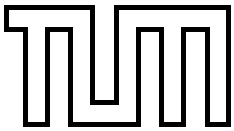
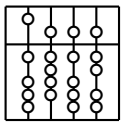


Technische Universität München
Fakultät für Informatik



Calibration of Endobronchial Ultrasound

Kalibrierung von endobronchialem Ultraschall

Diplomarbeit in Informatik

Philipp Dressel

Advisors: Prof. Dr. Nassir Navab
Prof. Kensaku Mori, Ph.D.

Supervisors: Dr. Marco Feuerstein
Tobias Reichl

Date: January 7, 2010

I assure the single handed composition of this diploma thesis only supported by declared resources.

Munich, January 8, 2010 Philipp Dressel

Contents

Abstract

Endobronchial Ultrasound (EBUS) has become a valuable tool for guidance during bronchoscopic interventions. In this thesis a calibration method and associated phantom are developed. This calibration allows for the estimation of the spatial relationship between the optical camera of a bronchoscope and the integrated ultrasound transducer, without the need for a tracking sensor.

Knowledge about this spatial relation, in conjunction with a model of the perspective projection of the camera, enables the mapping of points from the ultrasound into the camera image, and of points from the camera images to lines in the ultrasound image. A possible application for this is the compounding of freehand 3D ultrasound directly within a CT volume. For this, the position of the camera is first determined via an image based camera to CT registration, and the spatial relation is then applied to find the position of the ultrasound plane in CT coordinates.

The proposed method is based on an automatic pose estimation for the camera using a dot pattern. A method based on Z-fiducial pose estimation with hollow rubber tubes is used for the ultrasound plane.

After the required geometrical properties of the phantom are approximated, a precise specification is developed, a phantom is built and the resulting geometry is measured with an optical tracking system.

The achievable accuracy of the proposed calibration method is evaluated. The calibration is compared to another calibration method based on the established hand-eye and single-wall techniques using an electromagnetic tracker for the latter method. The new phantom based calibration is found to be more robust, producing an average transformation with a smaller backprojection error than the established techniques. Calibrations can also be performed much faster with the phantom and do not require a tracking system, thus rendering it an interesting alternative to the established methods.

Acknowledgements

I want to thank Prof. Nassir Navab and Prof. Kensaku Mori for offering me the opportunity to do the research for my thesis at Nagoya University in Japan. Further thanks go to my supervisors Dr. Marco Feuerstein and Tobias Reichl, who have been supporting me throughout the whole time with their advice and ideas.
I am also very grateful towards the Deutscher Akademischer Austauschdienst (DAAD) and Japan Society for the Promotion of Sciene (JSPS) for funding and managing my trip to Nagoya, Japan.
Special thanks also go to Matthew Clarke, who has been a good friend during some troublesome weeks in Japan. Finally, I want to thank Ralf Hartmann for proofreading my thesis and my family, Gela, Werner and Christine for their support throughout my entire studies.

Chapter 1

Introduction

This thesis deals with the development of a calibration procedure for endobronchial ultrasound. A method is developed to find the spatial relation between the embedded fibreoptic camera and the miniature ultrasound transducer of a bronchoscope. The proposed method requires no external tracking solution for the calibration itself. Knowledge about the spatial relation, together with a model of the perspective projection of the camera system allows for the mapping from points in the ultrasound image into the camera image, and from points in the camera image to lines in the ultrasound image.

This chapter details the medical background to interventions using bronchoscopes, ultrasonography and navigated bronchoscopies. Chapter 2 covers the technical background of optical camera calibration, tracked ultrasound calibration and the technological equipment used to perform these. Chapter 3 develops the problem statement of this thesis. The chapters 4, 5, 6, 7 contain the design and implementation of the phantom, the calibration method and the results and conclusions, respectively.

1.1 Cancer Statistics

Three measures are used to describe the burden of cancer on a population: incidence, mortality and prevalence [64]. Incidence is the amount of new cases occuring in the population in one year, expressed either as a total amount or a ratio of cases per 100.000 people. Mortality is the number of deaths induced by a cancer. If the fatality of a cancer is the proportion of cases who die, the mortality is the product of incidence and fatality. Finally, prevalence is the amount of people living with a certain cancer, at one point in time.
In a study on global cancer statistics for 2002, Parkin et al. [64] find a global incidence for all types of cancer of 10.9 million people, a mortality of 6.7 million people and a prevalence of 24.6 million people. Of these cases, lung cancer has the highest incidence with 1.3 million cases. The mortality is also the highest of all cancers with 1.18 million deaths. The ratio of incidence to mortality of 0.78 shows that this cancer also has a high fatality.

Computed Tomography (CT) is a valuable tool in diagnosing lung cancer. As a first stage of diagnosis, possibly cancerous nodules can be identified in a CT scan. To reach a

conclusive diagnosis, tissue samples must then be obtained, and depending on the location within the lung a bronchoscopy may be performed.

1.2 Bronchoscopy

In a bronchoscopy, a flexible endoscopic tool is inserted into a patient's lung without any incisions through the mouth or the nose. The bronchoscope has a fibreoptic camera system, which provides the operator with an inside view of the lung for orientation. It also has a working channel, through which various instruments can be inserted to perform various medical tasks. It is fundamentally important for the diagnosis, staging and subsequent treatment of cancer to obtain samples of the cancerous tissue for histological analysis [78]. This procedure is generally called biopsy. For lung cancer, the location of the lesions within the lung indicates the method to perform a biopsy. For peripheral lesions, e.g. those lying

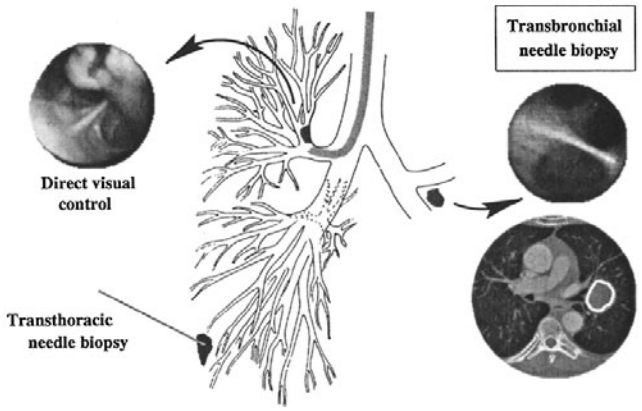


Figure 1.1: Different location scenarios for lung cancer lesions and indicated biopsy techniques. Image courtesy of Bricault et al. [10].

near the surface of the lung or on the chest wall, a transthoracic needle biopsy is indicated. Through a small incision, a needle is inserted into the lesion directly from the outside. This procedure is quick and simple, but it cannot reach lesions which are located deeper within the lung. Endoluminal invasive lesions are inside the lung, but are visible on the inner lung surface. A conventional bronchoscopy allows for a direct, optically controlled biopsy here. The lesions most difficult to reach are hilar or mediastinal, e.g. lying under the inner surface of the lung. The traditional procedures, mediastinoscopy and mediastinotomy are both fairly invasive prodcedures. Although these lesions cannot be seen on the camera image of a bronchoscope, a transbronchial needle aspiration (TBNA) can be performed with a 'blind' puncture at the right spot [78]. This requires a substantial amount of training [22]. Several approaches to provide guidance for transthoracic needle biopsies exist. The movement of the bronchoscope is often viewed under fluoroscopy, but in this case especially small solitary pulmonary nodes (SPNs) are not visible. There are special bronchoscopes with an embedded endobronchial ultrasound (EBUS) probe in use, which can help in the location of lesions not visible under fluoroscopy. Another aproach is the tracking of the bronchoscopes position with a computer system and provision of a navigation interface to the user.

1.2.1 Endobronchial Ultrasound

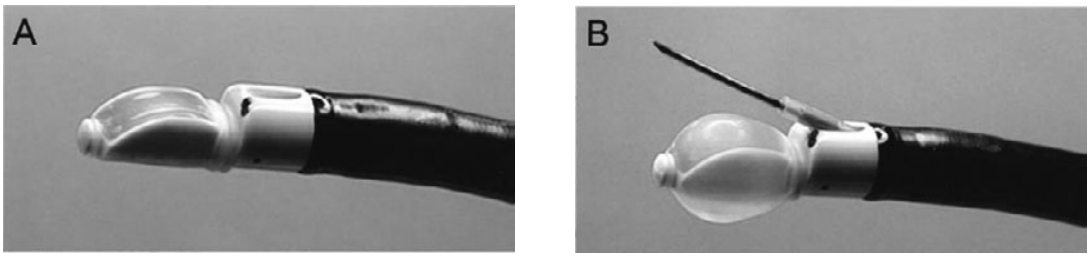


Figure 1.2: Endobronchial ultrasound (EBUS) tool tip containing transducer, camera, working channel exit (A). Image (B) shows inflated baloon and aspiration needle emerging from working channel. Images courtesy of Yasufuku et al. [84].

Recently, conventional camera-based bronchoscopes have been enhanced with the inclusion of a miniature ultrasound transducer. These devices help locating mediastinal or hilar lesions for accurate biopsy needle guidance. The ultrasound images can show lesions which may not be visible under fluoroscopy, and help to improve the success rates of biopsies. However, the manipulation of the probe inside the lung is difficult, and it is impractical to scan large areas of the lung. Thus, the approximate location of the lesions need to be known in advance. Also, the size of the instrument is greater due to the embedded ultrasound transducer, limiting the depth to which it can be inserted into the lung and thus the reachable positions.

Ultrasonography is one of the most commonly used imaging modalities in medicine. In its basic form, it provides two-dimensional images of a patients anatomy in real-time by emitting high-frequency sound waves and analyzing the echo. It requires neither the application of radiation nor special operating environments. Different anatomical regions of a patient can be imaged directly by simply moving the ultrasound probe. Compared to other well established imaging modalities such as Computed Tomography or Magnetic Resonance Imaging it is relatively cheap, and usually does not cause the patient any discomfort. It is well suited for imaging soft tissues, however since the sound waves do not penetrate bones well, these cannot be imaged, and, furthermore, may obstruct the view on tissue behind them. As a further limitation, ultrasound waves do not propagate well through air, requiring the probe to be in close contact to the surface, usually coupled with a water-based gel. Since the transducer of an EBUS cannot always easily be pushed against the lung surface, a baloon is often placed around it, which is then inflated with water.

1.2.2 Navigation

Locating lesions or other points of interest inside the lung can be supported by computer systems and CT or MRI datasets. Virtual bronchoscopy (VB) is a simulation of a real bronchoscopy (RB) on a computer terminal, controlled with a mouse or similar input device. The view of a virtual bronchoscope is generated from the three-dimensional datasets in real-time, as the virtual bronchoscope can be moved to examine a patient's anatomy [30, 70, 57, 38]. This allows the surgeon to observe the planned trajectory of the bronchoscope from all positions and directions. Quantitative measurements can be done on-the-fly and even the visualization of anatomy beyond the organ walls is possible. The logical extension

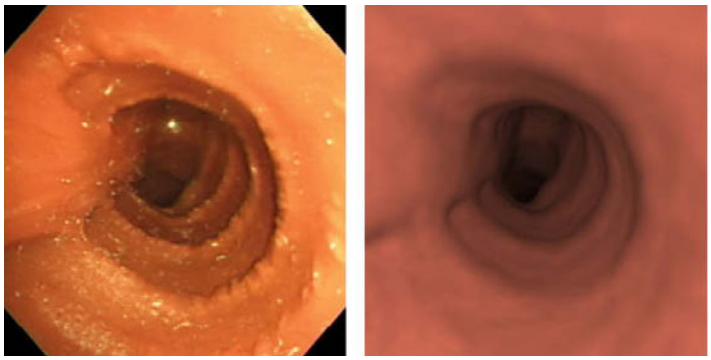


Figure 1.3: Real bronchoscopy image (RB, left) and CT-derived virtual bronchoscopy image (VB, right) used for registration. Image courtesy of Deguchi et al. [23].

to this system is to provide its additional information during the actual intervention [36, 37]. This requires the continuous tracking of the camera's position and orientation with respect to the coordinates of a preoperative CT scan. Additional information from this scan can then be fused into the camera image, and furthermore the bronchoscope tip's location can be shown on the CT slices, providing the surgeon with a live map for orientation.

The major challenge for navigated bronchoscopy systems is the accurate location of the bronchoscope with respect to the CT data. Electromagnetic tracking equipment can be used, since it does not require a line of sight to the tracking targets. (Refer to section 2.1 for details on tracking systems.) Measurement inaccuracies are a major challenge, resulting both from the characteristics of electromagnetic tracking and the distortion caused in the soft tissue during breathing. While this type of tracking is valuable for getting an approximate position, the accuracy is not sufficient for robust localization [72].

Bricault et al. [10] first proposed a system which used the registration of RB and VB images to find the camera's position and orientation. Here, the virtual camera is moved around, and the resulting images are compared to the real video image. By maximizing the similarity between the two, the position of the camera in the lung is found. This approach generates accurate position data, however it needs a good initial guess for the position and is sensitive to defining features in the video images, such as bifurcations. If the image is obscured through coughing or bubbles in the lung, or shows just a homogeneous pink lung wall, no position can be estimated. Furthermore, if the position is lost during quick

movement, recovery based on image registration alone may be difficult, since there is no longer a good initial guess for the position available.
Helferty et al. [35] later introduced a refined method for image registration based on mutual information, which is more robust against visible image features. Another system was proposed by Mori et al. [56] which also uses only RB and VB images. Their system used a two-step method, which first computes an approximate camera motion estimate using RB video analysis, followed by an accurate RB to VB image registration using mean squared errors or cross correlation algorithms. They later also developed a hybrid approach, in which the bronchoscope is tracked electromagnetically to provide initial guesses for a subsequent image based registration [58].
Merritt et al. [53] improved on the previous systems with a method capable of real-time image registration in unter $1/15$th of a second, allowing for greatly improved system response and accuracy. Another approach by Deligianni et al. [24] used a shape-from-shading, pq-based registration approach in combination with a deformable breathing model. They also used an EM tracking system to recover from situations where image registration is not possible, and an improved VB image renderer based on bidirectional reflectance distribution function (BDRF) recovering [17].

Chapter 2

Technical Background

This chapter depicts the technical resources and methods used to design, build and measure the proposed calibration technique. The first section describes methods for tracking objects in 3D space. Even though the actual calibration does not require a tracking system, both optical and electromagnetic tracking systems were used during design and construction of the phantom, as well as during the experiments. The second section details camera calibration methods. The third section covers ultrasound calibration, and the last section details the bronchoscope used in this thesis.

2.1 Tracking

This section explains optical and electromagnetic tracking in more detail. Tracking systems in general allow for the location of an object within a confined space. This localization is relative to a base unit's coordinate system which is often itself registered to another coordinate system. There are four different technologies commonly used, which are mechanical, acoustical, electromagnetic and optical tracking solutions. A detailed description of tracking technologies can be found in [54, 18]. In mechanical tracking, the position of an object attached to a flexible arm is computed from the angles of its joints. This tracking method is accurate, however one arm can only track one object [18].
Acoustical tracking is achieved by emitting ultrasound from speakers which is then received by a microphone. Using either a time of flight or phase shift approach, the distance between speaker and microphone can be computed, and for multiple distances a position can be triangulated [18].

2.1.1 Optical Tracking

Optical tracking systems use infrared light to locate features of a structure with a known geometrical configuration. If these features are detected through at least two cameras, the location and orientation of the structure in space can be accurately reconstructed. For this, the target structure needs to be located within a tracking volume, which is bound by the construction of the tracking device. The target structure itself can either be active or passive, and hybrid tracking systems exist which can locate both types of targets. The simultaneous

tracking of multiple targets is possible with both active and passive targets. Figure 2.1 shows the NDI Polatis Spectra tracking system, a sample object and the geometry of the tracking volume.

Active Markers

Active markers consist of at least three light emitting diodes operating in the infrared spectrum. A control unit successively powers on each diode, allowing for individual location through the cameras. Targets with active markers require a power source, and are thus often cumbersome to use. However, they allow for arbitrary geometrical placement of the diodes and do not suffer from deteriorating reflective surfaces as passive markers do.

Passive Markers

For the use of passive markers, an infrared illumination system is integrated with the camera's chassis. The target itself is built from IR reflective spherical markers. This has the advantage that the target does not require a power source and can thus be used more flexibly. However, since individual points of the target cannot be matched by switching them on and off as with active markers, the spheres need to be placed asymetrically, leaving no ambiguity about the orientation of the tool. Furthermore, if more than one target is to be tracked, each target needs to be of a unique configuration. Another drawback is a reduced accuracy if the reflective spheres are not cleaned properly [83].

West and Maurer [82] explain the challenges of target design. Optical tracking systems achieve a very good accuracy compared to other systems, are robust against the presence of metal and electromagnetic fields and are very tolerant to different lighting conditions. However, they require a continuous line of sight between at least two cameras and the target, otherwise tracking is interrupted. Some systems use more cameras than the minimum requirement of two to provide redundancy, but in any case applications need to be designed with the possibility for lost tracking. Localization of targets entirely inside a patient is also not possible with optical tracking.
A Polaris Spectra passive infrared tracker manufactured by Northern Digital Inc. was used in this thesis to measure the geometry of calibration phantoms. An overview of the system's accuracy under different conditions is given in [83, 27]. Generally, optical tracking systems range among the most accurate tracking methods available.

2.1.2 Electromagnetic Tracking

Electromagnetic tracking systems, first published by Raab et al. [67] measure the induced current in a sensor coil placed in an electromagnetic field. Both direct current (DC) and alternating current (AC) field variants exist, and both are generated by various types of transmitter units. This technology has the outstanding advantage of not requiring a line of sight to the target coils and thus being applicable easily inside a patient. However, it also suffers from accuracy usually an order of magnitude worse than optical tracking solutions. A further drawback is that the measurements are sensitive to distortion from metallic objects and

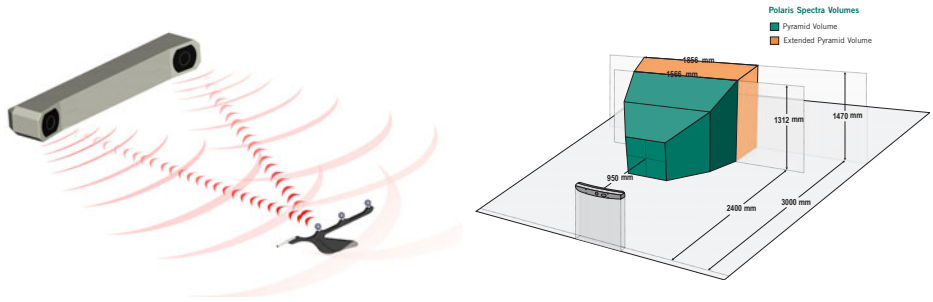


Figure 2.1: NDI Polaris Spectra infrared optical tracking system (left) and tracking volume dimensions (right). Images taken from Northern Digital Inc. documents [44] and [45].

foreign electromagnetic fields, both of which are common in an operating room [5]. Refer to Kindratenko [47] and Birkfellner et al. [5] for details on the physical effect that cause these distortions.
A Northern Digital Inc. Aurora alternating current electromagnetic tracking system was used for the experiments in this thesis. Furthermore, an Ascension Microbird direct current tracking system was also evaluated, but discarded in favor of the Aurora tracking system. This choice was based on the readily available sensors, their respective coil sizes and resulting accuracy characteristics.

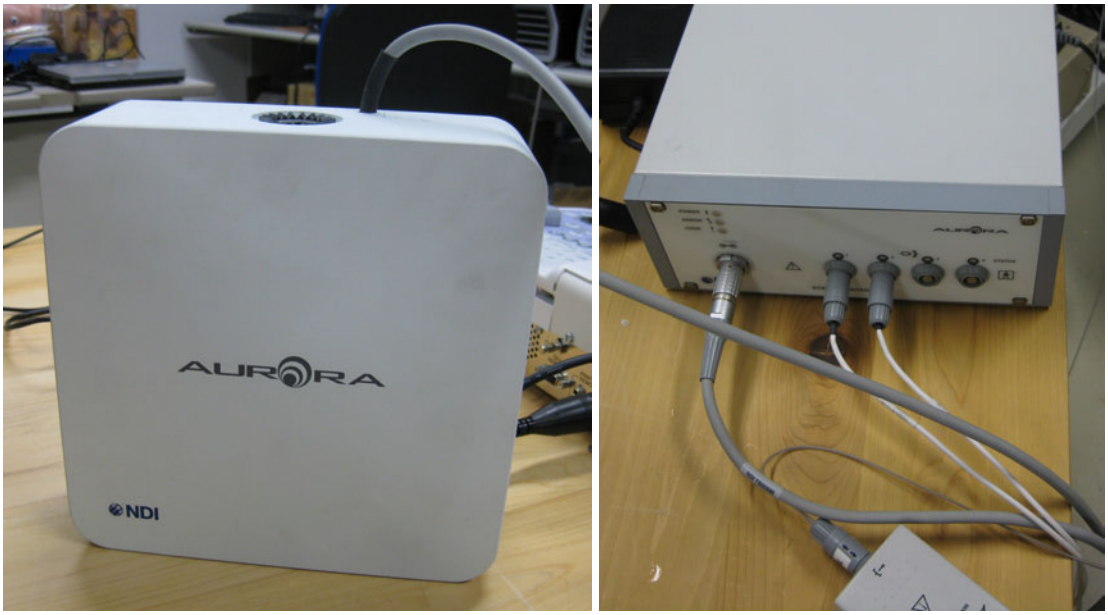


Figure 2.2: NDI Aurora electromagnetic tracking system field generator (left) and sensor connection box (right)

2.1.3 Pointer Tip Calibration

When a tool is located with a tracking system, usually the position and orientation of a specific point on the tool is of interest, as opposed to the location of the attached tracking target. If the tracked tool is a pointer, the hotspot calibration method provides a very simple and accurate way to compute the fixed relation between an attached tracking target and the pointer's tip. After calibration, this transformation can then be applied to the tracking target in real-time, resulting in the location of the tool's tip in the coordinate system of the tracker. The calibration is performed by placing the tip of the pointer in a mold or similar fixture, to lock its position in place relative to the tracking system's world coordinate system (e.g. the tracking device itself). The tool is then rotated around this pivoting point, covering as much of the hemisphere above the tip as possible. Both iterative and closed form solutions

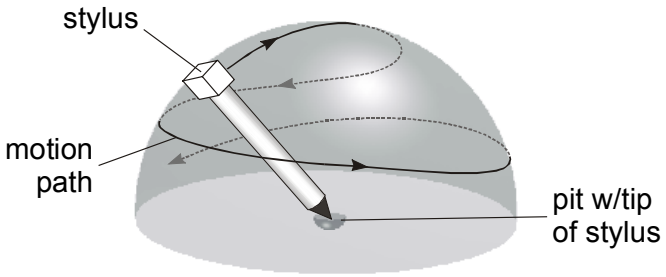


Figure 2.3: Hotspot pointer tip calibration. Image courtesy of Fuhrmann et al. [28].

to compute the target to tip transformation exist.

Iterative Solution

Fuhrmann et al. [28] present a solution in which the spatial relation is iteratively optimized by using the variance σ of the tip location offset for the measured target positions as a metric:

$$\sigma = \frac{1}{n-1} \sum_{i=1}^{n} (H_i \cdot p_t - \overline{p_w})^2 \tag{2.1}$$

Where H_i are the world to tracking target transformations, p_t is the desired tip location in target coordinates and $\overline{p_w}$ is the mean of all computed tip locations in world coordinates. This yields a least-squares fit solution for the tip offset p_t. To ensure numerical stability of the solution, the pointer movement needs to cover as much of the hemisphere as possible.

Closed Form Solutions

The implementation used in this thesis to compute the pointer tip location is based on a closed form solution [77]. The translations t_i and rotations R_i for the tracking target in world coordinates are assumed to fulfill the following equation with respect to the pointer tips' coordinates in world space p_w and tracking target space p_t:

$$p_w = (R_i \, t_i) \, p_t \tag{2.2}$$

A linear equation system $A \cdot x = b$ is constructed, with A composed of the rotation matrices R_i:

$$A = \begin{pmatrix} I & -R_1 \\ I & -R_2 \\ \vdots & \vdots \\ I & -R_n \end{pmatrix} \tag{2.3}$$

b is composed of the translation vectors t_i and $x = (p_w p_t)^T$. This equation system is then solved with a QR-decomposition.

2.2 Camera Calibration

The goal of camera calibration is to find a set of parameters that define the mapping beween points in 3D coordinates and their projection onto a two-dimensional image plane. One subset of these parameters is describing the internal geometric and optical characteristics of the camera (intrinsic parameters), while the other determines the camera position and orientation in a world coordinate system (extrinsic parameters) [75]. In computer vision, the origin of the local coordinate system of an image is usually assumed to be in the upper left corner of the image. With a pinhole camera model, the intrinsic camera parameters are defined as focal length in pixel dimensions α_x, α_y and principal point or image center x_0, y_0, also in pixels. For the unusual case of non-perpendicular image axes, a skew factor s is included, which remains $s = 0$ for most real-world cameras [33]. (However, if for example pictures of a picture are taken, the skew factor may change to $s \neq 0$.) The projection of a point onto the image plane in homogeneous coordinates up to scale λ is represented by the following formula:

$$\lambda \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \left[\begin{array}{c|c} K & 0 \end{array} \right] \cdot \begin{pmatrix} u \\ v \\ w \\ 1 \end{pmatrix} \tag{2.4}$$

where the camera matrix K is composed as:

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \tag{2.5}$$

Several generic algorithms for estimating the camera parameters have been proposed, with the method of [34] being the most widely adopted. A free implementation of this method is available as a MATLAB toolbox from [8]. It is based on the detection of corners in a black and white chessboard image which are matched to the known 3D geometry of the pattern. However, since the images produced by the miniature optics of a bronchoscope tend to be highly distorted, this method often has difficulties to correctly match the extracted points, making a calibration cumbersome and tedious. Several authors have studied methods for more effective calibration from highly distorted images [86, 73, 81], with the method of Wengert et al. being used for the calibration in this thesis. An implementation as a MATLAB toolbox for the method of Wengert et al. is available from [80].

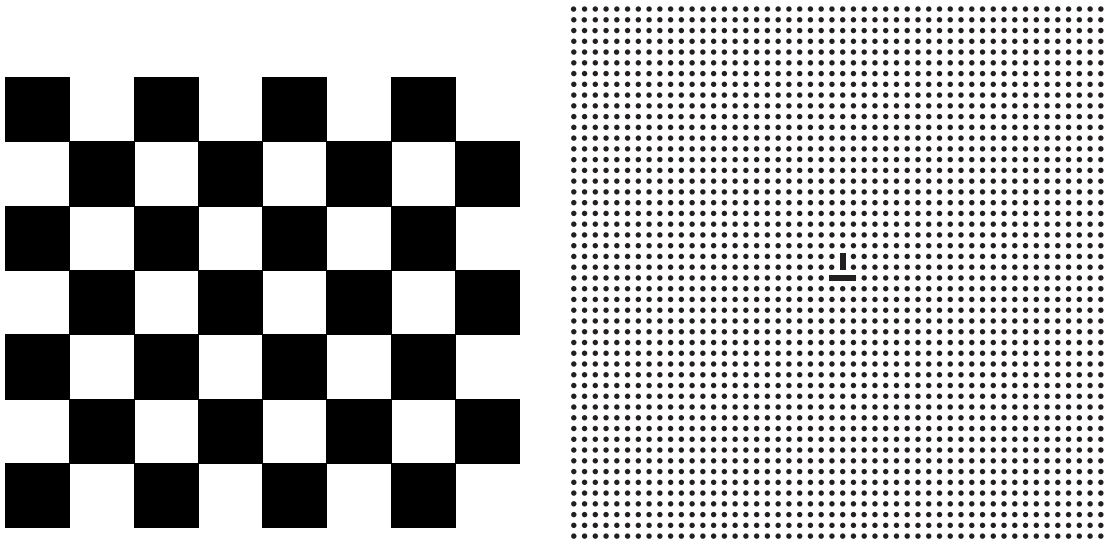


Figure 2.4: Calibration patterns used by [34] (left) and [81] (right). Images taken from the MATLAB implementations of [8] and [80].

2.2.1 Distortion Correction

The pinhole camera model is not accurately describing real-world camera optics, since camera lenses distort the image as the light passes through them. Two sets of parameters are used to model this distortion, the first for tangential distortion (p_1, p_2) complemented by several parameters for radial distortion $(k_1, k_2, ...)$. Usually, only two parameters k_i are enough to describe radial distortion. The equation to compute the undistorted coordinates p_u for any given point $p_d = (x, y)^T$ on the image plane is thus:

$$p_u = \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right) p_d + \begin{bmatrix} 2\,p_1\,x\,y + p_2(r^2 + 2x^2) \\ p_1\,(r^2 + 2y) + 2\,p_2\,x\,y \end{bmatrix} \tag{2.6}$$

Refer to Heikkilä et al. [34] for more information on image distortion and correction.

2.2.2 Underwater Camera Calibration

The optical properties of a camera system depend on which medium the rays reaching the lens are travelling through. If a camera is calibrated in air and placed underwater, the intrinsic camera parameters change. In computer vision, the pinhole camera model for perspective projection is widely used. This model does not take the possibility of different optical indices for fluids on either side of the lens (or pinhole) into account. Lavest et al. consider the more generalised thick lens optical model [50] for a homogeneous medium, and later expand this to the application of underwater cameras [49]. They describe both theoretical and practical concepts and find a relationship between calibrations performed in water and air, allowing the intrinsic camera parameters to be adapted when a camera is placed in a different medium.

Thick Lens Model

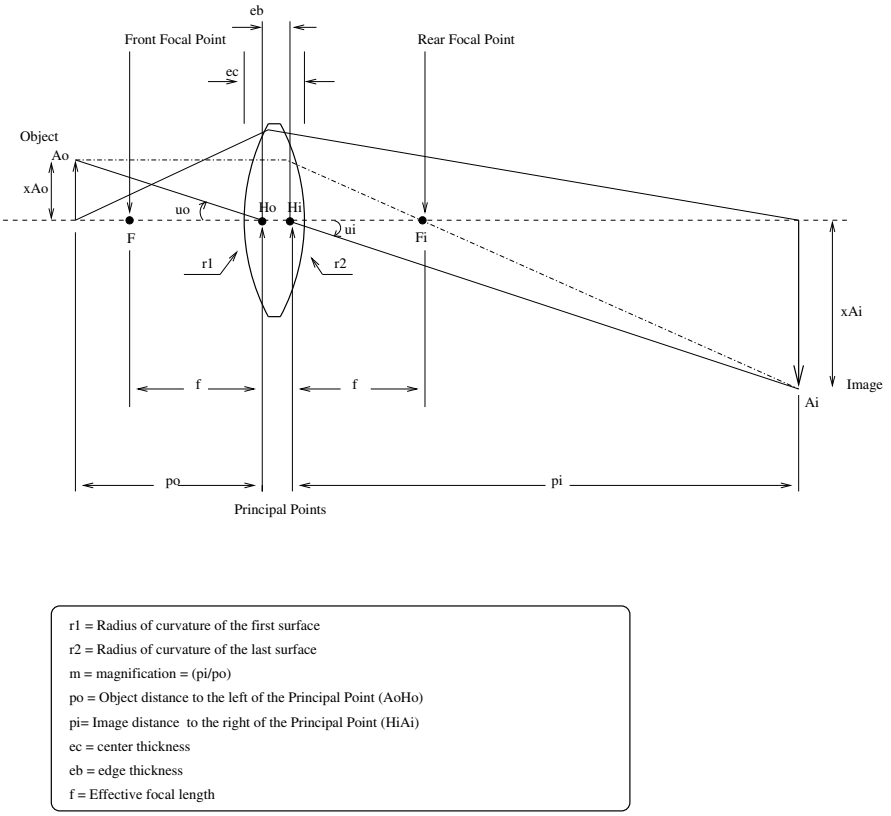


Figure 2.5: Thick lens model for homogeneous media. Image courtesy of Lavest et al. [49].

The point A_i is said to be the *conjugate* (or image) of the point A_o if the optical system makes the rays from A_o converge at the point A_i. Subsequently, all conjugate points for a set of distances p_o and p_i lie on a *conjugate plane*.
The ratio of the distances x_{A_i} and x_{A_o} of two conjugate points to the optical axis are the *transversal magnification*:

$$Gt = \frac{x_{A_i}}{x_{A_o}} \tag{2.7}$$

The *angular magnification* is given by the ratio of the angles u_i and u_o of two conjugate points:

$$Ga = \left(\frac{u_i}{u_o}\right)_{x=0,y=0} \tag{2.8}$$

The conjugate planes for which the transversal magnification is exactly one are the *principal planes*. The point where the optical axis intersects the principal planes are the *principal points*. The *nodal points* are analogously the two conjugate points for which the angular magnification is exactly one. For optical systems in a homogeneous medium, the focal and nodal points fall within the same location, which is consistent with the classic pinhole camera model.

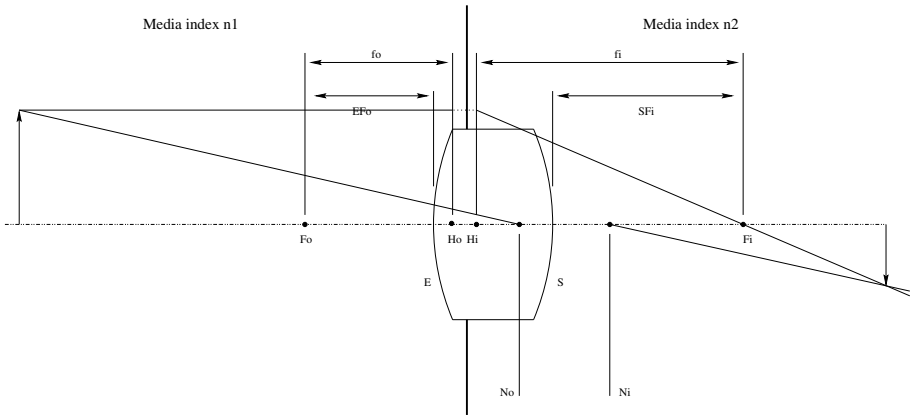


Figure 2.6: Thick lens model for different media. Image courtesy of Lavest et al. [49].

Inhomogeneous Media

Lavest et al. modify the thick lens model for homogeneous fluids to account for different refractive indices of the object (n_1), image media (n_2) and lens material (n). This implies that there are now two distinct focal lengths for the object and image side of the lengths, and nodal and principal points no longer coincide. The refractive indices govern the optical variables as follows:

$$k = \frac{n - n_1}{r_1} + \frac{n_2 - n}{r_2} - \frac{tc(n - n_1)(n_2 - n)}{n \cdot r_1 \cdot r_2} \tag{2.9}$$

$$k = \frac{n_1}{p_1} + \frac{n_2}{p_2} \tag{2.10}$$

$$f_o = \frac{n_1}{k} \tag{2.11}$$

$$f_i = \frac{n_2}{k} \tag{2.12}$$

$$EH_o = \frac{n_1 \cdot tc(n_2 - n)}{k \cdot n \cdot r_2} \tag{2.13}$$

$$SH_i = \frac{-n_2 \cdot tc(n - n1)}{k \cdot n \cdot r_1} \tag{2.14}$$

$$H_oN_o = \frac{n2 - n1}{k} \tag{2.15}$$

$$H_iN_i = \frac{n2 - n1}{k} \tag{2.16}$$

$$EN_o = EH_o + H_oN_o \tag{2.17}$$

$$SN_i = SH_i + H_iN_i \tag{2.18}$$

Implication for the Pinhole Model

In the pinhole model the focal length is usually defined as the distance between the image sensor and the optical center. The thick lens model for inhomogenous media implies that the

optical center is shifted to the fusion of the two nodal points which conserve the angular magnification. The focal length is thus equal to the distance N_iF_i if the object is at infinity.

$$f_o = N_iF_i = N_iH_i + H_iF_i = \frac{(n_1 - n_2)}{k} + \frac{n_2}{k} = \frac{n_1}{k} \tag{2.19}$$

$$f_o = n_1 \cdot \frac{r_2}{(n_2 - n)} \tag{2.20}$$

Since r_2, n_2 and n are constant for a given camera system, this implies that the focal length is only dependent on the refractive index of the outside medium. In the case of a camera calibrated in air that is placed in water, the factor was empirically detemined to be 1.333. This enlarged focal length is also matching the observed image magnification when a camera is submerged in water.

