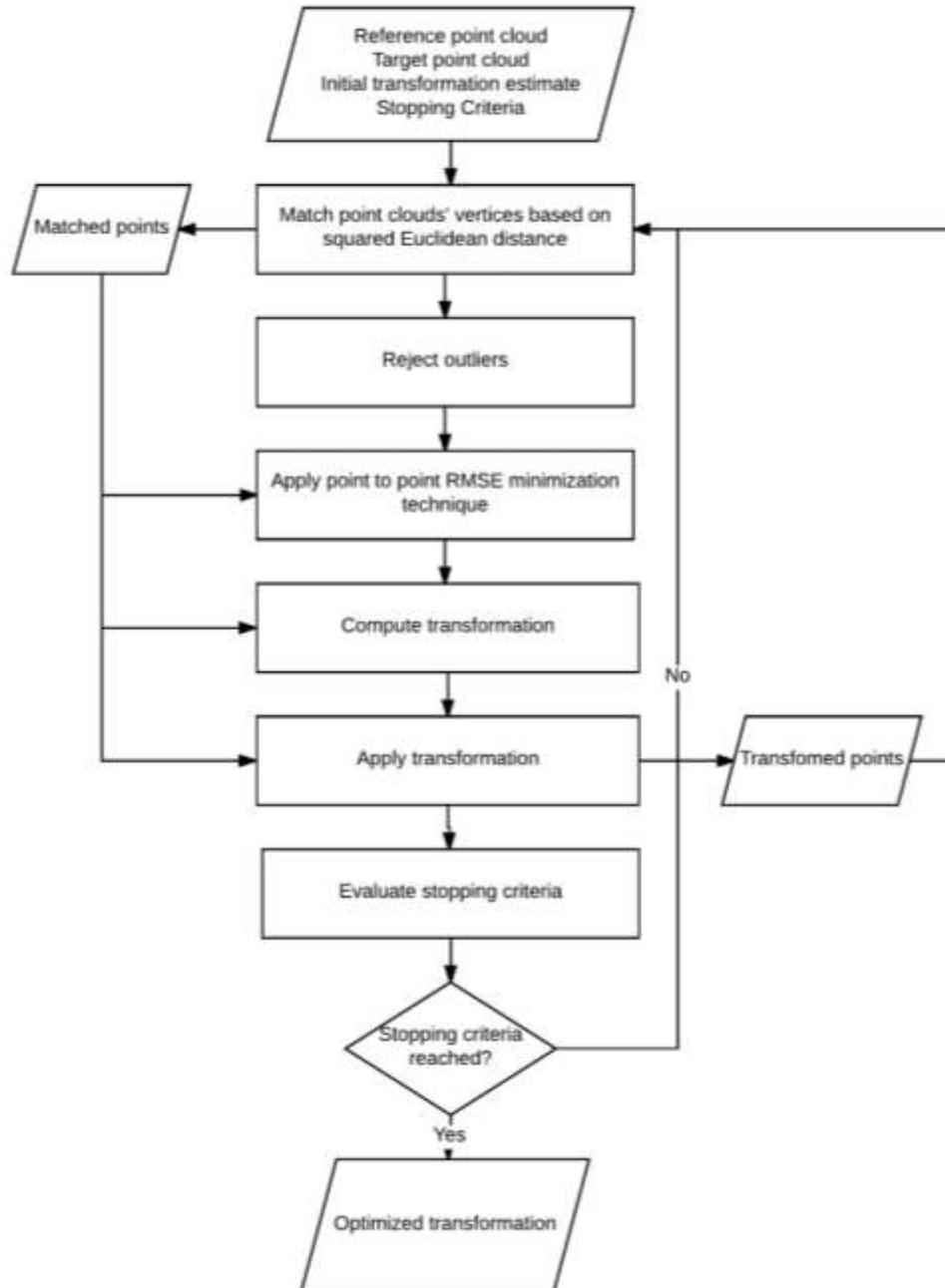
$$d = \sqrt{(u_1 - x_1)^2 + (v_1 - y_1)^2 + (w_1 - z_1)^2} \quad \text{Equation 6.1}$$

