Thesis report for obtaining the degree of Master of Science

Camera calibration for oblique viewing laparoscopes

Sjirk Gerard Snaauw

August 2017

Graduation committee

Chairman	Dr. Ir. F. van der Heijden
Medical supervisor	Prof. Dr. T.J.M. Ruers
Technical supervisor	Dr. Ir. F. van der Heijden
Process supervisor	Drs. P.A. van Katwijk
External member	Prof. Dr. A.A. Stoorvogel
Additional member	Dr. J.A. Nijkamp

Preface

The thesis report before you is the result of a year of hard work at the Netherlands Cancer Institute – Antoni van Leeuwenhoek, and of the many years of studying before that. I am very grateful to all of the people that supported me all of these years and helped me make the most of myself.

I would like to thank Jasper Nijkamp for providing me with the opportunity to get this first-hand experience in surgical navigation and camera calibration, and for his help during the past year. Together with Roeland Eppenga, they provided me with the day-to-day discussions that led to the insight needed to complete this research.

Annelies Loving has provided me with the support I needed at moments where I could not see how I could make this year a success. Her help showed me how I can improve myself in ways other ways than just research. I think it is very unfortunate that she cannot be a part of the end of this journey, but I am grateful to Paul van Katwijk for taking her place in my graduation committee.

Ferdi van der Heijden introduced me to most of my current research interests. It is his help and guidance that got me excited for these topics. There has not been a single meeting that he did not provide me with new insights and interesting questions that resulted in a better understanding and progression in my work. If only all supervisors would take as much time for their students as Ferdi.

I would also like to thank Anton Stoorvogel for taking the time and effort to take part in my graduation committee.

Of course, my friends and collogues cannot be omitted as they provided me with the highly appreciated distractions needed to keep me sane during the past year.

And last but not least, I would especially like to thank my family that always supports my no matter what I do. It is this support that provides me with the peace of mind and courage that are needed to take on any challenge I encounter. No words can ever describe my gratitude to them.

Summary

Chapter 1: Introduction
This chapter contains the clinical background of rectal cancer and its surgical options. Current efforts in improving surgical outcome are focused on surgical navigation. In the current implementation, surgical navigation can only be applied in an open surgical setting. As 85% of rectal cancer procedures are performed laparoscopically, the majority of patients cannot benefit from these developments. The rest of the chapter contains the technical background on used techniques, and shows how augmented reality can extend the use of surgical navigation to laparoscopic procedures. The chapter ends with the defined goals for this research in a first step towards application of augmented reality.

Chapter 2: Camera calibration
Intrinsic camera parameters are evaluated in order to verify current assumptions in literature, and to define a camera model for the used laparoscope. In addition to evaluation of current literature, we show that addition of decentering distortion to the camera model improves results.

Chapter 3: Hand-eye calibration
The position of the camera is related to the two optically tracked sensors attached the laparoscope. Behavior of the relation during rotation of the laparoscope is evaluated in order to define which reference sensor can best be used to model the position of the camera's pose. We show that the sensor attached to the cylinder of the laparoscope can provide the best results as it requires a simpler model, and, by being closer to the camera, it produces a lower tracking error.

Chapter 4: Delay estimation
A delay estimation procedure is developed to estimate the acquisition delay between the laparoscopic images and the optical tracking system. The procedure is based on phase differences in fitted sinusoidal patterns obtained by both systems. The pattern is generated by rotation of an object at constant angular frequency that can be tracked by both systems.

Chapter 5: Laparoscope calibration
The camera and hand-eye model are combined to evaluate the accuracy in a static environment. We show that the combined model produces accurate results on calibration data, but is not able to reproduces the results on validation data. The increase in error is determined to be caused by freedom of motion of the camera's image sensor within the laparoscope. As the image sensor cannot be tracked externally, it seems that generation of a calibration method for this specific laparoscope is not feasible.

Chapter 6: General discussion and conclusions
Obtained results are summarized to answer the research questions, and are compared to other literature. Finally, several recommendations are given to answer several questions that were raised by the obtained results.

Table of contents

CH1: Introduction

1.1 Clinical background

Since 1990 the incidence of colorectal cancer (CRC) in the Netherlands has more than doubled to over 15,500 new cases in 2015. Rectal cancer contributes 4,500 of these cases and went through a similar increase in incidence [1]. It is expected that the phased introduction of the colorectal screening program, started in 2013, will, at least temporarily, further increase the incidence of CRC in the Netherlands, especially that of early stage disease [2-4].

Treatment options of rectal cancer consist of surgery, radiotherapy, and systemic therapy, with surgery being the principal treatment leading to cure [5-11]. There are two main techniques available for rectal cancer surgery. The first technique is low anterior resection (LAR) consisting of excision of the proximal part of the rectum containing the tumor, followed by an anastomosis of the sigmoid to the remainder of the rectum. The second technique is abdominoperineal resection (APR), which is an extension of LAR for tumors in the distal rectum, where the anus is also removed and a permanent colostomy is made. Since the colostomy in APR results in a worse quality of life, LAR is the preferred treatment for tumors in the proximal two-thirds of the rectum with an overall share of 71% in 2015 [4]. A major improvement in surgery of rectal cancer came with the introduction of the total mesorectal excision (TME) in the early 90s. Instead of just removing the rectum, surgeons follow the mesorectum, which is the mesentery surrounding the rectum containing the rectum, blood vessels, fat, and pararectal lymph nodes. With TME, recurrence rates decreased from 25% to 5% and it has therefore become the gold standard in both LAR and APR [12, 13]. The subsequent step in improving rectal cancer surgery came with the introduction of laparoscopy in the early 00s. With laparoscopic surgery, patient have reduced blood loss, shorter hospital stay, and earlier return of bowel function [14]. The percentage of rectal cancer interventions performed laparoscopically has increased from 35% in 2009 to almost 85% in 2015 [4].

One of the ways success of surgery is measured is the circumferential resection margin (CRM). A positive CRM is defined as having a resection margin < 1 mm to the edge of the tumor and is associated with a hazard ration (HR) of 1.7 for reduced overall survival and a HR of 2.8 for developing distant metastasis when compared to a negative CRM. The HR for local recurrence in positive CRM ranges from 6.3 for treatment without, and 2.0 for treatment with neoadjuvant therapy compared to negative resection margins [15]. In 2015 a positive CRM was found in 5.0% of the patients undergoing rectal cancer surgery of which 4.5% had an irradical resection, meaning that there is no resection margin [4]. High positive CRM rates occur in subgroups such as patients with locally advanced disease, mainly with low rectal tumors (22%), and in patients who are operated with an APR (32%) [16-18]. Of all patients that underwent rectal cancer surgery 13% needed re-intervention within 30 days of the primary surgery due to complications [4]. Several studies have shown that local recurrence could have been lower if tumor resection is improved by, for example, excising wider around low rectal tumors [12, 19-21].

Vital structures surrounding the rectum can be damaged during surgery, leading to postoperative morbidity such as bladder and sexual dysfunction. After TME surgery, functional urinary problems arise in 24-32% of the patients due to visceral sacral nerve damage [22]. As a consequence, patients may experience voiding dysfunction, overflow incontinence, frequent lower urinary tract infections, and loss of bladder filling sensation. In addition, damage to the parasympatic nerve fibers can lead to disturbances in sexual function. Up to 30% of women and 45% of men experience sexual dysfunction after rectal surgery [22]. A nerve-sparing surgical approach is applied to minimize damage to the pelvic nerves [23]. However, this technique is difficult to perform due to the complex anatomy of the various neural branches.

In summary, surgery of rectal cancer is a challenging field where a balance needs to be found between reducing positive resection margins and preventing morbidity. Although major improvements have been achieved over the past 3 decades, subgroups of patients can still be identified which can profit from further technical developments.

1.2 Technical background

Imaging modalities such as MRI and CT are a valuable source of information for physicians. These images are used for diagnosis, treatment planning, surgical planning, and evaluation of the treatment. It is therefore surprising that these images are hardly used during surgery. Instead, surgeons rely on mental notes of the images to guide them through the procedure, or look at static images on a screen in the operating room. Image guided surgery (IGS), or surgical navigation (SN), is a technique that aims to provide surgeons access to the information contained in the images by displaying them in a dynamic way visibly to the surgeon during surgery. However, current implementations of SN are limited to open surgery while over 85% of rectal cancer surgeries are performed laparoscopically. SN needs to be expanded to the domain of laparoscopic surgery to allow surgeons to incorporate the imaging information in surgical decision making.

1.2.1 Surgical Navigation

Computer-assisted surgery (CAS) is a broad concept that describes the use of technology to create patient specific models for surgical simulation, surgical planning, and the use of these models during intervention. It is this interventional part of CAS what SN refers to. In SN, the patient’s anatomy is related to the pre-operative imaging by means of tracked sensors attached to the patient. Generally these sensors are placed on rigid anatomical landmarks, such as osseous structures, to minimize movement of the sensors during the intervention. Location of these sensors is then defined in the pre-operative imaging, followed by a registration procedure to match the pre-operative imaging to the patient’s position on the surgical table [24].

Tracked navigational tools are used to navigate in and around the registered patient. The tracked tool can be used to display orthogonal views of the patient’s CT or MRI scan at the current location of the tool, thereby providing the surgeon interactive access to the images. Pre-operative imaging can also be used to generate a 3D model of the patient containing the anatomy relevant to the procedure such as, tumor, critical structures surrounding the tumor, and osseous reference structures. In this case, the navigational tool and its motion are displayed within the 3D model of the patent, allowing localization of the tumor and critical structures with respect to the tool. The model does not only display what can be seen from the surface, but also relevant structures not directly visible to the surgeon. Visualization of these otherwise invisible structures helps the surgeon to navigate towards or around these structures and can thereby improve surgical outcome, or even enable surgeries that would otherwise not have been possible [24, 25].

Implementation of SN has mainly been focused around ear, nose, throat (ENT); neuro-; and orthopedic surgery, but is not limited to these areas [24]. SN systems can only match the patient’s model to the patient’s anatomy on the table if the position of relevant structures can accurately be described in relation to the sensors. In areas of deformable anatomy this requires a complex model, but when the position of relevant structures is more or less rigid in relation to the landmark a much simpler model can be used. In ENT, neuro-, and orthopedic surgery the relation between surrounding osseous structures and points of interest can often be described as rigid, explaining the interest and success of SN in these areas.

1.2.2 Tracking systems

Tracking systems used for SN can generally be divided in two categories, electromagnetic (EM) systems, and optical tracking systems (OTS). Optical systems consist of a stereoscopic camera, and infra-red reflecting spheres. Several of these spheres are attached to a rigid body to allow definition of a coordinate system with respect to the fixed geometry of the spheres. Infra-red light emitted by the camera system is reflected back by the spheres and detected by both image sensors. The stereoscopic view by the double camera system allows pose estimation of the rigid body containing the spheres with respect to the camera. Compared to EM systems, OTS is generally deemed more accurate and it has a larger working volume. However, it does require a direct line-of-sight between both cameras and the reflecting spheres. Tracking is lost as soon as the view of one of the cameras on one of the spheres is blocked. Higher accuracy and larger working volume make OTS ideal for tracking of objects outside of the patient, such as laparoscopes, where larger movements of these objects can be expected. Rigid body size and the direct line-of-sight requirement, however, make the system unfit for in-vivo use and tracking of the patient.

EM systems consist of a field generator and EM sensors. Coils in the field generator create an EM field that is detected by the sensors. Controlled variation of the EM field induces a signal in the sensor from which the position and orientation of the sensor is estimated with respect to the field generator. Only a small field can reliably be produced around the field generator, limiting the working volume of EM systems. A major limitation of EM systems is that magnetic fields are distorted by ferromagnetic materials. If accounted for this, EM systems provide a reliable, but slightly less accurate tracking solution compared to OTS. In abdominal surgery, tracking sensors need to be placed in areas where the surgeon would frequently obstruct a direct line-of-sight, making EM tracking the most suitable option.

1.2.3 Surgical navigation at the NKI/AVL

Several ongoing studies at the NKI/AVL focus on the implementation of SN to provide surgeons access to the valuable information contained in pre-operative imaging during surgery. The first study included seven patients and was target at malignancies in the pelvic area while preventing damage to surrounding tissue such as the ureters. The small size of a suspect lymph node makes it difficult to locate the node in a patient. This becomes even more challenging when the lymph node has decreased in size as a response to treatment leading up to the surgery. During this first study, twelve out of thirteen lymph nodes were found, and all tumors were removed radically. For two of these patients the surgeons indicated that radical resection was only possible due to navigation [25]. Since the first trial, other areas are included in the SN studies as well. Current ongoing SN trials are targeting liver, rectum, lymph nodes, bladder, kidney, and oral cavity tumors.

In areas of deformable anatomy, such as the rectum, liver, and tongue, the location of the tumor can move with respect to the sensors placed on the patient due to patient positioning, breathing, or handling of tissue by the surgeon. Several trials are started in which EM sensors are placed close to the tumor to track its motion during surgery. This allows updating of the 3D patient model for real-time visualization of the tumor motion during surgery. Currently, a wired sensor is used for tumor tracking, thereby limiting the possibilities of sensor placement. However, the first steps are made towards replacing the wired by wireless sensors.

Current implementations of SN require that the navigation tool can be placed directly on the organ for visualization of relevant surrounding anatomy. This requirement limits SN to 'open' procedures where there is direct access to the organs for the tool. Currently over 85% of rectal cancer surgeries are, at least partially, performed laparoscopically, meaning that the majority of patients cannot benefit from SN at this time [4]. The use of a navigation tool also requires the surgeon to pause the procedure in order to pick up the tool, point at the anatomy, and look on the screen to get the information, thereby temporarily diverting its attention away from the patient. Ideally the information SN offers is available on demand, without a tool, and within the surgical field.

1.2.4 Augmented reality

Augmented reality (AR) is defined as "an enhanced version of reality created by the use of technology to overlay digital information on an image of something being viewed through a device (as a smartphone camera)" [26]. According to this definition AR overlays digital information, like preoperative imaging, on a view of the world. This approach can be dated back to as early as 1938 when H. Steinhaus described a technique using x-rays to image a bullet inside the head of patient on a fluorescent screen and projecting its position back on the skull with a pointer [27]. In the following decades several revolutionary technological advances, such as the invention of imaging modalities as CT and MRI and the improvement of computers, resulted in an explosion in the field of AR research. The first head mounted display (HMD) was already created in 1968, but it took until the early 90s for computer technology to catch up and be able to produce images in real-time as needed for clinical application of the technique. This first system tracked the HMD and an ultrasound probe to visualize ultrasound images superimposed on a pregnant patient [28, 29]. Since then, AR has extensively been researched and used for education, neurosurgical interventions, and ENT surgery, using a variety of techniques such as HMDs, augmented optics, AR windows, endoscopes, and projections on to the patient. For a more elaborate introduction to the history of AR in medicine the reader is referred to the review of T. Sielhorst et al. [30].

AR can provide a solution to current limitations of SN. By tracking of the laparoscope and patient, as in SN, an overlay for the laparoscopic images can be created from the 3D patient model. This overlay of the patient model directly visualizes the information contained in the pre-operative imaging on the anatomy visible in the laparoscopic images. Direct visualization of the information eliminates the requirement of a navigational tool. As visualization is achieved by projecting the SN information directly on top of the laparoscopic images, the surgeon is also no longer required to divert its attention away from the surgical field to receive this information. Visualization of relevant information allows the surgeon to increase its understanding of the environment. During surgery, AR can be used to display critical information about the patient such as vital functions, location of a tumor, and the location of other critical structures. The projected information is not any different from that of SN, in fact, it is the same information displayed in a different way. AR should therefore be considered an extension of SN in our application. With the increasing use of laparoscopy for (rectal cancer) surgery, development of an AR solution can benefit many patients.

First implementations of AR in laparoscopic surgery are reported in the early 90s, again in the field of ENT- and neurosurgery where the anatomy can be described as rigid relative to externally tracked anatomical landmarks [30, 31]. Since then, the field has expanded to other areas such as oncological adrenal and liver surgery [32]. Under certain circumstances, these are again areas with a rigid transformation relative to external sensors, if placed appropriately. To our knowledge, no implementations have been reported in areas with considerable anatomical deformation, as is the case in rectal surgery. However, with the ongoing research of mobile tumor tracking, this should no longer be a limitation to surgical navigation.

1.2.5 Laparoscopy

Laparoscopes of the type Olympus ENDOEYE HD are used for this research, Table 1. Olympus, and other manufacturers, offer two different types of scopes. The first is a forward viewing scope that has its direction of view along the axis of the scope. The second scope offered is an oblique viewing scope that has its direction of view at a fixed angle to the axis of the scope, in this case 30˚. Both scopes consist of a lens system and sensor placed in the tip of the scope. The images recorded by the sensor are transmitted along the shaft of the scope to a processor for display. In the forward viewing scope, the lens system and sensor are fixed, resembling a camera placed on the tip of a stick. Every movement of the 'stick' results in the same predictable movement of the camera. In the oblique viewing scope on the other hand, the lens system and sensor are oriented independently. The sensor is magnetically coupled to the handle of the laparoscope [33]. This allows the surgeon to keep the desired onscreen orientation of the view on the patient while the lens can be rotated independently to direct the view of the scope in the direction of interest.

Laparoscopic surgery has drastically reduced recovery time, pain, hemorrhage, and infection rate in patients and enables surgery to patients that are not able to receive open surgery due to poor health. These are all excellent benefits to the patient, but have come at the cost of increased surgical complexity such as loss of sense of touch, a smaller range of motion, limited field of view (FOV), poor depth perception, and opposite movement of surgical tools with respect to the surgeons hand due to the pivot point (Fulcrum effect]. Augmented reality can help overcome some of these disadvantages. The limited FOV makes orientation within the patient challenging. If the surgeon is mistaken in orientation, the possibility of surgical complication increases. AR can ease orientation and decrease the likelihood of erroneous surgical decisions by visualization of relevant structures. This visualization does not need to be in the laparoscopic FOV, but can be extended to outside to FOV to help in navigation. Another enhancement AR can provide is an improved sense of depth. 2D images recorded by the laparoscope cannot be easily turned in to 3D, but with the combined knowledge of patient anatomy, laparoscope position, and surgical tool used it is possible to estimate the distance between tool and surface of an organ or tumor. Addition of an onscreen depth metric that uses color or numbers can give the surgeon a sense of depth or distance between tool and patient without visually experiencing the distance [32, 34]. Implementation of these AR benefits is however beyond the scope of this research.

Table 1: ***Laparoscope specification overview of the type ENDOEYE HD from Olympus***

	ENDOEYE HD 5 mm	ENDOEYE HD 10 mm
Field of view	80˚	90˚
Direction of view	0˚ / 30˚	0˚ / 30˚
Depth of field	12 – 200 mm	12 – 200 mmm
Working length	300 mm / 302 mm	335 mm / 330 mm
Distal end outer diameter	5.4 mm	10 mm

1.2.6 Camera calibration

Successful implementation of AR during laparoscopic interventions does not only require tracking of the laparoscope and patient, but also a model that describes how the laparoscopic image is formed from the anatomy in the FOV of the laparoscope. Camera calibration is a procedure extensively used in computer vision to characterize the properties of a camera. If the position of an object is known in relation to a calibrated camera, the projection of the object on the image plane can be simulated from the calibrated parameters. This simulation is used to generate the AR overlay for the laparoscopic images from pre-operative imaging where the position in relation to the camera is known from SN. The lens-system in the laparoscope does not project an exact copy of the scenery on to the sensor [31, 35]. Accurate fusion of laparoscopic images with the AR overlay requires the same projection in both modalities. The projections can be made similar by either distorting the AR overlay in the same way the laparoscope does, or by removing the distortions from the laparoscopic images [31, 36, 37]. In this research we focus on the first since we want to keep the amount of changes as small as possible.

1.3 Definition of mathematical notations

During this thesis several mathematical notations are used. In general, bold font capital letters represent matrices $(\mathbf{M})$, bold lowercase letters represent column vectors $(\mathbf{v})$, italic lowercase letters represent scalar values (s), and coordinate systems are denoted in capital non-bold font (CS). However, if the common notation of symbols in literature deviates from this convention, the convention in literature is used.

During surgical navigation, many different coordinate systems are used. An object known in one coordinate system can be expressed in another coordinate system if the relation between the two coordinate systems in known. The relation between two coordinate systems can be expressed by a homogeneous transformation matrix. Transformation matrices are denoted as ${}^{B}\mathbf{T}_{A}$, where the matrix expresses the pose of coordinate system A in coordinate system B. A transformation matrix consists of a rotation component, the 3x3 rotation matrix $\mathbf{R}$, and a translation component, the 3x1 translation vector $\mathbf{t}$.

$$ {}^{B}\mathbf{T}_{A} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{1.1} $$

In this notation, $\mathbf{0}$ is a 1x3 zero vector. Two transformation matrices can be chained together to describe the transformation from A to C if both are known in relation to B by

$$ {}^{C}\mathbf{T}_{A} = {}^{C}\mathbf{T}_{B} \cdot {}^{B}\mathbf{T}_{A} \tag{1.2} $$

Here, A is first transformed to B, followed by a transformation to C. With this transformation, point $\mathbf{p}_{A}$ in the coordinate system of A can be expressed in the coordinate system of C as point $\mathbf{p}_{C}$ by

$$ \begin{bmatrix} \mathbf{p}_{C} \\ 1 \end{bmatrix} = {}^{C}\mathbf{T}_{B} \cdot {}^{B}\mathbf{T}_{A} \cdot \begin{bmatrix} \mathbf{p}_{A} \\ 1 \end{bmatrix} \tag{1.3} $$

Here, $\begin{bmatrix} \mathbf{p}^{\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}}$ is the expression of point $\mathbf{p}$ in homogeneous coordinates. If the pose of a A is known in relation to B, the opposite relation is given by the inverse on the transformation matrix.

$$
{}^{A}\mathbf{T}_{B} = \left({}^{B}\mathbf{T}_{A} \right)^{-1} = \begin{bmatrix} \mathbf{R}^{-1} & -\mathbf{R}^{-1}\mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{1.4}
$$

As $\mathbf{R}$ is an orthonormal matrix, the expression can be simplified by replacing $\mathbf{R}^{-1}$ with $\mathbf{R}^{\mathrm{T}}$.

1.4 Research design

The ultimate goal is accurate real-time visualization of the tumor and other critical anatomical structures in the laparoscopic images during rectal cancer surgery. There are many challenges on the path to this goal of which some are already conquered, and others beyond the scope of this research. During this study, the focus is on achieving accurate projection of a moving object on to the image plane. Literature on oblique viewing laparoscopes is limited and there exists no literature for the type of laparoscope under investigation in this research. The assumptions made in literature for the camera and laparoscope models are investigated to establish if they are valid in our laparoscope for each parameter independently. Based on the observations, an attempt will be made to produce a calibration method for the laparoscope under investigation that allows real-time visualization. To achieve this, the following subgoals are defined:

1. Design of a camera calibration model for oblique viewing laparoscopes

2. Design of a model to relate the pose of the camera to a reference sensor on the laparoscope

3. Design of a delay estimation procedure between different sources used

4. Proof of concept where the first three subgoals are combined

5. Evaluation of the achieved results and their implications for clinic application

CH2: Camera calibration

2.1 Introduction

Camera calibration describes the relation between an object at known position in front of the camera and its projection on the image plane. The parameters obtained during camera calibration are unique to a camera. Even cameras of the same brand and type have slightly different calibration parameters due to small geometrical differences in the lens configuration. Of the two common types of rigid laparoscopes, the forward viewing scope can directly be calibrated using a standard calibration algorithm such as Tsai's, or Zhang's method [38-40]. However, in the oblique viewing laparoscope, the camera system does not have a fixed configuration. The angled view of the laparoscope has the advantage of a much larger field of view through rotation of the scope cylinder. While the scope is rotated, the image sensor keeps a fixed orientation relative to the handle. This independent rotation of image sensor and scope cylinder changes the calibration parameters for each angle, requiring a calibration model that describes the camera parameters as a function of the rotation angle.

Only a few groups have developed methods to describe the camera calibration for an oblique viewing laparoscope. In order to describe the camera parameters as a function of the rotation angle it is necessary to know the rotated angle. This is achieved by tracking of the movable parts of the laparoscope as described in the next chapter. Yamaguchi et al. [41] described the pose of the camera as a function of the rotation angle and kept all internal parameters of the camera constant. Wu et al. [42] improved the Yamaguchi method by keeping the pose of the camera fixed in relation to the scope cylinder, and rotating the image around the center of the image plane. De Buck et al. [43] modeled the camera pose in a similar way and extended the standard camera model by interpolation of internal camera parameters obtained at several angles to account for scope rotation. The most recent model by Liu et al. [44] improved on these methods by rotating the image around a rotation axis, defined as the center of principal points obtained from calibration at several angles, instead of rotating around the center of the image. Melo et al. [45] presented a method that relies on a wedge mark in the image to model the camera parameters to the rotation angle. However, this wedge mark is not available in the Olympus scopes used here and many other laparoscopes.

All of the proposed methods assume that some or all of the camera parameters are fixed during rotation. However, none of the authors evaluated what happens to the camera parameters during rotation to validate these assumption. Decentering distortion, a parameter extensively used in computer vision to correct for a specific type of lens distortion, is also excluded in all proposed methods. Here we aim to develop a camera calibration model that can accurately describe the camera as a function of the rotation. To do this, all parameters are evaluated independently and combined in to a single model.

2.2 Method

Camera calibration is performed using Zhang's method as implemented in Matlab R2017a [39, 40]. Camera calibration is performed at nine angles ranging from -120° to 120° with increments of 30°. At each angle, nine images of the standard Mathworks checkerboard pattern are captured. The checkerboard is printed on a flat board with pattern square size of 1.02 cm. The nine positions of the checkerboard are chosen to equally distribute corner points in the checkerboard over the image,

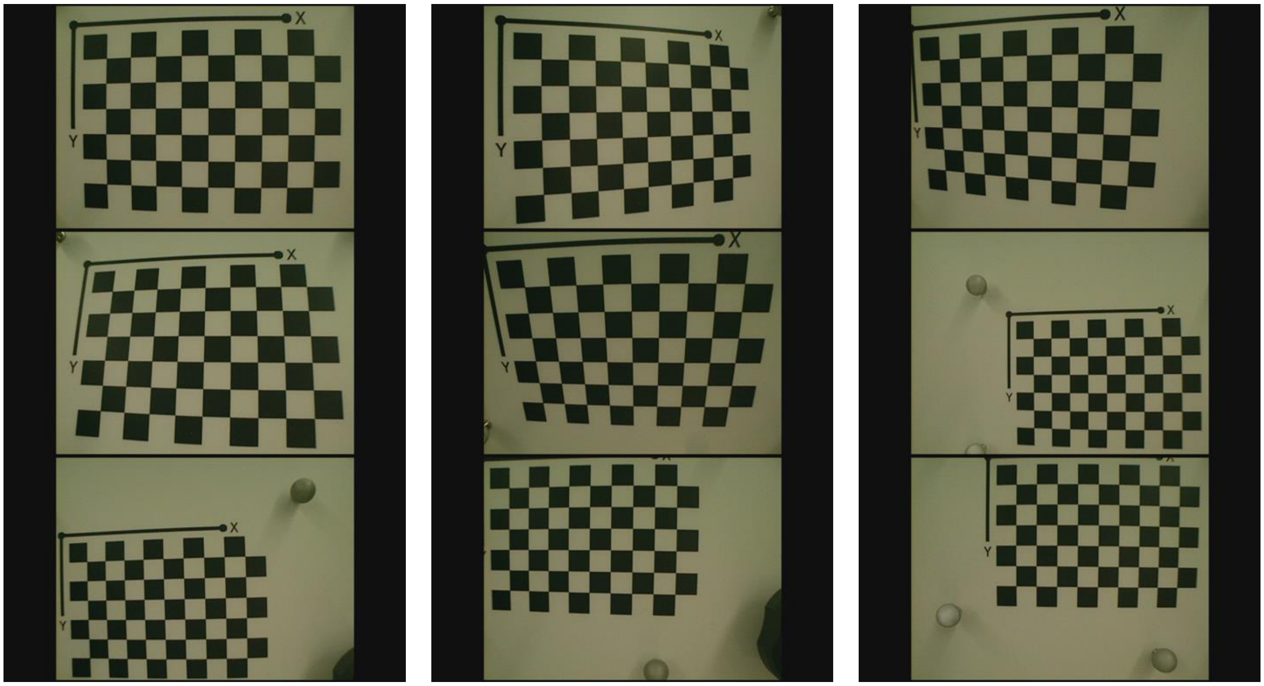


Figure 1: Nine calibration images captured for each of the nine angles camera calibration is performed at. Image poses are chosen to equally distribute the corner points over the entire field of view.

Figure 1. The laparoscope image processor has an edge enhancement function that is set to its lowest value to minimize the influence of image processing on the calibration.

<u>Camera model</u>

Let (x_p, y_p) be the normalized pinhole projection of a point after lens distortion, and (u_p, v_p) its corresponding point in pixel coordinates given by

$$\begin{bmatrix} u_p \\ v_p \\ 1 \end{bmatrix} = K \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} \tag{2.1}$$

With K the camera matrix containing the camera's intrinsic parameters.

$$K = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.2}$$

Here, f_x, f_y are the focal lengths in pixel dimensions, and (u_0, v_0) the principal point coordinates in pixel dimensions. If pixel axes are not orthogonal, K can be extended to include a skew parameter to correct for this. Coordinates are usually not defined in the coordinate system of the camera, but in some other coordinate frame that we call world. To apply (2.1), the world coordinates are first transformed to the camera coordinate system by

$$\lambda \begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = \left[\mathbf{R} \mid \mathbf{t} \right] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{2.3}$$

Here, λ is a scaling factor, and $\mathbf{R}, \mathbf{t}$ are the rotation and translation needed to transform a point $(x_w, y_w, z_w, 1)$ in homogenous world coordinates to its unnormalized position in the camera coordinate system, ignoring distortions. If we refer to world as CBI, and camera coordinate system as CC, the transform ${}^{CC}\mathbf{T}_{CBI}$ needed for hand-eye calibration in the previous chapter is defined by

$${}^{CC}\mathbf{T}_{CBI} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{2.4}$$

Intrinsic and extrinsic parameters $\{ f_x, f_y, u_0, v_0, \mathbf{R}, \mathbf{t} \}$ are estimated using Zhang's method as implemented in Matlab 2017a. See Appendix A for a more extensive derivation and interpretation of the camera model.

<u>Lens distortions</u>

Radial distortion of lenses causes a displacement of projected points along the radial lines from the principal point. Radial distortion for the first three radial components $\{k_1, k_2, k_3\}$ is given by

$$\begin{bmatrix} \Delta x_r \\ \Delta y_r \end{bmatrix} = \begin{bmatrix} x'_p \\ y'_p \end{bmatrix} \left(k_1 r_p^2 + k_2 r_p^4 + k_3 r_p^6 \right) \tag{2.5}$$

Here, (x'_p, y'_p) are the undistorted normalized points, and $r_p^2 = x'^2_p + y'^2_p$ the radius measured from the principal point. Decentering distortion by misalignment of the image sensor and optical axes of the lenses in the lens-system is given by

$$\begin{bmatrix} \Delta x_d \\ \Delta y_d \end{bmatrix} = \begin{bmatrix} 2x'_p y'_p & r_p^2 + 2x'^2_p \\ r_p^2 + 2y'^2_p & 2x'_p y'_p \end{bmatrix} \begin{bmatrix} J_1 \cos(\varphi) \\ -J_1 \sin(\varphi) \end{bmatrix} \tag{2.6}$$

Here, φ is the direction of the decentering distortion, it indicates the axis the lens is tilted on, and J_1 is the magnitude of distortion, an indication of the amount of lens tilting. The two are combined to get the parameters estimated in camera calibration.

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} J_1 \cos(\varphi) \\ -J_1 \sin(\varphi) \end{bmatrix} \tag{2.7}$$

With $\{p_1, p_2\}$ the decentering distortion parameters. Lens distorted normalized points are given by

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = \begin{bmatrix} x'_p \\ y'_p \end{bmatrix} + \begin{bmatrix} \Delta x_r \\ \Delta y_r \end{bmatrix} + \begin{bmatrix} \Delta x_d \\ \Delta y_d \end{bmatrix} \tag{2.8}$$

During camera calibration the distortion parameters $\{k_1, k_2, k_3, p_1, p_2\}$ are estimated as well. Besides displacement of projected points with respect to the principal point, decentering distortion also displaces the principal point itself. Displacement of the principal point due to decentering of a lens is given by

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} x'_0 \\ y'_0 \end{bmatrix} + \rho \begin{bmatrix} \sin(\varphi) \\ \cos(\varphi) \end{bmatrix} \tag{2.9}$$

Here, $(x_0, y_0), (x'_0, y'_0)$ are the distorted and undistorted principal point coordinates respectively, and ρ the amount of displacement due to decentering given by

$$\rho = 3c(\mu - 1)\varepsilon \tag{2.10}$$

With c the principal distance (calculated focal length), μ the index of refraction, and ε the tilt angle of the lens in radians given by

$$\varepsilon = \frac{2c\mu J_1}{(\mu - 1)(\mu + 1)} \tag{2.11}$$

In camera calibration, this shift of the principal point is ignored as its effect on the projection of points is compensated by a change in extrinsic parameters, only the distortions are influenced by this shift. The overall effect on distortion of points is marginal and can therefore be ignored [46]. For a more extensive derivation of the distortions, its parameters, and an interpretation of the distortions, see Appendix B.