Underwater Camera Distortion

With the reduction to a pinhole model, the problem of image distortion is apparent again. Lavest et al. show that the linear magnification of the image is also affecting the distortion in the same way [49]:

$$1.333(u + du) = u' + du' \tag{2.21}$$

With u being the the distorted image of a point in air and du the neccessary distortion correction to obtain a perfect perspective projection, and u' and du' their counterpart of a point in water. The coordinate system of these points needs to lie within the image center, so image coordinates must be offset with the principal point. This distortion scaling is valid for both radial and tangential distortion. Any point p in image coordinates that was projected using the camera matrix and distortion coefficients for air can be transformed into underwater camera coordinates p' through the following equation:

$$p' = \alpha + (1.333 \cdot (p - \alpha)) \tag{2.22}$$

Where α is the principal point.

Experimental Validation

Lavest et al. validate their theoretical model with calibrations performed both in air and under water for the same camera, finding a ratio for the respective focal length that matches the theoretical prediction very well [49]. They conclude that it is possible to calibrate a camera in the air and accurately predict the intrinsic camera parameters if the camera is placed in water.

2.2.3 Hand-Eye Calibration

The aim of a hand-eye calibration is to find the spatial relation between a camera and a position sensor rigidly attached to it. This may be the joint of a robot (from which the term hand-eye calibration originates) or a tracking target of a tracking system. The calibration is performed by moving the camera to different positions with distinct orientations, while

facing a calibration pattern. The features of the pattern are automatically detected, and if their geometry is known, the camera's relative position and orientation towards the pattern can be computed. The tracking target's position and orientation are recorded for the same positions.

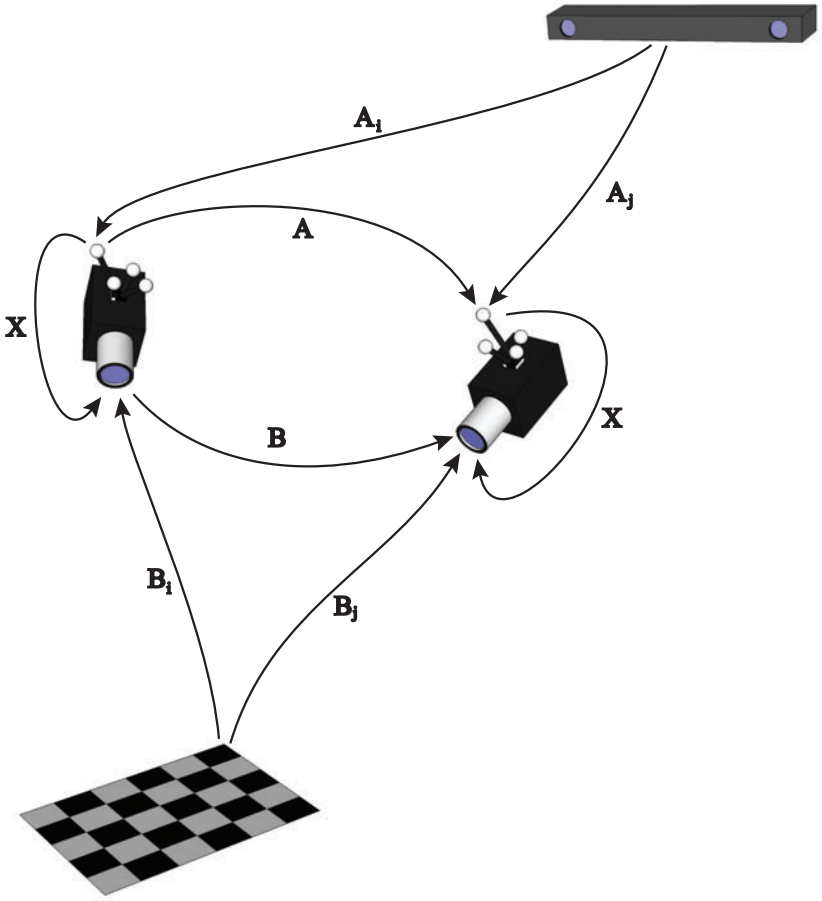


Figure 2.7: Hand-Eye calibration. The figure shows two consecutive positions for a single camera and the involved transformations.

If B denominates the transformation between two positions in the camera's coordinate system, A denominates the transformation between two positions in the tracker's coordinate system and X is the (constant) transformation between camera and sensor, the following linear equation system can be constructed:

$$A \cdot X = X \cdot B$$

Several methods for solving this equation have been proposed [76, 21, 16, 63]. In this thesis an implementation based on the approach by Tsai et al. [76] included in the calibration toolbox [80] is used.

2.3 Ultrasound Calibration

To produce freehand 3D ultrasound images with a standard 2D imaging probe, the trajectory needs to be measured as the probe is moved over the target region[41]. This allows the recorded ultrasound images to be correctly aligned in space in accordance with the probe's position at the time the images were acquired.
The following equation allows for the mapping of each pixel of a B-scan with image coordinates u, v and scaling from pixels to millimeters s_x, s_y to be mapped to a reconstruction volume:

$$ {}^{V}p = {}^{V}H_T \, {}^{T}H_S \, {}^{S}H_U \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix} \tag{2.23} $$

where ${}^{V}p$ is the location of a mapped pixel in the reconstruction volume in homogeneous coordinates. In the above notion, the matrix ${}^{B}H_A$ represents the homogeneous matrix that transforms from the coordinate system A to the coordinate system B. The involved coordinate systems in the above equation are U for the ultrasound plane, S for the position sensor, T for the tracking system and V for the reconstruction volume. The transformation ${}^{S}H_U$ is initially unknown and needs to be determined through spatial calibration. the transformation ${}^{T}H_S$ is provided by the tracking system in real-time as the probe is moved. Finally, the transformation ${}^{V}H_T$ is usually defined by the application to place the reconstruction volume size at a convenient location.

2.3.1 Spatial Calibration

The purpose of spatial ultrasound calibration is to find the rigid-body transformation ${}^{S}H_U$ between the ultrasound plane and a position sensor attached to the probe. While the spatial relation beween an ultrasound probe and position sensor could be measured physically, it is difficult to find the exact location of the position sensor's local coordinate system within it's coils. The image plane's origin is equally hard to measure physically. Extensive research has been conducted over the past years on the use of the images and position readings themselves to accurately estimate this relation. The equation governing the mapping between image pixels and the reconstruction volume 2.23 is used in a modified way to model the calibration process in general:

$$ {}^{P}p = {}^{P}H_T \, {}^{T}H_S \, {}^{S}H_U \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix} \tag{2.24} $$

The transformation ${}^{V}H_T$ is replaced by ${}^{P}H_T$ which transforms from the tracking system's coordinate system T to a defined origin of a scanned object P. This object, usually referred to as a phantom, has some known geometrical properties. This transformation is usually also initially unknown, albeit constant for all scanned images provided the phantom is not moved. Some methods use a second position sensor attached to the phantom to allow for movement of the phantom during calibration and eliminate the unknown transformation

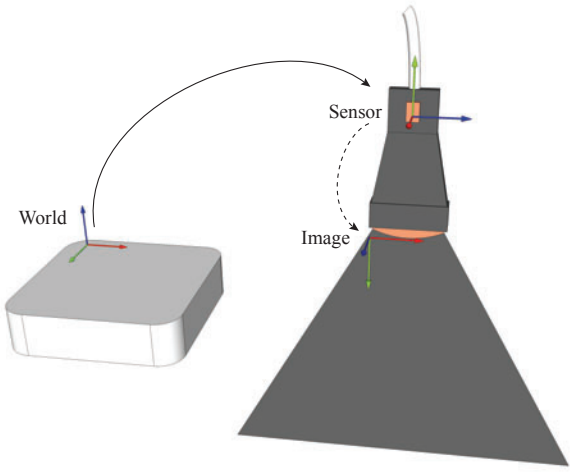


Figure 2.8: Spatial ultrasound calibration. The figure shows a tracking system and ultrasound probe with image plane and attached position sensor.

from the calibration process. The calibration procedure itself can be broken down into three distinct parts. These are data acquisition, feature localization and computation of the calibration parameters which define the transformation ${}^{S}H_{U}$.
Over the last years several calibration methods have been developed. While the general procedure and assumptions for ultrasound calibrations are similar, the individual steps involved in a calibration differ significantly for the different techniques. In general, four different classes of calibration phantoms exist. The cross-wire [3, 25, 32], three-wire [12], ball bearing [1] and ping-pong ball [9] phantoms are designed around matching single points in space to a location in the ultrasound image. Using a series of images taken from different positions the solution is then found through numerical optimization algorithms.
The second class of calibration phantoms works through imaging of a plane, producing a line in the ultrasound image. The line has two degrees of freedom, and by carefully imaging the plane such that all six degrees of freedom for the ultrasound probe are covered, enough constraints can be established to find a unique solution for the calibration. The single-wall phantom [66] in its various forms and the Cambridge phantom [66] are buit around this concept.
Another class of phantoms is build around the imaging of two-dimensional shapes. Provided that at least three non-linear features of known geometric properties can be found, the solution to the calibration can be computed from a single image. This holds the potential for very quick calibrations, however the alignment of the scan plane with the phantom is challenging. The 2D shape alignment [71] and precision mechanical device phantoms [29] are members of this class of methods. The z-fiducial phantom [19] can be classified as a 2D shape alignment phantom as well, since the fiducial positions estimated from the 3D fiducials form a virtual 2D phantom.
The last class of calibration methods are referred to as *phantomless calibration*, since they rely only on the use of a 3D pointer [59] or tracked rod [46], while the probe remains in a fixed position. The sections 2.3.4, 2.3.5, 2.3.6 and 2.3.7 explain *point-based*, *wall-based*, *shape-based* and *phantomless* calibration methods, respectively.

Data Acquisition

The incoming position readings from the tracking system and the images produced by the ultrasound system need to be received, timestamped and matched into pairs. Refer to section 2.3.3 for details on data matching. Depending on the calibration method, one or several data pairs need to be recorded and processed. The tracking data is usually received through a serial or USB interface, and accessed through the tracking system vendor's API. The ultrasound images are commonly transferred as an analog video signal from the ultrasound machine and digitalized by a framegrabber card. This method works with virtually any available ultrasound machine, but it comes at the price of a loss in quality due to the analog transmission.
An alternative is the digital transfer of the images, in which either the raw data from the probe or digital images are transferred. This however requires vendor- and model specific interfaces, and is thus rarely used.

Feature Detection

The defining features of the calibration phantom need to be segmented in the ultrasound image(s). Depending on the method and phantom, this process can be manually performed by the user, or automatically through image processing. The point based phantoms are generally challenging to segment automatically, because of the difficulty to distinguish them from image noise. While these features are mostly segmented manually [46, 25, 1], a semi-automatic algorithm was designed by Carr et al. [13] which requires only the selection of a region of interest. Lindseth et al. even developed a fully automatic feature detection [51] which detects several points in an image in a pyramidal shape. The detection of features for a Z-fiducial phantom was automated by placing a rubber membrane at the top, to help automatically identify possible locations for fiducial wires [41].
The lines produced by wall based methods are generally more easily detected, since a line contains a lot of redundant information. It is even possible to detect a line robustly if parts of it are invisible or distorted. Prager et al. detect lines automatically by tracing vertical lines at regular intervals from the top of the image. The image is first smoothed and then an edge detection operator is applied. When a certain treshold is exceeded along the lines, this point is assumed to lie on the topmost line in the image [66]. A line is then fitted to the points using the random sample consensus algorithm [26]. Rousseau et al. [68] chose a different approach in which the lines are detected by applying a Hough transform [39] to the images.

Parameter Computation

Determining the calibration parameters is the last step. This always involves finding the 6DOF rigid transformation from the ultrasound plane to the position sensor, and sometimes also the two scaling factors for converting image pixels into mm. Since the images and tracking measurements are noisy, so are the detected point sets, and there is generally no exact solution for a mapping between them. These overdetermined equation systems are thus solved for the minimum residual error. This can be achieved both by closed form solutions (i.e. the

method of Umeyama [79]) or by iterative algorithms of which the iterative closest point [85] and Levenberg-Marquard [55] algorithms are most common.

2.3.2 Speed of Sound

The speed of sound c_s in water is dependent on the water temperature and the presence of solutes. Ultrasound machines compute the images based on an assumed $c_s = 1540\frac{m}{s}$, which is the average speed of sound in human tissue. Some machines allow for the adjustment of the assumed c_s, but this is not common. If the speed of sound is actually slower than the assumed value, objects will appear farther away in the image. If is is higher, they will appear nearer than they actually are. This effect and the resulting distortions on phantoms were researched by Goldstein [31].

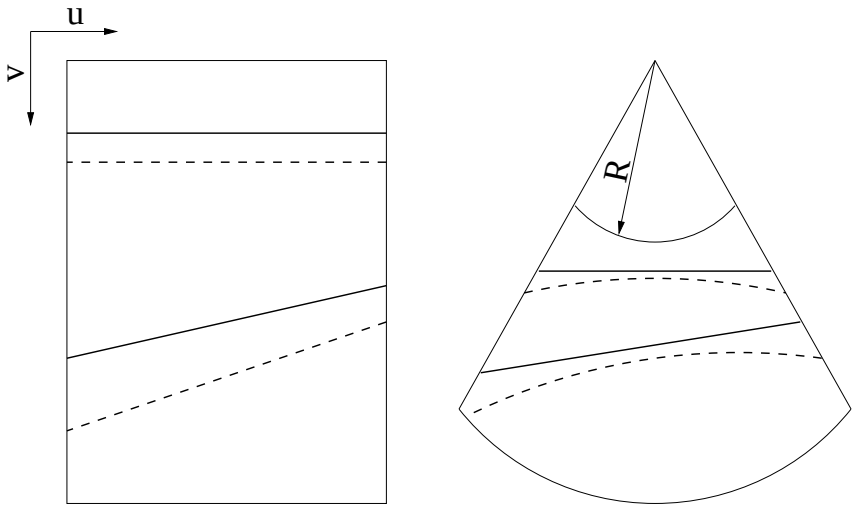


Figure 2.9: Ultrasound distortion due to speed of sound effect: the dotted lines show the distorted image of a plane, the solid lines are the actual locations. Image courtesy of Hsu et al. [40].

The speed of sound in pure water at room temperature is roughly $1490\frac{m}{s}$ [15], so objects will appear too far away from the transducer, leading to inaccurate measurements. There are several approaches to compensate this.
The speed of sound is dependent on the water temperature [4], and by raising it to $48°C$ the speed of sound can be matched to that of human tissue. Boctor et al. use this approach for their calibration method [6]. However, it is not trivial to keep a water bath at exactly this temperature and it may be inconvenient to manipulate a probe in this warm water.
Another way to raise the speed of sound is to add table salt (sodium chloride) in the right concentration to the water. Chen et al. describe the effect of different chemicals on the speed of sound in water [14]. Brendel et al. use this approach to compensate the speed of sound problem in their calibration [9].
The distortion can also be compensated after image acquisition. By estimating the speed of

sound as a function of the current water temperature [4], the ratio R of the actual speed of sound c_r versus the assumed speed of sound by the ultrasound machine c_a can be computed:

$$R = \frac{c_a}{c_r} \tag{2.25}$$

Based on this ratio any point in the image can be scaled relative to the center of the arc of a radial transducer in the following way:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{R}{\sqrt{x^2 - y^2}} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \tag{2.26}$$

Finally, Pagoulatos et al. [62] among others use gelatinous materials which directly mimick the properties of tissue. After the phantom is assembled these materials solidify, making subsequent manipulations of the phantom cumbersome.

2.3.3 Temporal Calibration

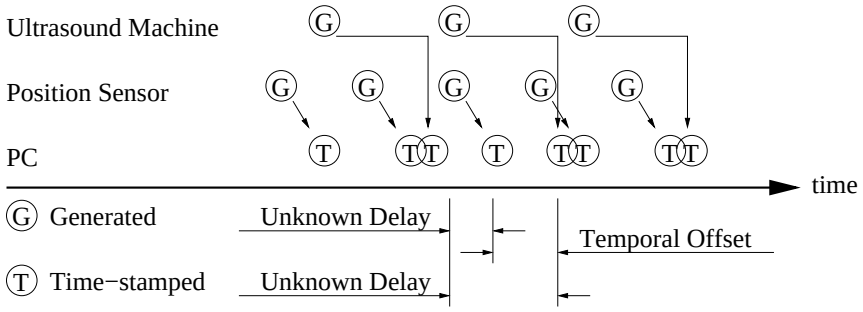


Figure 2.10: Temporal ultrasound calibration. Data is generated (G) by the ultrasound machine and tracker and timestamped at arrival on the maching (T). Image courtesy of Hsu et al. [41].

The position data and ultrasound images are generated by separate hardware without any synchronization. This means that usually both data streams are generated at different frequencies and the hardware devices produce an unknown delay during the processing and transmission of the data. In the common setup for a freehand 3D ultrasound system, a computer is used to combine the two streams of data, timestamping each individual position and image as it arrives on the computer. The aim of temporal calibration is to find the relative temporal offset between the two data streams, so that data pairs which were recorded at the same time can be found.
If the temporal calibration is performed before the spatial calibration, the location of the phantom is unknown at this time [65, 42]. The direct approach to this problem is to detect a sudden step in the input data streams, for example by quickly removing the probe from a surface [65, 52]. The temporal difference between the detected steps in the data streams is then used as the offset for the temporal calibration. This method can only be accurate up to $\Delta T = \frac{t_1+t_2}{2}$ with t_1 and t_2 the temporal resolutions of the image and tracking data

streams. Rousseau et al. found that for ultrasound images at PAL temporal resolution ($25Hz$) accompanied by tracking data of equal or greater acquisition rate are sufficient for clinical volumetric measurements [69].
Another approach is to scan a wire [11] or plane [74] while moving the probe in one direction. By approximating the location of the object directly from the images, the temporal calibration is then computed by finding the maximum correlation between the two data streams.
If the temporal calibration is performed after the spatial calibration, there is no temporal correction during the spatial calibration available. This can be compensated by holding the probe stationary for a while at each scan position to allow the data streams to match up [32, 40]. The known spatial location of the phantom is then used to compute the temporal offset between the data streams [60].

2.3.4 Point Based Calibration Methods

Cross-Wire Phantom

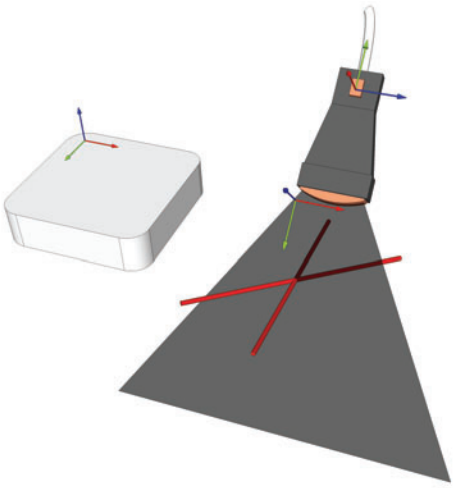


Figure 2.11: Cross-wire calibration phantom. The ultrasound probe is placed to image the crossing point of the two wires from different positions.

The cross-wire phantom, proposed by [3, 25, 32] aims to find the crossing of two wires in the ultrasound image. This point is also defined as the phantom's origin. A series of points is found for successive scans from different positions.
The transformation ${}^{S}H_{U}$ into the sensor's coordinate system which is valid for all these points can then be estimated through numerical optimization algorithms, where each segmented point should satistfy the following equation:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = {}^{P}H_{T}\,{}^{T}H_{S}\,{}^{S}H_{U} \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix} \tag{2.27}$$

The accuracy achievable with this method depends strongly on the accurate localization of the crossing point in the ultrasound images. This method also has a few substantial drawbacks. Since the segmentation of single points in an ultrasound image is usually not reliable, the points are mostly segmented manually [3, 1, 20], which is tedious and time consuming. Furthermore, aligning the ultrasound probe such that the crossing of the wires lies exactly within the ultrasound plane requires skill, since the ultrasound plane itself has a certain thickness. Finally, the positions from which the crossing is scanned need to be arranged in a configuration where the constraints on the optimization are sufficiently over-determined, leading to a unique solution [66].

Ball-Bearing Phantom

The ball-bearing phantom [1] is similar to the cross-wire phantoms in that it is used to identify a single point in the image for a series of successively taken scans from varying positions. Instead of scanning the crossing of two wires, a small ball bearing attached to a float is scanned. The same numerical optimization as with the cross-wire phantoms can be used to estimate the transformation and the same drawbacks apply to this method.

Ping-Pong Ball Phantom

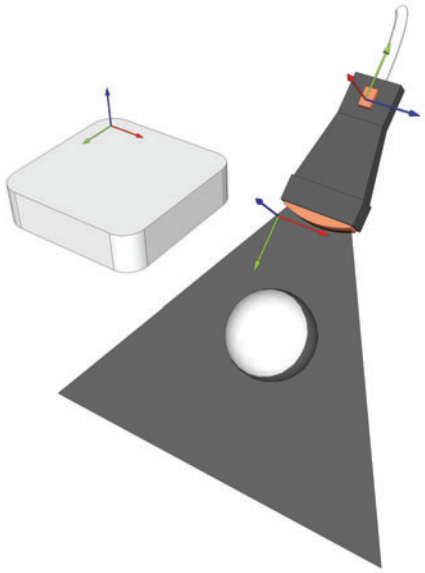


Figure 2.12: Ping-pong ball calibration phantom. The ultrasound probe images the liquid-filled ping-pong ball, producing a circle in the ultrasound image.

An extension to the class of single-point calibration has been proposed by Brendel et al. [9]. They image a ping-pong ball to counter the difficulties in holding the probe to image a very small point. The ball, when filled with liquid, produces a circle in the ultrasound image. By defining a region of interest, this circle can be segmented automatically with the Hough transformation. The main advantage of this method is that it does not require the ball's center to lie within the image plane. Instead, the elevational offset of the ball can be computed from the radius of the circle in the ultrasound image. The calibration is then governed by the following equation:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = {}^{P}H_{T}\,{}^{T}H_{S}\,{}^{S}H_{U} \begin{pmatrix} s_x u \\ s_y v \\ z \\ 1 \end{pmatrix} \tag{2.28}$$

where the ball's elevational offset to the ultrasound plane z is computed from the radius of the detected circle. This method achieved a greater accuracy than the cross-wire or ball-bearing phantoms. However, the acquisition of points in a non-degenerate configuration remains a challenge with this method.

Three-Wire Phantom

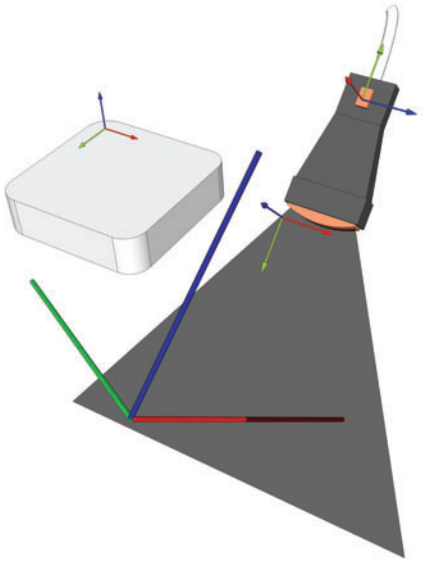


Figure 2.13: Three-wire phantom. Each of three mutually orthogonal wires is scanned along it's length.

Similar to the cross-wire technique, the three-wire calibration proposed by [12] involves segmenting points in scans of a wire. In this phantom, however, there are three wires placed in mutually orthogonal directions.
The origin of the wires is defined as the origin of the phantom's coordinate system, and each axis is assigned to one of the wires. Each wire is then scanned with multiple images, and the segmented points should satisfy:

$$W = {}^{P}H_{T}\,{}^{T}H_{S}\,{}^{S}H_{U} \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix} \tag{2.29}$$

where $W = (x\,0\,0\,1)^T$ for the x-axis, $W = (0\,y\,0\,1)^T$ for the y-axis and $W = (0\,0\,z\,1)^T$ for the z-axis. The advantage over the cross-wire method is that scanning along the length of a wire is easier than finding the exact crossing point of two wires. Carr et al. later also introduced a method for semi-automatic wire segmentation [13], thus improving the speed of this calibration.

2.3.5 Wall Based Calibration Methods

Single-Wall Phantom

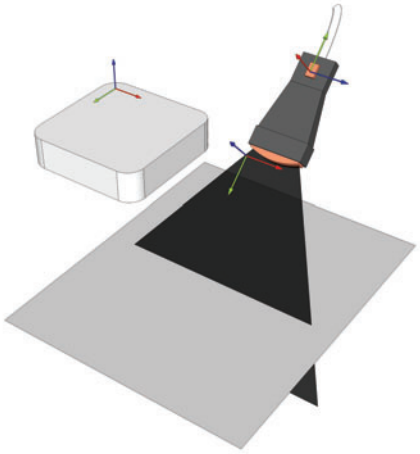


Figure 2.14: Single-wall phantom. The probe images a plane, producing a line in the ultrasound image.

The single-wall phantom was initially proposed by Prager et al. [66]. Scanning a planar phantom produces a straight line in the ultrasound image. This line can be segmented automatically in a robust way, giving two points (defining the line) for every B-scan.
The following equation then governs the calibration process:

$$\begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = {}^{P}H_{T}\,{}^{T}H_{S}\,{}^{S}H_{U} \begin{pmatrix} s_x u \\ s_y v \\ z \\ 1 \end{pmatrix} \tag{2.30}$$

This method requires careful positioning of the probe during the scans to cover all six degrees of freedom during the movement. Hsu et al. proposed a way to assess the scan quality during acquisition using the eigenvalue metric [40].

A particular challenge of this method involves the scanning from oblique angles to the plane, in which case the thickness of the ultrasound plane leads to wide lines in the images, which can lead to inaccuracies in the calibration.
There are several derivations from the original wall phantom which used the floor of a water tank. Rousseau et al. used a plexiglass plate [68]. Langø used a nylon membrane [48], producing better images at oblique angles by reducing specular sound reflections. The Cambridge phantom [66] is also an extension to the basic single-wall concept.

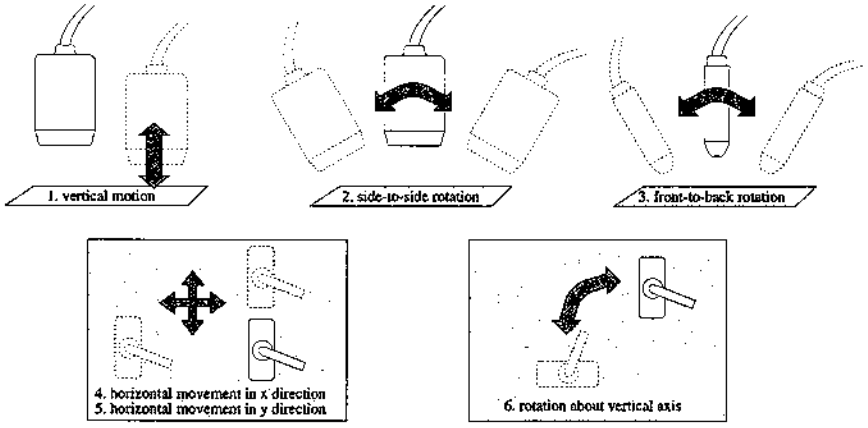


Figure 2.15: Required movement during single-wall calibration. Image courtesy of Prager et al. [66].

Cambridge Phantom

The Cambridge phantom consists of a clamp, which is fixed around the ultrasound probe, and a base consisting of two discs connected by a brass bar [66]. The top of the brass bar is located such that it lies in the center of both the ultrasound plane and the discs. Moving the phantom in the freedom granted by its construction, the top of the brass bar follows a virtual plane located above the real floor (elevated by the radius of the discs). Since the brass bar is both thin and by construction always facing the transducer, a much clearer line is produced in the image. The actual computation of the calibration is exactly the same as with the single-wall calibration.

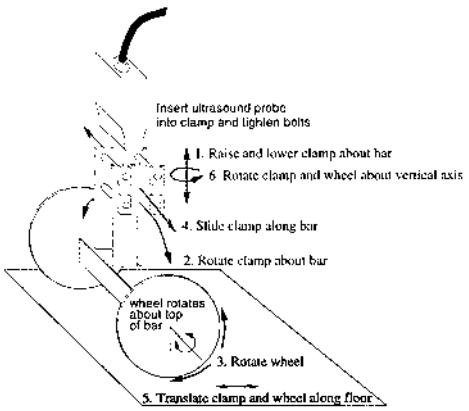


Figure 2.16: Cambridge phantom. Image courtesy of Prager et al. [66].

2.3.6 Shape Alignment Calibration Methods

2D Shape Alignment Phantom

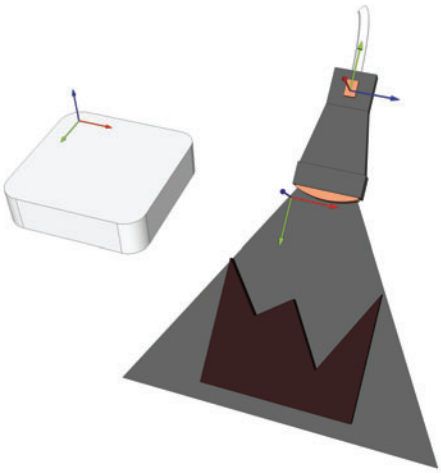


Figure 2.17: 2D shape alignment phantom. A thin board is aligned with the scan plane.

The 2D shape alignment phantom is intended to allow calibration by detecting multiple features in a single B-scan. If at least three features can be mapped, the calibration can be performed with a single image. The feature locations need to be known in world coordinates of the tracking system for this method. They can be localized with a 3D pointer (refer to 2.1.3) for every calibration, after which the phantom must not be moved. An alternative is to attach a tracking sensor to the phantom, in which case it can be moved during calibration.

Sato et al. [71] used a thin board which is aligned with the scan plane. Three corners are measured with a pointer and manually extracted in the ultrasound image. A drawback of this phantom is the difficulty to align it exactly with the ultrasound scan plane.

Precision Mechanical Instrument Phantom

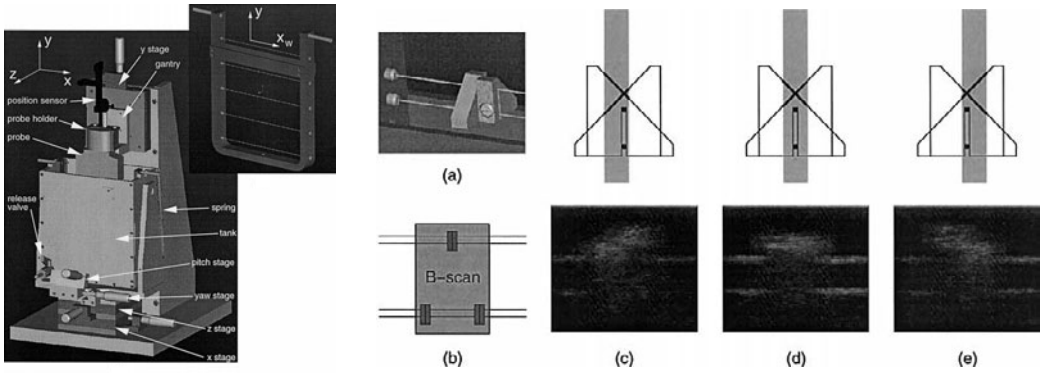


Figure 2.18: Precision mechanical instrument phantom. Images courtesy of Gee et al. [29]

Intending to overcome the difficulty in alignment, Gee et al. [29] designed a mechanical setup where the phantom can be aligned using micrometers.

Z-fiducial Phantom

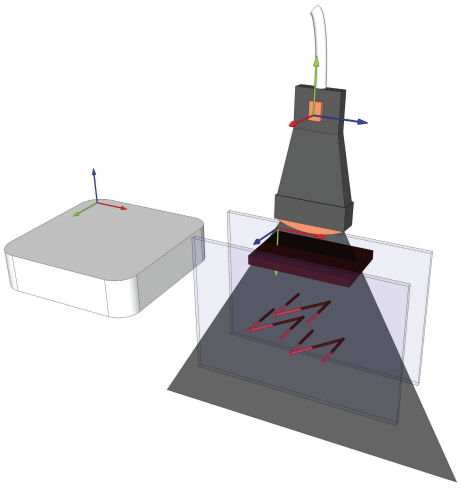


Figure 2.19: Z-fiducial phantom. At least three Z-shaped fiducials between two planes are scanned.

The Z-fiducial phantom allows for single B-scan calibration while avoiding the difficulties in alignment of the 2D shape phantom [19].
Wires or small tubes are run through holes $H_{1..6}$ of two parallel plates to form a Z-shape with the end points A, B, C, D. Scanning this fiducial produces three dots or small circles with the locations M, Z, N in the ultrasound image. The distances $\mid \overline{MZ} \mid$ and $\mid \overline{MN} \mid$ are directly measureable in the image. Since ΔBMZ and ΔCNZ are similar, the location of Z

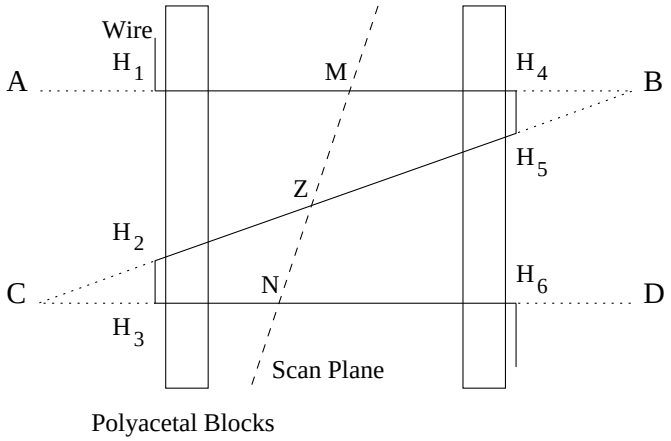


Figure 2.20: Z-fiducial geometry. Image courtesy of Hsu et al. [41]

can be computed as [41]:

$$Z = B + \frac{|\,\overline{BZ}\,|}{|\,\overline{BC}\,|}(C - B) = B + \frac{|\,\overline{MZ}\,|}{|\,\overline{MN}\,|}(C - B) \tag{2.31}$$

Thus, at least three Z-fiducials in a non colinear configuration are necessary to estimate the calibration transformation [2]. The definition of the phantom itself can be done with a 3D-pointer [62, 41] or a position sensor [7, 51].

2.3.7 Phantomless Calibration Methods

3D Pointer

In phantomless calibration, first published by Muratore et al. [59], the probe is locked in place in a water bath. A tracked pointer (refer to 2.1.3) is then moved to different positions on the ultrasound scan plane. The location of the tip in the image is segmented and the corresponding calibration between image and tracking system can then be estimated with different techniques [7].

3D Rod

The difficulty of placing the tip exactly in the scan plane is overcome by the method of Khamene and Sauer [46], who use a rod with a position sensor attached at each end. By taking a series of scans of the rod, the calibration parameters are constrained and estimated with numerical optimization.