(u_1, v_1, w_1) : Transformed point

(x_1, y_1, z_1) : Corresponding point in reference surface

Iterative Closest Point

Iterative Closest Point is a widely used method for aligning three dimensional models given an initial guess of the rigid transformation. In the ICP, the reference point cloud is kept fixed while the target point cloud is transformed to best match the reference with a rotation and translation. The algorithm iteratively revises the transformation needed to minimize an error metric, usually the distance from the target to the reference point cloud.



Flow Chart 7: ICP

6.2 Results

6.2.1 Surface Registration

The previously detailed registration approach was used for camera-camera surface registration and applied on multiple camera reconstructed surfaces in the aim of filling holes obtained in single point clouds. Figure 49 shows 2 point clouds obtained from 2 different pairs of images while Figure 50 shows the registration of these point clouds before and after applying ICP. The left image in Figure 50 demonstrates the registration based on the first estimate transformation obtained from the markers while the right image demonstrates the final ICP optimization.

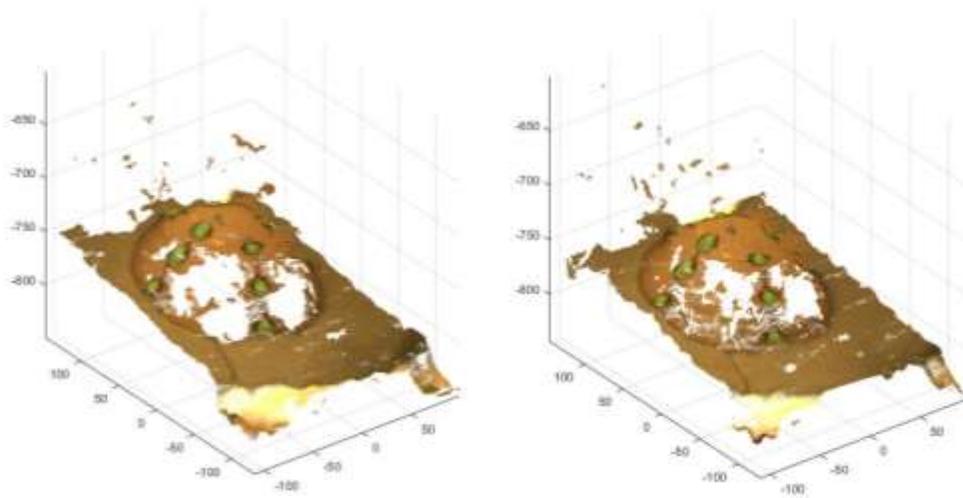


Figure 49: Two point clouds reconstructed from two different pairs of images

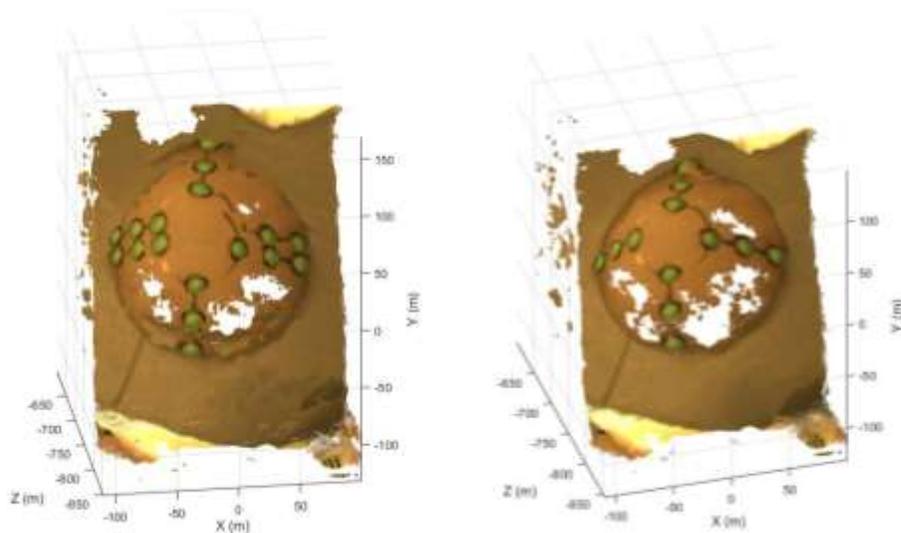


Figure 50: Camera-camera surface registration of 2 point clouds before ICP (left) and after ICP (right)