<u>Laparoscope lens-system model</u>

Camera calibration as described above will return parameters specific to the lens-system configuration used for collection of the calibration images. As there is no information available on the camera configuration present in the laparoscope, several assumptions are made to allow prediction of the behavior of the calibration parameters. Rotation of the scope changes the aiming direction of the scope, but does not rotate the image. This means that the sensor and the outer lens are not contained in one compound lens-system. Therefore the assumption is made that the system consists of two compound lens-systems, Figure 2. The outer compound system is referred to as Optics, it contains the lenses with the optical axis in the oblique view direction, and a prism that changes the direction of the optical axis from the oblique axis to one parallel with the axis of rotation. The inner compound system is referred to as Camera, it consists of the image sensor and lenses with the optical axis parallel to the rotation axis. Camera has a fixed pose in relation to the handle of the scope, while Optics is attached to the outside of the scope and rotates when the scope shaft is rotated. As the two systems move independently during rotation, the parameters will change depending on the amount of rotation.

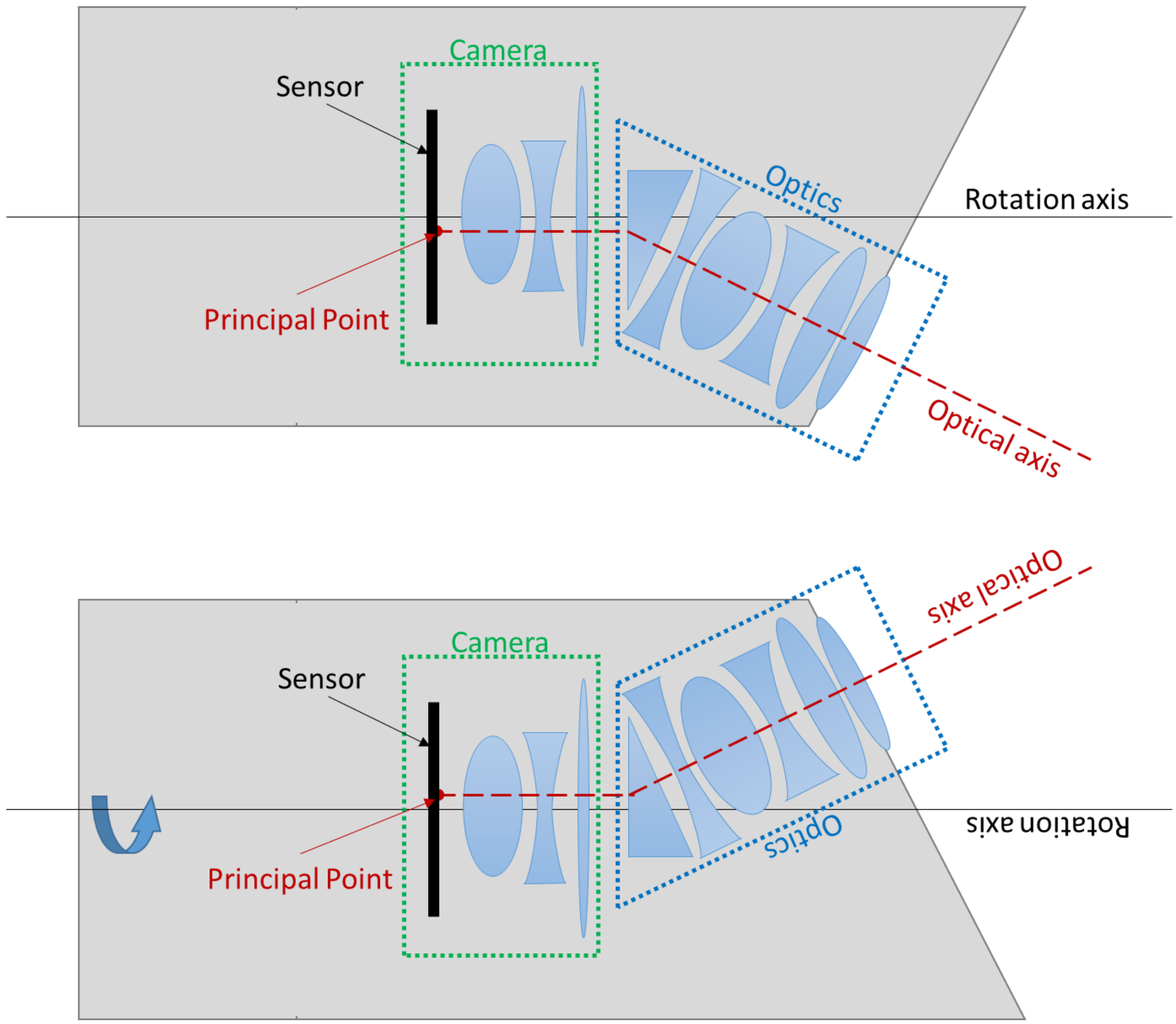


Figure 2: Model of the camera in the laparoscope. The total lens-system can be separated in to two compound lens-systems. One of the compound lenses contains the camera sensor and several lenses and is referred to as Camera. Ideally the sensor is orthogonal to the rotation axis, however, this does not need to be the case. The other compound lens-system, referred to as Optics, contains several lenses and a prism to align the optical axis of the oblique viewing part of the scope with the Camera system. Results of camera calibration are the combination of both compound systems. Independent rotation of the parts can influences camera parameters such as, a displacement of the principal point.

<u>Effects of rotation on calibration parameters</u>

Principal point position is determined by the optical axis. If the optical axis does not coincide with the rotation axis, rotation of the scope will, in an undistorted lens-system, result in a circular motion of the principal point around the point where the rotation axis intersects the image sensor. As decentering distortion also changes the location of the principal point, the actual pattern of the principal point due to rotation can deviate from a circular pattern.

Focal distances are determined by the pixel dimensions and distance between sensor and lenses. The assumption is made that distance between sensor and lenses does not change as this would have unwanted noticeable changes in the image that are not observed in the used system. Since the camera pose is defined by the axis of the image sensor, the pixel dimensions do not change either. Focal distances are therefore expected to be constant.

Extrinsic parameters depend on the optical axis coming out of the scope. As the direction of the optical axis is changed by rotation of the scope the extrinsic parameters will change accordingly. These changes are modeled by hand-eye calibration in the next chapter.

Radial distortion is radially symmetrical around the principal point. Radial distortion parameters are therefore expected to be constant with the center of distortion changing according to the movement of the principal point. As radial distortion depends on the focal length, the assumption will only hold if the focal lengths are indeed constant.
Decentering distortion originates from lens misalignment. As we have defined two compound lens-systems, the total decentering distortion is described by a sum of three separate distortions. Each compound system has its own internal decentering distortion, and there is an external decentering distortion between the compound systems. Magnitude of the internal distortions do not change as the compound systems are unaffected by rotation. Direction of the internal distortion of Optics is determined by rotation, and that of Camera is fixed as the compound system also holds the sensor. External distortion originates from the decentering between the compound systems. As the tilt and displacement between the two systems changes during rotation, magnitude and direction change with rotation. Decentering distortion parameters as a function of the rotation angle θ_{view} is given by

$$\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} p_1^C \\ p_2^C \end{bmatrix} + \mathbf{R}\left(\theta_{view}\right)\begin{bmatrix} p_1^O \\ p_2^O \end{bmatrix} + \begin{bmatrix} p_1^{CO}\left(\theta_{view}\right) \\ p_2^{CO}\left(\theta_{view}\right) \end{bmatrix} \tag{2.12}$$

Here, p^C, p^O, p^{CO} are the distortions due to Camera, Optics, and interaction between the two compound systems respectively, and $\mathbf{R}\left(\theta_{view}\right)$ a 2D rotation matrix corresponding to the angle of rotation. By ignoring the external distortion, $\left\{p_1, p_2\right\}$ describe a circle during rotation. The added external distortion will cause a deviation from this circular path.

<u>Angle dependent modelling</u>
Focal lengths and radial distortion parameters are estimated by finding the value that fits best to all calibration angles. Camera calibration returns a value and uncertainty estimation for each of the parameters at each angle. From the estimated values and uncertainties a Gaussian profile can be created for each angle. Gaussian profiles for all angles are normalized with respect to area and summed, the peak value is set as the estimated parameter for the focal lengths and radial distortions.

Ignoring the effects of the external component of decentering distortion, the principal point and decentering distortions parameters will describe a circle during rotation of the scope. Depending on the magnitude of the external decentering component, the actual path described can be closer to an ellipse. As the external decentering distortion depends on the relative positions and lens tilting between the two, its true contribution is complex and hard to model with little data. Therefore, an ellipse is fitted to the principal point and decentering distortion parameters, both estimated at several angles.

Reprojection error

Accuracy of the model is assessed by comparison of the pixel reprojection error obtained using parameters given by the model, and those obtained during single angle camera calibration. Reprojection error within an image is given as the root-mean-square (RMS) of the distances between projected points p_{proj}, and detected points in the image p_{det}.

$$\text{Reprojection error} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} d\left(p_{proj}^{n}, p_{det}^{n}\right)^2} \tag{2.13}$$

Here, $d(\cdot,\cdot)$ is the Euclidean pixel distance, and N the numbers of corner points in an image.

2.3 Results

Calibration is performed with and without including decentering distortion in the model, at angles ranging from -120° to +120° with increments of 30°. The angles are updated using rotation calibration as performed in the next chapter where 0° is set as the reference angle for the rest. At each angle, nine images are captured, Figure 1. Image positions are chosen to get a good distribution of calibration points over the image plane, Figure 3.

Focal lengths and radial distortion

When decentering distortion is included in the model, estimated parameters have a similar value over all angles and their standard deviations are narrower, Figure 4. This allows the focal lengths and radial distortion parameters to be approximated by a single value for all angles. When decentering distortion is

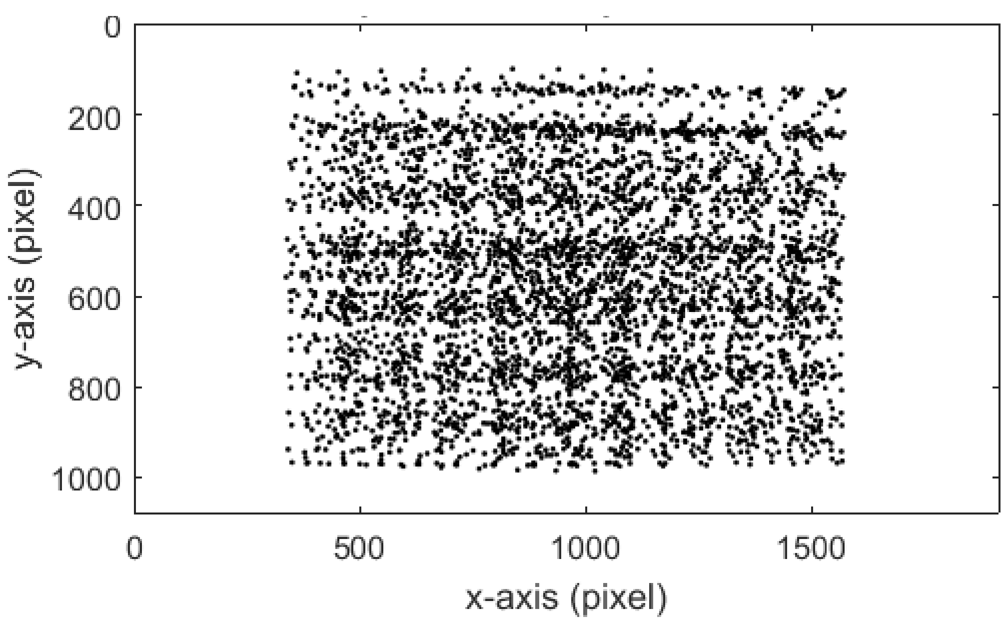


Figure 3: Scatterplot of all found corner points of the checkerboard over nine angles and nine images corresponding to the locations shown in Figure 1. Large empty areas on the side are where the images are black. Chosen pattern positions place the entire checkerboard within the image. No corners are detected on the outer edge of the pattern. This creates a white bar at the top and bottom of the image where no corners are detected.

excluded from the model, parameters have a larger spread and standard deviation. For the focal lengths the spread is large enough that it is not possible to choose a single value for the model. Radial distortion parameters have a larger standard deviation as their number increases, corresponding to a decrease in influence on the distortion by the larger parameters.

Principal point

An ellipse is fitted to the principal points obtained by camera calibration of the nine angles, Figure 5. Model based locations are obtained by rotation of the principal point around the rotation axis, defined as the center of the ellipse, using the angles acquired by optical tracking of the laparoscope. Model positions are estimated in reference to angle 0° which is assumed to be correct. Displacement effects of decentering distortion on the principal point is investigated by correcting the measured principal point for decentering distortion. For correction, refractive index is assumed to be equal to glass (1.5), and sensor size is assumed to be 1/3.2" (4.54x3.42 mm) to leave enough space in the tip of the scope for other components and light fibers to pass the sensor. Principal point positions for calibration without decentering distortion have marginal differences compared to calibration with decentering distortion.

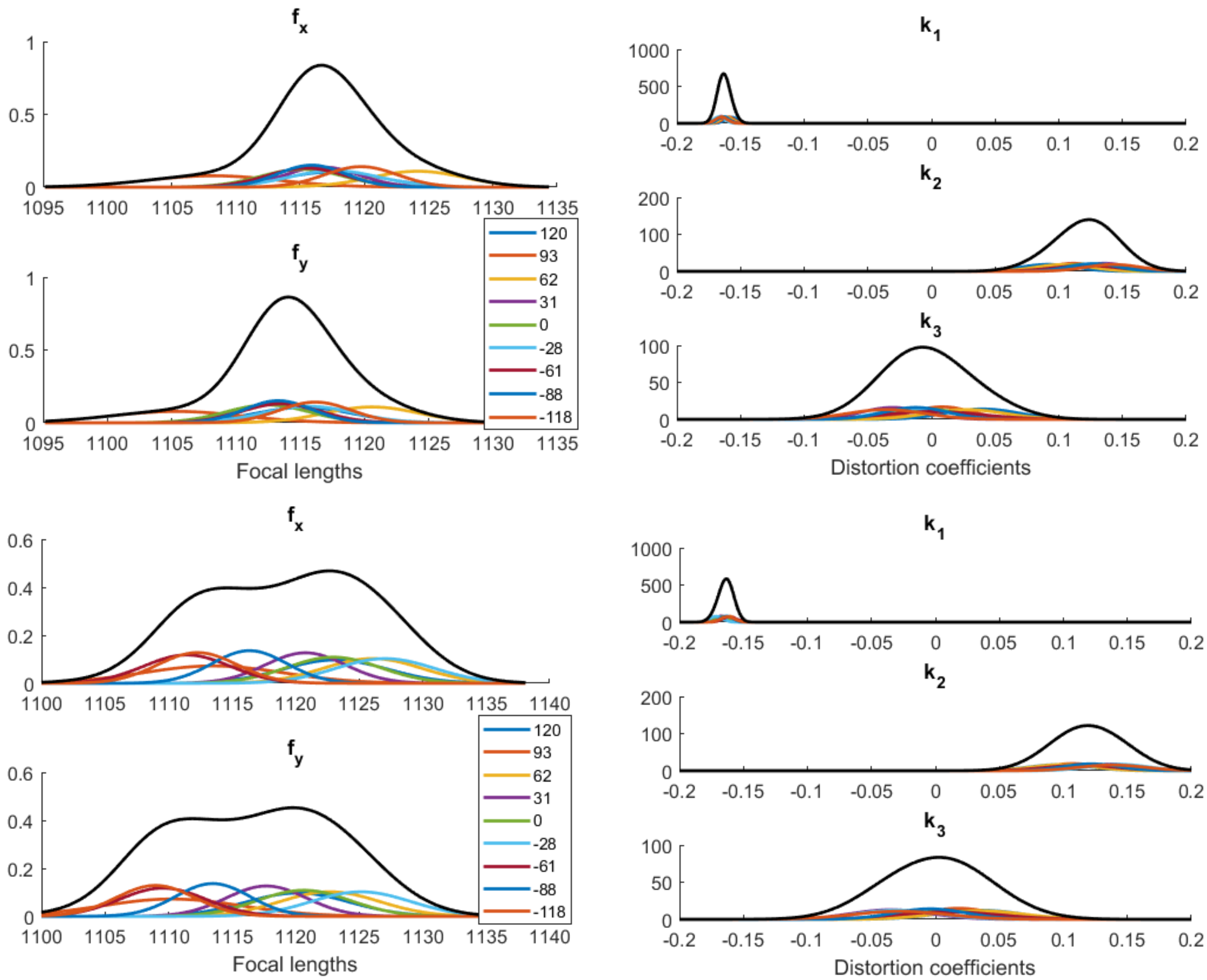


Figure 4: Gaussian profiles for focal lengths and radial distortion coefficients obtained by camera calibration for the nine calibration angles. Top images are acquired when decentering distortion is included in the model, bottom without decentering distortion.

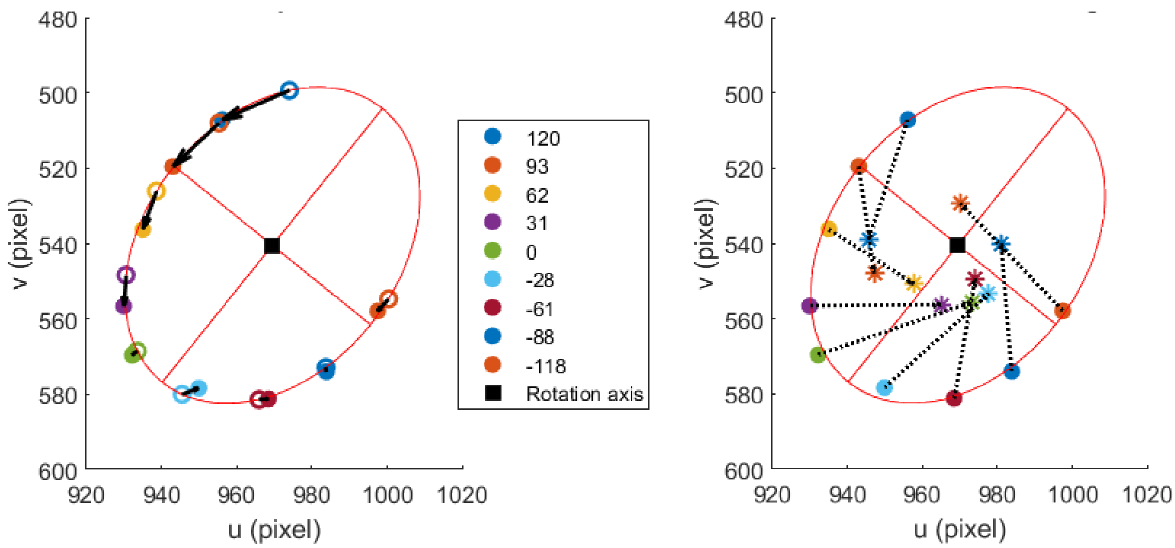


Figure 5: Left, ellipse fitted on principal points obtained during camera calibration (solid), and the locations estimated with the model (open). Right, estimated positions of the obtained principal points corrected for the displacement imposed on the principal points by decentering distortion (star). For correction, the sensor size was assumed to be 4.54x3.42 mm, and index of refraction μ = 1.5.

<u>Decentering distortion</u>

Again, an ellipse is fitted to the decentering distortions parameters obtained by camera calibration, Figure 6. Ellipse axis length ratio is large, indicating a large influence by the external component of decentering distortion. Decentering is modeled in the same way as for the principal point. Model based decentering is estimated by rotation of the decentering at 0° around the rotation axis with the angle measured by navigation. Model errors are large for larger angles suggesting an ellipse might not be the best model.

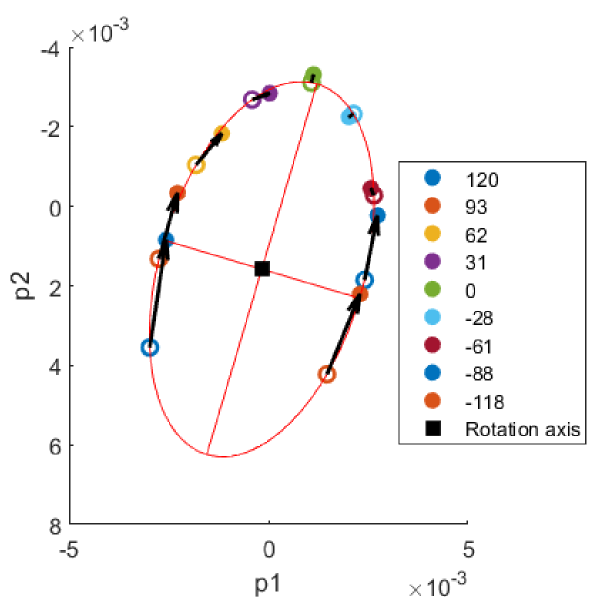


Figure 6: Decentering distortion for the nine angles. An ellipse is fitted to the measured decentering distortion (solid) to model the decentering distortion as a rotation around the center of the ellipse (open) with reference to 0°. Larger angles show a large error in the modeled decentering distortion.

Reprojection error

Model based and calibration based reprojection errors are compared with and without decentering distortion included, Figure 7. Reprojection error per angle is given as the mean RMS of all images for that angle. Extrinsic parameters are re-estimated to account for the changes in camera parameters due to modelling before the reprojection error is determined. Calibration based errors are consistent and angle independent. Method based comparison shows inclusion of decentering distortion results in a lower error in all cases with the exception of one angle in the model based method. At the reference angle, the model based reprojection error including decentering distortion is lower than the calibration based reprojection error without decentering included. For all methods, the mean reprojection error is lower than 1 pixel.

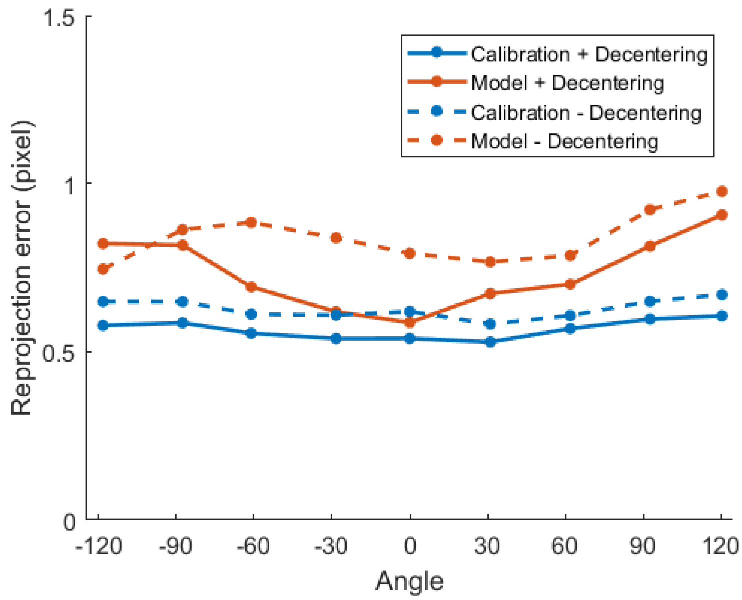


Figure 7: Reprojection errors using parameters obtained by calibration and model based method, with and without decentering distortion included. Model based errors show some angle dependency and are overall larger than the calibration based method. At reference angle 0° the model based method with decentering distortion has a lower error compared to the calibration based method without decentering distortion.

2.4 Discussion

Inclusion of decentering distortion in the model results in overall better performance. It also allows focal lengths and radial distortion to be approximated to a single value. As the true focal length is unlikely to change during rotation, the spread in focal lengths found without decentering distortion can be explained as a compensation mechanism for the effects of unmodeled decentering. Radial distortions show a narrow peak for k_1 increasing in width to k_3. This is to be expected as the effects of radial distortion decreases with an increase in coefficient number. As the first coefficient has the most influence, its estimated value needs to be more precise compared to the other coefficients to correctly describe the distortion. For this particular laparoscope the radial distortion can be described by two parameters instead of three as the third coefficient is nearly zero. As this is not the case for all laparoscopes, the model can best be approximated by three coefficients for generality.

Principal point modeling is accurate for zero and negative angles. Modeled principal points have a positional discrepancy with the measured principal point, increasing in size as the angle between modeled point and reference point zero increases. Part of this error can be explained by the incorrect angle determined with navigation as described in the next chapter. The one-sidedness of incorrect principal point estimation is caused by assuming the reference angle is correct. One or two pixels offset against the clock for the reference principal point results in an error distribution similar to that of decentering distortion. Correction for decentering distortion shows a significant influence of decentering distortion on the position of the principal point. A consequence of this is that a model of the principal point needs to include the effects of decentering distortion on the path described by the principal point. If the magnitude of decentering distortion has a large variation, the path described by the principal point will show a similar effect. The overall shape change of the path depends on the radius of the rotation described for the undistorted principal points, and the magnitude of decentering distortion. If decentering is low in comparison to the undistorted path, principal point can be approximated by a circle. As the effects of decentering seem to be larger than the undistorted radius, its effects cannot be ignored and the principal point cannot be described by a circular path around the rotation axis.

Decentering distortion model shows a strong elliptical path indicating a significant contribution by the external decentering distortion component. However, estimated distortions based on the model show a large difference compared to the calibrated distortion with a magnitude of decentering twice as large for the model compared to calibrated at ±120°. This could either be due to an ellipse not being the right model, or that the decentering interaction between the two compound lens-systems is not linear to the rotation angle. A definitive answer to this question requires more measurements, and will be investigated in chapter 5.

Average reprojection error per method and angle is lower than one pixel in all cases. For the calibration based method the error is consistent over all angles indicating that by correctly modeling all parameters to the angle it is possible to obtain an angle-independent reprojection error. In both methods the reprojection error when excluding decentering distortion in the model is slightly higher than for the included case. This low error is obtained by compensating for the lack of decentering distortion in the model with adjustment of the other parameters as can be seen in the spread of the focal length, and to a lesser extent in radial distortion. If the focal length and radial distortion parameters are fixed, reprojection errors in the model based method without decentering distortion describes a sinusoid due to a lack of this compensatory mechanism.

In the decentering model based method, the error is angle dependent. At the reference angle, the error for the model based method is even lower than the error obtained using calibration for the case without decentering, increasing to 1.5 times the error of the calibrated case in angles ±120°. The angle dependency is due to incorrect modelling of the decentering distortion. The error in principal point does not contribute to the reprojection error as the error at ±90° is the same while the principal point has no error for -90° but does have an error in the +90° case. This error is independent on principal point here due to re-estimation of the extrinsic parameters based on the model. A shift in principal point is corrected for by a similar shift of the calibration object in the coordinate system of the camera. This suggests that small errors in principal point position can be corrected for by the hand-eye calibration model. To do this, it is necessary to use the extrinsic parameters obtained with the model based intrinsic and distortion parameters, not those obtained during calibration, to accurately describe the hand-eye calibration as a function of the angle. If the extrinsic parameters are not re-estimated, the reprojection error increases to 5-25 pixels for both methods (not shown).

2.5 Conclusion

Obtained results suggest that the camera model can be described with a constant focal length and radial distortion at all angles if decentering distortion is modeled correctly. Principal point heavily depends on decentering distortion. If the distortion is large, the effects of decentering distortion need to be included in the model to accurately describe principal point position. However, errors in principal point position can be compensated for by hand-eye calibration. This requires feeding the extrinsic parameters corresponding to the parameters of the model based method to the hand-eye calibration procedure, and not the extrinsic parameters obtained during camera calibration. Decentering distortion is shown to have a large influence on all parameters regardless of its small magnitude. It is therefore necessary to accurately model the decentering distortion as a function of the rotation angle. The current model of an ellipse is not sufficient to correctly describe decentering distortion. This is likely due to the complex decentering distortion between the two compound lens-systems of the camera during rotation. In chapter 5 these effects will further be investigated. In all cases the reprojection error is less than one pixel. On an image size of 1440x1080 pixels a reprojection error of less than a pixel is not discerning. How the reprojection errors evolve if the extrinsic parameters are not estimated from the image, but obtained by navigation, will be the true test of this model and are investigated in chapter 5.

CH3: Hand-eye calibration

3.1 Introduction

With AR we aim to create a virtual image of a tumor that is projected in the correct position of the patients anatomy as visible on the image captured by the laparoscope. This requires the that the position of the tumor in known in relation to the camera. The position of the patient and tumor is known from tracking sensors placed for surgical navigation. The effective position of the camera is located somewhere in the tip of the laparoscope. Tracking sensors are placed on the laparoscope, followed by a registration to procedure to relate the effective position of the camera to the tracking sensor. This registration procedure is referred to as hand-eye calibration. In oblique viewing laparoscopes, the camera's lens-system can move independently from the image sensor. As the camera pose is defined by the combination of lens-system and image sensor, two sensors are placed on the laparoscope to track the individual movements. One sensor is attached to the handle and has a fixed relation to the camera's image sensor, while a second sensor is attached to the cylinder of the laparoscope to track the motion of the lens-system.

Either one of the sensors attached to the laparoscope can serve as reference for tracking of the camera, and both applications have been described in the few literature references available on the topic. Yamaguchi et al. [41] were the first to attempt calibration of an oblique viewing laparoscope. They attached an optical sensor to the handle, and a rotary encoder was used to track the relative rotation between the moving parts. The proposed method was rather complex as it required estimation of five parameters for the hand-eye calibration model. All other methods use the sensor attached to the scope cylinder as reference for tracking of the camera [42-44]. Both electromagnetic and optical systems have been used for this purpose. These methods assume a fixed relation between the optical axis of the laparoscope and the reference tracking sensor attached to the scope cylinder. This assumption is used to create a projection of an object using the standard camera model, initially ignoring the orientation of the image sensor. After projection, a rotation correction is applied to align the orientation of the projected image with the orientation of the image sensor. Rotation correction is achieved by rotating the image around a point in the image plane. The point around which the projected image is rotated differs per method and is either the center pixel of the image plane, the principal point (one of the intrinsic camera parameters), or a point in the image plane around which the principal point rotates. These methods are relative simple compared to the reference on the handle as they only require one or two parameters to be estimated for the hand-eye calibration model.

Current choices for the reference sensor in hand-eye calibration methods are based on the hardware available for tracking, and simplicity of the model. All methods are evaluated by the overall reprojection of the combined camera model and hand-eye calibration model, but none of the authors evaluated the hand-eye calibration itself. Here we will investigate what sensor can best serve as a reference for hand-eye calibration, and what model best describes the relation between the reference sensor and camera in the tip of the laparoscope. This is achieved by inspection of the hand-eye transformations in several laparoscope configurations to gain insight in the differences between hand-eye transformations with respect to both reference sensors.

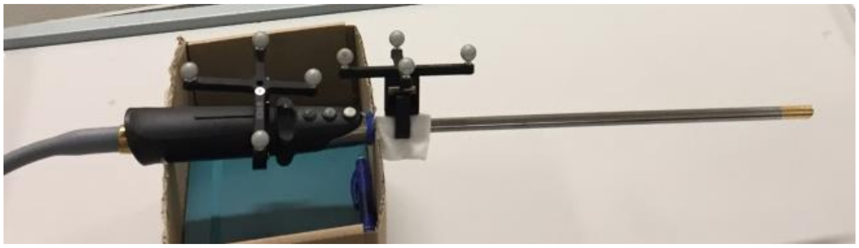
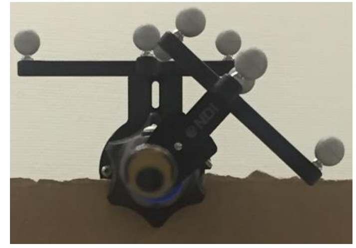
 

Figure 8: Top, side and front view of the Olympus EndoEye HD 10 mm with an oblique view angle of 30° and two optical frames attached for navigation.

3.2 Methods

Laparoscope

An Olympus EndoEye HD 10 mm oblique viewing laparoscope is used for this study, Figure 8. The scope's viewing direction is 30 degrees away from the scope's cylinder axis. Olympus' chip-on-the-tip technology allows the camera's sensor in the tip to keep a fixed pose with respect to the handle while the scope's cylinder can be rotated to change the direction the camera is aimed at. In the camera, two coordinate systems are defined. The first is the coordinate system of the image sensor and will be referred to as CCD, the second is the coordinate system of the camera itself with its origin in the 'pinhole' of the camera and is referred to as CC, Figure 12. CC is the coordinate system used in the camera model for projection of an object and is the coordinate system of interest in hand-eye calibration.