2.4 Olympus Bronchoscope

The bronchoscope used throughout this thesis is produced by Olympus, Model UC260F-OL8. It features a fibreoptic camera and illumination system with a field of view of $80°$. This model also features and endobronchial ultrasound (EBUS) transducer, capable of acquiring ultrasound images from within the lung. If the transducer cannot be pressed against the lung surface, an optional inflatable baloon system is integrated, which can bridge the gap between transducer and lung wall in such cases. The transducer has a slightly forward-swept fiel of view of $50°$, and is capable of scan depths up to $9cm$.

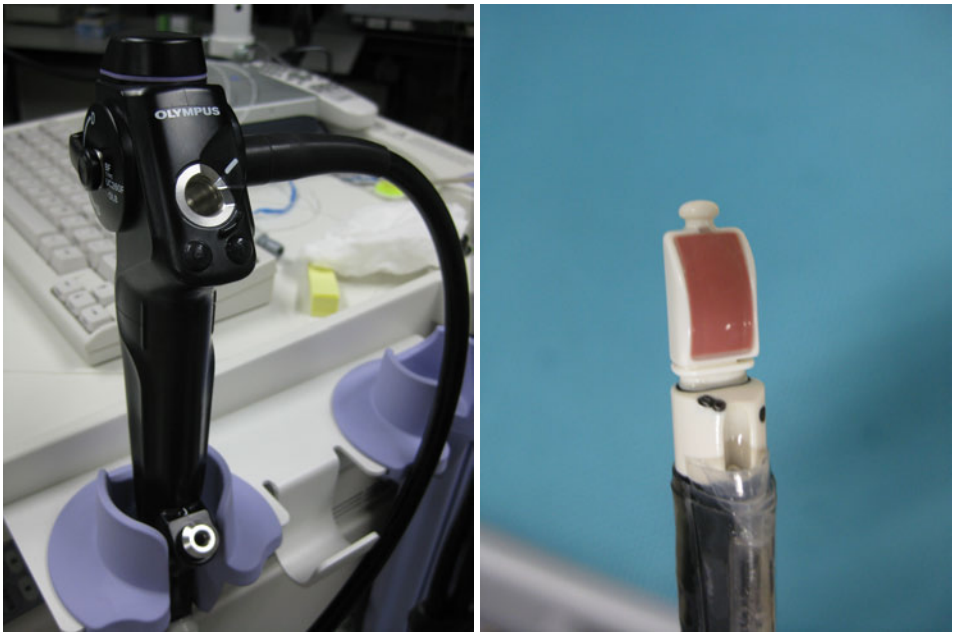


Figure 2.21: Olympus bronchoscope. Left image shows the operator's interface, right side the tip of the flexible tube containing the optics and ultrasound transducer.

Chapter 3

Problem Statement

The aim of this thesis is to find the exact spatial relation between the optical camera C and the ultrasound plane U of an EBUS bronchoscope. This rigid relation ${}^{U}H_{C}$ can be computed in a one-time calibration procedure for subsequent use by navigation systems. The computation of the six degrees of freedom variables (three translation and three rotation) should be done from the imaging of a phantom with the camera and ultrasound. The additional parameters neccessary to model the camera and ultrasound (intrinsic camera parameters, distortion parameters, ultrasound scaling) are computed in a different step with existing methods. The resulting transformation, expressed as a homogeneous transformation matrix should satisfy the following equation:

$$\lambda \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \left[\begin{array}{c|c} K & 0 \end{array} \right] \cdot {}^{C}H_{U} \cdot \begin{pmatrix} s_x \cdot u \\ s_y \cdot v \\ 0 \\ 1 \end{pmatrix} \tag{3.1}$$

With u,v the position of a pixel in the ultrasound image, s_x,s_y the pixel to mm scaling of the ultrasound plane, ${}^{C}H_{U}$ the rigid transformation, K the camera matrix and x,y a pixel in the camera image, up to scale λ.

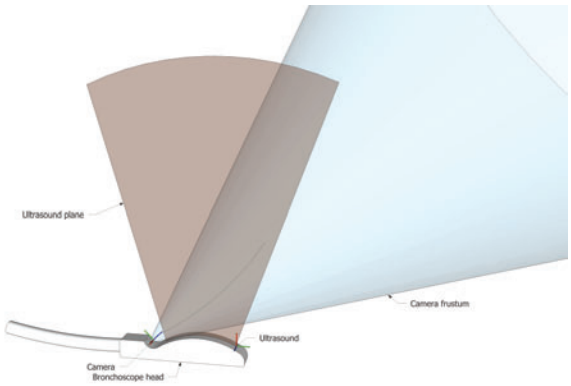


Figure 3.1: EBUS bronchoscope and involved coordinate systems.

Chapter 4

Design of the Phantom

In Chapter 2, ultrasound and camera calibration methods were detailed. The process of designing and building the calibration phantom is based on these methods. The goal of the phantom, as described in the problem statement in Chapter 3, is to allow a direct calibration of the spatial relation between camera and ultrasound coordinate system of an EBUS bronchoscope. While the final design is capable of computing the calibration from a single pair of frames from camera and ultrasound, the earlier models were based on approaches requiring a series of image pairs.

The design of the phantom went through multiple iterations. Since the calibration must be possible without external tracking, the phantom needs to be visible in the camera and ultrasound images at the same time. The fields of view of the two overlap slightly, so the early approaches were based on the idea of imaging the same portion of an object with both modalities.

4.1 Single Wall Based Approach

The first phantom design was inspired by the ease of construction and automatic segmentation capabilities of the wall phantoms. If an optical pattern could be printed on a membrane, continuous pose estimations could be performed with the camera while scanning the plane, with these readings effectively replacing the data stream of a position sensor. This design promised to require little change to the existing method of single wall calibration, and would require only the implementation of an automatic line segmentation [66] and automatic pose estimation for the camera [81], both of which are well described. This approach was ultimately rejected for two reasons. First, the optical pattern for the camera pose estimation needs to be very precise, to allow accurate poses to be computed. The quality offered by office laser printers is usually sufficient, but printing directly on a nylon membrane is not possible with this accuracy. The workaround of printing the pattern on decals and attaching these to the membrane lead to substantial artifacts in the ultrasound image of the line, reducing detection robustness. The second reason this approach was rejected is that it proved very difficult to cover the 6 degrees of freedom of the movement for the ultrasound probe, while still keeping the optical pattern in the view of the camera to provide

position readings. These two characteristics of the system, the camera severely restricting the movement possibilities and the ultrasound calibration requiring good coverage of the possible movement are directly contradictory. While a construction providing the camera with a view on an extended pattern would be possible, such a construction was judged to be ineffective due to complexity.

4.2 Geometry Approximation

Since the construction of a phantom proved to be more difficult than initially anticipated, the geometrical setup of the EBUS system was approximated and subsequent phantom designs were based on this data. In particular, the characteristics of the overlapping area of the camera frustum and ultrasound plane are not very intuitive. The goal of this approximation was to model a cone, representing the frustum of the spherical camera image, and the relative position of the ultrasound plane. A visually accurate representation would then serve as a tool in designing subsequent phantoms for highly accurate calibration.

4.2.1 Ultrasound Plane Geometry

The geometry of the ultrasound plane can be directly measured with the device itself. All ultrasound machines offer the option to measure arbitrary distances in the image, making the measurement of the scan area a simple task. Furthermore, the ratio of the pixels of the ultrasound image to distances in mm can be easily computed. The geometry of the ultrasound

	x	y
Scaling	0.215277	0.236392

Table 4.1: Ultrasound scaling values, mapping pixels to millimeters

can be modeled as the segment of a circle. The greyed area in figure 4.1 is the ultrasound plane. The outer lengths of the image $\overline{x_2\, x_3}$, $\overline{x_1\, x_4}$, $\overline{u\, v}$ have a length of $9\, cm$ if the ultrasound machine is set to maximum imaging depth. The width of the image $\overline{x_4\, x_3}$ is $8.1cm$. The segment angle α is $25°$.

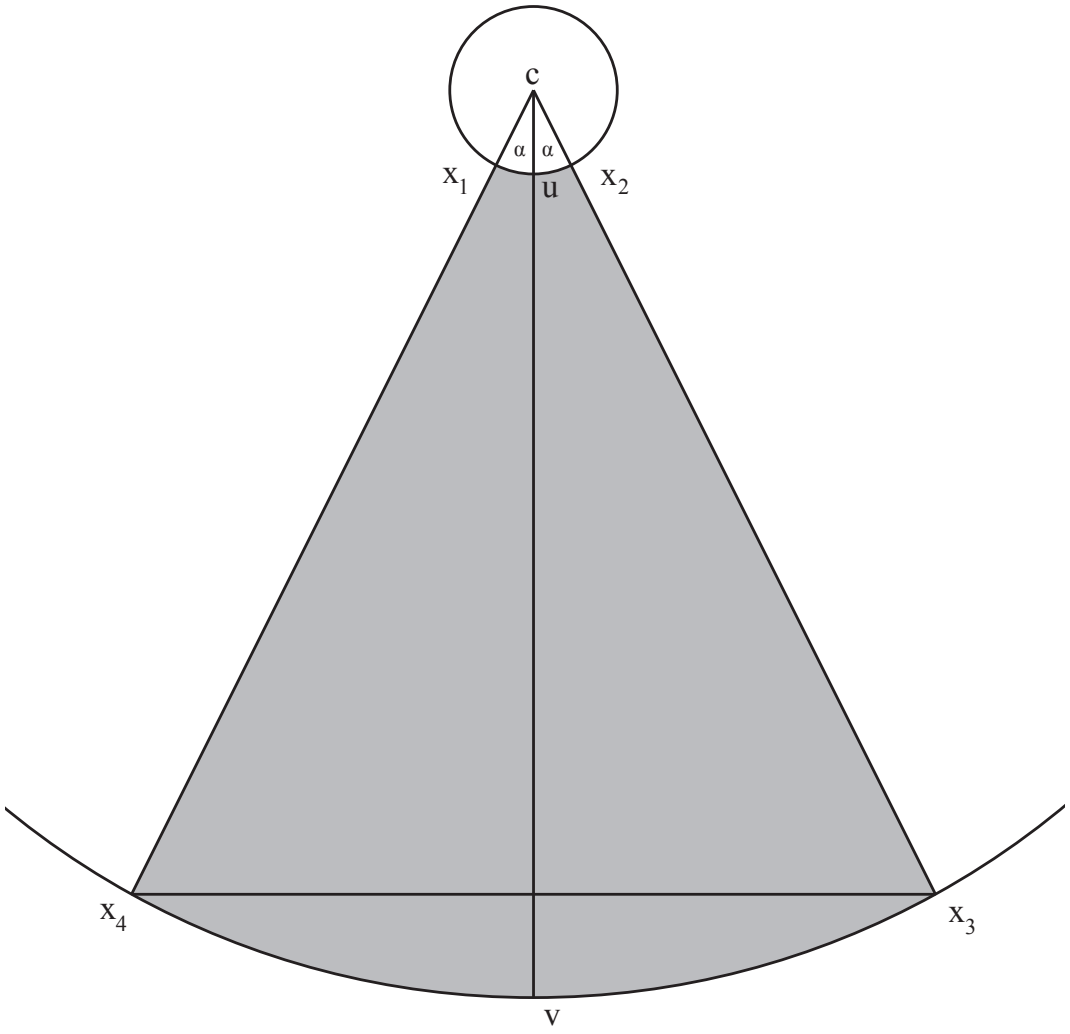


Figure 4.1: Ultrasound plane geometry.

4.2.2 Chopstick Calibration

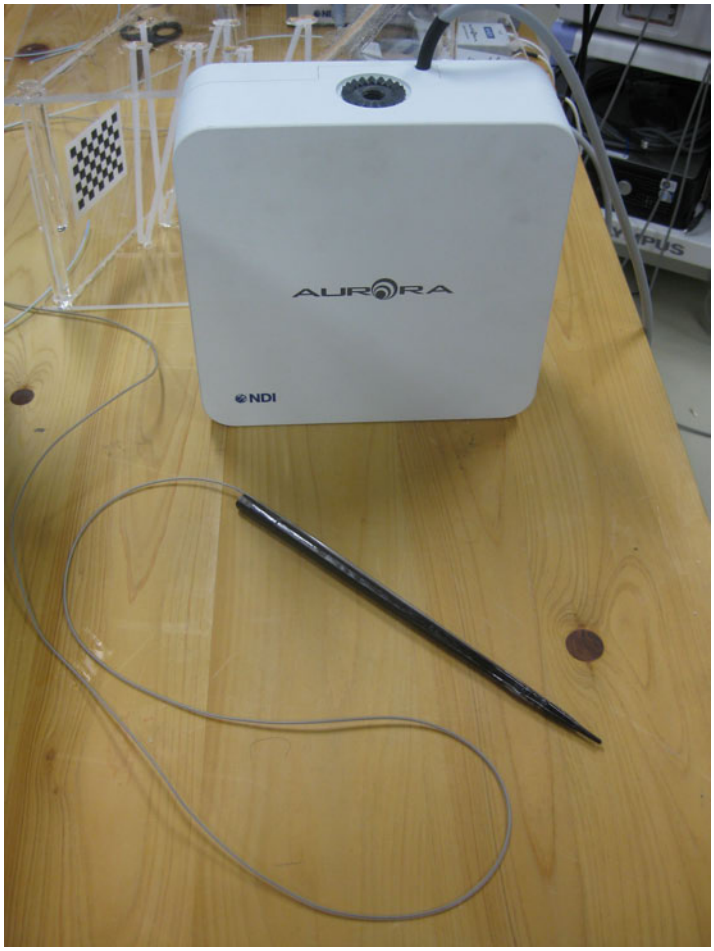


Figure 4.2: Chopstick pointer used for coarse calibration.

The initial approach for a coarse approximation was based on the hotspot calibration of a Japanese chopstick with an attached 6DOF electromagnetic tracking sensor. With the bronchoscope locked in place with a plastic clip, the chopstick's tip was moved along the left and right outer edge of the ultrasound plane, taking a series of position readings. Since the geometry of the ultrasound plane could easily be determined by using the integrated measurement tool of the device, the resulting plane could be fitted to the position readings in MATLAB.
To locate the camera frustum, the tip of the chopstick pointer was placed on the outer edge of the image at a distance close to the lens. On a subsequent move around the edge of the image circle, multiple position readings were saved. This procedure was repeated with increasing distance, resulting in a point cloud to which a cone could be fitted.

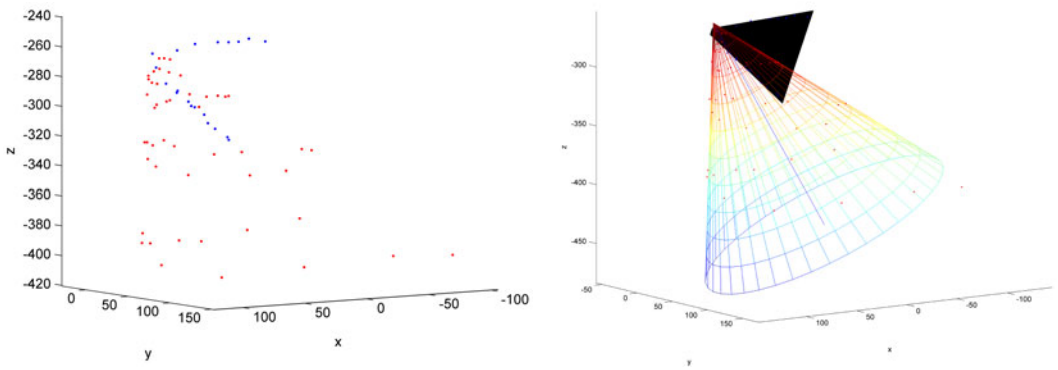


Figure 4.3: Chopstick calibration. Recorded point cloud (left) and fitted cone and plane (right).

4.2.3 Camera Calibration and Single Wall

Since the simple chopstick calibration proved to be not accurate enough, the combination of the highly accurate hand-eye calibration for the camera and a single wall calibration for the ultrasound were performed. While this calibration is subject to the error introduced by the EM tracking system involved, it was expected to provide a good approximation of the final calibration quality. To model the camera frustum cone, the intrinsic parameters of the camera needed to be determined as well. The calibration toolbox implementation of [8, 80] were used. The images of the calibration grid were subjected to a histogram equalization and median filtering (see figure 4.4) before being processed with the calibration toolboxes.

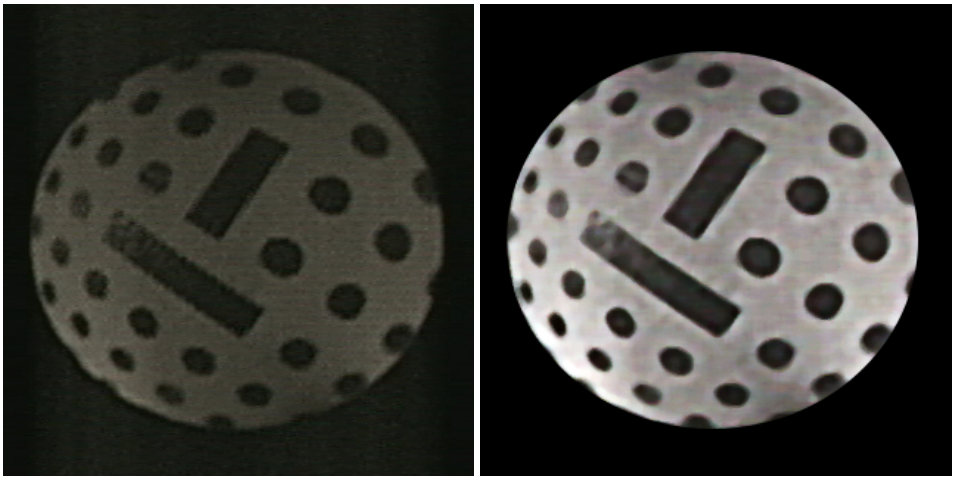


Figure 4.4: Pattern used for camera calibration. The left image is raw, the right image is preprocessed with histogram equalization and median filtering.

	x (pixels)	y (pixels)
Focal length	218.08129 ± 4.53966	198.94756 ± 4.35092
Principal point	161.07441 ± 2.53894	178.09913 ± 2.46174

Coefficient	-0.43073	0.15188	0.00372	0.00137
Uncertainty	0.02240	0.02217	0.00270	0.00251

Table 4.2: Intrinsic camera calibration results and distortion coefficients

The hand-eye calibration resulted in a backprojection error of 7.4746 pixels, which is an acceptable value for a miniature endoscopic camera. Combined with the transformation obtained by the single-wall calibration, which was just verified by backprojecting a line at this point, the relationship shown in figure 4.5 was computed.

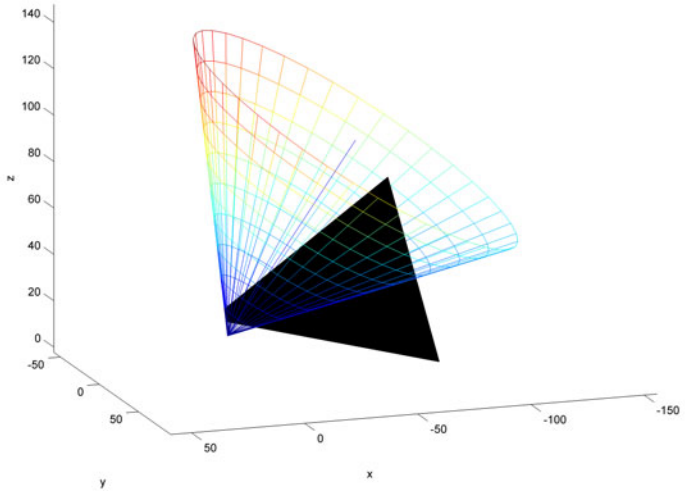


Figure 4.5: Resulting geometry for camera cone and ultrasound plane obtained from combined hand-eye and single-wall calibration.

4.3 Z-fiducial Based Approach

It became apparent with the approximate calibration described above that the overlapping area of the ultrasound plane and camera frustum is small. Imaging the same object with both modalities is thus not only technically difficult, but also wasting a significant amount of image space. If two different objects are imaged by the camera and ultrasound, much more of the image area can be used for the calibration. The approach of essentially imaging two different phantoms has several design implications. First, the two phantoms need to be rigidly attached to a common base, their spatial relationship needs to be known in advance (necessitating a calibration on its own) and due to the way camera and ultrasound are aligned, imaging will only be possible from a constrained set of positions. The last characteristic virtually rules out calibration approaches which rely on imaging from several different positions. The camera pose estimation can easily be performed from a single frame with accurate results. For the ultrasound image, the natural choice in this combination is a z-fiducial

based phantom. It allows for a pose estimation from a single image, while permitting some freedom in the probe's positioning. With the right geometric setup it is possible to construct a phantom which allows simultaneous imaging of an optical pattern and a set of z-fiducials, while permitting the probe and camera to be moved within a few centimeters. A single pair of images is then sufficient to compute a calibration, with the option to take a short series of image pairs from slightly different positions.

4.3.1 Initial Phantom Design

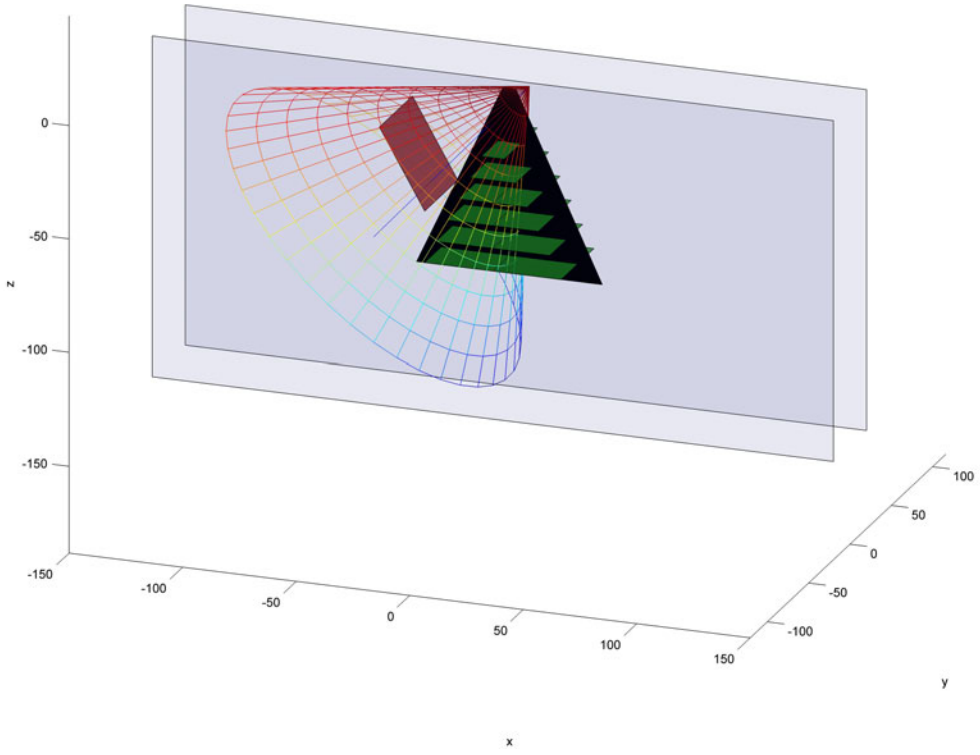


Figure 4.6: Initial phantom design derived from approximate calibration. The red plane is the base for the optical pattern, the green planes represent z-fiducials.

The phantom base is formed by two plates, placed 40 mm apart. Six layers of Z-fiducials and the optical pattern are placed in between the plates. Since it would obstruct the view of the camera, no rubber membrane is placed atop the fiducials to aid the automatic segmentation. Hollow tubes filled with water are used instead of wires. The resulting circles in the ultrasound image are then automatically segmented without a reference. Using a 2D local coordinate system on the plates (refer to figure 4.7), the positions of the holes through which the Z-fiducials are run are shown in table 4.3.

The first two levels have just one fiducial, while the subsequent levels have two each. The optical pattern is placed at the coordinates $[-40, -30]$, and rotated $60°$ clockwise from the x-axis. The pattern proposed by Wengert [81] is used, the grid dots on the pattern have a spacing of $5\,mm$.

Level	Hole 1	Hole 2	Hole 3
1	$[-5, 40]$	$[5, 40]$	-
2	$[-10, 30]$	$[10, 30]$	-
3	$[-15, 20]$	$[0, 20]$	$[15, 20]$
4	$[-20, 10]$	$[0, 10]$	$[20, 10]$
5	$[-25, 0]$	$[0, 0]$	$[25, 10]$
6	$[-30, -10]$	$[0, -10]$	$[30, -10]$

Table 4.3: Phantom fiducial hole coordinates in mm on the base planes.

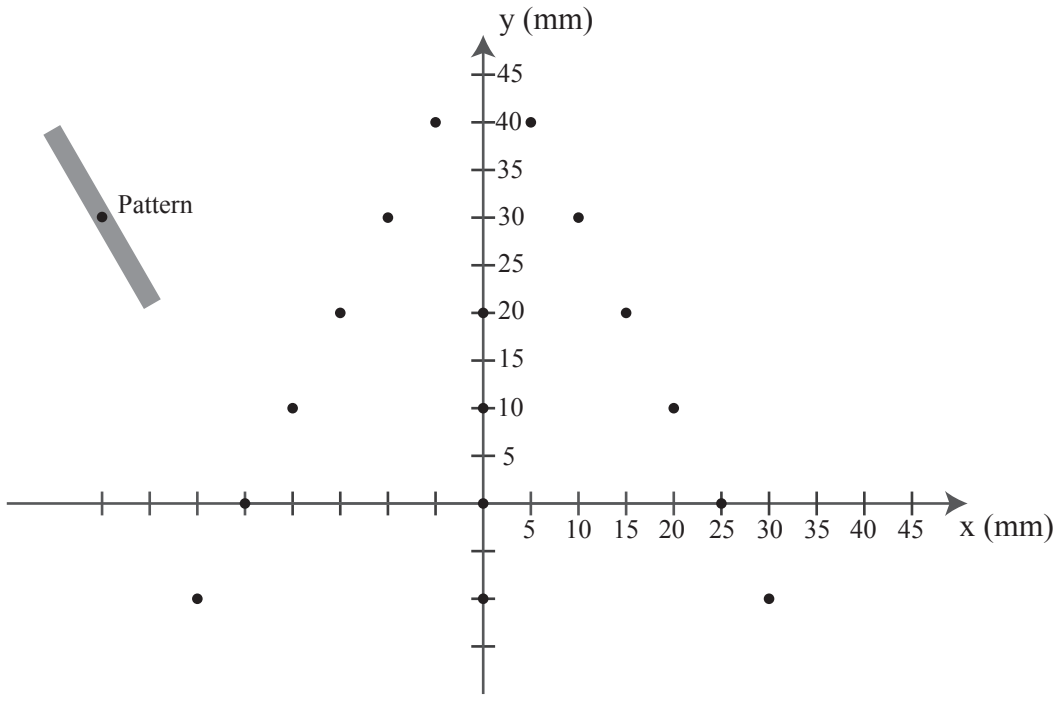


Figure 4.7: Phantom base plane. The dots mark the drill hole positions for the Z-fiducials.

Chapter 5

Implementation of the Phantom

The phantom is constructed from two acrylic glass plates, measuring $300\,mm$ by $150\,mm$ with a thickness of $5\,mm$. Four hollow acrylic cylinders with a height of $40\,mm$ are used as spacers, held in place with metal screws. Another $50\,mm \cdot 40\,mm \cdot 5\,mm$ acrylic block is used to hold the optical pattern. It is glued to just one of the plates, allowing for non-destructive disassembly of the phantom. The optical pattern is printed with an office laser printer to a waterproof and adhesive decal foil. The Z-fiducials are built from a hollow rubber tube with a diameter of $3\,mm$ and a wall thickness of roughly $1\,mm$. The holes for the fiducials are drilled using a standard handheld drilling machine. Holes which only hold one tube have a diameter of $3\,mm$ while those which hold two tubes are $5\,mm$ in size. The tube is run through all the holes in one piece, filled with water and then tightened. Cutoffs from a plastic drinking straw are used to hold the tube in place.

5.1 Occlusion Correction

During the early experiments, it quicky became apparent that the ultrasound does not propagate through the rubber tube as well as anticipated. This caused the upper layers to occlude the lower layers, effectively leaving only the top two layers of fiducials visible in the ultrasound image. The initial design required the sound to pass through all previous layers to reach any given layer, thus making rather inefficient use of the field of view of the probe. A second design iteration which uses only four layers of fiducials, and only one fiducial per layer was thus conceived. The new design allows for the ultrasound waves to reach any fiducial without having to pass another fiducial from a higher layer. The total extents of the fiducials were also reduced to allow more flexibility for the probe movement. The new fiducial positions are shown in table 5.1.

5.2 Measurement of Resulting Geometry

After the construction of the phantom, the precise location of the drill holes (and thus fiducials) and the optical pattern needs to be measured. The spatial relation between optical pattern and fiducials is required for the calibration. Knowledge about the exact z-fiducial

Level	Hole 1	Hole 2
1	$[-15, 30]$	$[15, 35]$
2	$[-15, 25]$	$[15, 25]$
3	$[-15, 15]$	$[15, 15]$
4	$[-5, 5]$	$[25, 5]$

Table 5.1: Phantom fiducial hole coordinates in mm on the base planes.

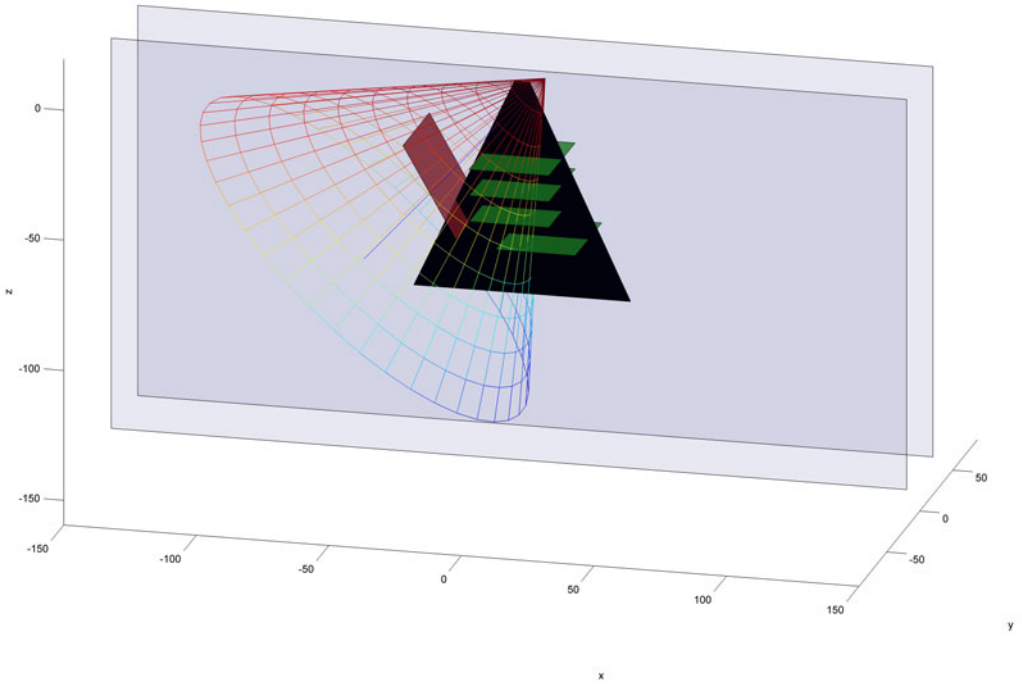


Figure 5.1: Corrected phantom design, avoiding Z-fiducials occluding each other. The red plane is the base for the optical pattern, the green planes represent Z-fiducials.

layout will increase the overall accuracy of the calibration. The optical pattern coordinate system is used as the base coordinate system of the phantom, and the fiducial locations are computed in this coordinate system.
An NDI Polaris passive optical infrared tracking system was used to measure the phantom with high accuracy. A custom-built pointer was hotspot calibrated and subsequently used to measure the phantom. The hotspot calibration was performed with an uncertainty of 0.71 mm, allowing for precise localisation of points in space.

Each drill hole was measured ten times, the values being averaged. Each point on the outer edges of the optical pattern was measured five times, the values also being averaged. Table 5.2 shows the measured point locations in ideal local pattern coordinates. These were subsequently used for a point based registration with the averaged point positions measured in the tracking coordinate system, which are shown in table 5.3. Table 5.4 shows the standard deviations for the measured point positions.
The values measured for the individual Z-fiducial layers are shown in table 5.5, already

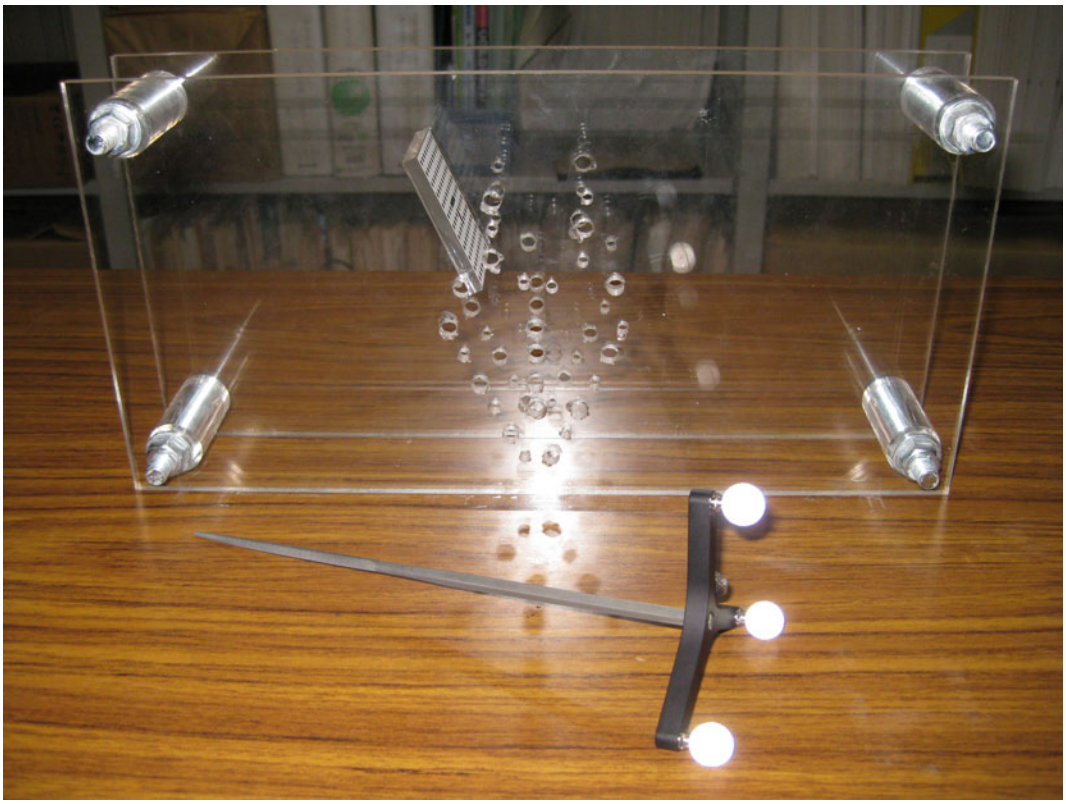


Figure 5.2: Calibration pointer in front of the partly assembled phantom.

converted into pattern coordinates with the transformation obtained through the registration of the pattern points. The standard deviations for the averaged points is shown in table 5.6. These Z-fiducial locations are later used by the calibration software to compute the pose of the ultrasound directly in the coordinate system of the optical pattern.

Point 1	Point 2	Point 3	Point 4
$\begin{bmatrix} 15 \\ -5 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -15 \\ -5 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -15 \\ 35 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 15 \\ 35 \\ 0 \end{bmatrix}$

Table 5.2: Ideal pattern point positions in mm, in pattern coordinates.