The registration was finally applied to seven different camera-camera reconstructed point clouds. Outliers were excluded from the registration by setting a threshold to the RMSE. From this experimentation, only one outlier out of 7 point clouds was obtained. Figure 51 shows the final camera-camera surface registration result.

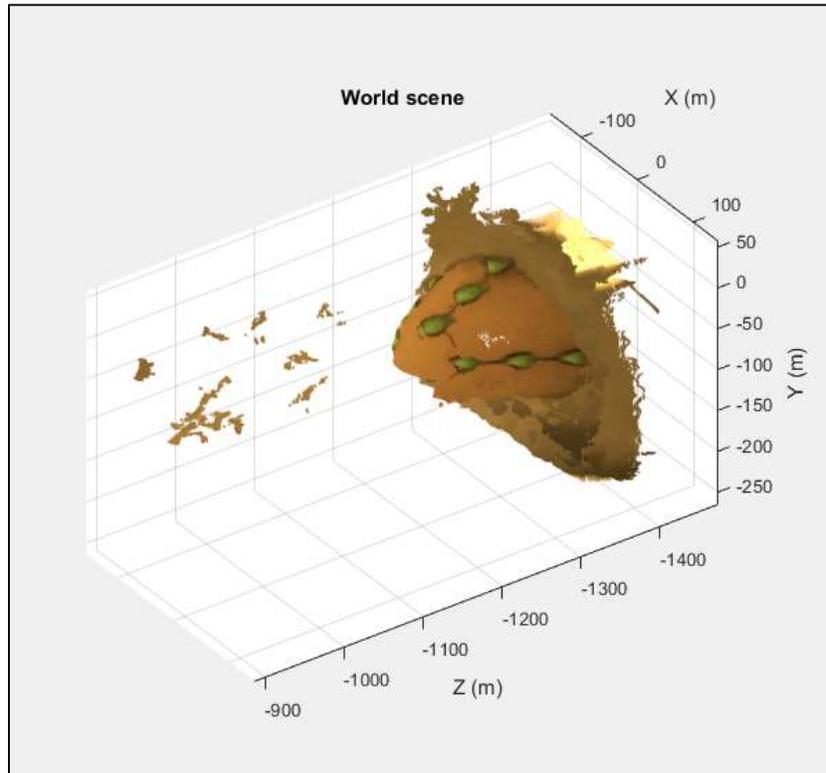


Figure 51: Camera-camera registration

The same algorithm was then used for MRI-camera surface registration in order to align the MRI surface with the camera surface, the output of camera-camera registration. The results of the final registration are illustrated in Figure 52 and Figure 53.