Hand-eye calibration of the laparoscope

During surgery, the pose of CC is determined based on tracking of the rigid bodies attached to the handle (H) and scope's cylinder (S) denoted by transforms ${}^{OTS}\mathbf{T}_H$ and ${}^{OTS}\mathbf{T}_S$. Here, OTS refers to the coordinate system of the optical tracking system; see Figure 9 and Figure 11. Hand-eye calibration is performed to estimate the transformation between coordinate systems S or H and that of CC. There is no way of directly measuring CC by the OTS. Hand-eye calibration is therefore performed with a checkerboard pattern of which the pose is estimated in relation to CC during camera calibration as explained in the previous chapter. By attaching optical reflecting spheres to the checkerboard pattern, its pose can also be estimated by the OTS. As the pose of the checkerboard pattern is now known in relation to CC and the OTS, it can indirectly relate CC to the OTS, Figure 9.

For hand-eye calibration two extra coordinate system and three extra transformations need to be defined. CB defines the coordinate system of the checkerboard as tracked by the OTS, given by ${}^{OTS}\mathbf{T}_{CB}$, and CBI defines the coordinate system of the checkerboard pattern itself. ${}^{CBI}\mathbf{T}_{CC}$ denotes the transformation from the checkerboard image to the camera coordinate system, and ${}^{CB}\mathbf{T}_{CBI}$ the transformation between the checkerboard pattern and the optical sensor attached to the board. CC can now be expressed in relation to H or S by

$$\begin{aligned} {}^{H}\mathbf{T}_{CC} &= \left({}^{OTS}\mathbf{T}_H\right)^{-1} \cdot {}^{OTS}\mathbf{T}_{CB} \cdot {}^{CB}\mathbf{T}_{CBI} \cdot {}^{CBI}\mathbf{T}_{CC} \\ {}^{S}\mathbf{T}_{CC} &= \left({}^{OTS}\mathbf{T}_S\right)^{-1} \cdot {}^{OTS}\mathbf{T}_{CB} \cdot {}^{CB}\mathbf{T}_{CBI} \cdot {}^{CBI}\mathbf{T}_{CC} \end{aligned} \tag{3.1}$$

Here, ${}^{H}\mathbf{T}_{CC}$, ${}^{S}\mathbf{T}_{CC}$ are the transforms to CC from H and S respectively.

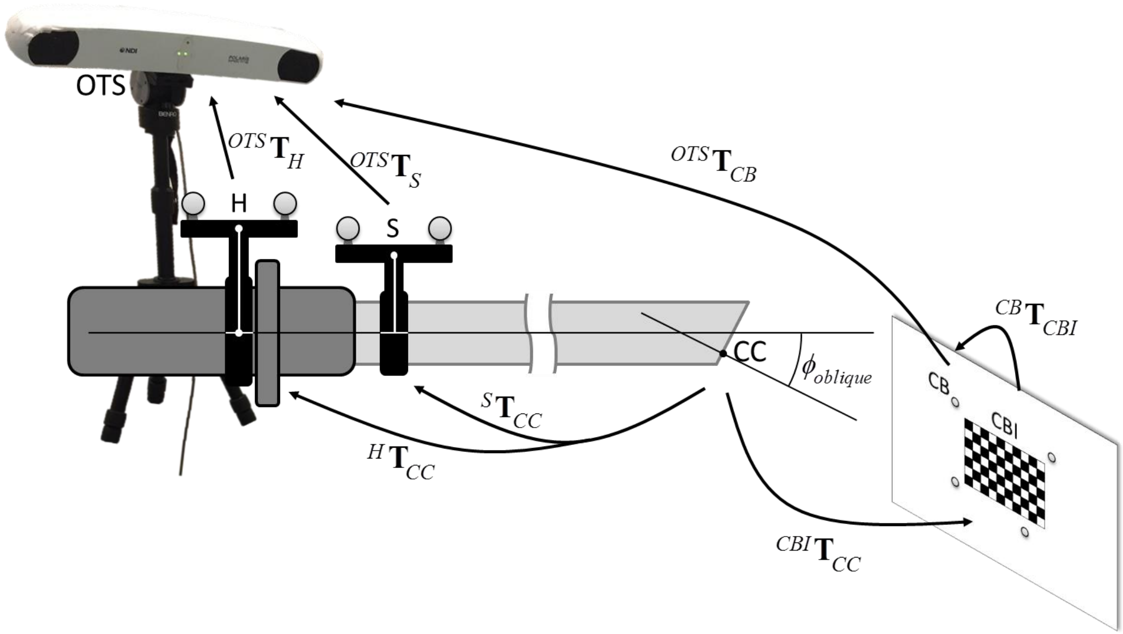


Figure 9: Hand-eye calibration is performed using an object that can be tracked by both the camera and OTS. $^{H}T_{CC}$, $^{S}T_{CC}$ are the target transforms of hand-eye calibration. H and S can be expressed in relation to the checkerboard CB by the OTS. The checkerboard pattern CBI is known in relation to CC through camera calibration. The two systems can be linked together by determination of the fixed transform between CB and CBI.

Estimation of $^{CB}\mathbf{T}_{CBI}$

Transform between checkerboard image and optical sensor on the board (CBI and CB) is determined with a point registration. Let $\mathbf{p}^{i}_{CBI}, \mathbf{p}^{i}_{CB}\ (i = 1, 2, 3, \ldots, n)$ be column vectors of the corner point positions in the checkerboard with respect to CBI and CB respectively where the *i*th corner of CBI corresponds with the *i*th corner of CB. Corner points positions in CBI are known from the geometry of the checkerboard, Figure 10. Corner points in CB are determined with point measurements using an optically tracked pointer. $\mathbf{p}_{CB}$ are defined by the translation vector in $^{CB}\mathbf{T}_{pointer}$. The two point sets are registered using the Procrustes algorithm. First, the centroid of both point sets is shifted to the origin of their respective coordinate system

$$\begin{aligned} \mathbf{p}'^{i}_{CBI} &= \mathbf{p}^{i}_{CBI} - \overline{\mathbf{p}}_{CBI} \\ \mathbf{p}'^{i}_{CB} &= \mathbf{p}^{i}_{CB} - \overline{\mathbf{p}}_{CB} \end{aligned} \tag{3.2}$$

Here $\overline{\mathbf{p}}$ is the centroid, or mean coordinate of the set of points. Rotation is found using the Kabsch algorithm. The 3x3 cross-covariance matrix $\mathbf{C}$ between the point sets is calculated by

$$\mathbf{C} = \sum_{i=1}^{n} \mathbf{p}'^{i}_{CBI} \cdot \mathbf{p}'^{i\ \mathrm{T}}_{CB} \tag{3.3}$$

From the cross-covariance matrix the optimal rotation matrix $\mathbf{R}$ between the two point sets is determined using singular value decomposition

$$\mathbf{C} = \mathbf{USV}^{\mathrm{T}} \tag{3.4}$$

$$\mathbf{R} = \mathbf{VU}^{\mathrm{T}} \tag{3.5}$$

We require rotation for a right handed coordinate system, meaning $\det(\mathbf{R}) = 1$. If the determinant is -1, the rotation matrix is recalculated using

$$\mathbf{R} = \mathbf{V} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \mathbf{U}^{\mathrm{T}} \tag{3.6}$$

Translation between the coordinates system is given by

$$\mathbf{t} = -\mathbf{R} \cdot \overline{\mathbf{p}}_{CBI} - \overline{\mathbf{p}}_{CB} \tag{3.7}$$

Using this rotation and translation the transformation matrix from CB to CBI is defined as

$$^{CB}\mathbf{T}_{CBI} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{3.8}$$

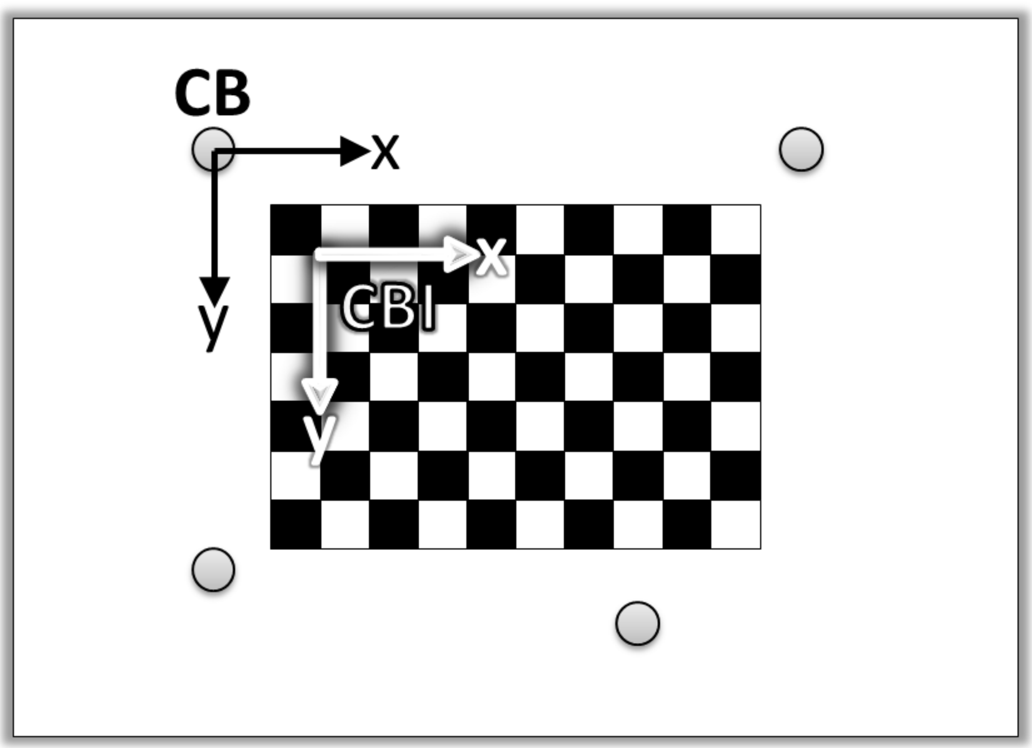


Figure 10: Checkerboard pattern has an even number of squares on one axis and uneven number on the other. This allows definition of a coordinate system CBI with its origin in the first corner within the image where the top-left square is black. For a right-handed coordinate system, the z-axis points in to the plane. CB tracked by the OTS is defined from the sphere closest to the origin of CBI.

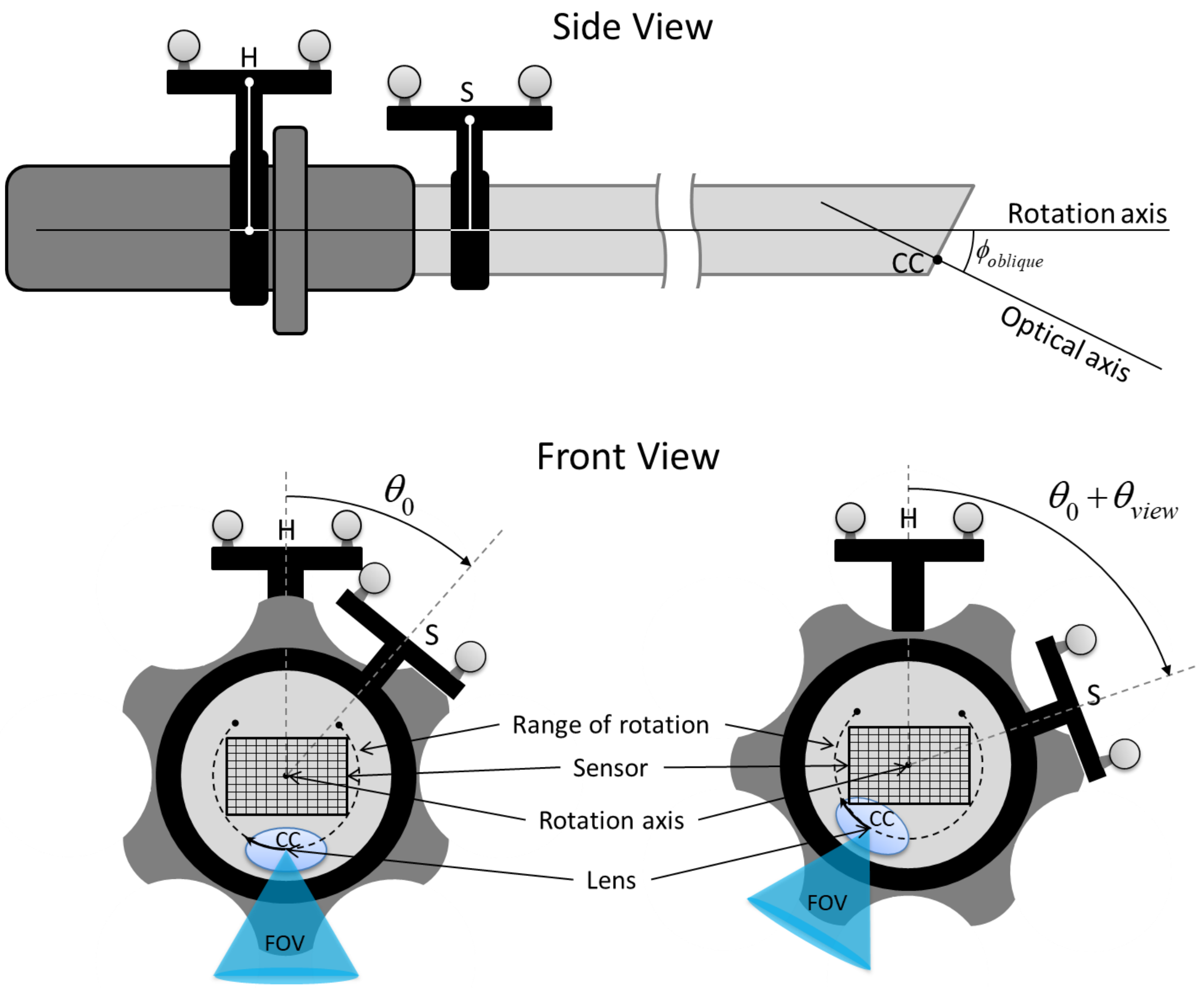


Figure 11: Model of the laparoscope used. Aim of laparoscope navigation is knowing the pose of the camera coordinate system CC in the OTS coordinate system. H and S are rigid bodies tracked by the OTS. H is attached to the handle and has a fixed relation to the camera sensor. S is attached to the scope and has a fixed relation to the orientation of the outer camera lens. θ_0 is the angle around the rotation axis between H and S at reference orientation. Other orientations of the outer camera lens are given by $\theta_0 + \theta_{view}$.

Rotation calibration

The two optical rigid bodies attached to the laparoscope, H and S given by the OTS as ${}^{OTS}\mathbf{T}_H$ and ${}^{OTS}\mathbf{T}_S$, are used to determine the orientation of the camera coordinate system, Figure 11. S can be expressed with respect to H by transformation

$$ {}^{H}\mathbf{T}_S = \left({}^{OTS}\mathbf{T}_H \right)^{-1} \cdot {}^{OTS}\mathbf{T}_S \tag{3.9} $$

By rotating the scope, the origin of S describes a circular path in the coordinate system of H, Figure 14. Rotation between the two parts of the laparoscope is given by the angle between the two tracking sensors around a rotation axis going through the center of the circle. Definition of the rotation axis requires a direction of the axis, and a point on the axis. The direction is given by the normal to the circular path described by S in H, and the center of the circle is used as the required point on the axis. The center and normal of the circle are found by solving a non-linear least squares problem using the Levenberg-Marquardt method as described in appendix C. The angle from H to S around the rotation axis is given by θ_r. Reference orientation of the laparoscope is defined as the oblique view aimed downwards $\theta_r = \theta_0$. Viewing angle θ_{view} of the scope with respect to the reference angle is given by

$$\theta_{view} = \theta_r - \theta_0 \tag{3.10}$$

Results are validated by comparing the measured angle to several manually set angles.

<u>Rotation angle dependent hand-eye modeling</u>
Rotation of the scope with respect to the handle changes the hand-eye transformation. This change requires modelling of ${}^{H}\mathbf{T}_{CC}$, ${}^{S}\mathbf{T}_{CC}$ as a function of the angle θ_{view}. As CC is defined by camera calibration, its orientation is defined in relation to the orientation of the image sensor (CCD). In a forward viewing scope, or any standard camera for that matter, the orientation of CC and CCD are very similar and can be described by a single fixed transformation. In the standard camera model; the x-, y-, and z-axis of CC are usually parallel to their counterparts in CCD. The only exception in orientation is the possible inversion of axes directions, Figure 12.

In the oblique viewing scope, the orientation of CC is changed to point the optical (z-)axis away from the z-axis of CCD. Rotation of the scope then results in precession of the optical axis around the rotation axis of the scope. During rotation, the orientation of CCD is still fixed with respect to the handle. As the axes of CC are defined by the orientation of CCD, rotation of the scope does not cause the orientation of the x- and y-axis to rotate around the rotation axis. Instead, the x-axis of CC rotates in a plane parallel to the yz-plane of CCD, and the y-axis of CC rotates in a plane parallel to the xz-plane of CCD. In other words, rotation of the scope causes a precession of the optical axis around the rotation axis but does not involve a rotation of CC around its z-axis, Figure 12. Technically, the motion is not a precession as there is no rotation around the z-axis, but it will be referred to as such.

Modeling of ${}^{H}\mathbf{T}_{CC}$ as a function of θ_{view} involves definition of the position of CC in relation to H. The position should either be fixed, or if the origin of CC does not lie on the rotation axis, describe a circular motion around the rotation axis of the scope. Modeling of the orientation is rather complex as none of axes has a fixed relation to H. As CCD is fixed in relation to H, the x- and y- axis of CC will each describe a motion in a plane of H during rotation. Correct definition of the x- and y-axis of CC as a function of the angle will yield a function that gives the precession of the z-axis.

S is assumed to have a fixed relation with the optical axis in the tip of the scope. Orientation modeling of ${}^{S}\mathbf{T}_{CC}$ requires rotation around the z-axis of CC in the opposite direction of the scope's rotation to correct the xy-plane of CC in relation to CCD. As the origin of CC is not expected to coincide with the rotation axis of the scope, the translation from S to CC describes a circle during rotation.

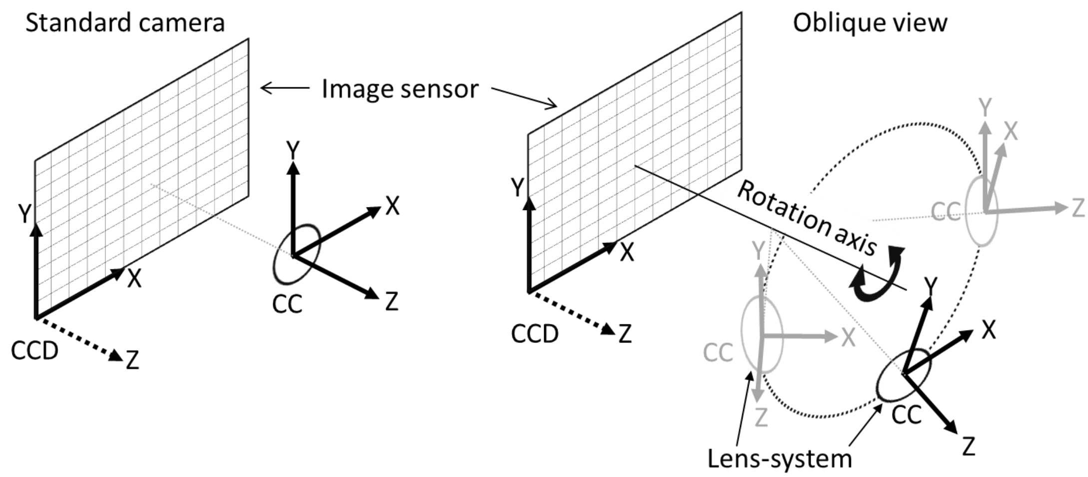


Figure 12: Coordinate systems within two simplified camera models where only the image sensor and outer lens are shown. CCD is the coordinate system of the image sensor with its x- and y-axis parallel to the pixels. CC is the camera coordinate system with its origin in the 'pinhole' of the camera. In a standard camera there is fixed relation between CC and CCD. In an oblique viewing laparoscope, the pose of CCD is assumed fixed in relation to the handle and does not change during rotation of the scope. The z-axis of CC is the optical axis of the camera and has an angle with the z-axis of CCD. During rotation, the optical axis describes a precession around the rotation axis as scope's cylinder. As CCD determines the orientation of CC, rotation of the scope does not result in rotation of CC around it's z-axis. In the models, CC is shown on the outer lens of the camera. However, the effective position of CC is located somewhere in the lens-system between the outer lens and the image sensor.

Inverse of the hand-eye calibration $\left({}^{CC}\mathbf{T}_H, {}^{CC}\mathbf{T}_S\right)$ expresses the position of H and S in relation to the origin of CC. If the hand-eye calibration can be modeled as a function of the rotation angle, the rotation angle between two orientations of S or H expressed in CC should have the same rotation angle as found with navigation for S expressed in H. The same strategy is applied as in rotation calibration to find the rotation axis of S/H expressed in CC. The rotation angle between hand-eye calibration poses is compared to the rotation angle found for S expressed in H.

Hand-eye calibration is performed at nine rotation angles of the scope ranging from -120° to +120° with increments of 30°. The results of hand-eye calibration for ${}^{H}\mathbf{T}_{CC}$ and ${}^{S}\mathbf{T}_{CC}$ are inspected to get an insight in which of the two can best be used, and how to model hand-eye calibration as a function of the rotation angle.

3.3 Results

Point set registration of ${}^{CB}\mathbf{T}_{CBI}$

A standard Mathworks checkerboard pattern is printed on a board for the experiments. The pattern has 10x7 squares and has a square edge length of 1.02 cm, Figure 10. Of the 54 corners in the pattern, 43 are measured with respect to the reference coordinate frame attached to the printed board using an optically tracked pointer. One of the corners is left out for visual inspection of the results, Figure 13. The root mean square error of registration is less than 0.4 mm for all 43 points.

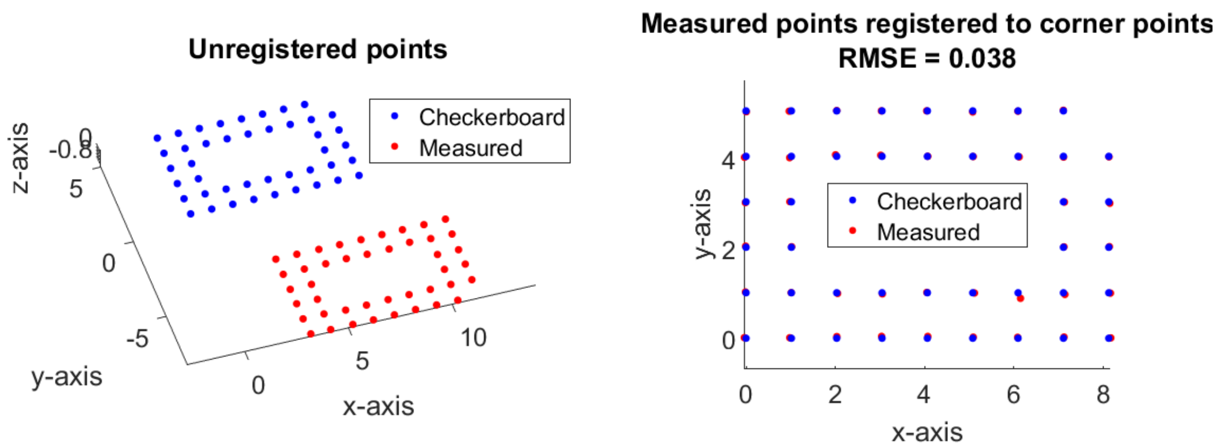


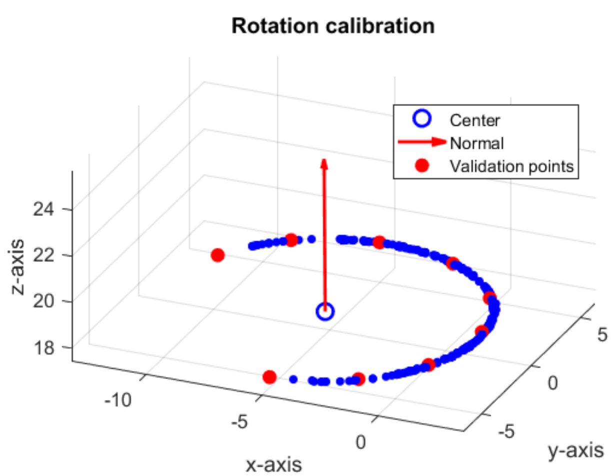
Figure 13: Registration of point clouds between CB and CBI. Red are the measured points in coordinates of CB. Blue are known points of CBI. The root mean square error of registration <0.4 mm.

<u>Rotation calibration</u>

The rotation axis defined by the normal and center of the rotation calibration results is shown in Figure 14. The nine angles used for hand-eye calibration are estimated based on the gear shaped grip on the handle. Five of the peaks on the gear correspond to the 0°, ±60°, and ±120° angles. The remaining angles correspond to the notches in between the peaks. These estimated angles are compared to the measured angles by setting the measured angle at 0° as the reference angle.

<u>Hand-eye calibration</u>

Pure translation and orientation components of CC in the coordinate frames of H are shown in Figure 15. The optical axis describes a precession similar to the rotation of the scope as expected. However, the x- and y-axis were expected to move in a plane during rotation as the sensor orientation should be fixed in relation to the handle. Translations are expected to be fixed, or describe a circular motion. This is clearly not the case as translations show two clusters in the data. One corresponding to the positive angles, and one to the zero and negative angles.



Estimated	Measured
120	120
90	93
60	62
30	31
0	0
-30	-28
-60	-61
-90	-88
-120	-118

Figure 14: Center and normal of circular path described by S in H. Nine orientations were recorded for which θ_{view} was estimated. The table on the right shows the measured angles using navigation for the estimated angles. Here, the measured angle at 0° was set equal to that of the estimated angle.

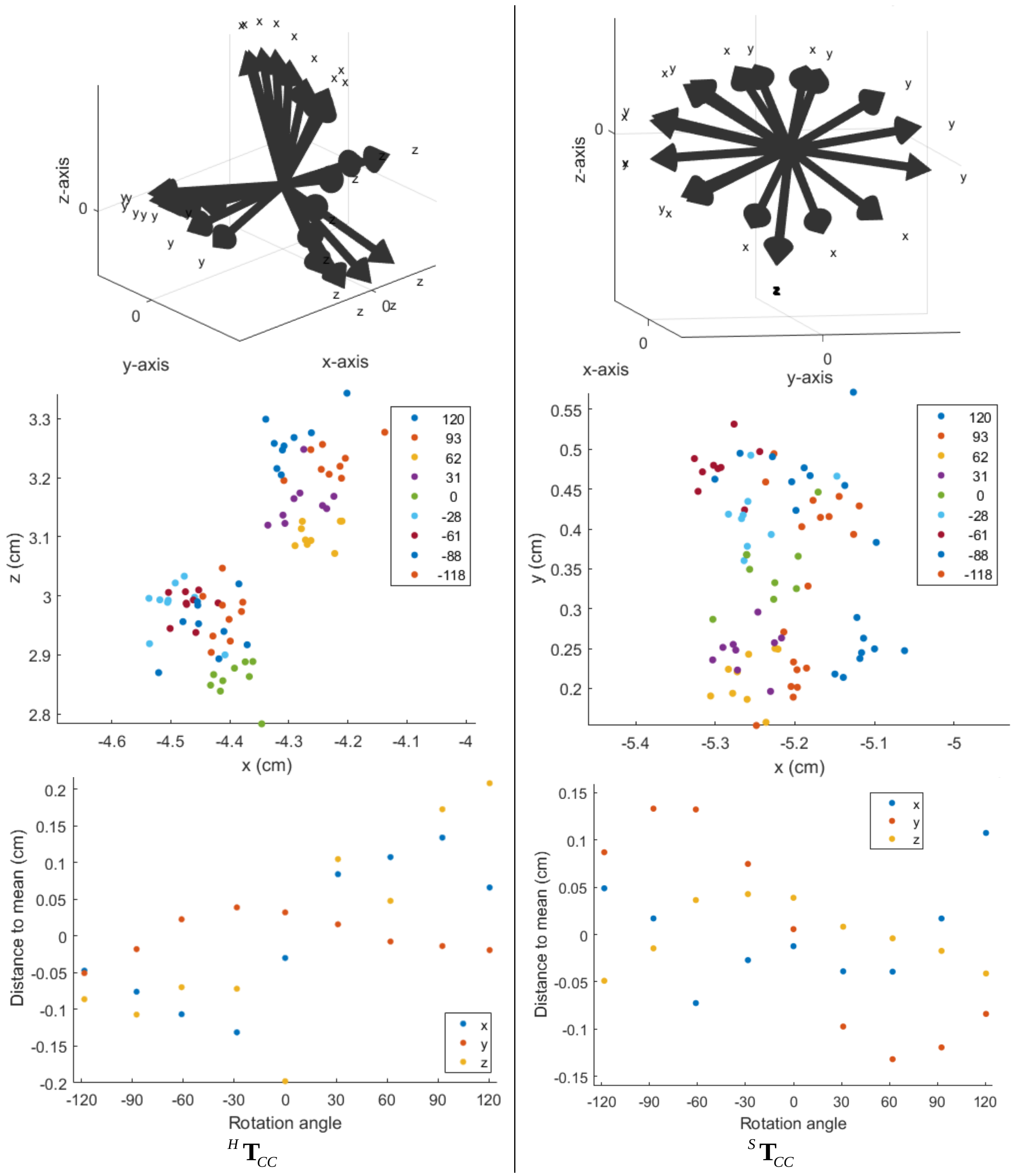


Figure 15: Hand-eye calibration results with reference to the tracking sensors attached to the handle (H) in the left column, and to the scope's cylinder (S) in the right column. Top row, hand-eye orientation components; middle row, scatter plot of translation components; bottom row, average translation component per axes and per angle shifted to zero mean. The optical (z-)axis describes a precession for H, and is fixed for S. All x- and y-axes almost lie in a plane for H. Translation components of H show two distinct clusters. The translation components resemble a circle with an indentation between -30° to 30° for S. Mean translations for individual axes show a sinusoidal motion during rotation in the y-axis. The other axes do not complete the second coordinate for a circular motion.

Translation and orientation of CC in the coordinate frame of S are shown in Figure 15. Orientation of the optical axis is constant with respect to S as expected. As there are six hand-eye calibration combinations that have a rotation angle difference of approximately 90°, the combined orientations should show collinearity of several of the x- and y-axes. However, there is a clear angular difference visible between the axes that are expected to be collinear. The translation components are expected to describe a circular path. There is clear sinusoidal motion visible in the y-axis, but the other axes do not complete the second coordinate to obtain a circular motion. The scatter plot resembles a circle with an indentation between -30° to 30°.

<u>Navigation versus inverse hand-eye based rotation angle</u>

Translations for the inverse of the hand-eye calibrations (from CC to H and S) describe a circle as expected, Figure 16. The rotation angle between different scope configurations is determined with the same rotation calibration method as used for rotation between the optical tracking sensors. Hand-eye calibration rotation angles for both methods deviate from the true rotation angle, as determined with navigation, with a maximum absolute difference of 7.8°, Table 2.

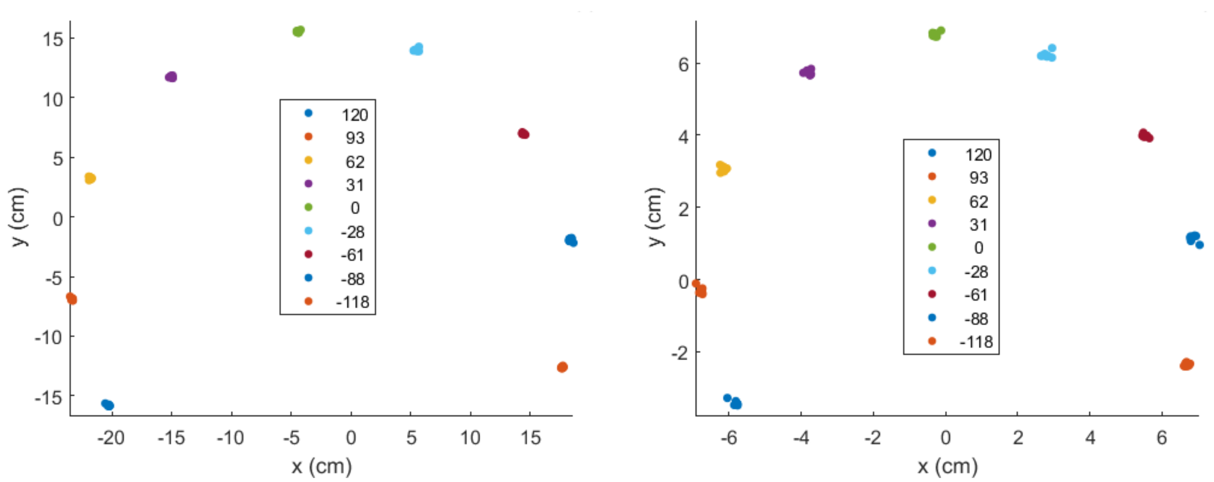


Figure 16: Positions of H (left) and S (right) in reference to CC from the inverse hand-eye calibrations. The positions describe a clear circular path during rotation. Applying rotation calibration to these coordinates gives the rotation between hand-eye coordinates with reference to 0°, Table 2.