Point 1	Point 2	Point 3	Point 4
$\begin{bmatrix} -128.7 \\ 7.2 \\ -1455.8 \end{bmatrix}$	$\begin{bmatrix} -128.7 \\ 37.5 \\ -1455.0 \end{bmatrix}$	$\begin{bmatrix} -89.9 \\ 37.0 \\ -1447.2 \end{bmatrix}$	$\begin{bmatrix} -89.8 \\ 7.0 \\ -1447.9 \end{bmatrix}$

Table 5.3: Measured pattern point positions in mm, in the tracking system's coordinates.

Point 1	Point 2	Point 3	Point 4
$\begin{bmatrix} 0.2968 \\ 0.2861 \\ 0.2950 \end{bmatrix}$	$\begin{bmatrix} 0.1455 \\ 0.2100 \\ 0.1892 \end{bmatrix}$	$\begin{bmatrix} 0.2846 \\ 0.3387 \\ 0.3263 \end{bmatrix}$	$\begin{bmatrix} 0.1692 \\ 0.4162 \\ 0.3201 \end{bmatrix}$

Table 5.4: Standard deviations for pattern point positions in mm, in the tracking system's coordinates.

Level	Hole A	Hole B	Hole C	Hole D
1	$\begin{bmatrix} 23.4655 \\ 13.7154 \\ 14.7131 \end{bmatrix}$	$\begin{bmatrix} -20.0027 \\ 13.6709 \\ 14.7082 \end{bmatrix}$	$\begin{bmatrix} 24.712 \\ 25.4921 \\ 44.0669 \end{bmatrix}$	$\begin{bmatrix} -20.5523 \\ 25.3027 \\ 44.2794 \end{bmatrix}$
2	$\begin{bmatrix} 24.6690 \\ 22.8151 \\ 11.6073 \end{bmatrix}$	$\begin{bmatrix} -20.5233 \\ 21.3386 \\ 9.3278 \end{bmatrix}$	$\begin{bmatrix} 24.5425 \\ 35.0974 \\ 40.5261 \end{bmatrix}$	$\begin{bmatrix} -20.1233 \\ 34.2455 \\ 38.4440 \end{bmatrix}$
3	$\begin{bmatrix} 24.2148 \\ 31.0329 \\ 5.8097 \end{bmatrix}$	$\begin{bmatrix} -20.4786 \\ 31.3890 \\ 5.6764 \end{bmatrix}$	$\begin{bmatrix} 24.6748 \\ 44.8387 \\ 34.7115 \end{bmatrix}$	$\begin{bmatrix} -20.5846 \\ 44.5630 \\ 34.2434 \end{bmatrix}$
4	$\begin{bmatrix} 24.5040 \\ 44.2274 \\ 13.7867 \end{bmatrix}$	$\begin{bmatrix} -20.3325 \\ 44.2135 \\ 12.1654 \end{bmatrix}$	$\begin{bmatrix} 24.3622 \\ 57.8136 \\ 41.7263 \end{bmatrix}$	$\begin{bmatrix} -20.2944 \\ 57.3632 \\ 40.1816 \end{bmatrix}$

Table 5.5: Measured drilling hole positions in mm, in the coordinate system of the optical pattern.

Level	Hole A	Hole B	Hole C	Hole D
1	$\begin{bmatrix} 0.3871 \\ 0.6820 \\ 0.3078 \end{bmatrix}$	$\begin{bmatrix} 0.4834 \\ 0.3739 \\ 0.1420 \end{bmatrix}$	$\begin{bmatrix} 0.3836 \\ 0.2967 \\ 0.3068 \end{bmatrix}$	$\begin{bmatrix} 0.3902 \\ 0.3648 \\ 0.3489 \end{bmatrix}$
2	$\begin{bmatrix} 0.5771 \\ 0.2856 \\ 0.3975 \end{bmatrix}$	$\begin{bmatrix} 0.3628 \\ 0.2752 \\ 0.4020 \end{bmatrix}$	$\begin{bmatrix} 0.3447 \\ 0.2846 \\ 0.4376 \end{bmatrix}$	$\begin{bmatrix} 0.4497 \\ 0.2272 \\ 0.2896 \end{bmatrix}$
3	$\begin{bmatrix} 0.4786 \\ 0.7005 \\ 0.5445 \end{bmatrix}$	$\begin{bmatrix} 0.4832 \\ 0.3881 \\ 0.3315 \end{bmatrix}$	$\begin{bmatrix} 0.2531 \\ 0.2723 \\ 0.4349 \end{bmatrix}$	$\begin{bmatrix} 0.5718 \\ 0.4277 \\ 0.3410 \end{bmatrix}$
4	$\begin{bmatrix} 0.3982 \\ 0.2294 \\ 0.2902 \end{bmatrix}$	$\begin{bmatrix} 0.4205 \\ 0.2813 \\ 0.4434 \end{bmatrix}$	$\begin{bmatrix} 0.5188 \\ 0.3701 \\ 0.5579 \end{bmatrix}$	$\begin{bmatrix} 0.5567 \\ 0.3287 \\ 0.4911 \end{bmatrix}$

Table 5.6: Measured drilling hole position uncertainties in mm.

Chapter 6

Calibration Method

The origin of the optical pattern is defined as the origin of the entire phantom. The positions of the Z-fiducials are also transformed into these coordinates after their measurement. Pose estimations performed on the optical pattern and the Z-fiducials thus have a common coordinate system. Defining the transformation for the camera as ${}^{P}H_{C}$ and the transformation for the ultrasound plane as ${}^{P}H_{U}$, equation 3.1 can be reformulated as:

$$\lambda \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \left[\begin{array}{c|c} K & 0 \end{array} \right] \cdot {}^{P}H_{C}^{-1} \cdot {}^{P}H_{U} \cdot \begin{pmatrix} s_x \cdot u \\ s_y \cdot v \\ 0 \\ 1 \end{pmatrix} \tag{6.1}$$

The projection matrix K is built from the parameters calibrated in air, and the distortion is also performed with the coefficients for air. After the projection, points $p_u = (x, y)^T$ need to be first distorted (see subsection 2.2.1) and the distorted point p_d then scaled to compensate for the magnification effect of the underwater camera in accordance with equation 2.21 for the final point coordinates p_f:

$$p_d = \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right) p_u + \begin{bmatrix} 2\,p_1\,x\,y + p_2(r^2 + 2x^2) \\ p_1\,(r^2 + 2y) + 2\,p_2\,x\,y \end{bmatrix} \tag{6.2}$$

$$p_f = \alpha + (1.333 \cdot (p_d - \alpha)) \tag{6.3}$$

Where α is the principal point of the camera. This scaling is sufficient to fully model the submersion of the camera in water.

6.1 Averaged Transformations

A transformation can be computed from a single pair of frames. To increase the stability of the calibration, several frames may be taken in succession and the resulting transformation matrices averaged. While not significantly changing the required time to capture the calibration input data, this can reduce the influence of small errors in the pose estimations, and result in a more stable calibration result.

6.2 Software Implementation

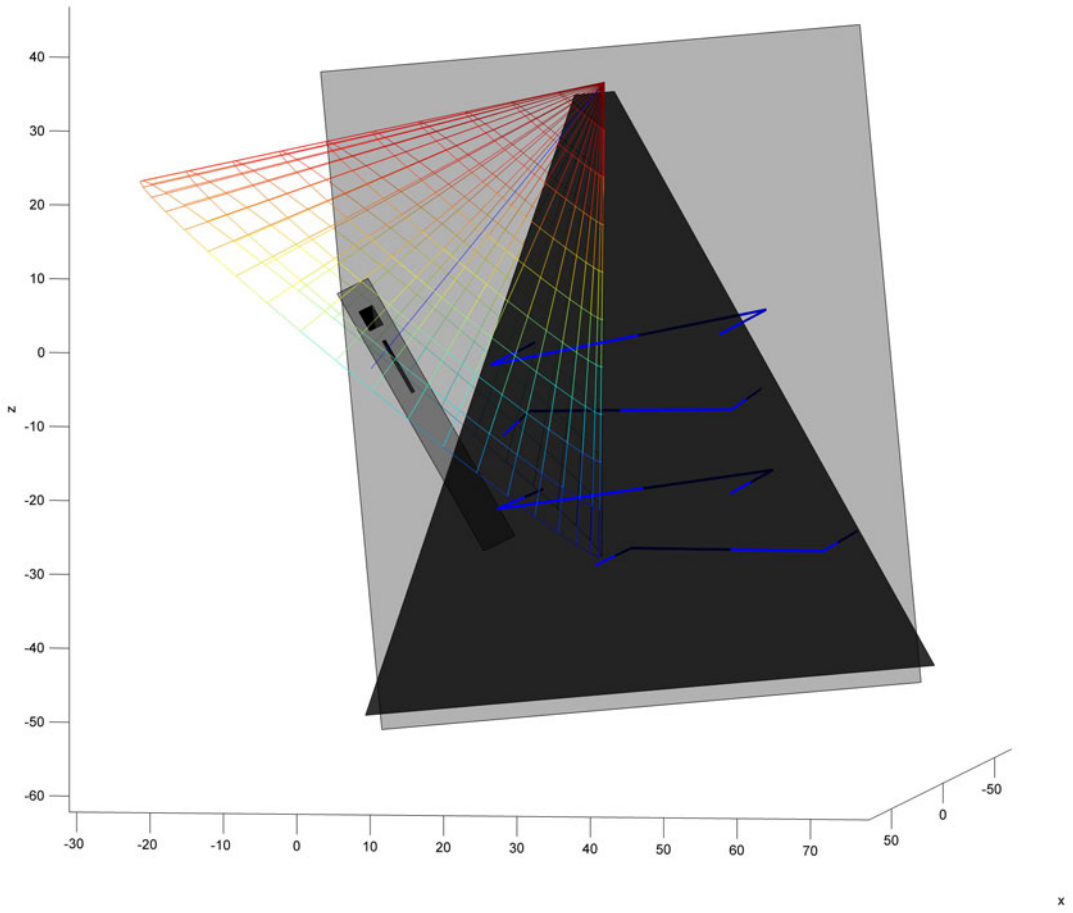


Figure 6.1: Estimated poses from optical pattern and Z-fiducials, representing the final calibration.

6.2.1 Camera Pose Estimation

The feature detection for the camera image is based on the algorithm proposed by Wengert [81]. The pattern is used unmodified (except for scaling) and the feature matching algorithm is analogous to the originally proposed version. The preprocessing and blob extraction before the feature matching was adapted to the underwater, real-time environment of the calibration procedure. The color video image from the camera is first converted to grayscale. It is then subjected to a histogram equalization, followed by an (inverse) adaptive thresholding to produce a binary image containing the dots and bars of the pattern. The adaptive threshold allows for a robust extraction of the features even under the usually very inhomogeneous illumination caused by the bronchoscope's light source. Blobs are then extracted from the binary image. Before being processed, the detected points are scaled as if

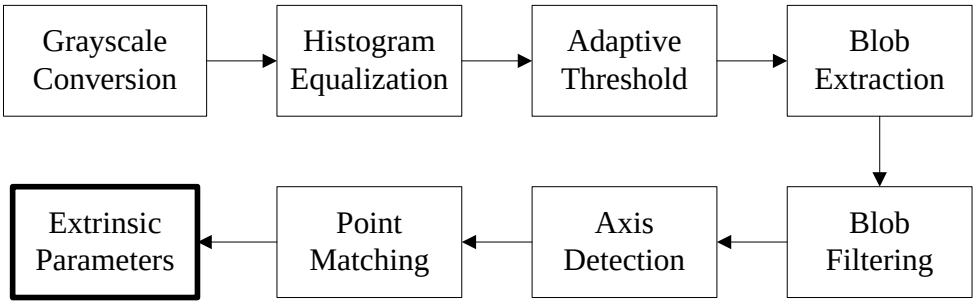


Figure 6.2: Camera feature extraction algorithm outline.

they were in an air environment, so that the intrinsic parameters calibrated in air can be used for the pose estimation (see section 2.2.2 for details). The two large bars defining the local coordinate system of the pattern can be identified through their large ellipticity, defined as the ratio of their major and minor axis length. For the bars, this ratio is greater than two, while the circles have a value of roughly one. The two bars are then distinguished by their area. Once the bars are detected, an iterative search is performed, matching the circles of the pattern to its local coordinate system. Once the point correspondences are established, a projective transformation between all points is computed using the OpenCV library [43].

6.2.2 Ultrasound Pose Estimation

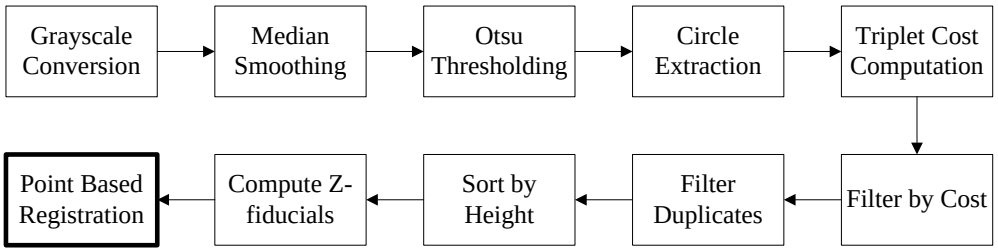


Figure 6.3: Ultrasound feature detection algorithm outline.

The ultrasound image is first converted to grayscale, smoothed with a median filter and subjected to a binary thresholding using the method of Otsu [61]. From the resulting binary image, all contours are extracted using the OpenCV library's `cvFindContours` method [43]. The extracted contours are then filtered based on an area threshold to eliminate small artifacts. From the resulting list of candidates, every possible triplet is assembled and associated with a cost. The cost function is defined as the sum of squared distances of the points on the image y-axis. For any given triplet of y-coordinates c_1, c_2, c_3 the cost is thus:

$$C(c_1, c_2, c_3) = (c_1 - c_2)^2 + (c_1 - c_3)^2 + (c_2 - c_3)^2 \tag{6.4}$$

Points are then filtered with a cost threshold. No detected contour can be part of two fiducials, but all points are part of multiple triplets. To correct this, the triplets are iterated

through, starting from the lowest cost triplet. Any subsequent triplet containing one of the points of the current triplet is discarded, ensuring all triplets contain unique points. Finally, the remaining triplets are sorted by their mean y-position, and the points of the triplets are sorted by their x-position. The n triplets with the lowest cost now contain, in order, the points of the n Z-fiducials. Using the method described in section 2.3.6, n points in 3D space can be computed from the fiducials. These points are registered to the phantom geometry with a point based registration based on the method of Umeyama [79], yielding the desired pose of the ultrasound probe.

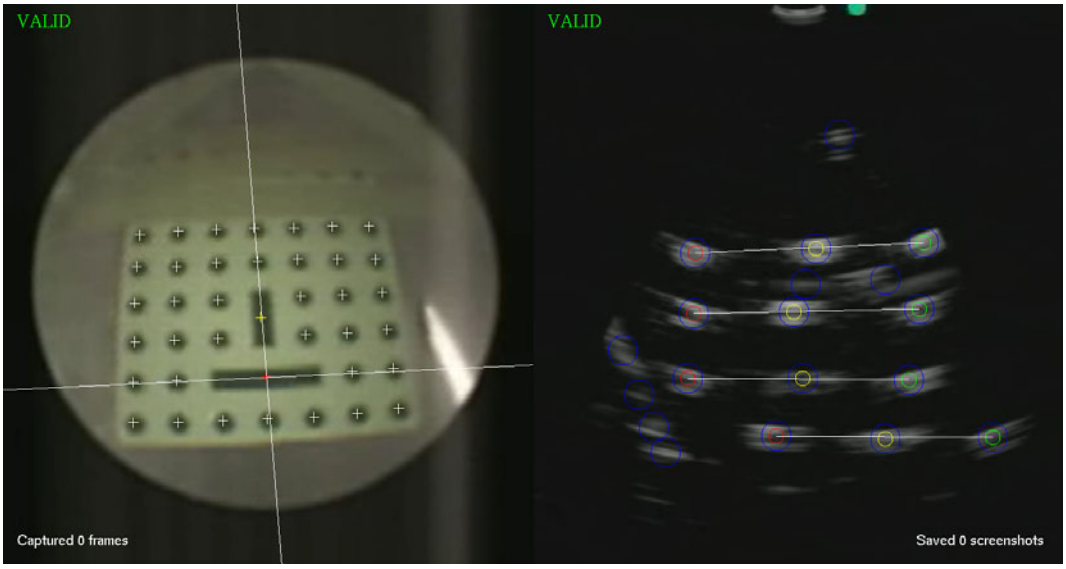


Figure 6.4: Realtime visualization of the detected pattern features in the camera image (left) and in the ultrasound image (right).

Chapter 7

Experiments and Results

This chapter describes the experimental validation of the proposed calibration method. It is compared to a combination of hand-eye and single-wall calibrations, which were performed using an NDI Aurora tracking system. To evaluate the accuracies, a series of image locations of a probe tool was backprojected from ultrasound images into camera images. These positions are then compared to the actual position of the probe in the camera image.

7.1 Intrinsic Camera Parameters

The intrinsic camera parameters are necessary both for the pose estimation during the phantom based calibration and for the hand-eye calibration, and were thus computed first. Five series of 15 images were taken in air to compute the intrinsic and distortion parameters of the camera. The automatic camera calibration toolbox [80] pattern (see figure 2.4) was used for the calibration. The estimated parameters were subsequently used for both the hand-eye based and the phantom based calibrations. The results of the calibrations are presented in tables 7.1 and 7.2. The numerical uncertainties are approximately three times the standard deviations. While they initially seem to be high, they fall within the expected range for a camera system with low resultion, high distortion and inhomogeneous illumination of the pattern. Series one and five may be considered outliers, since they have significantly larger error estimations than series two, three and four.

Series	Focal Length	Principal Point
1	$[217.305\ 198.62] \pm [10.336\ 8.424]$	$[161.876\ 185.055] \pm [6.752\ 11.217]$
2	$[217.614\ 197.864] \pm [1.871\ 1.403]$	$[158.549\ 186.36] \pm [1.774\ 2.479]$
3	$[217.630\ 199.144] \pm [1.945\ 1.743]$	$[159.748\ 182.782] \pm [1.938\ 2.3]$
4	$[216.977\ 198.207] \pm [1.296\ 1.164]$	$[160.0\ 182.256] \pm [1.457\ 1.819]$
5	$[214.467\ 196.76] \pm [5.322\ 4.653]$	$[155.902\ 173.708] \pm [5.483\ 6.564]$

Table 7.1: Intrinsic camera calibration results for a series of five calibrations

Series 1	k_1	k_2	p_1	p_2
Coefficient	-0.44568	0.15471	-0.00311	0.00157
Uncertainty	0.04022	0.03507	0.01669	0.00536
Series 2	k_1	k_2	p_1	p_2
Coefficient	-0.42755	0.14222	-0.00546	0.00364
Uncertainty	0.01021	0.00929	0.00335	0.00118
Series 3	k_1	k_2	p_1	p_2
Coefficient	-0.43375	0.15173	-0.00092	0.00371
Uncertainty	0.01033	0.01027	0.00293	0.00146
Series 4	k_1	k_2	p_1	p_2
Coefficient	-0.43002	0.14834	-0.00074	0.00306
Uncertainty	0.00639	0.00686	0.00166	0.00102
Series 5	k_1	k_2	p_1	p_2
Coefficient	-0.43073	0.15188	0.00372	0.00137
Uncertainty	0.03414	0.04829	0.00726	0.00438

Table 7.2: Undistortion coefficients for five measurement series

7.2 Verification Method

To measure the quality of the calibrations, the tip of a probe with a 65 mm rigid metal tip was placed in the ultrasound image such that is is also visible in the camera image. After a manual segementation of the position in the ultrasound image, the location of the tip is backprojected into the camera image, distorted and scaled to a water based camera system using equations 6.1, 6.2 and 6.3, respectively. Through another manual segmentation of the actual position of the probe in the camera image, an error estimation in pixels for the backprojection can be computed. A series of 20 probe tip location pairs in ultrasound and camera images was taken in advance to the calibrations to provide a common dataset for verification. Figure 7.1 shows the point locations for camera and ultrasound in image coordinates. Since the points can only lie within the overlapping area of camera and ultrasound image, they are intentionally not distributed throughout the entire image spaces.

As a quality indicator the cumulative pixel error for all $n = 20$ point pairs p_i, q_i is computed as the sum of their euclidean distances:

$$cpe_e = \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{2} \left(p_j^i - q_j^i\right)^2} \tag{7.1}$$

7.2.1 Ultrasound Beam Thickness

The ultrasound plane is not a true plane in the mathematical definition since it has a thickness. Due to the miniature sizes, this thickness covers a substantial area in the camera image, where a mathematical plane would only show as a line. It is very difficult to place the tip of the probe exactly in the center of the ultrasound plane. The recorded probe positions in the camera image can thus deviate orthogonally from the location of the backprojected

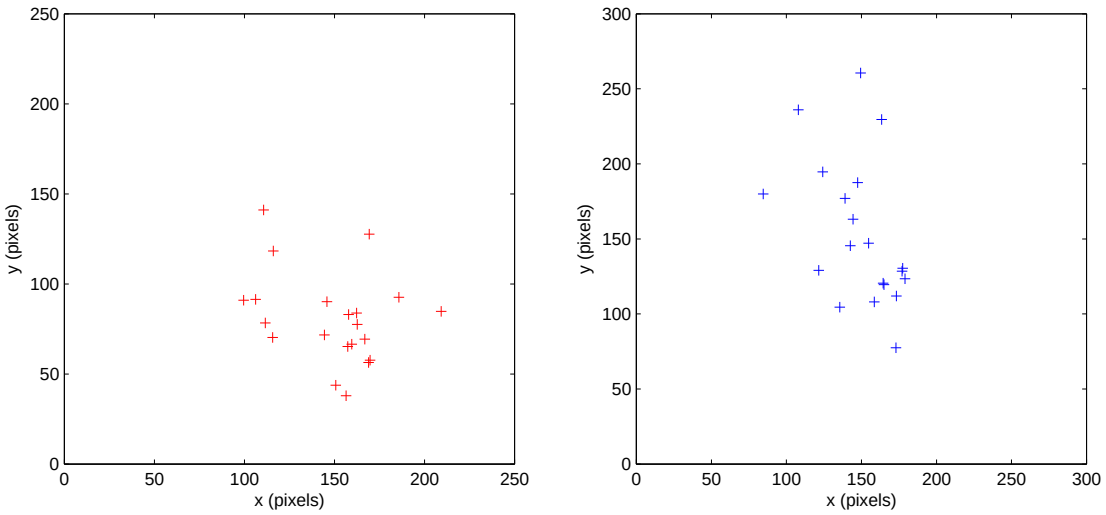


Figure 7.1: Probe tip positions used for verification in camera image (left) and ultrasound image (right)

ultrasound plane. This leads to high cpe_e error values, although the backprojection might actually be of good quality. For this reason, in addition to the euclidean distances between the points cpe_e their component parallel to the ultrasound plane is considered during the experiments, approximated by their image height y:

$$cpe_h = \sum_{i=1}^{n} \left| \; p_y^i - q_y^i \; \right| \tag{7.2}$$

This measure can give a better estimate of the actual backprojection quality since it eliminates errors caused by the validation method itself, albeit at the price of only considering the image height as opposed to a full two-dimensional parameter space.

7.3 Phantom Calibration

To evaluate the quality of the phantom based automatic calibration, a series of 25 camera and ultrasound image pairs was taken. From the ultrasound images, 25 Z-fiducial based pose estimations ${}^{P}H_{U}^{i}$ were computed. From the camera images, pose estimations were performed for each image with all five intrinsic camera parameter sets computed beforehand, resulting in $25 \cdot 5$ camera poses ${}^{P}H_{U}^{j}$.
These two sets of transformations were combined into 125 calibration matrices ${}^{C}H_{U} = {}^{P}H_{C}^{-1} \cdot {}^{P}H_{U}$, matching the 125 camera poses to their respective ultrasound pose counterpart.
Finally, groups of five transformations ${}^{C}H_{U}$ were averaged, resulting in 25 final calibration matrices. The groups were chosen such that all five transformations are always based on the same intrinsic parameter set. This averaging of the transformation matrices is performed to decrease the influence of small errors in the pose estimations and increase the robustness of the calibration. Since this effectively only requires capturing five frames in succession as opposed to just one, this has virtually no significance for the time it takes to acquire the calibration data.
Figure 7.2 shows the backprojection for the calibration with the smallest cpe_h. Table 7.4 shows the cpe_h for all 25 calibrations, and figure 7.3 shows a bar graph representation of the same data.
Table 7.3 shows the statistical results for all 25 phantom calibration based backprojections. The root mean square of the cpe_e and cpe_h values of all backprojections give an estimate on the overall calibration quality. These values, containing the sum of the errors for the $n = 20$ point pairs, divided by n show the global root mean squared backprojection error per point pair for all calibrations. This value gives an indication on the expected rms error for a single backprojection. The statistics also contain the standard deviation of the error measures, providing insight into the robustness of the calibration method, and the minimum and maximum error values.

$rms(cpe_h)$	135.0 px
$rms(cpe_h)/n$	6.8 px
$rms(cpe_e)$	883.5 px
$rms(cpe_e)/n$	44.2 px
$std(cpe_h)$	8.7 px
$std(cpe_e)$	243.0 px
$min(cpe_h)$	122.1 px
$min(cpe_e)$	599.2 px
$max(cpe_h)$	156.3 px
$max(cpe_e)$	1150.4 px

Table 7.3: Statistics for the cumulative pixel error values cpe_e and cpe_h for the n backprojected point pairs using the phantom based calibration results.

Series / Calibration	1	2	3	4	5
1	134.4025	143.0106	125.2891	132.6114	127.3032
2	142.4239	156.2790	135.2819	132.8594	129.0526
3	133.4166	150.7048	127.8473	133.9132	126.7974
4	136.8322	153.0816	130.6254	133.1631	127.6720
5	139.6033	137.6144	122.0989	129.1382	128.0037

Table 7.4: Cumulative backprojection height error cpe_h in pixels for phantom based calibration

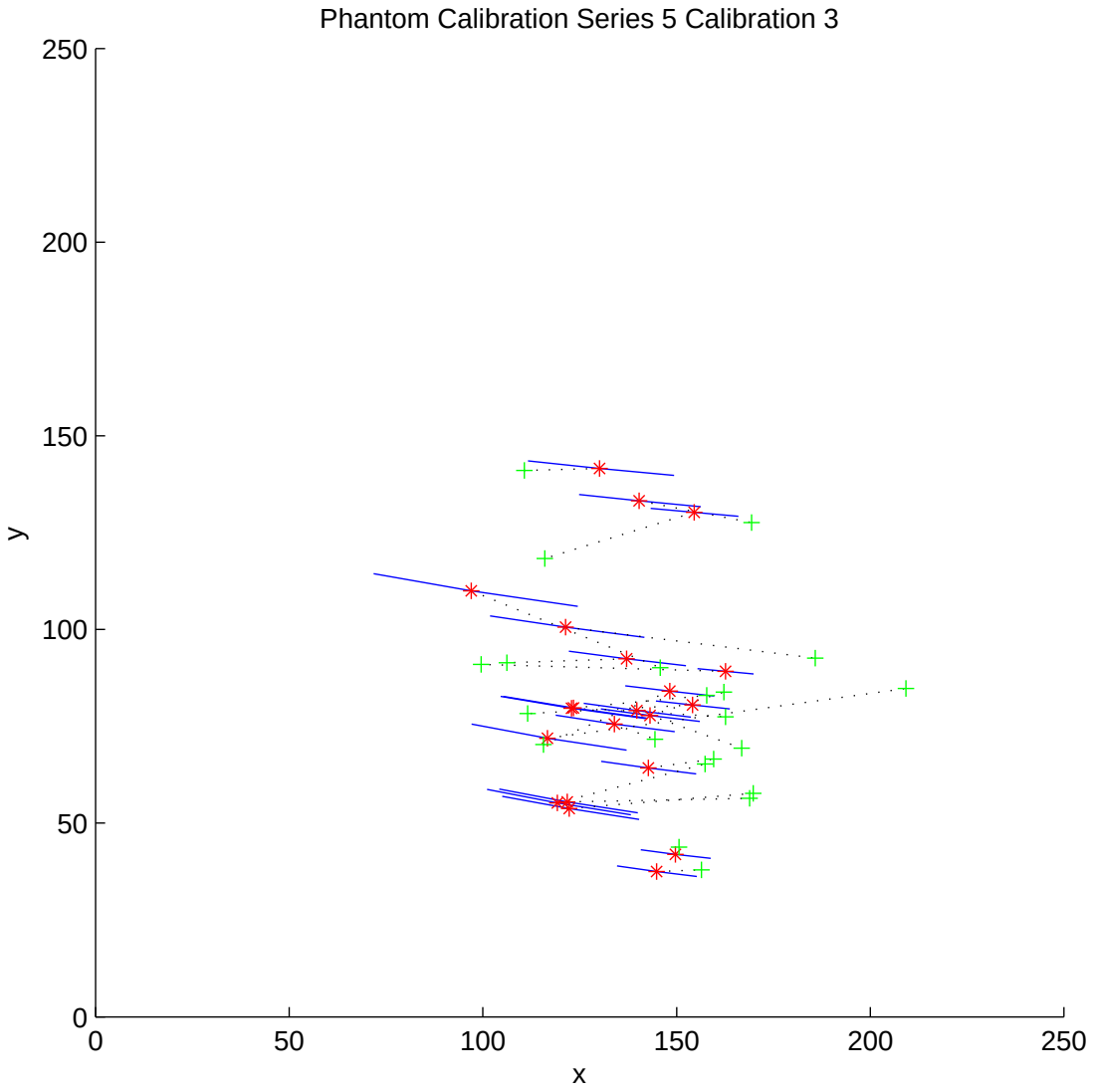


Figure 7.2: Best backprojection of the phantom based calibrations. The pluses show the tip locations in the camera image, the asterisks show the backprojected points from the ultrasound image. The bars on the backprojected points show the estimated ultrasound beam thickness at this position, and the dashed lines show the point pairs.

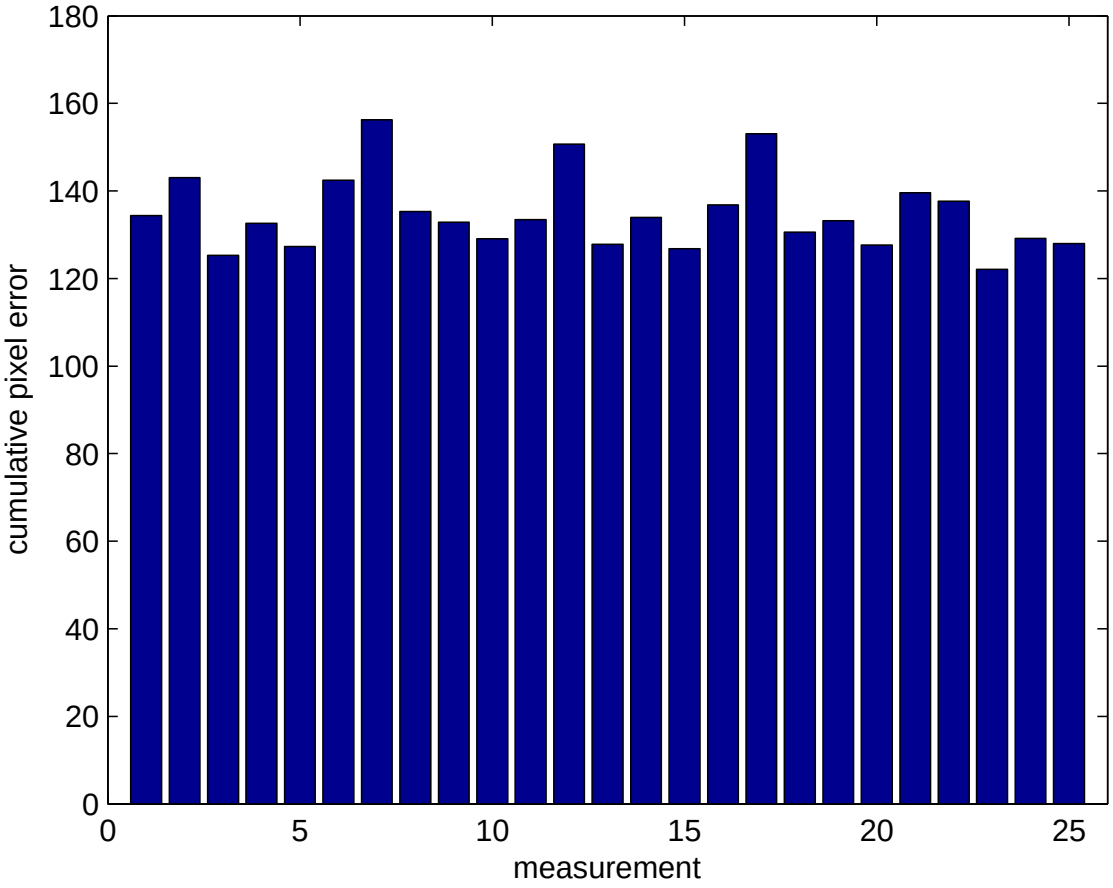


Figure 7.3: Phantom based calibration accuracies.

As can be seen from the data, the $rms(cpe_e)$ measure is much higher than the $rms(cpe_h)$ measure at a ratio of $\frac{rms(cpe_e)}{rms(cpe_h)} = 6.5$, due to the ultrasound beam thickness discussed in subsection 7.2.1. While the backprojection error per point pair for the height component $rms(cpe_h)/n$ may be considered acceptable at 6.8 pixels for an endoscopic camera and miniature ultrasound system, the euclidean distance error measure is very high at $rms(cpe_e)/n = 44.2$ pixels.

7.4 Hand-Eye and Single Wall Calibration

A hand-eye calibration was performed from the same images as the intrinsic camera calibrations using an attached NDI Aurora 6DOF position sensor. Since this calibration is performed in air, there is no need for adaptation to a water based camera system with this method. Table 7.5 shows the backprojection errors for the hand-eye calibrations, which is the mean euclidean distance in pixels between the backprojected points and their original segmentation in the source images.
A single-wall calibration was also performed five times, with table 7.6 showing the qualitative calibration results obtained by backprojecting a line onto the live ultrasound image stream and backprojecting a segmented pointer tip in captured ultrasound images.

Calibration	Backprojection Error (pixels)
1	8.4880
2	5.2993
3	9.9622
4	5.8400
5	6.8718

Table 7.5: Backprojection error for hand-eye calibration series

Calibration	Line-to-Line Quality	Pointer Tip Quality
1	OK	10 mm deviation
2	OK	OK
3	OK	OK
4	OK	OK
5	Best	Best

Table 7.6: Line and pointer backprojection quality for single-wall calibrations

The five hand-eye and five single-wall calibration transformations were combined in all permutations to guarantee the pairing of the best of both calibrations. Each of the resulting 25 transformation matrices was then used for a backprojection of the verification data in conjunction with the matching intrinsic parameters.
Figure 7.4 shows the backprojected point pairs for the calibration result with the lowest cpe_h error. Table 7.7 shows the cpe_h measurements for all 25 backprojections. Figure 7.5 shows a bar graph representation of the cpe_h values. Finally, table 7.8 show the results and statistics from the backprojections, analogous to the statistics of the phantom based calibration.