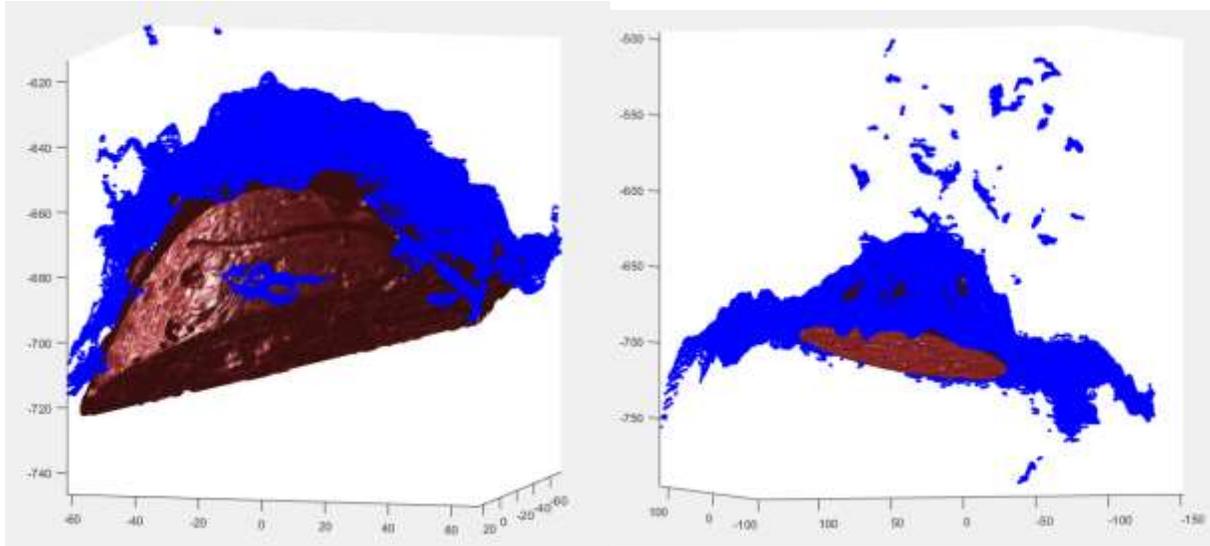


Figure 52: MRI-camera registration before ICP

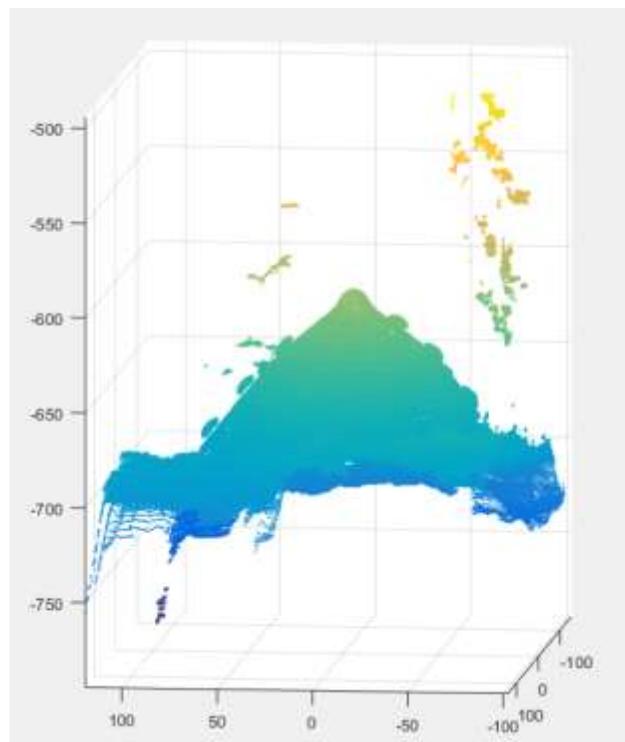


Figure 53: MRI-camera matching after ICP

6.2.2 Validation

Table 7 shows the RMSE obtained from the camera-camera registration, before and after applying the ICP as well as before and after excluding the outlier from the registration. Before applying ICP, the error is calculated from the transformed markers only since the point correspondences of the whole volume are not yet known. After applying ICP, the error is calculated from the transformation applied on the whole volume. These errors are not to be compared.