Navigation	Hand-eye Handle	Difference	Hand-eye Scope	Difference
120	113	-7.8	117	-3.6
93	87	-5.8	89	-3.5
62	60	-2.3	60	-1.7
31	31	-0.1	31	-0.3
0	0	0	0	0
-28	-29	-0.4	-26	2.0
-61	-62	-0.8	-56	4.6
-88	-89	-1.3	-83	4.6
-118	-117	1.1	-112	6.1

Table 2: Rotation angles obtained by navigation, rotation in hand-eye calibration, and the difference between the two. In both cases 0° is set as the reference angle. In both cases the rotation angle of hand-eye calibration deviates from the true/navigation based rotation.

3.4 Discussion

Point cloud registration using the Procrustes algorithm has an RMSE of <0.4 mm. As the RMSE of a single reflecting sphere as detected by the OTS can be up to 0.25 mm, the current results are at the limit of what is possible with the used method. Visible inspection of the results show a good match between the two point clouds with only one clear mismatch caused by a measurement error. Removal of the incorrectly measured point lowers the RMSE to 0.033 cm, but does not change the registration results. The method therefore produces accurate results and is robust to at least one outlier.

Rotation calibration produces angles consistent with the input angles. In comparison to the reference angle, six out of eight evaluation angles have an angle 1-3° larger than the input angle. This suggests that reference angle 0° should be 1°. As this is just an offset in reference angle, it does not change the accuracy of rotation calibration. This method is already shown to be accurate and the results are therefore only to validate correct implementation. As there is no pattern visible in angular differences, the inaccuracies are attributed to errors in the estimated angles by manual positioning.

Hand-eye calibration results differ from what is to be expected based on the information available on the laparoscope. One possible explanation could be a defect in the laparoscope. During the experiments a strong unnatural resistance was felt when rotation the scope around 0°. This resistance was not present in other laparoscopes used during the initial experimental stage of which no results are shown here. When rotating around the reference angle the scope's cylinder felt to be forced downward in the reference angle direction. This defect would also explain the two clusters in hand-eye calibration translations with respect to the handle. During measurements, first the positive angles were measured, then the scope was rotated back to the reference angle and then on to the negative angles. At the positive angles the scope's cylinder was in one configuration with respect to the handle, while rotating it back over the reference angle put it in another configuration. This could also be the cause of the circle indentation in scope translation, but the mechanisms of this are not understood, yet. Results obtained from these experiment suggest that hand-eye calibration can best be performed in relation to a sensor attached to the scope's cylinder as it produces a simpler model, is closer to the scope's tip, and its relation to the camera coordinate system is less influenced by tilting between the handle and cylinder. The proximity to the scope's tip is of benefit since a small angular error in orientation estimation can result in a large error of a distant point expressed in the reference frame. As the reference frame on the scope is closer to the camera coordinate system than the handle, this arm effect of the orientation error is smaller.

In the current understanding of the laparoscope it is assumed that the orientation of the camera sensor is fixed in relation the handle. Hand-eye calibration orientations suggest that this is not the case. The deviation of the x- and y-axis during motion from a plane in the hand-eye calibration from the handle show that the sensor orientation is changed with respect to the handle. Rotation differences between navigation and those extracted from the inverse hand-eye calibration suggest the same. The camera in the tip of the scope is presumed to be magnetically coupled to the handle [33]. It is possible that this coupling does not completely fix the orientation of the sensor and the orientation is affected by rotational friction in the scope resulting in hysteresis. If this is the case, it needs to be investigated whether this rotation offset is a defect, or that its behavior is predictable. If the behavior is unpredictable, it is very difficult to accurately determine the pose of CC based on tracking of the outside of the scope alone.

3.5 Conclusion

Point cloud registration of the checkerboard pattern and rotation calibration methods are performed using optical tracking. Both provide results close to what can be expected with the accuracy of optical tracking. These results are sufficient for our purpose, and can be implemented as is for navigation and subsequent projection of the patient model on the laparoscopic images.

Hand-eye calibration does not behave as expected. Rotation of the scope does not have the same rotation in orientation of the camera. As a result, the orientation of the image plane does not correspond with what is to be expected based on navigation of optical trackers attached to the laparoscope. Further research is needed to investigate if this is due to the specific laparoscope used for these experiments, or if this is a property of the Olympus EndoEye HD. This study should also include experiments to investigate if the angular offset is predictable, and thus can be included in the model for hand-eye calibration. These questions are looked in to in chapter 5.

Translation from the laparoscope attached reference frames to the camera coordinate system does not behave as a rotation of the scope's cylinder around a fixed rotation axis through the handle. Here, the behavior is expected to be due to a defect in the particular laparoscope used. To verify this, experiments will be repeated with another laparoscope in chapter 5. To eliminate effects of variable configuration between scope's cylinder and handle, and to minimize the effects of errors in pose estimation of the reference frame, the optical reference frame attached to the scope's cylinder is most suited for hand-eye calibration.

CH4: Delay estimation

4.1 Introduction

Real-time visualization of objects in the laparoscopic images requires input from two different sources. Laparoscopic images are acquired by a frame grabber connected to the computer, while position and orientation data is acquired by means of an optical tracking system (OTS). Data from both systems needs to be synchronized to produce a visualization that accurately displays motion of the projected anatomy within the images. If there is a delay between the data streams, the projected anatomy will either move ahead or lag behind the anatomy seen through the laparoscope. If we assume the delay to be constant it is possible to delay one of the systems to synchronize with the other.

Sampling rate of both systems is 30 Hz, meaning that the time between measurements is ~33 ms. Cross-correlation is the easiest way to estimate the delay between two signals. However, since we are working with discrete time signals, cross-correlation can provide accurate results up to half the sampling interval. With a sampling interval of ~33 ms this would result in an accuracy up to ~17 ms. Since we are aiming for an accuracy of less than 5 ms, cross-correlation is not a suitable option. The proposed method for delay estimation makes use of an object that is rotated at a constant angular frequency to allow fitting of a sinusoidal function to both of the coordinates. The delay between both systems can be estimated from the phase difference between the fitted sinusoidal functions.

4.2 Method

Delay estimation setup consists of an object that can be tracked by both systems, in this case a rigid body for the OTS, and can be rotated at a constant angular velocity around a fixed point, Figure 19. Rotation of the frame can be described as a 2D circular motion around the center of rotation, Figure 17. Stationary position of the object can be expressed as

$$\mathbf{p} = \begin{bmatrix} p_x \\ p_y \end{bmatrix} = \begin{bmatrix} r \cdot \cos(\theta_s) + c_x \\ r \cdot \sin(\theta_s) + c_y \end{bmatrix} \tag{4.1}$$

Here, $\mathbf{p}$ is the position of the object, r the radius of circle, θ_s the stationary phase relative to the positive x-axis, and $\mathbf{c}$ the center of rotation. If the axes of the coordinate systems in the image and OTS are aligned, the stationary phase will be equal for both. However, if the axes are not aligned the phases will differ between both systems with an angle equal to the rotation between the coordinate systems.

$$\Delta\theta_s = \theta_{s,I} - \theta_{s,OTS} \tag{4.2}$$

Here, $\Delta\theta_s$ is the phase difference between the image and optical tracking system which are denoted with subscripts I and OTS respectively.

If the object is rotated at a constant velocity, the position can be described as a function of time by

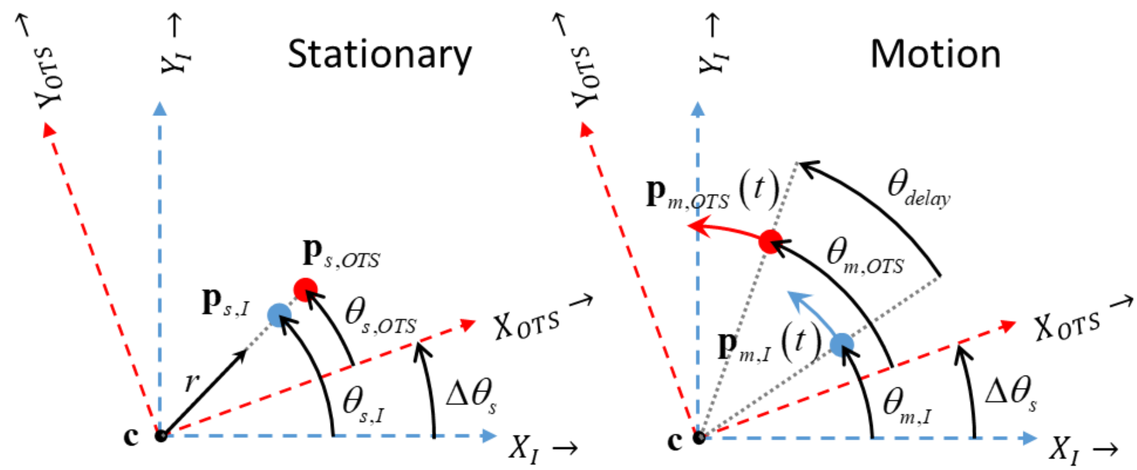


Figure 17: Phase differences between modalities of a rotating object. Phase difference during motion is a measure for the amount of acquisition delay between the image and OTS. The phase difference in motion needs to be corrected for the rotation between coordinate systems of both modalities before the delay can be estimated. This rotation is given by the phase difference when there is no motion.

$$\mathbf{p}(t) = \begin{bmatrix} r \cdot \cos(\omega \cdot t - \theta_m) + c_x \\ r \cdot \sin(\omega \cdot t - \theta_m) + c_y \end{bmatrix} \tag{4.3}$$

Here, ω is the angular velocity, t the time, and θ_m the phase where subscript m denotes motion. If there is no acquisition delay between the OTS and the image, the phase difference between modalities in motion is the same as the stationary phase difference. If these phase differences are not equal it is possible to find the delay between both systems from the difference in phase differences.

$$\Delta\theta_m = \theta_{m,I} - \theta_{m,OTS} = \Delta\theta_s - \theta_{delay} \tag{4.4}$$

Here, $\Delta\theta_m$ is the phase difference in motion between the image and OTS, and θ_{delay} the phase delay of OTS with respect to the image. Phase delay is given by the difference between phase differences.

$$\Delta\theta_s - \Delta\theta_m = \theta_{delay} \tag{4.5}$$

Since we are interested in a time delay, and not a phase delay this needs to be converted to time

$$\Delta t = \frac{\theta_{delay}}{\omega} \tag{4.6}$$

Parameter estimation

To estimate the delay we need to find $\theta_{s/m}$, ω. If the object is in motion, this can be achieved by fitting a sinusoidal curve to the x or y coordinate of the object versus time, Figure 18.

$$f(t) = r \cdot \sin(\omega \cdot t - \theta_m) + c \tag{4.7}$$

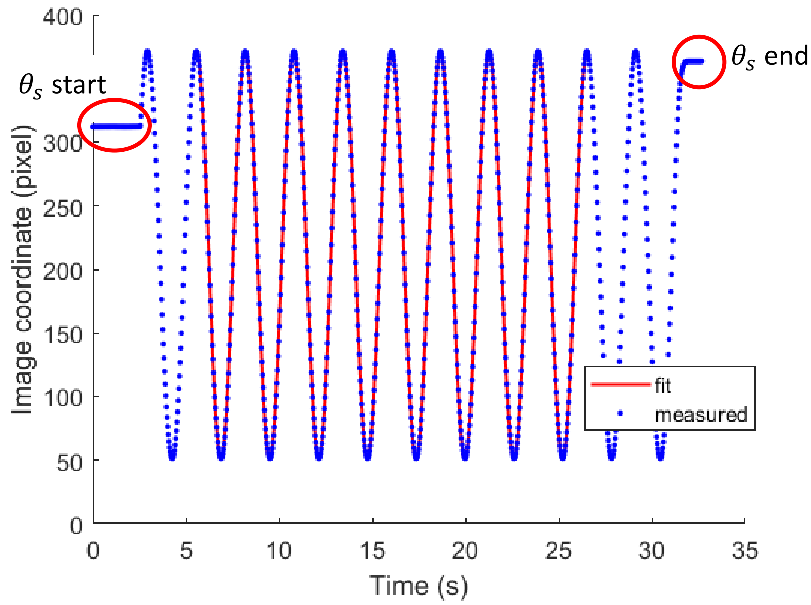


Figure 18: Sinusoidal curve fitting for the coordinates of one of the axes in the image. Circles indicate the stationary phase before and after motion.

Fit results give an estimate of θ_m and ω. Using the center and radius of the fit, and the stationary position $\mathbf{p}$ the stationary phase can be estimated by

$$\theta_s = \sin^{-1}\left(\frac{f-c}{r}\right) \tag{4.8}$$

To perform the procedure described above, we first need to estimate the positions $\mathbf{p}(t)$ for both modalities. The OTS already gives us a position in 3D space of the object. For the camera we need to find the position of the object in each image. The procedure can be split in four parts, data acquisition, pre-processing, position estimation, and delay estimation. Delay estimation is already discussed above, the other three will now be explained in the order they occur in.

Data acquisition

Acquisition for both modalities is straight forward. The OTS needs to be placed at a position where it can see the optical frame during the entire rotation. The camera, in this case a webcam, is positioned with the optical axis parallel to the rotation axis of the object, Figure 19. This is needed to prevent oblique projection of the circular motion turning it in to a spherical trajectory. The data for both modalities is acquired and time stamped by Plus Server via OpenIGTLink [47, 48].

Pre-processing and position estimation

<u>Optical tracking system</u>

The only pre-processing step required for the 3D OTS data is eliminating one dimension to get a 2D circular motion. The OTS gives a position and orientation for each point of the rigid body during motion. By finding the center of a rotating point it is possible to get the axis of rotation, see appendix C. By aligning the axis of rotation with the z-axis of the OTS, the plane rotation takes place in is placed parallel to xy-plane of the OTS, allowing removal of the z-axis to get the 2D coordinates of the rotating point.

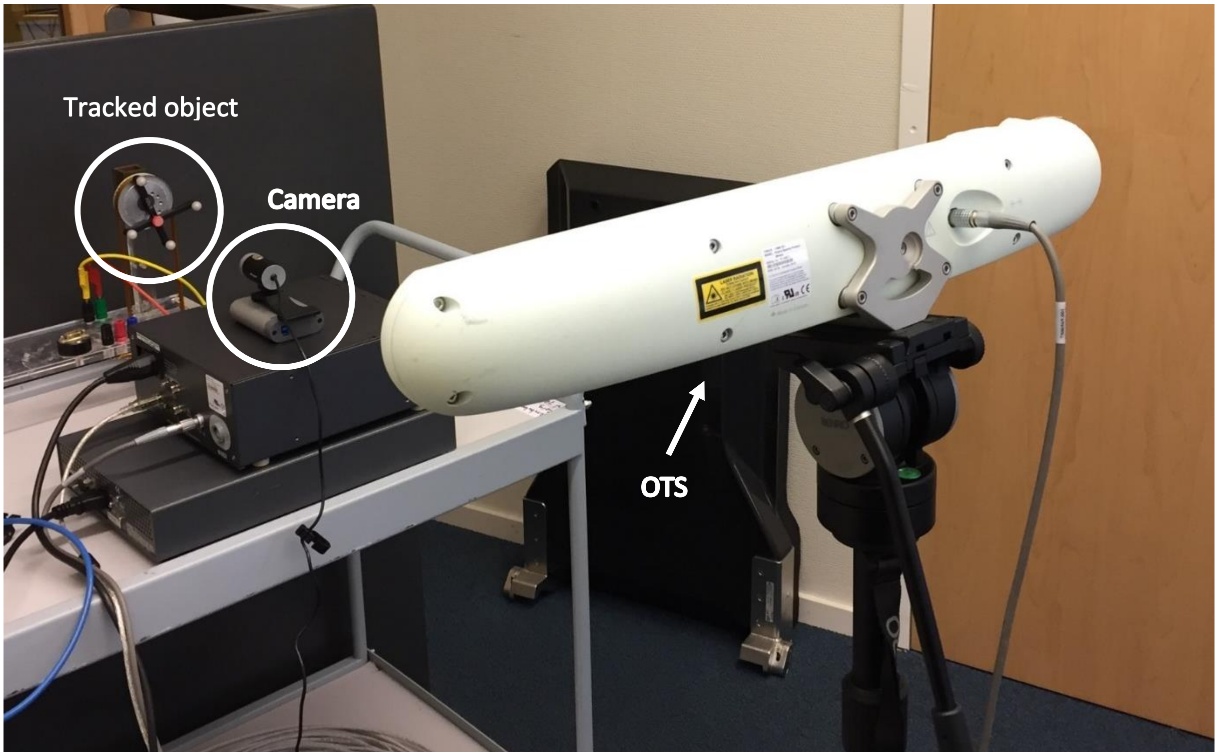


Figure 19: Experimental setup. Tracked object is connected to a motor for rotation and placed in front of a dark background to prevent the reflecting spheres from disappearing in the background. A webcam is placed in front of the object with the optical axis parallel to the object's axis of rotation. The OTS is placed where it can track the object during the entire rotation.

Image

Pre-processing of the images consists of correction for lens distortion, converting to grayscale, and defining a region of interest (the predicted trajectory of the object) where the object can be found, Figure 20. This last step is taken to reduce the number of specular reflections in the image that can simulate the object resulting in a wrong estimate of the position, and to speed up processing time.

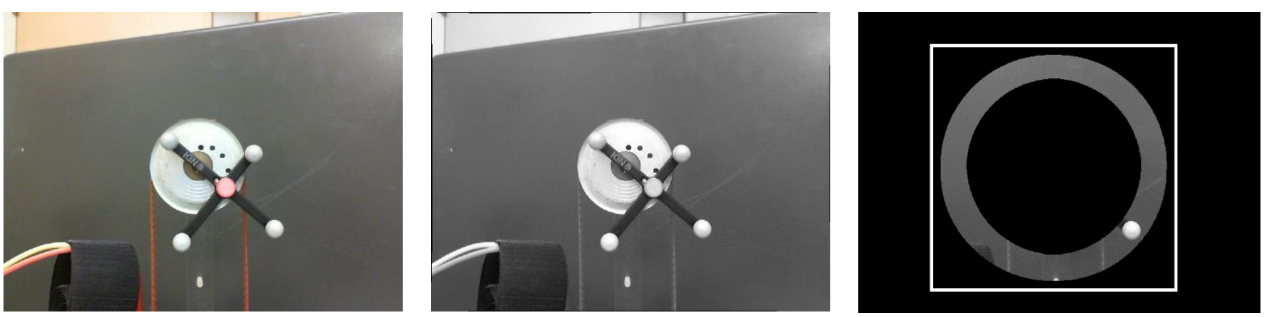


Figure 20: From left to right, original image, grayscale and lens correction, and masking and cropping of the image. As can be seen in the last image, the bottom right sphere is the sphere that is tracked. The coordinates of the object are expressed by the OTS with respect to the same sphere.

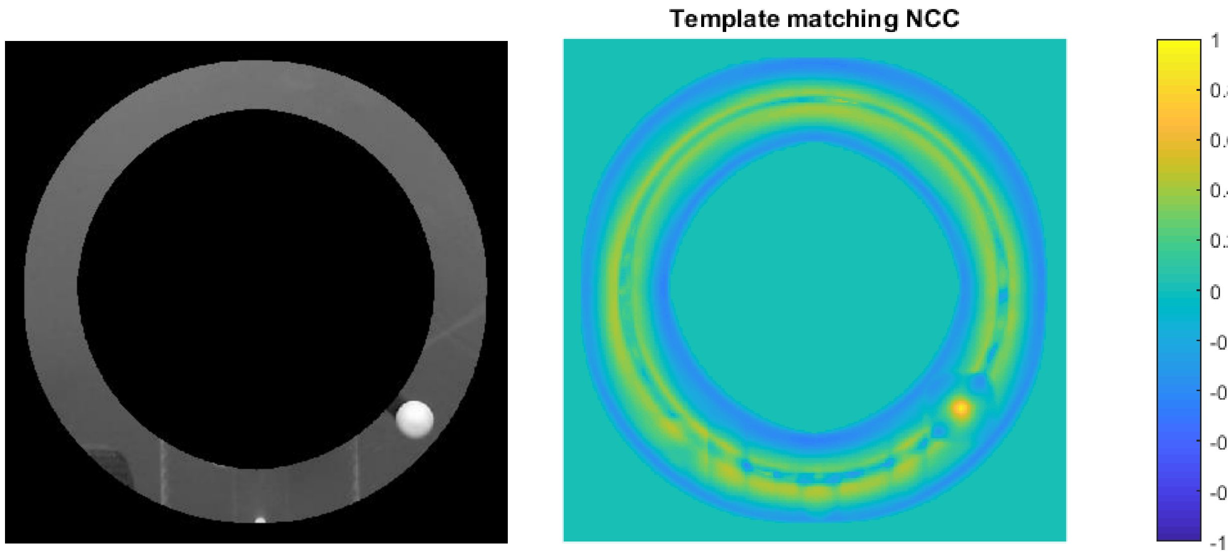


Figure 21: Left, cropped image with tracked object and masked expected trajectory of motion. Right, NCC map. High values indicate the template is very similar to the image at that position, while negative values indicate the opposite. Cross-correlation cannot be performed on a patch with constant values, as is the case for the masked out area. If the patch is constant, NCC returns zero.

The position of the object, in this case the sphere furthest away from the center of rotation, is found by template matching using normalized cross-correlation (NCC). In NCC the template is slid over the image to create a map that shows the correlation between the template and the current position on the image. In this map we can find the position of the best match between template and image. The best match is assumed to be the location of the sphere being searched Figure 21. Sphere position is updated by quadratic curve fitting to get a sub-pixel position estimation.

4.3 Experiments

4.3.1 Curve fitting

Curve fitting is evaluated by performing five measurements and fitting over five periods (rotations) of the moving object in each measurement. In starting and stopping of the rotation the angular velocity will not be constant. To account for this changing velocity an extra period is recorded at the start and end of motion that is excluded from curve fitting. Data is recorded for a short period before and after motion to estimate the phases in rest using the parameters found by curve fitting. Fit results are shown in Table 3.

Camera and OTS *radii* fits are very consistent with a standard deviation < 0.1% of the radii. Since the object describes a circular motion, the radii in x- and y-direction are expected to be of equal length. This is the case for OTS, but the radius found in the y-direction of the camera is slightly shorter than the one found in the x-direction. This is probably due to placement of the camera with respect to the rotating object. If the camera's optical axis is not parallel to the object's axis of rotation, the circular motion is projected as an ellipse on the image plane. The ellipsoid can be corrected back to a circle. However, since the difference is in the order of 0.1% of the radius no further processing was performed.

The *angular frequency* for the last two measurements is different from the first three. This is likely due to limitations of the equipment used for rotating the object. This change in frequency makes it difficult to compare angular frequency between measurements. However, since the camera and OTS captured the motion at the same time, their angular frequencies should be the same within a measurement. The difference in angular frequency between both modalities is less than a milliradian per second for each measurement. With a standard deviation for the measurement differences of 0.02% of the average angular frequency, it very likely that the difference between measurements is due to the equipment and not fit inaccuracy.

Phases in motion are shown as a *phase difference* between the camera and OTS. Since the phase difference in motion depends on angular frequency it is not useful to compare radial phase differences over measurements. Dividing the phase difference by the measurement's angular frequency removes this dependency and gives a more meaningful result. For both of the axes, the phase difference in motion shows a delay of the OTS compared to the image of 13.5 ms with a standard deviation of less than 2 ms. Delay difference between x- and y-axis should the same within one measurement. Comparison of the intra and inter measurement standard deviation (0.7 ms vs 1.7 ms) suggests that the delay is not consistent over all measurements. This could be due to how the data is processed in Plus Server, or small timing differences at the start acquisition. However, with only five measurements, statistics are not strong enough to draw any real conclusions. If there is a delay difference between measurements, the data suggests that it is in the order of a ms and therefore acceptable to our cause.

Coordinates for the *center of rotation* are consistent between measurements for the camera and OTS with a standard deviation in the order of 0.1% of the radius. Camera standard deviation of the y-axis is three times larger than that in the x-direction. This is likely due to the same ellipsoid projection as in radius fit. Ellipsoid shape cannot be perfectly fitted by a single sinusoid if the major and minor axis of the ellipsoid are not parallel to the x- and y-axis used for fitting. As a result there is a small spread in the

Measurement #	Source	1	2	3	4	5	Mean ± SD
Amplitude / Radius [pixel, mm]	I_X	160.1	160.0	159.9	160.0	160.0	160.0 ± 0.07
	I_Y	159.8	159.9	159.9	159.7	160.0	159.8 ± 0.12
	OTS_X	63.6	63.6	63.6	63.5	63.5	63.6 ± 0.02
	OTS_Y	63.6	63.6	63.6	63.6	63.6	63.6 ± 0.02
Angular Frequency [rad/s]	I_{XY}	2.40	2.40	2.40	2.44	2.46	2.42 ± 0.03
	OTS_{XY}	2.40	2.40	2.40	2.44	2.46	2.42 ± 0.03
	I-OTS	-0.0004	0.0003	0.0008	0.0007	0.0003	0.0003 ± 0.0005
Phase Difference [mrad]	I_X-OTS_X	-25.3	-32.1	-32.4	-38.4	-35.6	-
	I_Y-OTS_Y	-25.9	-33.2	-32.1	-35.5	-37.3	-
Phase Difference [ms]	I_X-OTS_X	-10.5	-13.4	-13.5	-15.7	-14.5	-13.5 ± 1.9
	I_Y-OTS_Y	-10.8	-13.8	-13.4	-14.5	-15.2	-13.5 ± 1.7
	X-Y	0.3	0.4	-0.1	-1.2	0.7	0.0 ± 0.7
Offset / Center of Rotation [pixel, mm]	I_X	211.0	210.9	211.0	210.9	210.9	210.9 ± 0.05
	I_Y	208.9	208.9	208.6	208.8	209.1	208.9 ± 0.16
	OTS_X	291.4	291.4	291.4	291.4	291.5	291.4 ± 0.03
	OTS_Y	10.1	10.2	10.1	10.3	10.1	10.2 ± 0.07

Table 3: Curve fitting results for 5 measurements. Origin of the data is denoted by the sources I and OTS that represent image and optical tracking respectively. Source subscripts refer to the axes the data is from. Due to differences in angular frequency the radial phase differences cannot be compared between measurements. Instead, phase differences are compared in milliseconds.

center coordinates between measurements. Standard deviations of the OTS show a similar difference. In pre-processing the OTS data, coordinates are transformed from 3D to 2D by aligning the axis of rotation with one of the axes of the coordinate frame, followed by removal of that axis. In the results shown here, the axis of rotation is only estimated in the first measurement and used to pre-process all five measurements. However, when the axis of rotation in the OTS data is estimated for each measurement independently, the standard deviation increases to 1-2% of the radius. Incorrect estimation for the axis of rotation results in an ellipsoid shape of the 2D motion. This ellipsoid shape is the reason OTS shows a similar variation as the camera for center coordinates. The increase in spread, if pre-processing is applied for each measurement independently, shows that the rotation axis estimation is not robust.

4.3.2 Stationary phase differences

With the parameters found in curve fitting the stationary phases can be estimated using (4.8). The coordinates at start and end are averaged over ten samples to limit the effects of measurement noise. Due to the ellipsoid shape of the rotation by processing or acquisition, or due to measurement noise it happened several times that the radius was larger than the radius found by curve fitting. As a result, the fraction in the inverse sine could become larger than one and returned a complex phase. The atan2 function was therefore used to find the phases at the start and end of a measurement to solve this problem. Another option to solve this problem was recalculating the radius from the center coordinates and coordinates found at the start and end of a measurement. Since both options will give the same answer, using the atan2 function is computationally more effective.

Phase differences and radii at the start and end of a measurement can be found in Table 4. In the absence of motion, the phase differences are independent of a measurement's angular frequency. Phase differences can therefore best be compared in radians, or in this case milliradians. The table also shows the phase difference in milliseconds calculated using the measurement's angular frequency. Phase differences in milliseconds can be used to compare the phase differences with those of curve fitting.

The setup was not changed between measurements and therefore the setup at the end of one measurement was the same as the start of the next. Looking at the phase differences we can clearly see this relationship between measurements. The relationship is also visible in the radius of the camera. A

Measurement #		1	2	3	4	5	Mean ± SD
Phase Difference [mrad]	Start	19.3	20.0	2.2	11.9	3.5	11.2 ± 8.0
	End	18.6	-1.4	13.4	6.5	18.5	
Phase Difference [ms]	Start	8.0	8.4	0.9	4.9	1.4	-
	End	7.8	-0.6	5.6	2.7	7.5	
Camera Radius [pixel]	Start	161.6	161.3	158.8	161.0	158.3	160.2 ± 1.5
	End	161.3	158.8	161.3	158.2	161.7	
OTS Radius [mm]	Start	63.6	63.4	63.8	63.7	63.5	63.7 ± 0.2
	End	63.5	63.9	63.9	63.7	63.5	

Table 4: Radii and phase difference at start and end of measurements. Radii are consistent for OTS over all measurements with a standard deviation of 0.2 mm, the accuracy that can be expected from the OTS. Variation of camera radii are slightly larger if compared to the radius. There is a clear correlation visible between the camera radius and phase difference. A shorter radius has a smaller phase difference and vice versa. Stationary phase differences are independent of angular frequency and therefore shown in milliradians.

shorter radius results in a lower phase difference while a larger radius results in a larger phase difference. Changes in radial magnitude are due to elliptical trajectory found of the moving object. Radial length at start and end depends on their position in the ellipse. For the OTS, the same relationship between a measurements start and end radius is visible, but it doesn't seem to have any correlation with the phase differences. In contrast to the fit parameters, the elliptical shape of the motion does have an influence on the phase differences at start and end of motion. To reduce this effect the shape of the motion can be corrected back to a circle.

4.3.3 Delay estimation

From the phase differences between rest and motion the delay in milliseconds is estimated using (4.5) and (4.6), Table 5. There is still a relationship visible between the end delay of a measurement and the start delay of the next measurement. With a standard deviation of 3.1 ms the variation is well within an acceptable range for our purpose. Within a single measurement the delay should be the same comparing start to end. In the curve fitting results of phase difference in motion we saw that the difference within a measurement was less than a ms, while it increased to 6 ms comparing the estimated delay at start and end. This increase in variation is due to stationary phase estimation, and is by far the largest source of variation in the delay estimation process.

<table>
<tr><th colspan="2">Measurement #</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>Mean ± SD</th></tr>
<tr><td rowspan="3">Delay [ms]</td><td>Start</td><td>18.5</td><td>21.7</td><td>14.4</td><td>20.6</td><td>15.9</td><td rowspan="2">18.2 ± 3.1</td></tr>
<tr><td>End</td><td>18.5</td><td>13.2</td><td>19.0</td><td>17.2</td><td>22.7</td></tr>
<tr><td>Start-End</td><td>0.0</td><td>8.5</td><td>-4.6</td><td>3.4</td><td>-6.8</td><td>0.1 ± 6.1</td></tr>
</table>

Table 5: Estimated delay for five measurements using stationary phases at start and end of motion, and the difference between estimated delay in a single measurement.

4.3.4 Influence of template selection

Template selection is the only manual part in the delay estimation process. To evaluate the effects of template selection, five templates are used to get the coordinates from each of the five measurements. Templates are selected from the first frame from each of the five measurements, Figure 22. The five acquired sets of coordinates for each measurement are used for curve fitting and estimation of the stationary phases.

Standard deviations of radius, angular frequency, and phase difference within a measurement for curve fitting results (not shown) are an order of magnitude smaller than the standard deviations between measurements using a single template as stated in Table 3. Template selection can therefore be assumed to have negligible effects on these parameters.