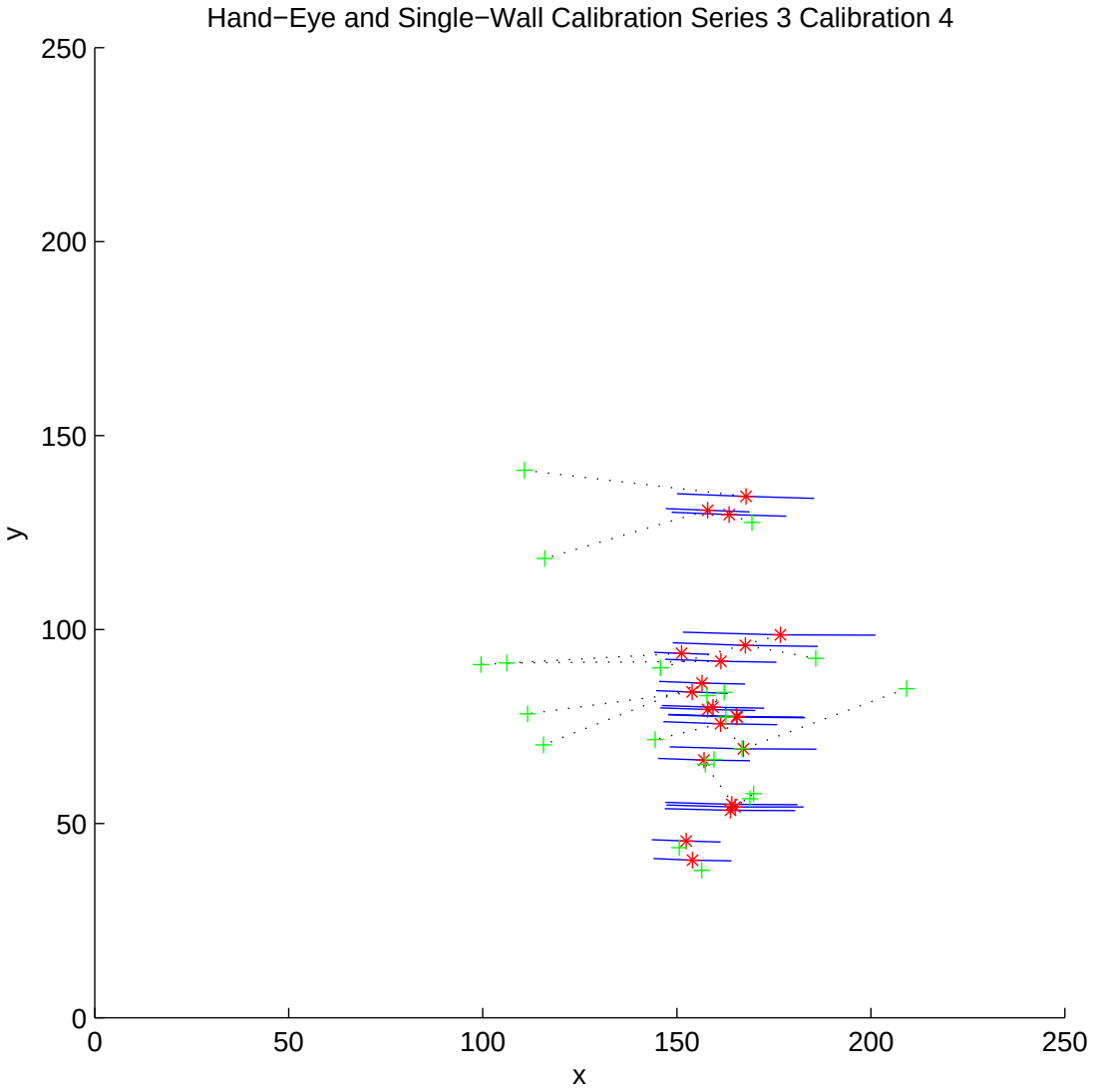


Figure 7.4: Best backprojection of the hand-eye and single-wall based calibrations. The pluses show the tip locations in the camera image, the asterisks show the backprojected points from the ultrasound image. The bars on the backprojected points show the estimated ultrasound beam thickness at this position, and the dashed lines show the point pairs.

Series / Calibration	1	2	3	4	5
1	1677.6	300.93	468.78	448.31	606.90
2	1300.1	173.37	158.69	144.24	264.24
3	1152.0	198.81	127.95	119.26	196.18
4	1315.2	170.41	163.61	146.99	270.37
5	1324.5	164.28	170.65	153.47	276.03

Table 7.7: Cumulative backprojection height error cpe_h for hand-eye and single-wall calibrations.

$rms(cpe_h)$	656.0 px
$rms(cpe_h)/n$	32.8 px
$rms(cpe_e)$	951.0 px
$rms(cpe_e)/n$	47.6 px
$std(cpe_h)$	477.5 px
$std(cpe_e)$	385.2 px
$min(cpe_h)$	119.3 px
$min(cpe_e)$	479.3 px
$max(cpe_h)$	1677.6 px
$max(cpe_e)$	1937.4 px

Table 7.8: Statistics for the cumulative pixel error values cpe_e and cpe_h for the n backprojected point pairs using the hand-eye and single-wall calibration results.

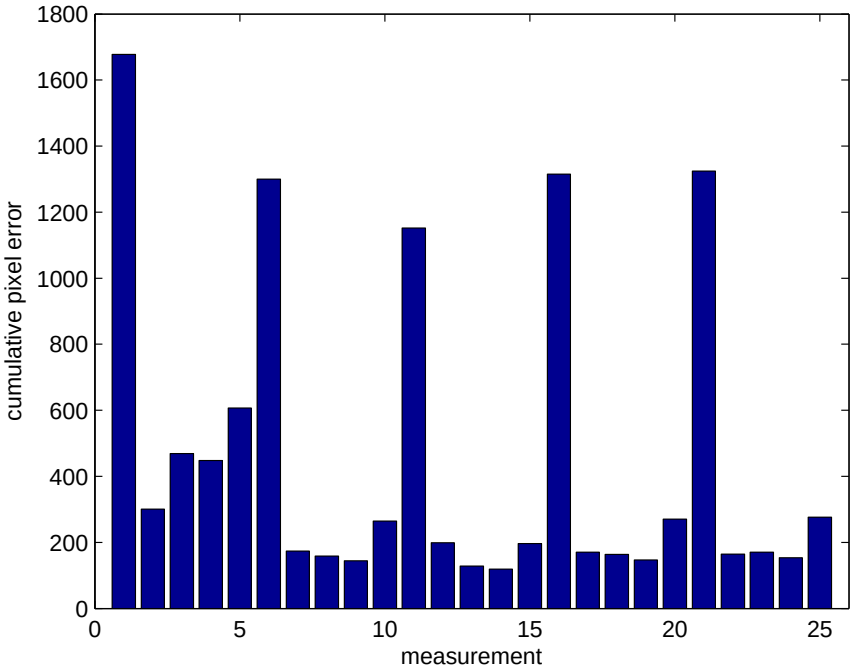


Figure 7.5: Hand-eye and single-wall based calibration accuracies.

Compared to the phantom based calibration, the $rms(cpe_e)$ and $rms(cpe_h)$ values are larger, while their ratio is smaller at $\frac{cpe_e}{cpe_h} = 1.4$. If the best in class reprojection errors $min(cpe_h)$ are considered, both methods give similar results at 119.3 pixels for the hand-eye and single-wall method, and 122.1 pixels for the phantom based method.
The high standard deviations for the reprojection error measures suggest that the method is more fragile than the phantom based calibration. Especially in the bar graph representation of the calibrations, a pattern of outstandingly high error values can be observed. These particular calibrations are all based on the first single-wall calibration which also produces a 10 mm deviation in the pointer tip validation. It is thus likely that this calibration is inaccurate and subsequently causing the high backprojection errors. However, even if this calibration is considered an outlier, the cpe_h errors are still deviating significantly more than with the phantom based calibrations, as can be seen in the bar graph representations.

7.5 Conclusion

The best calibration results of each method, defined as the calibration with the lowest cpe_h, are on par at 119.3 pixels for the hand-eye and single-wall method, and 122.1 pixels for the phantom based method. It is likely that this maximum accuracy is limited by the error of the ultrasound calibrations, since this modality has a lower resolution and more noise than the camera system.

The phantom based calibration is clearly more robust, supported by the significantly lower average error values and standard deviations. In conjunction with the fact that the phantom based calibration can be performed within a very short time, the robust results can be considered an advantage over the hand-eye and single-wall approach. Furthermore, a tracking system is required only during phantom construction, which may be an advantage depending on the intended application.

The goal of this thesis to develop a calibration method for finding the spatial relation between the camera system and ultrasound transducer has been reached. While the accuracy achieved by the method is not enough for image fusion applications, it achieves an accuracy similar to the established hand-eye and single-wall calibrations, while being faster and more robust.

7.6 Future Work

The most deteriorating influence for the accuracy during the phantom based calibration is most probably the ultrasound pose estimation. As can be seen in figure 6.4, the hollow tubes appear as large smeared circles, leaving a large margin for segmentation errors. Furthermore, the four fiducials provide only one redundant point for the point based registration, leaving little capabilities for error compensation.

If an improved phantom with more and smaller fiducials could be constructed the calibration accuracy will improve, as found by Pagoulatos et al. [62]. For this, a segmentation reference similar to the rubber plate used by Hsu et al. [41] would be required. The problem of the obstructed camera view caused by this plate could be overcome by placing the reference at the bottom of the phantom, as opposed to [41].

Another challenge is the magnification effect caused by placing the camera underwater. This leads to comparatively few feature points from the pattern being visible, adversely affecting the pose estimation quality. A more thorough evaluation of the characteristics of this camera system may lead to an improved pattern and pose estimation algorithm.

7.7 Applications

A possible use for the calibration data is freehand 3D ultrasound. In a navigation system capable of image based camera to CT registration, the precise position of the camera inside a CT coordinate system is available. Using the calibration transformation, ultrasound scans could then be compounded directly inside this coordinate system.

Bibliography

[1] D.V. Amin, T. Kanade, B. Jaramaz, A.M. DiGioia, C. Nikou, R.S. LaBarca, and JE Moody. Calibration method for determining the physical location of the ultrasound image plane. *Lecture Notes in Computer Science*, pages 940–947, 2001.

[2] KS Arun, TS Huang, and SD Blostein. Least-squares fitting of two 3-D point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):698–700, 1987.

[3] CD Barry, CP Allott, NW John, PM Mellor, PA Arundel, DS Thomson, and JC Waterton. Three-dimensional freehand ultrasound: image reconstruction and volume analysis. *Ultrasound in medicine & biology*, 23(8):1209–1224, 1997.

[4] N. Bilaniuk and G.S.K. Wong. Speed of sound in pure water as a function of temperature. *The Journal of the Acoustical Society of America*, 93:2306, 1993.

[5] W. Birkfellner, F. Watzinger, F. Wanschitz, G. Enislidis, C. Kollmann, D. Rafolt, R. Nowotny, R. Ewers, and H. Bergmann. Systematic distortions in magnetic position digitizers. *Medical Physics*, 25:2242, 1998.

[6] E.M. Boctor, A. Jain, M.A. Choti, R.H. Taylor, and G. Fichtinger. A rapid calibration method for registration and 3D tracking of ultrasound images using spatial localizer. *SPIE Medical imaging2003, San Diego*, pages 15–20, 2003.

[7] L.G. Bouchet, S.L. Meeks, G. Goodchild, F.J. Bova, J.M. Buatti, and W.A. Friedman. Calibration of three-dimensional ultrasound images for image-guided radiation therapy. *Physics in Medicine and Biology*, 46(2):559–578, 2001.

[8] J. Bouguet. Camera calibration toolbox for matlab. `http://www.vision.caltech.edu/bouguetj/calib_doc/`, December 2009.

[9] B. Brendel, S. Winter, and H. Ermert. A simple and accurate calibration method for 3D freehand ultrasound. *Biomedizinische Technik*, 49(Suppl 2):872–873, 2004.

[10] I. Bricault, G. Ferretti, and P. Cinquin. Registration of real and CT-derived virtual bronchoscopic images to assist transbronchial biopsy. *IEEE Transactions on Medical Imaging*, 17(5):703–714, October 1998.

[11] M. Burcher. *A force-based method for correcting deformation in ultrasound images of the breast*. University of Oxford, 2002.

[12] J. Carr. Surface reconstruction in 3D medical imaging. *University of Canterbury, Christchurch, New Zealand*, 1996.

[13] J.C. Carr, J.L. Stallkamp, M.M. Fynes, A.H. Gee, R.W. Prager, G.M. Treece, C. Overton, and L. Berman. Design of a clinical free-hand 3D ultrasound system. In *Proceedings of SPIE*, volume 3982, page 14, 2000.

[14] C.T. Chen, L.S. Chen, and F.J. Millero. Speed of sound in NaCl, MgCl, NaSO, and MgSO aqueous solutions as functions of concentration, temperature, and pressure. *The Journal of the Acoustical Society of America*, 63:1795, 1978.

[15] C.T. Chen and F.J. Millero. Reevaluation of Wilson's sound-speed measurements for pure water. *J. Acoust. Soc. Am*, 60(6), 1976.

[16] J.C.K. Chou and M. Kamel. Quaternions Approach to Solve the Kinematic Equation of Rotation AaAx= AxAb of a Sensor Mounted Robotic Manipulator. In *IEEE International Conference on Robotics and Automation*, 1988.

[17] A.J. Chung, F. Deligianni, P. Shah, A. Wells, and G.Z. Yang. Patient-specific bronchoscopy visualization through BRDF estimation and disocclusion correction. *IEEE transactions on medical imaging*, 25(4):503, 2006.

[18] P. Cinquin, F. Gerritsen, and D.G. von Keyserlinck. Computer assisted medical interventions. *Health in the New Communications Age: Health Care Telematics for the 21st Century*, page 128, 1995.

[19] R.M. Comeau, A. Fenster, and T.M. Peters. Integrated MR and ultrasound imaging for improved image guidance in neurosurgery. In *Proceedings of SPIE*, volume 3338, page 747, 1998.

[20] S. Dandekar, Y. Li, J. Molloy, and J. Hossack. A phantom with reduced complexity for spatial 3-D ultrasound calibration. *Ultrasound in medicine & biology*, 31(8):1083–1093, 2005.

[21] K. Daniilidis. Hand-eye calibration using dual quaternions. *The International Journal of Robotics Research*, 18(3):286, 1999.

[22] F.R. de Castro, F.D. López, G.J. Serdà, A.R. López, J.F. Gilart, and P.C. Navarro. Relevance of training in transbronchial fine-needle aspiration technique. *Chest*, 111(1):103, 1997.

[23] Daisuke Deguchi, Kensaku Mori, Marco Feuerstein, Takayuki Kitasaka, Calvin R. Maurer, Jr., Yasuhito Suenaga, Hirotsugu Takabatake, Masaki Mori, and Hiroshi Natori. Selective image similarity measure for bronchoscope tracking based on image registration. *Medical Image Analysis*, 13(4):621–633, August 2009.

[24] F. Deligianni, A.J. Chung, and G. Yang. Nonrigid 2-D/3-D registration for patient specific bronchoscopy simulation with statistical shape modeling: Phantom validation. *IEEE transactions on medical imaging*, 25(11):1462, 2006.

[25] PR Detmer, G. Bashein, T. Hodges, KW Beach, EP Filer, DH Burns, and DE Strandness Jr. 3D ultrasonic image feature localization based on magnetic scanhead tracking: in vitro calibration and validation. *Ultrasound in medicine & biology*, 20(9):923, 1994.

[26] M. Fischler and R. Bolles. Random Sample Consensus: A Paradigm for Model Fitting With Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[27] D.D. Frantz, S.R. Kirsch, and A.D. Wiles. Specifying 3D Tracking System Accuracy. In *Workshop Bildverarbeitung fur die Medizin, Berlin*, 2004.

[28] AL Fuhrmann, R. Splechtna, and J. Přikryl. Comprehensive calibration and registration procedures for augmented reality. In *Immersive Projection Technology and Virtual Environments 2001: Proceedings of the Eurographics Workshop in Stuttgart, Germany, May 16-18, 2001*, page 219. Springer Verlag Wien, 2001.

[29] A.H. Gee, N.E. Houghton, G.M. Treece, and R.W. Prager. A mechanical instrument for 3D ultrasound probe calibration. *Ultrasound in medicine & biology*, 31(4):505–518, 2005.

[30] B. Geiger and R. Kikinis. Simulation of endoscopy. *Lecture Notes in Computer Science*, 905:277–281, 1995.

[31] A. Goldstein. The effect of acoustic velocity on phantom measurements. *Ultrasound in medicine & biology*, 26(7):1133–1143, 2000.

[32] M.J. Gooding, S.H. Kennedy, and J.A. Noble. Temporal calibration of freehand three-dimensional ultrasound using image alignment. *Ultrasound in medicine & biology*, 31(7):919–927, 2005.

[33] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge Univ Pr, 2003.

[34] J. Heikkila and O. Silven. A four-step camera calibration procedure with implicit image correction. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1106–1112. INSTITUTE OF ELECTRICAL ENGINEERS INC (IEEE), 1997.

[35] JP Helferty and WE Higgins. Technique for registering 3D virtual CT images to endoscopic video. In *Image Processing, 2001. Proceedings. 2001 International Conference on*, volume 2, 2001.

[36] JP Helferty, AJ Sherbondy, AP Kiraly, and WE Higgins. Computer-based system for the virtual-endoscopic guidance of bronchoscopy. *Computer Vision and Image Understanding*, 108(1-2):171–187, 2007.

[37] W.E. Higgins, J.P. Helferty, K. Lu, S.A. Merritt, L. Rai, and K.C. Yu. 3D CT-video fusion for image-guided bronchoscopy. *Computerized Medical Imaging and Graphics*, 2007.

[38] L. Hong, S. Muraki, A. Kaufman, D. Bartz, and T. He. Virtual voyage: Interactive navigation in the human colon. In *Proc. of ACM Siggraph*, pages 27–34. Citeseer, 1997.

[39] PVC Hough. Machine analysis of bubble chamber pictures. In *International Conference on High Energy Accelerators and Instrumentation*, volume 73, 1959.

[40] P.W. Hsu, R.W. Prager, A.H. Gee, and G.M. Treece. Rapid, easy and reliable calibration for freehand 3D ultrasound. *Ultrasound in medicine & biology*, 32(6):823–835, 2006.

[41] P.W. Hsu, R.W. Prager, A.H. Gee, and G.M. Treece. Real-time freehand 3D ultrasound calibration. *Ultrasound in Medicine & Biology*, 34(2):239–251, 2008.

[42] QH Huang, YP Zheng, MH Lu, and ZR Chi. Development of a portable 3D ultrasound imaging system for musculoskeletal tissues. *Ultrasonics*, 43(3):153–163, 2005.

[43] Intel Inc. The OpenCV Open Source Computer Vision Library. `http://opencvlibrary.sourceforge.net`, November 2009.

[44] Norhtern Digital Inc. How does the ndi polaris system work. `http://www.ndigital.com/medical/documents/polaris/polarishdtw_datasheet.pdf`, November 2009.

[45] Norhtern Digital Inc. Polaris spectra measurement volume. `http://www.ndigital.com/medical/documents/polaris/polarisspectra/Spectra_meas_volume_datasheet.pdf`, November 2009.

[46] A. Khamene and F. Sauer. A novel phantom-less spatial and temporal ultrasound calibration method. *Lecture Notes in Computer Science*, 3750:65, 2005.

[47] V.V. Kindratenko. A survey of electromagnetic position tracker calibration techniques. *Virtual Reality*, 5(3):169–182, 2000.

[48] T. Langø. *Ultrasound guided surgery: image processing and navigation*. PhD thesis, Norwegian University of Science and Technology, Faculty of Information Technology, Mathematics and Electrical Engineering, 2001.

[49] J. Lavest, G. Rives, and J. Lapreste. Underwater camera calibration. *Lecture notes in computer science*, pages 654–668, 2000.

[50] JM Lavest, G. Rives, and M. Dhome. 3D reconstruction by zooming. *Intelligent Autonomous System, Pittsburg*, 1993.

[51] F. Lindseth, G.A. Tangen, T. Langų, and J. Bang. Probe calibration for freehand 3-D ultrasound. *Ultrasound in medicine & biology*, 29(11):1607–1623, 2003.

[52] S. Meairs, J. Beyer, and M. Hennerici. Reconstruction and visualization of irregularly sampled three-and four-dimensional ultrasound data for cerebrovascular applications. *Ultrasound in medicine & biology*, 26(2):263–272, 2000.

[53] S.A. Merritt, L. Rai, and W.E. Higgins. Real-time CT-video registration for continuous endoscopic guidance. In *Proceedings of SPIE*, volume 6143, page 614313, 2006.

[54] K. Meyer, H.L. Applewhite, and F.A. Biocca. A survey of position trackers. In *Presence: Teleoperators and Virtual Environments (ISSN 1054-7460), vol. 1, no. 2, p. 173-200.*, volume 1, pages 173–200, 1992.

[55] J.J. More. The Levenberg-Marquardt algorithm: implementation and theory. *Lecture notes in mathematics*, 630:105–116, 1977.

[56] K. Mori, D. Deguchi, J. Sugiyama, Y. Suenaga, J. Toriwaki, C. R. Maurer, H. Takabatake, and H. Natori. Tracking of a bronchoscope using epipolar geometry analysis and intensity-based image registration of real and virtual endoscopic images. *Medical Image Analysis*, 6(3):321–336, September 2002.

[57] K. MORI, A. URANO, J. HASEGAWA, J. TORIWAKI, H. ANNO, and K. KATADA. Virtualized Endoscope System–An Application of Virtual Reality Technology to Diagnostic Aid–. *IEICE TRANSACTIONS on Information and Systems*, 79(6):809–819, 1996.

[58] Kensaku Mori, Daisuke Deguchi, Kenta Akiyama, Takayuki Kitasaka, Calvin R Maurer, Yasuhito Suenaga, Hirotsugu Takabatake, Masaki Mori, and Hiroshi Natori. Hybrid bronchoscope tracking using a magnetic tracking sensor and image registration. volume 3750, pages 543–550, 2005.

[59] D.M. Muratore, R.L. Galloway, et al. Beam calibration without a phantom for creating a 3-D freehand ultrasound system. *Ultrasound in medicine & biology*, 27(11):1557–1566, 2001.

[60] M. Nakamoto, Y. Sato, K. Nakada, Y. Nakajima, K. Konishi, M. Hashizume, and S. Tamura. A temporal calibration method for freehand 3D ultrasound system: a preliminary result. In *International Congress Series*, page 1256:1365, 2003.

[61] N. Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11:285–296, 1975.

[62] N. Pagoulatos, D.R. Haynor, and Y. Kim. A fast calibration method for 3-D tracking of ultrasound images using a spatial localizer. *Ultrasound in medicine & biology*, 27(9):1219–1229, 2001.

[63] FC Park and BJ Martin. Robot sensor calibration: solving AX= XB on the Euclidean group. *IEEE Transactions on Robotics and Automation*, 10(5):717–721, 1994.

[64] D.M. Parkin, F. Bray, J. Ferlay, and P. Pisani. Global cancer statistics, 2002. *CA: a cancer journal for clinicians*, 55(2):74, 2005.

[65] R.W. Prager, A. Gee, and L. Berman. Stradx: real-time acquisition and visualization of freehand three-dimensional ultrasound. *Medical Image Analysis*, 3(2):129–140, 1999.

[66] RW Prager, RN Rohling, AH Gee, and L. Berman. Rapid calibration for 3-D freehand ultrasound. *Ultrasound in medicine & biology*, 24(6):855–869, 1998.

[67] FH Raab, EB Blood, TO Steiner, and HR Jones. Magnetic position and orientation tracking system. *IEEE Transactions on Aerospace and Electronic Systems*, pages 709–718, 1979.

[68] F. Rousseau, P. Hellier, and C. Barillot. Confhusius: A robust and fully automatic calibration method for 3D freehand ultrasound. *Medical Image Analysis*, 9(1):25–38, 2005.

[69] F. Rousseau, P. Hellier, and C. Barillot. A novel temporal calibration method for 3-D ultrasound. *IEEE transactions on medical imaging*, 25(8):1108, 2006.

[70] GD Rubin, CF Beaulieu, V. Argiro, H. Ringl, AM Norbash, JF Feller, MD Dake, RB Jeffrey, and S. Napel. Perspective volume rendering of CT and MR images: applications for endoscopic imaging. *Radiology*, 199(2):321, 1996.

[71] Y. Sato, M. Nakamoto, Y. Tamaki, T. Sasama, I. Sakita, Y. Nakajima, M. Monden, and S. Tamura. Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization. *IEEE Transactions on Medical Imaging*, 17(5):681–693, 1998.

[72] A. Schneider, H. Hautmann, H. Barfuss, T. Pinkau, F. Peltz, H. Feussner, and A. Wichert. Real-time image tracking of a flexible bronchoscope. volume 1268, pages 753–757, 2004.

[73] R. Shahidi, MR Bax, CR Maurer Jr, JA Johnson, EP Wilkinson, B. Wang, JB West, MJ Citardi, KH Manwaring, and R. Khadem. Implementation, calibration and accuracy testing of an image-enhanced endoscopy system. *IEEE Transactions on Medical imaging*, 21(12):1524–1535, 2002.

[74] G.M. Treece, A.H. Gee, R.W. Prager, C.J.C. Cash, and L.H. Berman. High-definition freehand 3-D ultrasound. *Ultrasound in medicine & biology*, 29(4):529–546, 2003.

[75] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *Radiometry*, page 221, 1992.

[76] RY Tsai and RK Lenz. A new technique for fully autonomous and efficient 3D robotics-hand/eye calibration. *IEEE Transactions on Robotics and Automation*, 5(3):345–358, 1989.

[77] M. Tuceryan, DS Greer, RT Whitaker, DE Breen, C. Crampton, E. Rose, and KH Ahlers. Calibration requirements and procedures for a monitor-basedaugmented reality system. *IEEE Transactions on Visualization and Computer Graphics*, 1(3):255–273, 1995.

[78] JF Turner Jr and K.P. Wang. Staging of Mediastinal Involvement in Lung Cancer by Bronchoscopic Needle Aspiration: Pro Bronchoscopic Needle Aspiration. *Journal of Bronchology & Interventional Pulmonology*, 3(1):74, 1996.

[79] S. Umeyama. Least-squares estimation of transformation parameters between twopoint patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380, 1991.

[80] C. Wengert. Fully automatic camera and hand to eye calibration. `http://www.vision.ee.ethz.ch/~cwengert/calibration_toolbox.php`, December 2009.

[81] C. Wengert, M. Reeff, P.C. Cattin, and G. Székely. Fully automatic endoscope calibration for intraoperative use. *Bildverarbeitung für die Medizin*, pages 419–23, 2006.

[82] J.B. West and C.R. Maurer. Designing optically tracked instruments for image-guided surgery. *IEEE Transactions on medical imaging*, 23(5):533–545, 2004.

[83] A.D. Wiles, D.G. Thompson, and D.D. Frantz. Accuracy assessment and interpretation for optical tracking systems. In *Proc. SPIE*, volume 5367, pages 421–32, 2004.

[84] Kazuhiro Yasufuku, Masako Chiyo, Yasuo Sekine, Prashant N Chhajed, Kiyoshi Shibuya, Toshihiko Iizasa, and Takehiko Fujisawa. Real-time endobronchial ultrasound-guided transbronchial needle aspiration of mediastinal and hilar lymph nodes. *Chest*, 126(1):122–128, July 2004.

[85] Z. Zhang. Iterative point matching for registration of free-form curves. *Int. J. Comp. Vis*, 7(3):119–152, 1994.

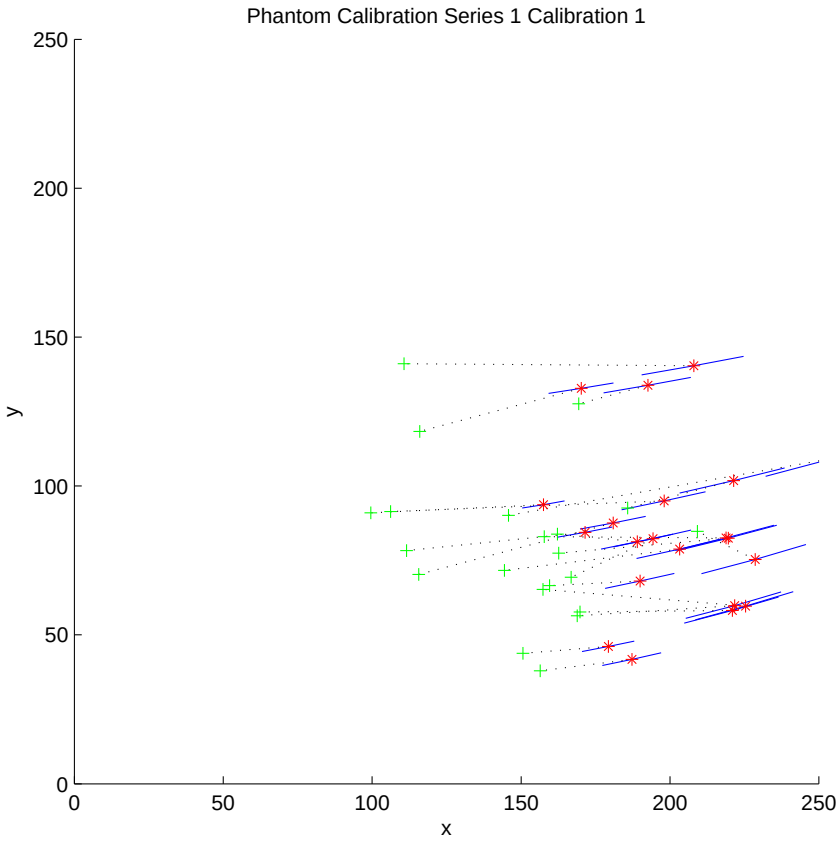
[86] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *International Conference on Computer Vision*, volume 1, pages 666–673, 1999.

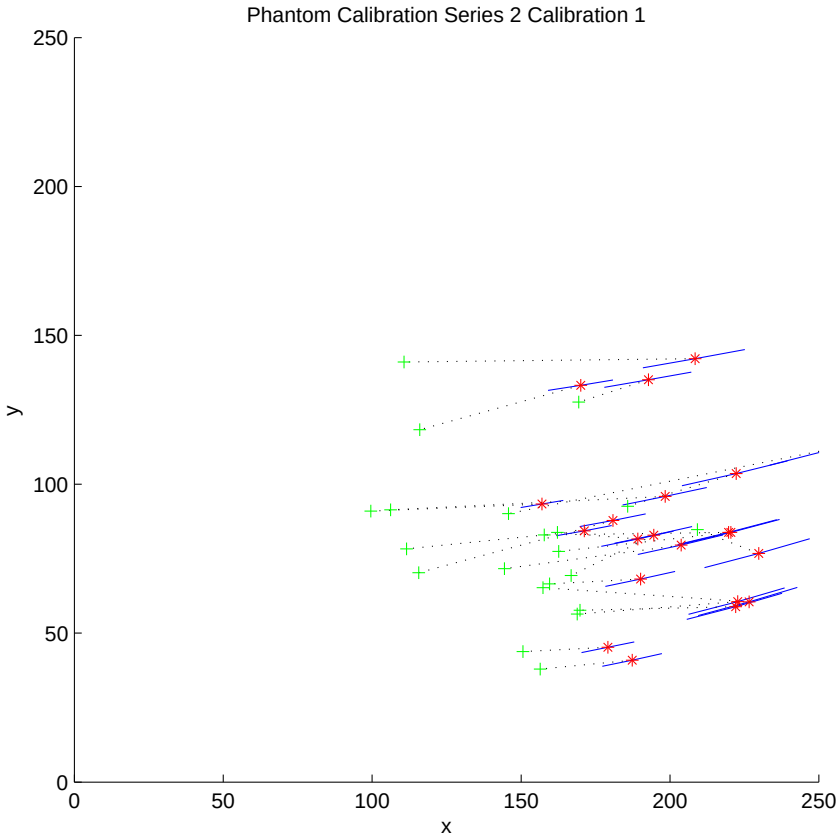
List of Figures

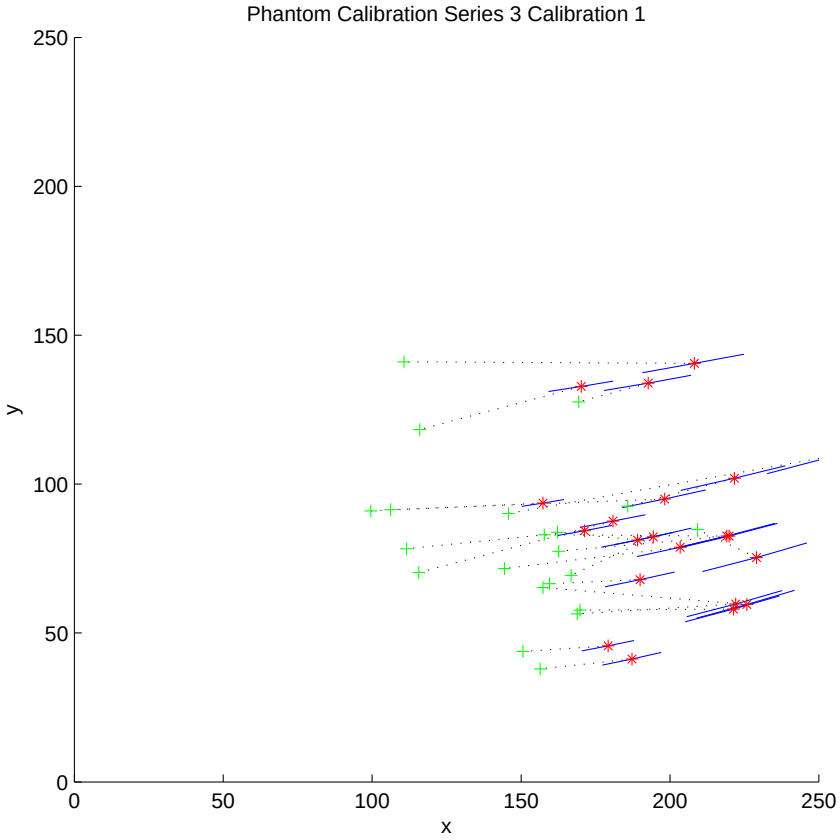
List of Tables

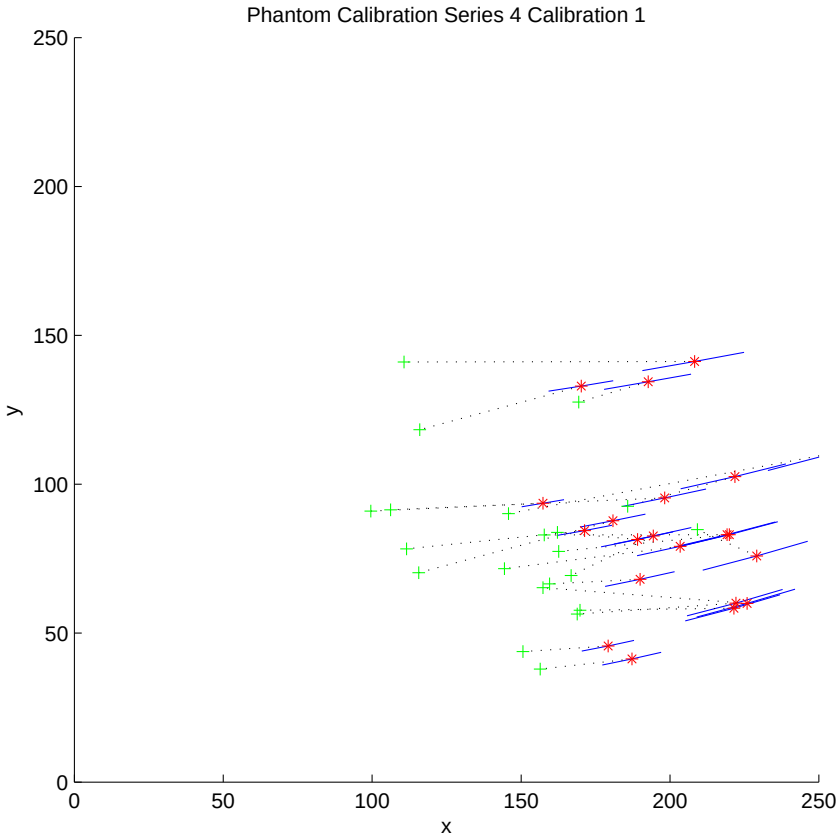
Appendix

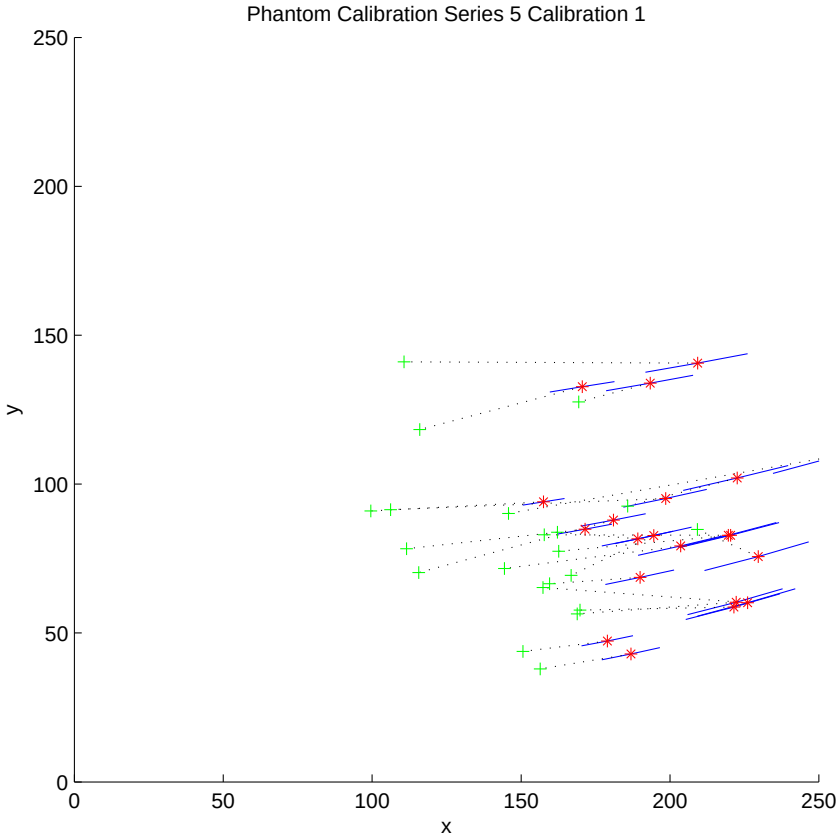
Phantom Based Calibration Backprojections