Camera-Camera Surface Registration (Outlier included)		
	RMSE before ICP (mm)	RMSE after ICP (mm)
X direction	0.97	
Y direction	0.63	
Z direction	4.23	
Mean RMSE	1.94	
Camera-Camera Surface Registration (Outlier excluded)		
	Mean RMSE before ICP (mm)	RMSE after ICP (mm)
X direction	0.85	
Y direction	0.53	
Z direction	4.42	
Mean RMSE	1.93	

Table 7: Registration Validation

6.2.3 Discussion

Concerning the camera-camera surface reconstruction, the visual results clearly show that the overall registration helps fill the holes in the point clouds due to unreliable points by merging several point clouds. It is also observed that the ICP optimizes the registration to a huge extent, eliminating the shift seen in the left image in Figure 50. As for the RMSE computation, the error computed from the markers does not vary to a large extent before and after eliminating the outliers. This is expected because transforming the already detected markers from two point clouds to each other is not a complex process. However, when it comes to the error computation of the complete volume registration, the error is reduced to half its value when the outliers are rejected from the registration. For that, it is important to exclude the outliers.

Conclusion and Future Recommendations

This study was conducted in the aim of testing the applicability and accuracy of MRI-Camera surface registration in the context of the European MURAB project with the hope of proposing an initial camera setup that could be further elaborated and fixed to the KUKA arm end effector.

In the first phase, MRI surface reconstruction was performed and the 3D coordinates of breast phantom markers were localized in anatomical space using centroid detection of the blobs. An average error of 0.5 mm in the marker localization was considered as acceptable.

In the second phase, a converged camera setup was designed, taking into consideration the field of view, robot working distance, and acceptable blur circle radius. It was implemented using rapid 3D printed prototyping.

In the third phase, camera surface reconstruction was performed, relying on stereo vision concepts and point cloud reconstruction. A color-based method was proposed in the aim of localizing the 3D coordinates of breast phantom markers in real space.

In the fourth phase, a semi-automatic, marker guided, registration method based on the ICP was proposed. Due to the close position required for ICP initialization, a first estimate rigid transformation was computed by marker matching. The algorithm was applied to 6 camera reconstructed point clouds in the aim of merging them together and filling any holes resulting from unreliable points in individual point clouds. It was also applied in the merging of MRI and camera reconstructed surfaces to fulfill the final goal of the project. The experimentation results showed that the accuracy of the algorithm (average of 3.39 mm for camera-camera registration) is reasonably acceptable for a first prototype and competes with an existing Kinect-based CT-camera registration approach with mean target positioning error of 5.23 mm [10].

The results obtained were reasonable and satisfactory for the number of experiments we were capable to conduct during a 3 months period of time. For the future, the camera setup is recommended to be mounted on the KUKA arm end effector and further tested. The use of more practical markers that could be more accurately localized also promises to improve the error of the registration. It is also predicted to elaborate the algorithm for this project in the aim of detecting lesions and evaluating the error in that case.

The registration of MRI and Camera reconstructed surfaces offers countless promises in what concerns combining the accuracy of MRI detectability and camera guidance for precise needle steering and a robust breast biopsy. In this context, a novel study was conducted and presented an impressive potential for clinical application of MRI-camera registration. Finally, this project lays groundwork for the development of a final camera setup design and clinically approved MRI-camera registration software that could lead towards revolutionizing the integration of medical imaging in robotic systems for cancer diagnostic operations.