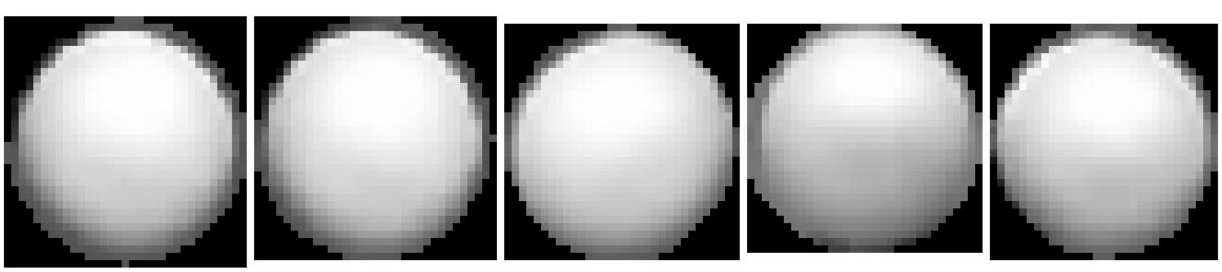


Figure 22: Templates selected from the first image of each of the five measurements. All five templates are used on all five measurements to evaluate the effects of template selection on the coordinates of the moving object in the image.

Template selection does have an influence on the coordinates found of the rotating object as it can have a fixed translation for all coordinates from one template to another. In Figure 23 the effects on the found coordinates of the stationary positions at start and end, and the center of rotation can be seen. The mean for start, end, and center of rotation coordinates is set to zero within a measurement to get all coordinates in the same window. It is clear that template selection has a significant influence on the coordinates found. It is also clear that the influence is constant over all measurements. By subtracting the center of rotation from the start and end coordinates this bias can be removed, Figure 24. Subtracting the center of rotation does not fully remove the bias as there is still a pattern visible, but the spread is now less than 0.1 pixel in all directions, or 0.1% of the radius, the same order as the standard deviation of the radius. As template selection has no influence on the angular frequency and phase difference in motion, and since the influence it does have on the parameters for stationary phase difference is self-correcting, the overall effect of template selection on delay estimation seems to be negligible based on the performed experiments.

4.3.5 Validation

Validation is performed by introducing a delay between the camera and OTS data. Plus server allows the addition of a delay to one of the sources to synchronize the data. This functionality is used create a delay between both sources and test the delay estimation process. For validation, delays of 0 (no added delay), 20, 50, 100, and 200 ms are added to the OTS data. Each delay is measured five times, Figure 25.

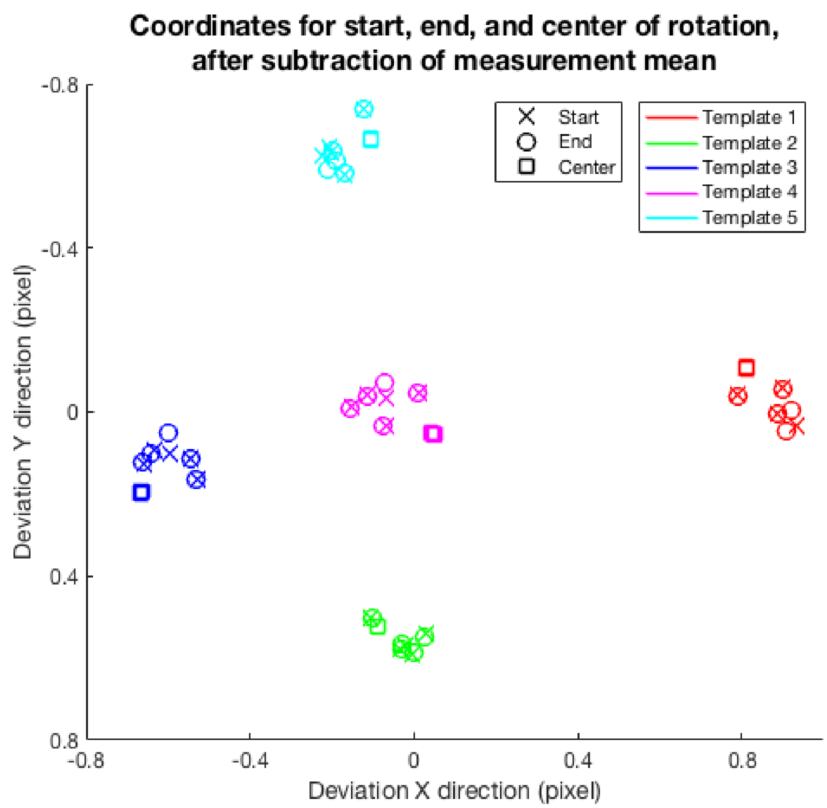


Figure 23: Coordinates for the object before and after motion, and center of rotation from curve fitting for five measurements using five templates in each measurement. All coordinates are shifted to the same window by subtracting the measurement mean from the parameters. The selected template has a clear and predictable influence on the coordinates found.

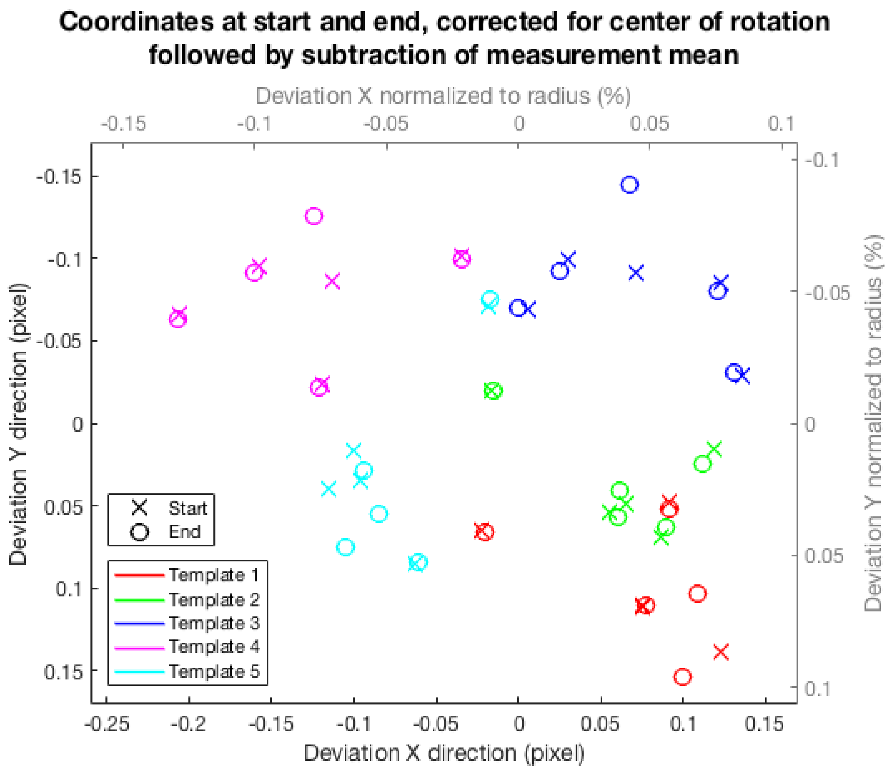


Figure 24: After subtraction of the center of rotation from the start and end coordinates of the object the bias is mostly removed. Remaining variation is in the same range as the radius standard deviation.

Variation in the estimated delays is larger for delays of 20, 50, and 200 ms compared to that of no added delay. This could be because a delay of 100 ms is a delay of exactly 3 data points at an acquisition frequency of 30 Hz, not requiring any interpolation of the data. If this were the case, the same would be expected of the spread at 200 ms, which is a delay of exactly 6 data points. No conclusive answer can be given on the cause of the low spread at 100 ms, if there even is a cause except chance, but it is worthwhile to investigate if a better approximation is needed.

Trend lines are fitted for the delay with reference to the start and end of measurements. Both trend lines show a slope of nearly one with an offset of just over 15 ms. Even though there is a spread in the estimations, the delay estimation process is able to accurately find the added delay in the measurements.

Inspection of the phase difference based on motion alone suggests that phase fitting is the cause of spread in the estimated delay, Figure 26. As the standard deviation of differences between phase differences of a single measurement is less than a millisecond, it is also possible that the true delay is variable.

4.4 Discussion

Ellipsoid shape of the motion does not have a relevant influence on the fitted parameters. It does influence the phase estimation before and after motion and is in the current method by far the largest source of variation in delay estimation within a measurement. If the current delay estimation is to be improved, the phases in rest need to be estimated more accurately.

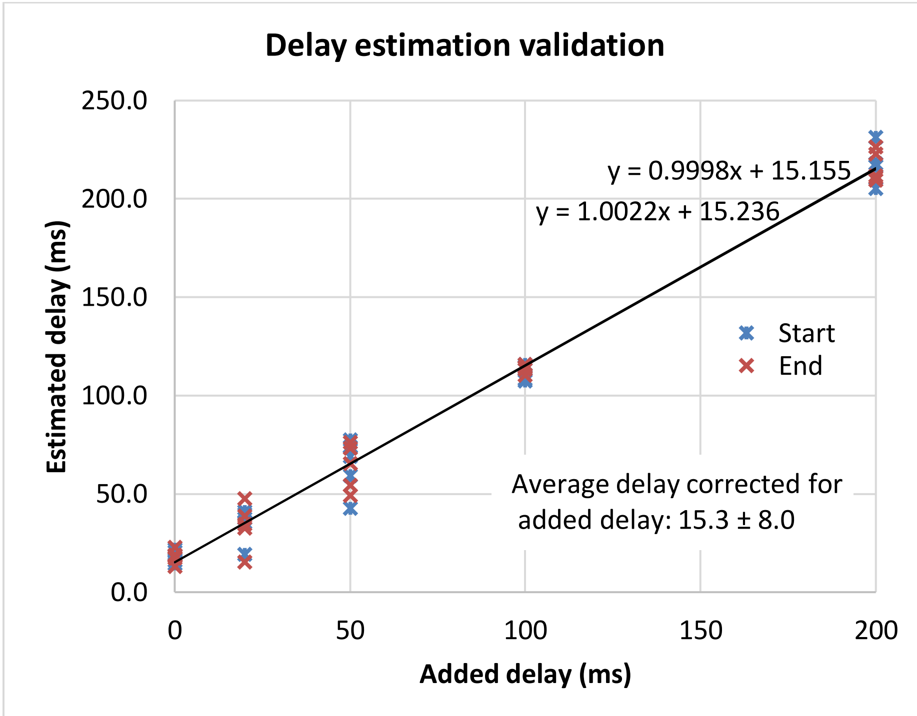


Figure 25: Delay estimation with an added delay of 0, 20, 50, 100, and 200 ms, and five measurements per added delay time. The spread for measurements with an added delays is larger than those without. Fitted trend lines have a slope of nearly one meaning the delay estimation process has accurately estimated the delay.

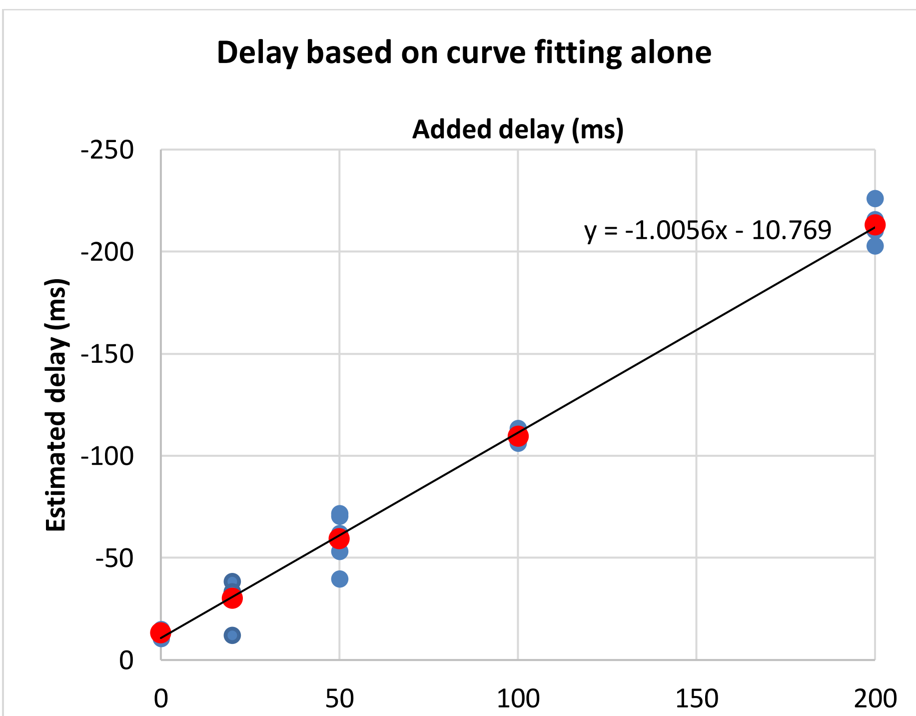


Figure 26: Delay based on phase difference during motion alone shows the same spread as the delay estimation corrected for stationary phase difference. This suggest that phase fitting is inaccurate, or that the true delay is variable.

Until now, the variations in estimation of radius, center coordinates, and phases in rest are presumed originate from ellipsoid shape of the registered motion due to incorrect rotation axis estimation, or projections under an angle on the image plane. Ellipsoidal motion can explain some of the variation seen in the measurements, but it does not account for the relatively large standard deviation in estimated radii for the stationary phases in the images. Standard deviations of fitted radii and center coordinates are in the same range for the OTS and camera if expressed as a percentage of the modality's corresponding radius. The same goes for radii found in stationary phase estimation for the OTS. The relatively large variation in radii for the phases in rest of the camera can be an effect of the template used. During template selection a circle is placed around the object to be tracked. In the current implementation of the normalized cross-correlation it is necessary to have a rectangular template. To fulfill this requirement, the selected patch is made rectangular by filling the patch with zero values around the selected semi-circular area, Figure 22. The mask created in pre-processing the images also has zero values on areas that are excluded from the predicted path of motion. Zero values in the mask and template will have a high correlation, and allow for inaccurate coordinate estimation of the object. In Figure 21, the effect of the zero patches can be seen. Throughout the entire predicted path of motion a high correlation is found. As the object is obviously only in one point in the image, the high correlation in the rest of the image is due to overlapping zero patches. The amount of overlap between zeros from template and mask will determine the size of the error in coordinate estimation. If there is direction in the template that has more zero values than the other, the template will have a preferred axis of deviation from the true coordinates. In the current experiments these effects of template selection are not evaluated. The current technique can be improved by dilating the mask to prevent overlap between the zero patches in the template and mask. As the legs of the rigid body that holds the tracked sphere has a dark color, the overlap with the zero patch in the template will draw the estimated coordinates closer to the center of rotation on the diagonal axis ($y = \pm x$) compared to when there is little overlap with the dark legs, such as when the template is on one of the axis with its origin in the center of rotation. If the delay estimation needs to be improved I recommend investigating the accuracy of finding the coordinates of the spherical object using Hough transforms to circumvent the limitations and cumbersome process of the current technique.

In a study on the effects of visual delay on dexterity and accuracy in a laparoscopic setting the upper limit the human brain can correct for is found to be 0.25 sec [49]. However, this is for a delay between hand movement and visual feedback. In our case, a delay between the two data sources can create visual stimuli with conflicting messages. Delays between the two sources the brain is able to process correctly will therefore be significantly less. Given the upper limit, it is possible to delay both sources in order to have enough time to generate the patient model using the navigation, and accurately project it on to the laparoscopic image before presenting the view to the surgeon on screen.

Delay estimation is performed to allow accurate projection of the patient model on to the laparoscopic images. The presented method allows estimation of the delay with a standard deviation of 3.1 ms for five measurements, or 8.0 ms for twenty five measurements in which twenty measurements had an added delay. Calculated standard deviations assume the highly unlikely case that there is no variation in true delay between measurements, meaning that the delay estimation method is probably more accurate than shown here. The experiments performed show that the delay can be estimated accurate enough for our purpose.

CH5: Laparoscope calibration

5.1 Introduction

During the previous chapters, all steps needed to prepare the laparoscope for real-time tracking and visualization were investigated separately. Before the technique can be used in the clinic, all parts need to be combined in to one model. In this chapter, the camera calibration and hand-eye calibration methods are combined and the results are evaluated in a stationary setting where the delay between the laparoscopic images and tracking of the components does not play a role.

The laparoscope under investigation during this thesis is the Olympus EndoEye HD 10 mm with an oblique angle of 30°. The scope's camera is assumed to consist of two compound lens-systems. The inner compound lens-system holds the image sensor and is assumed to have fixed pose in relation to the handle of the scope. To change the direction the scope is looking at, the cylinder attached to the outer lens-system can be rotated around its axis. To keep the model simple, it is assumed that the only motion possible between the two lens-systems is a rotation around a fixed axis through both lens-systems.

Camera calibration results of chapter 2 indicate that the focal length and radial distortion parameters can be assumed fixed. These parameters are estimated from a sum of Gaussians where the mean and standard deviation of camera calibration at a single angle is used to generate one Gaussian curve with an area under the curve of one. The value where the sum of Gaussians from all angles obtains its maximum is set as the fixed value for each of the parameters. Principal point and decentering distortion parameters describe a circular motion during rotation of the scope. These parameters are estimated by fitting an ellipsoid through the parameter values obtained during calibration at several angles. By setting the value obtained in the standard configuration of the laparoscope (oblique view directed downwards) as reference, the values at other angles during rotation are estimated by rotation along the fitted ellipse. We have seen that this ellipsoid model is not perfect, but as the error stays below one pixel RMS its accuracy is sufficient for our purpose.

Two optical tracking sensors are attached to the laparoscope for navigation and measuring the scope's rotation angle. One sensor is attached to the handle and has a fixed relation to the image sensor while the sensor attached to the scope's cylinder has a fixed relation to optical axis of the camera. The sensor attached to the scope's cylinder can best be used as reference for the camera coordinate system as it requires a simpler model and its proximity relative to the other sensor will result in a smaller tracking error. Results of chapter 3 suggest that the hand-eye calibration can be modeled as a circular motion of the translation component, with a rotation of the camera coordinate system around the optical axis in opposite direction of the scope's rotation. Hand-eye calibration is performed using a checkerboard that is tracked relative to the optical sensors attached to the scope, and of which the position is known in relation to the camera coordinate system by camera calibration. The proposed camera model does not produce the exact same results as found by camera calibration. To account for this, the extrinsic camera parameters that describe the relation between the checkerboard and camera coordinate system are re-estimated based on the model and used for hand-eye calibration.

Comparison of the rotation angle obtained by navigation using the sensors attached to the scope, and by rotation angle obtained from hand-eye calibration results, suggested that the image sensor in the laparoscope is not completely fixed in relation to the handle. The translation component of the hand-eye calibration deviated from a circular motion in chapter 3, however, this is expected to be due to a defect in the scope used for the experiments. If this assumption holds, and how the rotation offset influences the overall reprojection result is investigated in this chapter.

In this chapter, we will first look at the overall reprojection results of the combined hand-eye and camera calibration model. Then, the individual components of camera calibration and hand-eye calibration, including the angle dependent model, are re-evaluated to see how the results came to be. Finally, the extrinsic parameters obtained by navigation are compared with the extrinsic parameters extracted from validation images using the camera model to determine the causes of the overall reprojection errors.

5.2 Navigation based reprojection of checkerboard

For the experiments, two Olympus laparoscopes of the same type are used. Results for both laparoscopes are evaluated simultaneously. *During this chapter, images on the left correspond to laparoscope 1, and images on the right correspond to laparoscope 2, unless stated otherwise.*

Camera calibration is performed at nine angles ranging from -120° to +120° with increments of 30° by capturing nine images of the checkerboard pattern in the same positions as in chapter 2. From the results, the proposed angle dependent camera model is created. Extrinsic parameters of all calibration images are re-estimated using our camera model. These extrinsic parameters are used to define all hand-eye transformations from the camera coordinate system to the optical tracking sensor attached to the scope's cylinder. An angle dependent hand-eye calibration model, explained in section 5.4, is generated from the results and used to estimate the extrinsic parameters between the checkerboard pattern in validation images and the camera coordinate system. These navigation based extrinsic parameters are used in combination with the camera model to reproject the 54 checkerboard points on to the image plane and compared to the true points as detected in the image. Reprojection error is expressed as the Euclidean root mean square error of all points in a single image. For each laparoscope, 30 validation images are captured, Figure 27.

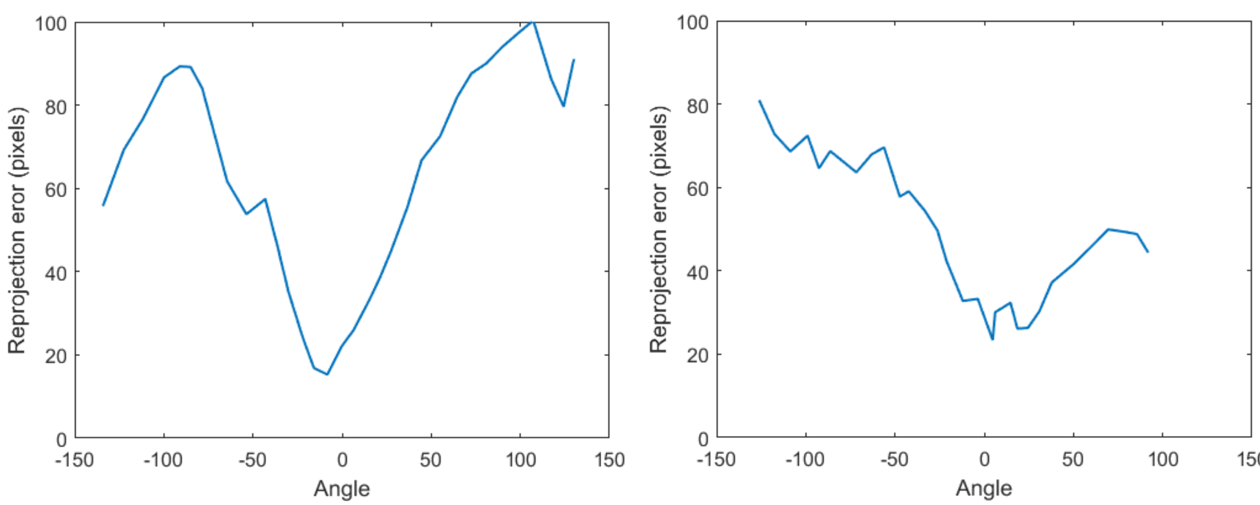


Figure 27: RMS error for reprojection of checkerboard point on image using combined camera and hand-eye calibration model. Thirty validation images are captured per laparoscope. The maximum error of 100 pixels corresponds to an error of ~10% of the image size.

Reprojection error ranges between 15-100 pixels where the lowest error occurs around 0° and the highest error around 100°. An error of 100 pixels on image size of 1440x1080 corresponds to a displacement of approximately 10% of the short axis of the image. In the camera calibration model, the parameters are expressed in relation to the reference angle 0°, giving a possible cause for the error distribution. However, the camera parameter errors are small in the model, and the reprojection error stayed below one pixel if the extrinsic parameters are estimated based on the model. Even though this was a doubling of the error compared to the standard camera model, it does not explain why the error increases to 100 pixels here.

All intrinsic and extrinsic parameters are evaluated again to determine why the model produces an error of this magnitude, and why the maximum error occurs at 100°, and not at 180°.

5.3 Intrinsic camera parameters

5.3.1 Reprojection error

The error of camera calibration is compared to the error obtained after applying our model. Similar to previous test, both methods are compared with decentering distortion in- and excluded from the camera model, Figure 28. Since a change in intrinsic parameters imposes a change on origin of the camera coordinate system, the extrinsic parameters are re-estimated for our model before point reprojection. As expected, the model based reprojection error is higher than the calibration based error and the inclusion of decentering distortion improves the results. However, the errors are still in the order of a pixel, and do not explain why the error increases up to 100 pixels in the navigation based method.

5.3.2 Focal length

In the proposed camera model, and all currently existing models, it is assumed that the focal lengths can each be approximated by a single value. The reasoning behind this is that a changing focal length would result in a noticeable, but unobserved, change in the images produced by the laparoscope during rotation. As the focal length describes the distance between the image plane and lenses, a change in focal length would mean that the distance between the two compound lens-systems changes during rotation. Results for focal length are shown in Figure 29.

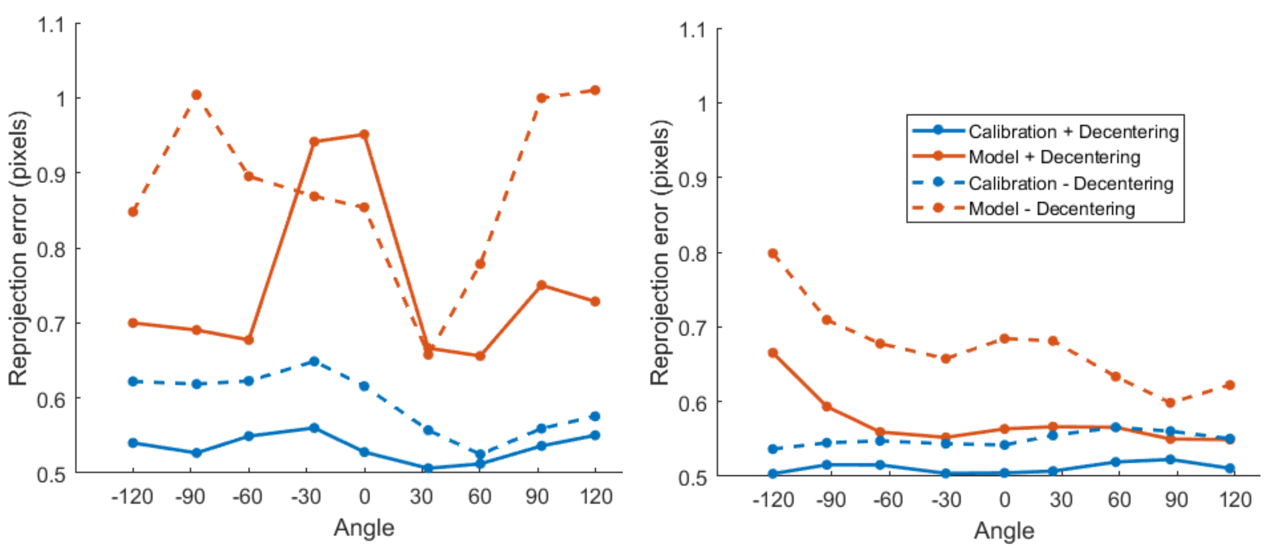


Figure 28: Reprojection error obtained during camera calibration and with the proposed model. Both methods are evaluated with decentering distortion in- and excluded from the camera model.

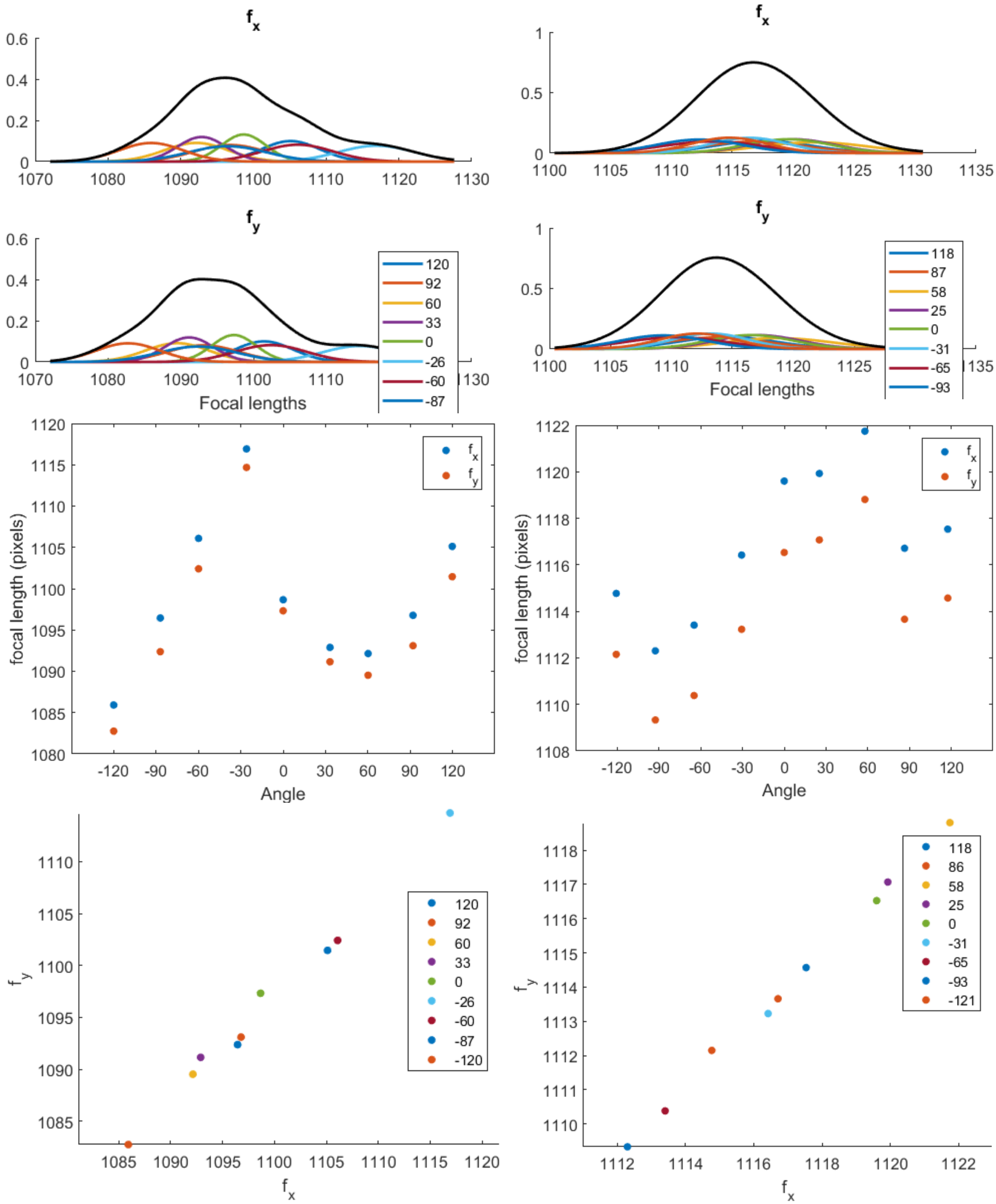


Figure 29: Top to bottom: sum of Gaussians for focal length calibration results at different angles; focal length parameters per angle; focal length x vs. y. In the model, it is assumed that the focal length can be approximated by a fixed value at all angles. This is clearly not the case as there is a 40 pixel difference, or ~4% of the average value, in laparoscope 1 during rotation.

From the Gaussian profiles and their sum it is clear that the estimated values for focal length are not consistent for laparoscope 1. There is 40 pixel range in estimated values for laparoscope 1. On a focal length of approximately 1000 pixels, this range corresponds to 4% of the total focal length. The values for laparoscope 2 show a spread of 15 pixels, similar to what was observed in chapter 2. Here, the variation was assumed to be due to calibration inaccuracies. However, if calibration is the cause of variation, the ratio f_x / f_y is also expected to change. In the plots of f_x vs f_y it can be seen that this is not the case. This means that the variation is due to a physical change in distance between the two compound lens-systems in the laparoscope. If focal length values are not consistent, they should be modeled as a function of the rotation angle. This requires the parameters to be predictable and describe a cyclical change as a function of the rotation angle. The rotation angle versus focal length plots show a sinusoid pattern for laparoscope 2. However, laparoscope 1 only shows a similar pattern for the positive angles, and the negative angles have a linear change.

The large unpredictable change of the focal lengths during rotation of laparoscope 1 suggest that the compound lens-systems not only rotate around a single axis, but are also permitted to move along this axis. This freedom of motion along the optical axis invalidates one of the assumptions for our model. As the reprojection error of our model with a fixed focal length stays below one pixel, the overall effect of this varying focal length on reprojection is not that large. However, the change in focal length also changes where the origin of the camera coordinate system is located within the camera. This topic will be discussed in the hand-eye calibration section.

5.3.3 Radial distortion

In the proposed camera model, and all currently existing models, it is assumed that the radial distortion parameters can each be approximated by a single value. This assumption is based on the assumption that the focal lengths are fixed. From the derivation of radial distortion, Appendix B, we know that the amount of distortion is proportional to the cubed field angle. The field angle is given by arctangent of the distance from a point on the image plane to the principal point, divided by the focal length. As the focal length changes, we already know that the radial distortion parameters will change, Figure 30.

The sum of Gaussians profiles clearly do not resemble a single Gaussian for any of the distortion coefficients. The width of the peaks still increases with coefficient number, but there is not a single value anymore that can be distinguished as the optimal value. In all profiles there seem to be two peaks where the values between the peaks differ by a factor of two.

If the radial distortion is to be modeled as a function of the rotation angle, the pattern preferably describes a cyclical motion or any other pattern that can be described by a simple function. From the normalized coefficient versus rotation angle plots it is clear that there is no simple function that can describe the change of the coefficient values. Since the radial distortion is a function of the focal length, one could expect to find a relation between the two. However, this relation is not visible in the normalized coefficient plots. This is because rotation and shift between the two compound lens-systems is not around and along the optical axes of both systems. Besides the relation to the focal length, radial distortion also depends on several other lens coefficients. As the rotation and motion are not around and along the optical axis, the interaction between the lens coefficients also changes during rotation.