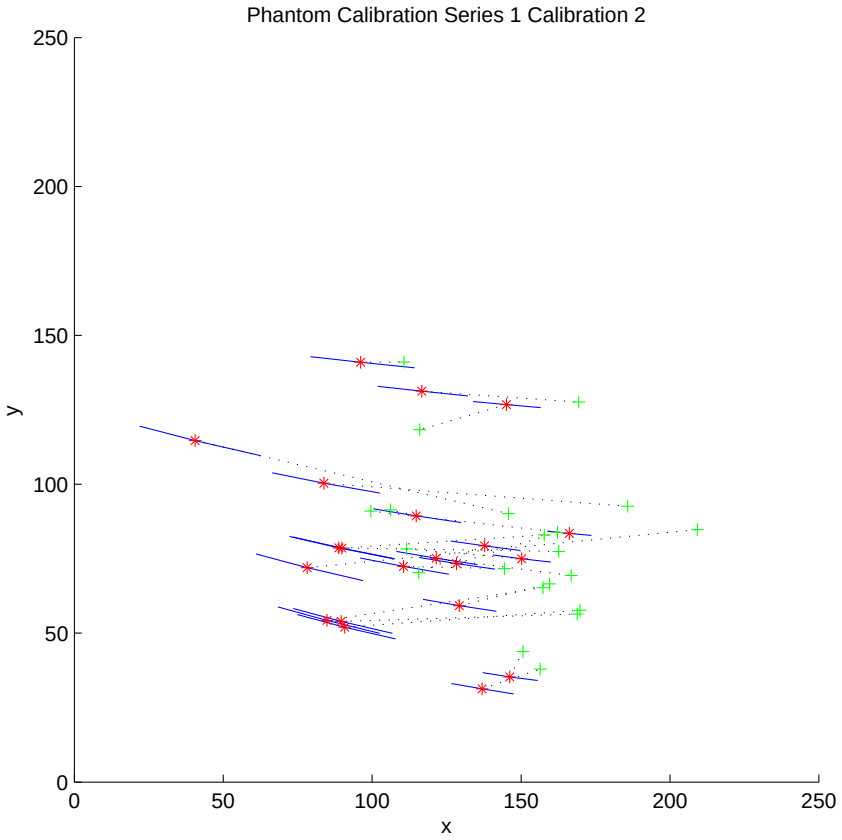


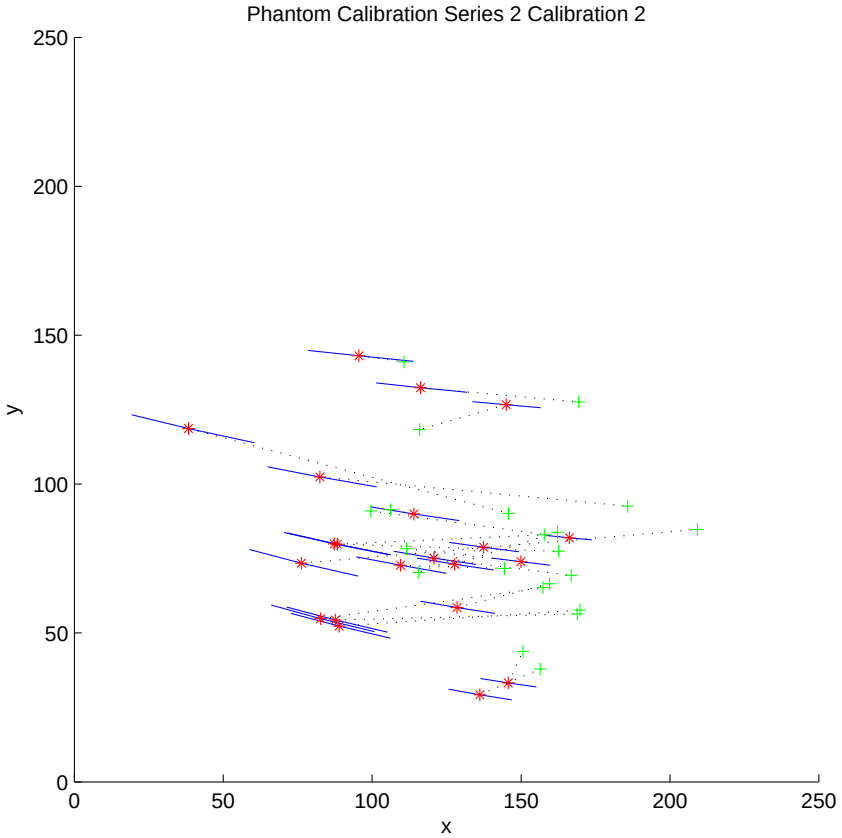

Phantom Calibration Series 2 Calibration 1
250
200
150
100
50
0
y
0
50
100
150
200
250
x

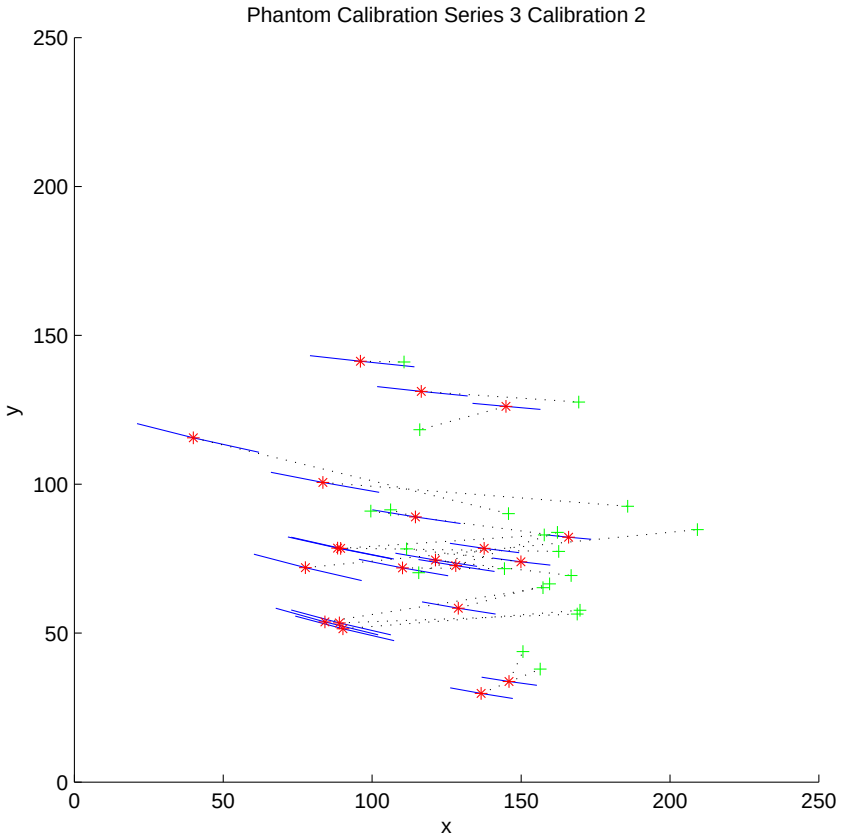

Phantom Calibration Series 3 Calibration 1
250
200
150
100
50
0
y
0
50
100
150
200
250
x

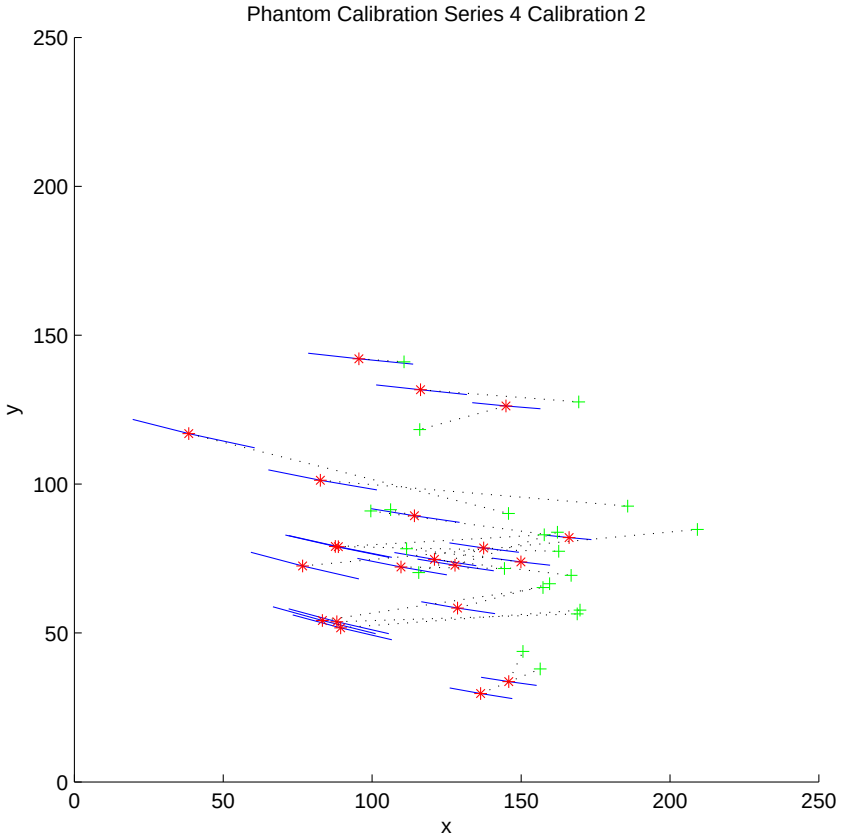

Phantom Calibration Series 4 Calibration 1
250
200
150
100
50
0
y
0
50
100
150
200
250
x

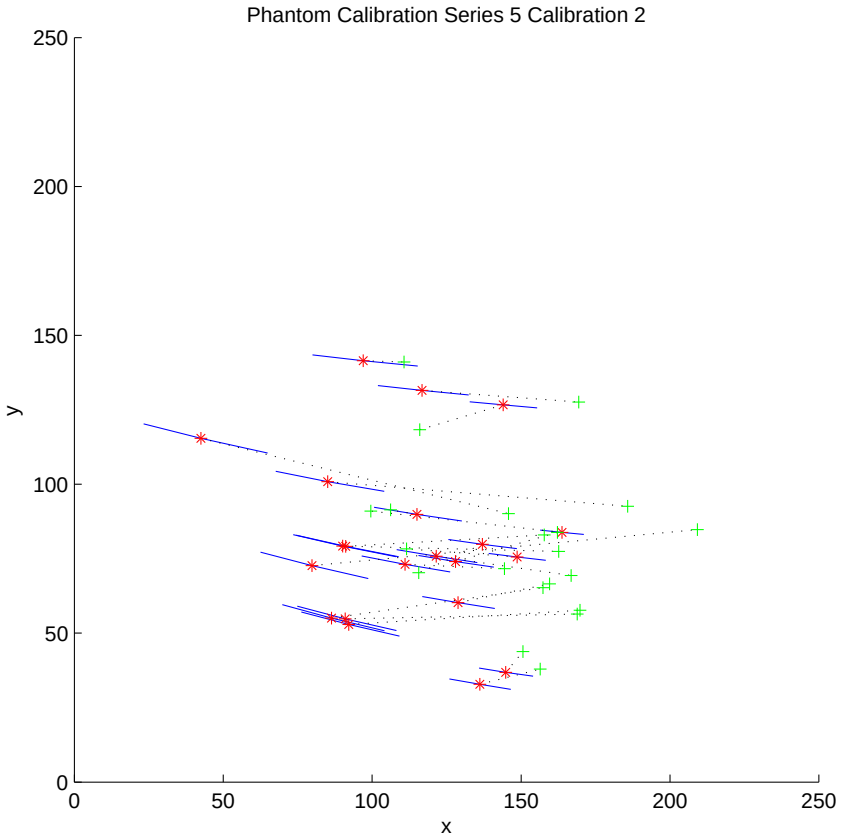

Phantom Calibration Series 5 Calibration 1
250
200
150
100
50
0
y
0
50
100
150
200
250
x

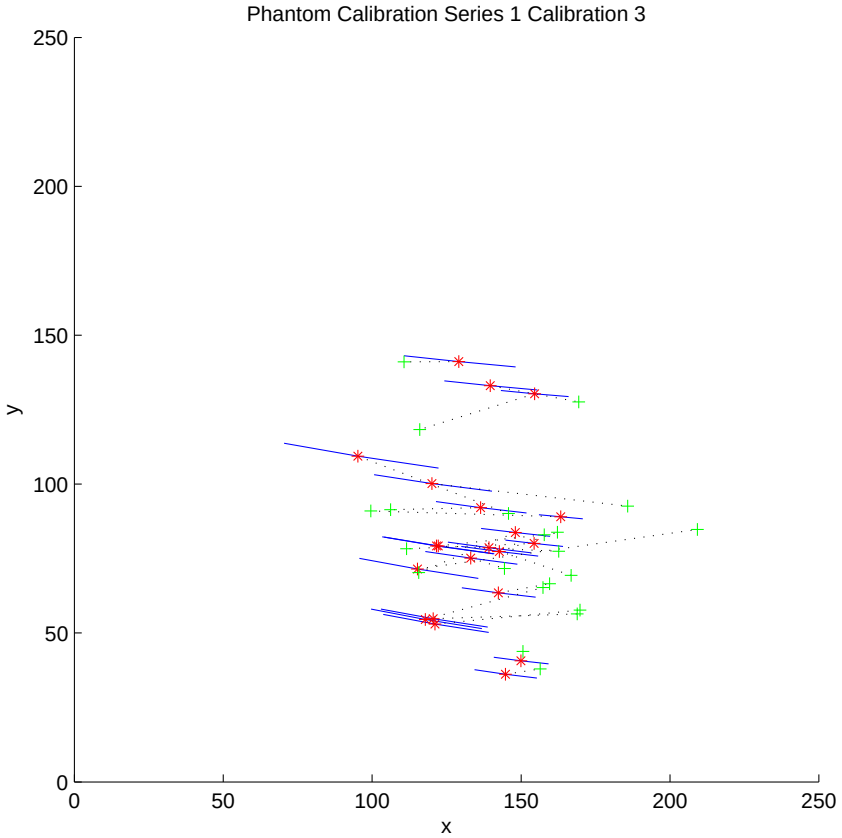

Phantom Calibration Series 1 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

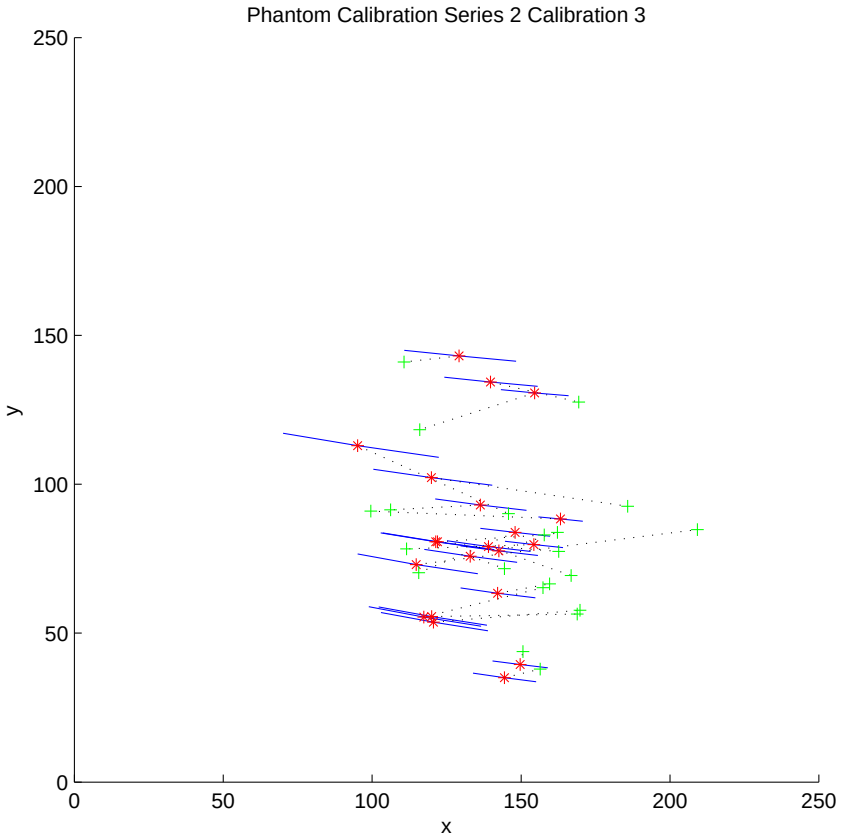

Phantom Calibration Series 2 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

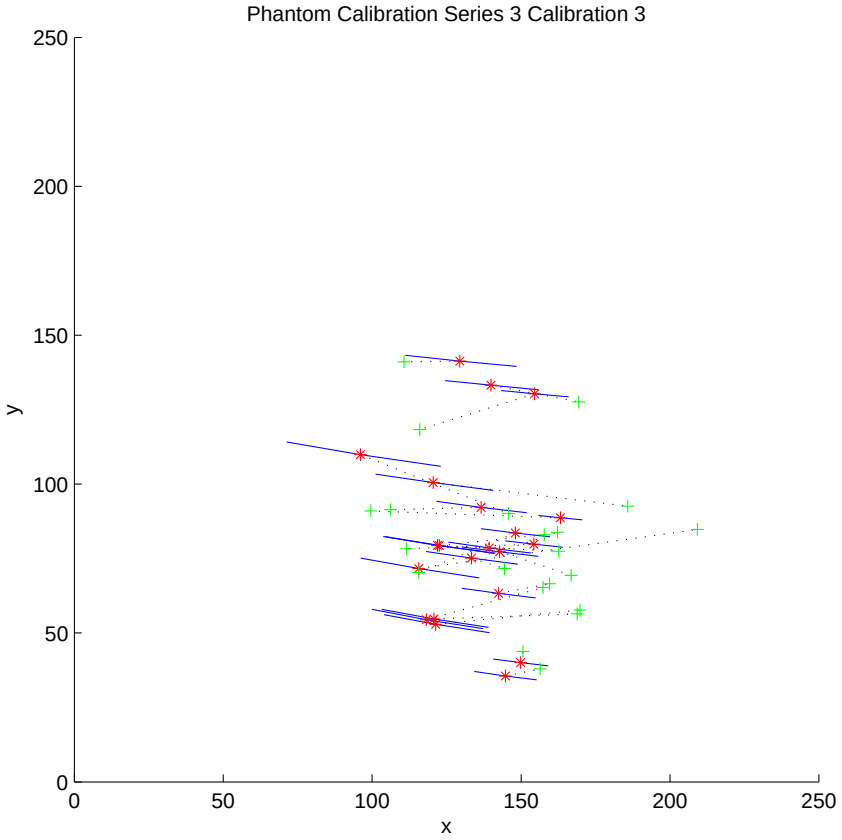

Phantom Calibration Series 3 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

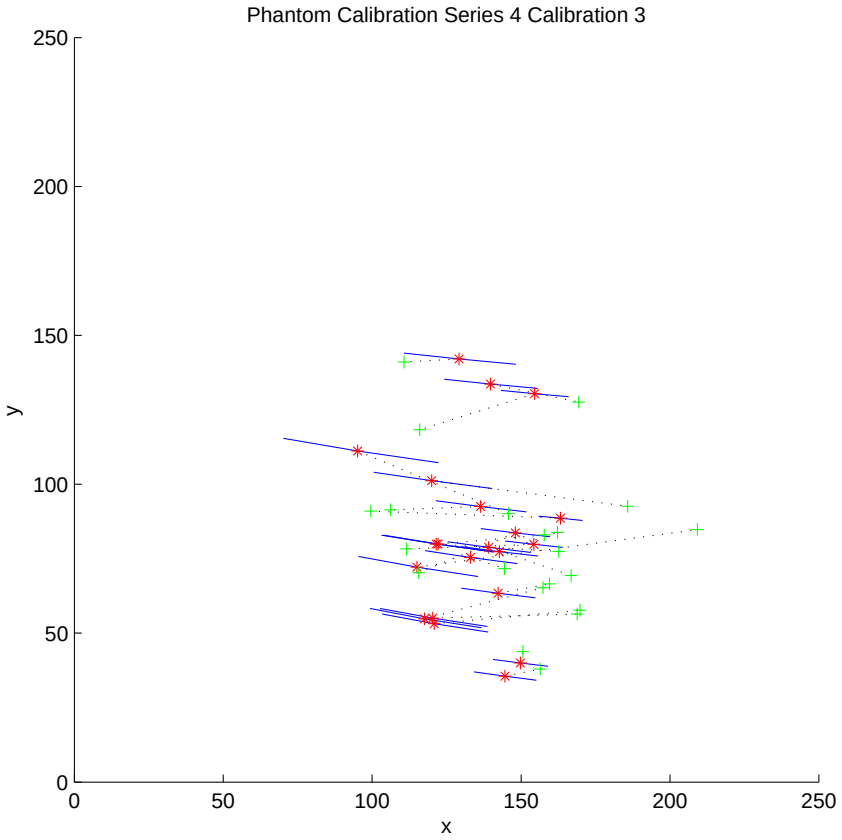

Phantom Calibration Series 4 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

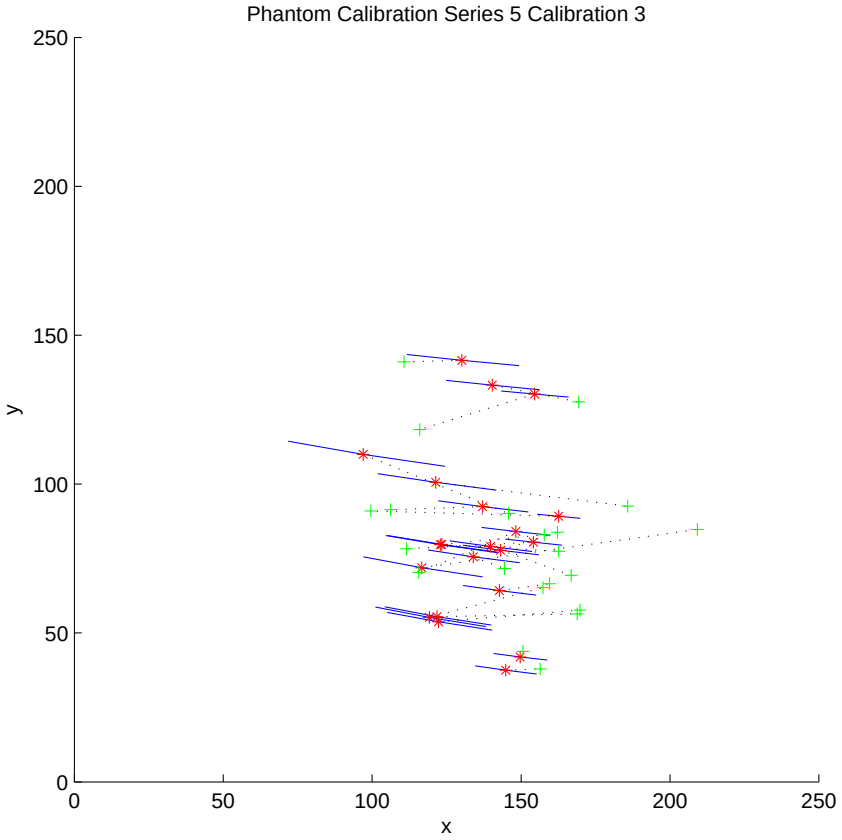

Phantom Calibration Series 5 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

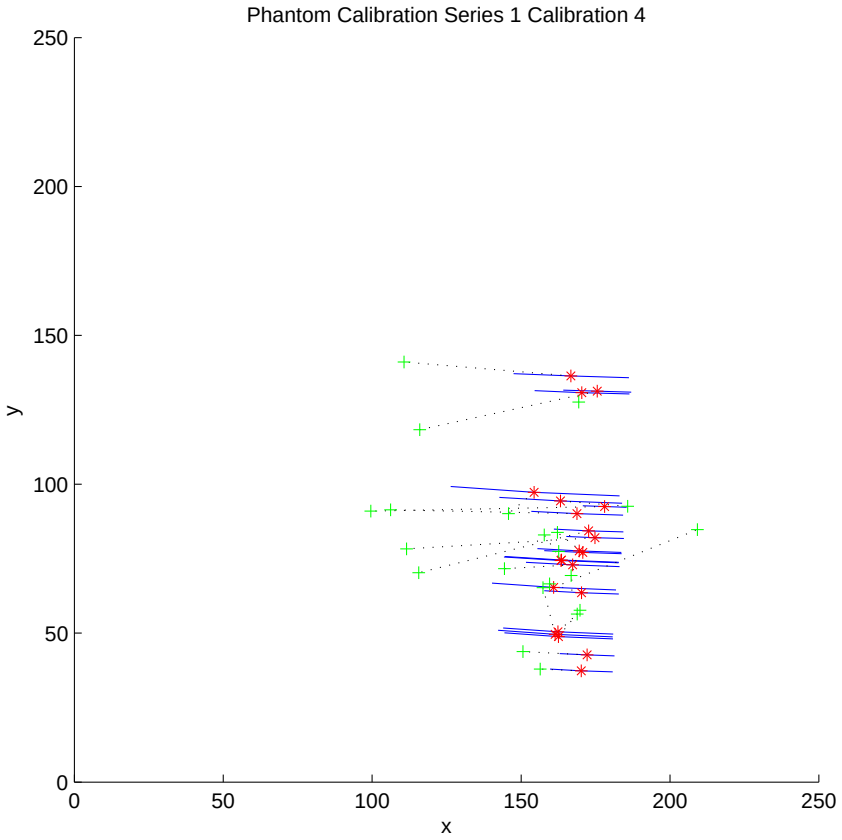

Phantom Calibration Series 1 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

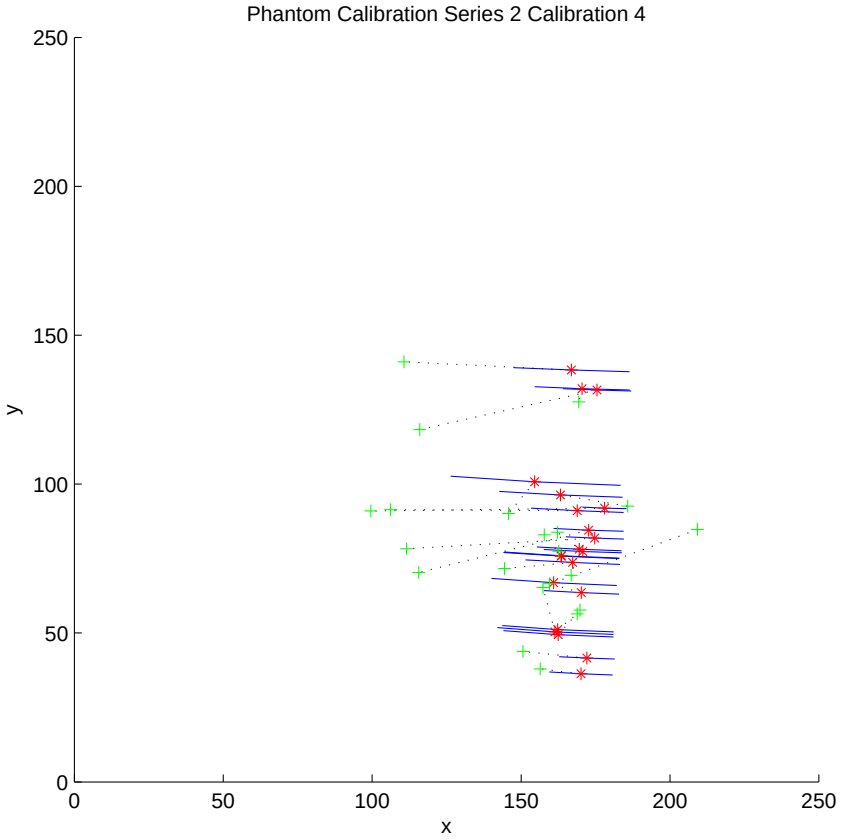

Phantom Calibration Series 2 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

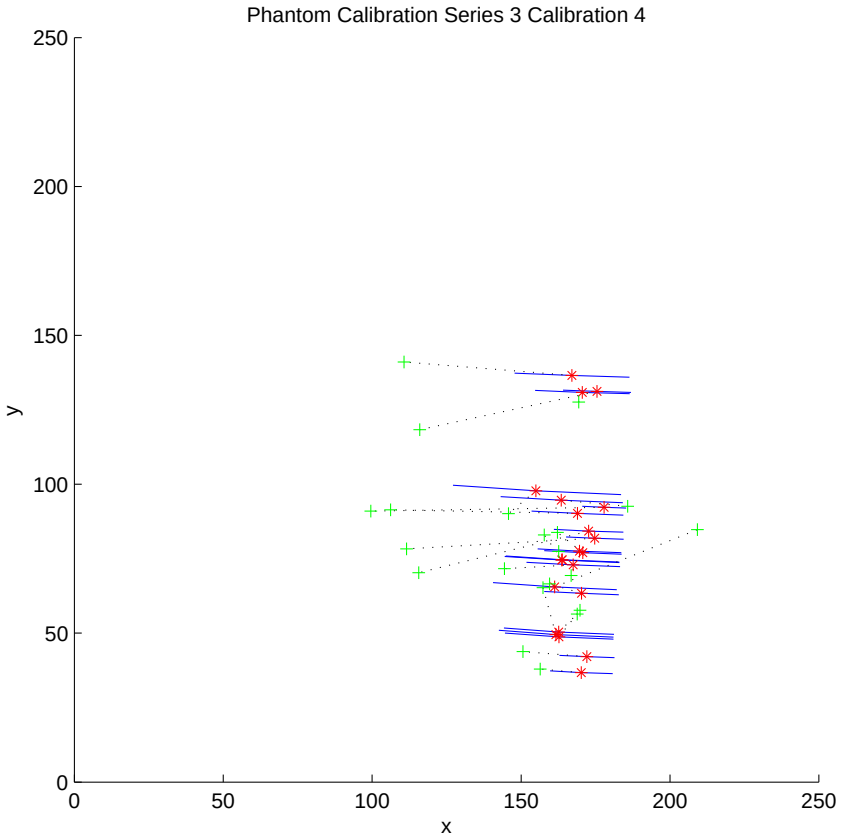

Phantom Calibration Series 3 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

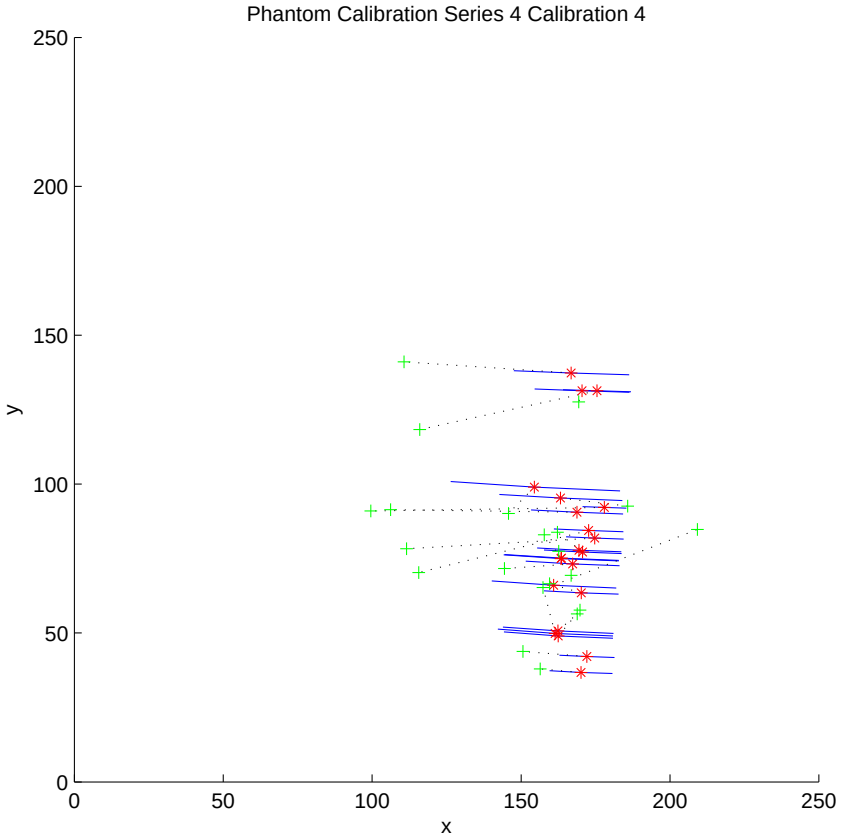

Phantom Calibration Series 4 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