Bibliography

- [1] J. S. Uiterkamp, "Design and implementation of autonomous robotic scanning of the breast," 2016.
- [2] "Breast Cancer Facts," 2016. [Online]. Available: <http://www.nationalbreastcancer.org/>.
- [3] "Breast Cancer Biopsy," 2016. [Online].
- [4] Priester, A. M., S. Natarajan and M. Culjat, "Robotic ultrasound systems in medicine," 2013.
- [5] V. G. Mallapragada, N. Sarkar and T. K. Podder, "Robot-assisted real-time tumor manipulation for breast biopsy," 2009.
- [6] F. El-Gamal, M. Elmogy and A. Atwan, "Current trends in medical image registration and fusion," *Egyptian Informatics Journal*, vol. 17, no. 1, pp. 99-124, 2016.
- [7] S. Nicolau, A. Garcia, X. Pennec, L. Soler and N. Ayache, "An augmented reality system to guide radio—frequency tumour ablation," *Computer Animation and Virtual Worlds*, vol. 16, no. 1, pp. 1-10, 2005.
- [8] L. Maier-Hein, A. Tekbas, A. Seitel, A. Müller, F. Pianka, S. Sätzl, S. Schawo, B. Radeleff, R. Tetzlaff and A. Tetzlaff, "In vivo accuracy assessment of a needle-based navigation system for CT-guided radiofrequency ablation of the liver," *Medical Physics*, vol. 35, p. 5385–96, 2008.
- [9] Cash D M, Chapman W C, Sinha T K, Terawaki H, Dawant B M, Galloway R L and Miga M, "Incorporation of a laser range scanner into image-guided liver surgery: surface acquisition, registration, and tracking," *Medical Physics*, vol. 30, p. 1671–82, 2003.
- [10] Deqiang Xiao, Huoling Luo, Fucang Jia, Yanfang Zhang, Yong Li, Xuejun Guo, Wei Cai, Chihua Fang, Yingfang Fan, Huimin Zheng and Qingmao Hu, "A Kinect™ camera based navigation system for percutaneous abdominal puncture," *Physics in Medicine and Biology*, vol. 61, no. 15, 2016.
- [11] Kilgus T, Heim E, Haase S, Prüfer S, Müller M, Seitel A, Fangerau M, Wiebe T, Iszatt J and Schlemmer H-P, "Mobile markerless augmented reality and its application in forensic medicine," *International Journal for Computer Assisted Radiology and Surgery*, vol. 10, p. 573–86, 2015.
- [12] S. A, "Towards markerless navigation for percutaneous needle insertions," *The International Journal for Computer Assisted Radiology and Surgery*, vol. 11, p. 107–17, 2016.
- [13] S. Trad, "Automatic segmentation method for lung cancer detection," 2014.
- [14] Slicer, "Coordinate systems," [Online]. Available: www.slicer.org. [Accessed 2016].
- [15] "Matrix Vision," [Online]. Available: <https://www.matrix-vision.com/USB2.0-single-board-camera-with-housing-mvbluefox-igc.html>. [Accessed 10 5 2017].
- [16] T. Pribanic, M. Cifrek and S. Tonkovic, "The Choice of Camera Setup in 3D Motion Reconstruction Systems," in *The 22nd Annual EMBS International Conference*, Chicago IL, 2000.
- [17] "Palisade," [Online]. Available: http://www.palisade.com/risk/monte_carlo_simulation.asp. [Accessed 12 05 2017].

- [18] P. Sheehy and E. Martz, "Minitab," [Online]. Available: <http://www.minitab.com/en-us/Published-Articles/Doing-Monte-Carlo-Simulation-in-Minitab-Statistical-Software/>. [Accessed 12 05 2017].
- [19] MATLAB, "Mathworks," MATLAB, [Online]. Available: <https://nl.mathworks.com/help/vision/ug/stereo-camera-calibrator-app.html>. [Accessed 20 April 2017].
- [20] F. v. d. Heijden, "3D surface reconstruction," *University of twente - RAM*, p. 3, 2016.
- [21] "Wikipedia," [Online]. Available: https://en.wikipedia.org/wiki/Gaussian_blur. [Accessed 12 05 2017].
- [22] [Online]. Available: <http://homepages.inf.ed.ac.uk/rbf/HIPR2/gsmooth.htm>. [Accessed 12 05 2017].
- [23] E. (. Bukai, "Design of a novel needle steering manipulator for minimally invasive guided breast biopsy," 2016.
- [24] Mathworks. [Online]. Available: www.mathworks.com. [Accessed 2016].
- [25] O. Zettinig, A. Shah, C. Hennersperger, M. Eiber, C. Kroll, H. Kubler, T. Maurer, F. Milletari, R. Julia, C. Schulte zu Berge, S. Enno, B. Frisch and N. Navab, "Multimodal image-guided prostate fusion biopsy based on automatic deformable registration," *International Journal of Computer Assisted Radiology and Surgery*, vol. 10, no. 12, p. 1997–2007, 2015.