From the results obtained here, it does not seem possible to model the radial distortion as a function of the rotation angle. However, as the reprojection error using our model is small, the initial assumption of a fixed value still seems accurate enough to describe radial distortion.

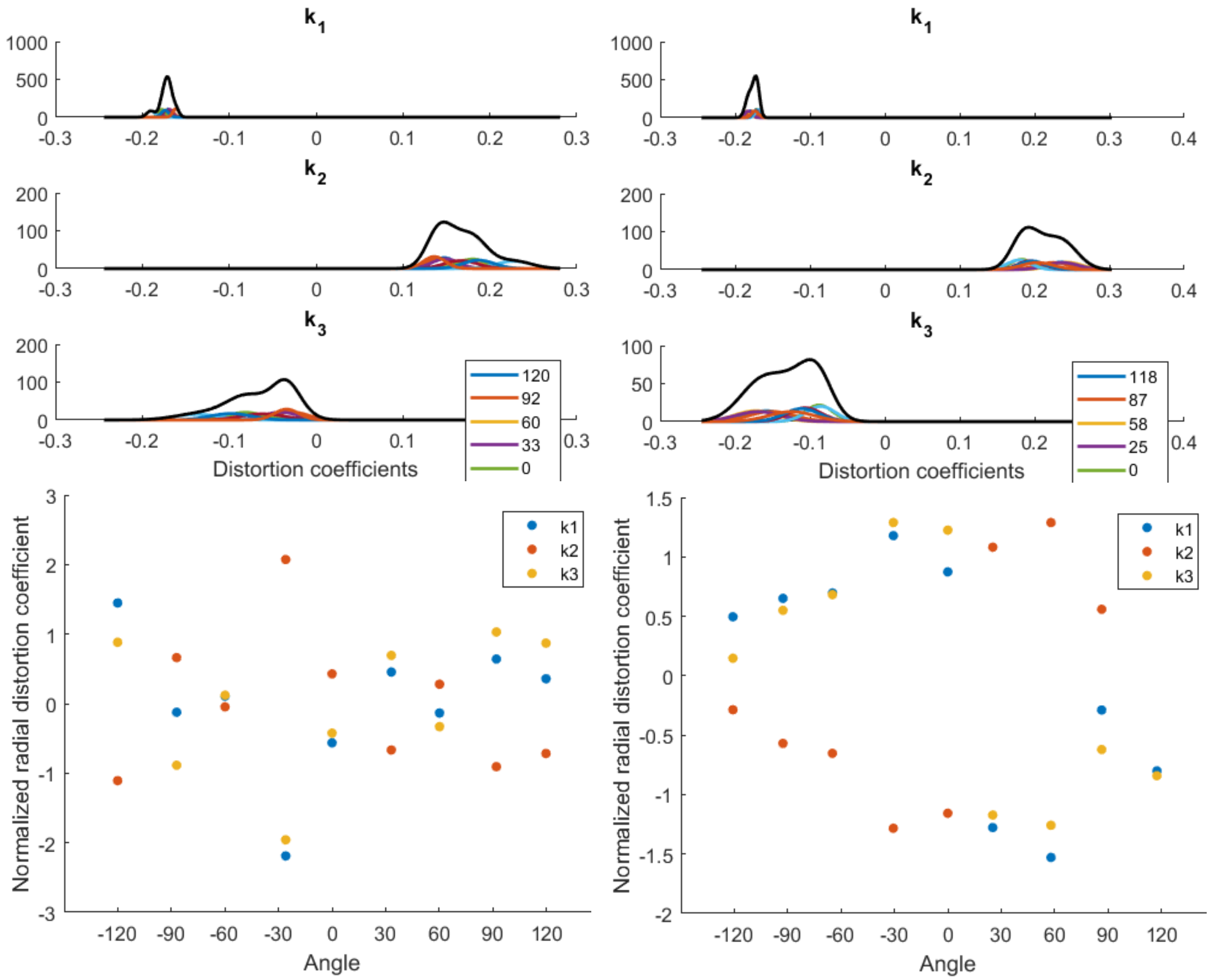


Figure 30: Top, sum of Gaussians for radial distortion coefficients. Bottom, radial distortion parameters normalized to zero mean and standard deviation one per coefficient versus the calibration angle. For both laparoscopes, the distortion parameters change during rotation, violating our assumption of a fixed value. In the bottom plots it is visible that the change cannot be described by a simple function.

5.3.4 Principal point

The path on the image plane described by the principal point during rotation of the laparoscope can be modeled as an ellipsoid. The principal point rotates around a point where the rotation axis of the laparoscope intersects the image plane. The path of the principal point depends on the distance to the rotation axis, the angle between the image plane and rotation axis, and on the decentering distortion. Results for the principal point modeling are shown in Figure 31.

Principal points for laparoscope 1 resemble a circle while those for laparoscope 2 plot an ellipsoid. The circular shape of laparoscope 1 clearly shows an angular offset in one direction between the reference angle at 0° and the other angles. This is most likely due to an error in estimation of the principal point at reference angle 0°. In the ellipsoid of laparoscope 2, we can see that the angular error in the model increases along the long axis of the ellipsoid, similar to that observed for the decentering distortion in chapter 2. As the error in principal point can partially be corrected by the extrinsic parameters, the results obtained here are sufficient for modeling of the principal point.

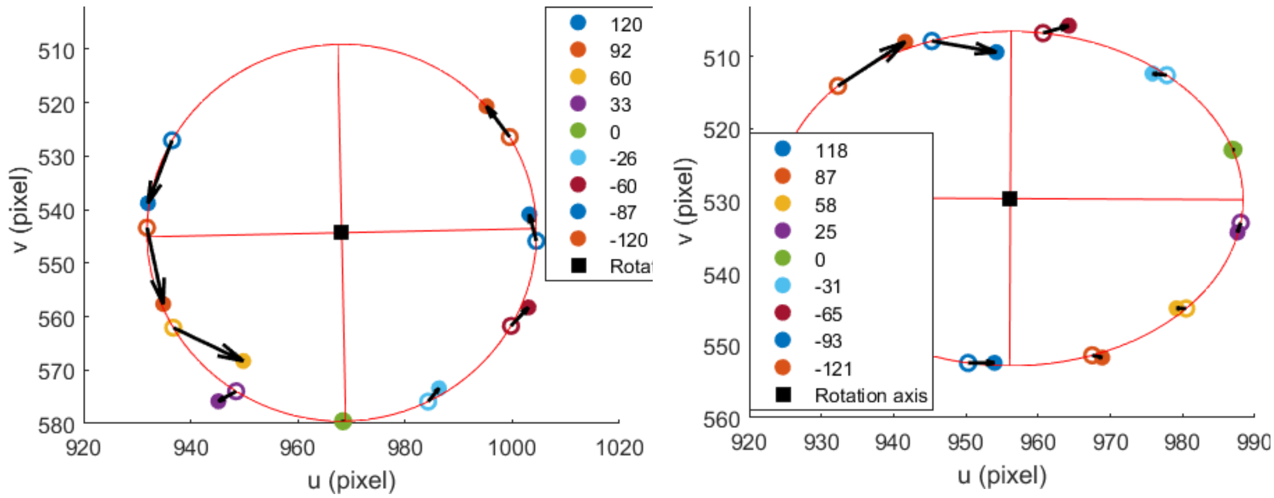


Figure 31: Model and calibration based principal point positions for the different calibration angles. Solid circles are the measured principal points, open circles the principal points estimated with the model, and the black square the center of the ellipsoid. Left shows an angular offset originating from an estimation error in the reference angle at 0°. The error on the right increases in size along the long axis of the ellipse.

5.3.5 Decentering distortion

Decentering distortion consists of a direction and magnitude that together can be modeled as a function of the rotation angle by an ellipsoid. The decentering distortion is a combination of the decentering distortion of the two compound lens-systems and the interaction between the two. It is assumed that the center of the ellipsoid indicates the direction and magnitude of the inner compound lens-system that holds the image plane. The path described by the decentering distortion relative to the center of rotation is a sum of the outer lens-system and the interaction between the two systems, Figure 32.

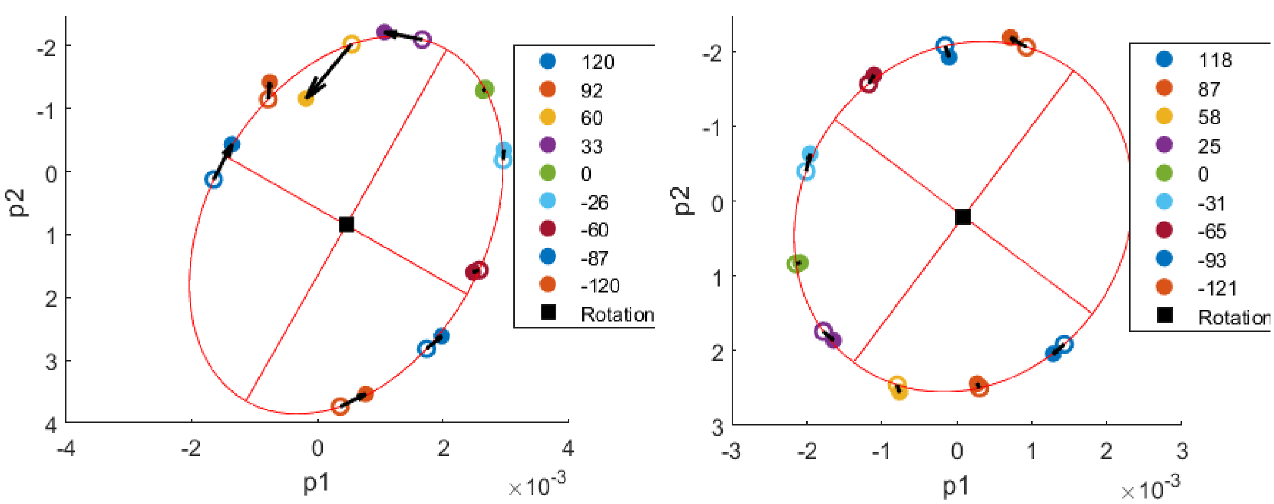


Figure 32: Model and calibration based decentering distortion parameters. Solid circles are the measured decentering distortion parameters, open circles the distortion estimated with the model, and the black square the center of the ellipsoid. In the strong ellipsoid shape on the left there is a clear error along the long axis of the ellipse. In the image on the right the shape is close to a circle and the model can better approximate the decentering parameters.

In the decentering distortion parameters there are again an ellipsoid and circular shape visible. Where the principal point had a circular path for laparoscope 1, the decentering distortion has an ellipsoidal path and vice versa for laparoscope 2. Just as in the principal point model of laparoscope 2, and the decentering distortion model of chapter 2, the error of the estimated values increases along the long axis of the ellipsoid. Calibrated values for laparoscope 2 are close to a circular shape and show that the model is a good approximation for the calibrated values.

5.3.6 Discussion

All intrinsic and distortion parameters included in the model are re-evaluated to test if the assumptions and approximations of the model still hold in attempt to find the origin of the large navigation based reprojection error. Reprojection error for the model based method is approximately 30% larger than the calibration based method. The reprojection error stays below one pixel in all cases when decentering distortion is included in the model. It therefore seems unlikely that the errors in the model are the cause of the large errors in the overall reprojection.

Focal lengths were assumed to be fixed during rotation. Results show that this is not the case as there is a variation of 4% of the total focal length visible in laparoscope 1. This can only be caused by a variable distance between the two compound lens-systems in the tip of the laparoscope. As the radial distortion is proportional to the cubed focal length, the assumption that distortion parameters are fixed also does not hold. Even though the assumptions are incorrect, the reprojection error obtained using these assumptions show that the effects on the error are not that concerning. One effect of the variable focal length not discussed is the change imposed on the extrinsic parameters. This will be looked in to further in the next section.

Ellipsoid fitting for principal point and decentering distortion is still deemed accurate enough for our purpose. The results are not perfect, but the errors in the model can be compensated for by the extrinsic parameters. The ellipsoid model achieves accurate results when the actual shape of the parameters resembles a pattern close to a circle. If the pattern changes to an ellipsoid with a strong difference in short and long axis, the error in values generated by the model increase in size along the long axis. As the effects on the reprojection error are small, and as the approximated values are much closer to the calibrated value than assuming a fixed value for all angles, the model is not changed.

5.4 Hand-eye calibration

During the hand-eye calibration experiments in chapter 3, the transformations between the reference sensors and camera coordinate system expressed unexpected behavior due to a suspected defect in the laparoscope used. It was not possible to define the hand-eye model as a function of the rotation angle because of this. Here, we will look in to the hand-eye transformations again, and investigate the effects of intrinsic camera parameters on the hand-eye transformations. Then, a hand-eye transformation model will be defined as a function of the rotation angle. All transformations are inspected relative to the optical tracking sensor attached to the scope’s cylinder.

5.4.1 Hand-eye transformations using different extrinsic parameter sources

During hand-eye calibration, four transformations are multiplied to get a transformation from the coordinate system of the reference sensor to the camera coordinate system. Two of the transforms are obtained using the optical tracking system, one is a fixed transformation that is obtained as defined in chapter 3, and the final transformation is between the checkerboard pattern and the camera coordinate system and is obtained during camera calibration. This last transformation is generated from the extrinsic parameters, and is obtained in combination with the intrinsic camera parameters during camera calibration. The intrinsic and extrinsic parameters from a set that together describe how a point in front of the camera is projected on to the image plane. In our model, the intrinsic parameters differ from the calibrated parameters. This means that the extrinsic parameters also need to be corrected to account for this change. Here we evaluate the effect of our intrinsic model on the hand-eye transformations. Figure 33 shows the hand-eye transformation obtained by re-estimating the extrinsic parameters using our model. Figure 34 shows the hand-eye transformations using the extrinsic parameters found during camera calibration. It is important to note that the intrinsic parameters self are not used for calibration.

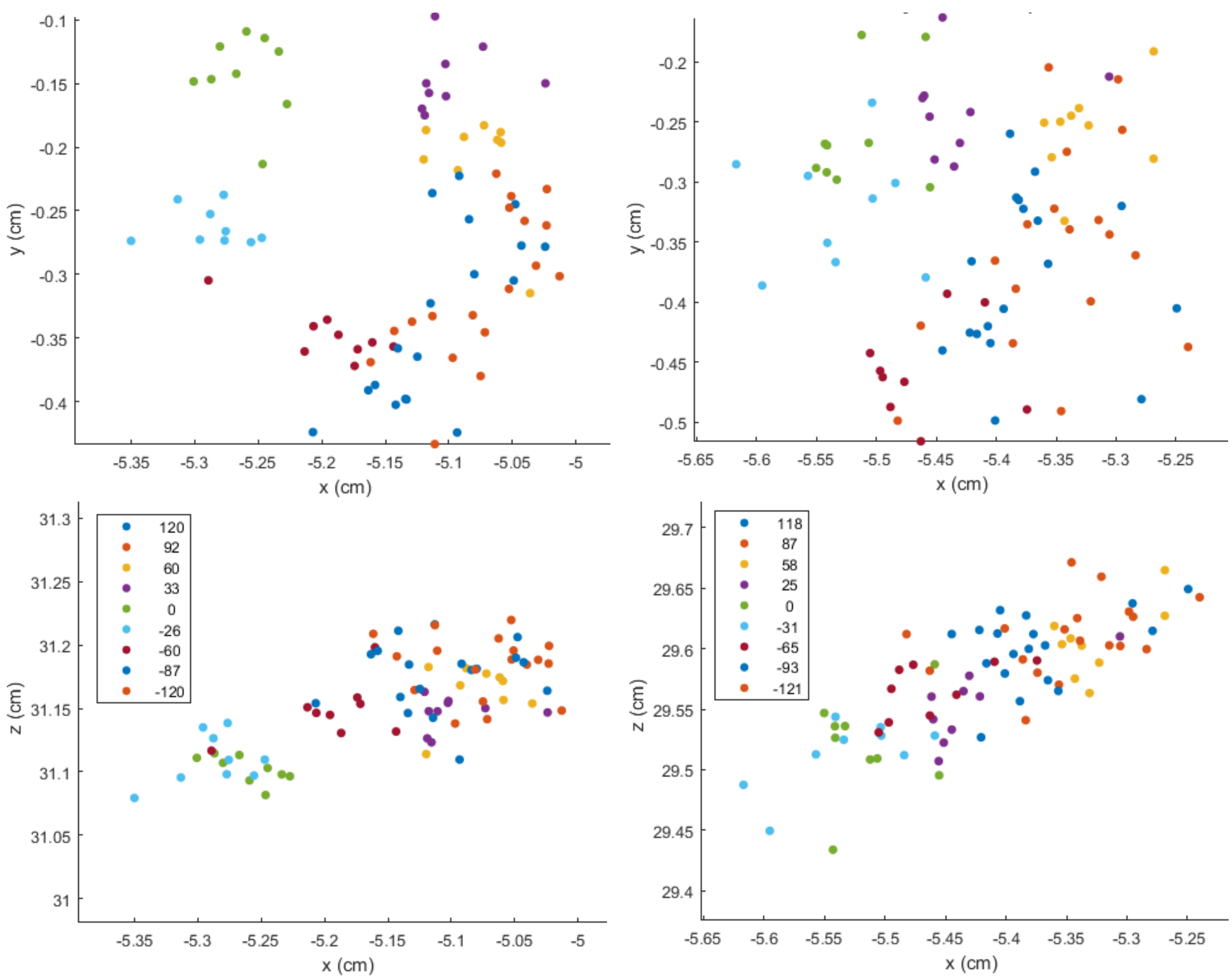


Figure 33: Hand-eye transformation translation components projected on the XY- and XZ-plane using the extrinsic parameters found using <u>our model</u>. Each rotation angle has nine hand-eye transformations. The spread is approximately 1 mm in all directions per angle. In the top row there is a clear circular pattern visible on the left. On the right, the pattern is less obvious as the points for all angles are closer together, but it is still present. The angle between successive angles of the nine calibration sets is approximately 30°. However, the angle between two successive sets around a fictive point in the center of the circle is clearly not constant. Around reference angle 0°, the angular distance between sets is larger than at distant angles. The circle is closed in the 240° degrees that are measured. In the XZ-plane it can be seen that the transforms are all in a plane with a 1 mm spread.

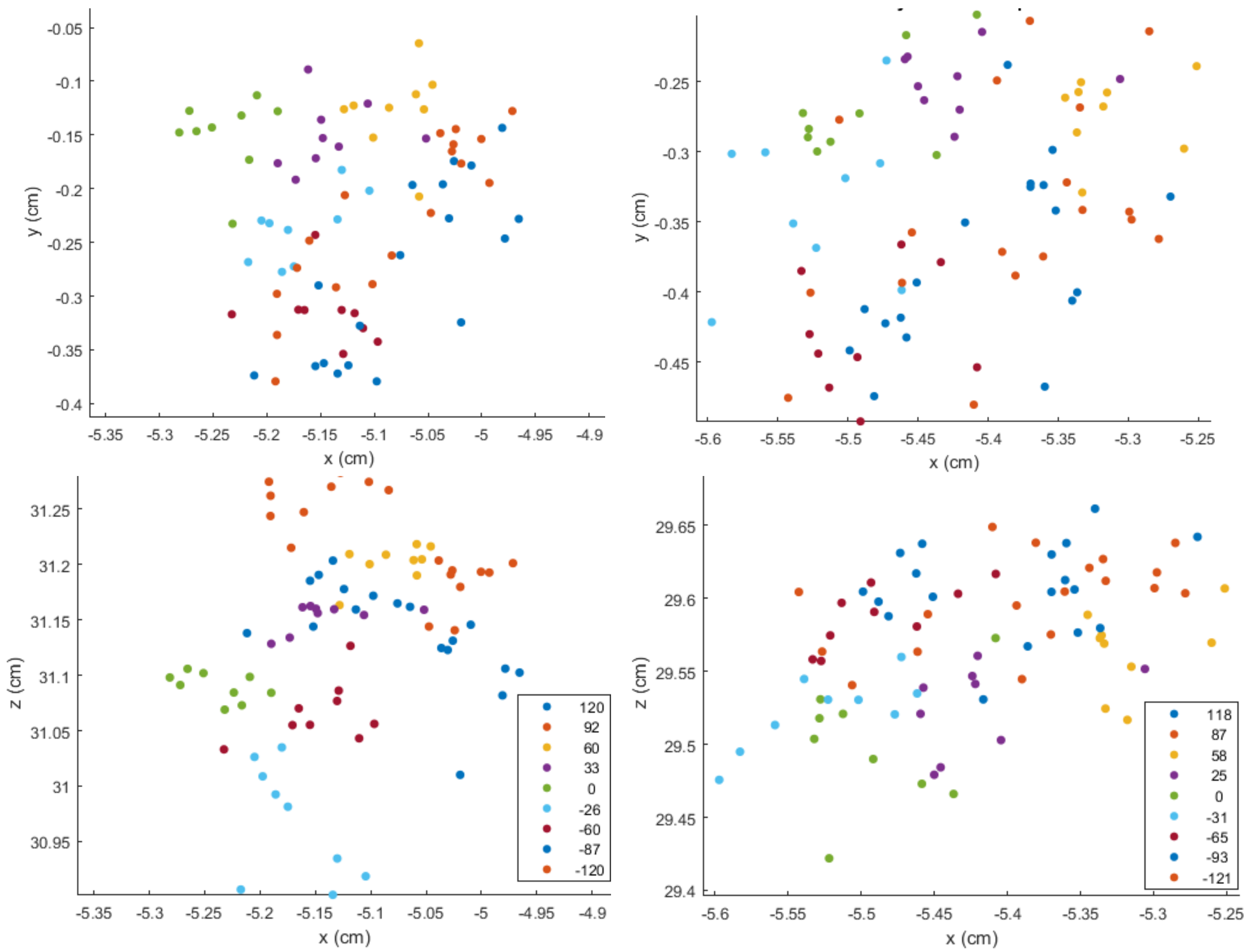


Figure 34: Hand-eye transformation translation components projected on the XY- and XZ-plane using the extrinsic parameters found during camera calibration. Each rotation angle has nine hand-eye transforms. For each angle, the hand-eye translations are sort of clustered, but neither of the laparoscopes shows a clear pattern. From the transforms in Figure 33 we know that the optical axis lies diagonally in the XZ-plane. For laparoscope 1, the range of the points along the optical axis is approximately 5 mm due to the large variation in focal lengths. For laparoscope 2, the variation along the optical axis in approximately 1.5 mm.

A change in focal length and principal point has a direct influence on the extrinsic parameters. The focal length influences the distance along the optical axis, and the principal point influences the position in the plane orthogonal to the optical axis. In the scatterplots, the optical axis is diagonally downwards in the XZ-plane. This can best be seen in the plots of Figure 33 where the optical axis is perpendicular to the plane of rotation. In the calibration based transformations of laparoscope 1 there is a spread of 5 mm along the optical axis due to the variation in the calibrated focal lengths. By fixing the focal lengths to a single value, the hand-eye transformations are forced in to a plane with a spread of around 1 mm. A similar behavior can be seen in the spread for scope 2. Here, the spread in the calibration based method is smaller as the variation in focal lengths is smaller, but still improves by fixing the focal lengths.

Fitting the principal point to an ellipsoid shape imposes a similar behavior on the hand-eye translations in the plane orthogonal to the optical axis during rotation. For laparoscope 1 the calibrated method creates a shape that resembles a triangle instead of a circle. In the hand-eye translations obtained using the model, the shape is close to a circle. For laparoscope 2, the behavior is similar but less obvious due to a smaller radius of the circle resulting in an overlap between the clusters per angle. For both laparoscope scatterplots in the model based method it can be seen that the rotation of the clusters around a fictive point in the center of the circular motion is not equal to the rotation angle of the laparoscope. The amount of rotation between successive angles is much larger close to the reference angle 0° than it is at distant angles. In both cases, the circle closes in the 240° used for camera calibration. This means that the translation component of the hand-eye transformation cannot be modeled by a simple rotation around a central axis.

Hand-eye calibration results obtained with the extrinsic parameters re-estimated based on our model not only improve the hand-eye translations, but are also necessary to create a pattern in that can be modeled by a simple function. In the ideal case, all translation components at a single calibration angle coincide in a single point. The results obtained here have a spread of around 1 mm in all directions. This spread can be a result of tracking inaccuracies of the optical sensors, and of errors in estimated extrinsic parameters. This spread will have an significant influence on the reprojection error no matter the origin.

Orientation of the hand-eye transformation does not change if the intrinsic parameters are changed. In all hand-eye transformations, the optical axis pointed in the same direction with a negligible variation. The orientation between the corresponding transformation from the calibration based method and the model based method were similar as expected. Based on the orientation results there is no preference for a method to use in the hand-eye model.

5.4.2 Rotation angle dependent hand-eye model

Above we have seen that the hand-eye transformations can best be created using the extrinsic parameters obtained using our model. The orientation and translation for the hand-eye transformations are modeled independently.

The translation component is determined by first finding the plane the rotation takes place in and the axis of the rotation. This is achieved using the same method as used for rotation calibration of the laparoscope. All translation components are assumed to lie in a plane and describe a circle around a point in that plane, Figure 33. The normal to the plane (circle), and the center of the circle are found by solving a non-linear least squares problem, Appendix C. The average translation per calibration angle is determined, and the rotations of these average positions around the rotation axis are determined with respect to the position at reference angle 0°, Figure 36. As the rotation of hand-eye positions does not move at the same speed around the rotational axis as the rotation of the scope, the amount of hand-eye rotation is fitted to the scope's rotation angle using a Fourier sum, Figure 35.

$$\gamma_{HE} = a_0 + a_1 \cos\left(\omega \cdot \theta_{view}\right) + b_1 \sin\left(\omega \cdot \theta_{view}\right) \tag{5.1}$$

Here, γ_{HE} is the hand-eye rotation angle, θ_{view} the scope's rotation angle relative to the reference angle, and $\{a_0, a_1, b_1, \omega\}$ the fitted values. The position in hand-eye transformations can then be modeled by rotation of a point at a fixed radius to the rotation axis with an angle as given by (5.1). Results for the model based position, and calibrated position are shown in Figure 36.

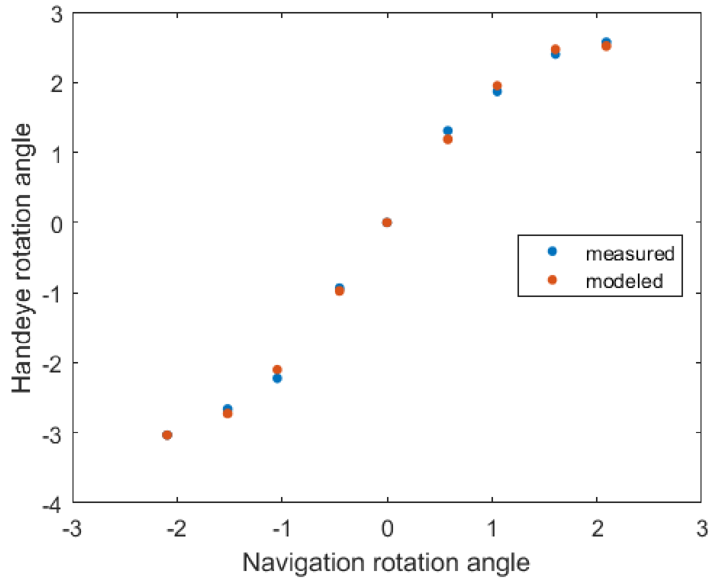


Figure 35: Fourier sum fit results of the of the scope's rotation angle to the rotation angle of the calibrated hand-eye translation components. The fit is a decent approximation of the true hand-eye angle.

Hand-eye translation modeling has a mean error of 0.5 mm compared to the averaged translation per angle. As the spread of the translations within an angle is 1 mm on average, the error of the model is within an acceptable range. Optical tracking errors will influence this error during use and can lower or increase the error as a result.

Orientation of the hand-eye transformation is modeled by rotating the orientation of the measured hand-eye transformation at reference angle 0° around the optical (z-)axis. The amount of rotation around the optical axis is equal but in opposite direction of the rotation of the laparoscope.

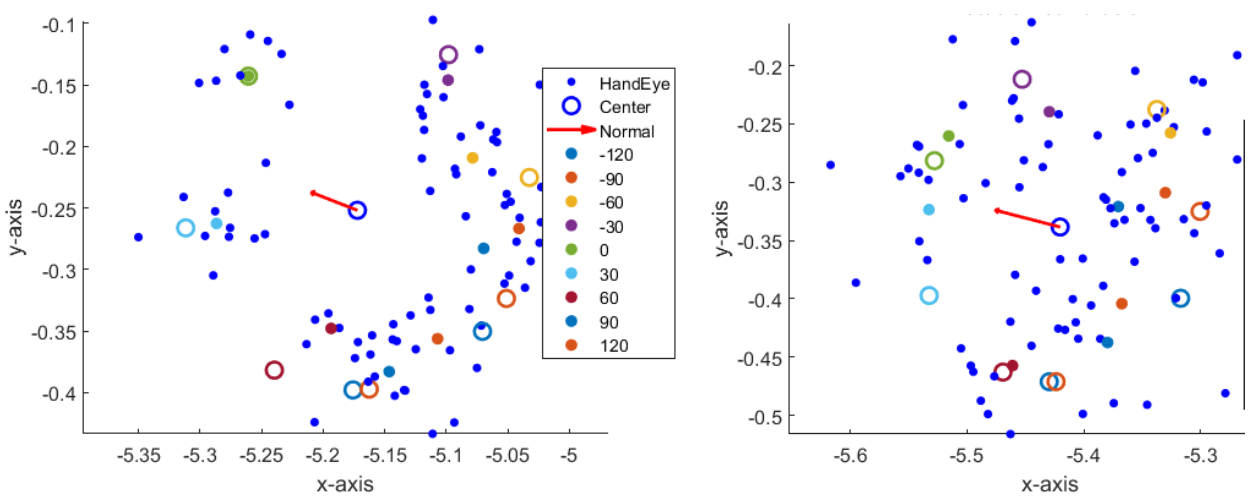


Figure 36: Model results for hand-eye translation components. Blue are the hand-eye positions from all images at nine angles. Open blue circle is the center of rotation, and red arrow is the normal that together from the rotation axis. Solid dots are the average positions per angle, and in open circles the modeled based position of the average positions. The variation of the positions per angle is about 1 mm. The modeled position deviates from the measured position with 0.5 mm on average, meaning that the modeled positions are not perfect, but within the measured variation.

5.5 Validation

The overall reprojection error using navigation is based on 30 validation images per laparoscope. For laparoscope 1, the calibration was performed from +120° to -120° followed by acquisition of the validation images from negative to positive rotation angles. For laparoscope 2, the procedure was repeated in opposite direction of rotation. By inverting the rotation direction in the acquisition procedure we aim to get more information on the rotational offset observed in chapter 3.

The reprojection error in the validation images is first evaluated using extrinsic parameters estimated from the image using our model, Figure 37. If the reprojection error of the validation image is compared to the reprojection error in the calibration images using our model, it can be seen that the error in the validation images is slightly larger in most cases but there are a few angles where the error is even lower. This indicates that the error observed using navigation is not caused by the intrinsic model and can be accounted to the extrinsic parameters in the hand-eye transformation.

The influence on the overall reprojection error is evaluated for translation and orientation separately. This is achieved by determination of the reprojection error where the hand-eye transformations using navigation are corrected for position, and for rotation around the optical axis. The position error is corrected by setting the translation of navigation based extrinsic parameters equal to the translation from the image based extrinsic parameters. The positional error is then defined as the distance between the navigation and image based extrinsic translations. The rotational error around the optical axis is determined by finding the geometrical transformation between the reprojected points in the image and the detected points in the image. The error in orientation is corrected by setting the orientation in the navigation based extrinsic parameters equal to that of the image based extrinsic parameters, Figure 38.