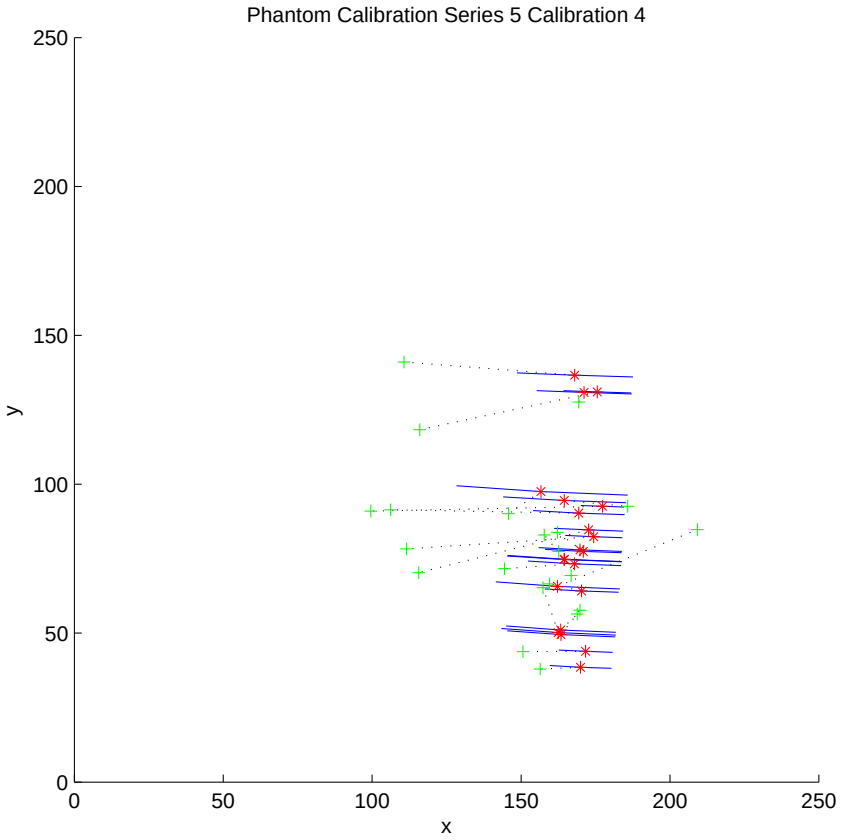

Phantom Calibration Series 5 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

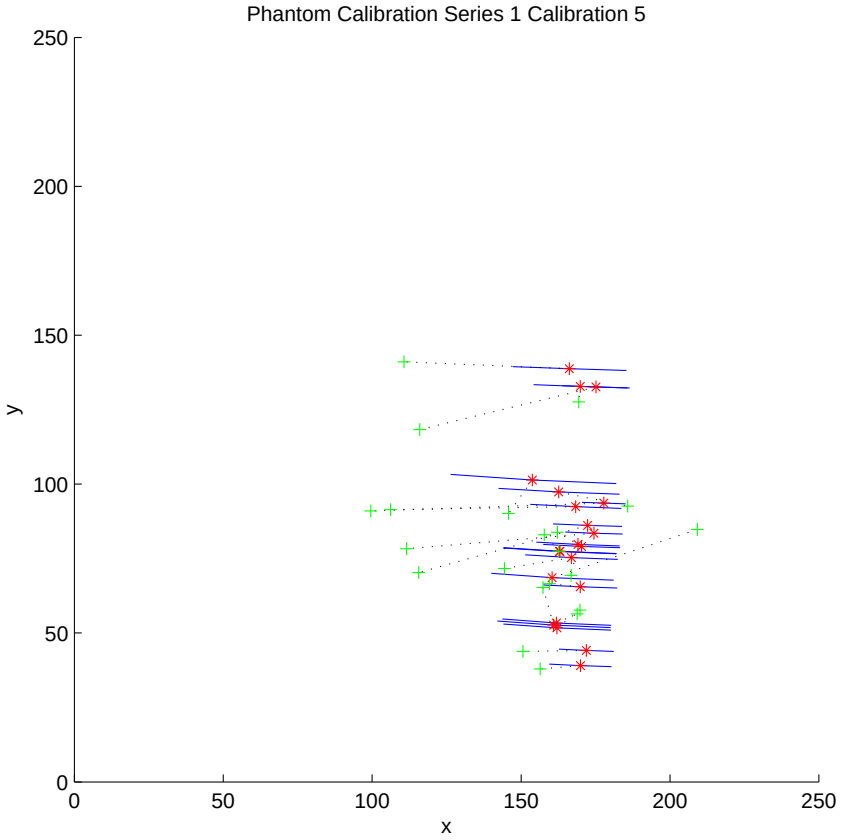

Phantom Calibration Series 1 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x

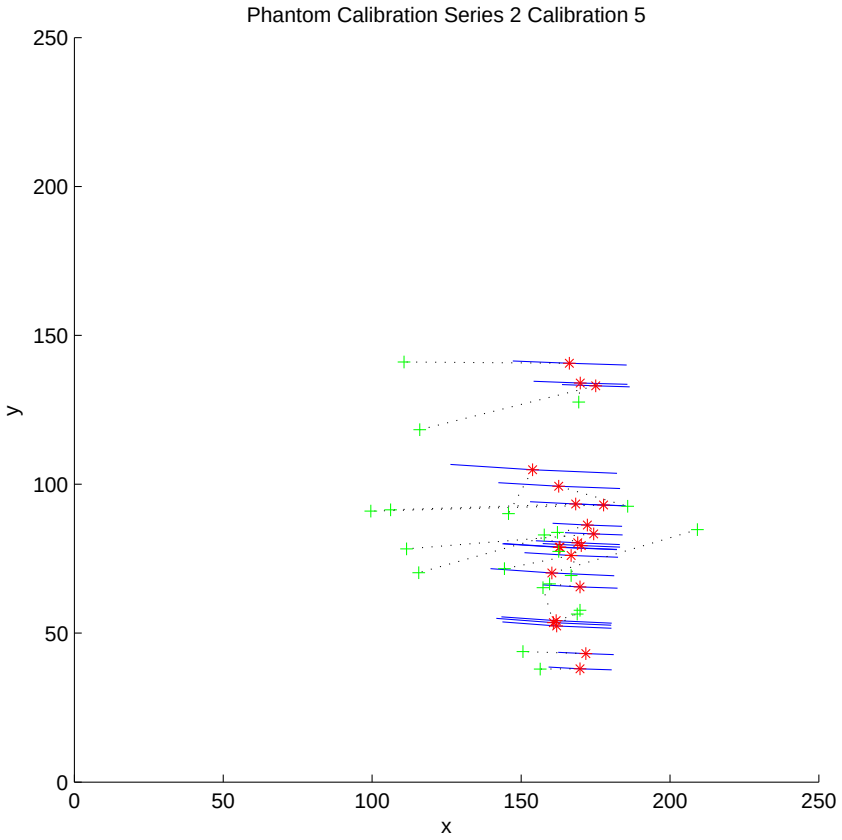

Phantom Calibration Series 2 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x

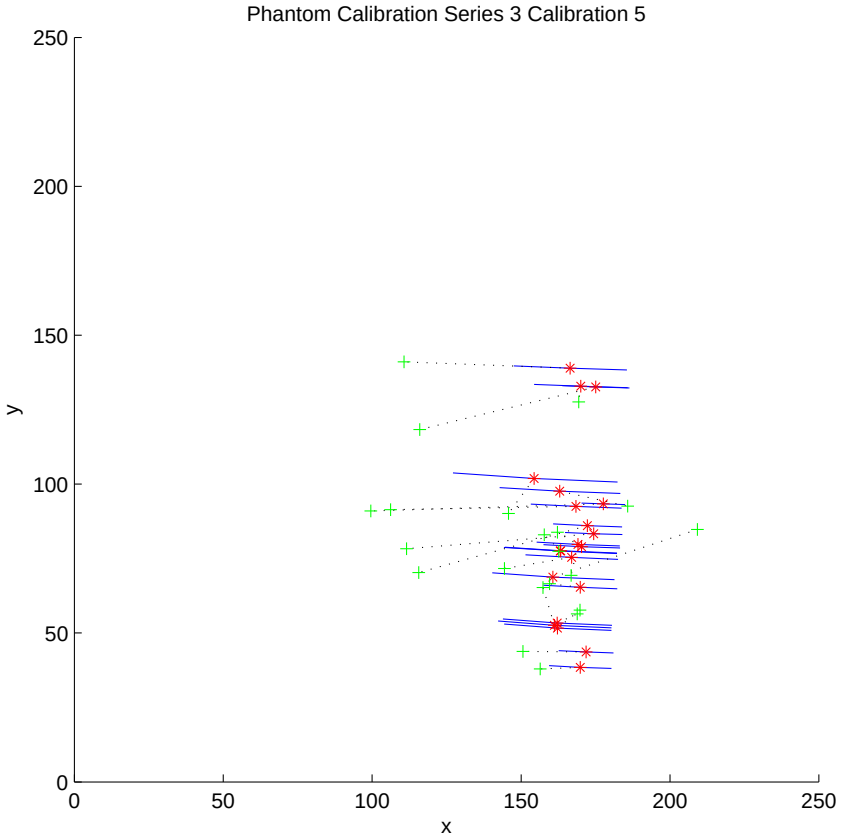

Phantom Calibration Series 3 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x

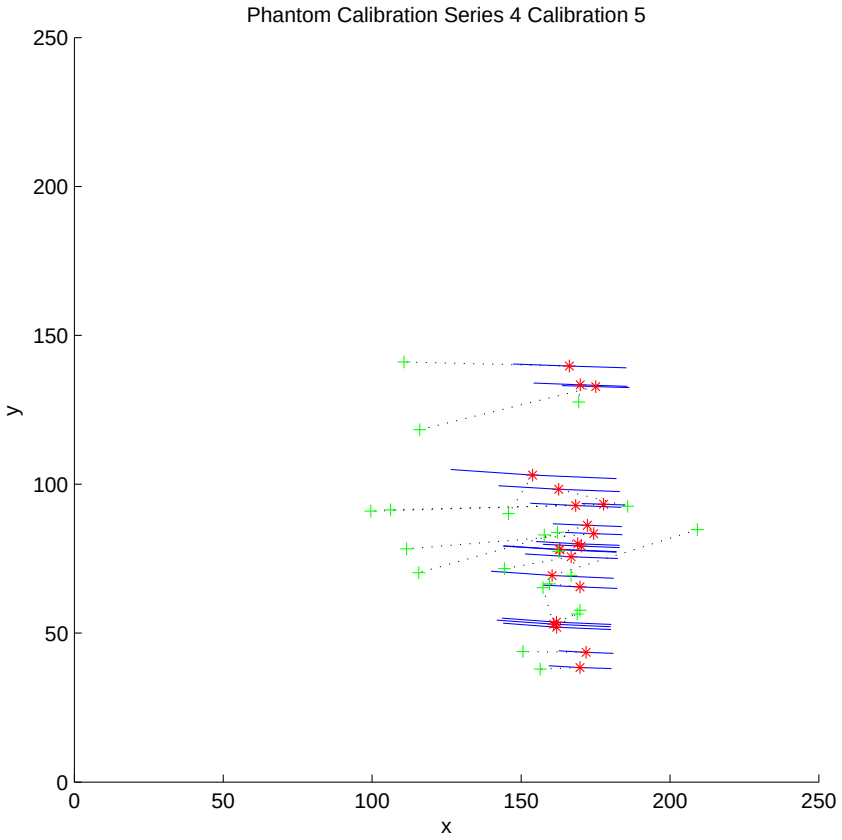

Phantom Calibration Series 4 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x

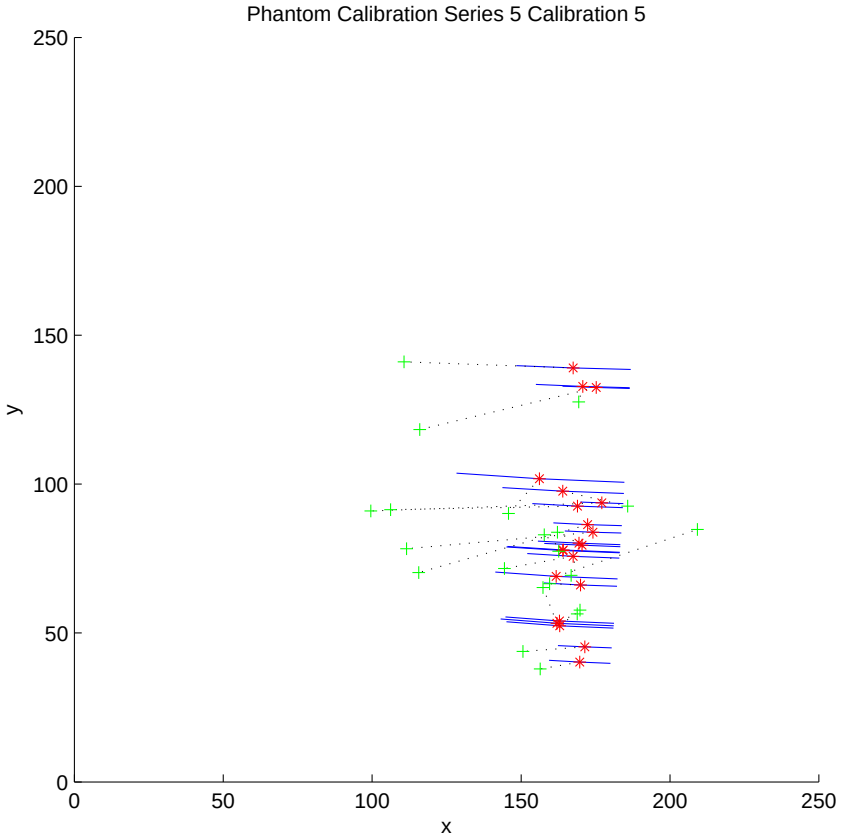

Phantom Calibration Series 5 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x


Phantom Calibration Series 1 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x

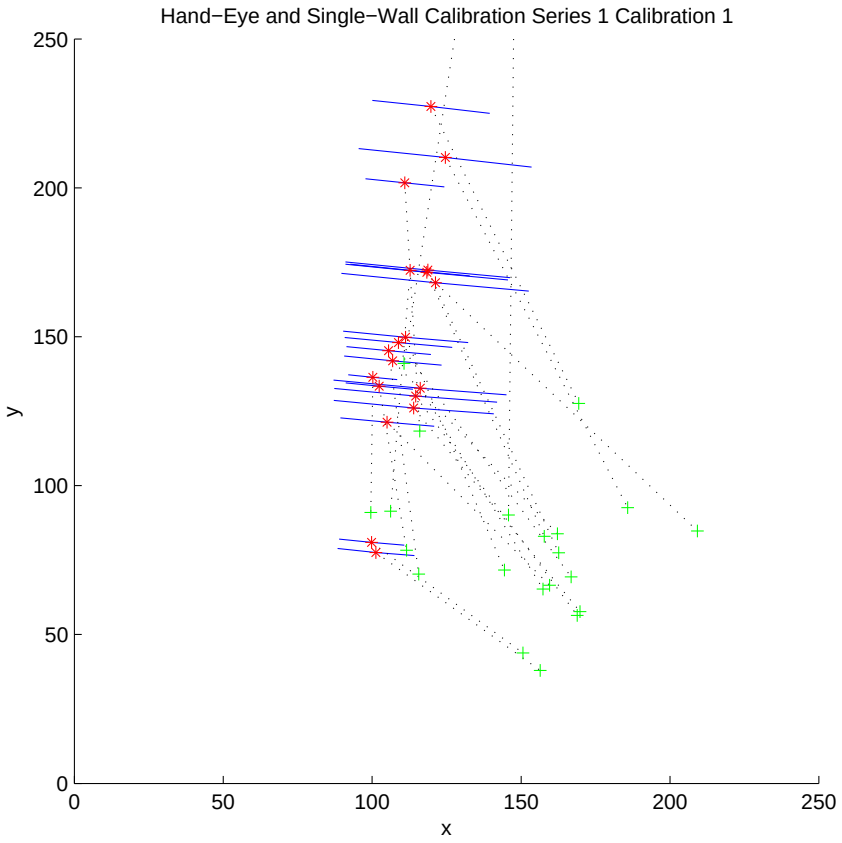

Phantom Calibration Series 2 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x

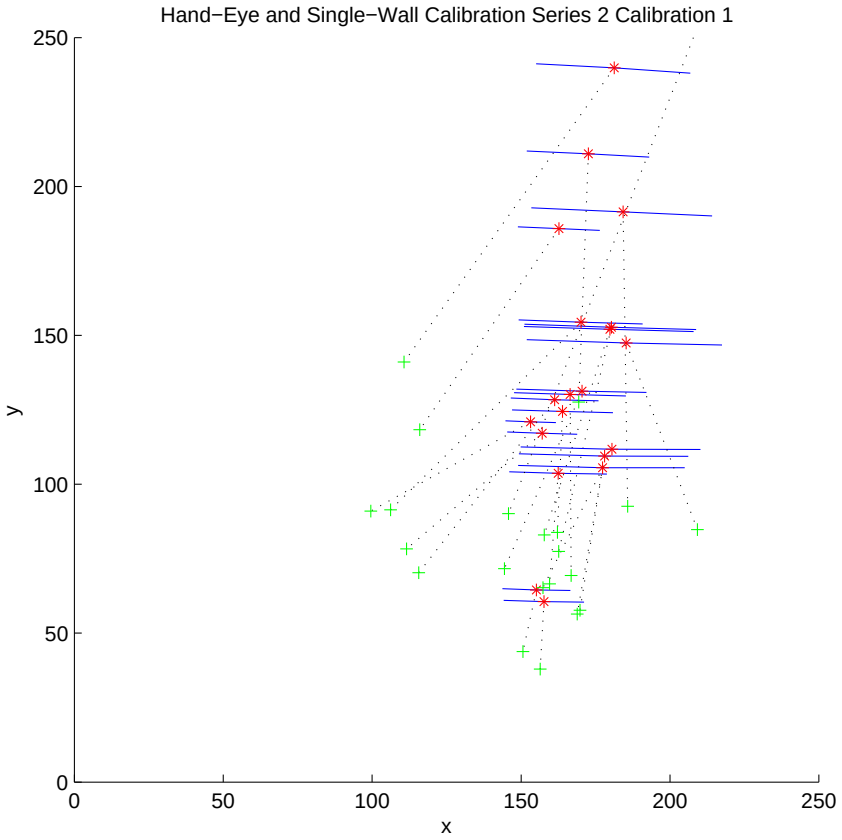

Phantom Calibration Series 3 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x

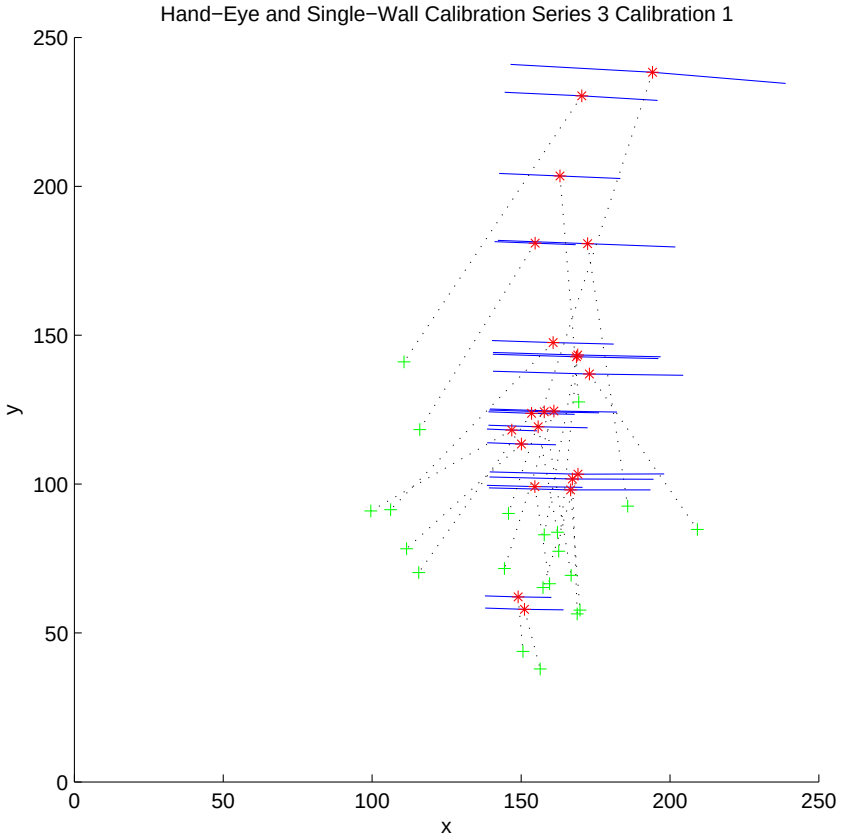

Phantom Calibration Series 4 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x

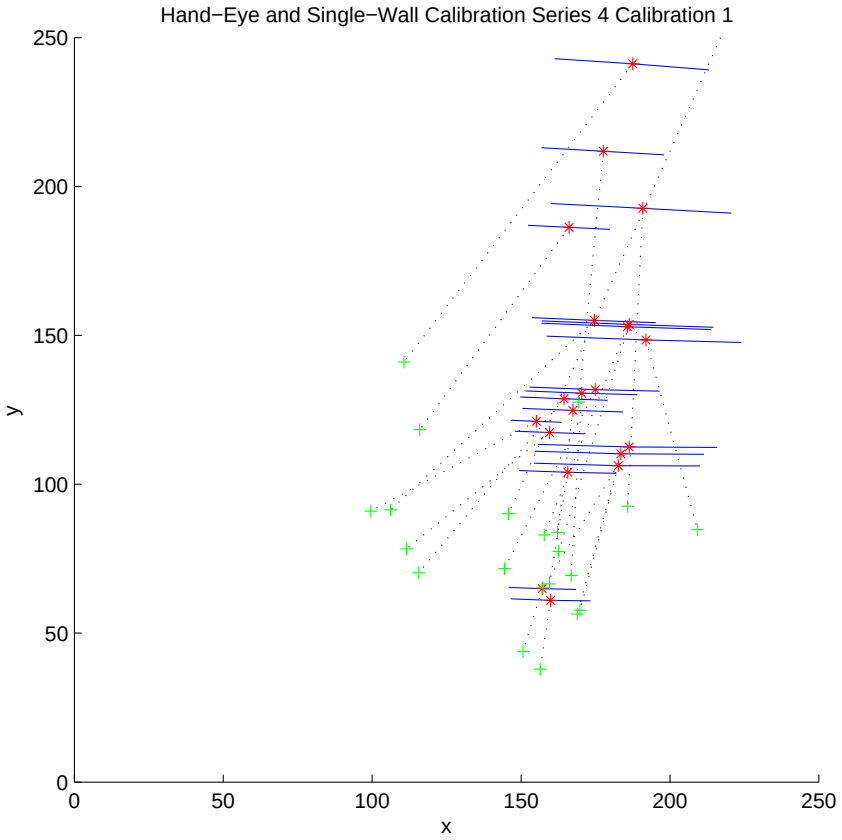

Phantom Calibration Series 5 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x

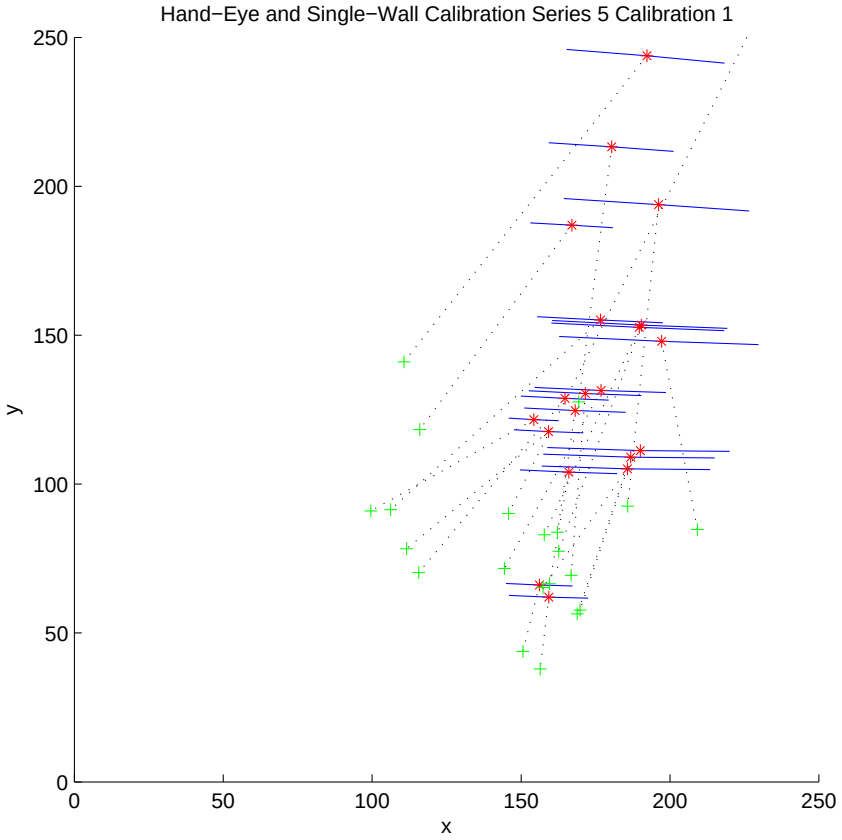
Hand-Eye and Single-Wall Based Calibration Backprojections

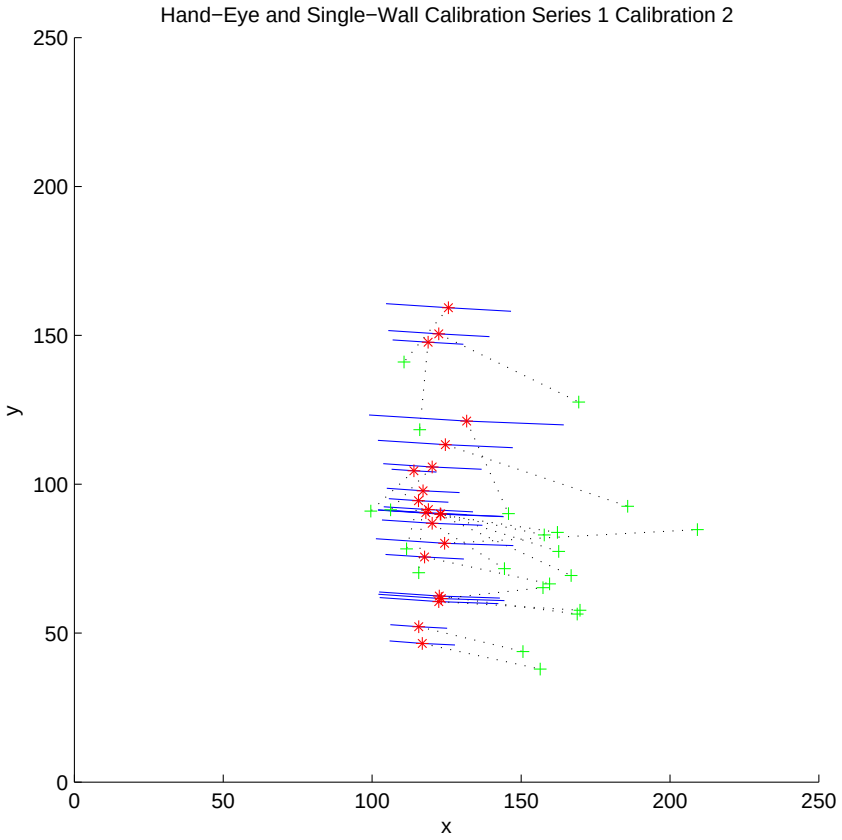


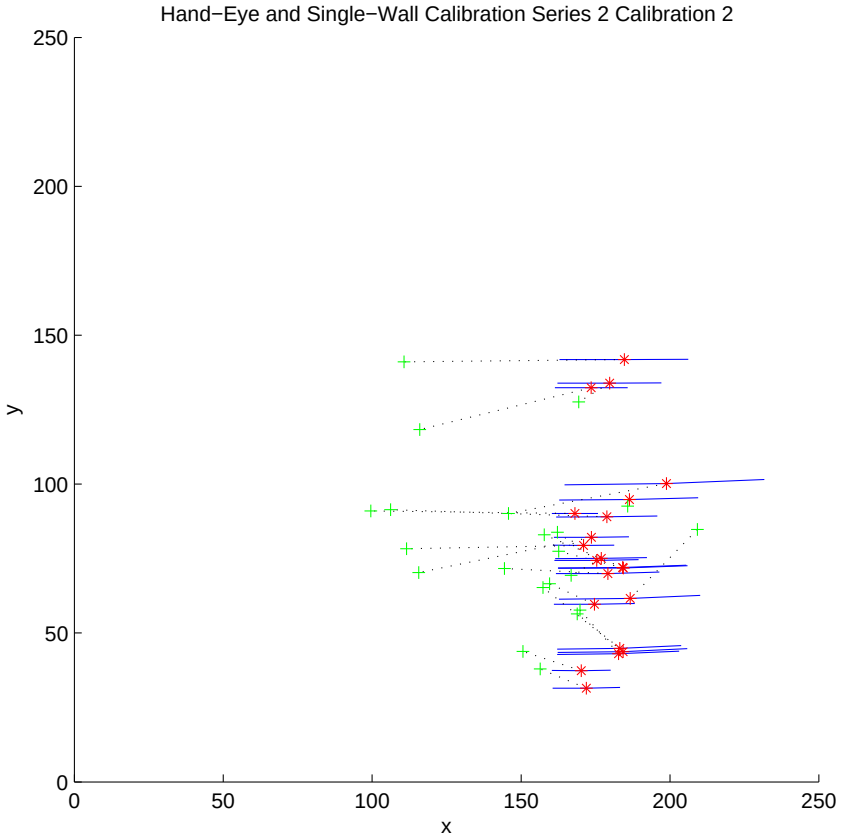

Hand−Eye and Single−Wall Calibration Series 2 Calibration 1
250
200
150
100
50
0
y
0
50
100
150
200
250
x

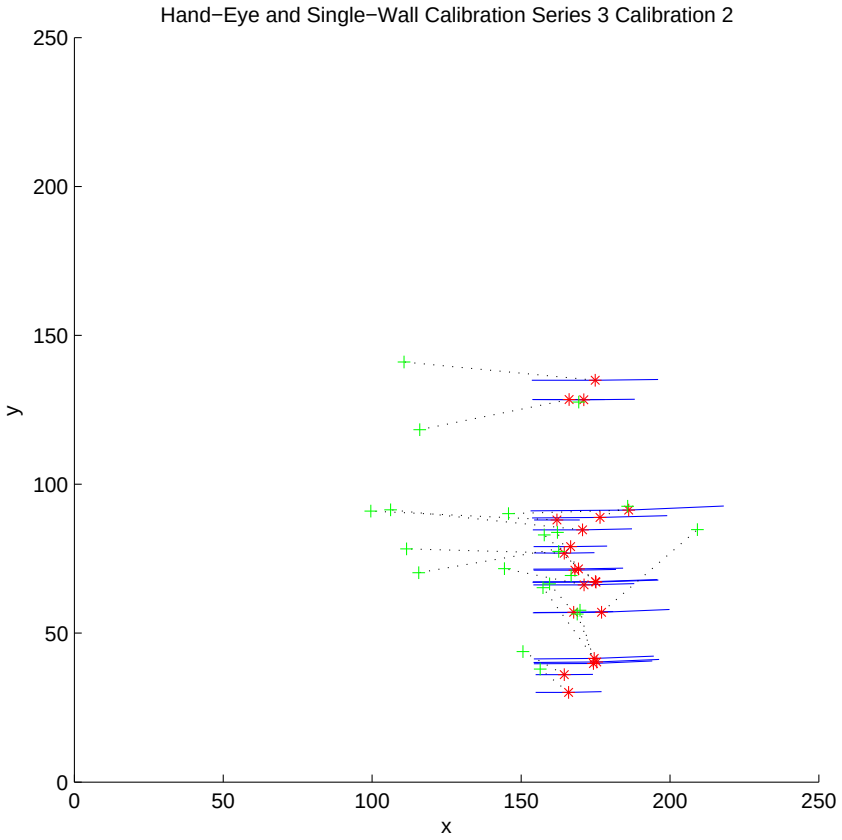

Hand−Eye and Single−Wall Calibration Series 3 Calibration 1
250
200
150
100
50
0
y
0
50
100
150
200
250
x

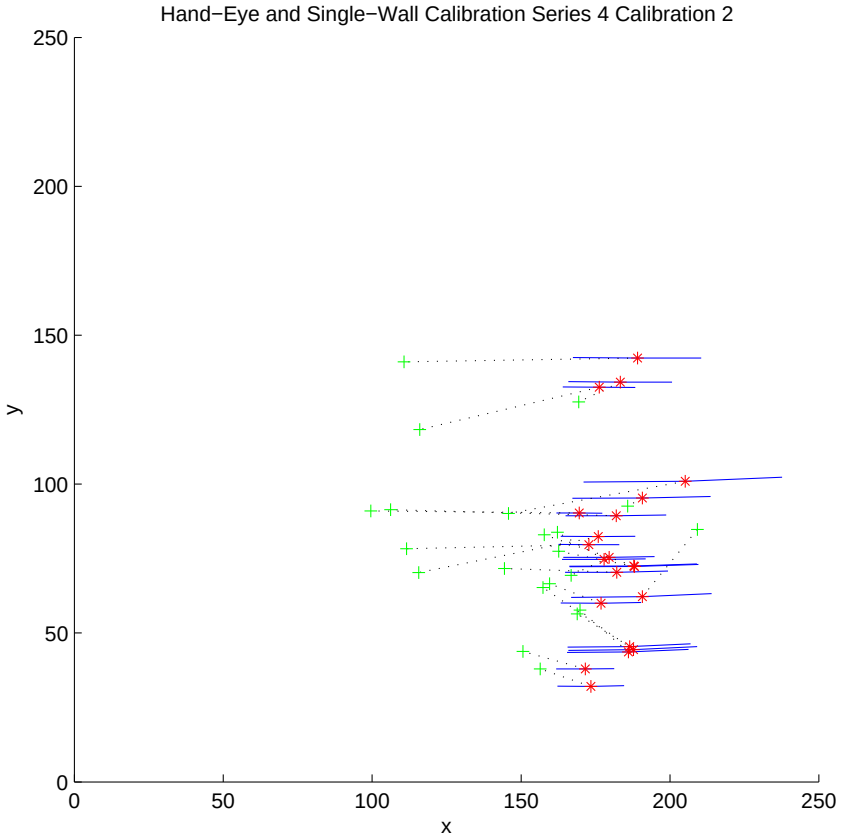

Hand−Eye and Single−Wall Calibration Series 4 Calibration 1
250
200
150
100
50
0
y
0
50
100
150
200
250
x

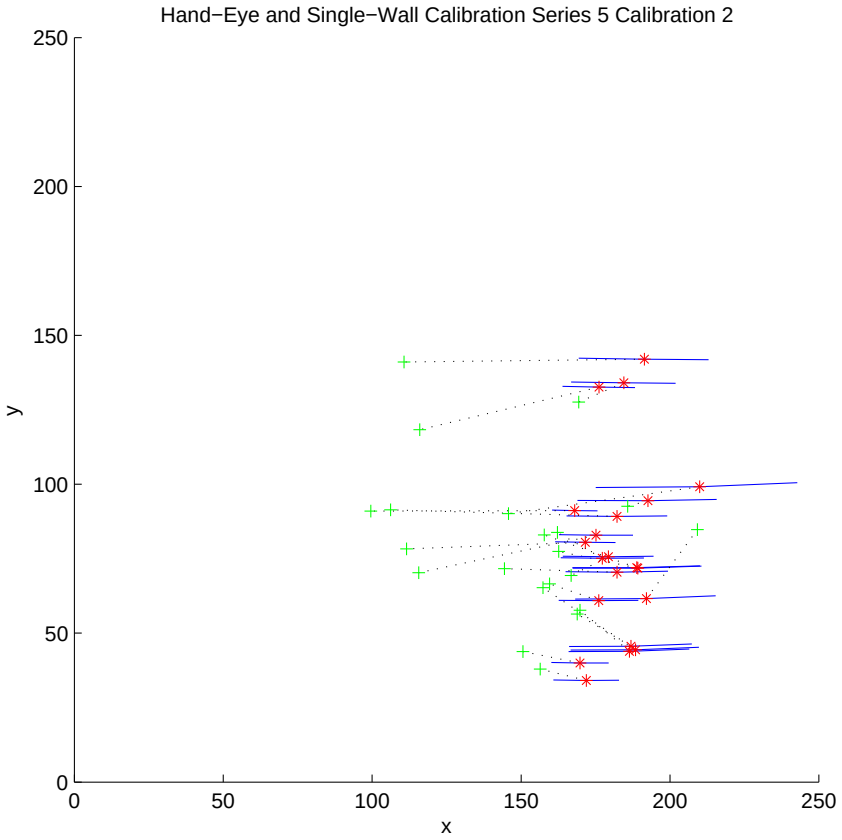

Hand−Eye and Single−Wall Calibration Series 5 Calibration 1
250
200
150
100
50
0
y
0
50
100
150
200
250
x