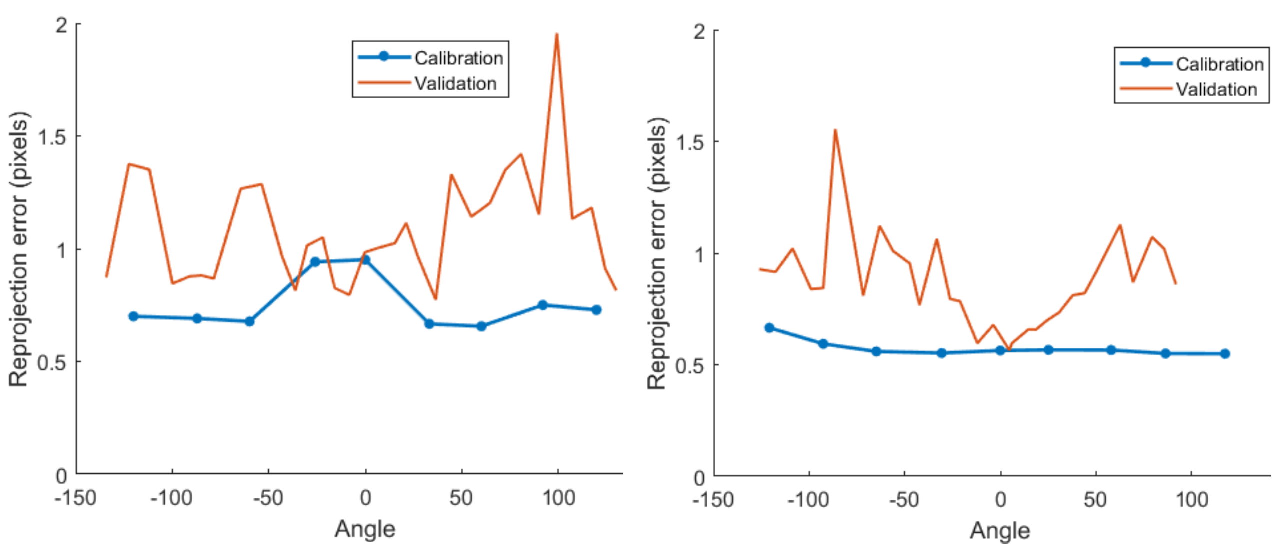


Figure 37: Reprojection error where extrinsic parameters are obtained from the images using our model. In blue, the error obtained from the calibration images (average of nine per angle). In orange, the RMSE on the validation images.

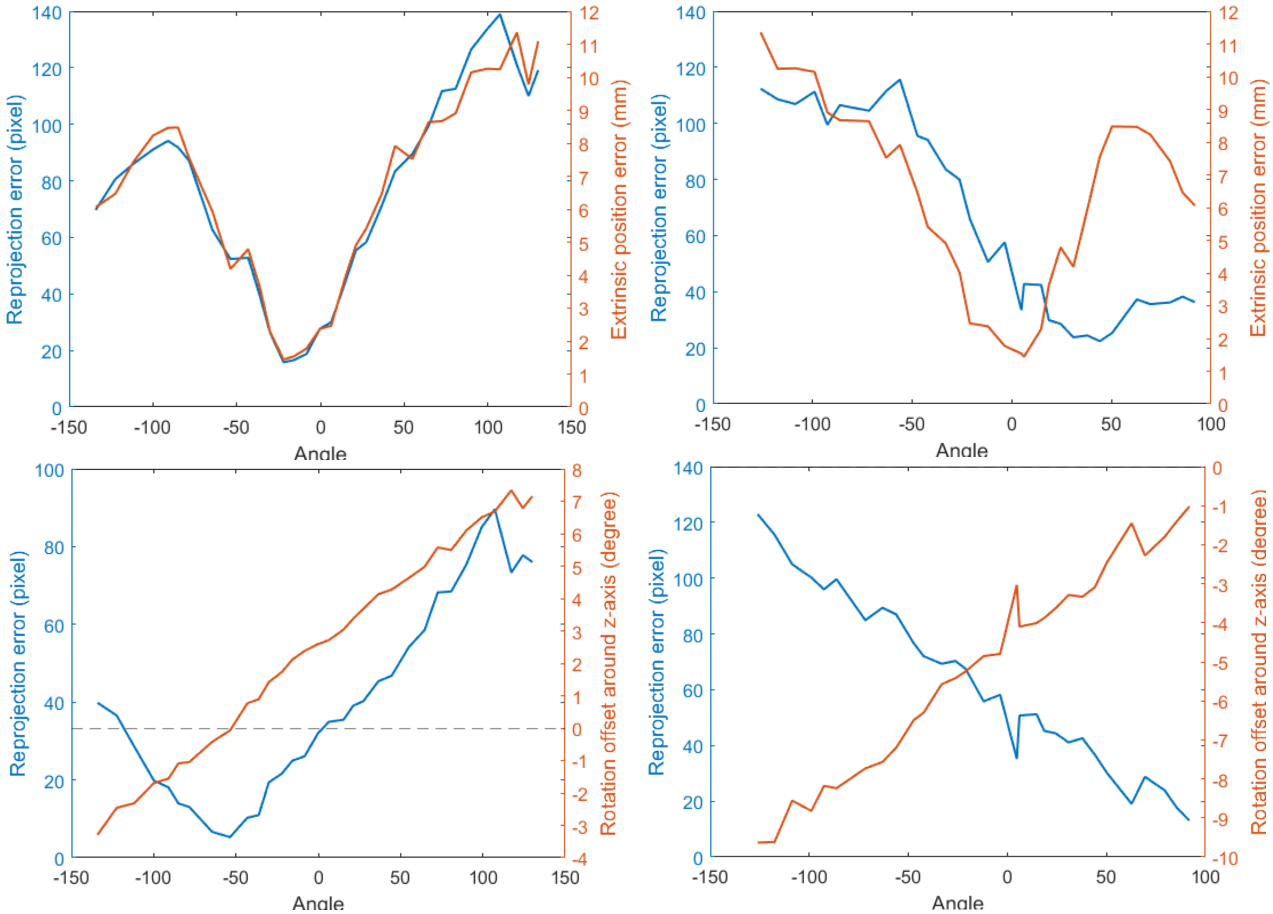


Figure 38: Top, reprojection error corrected for rotational offset around the optical axis in blue, and the extrinsic position error in orange. Bottom, reprojection error corrected for extrinsic position error in blue, and the rotational offset around the optical axis in orange.

If the orientation is corrected, the shape of the reprojection error is similar to that of the extrinsic position error. In laparoscope 1, the correlation between the two is almost perfect. In laparoscope 2 there is an angular shift between the two graphs due to the rotation offset in the translation component of the hand-eye model as can be seen in Figure 36. If the position is corrected, there is positive correlation for laparoscope 1, and a negative correlation for laparoscope 2 between the remaining reprojection error and the rotation error around the optical axis. This reversal of correlation is due to the way the data is acquired for both laparoscopes. By reversing the direction of acquisition, the rotational offset is in opposite directions. For both laparoscopes, it can be seen that reprojection error drops to 10 pixels where the angular error is zero. For both laparoscopes the slope is constant with a rotation error of approximately 8° over a scope rotation angle of 200°. It is clear that there is still a large error present in both rotation and position of the hand-eye model. Solving just one of the two does not lower the error, in fact, in three of the four cases the reprojection error even increases if just one is solved.

The rotation offset between the angle measured with navigation and the image plane can be explained by rotation of the compound lens-system containing the image sensor within the laparoscopes. It was assumed that the image sensor was fixed in relation to the handle by magnetic coupling, but the results show that there is still rotation possible. Unfortunately this rotation offset is not constant as this would result in a zero rotation error at the reference angle 0°. It seems that the offset is caused by friction within the laparoscope leading to hysteresis of the rotation.

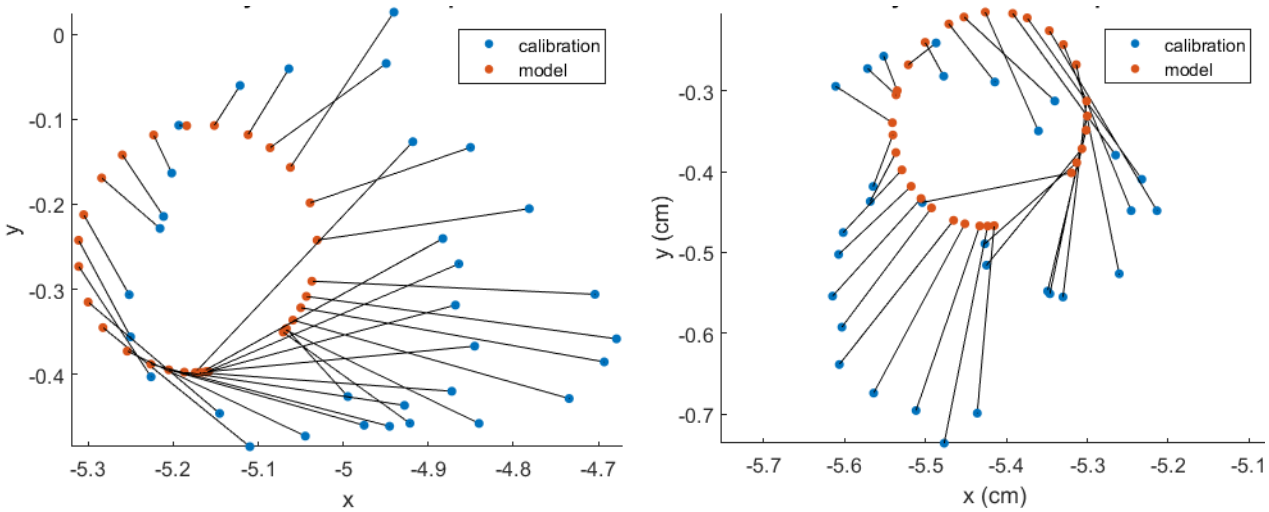


Figure 39: Position of the origin of the camera coordinate system obtained from hand-eye calibration. In orange, the positions obtained using navigation and the hand-eye model. In blue, the positions obtained using extrinsic parameters extracted from the validation images using our model. There is also an error in the dimension not shown here.

Rotation of the image sensor in the laparoscope does not explain the observed position error. In an effort to uncover the cause of the positional error, the model based hand-eye translations are compared to the navigation based hand-eye translations for the validation images, Figure 39. For both laparoscopes, the position errors at the reference angle are small and rapidly increase from there. The circular shape of the points generated by the model should be rotated around the position at the reference angle (opposite to the opening in the model based circle) to get a better match between the two sets. The direction the model based positions should be rotated in is in opposite directions for both laparoscopes, just as the acquisition order of the images. Changing the direction of rotation between acquisition of calibration and validation images, has moved the origin of the camera coordinate system. It seems that the position of the camera sensor in the laparoscope has changed by inverting the rotation direction. If we still assume the outer compound lens-system fixed to the laparoscope's cylinder, the change in camera position can only be caused by a movement of the inner compound lens-system containing the image sensor. We have already observed hysteresis in the rotation and a changing focal length. Both are the result of a movement of the image sensor in relation to the handle. Considering the observed movements, it is likely that other movements, such as tilting of the image sensor, are also possible. The small error in position around the reference angle, in both laparoscopes evaluated here, and in the laparoscope evaluated in chapter 3 (results not shown), suggest that the camera components have a default resting state the camera falls back to in the reference configuration.

5.6 Discussion

We have created a model to describe the hand-eye transformation, between the coordinate system of the optical reference sensor and the camera coordinate system, as a function of the scope's rotation angle. The reprojection error of the combined intrinsic camera and hand-eye model ranged from 20-100 pixels RMS for two different laparoscopes. The intrinsic camera model, and the hand-eye model are re-evaluated to find the source of the large reprojection error.

In the intrinsic camera model it was assumed that focal lengths are fixed. Here we have shown that this is not the case. The size and unpredictability of the change in focal length during rotation suggest that the distance between the two compound lens-systems in the laparoscope can change. As a consequence, radial distortion can also change during rotation as it is a function of the focal length. Modeling of the principal point and decentering distortion by an ellipsoid provides a good approximation of the calibrated values. If the fitted ellipsoid is close to a circle, the approximation is more accurate compared to a strong ellipsoidal shape. In the fitted functions, the rotation angle measured by navigation was used for visualization of the modeled values. As the rotation of the image plane is less than the navigation based rotation angle, the model fits better than shown here. The reprojection error for the calibration and validation images is close to one if the extrinsic parameters are estimated from the image.

Fixing the focal length to a single value, and modeling of the principal point by an ellipse, creates hand-eye translations that resemble a circle in a plane. The translations obtained from the original calibrated values produce a distribution that cannot be modeled. The intrinsic camera model is therefore also necessary for hand-eye calibration. The spread in hand-eye positions is in the order of 1 mm in all directions per angle. Ideally, all of the points per angle would coincide in a single point. However, due to measurement errors, there will always be some spread in the positions. The translation component is modeled as a rotating point around a fixed radius to the rotation axis in the center of the circle. The amount of rotation around the rotation axis is not linear to the navigation based rotation and is therefore approximated by a Fourier sum. The mean deviation between the average position per angle and its modeled position is 0.5 mm.

Re-evaluation of the intrinsic and hand-eye model show that the modeled values are a good approximation of the measured values. Based on these results, the overall reprojection errors are expected to be lower and have a smaller range than obtained with the validation images. The difference between the extrinsic parameters obtained from the images and from the hand-eye model show that there is a significant rotation and position error in the hand-eye model. Rotation around the optical axis shows hysteresis of the rotation of the image plane. On a rotation of 200°, the rotation of the image plane is constantly 4% less than the rotation measured using navigation. Since the rotation offset is very consistent, it might be possible to model the hysteresis by including the history of rotation. However, this would require the optical tracking system to record every single rotation of the scope to be accurate. This might be possible in an experimental setup, but in a clinical setting this is not realistic. Changing the rotation direction between acquisition of calibration and validation images creates a movement of the camera coordinate system in the laparoscope. As a result, there is position error of up to 10 mm in the extrinsic parameters obtained using navigation. Rotation offset (hysteresis), displacement of the camera coordinate system, and changing focal length can all be explained by a movement of the inner compound lens-system including image sensor relative to the handle of the laparoscope. The unexpected hand-eye translation behavior that could not be explained in chapter 3 can also be caused by this.

In order to create model for the laparoscope it is necessary to know the pose of both compound lens-systems in the tip of the laparoscope. The outer lens-system is still assumed fixed in relation to the scope's cylinder and can be tracked by an optical sensor attached to this cylinder. As the inner lens-system is not fixed in relation to the handle, as initially assumed, there is no option to track its pose based on anything accessible on the outside of the laparoscope. This is the result of the configuration in chip-on-the-tip laparoscopes. In conventional systems, the laparoscope consists of an interchangeable lens-system that is attached to a camera-head. This conventional setup allows external tracking of the individual components of the laparoscope, thereby eliminating the source of our problem.

Three different laparoscopes were used in the experiments performed during this research. All laparoscopes showed that the image sensor is not fixed in relation to the handle, thereby making it impossible to track the two components of the camera system. However, none of the laparoscopes was calibrated twice. It is possible that the movement between the two components is predictable and can be correlated to the rotation of the laparoscope. Further research will need to be conducted to verify if this correlation exists. If it exists, it would still be difficult to model the behavior based on optical tracking as the tracking system would need to record every single rotation of the scope. Including rotation history in the model would also make the model very complex and prone to errors.

5.7 Conclusion

In this chapter we have seen that the initial assumption of a fixed focal length during rotation of the laparoscope does not hold. As a consequence, the radial distortion coefficients are also not fixed. However, by fixing the focal length and radial distortion parameters to a single value, and by approximating the principal point and decentering distortion parameters by an ellipse, the translation between the camera coordinate system and coordinate system of the reference sensor is forced to behave as a circular motion in a plane during rotation of the laparoscope. Despite its inaccuracies, the camera model produces a reprojection error in the order of only 1 pixel RMS. The low error and circular motion of the hand-eye translation make the proposed model a good approximation of the camera. Rotation of the hand-eye translation component along the circle is not linear to the rotation between the two parts of the laparoscope. Hand-eye translation rotation angle can be modeled as a function of the scope's rotation angle using a Fourier sum.

Changing the direction of the scope rotation results in a substantial change of the hand-eye transformations. This is caused by a displacement of the camera coordinate system. The only explanation we see for a displacement of the camera coordinate system is a displacement of the image sensor within the laparoscope. In the model, it was assumed that only motion possible between the two camera components is a rotation around a fixed axis through the two components. This freedom of motion would require a significantly more complex model for the laparoscope. As the pose of the inner compound lens-system cannot be measured in relation to anything accessible on the outside of the laparoscope, it does not seem feasible to create a simple model for the laparoscope using external tracking.

CH6: General discussion and conclusions

In the few literature references available on camera calibration of oblique viewing laparoscopes, several assumptions are made on the behavior of the intrinsic and extrinsic camera parameters. All laparoscope configurations used in these references consist of a conventional laparoscope setup with a camera head and separable optics system. The Olympus EndoEye HD 10 mm laparoscope used during this study is an all-in-one system with the image sensor located in the tip of the laparoscope. To our knowledge, there are no literature references available of a (working) calibration method for oblique viewing laparoscopes with this chip-on-the-tip technology. The behavior of intrinsic and extrinsic camera parameters during rotation of the laparoscope were evaluated to verify if the assumptions made for the conventional laparoscopic setup are still valid.

In literature, focal length and radial distortion coefficients are assumed to be fixed, and the principal point is either described as a fixed point, or as a circular motion around a point on the image plane. We have shown that, for our laparoscope, the focal length does change during rotation. Radial distortion therefore also changes as it is a function the focal length. Our observations showed that the principal point behavior can best be modeled as an ellipse instead of a circle.

Decentering distortion is not included in any of the available references. Decentering distortion originates from misalignment of lens elements. As the total lens-system of the laparoscope consists of two compound lens-systems rotating relative to another, decentering distortion also changes during rotation. The estimated decentering parameters describe a magnitude and direction that can also be modeled by an ellipse during rotation. We have shown that the addition of decentering distortion improves the reprojection results of the camera model.

Even though the assumption of a fixed focal length and radial distortion is false, it is used in the designed model as it leads to better extrinsic camera parameters needed for hand-eye calibration. The proposed camera model has a reprojection error on calibration images of 0.5-1 pixel RMS. Applying the camera model to validation images increased the reprojection error to a mean of 1 pixel RMS. This increase can, at least partially, be contributed to errors in estimation of the rotation angle, leading to incorrect intrinsic parameters used to represent the camera during reprojection. An average error in the order of 1 pixel RMS is sufficient for any AR purposes of this laparoscope.

Two optical sensors are used to track the motion of the laparoscope. One is attached to the handle, and the other to the cylinder of the laparoscope. Rotation angle of the laparoscope is determined by solving a non-linear problem to find the center and axis of rotation.

The sensor attached to the cylinder of the laparoscopes is used as reference for the camera coordinate system as it has a fixed relation to the optical axis, and it being closest to the tip of the laparoscope produces smaller tracking errors. If our intrinsic camera model is used, the origin of the camera coordinate system produces a circular motion in the coordinate system of the reference sensor during rotation. The rotation angle of this circular motion was modeled using a Fourier sum to correct for the non-linearity in relation to the laparoscope's rotation angle. Origin of the camera coordinate system was modeled as a point rotating at a fixed radius around the center point of the circular motion with the axis of rotation normal to the center of the circular motion.

As the direction of the optical axis is fixed in relation to the reference sensor, the orientation of the camera coordinate system could be modeled by a rotation around the optical axis with an angle equal but in opposite direction of the rotation angle of the laparoscope. In the model, the orientation obtained in the default configuration of the laparoscope (oblique view directed downwards) was used as reference for the orientation at other viewing angles of the laparoscope.

The intrinsic camera model and hand-eye translation model have shown to be a good approximation of the truth on the calibration data of multiple laparoscopes. However, the camera orientation model showed hysteresis of rotation, and the overall model could not produce accurate results on validation images. In the validation images, the reprojection error increased to a maximum of 100 pixels RMS. The cause of this error was determined to be motion of the image sensor within the laparoscopes by changing the direction of scope rotation. Movement of the image sensor causes the camera coordinate system to move, making the hand-eye translation model incorrect. This freedom of motion permits changing focal lengths, hysteresis of the rotation, and the observed movements in camera coordinates.

Producing an accurate model for the laparoscope requires the pose of both parts of the camera system to be tracked. This was attempted by placing two optical sensors on the laparoscope. The sensor on the handle of the laparoscope was assumed to be fixed in relation to the image sensor, but this assumption is proven to be incorrect. As there is no other place accessible from the outside of the laparoscope to track the image sensor, it does not seem possible to create a model for this specific type of laparoscope based on external tracking. However, none of the laparoscopes was calibrated more than once. Repeating the calibration multiple times, and closer inspection of the behavior of the camera coordinate system, may reveal a pattern in the behavior of the image sensor in relation to rotation of the scope. This pattern can then be used to predict the behavior of the image sensor during rotation. This would require the rotation history of the laparoscope to be included in the model, making it substantially more complex. Inclusion of the rotation history would also require all rotations of the laparoscope to be measured by the optical tracking system, a challenging task that limits the usability in a clinical setting.

If we compare our laparoscope to the conventional laparoscopic setups used by other authors, we see that the issues in tracking are specific to our chip-on-the-tip system. In conventional setups, the image sensor is located and fixed within the camera head. This configuration allows accurate pose estimation of the image sensor and lens-systems, and therefore allows the laparoscope to be modeled by external tracking. However, most of the literature references are between 10-15 years old. Since the initial publication, none of the authors have released a follow up publication or shown a clinical application of the system. One of the possible reasons for this is that the issues experienced in our system are also present in the conventional system if used in the clinic. During clinical use of the laparoscope, the scope is introduced into the patient through a trocar that provides the pivot point for changing the laparoscopes view to the general direction of interest. The abdominal wall can be several centimeters thick. Pivoting of the laparoscope in the abdominal wall therefore requires force to be put on the camera head of the conventional laparoscope. This force can cause the camera head to tilt and maybe even move in relation to the optical system, leading to similar issues as experienced in our system. However, as there is no literature available on this topic, this explanation is no more than an assumption based on extrapolation of our results.

All experiments performed on the laparoscope were in a static setting where the acquisition delay between the laparoscope and optical tracking system played no role. Real-time AR application in a dynamic setting requires the delay between both systems to be known. A delay estimation procedure was developed that can accurately estimate the delay. In the procedure, an object is rotated at a constant angular frequency and tracked by both systems. In the images, the position of the object is estimated with subpixel accuracy using a template and normalized cross-correlation. A sinus was fitted to the constant rotation in both systems and the delay was determined by the phase delay between the two. The system was validated by introducing delays of up to 200 ms to the optical system. The system can estimate the delay accurate up to 5 ms. However, it was not validated if the delay between the two systems could be reduced to zero.

6.1 Recommendations

A lot of time and effort is put into the creation of a camera model. This has resulted in a detailed description of the behavior of the intrinsic camera parameters during rotation of the laparoscope. However, the camera model produces a reprojection error in the order of a pixel, even if the parameters are modeled incorrect. The role of the intrinsic model is therefore irrelevant on the scale of the currently obtained total reprojection error. In further attempts are made to create a laparoscope calibration method, the focus should therefore be aimed at describing the behavior of the image sensor's pose. The intrinsic parameters should only be evaluated to obtain extrinsic parameters that allow prediction of image sensor's behavior.

None of the laparoscopes used were calibrated more than once. This provided a lot of information on the functioning of the laparoscope itself, but did not provide an answer to how the laparoscope behaves over time. The calibration and validation procedure should be extended to include multiple calibrations, and rotation in both directions to investigate if there is a pattern in the behavior.

As the issues seem to be specific to the laparoscopes used, replacing the laparoscope could provide a solution. Olympus recently provided us with a newer version of the laparoscope, the EndoEye HD II. It would be worth wile to investigate if this laparoscope has similar properties as its predecessor, or that the image sensor is indeed fixed in relation to the handle in the new scope.

The delay estimation procedure has shown to estimate the accuracy of the delay up to 5 ms. However, the procedure was validated using a webcam, not a laparoscope, and the total delay ranged from 15-215 ms. It should be investigate if the procedure is still accurate when a laparoscope and other acquisition hardware is used, and if the delay can be reduced to zero ms.

Appendices

A. Standard camera model

To predict where an object will be visible in an image it is necessary to know the relationship between points in space and the projection on the image plane. The position of an object in space with respect to the camera can be determined by navigation. For the projection on to the image a camera model needs to be described. Ignoring the effects of lenses, the pinhole model gives an accurate description of image generation from a scene. This appendix first describes the pinhole model, followed by conversion of the distance metric to pixel dimensions. The information and images in this appendix are based on the book multiple view geometry in computer vision by Hartley and Zisserman [50].

Pinhole model

We consider a Euclidean coordinate system with a pinhole at its origin C. The *image plane*, or focal plane is located at a distance f from the pinhole in the Z-direction. Let $\mathbf{X} = (X, Y, Z)^T$ be a point in space. This point is projected on to the image plane at $\mathbf{x}$ where the line that connects point $\mathbf{X}$ to the origin intersects with the image plane. By similar triangles it is easy to see that the point $(X, Y, Z)^T$ is mapped to point $(fX/Z,\ fY/Z,\ f)^T$ on the image plane. By ignoring the final coordinate we can determine the position on image plane to be $\mathbf{x} = (fX/Z,\ fY/Z)^T$, Figure 40.

$$(X, Y, Z)^T \mapsto (fX/Z,\ fY/Z)^T \tag{7.1}$$

(7.1) describes the mapping of a point in Euclidean 3-space $\mathbb{R}^3$ to a point on the image plane in Euclidean 2-space $\mathbb{R}^2$. The pinhole is the point that all points in space are projected to. It is also called the optical center, or camera center. The XY-plane of the camera center is called the principal plane and it is parallel to the image plane. The line perpendicular to the principal plane at the origin is the principal axis. The principal axis intersects the image plane at the principal point $\mathbf{p}$.

The mapping in (7.1) can easily be performed in homogenous coordinates as the matrix multiplication

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{7.2}$$

Until now we have assumed the origin of the image plane to be located at the principal point. This does not need to be the case. For example, if the image plane is a camera sensor, the origin of the plane is usually defined in a corner of the sensor, Figure 41. If the principal point is located at $(x_0, y_0)^T$, in the image plane the mapping can be updated to

$$(X, Y, Z)^T \mapsto (fX/Z + x_0,\ fY/Z + y_0)^T \tag{7.3}$$

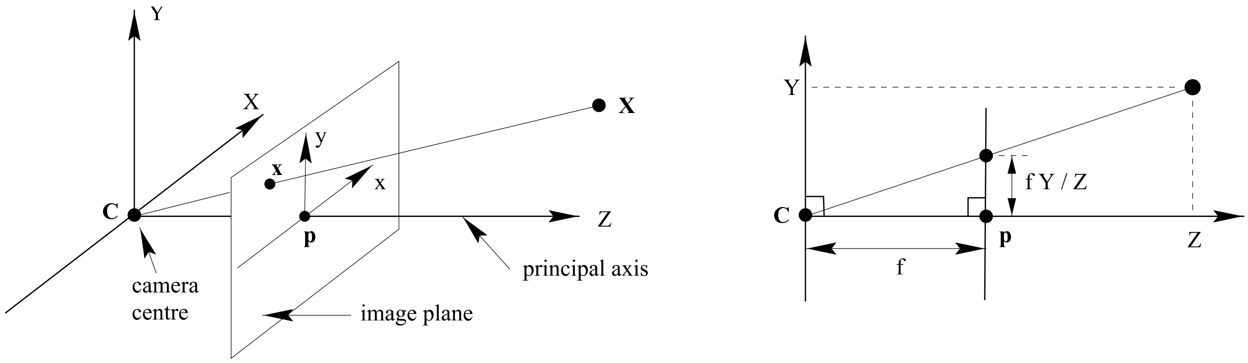


Figure 40: Point X in space is projected on the image plane at x where the line between X and pinhole C intersects the image plane. Coordinates for the position of point x on the image plane can be found using similar triangles. Optical or principal axis is in the Z-axis of the pinhole coordinate system. The image plane is perpendicular to the principal axis with its x- and y-axis parallel to the X- and Y-axis of the pinhole coordinate system. Intersection of the optical axis with the image plane is referred to as the principal point.

Or in terms of homogeneous coordinates

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX + Zx_0 \\ fY + Zy_0 \\ Z \end{pmatrix} = \begin{bmatrix} f & & x_0 & 0 \\ & f & y_0 & 0 \\ & & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{7.4}$$

This matrix multiplication can also be written as

$$\mathbf{x} = \mathrm{K}\left[\mathrm{I} \mid \mathbf{0}\right]\mathbf{X} \tag{7.5}$$

With I the 3x3 identity matrix, $\mathbf{0}$ a 1x3 null vector, $(X, Y, Z, 1)^T$ the point $\mathbf{X}$ given in homogenous coordinates, and K defines the *camera calibration matrix*, containing the intrinsic parameters of the camera calibration.

$$\mathrm{K} = \begin{bmatrix} f & & x_0 \\ & f & y_0 \\ & & 1 \end{bmatrix} \tag{7.6}$$

In the current mapping we assume that $\mathbf{X}$ is defined in the coordinate system of the camera with its center at $\mathbf{C}$. If $\mathbf{X}$ is defined in terms of a different Euclidean coordinate system, let's call it the world coordinate system and give it origin $\mathbf{O}$, we first need to transform point $\mathbf{X}$ from world coordinates to camera coordinates, Figure 41. The transformation in homogenous coordinates is given by

$$\mathbf{X}_{cam} = {}^{CC}T_w \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{X} \tag{7.7}$$

Here ${}^{CC}T_w$ is the transformation matrix that relates the world coordinate system to the camera coordinate system with $\tilde{\mathbf{C}}$ the position of camera center in world coordinates, and $\mathbf{R}$ the rotation matrix to get from world frame to camera frame. Combining (7.7) with (7.5), and setting $-\mathbf{R}\tilde{\mathbf{C}} = \mathbf{t}$ gives

$$\mathbf{x} = \mathrm{K}\left[\mathbf{R} \mid \mathbf{t}\right]\mathbf{X} \tag{7.8}$$

$\mathbf{R}, \mathbf{t}$ are the extrinsic parameters describing the pose of the camera coordinate system in the world coordinate system. The intrinsic and extrinsic parameters can be combined in to a 3x4 *camera projection matrix* P that relates points in space to its projection on the image plane

$$\mathbf{x} = \mathrm{P}\mathbf{X} \tag{7.9}$$

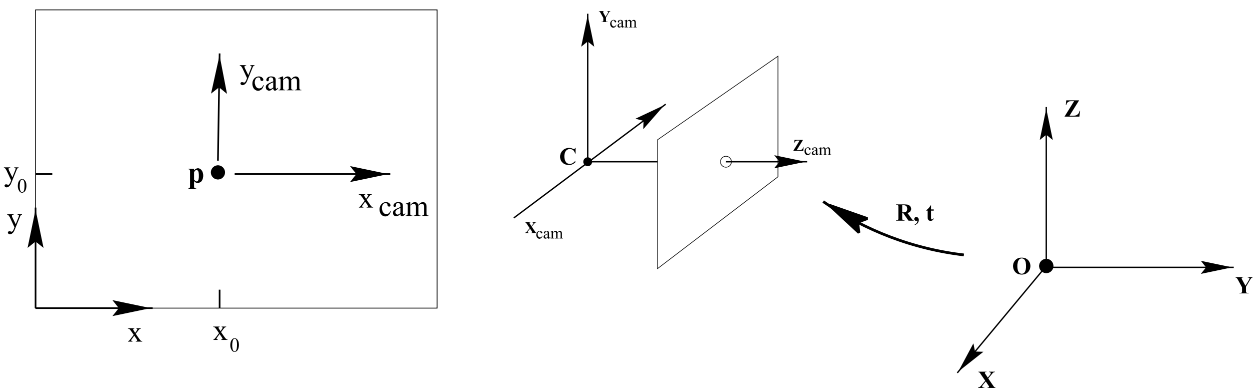


Figure 41: Left, with the origin of the sensor coordinate system in the bottom left corner, the principal point is located at pixel coordinates (x_0,y_0). Right, pose of the object defined in world coordinated with origin O is transformed to coordinates in the camera coordinate system using the extrinsic parameters R, t.

Pixel dimensions

In the pinhole model derived above the axes have equal scales in all directions. In dealing with images as captured by a camera it is easier to work in pixel or image coordinates. The number of pixels per unit distance in the x- and y-direction are given by m_x, and m_y respectively. An unequal scale factor can be applied in the orthogonal directions of the image plane. This scale factor expresses the focal length in terms of pixel dimensions in x- and y-direction as $\alpha_x = m_x f$, and $\alpha_y = m_y f$ respectively. Similarly the principal point is given in pixel dimensions by $\mathbf{p} = \left(m_x x_0, m_y y_0\right) = \left(u_0, v_0\right)$. The new camera calibration matrix is given by

$$\mathrm{K}=\begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix}\begin{bmatrix} f & & x_0 \\ & f & y_0 \\ & & 1 \end{bmatrix}=\begin{bmatrix} \alpha_x & & u_0 \\ & \alpha_y & v_0 \\ & & 1 \end{bmatrix} \tag{7.10}$$

Here we assume the axes of the pixels to be orthogonal. This is usually the case, however, if this assumption does not hold, a *skew* parameter needs to be added to the camera calibration matrix, Figure 42. Skew is defined as $s = tan(\gamma)/m_y$

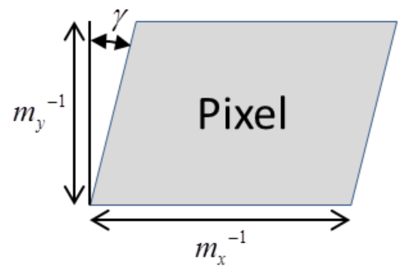


Figure 42: If the pixel axes are not orthogonal a skew parameter s needs to be added to the camera matrix.

$$\mathbf{x}=\begin{bmatrix} \alpha_x & s & u_0 \\ & \alpha_y & v_0 \\ & & 1 \end{bmatrix}[\mathbf{R} \mid \mathbf{t}]\mathbf{X}=\mathrm{K}[\mathbf{R} \mid \mathbf{t}]\mathbf{X}=\mathrm{P}\mathbf{X} \tag{7.11}$$

In words the projective mapping of (7.11) can be explain as, a transformation of point $\mathbf{X}$ from the world frame to the camera frame, followed by a mapping to the image plane expressed in pixel coordinates.