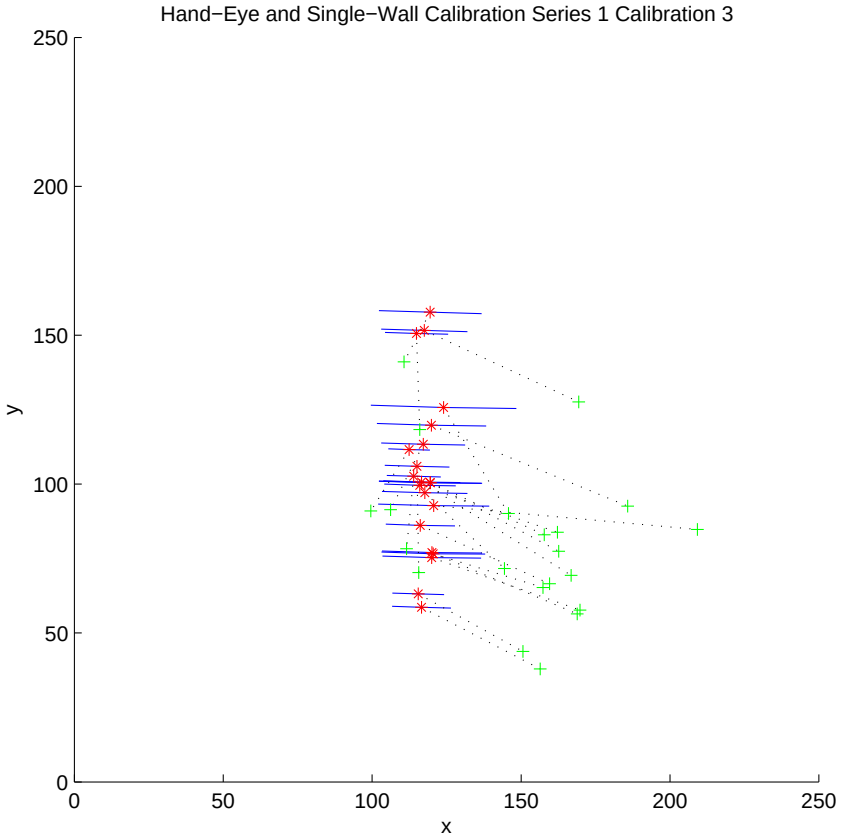

Hand−Eye and Single−Wall Calibration Series 1 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

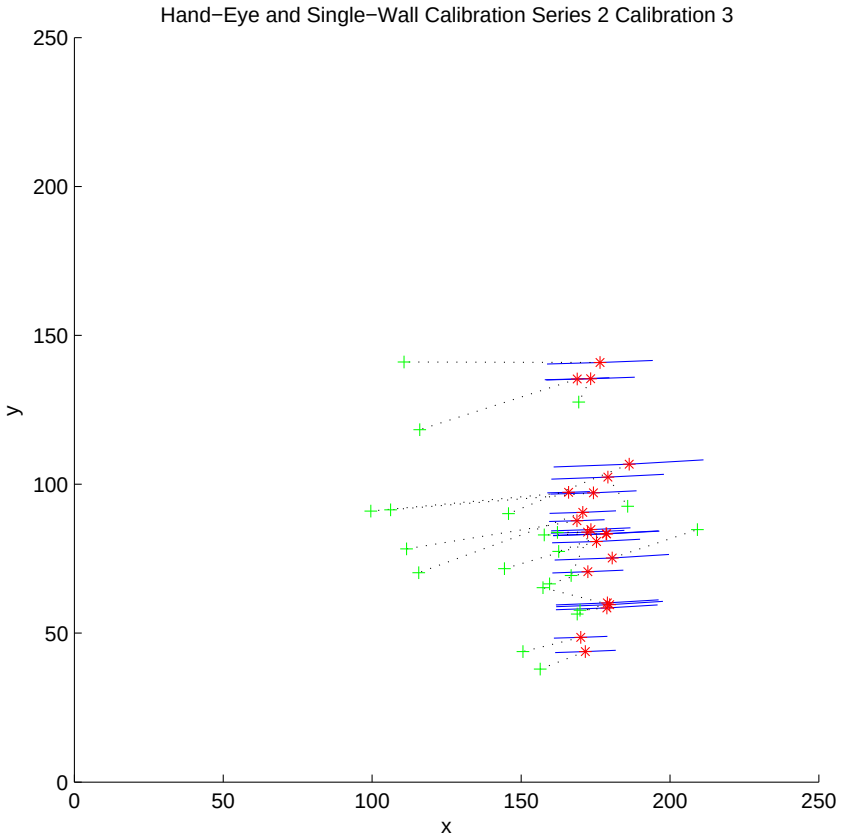

Hand−Eye and Single−Wall Calibration Series 2 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

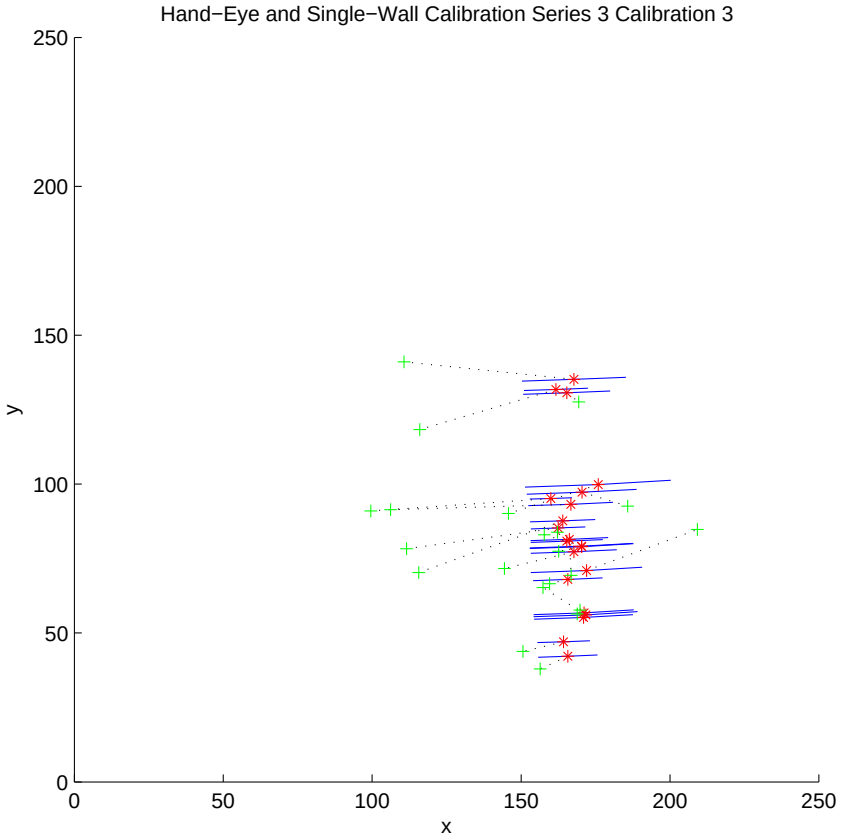

Hand−Eye and Single−Wall Calibration Series 3 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

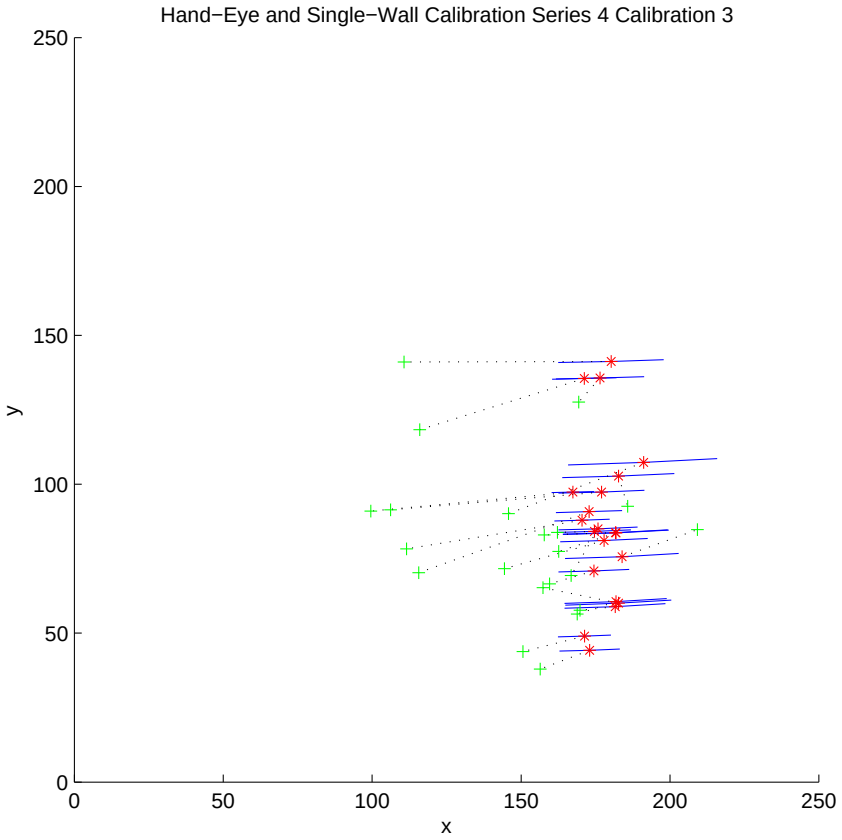

Hand−Eye and Single−Wall Calibration Series 4 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

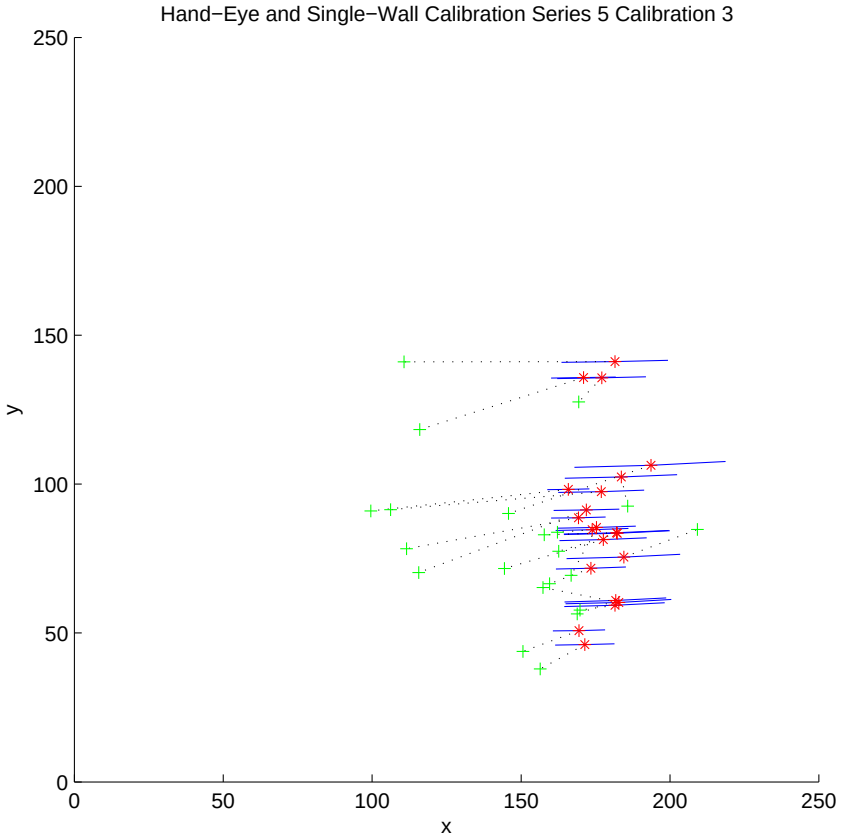

Hand−Eye and Single−Wall Calibration Series 5 Calibration 2
250
200
150
100
50
0
y
0
50
100
150
200
250
x

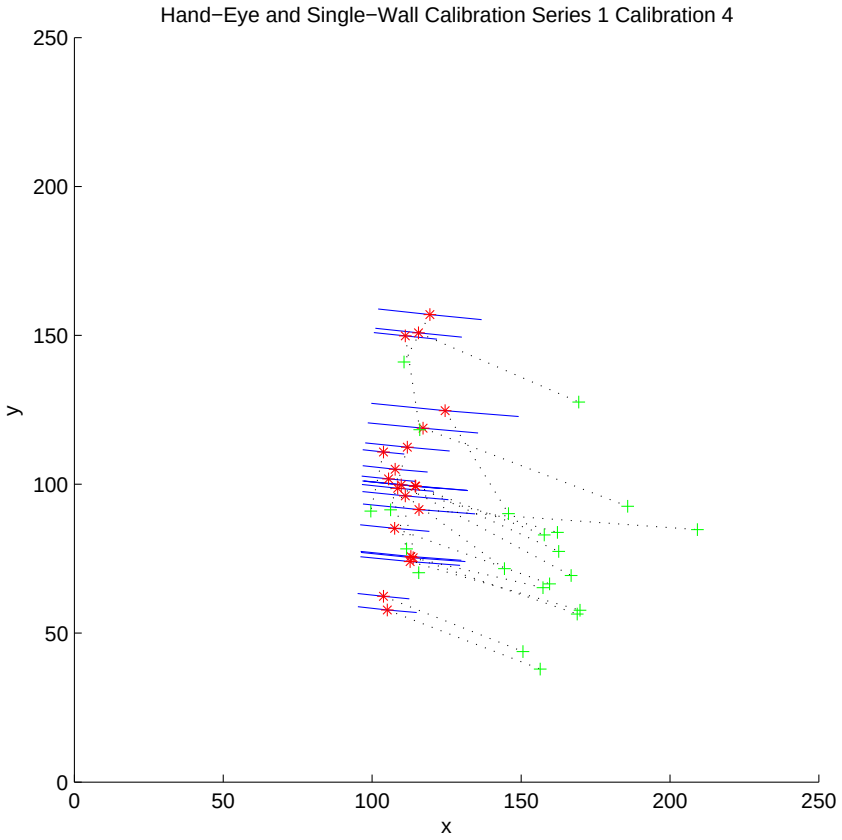

Hand−Eye and Single−Wall Calibration Series 1 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

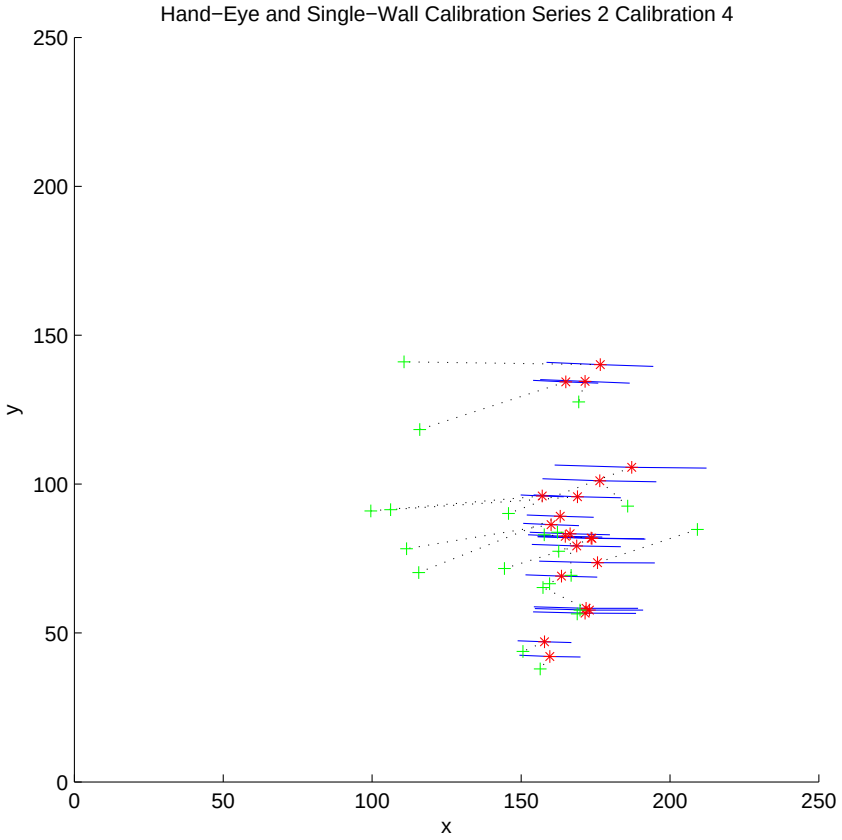

Hand−Eye and Single−Wall Calibration Series 2 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

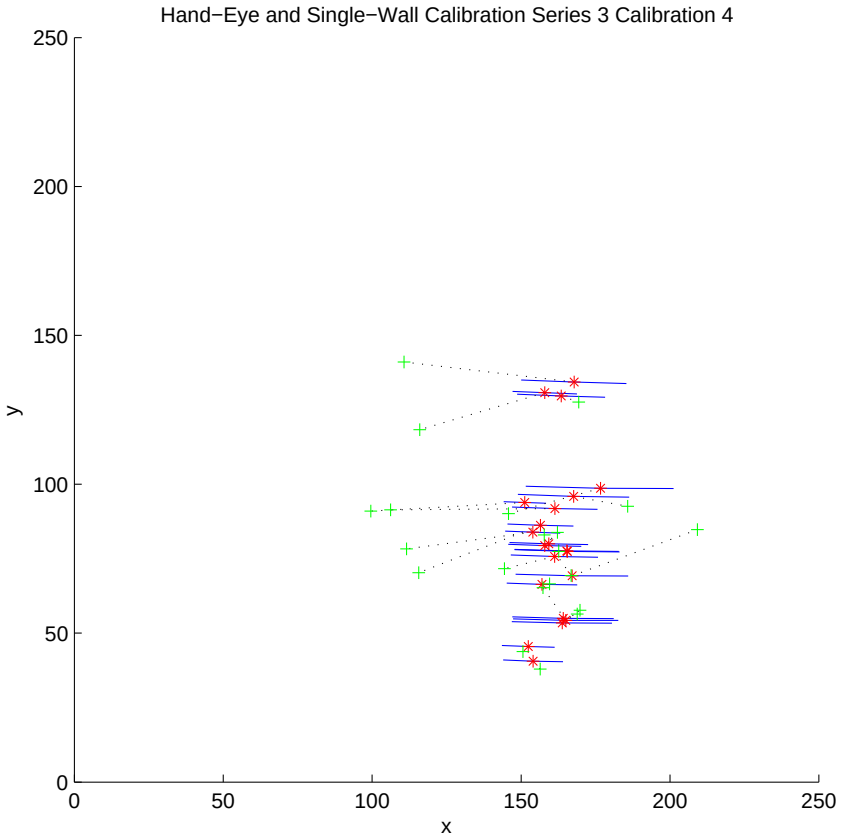

Hand−Eye and Single−Wall Calibration Series 3 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

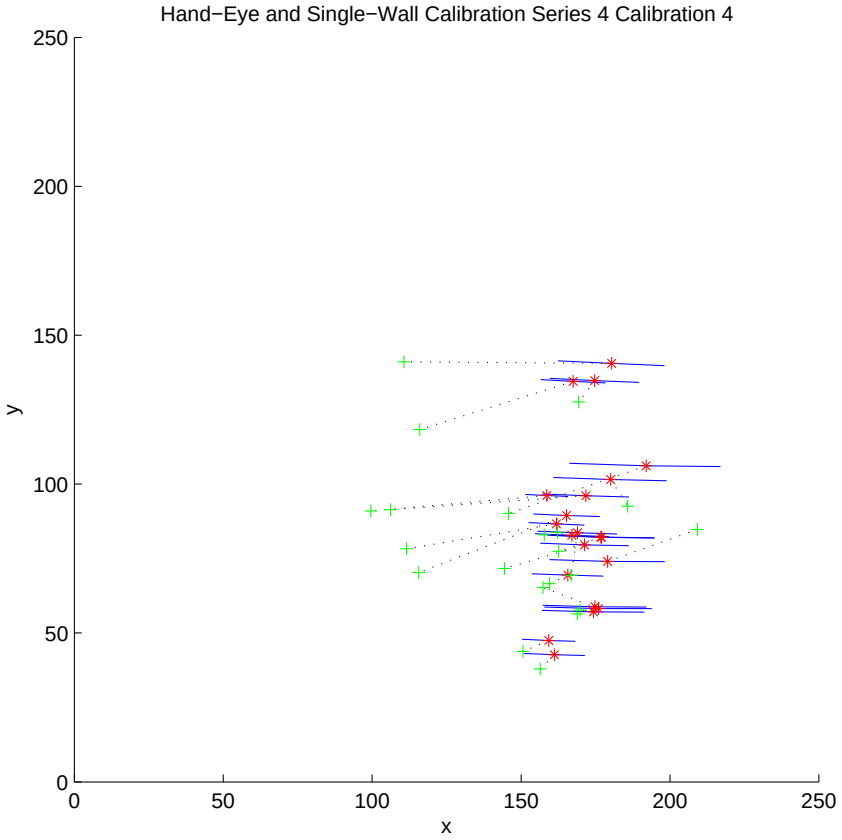

Hand−Eye and Single−Wall Calibration Series 4 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

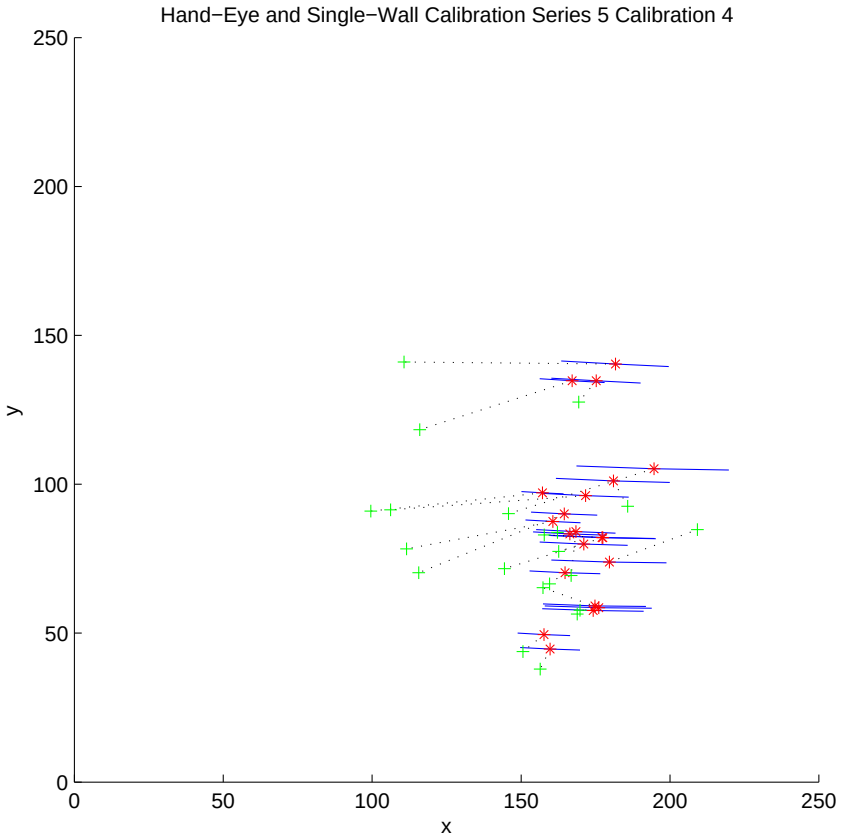

Hand−Eye and Single−Wall Calibration Series 5 Calibration 3
250
200
150
100
50
0
y
0
50
100
150
200
250
x

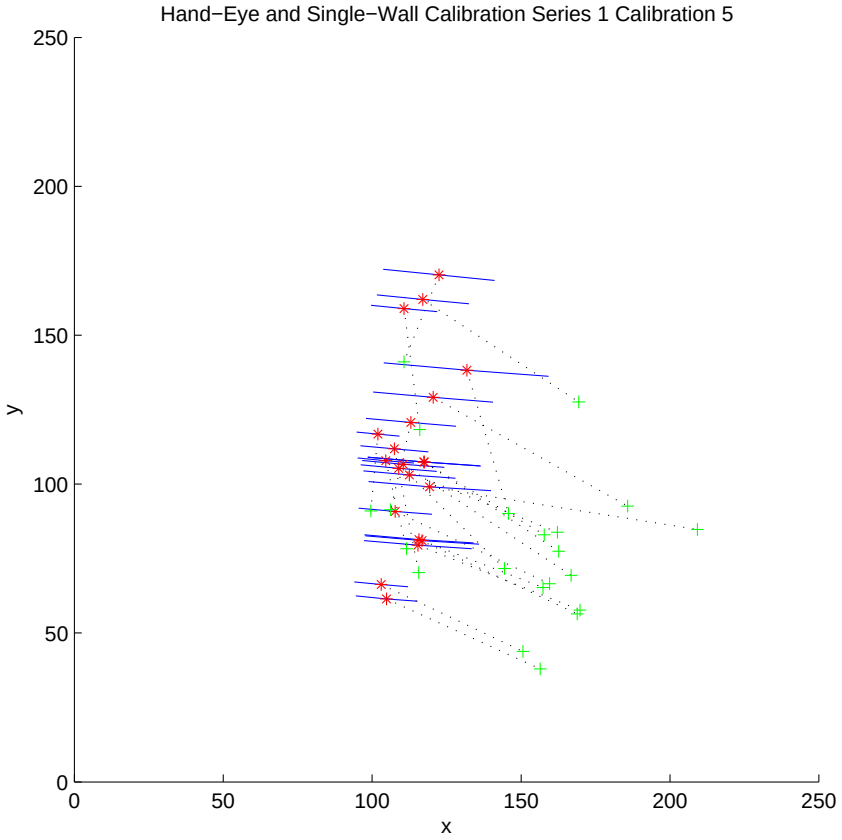

Hand−Eye and Single−Wall Calibration Series 1 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x

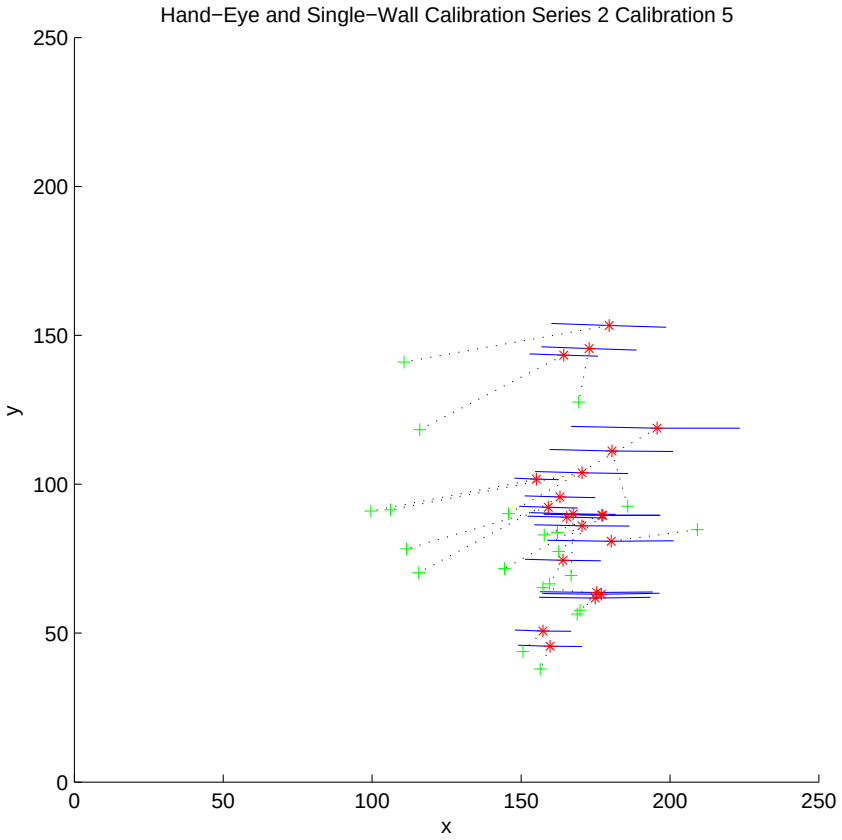

Hand−Eye and Single−Wall Calibration Series 2 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x

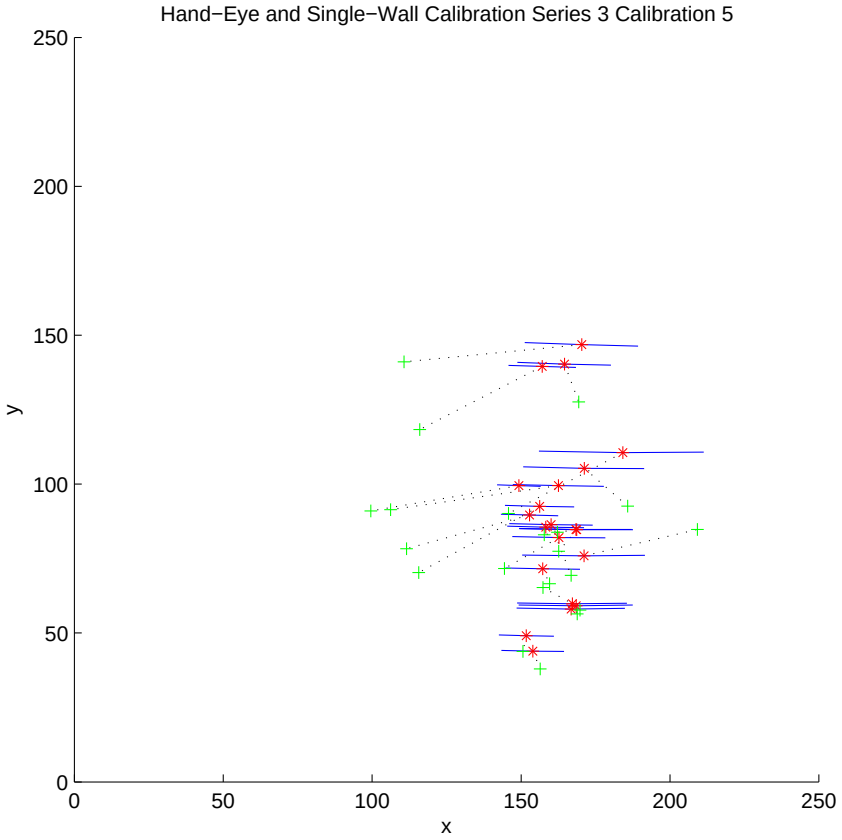

Hand−Eye and Single−Wall Calibration Series 3 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x

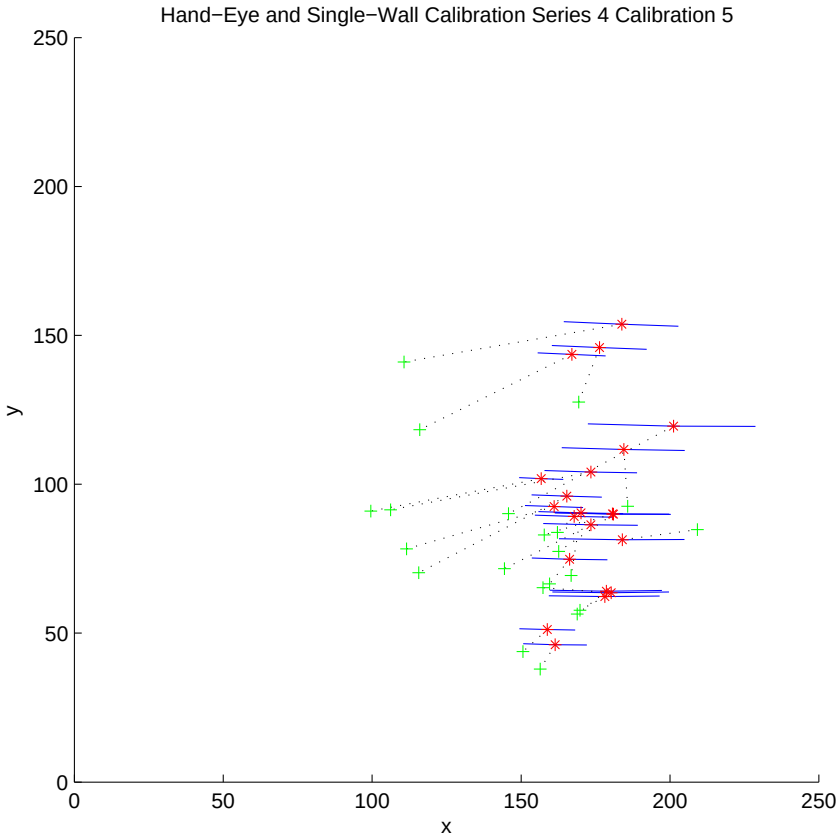

Hand−Eye and Single−Wall Calibration Series 4 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x

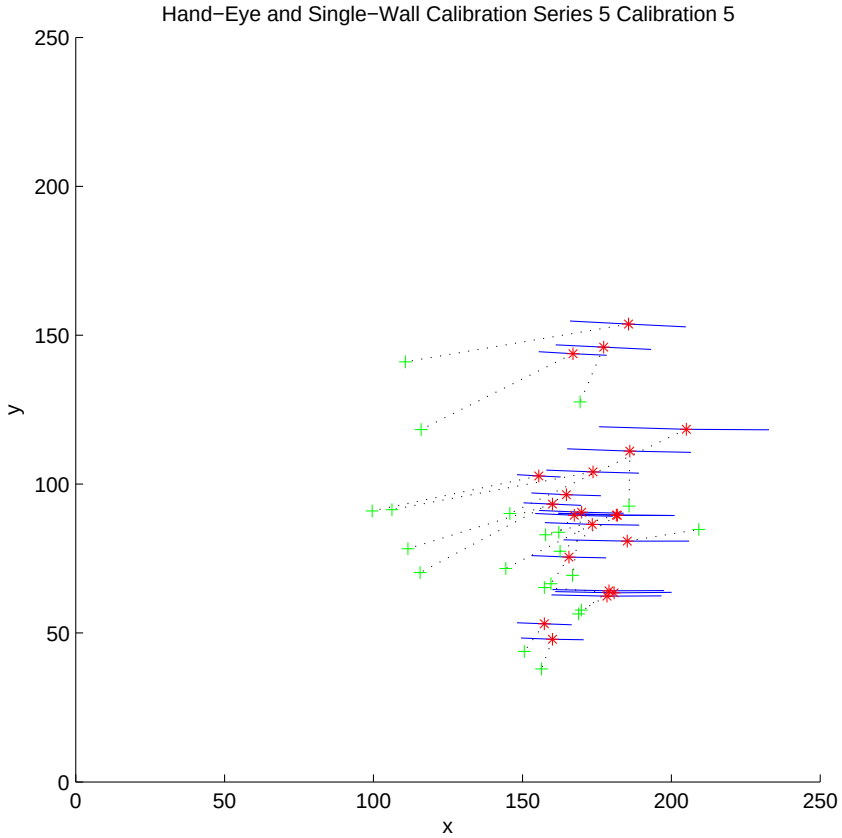

Hand−Eye and Single−Wall Calibration Series 5 Calibration 4
250
200
150
100
50
0
y
0
50
100
150
200
250
x


Hand−Eye and Single−Wall Calibration Series 1 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x


Hand−Eye and Single−Wall Calibration Series 2 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x


Hand−Eye and Single−Wall Calibration Series 3 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x


Hand−Eye and Single−Wall Calibration Series 4 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x


Hand−Eye and Single−Wall Calibration Series 5 Calibration 5
250
200
150
100
50
0
y
0
50
100
150
200
250
x