Distortions

In the pinhole model we ignored the effects of lenses on the projection of an object on the image plane. Even the most expensive and complicated lens-systems are to some extent affected by one or more types of optical aberrations. Depending on the intended application of a camera it is necessary to correct for these aberrations. In this case the application is accurately superimposing the image of a tumor on the images acquired by a laparoscope. Assuming a typical lens-system in which all types of aberrations are relative small, the only aberration we are interested in is geometrical distortion. Geometrical lens distortions are explained in the next appendix.

B. Lens distortions

In the field of optics it is common to make use of the paraxial approximation to describe a lens-system. In the paraxial approximation the assumption is made that for small angles $\sin\theta \approx \theta$ in the propagation of rays through the lens. While this assumption holds reasonably, a better approximation is achieved when higher order components of the Maclaurin series for the sine are included, (8.1).

$$\sin\theta = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\theta^{2n+1} = \theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \ldots \tag{8.1}$$

For small angles the contribution of the higher order components is minimal, but still enough to cause aberrations in the image. Analytical derivation of the five aberrations originating from the third-order component of the series expansion were published by Seidel in 1857. Of the five third order aberrations, spherical, coma, astigmatism, field curvature, and distortion; the last one is the only one of geometrical interest. The geometrical distortion by this aberration component causes a displacement of a point on the image plane. In order to accurately project the patient model on to the laparoscopic images it is necessary to account for this displacement by either correcting distortions in the image, or distorting the patient model to match the image.

Radial distortion

Geometrical distortion as derived by Seidel is, in camera calibration, usually referred to as radial distortion because it is radially symmetrical from the principal axis outwards, if the lens is radially symmetrical. Radial distortion Δr is given by

$$\Delta r = gd\alpha^3 \rho \cos\varphi \tag{8.2}$$

Here, g is a lens specific aberration coefficient, d the pupil radius, ρ the relative pupil radius (between 0 and 1), α the field angle (the angle between the chief ray an the optical axis), and φ the pupil angle measured from the positive x-axis, Figure 43. Component $\rho\cos\varphi$ are the pupil coordinates (ρ, φ) that indicate the position where a specific ray intersects the aperture in the optical system. In radial distortion this component guarantees that all rays originating from a single point on the object will focus at the same point on the image plane no matter where the ray intersects the aperture.

Without loss of generality we can rewrite the radial displacement of a point on the image plane as

$$\Delta r = b\alpha^3 = b\arctan^3\left(\frac{r}{f}\right) \tag{8.3}$$

Here, b is a constant, f the distance between aperture and image plane, and r the distance from the principal point to the undistorted image point.

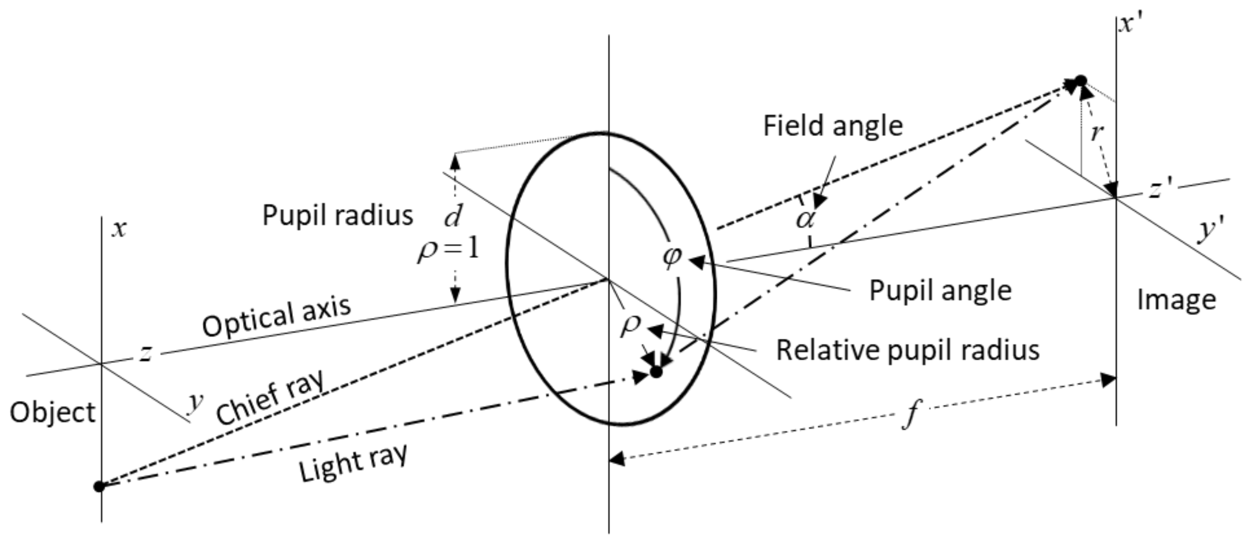


Figure 43: A ray from the object passes through the pupil at polar coordinates (ϕ,ρ), and intersects the image plane at a distance r from the principal point. The image plane is positioned at a distance f to the lens. The field angle is the angle between the optical axis and chief ray, the ray that passes through the optical center.

The arctan can be rewritten as a series expansion with f a constant

$$\arctan\left(\frac{r}{f}\right)=\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1}\left(\frac{r}{f}\right)^{2n+1}=\sum_{n=0}^{\infty}K_n r^{2n+1},\ K_n=\frac{(-1)^n}{(2n+1)\,f^{2n+1}} \tag{8.4}$$

Combining (8.3) and (8.4) gives the distortion as a distance from the ideal para-axial point along the radial line coming from the principal point.

$$\Delta r=b\left(\sum_{n=0}^{\infty}K_n r^{2n+1}\right)^3=\sum_{n=1}^{\infty}k_n r^{2n+1}=k_1r^3+k_2r^5+k_3r^7+\ldots \tag{8.5}$$

Here, coefficients k_n are a product of b and the cubed arctan coefficients. This distortion can also be given in separate x, y coordinates on the image plane given that $x=r\cos\phi,\ y=r\sin\phi$

$$\begin{bmatrix}\Delta x_r\\ \Delta y_r\end{bmatrix}=\Delta r\begin{bmatrix}\cos\phi\\ \sin\phi\end{bmatrix}=\frac{\Delta r}{r}\begin{bmatrix}x\\ y\end{bmatrix}=\left(k_1r^2+k_2r^4+k_3r^6+\ldots\right)\begin{bmatrix}x\\ y\end{bmatrix}=K(r)\begin{bmatrix}x\\ y\end{bmatrix} \tag{8.6}$$

Radial distortion is given by an endless series expansion, but is generally approximated in camera calibration by the first two or three coefficients, setting the rest to zero.

The overall effect of radial distortion is a change in magnification with distance to the principal point. As a consequence the amount of distortion increases with distance to the principal point, and the largest effects can be seen on the edges of the image. Depending on the type of lens, aberration coefficient g can be positive or negative. The sign of the aberration coefficient determines if a distorted point is closer

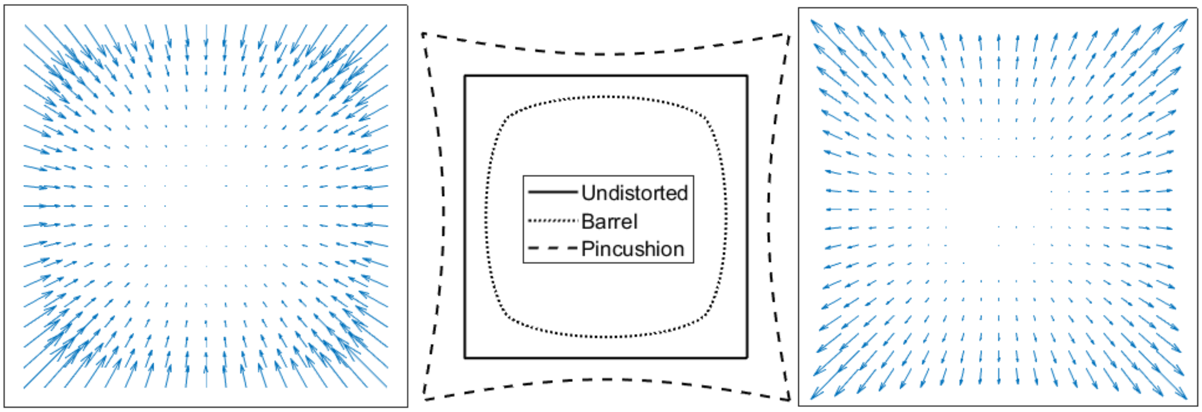


Figure 44: Left, barrel distortion; right, pincushion distortion. The displacement vectors start at the ideal undistorted point, and end at the distorted position. It can easily be seen that the amount of distortion increases with distance to the image center, or principal point. For simplicity, only the first distortion coefficient is used.

or further away from the principal point compared to the ideal undistorted point. If the distortion causes a decrease in magnification, barrel distortion can be observed. An increase in magnification along the radial lines will result in pincushion distortion, Figure 44.

Decentering distortion

Aberrations as derived by Seidel originate from the curvature of a lens, but do not account for lens-placement. However, it is difficult, if not impossible, to perfectly align all lenses and the image plane. If a lens is slightly tilted, so the optical axes of the lenses are no longer collinear, the projected points of the object on the image plane suffer a displacement with regard to the ideal situation, Figure 45. The analytical derivation of geometrical distortion due to decentering of lenses was first published by Conrady in 1919 [51]. His work has long gone unnoticed as the community used the thin prism model to correct for distortions created by decentering of lenses. His work was rediscovered by Brown who showed that the thin prism model and decentering model were analytically the same under certain conditions. However, since the decentering model of Conrady originated from an analytical derivation by ray tracing it gives a better description of the distortion. Brown therefore advocated to abandon the thin prism model and only use the decentering model [46]. In literature the model is usually referred to as the Brown-Conrady model.

Decentering distortion Δd consists of two components. It has a radial component Δr that acts along the radial lines from the principal point, and a tangential component Δt that acts perpendicular to the radial lines, Figure 46. Unlike radial distortion, decentering distortion is not radially symmetrical, but its magnitude and direction depend on the amount and axis of lens tilting. It is important to note that the radial component in decentering distortion is not the same as the radial distortion as derived by Seidel. This naming ambiguity has resulted in decentering distortion to be referred to by some as tangential distortion. Decentering distortion is given by

$$\left.\begin{aligned}\Delta r &= 3P(r)\sin(\theta-\theta_0)\\ \Delta t &= P(r)\cos(\theta-\theta_0)\end{aligned}\right\}\Delta d = \Delta r + \Delta t \tag{8.7}$$

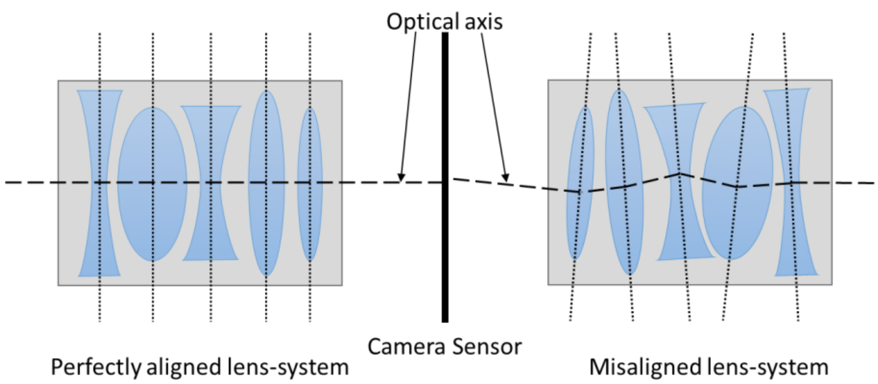


Figure 45: In an ideal case all lenses in the lens-system have their optical axes aligned, and the principal planes are parallel to the camera sensor. In a misaligned, or decentered lens-system, projected points suffer a displacement due to this decentering.

Here, $P(r)$ is the tangential profile, θ the angle from the positive x-axis, and θ_0 the angle to the axis of maximum tangential distortion, the axis on which the lens is tilted, Figure 46. Decentering distortion can be expressed in Euclidean coordinates given that $x = r\cos\theta,\ y = r\sin\theta$ for the radial component, and noting that the radial component precedes the tangential component by 90 degrees ny

$$\begin{bmatrix} \Delta x_d \\ \Delta y_d \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} \Delta r \\ \Delta t \end{bmatrix} = P(r) \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} 3\sin(\theta-\theta_0) \\ \cos(\theta-\theta_0) \end{bmatrix} \tag{8.8}$$

$$\begin{bmatrix} \Delta x_d \\ \Delta y_d \end{bmatrix} = \frac{P(r)}{r^2} \begin{bmatrix} 2xy & r^2+2x^2 \\ r^2+2y^2 & 2xy \end{bmatrix} \begin{bmatrix} \cos\theta_0 \\ -\sin\theta_0 \end{bmatrix} \tag{8.9}$$

With $r^2 = x^2 + y^2$. The tangential profile as derived by Conrady is given by

$$P(r) = p_3 V^2 \tag{8.10}$$

Here p_3 is a constant, and V the field angle (same as α in radial distortion). $P(r)$ can be rewritten in a similar form as (8.4) to get a series expansion for the tangential profile.

$$P(r) = p_3 V^2 = p_3 \arctan^2\left(\frac{r}{f}\right) = \sum_{n=1}^{\infty} J_n r^{2n} = J_1 r^2 + J_2 r^4 + J_3 r^6 + \ldots \tag{8.11}$$

Here coefficients J_n are a product of p_3 and the squared arctan coefficients. In camera calibration, all coefficients larger than J_1 are usually set to zero to get an approximation of the decentering distortion. Combining (8.9) with (8.11) results in the standard form used in camera calibration.

$$\begin{bmatrix} 2xy & r^2+2x^2 \\ r^2+2y^2 & 2xy \end{bmatrix} \begin{bmatrix} J_1\cos\theta_0 \\ -J_1\sin\theta_0 \end{bmatrix} = \begin{bmatrix} 2xy & r^2+2x^2 \\ r^2+2y^2 & 2xy \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \tag{8.12}$$

It is important to note that there is no clear convention for the order of numbering p_1 and p_2. One should always check if the parameters are defined as they are here, or with the numbers reversed.

Lens distorted points

By combination of the radial and decentering distortion, the distorted position $p_{distorted}$ of an ideal undistorted point p_{ideal} on the image plane is given by

$$p_{distorted} = p_{ideal} + \Delta r + \Delta d \tag{8.13}$$

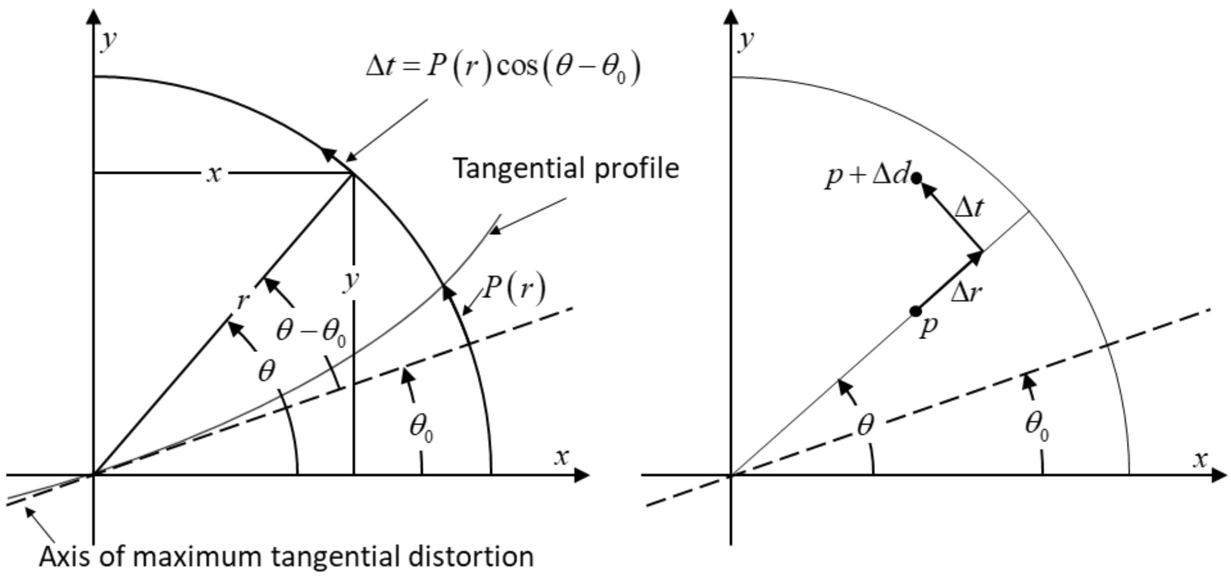


Figure 46: Decentering distortion. Left, an adaptation of Brown's image illustrating the geometrical significance of parameters defining tangential distortion [46]. Right, illustrates decentering distortion of point p to point p + Δd.

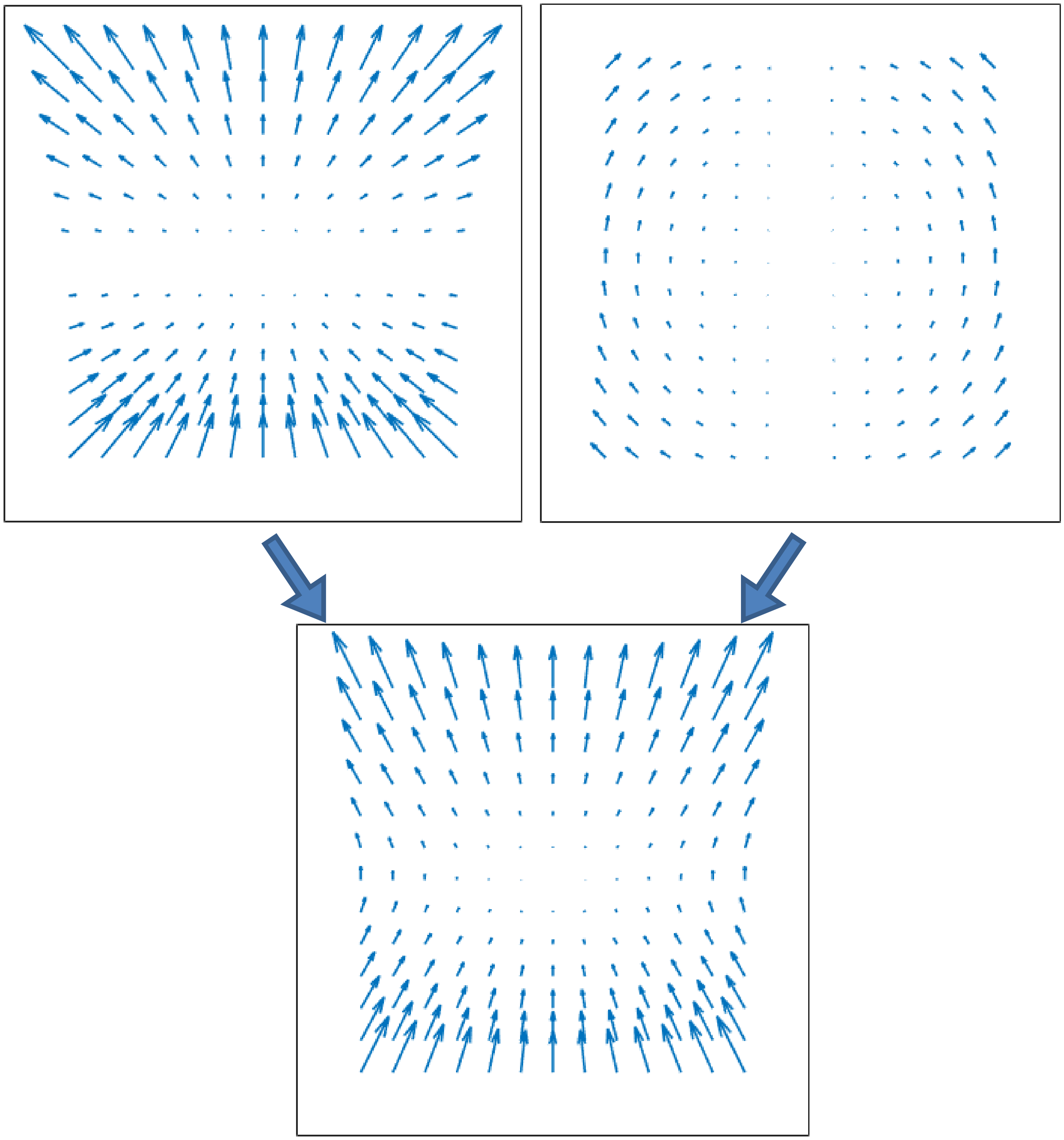


Figure 47: Top, radial and tangential components of decentering distortion at $\theta_0 = 0$. Distortion vectors point from the ideal undistorted points to distorted position of the points. The magnitude of the radial component is three times as large as the tangential component. Bottom, summation of the components gives the total decentering distortion.

C. Center and axis of a rotating 3D point

A circular motion in 3D can be described by the center and axis of rotation. Here we solve a non-linear least squares problem to find the center and axis corresponding to a set of 3D coordinates, ignoring orientation of the point. As the cost function of the method gives more weight to larger errors, it is important to remove large outliers before the problem is solved.

Rotation axis direction is determined by using the principal axis method to find the normal $\mathbf{n}$ to the data. Let $\mathbf{r}^i \left(i = 1, 2, 3, \ldots, n\right)$ be the *i*-th measured 3D position, and $\bar{\mathbf{r}}$ the measurements mean coordinates. Then, $\mathbf{r}_g^i = \mathbf{r}^i - \bar{\mathbf{r}}$ denotes the coordinates shifted to their gravity center. Normal is determined by the eigenvector corresponding to the lowest eigenvalue of the 3x3 matrix $\sum_{i=1}^{n} \mathbf{r}_g^{i\,\mathrm{T}} \cdot \mathbf{r}_g^i$, and is found using singular value decomposition. Letting ρ be the radius of the circle, the circle center $\mathbf{c}$ is found by minimizing

$$\sum_{i=1}^{n} \left\{ \rho^2 - \left|\mathbf{r}^i - \mathbf{c}\right|^2 + \left(\mathbf{n} \bullet \left(\mathbf{r}^i - \mathbf{c}\right)\right)^2 \right\}^2 \tag{9.1}$$

Here • is the inner product. This non-linear least squares problem can be solved using the Levenberg-Marquardt method [41].

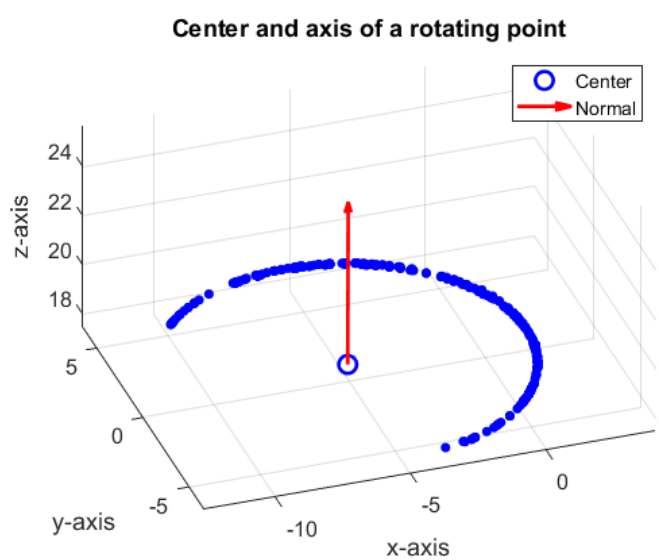


Figure 48: Center and axis of a set of points describing a circle obtained using non-linear least squares.

Bibliography

1. *Cijfers over kanker*. Integraal Kankercentrum Nederland 2015 [cited 2016 October 26]; Available from: http://www.cijfersoverkanker.nl/selecties/Dataset_1/img58108d449f1d6.
2. Hol, L., et al., *Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy.* Gut, 2010. **59**(01): p. 62-68.
3. van Maag, D.d.N.V., et al., *Nederlandse Richtlijn Coloscopie Surveillance.* 2013.
4. Auditing, D.I.f.C., *Jaarrapportage 2015*.
5. Boulis-Wassif, S., et al., *Final results of a randomized trial on the treatment of rectal cancer with preoperative radiotherapy alone or in combination with 5-fluorouracil, followed by radical surgery trial of the european organization on research and treatment of cancer gastrointestinal tract cancer cooperative group.* Cancer, 1984. **53**(9): p. 1811-1818.
6. Fisher, B., et al., *Postoperative adjuvant chemotherapy or radiation therapy for rectal cancer: results from NSABP protocol R-011.* Journal of the National Cancer Institute, 1988. **80**(1): p. 21-29.
7. Gerard, A., et al., *Preoperative radiotherapy as adjuvant treatment in rectal cancer. Final results of a randomized study of the European Organization for Research and Treatment of Cancer (EORTC).* Annals of surgery, 1988. **208**(5): p. 606.
8. Martling, A.L., et al., *Effect of a surgical training programme on outcome of rectal cancer in the County of Stockholm.* The Lancet, 2000. **356**(9224): p. 93-96.
9. Wibe, A., et al., *A national strategic change in treatment policy for rectal cancer—implementation of total mesorectal excision as routine treatment in Norway. A national audit.* Diseases of the colon & rectum, 2002. **45**(7): p. 857-866.
10. Kapiteijn, E., H. Putter, and C. Van de Velde, *Impact of the introduction and training of total mesorectal excision on recurrence and survival in rectal cancer in The Netherlands.* British journal of surgery, 2002. **89**(9): p. 1142-1149.
11. Peeters, K.C., et al., *The TME trial after a median follow-up of 6 years: increased local control but no survival benefit in irradiated patients with resectable rectal carcinoma.* Annals of surgery, 2007. **246**(5): p. 693-701.
12. MacFarlane, J., R. Ryall, and R. Heald, *Mesorectal excision for rectal cancer.* The lancet, 1993. **341**(8843): p. 457-460.
13. Enker, W.E., *Total mesorectal excision–the new golden standard of surgery for rectal cancer.* Annals of medicine, 1997. **29**(2): p. 127-133.
14. van der Pas, M.H., et al., *Laparoscopic versus open surgery for rectal cancer (COLOR II): short-term outcomes of a randomised, phase 3 trial.* The lancet oncology, 2013. **14**(3): p. 210-218.
15. Nagtegaal, I.D. and P. Quirke, *What is the role for the circumferential margin in the modern treatment of rectal cancer?* Journal of Clinical Oncology, 2008. **26**(2): p. 303-312.
16. Gosens, M.J., et al., *Circumferential margin involvement is the crucial prognostic factor after multimodality treatment in patients with locally advanced rectal carcinoma.* Clinical Cancer Research, 2007. **13**(22): p. 6617-6623.
17. Shihab, O.C., et al., *Patients with low rectal cancer treated by abdominoperineal excision have worse tumors and higher involved margin rates compared with patients treated by anterior resection.* Diseases of the Colon & Rectum, 2010. **53**(1): p. 53-56.
18. Kennecke, H., et al., *Predictors of positive radial margin status in a population based cohort of patients with rectal cancer.* Current Oncology, 2008. **15**(2): p. 56-61.

19. Marijnen, C., et al., *Radiotherapy does not compensate for positive resection margins in rectal cancer patients: report of a multicenter randomized trial.* International Journal of Radiation Oncology* Biology* Physics, 2003. **55**(5): p. 1311-1320.
20. den Dulk, M., et al., *Risk factors for adverse outcome in patients with rectal cancer treated with an abdominoperineal resection in the total mesorectal excision trial.* Annals of surgery, 2007. **246**(1): p. 83-90.
21. Kusters, M., et al., *Patterns of local recurrence in rectal cancer; a study of the Dutch TME trial.* European Journal of Surgical Oncology (EJSO), 2010. **36**(5): p. 470-476.
22. Lange, M.M. and C.J. van de Velde. *Long-term anorectal and urogenital dysfunction after rectal cancer treatment*. in *Seminars in Colon and Rectal Surgery*. 2010. Elsevier.
23. Celentano, V., et al., *Prospective study of sexual dysfunction in men with rectal cancer: feasibility and results of nerve sparing surgery.* International journal of colorectal disease, 2010. **25**(12): p. 1441-1445.
24. Mezger, U., C. Jendrewski, and M. Bartels, *Navigation in surgery.* Langenbeck's Archives of Surgery, 2013. **398**(4): p. 501-514.
25. Nijkamp, J., et al. *Image-guided navigation surgery for pelvic malignancies using electromagnetic tracking*. in *SPIE Medical Imaging*. 2016. International Society for Optics and Photonics.
26. Merriam-Webster, *Merriam-Webster's collegiate dictionary*. 2004: Merriam-Webster.
27. Steinhaus, H., *Sur la localisation au moyen des rayons x.* Comptes Rendus de L'Acad. des Sci, 1938. **206**: p. 1473-1475.
28. Bajura, M., H. Fuchs, and R. Ohbuchi. *Merging virtual objects with the real world: Seeing ultrasound imagery within the patient*. in *ACM SIGGRAPH Computer Graphics*. 1992. ACM.
29. Sutherland, I.E. *A head-mounted three dimensional display*. in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*. 1968. ACM.
30. Sielhorst, T., M. Feuerstein, and N. Navab, *Advanced medical displays: A literature review of augmented reality.* Journal of Display Technology, 2008. **4**(4): p. 451-467.
31. Shahidi, R., et al., *Implementation, calibration and accuracy testing of an image-enhanced endoscopy system.* IEEE Transactions on Medical imaging, 2002. **21**(12): p. 1524-1535.
32. Nicolau, S., et al., *Augmented reality in laparoscopic surgical oncology.* Surgical oncology, 2011. **20**(3): p. 189-201.
33. Wieters, M., *Contact-free magnetic coupling for an endoscope*. 2013.
34. Fuchs, H., et al. *Augmented reality visualization for laparoscopic surgery*. in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 1998. Springer.
35. Khadem, R., et al. *Endoscope calibration and accuracy testing for 3d/2d image registration*. in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2001. Springer.
36. De Buck, S., et al. *A system to support laparoscopic surgery by augmented reality visualization*. in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2001. Springer.
37. Smith, W.E., N. Vakil, and S.A. Maislin, *Correction of distortion in endoscope images.* IEEE Transactions on Medical Imaging, 1992. **11**(1): p. 117-122.
38. Tsai, R., *A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses.* IEEE Journal on Robotics and Automation, 1987. **3**(4): p. 323-344.
39. Zhang, Z., *A flexible new technique for camera calibration.* IEEE Transactions on pattern analysis and machine intelligence, 2000. **22**(11): p. 1330-1334.
40. Bouguet, J.-Y., *Camera calibration toolbox for matlab.* 2004.

41. Yamaguchi, T., et al., *Development of a camera model and calibration procedure for oblique-viewing endoscopes.* Computer Aided Surgery, 2004. **9**(5): p. 203-214.
42. Wu, C., B. Jaramaz, and S. Narasimhan, *A full geometric and photometric calibration method for oblique-viewing endoscopes.* Computer Aided Surgery, 2010. **15**(1-3): p. 19-31.
43. De Buck, S., et al., *Evaluation of a novel calibration technique for optically tracked oblique laparoscopes.* Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007, 2007: p. 467-474.
44. Liu, X., C.E. Rice, and R. Shekhar, *Fast calibration of electromagnetically tracked oblique-viewing rigid endoscopes.* International Journal of Computer Assisted Radiology and Surgery, 2017: p. 1-11.
45. Melo, R., J.P. Barreto, and G. Falcao, *A new solution for camera calibration and real-time image distortion correction in medical endoscopy–initial technical evaluation.* IEEE Transactions on Biomedical Engineering, 2012. **59**(3): p. 634-644.
46. Brown, D.C., *Decentering distortion of lenses.* Photometric Engineering, 1966. **32**(3): p. 444-462.
47. Lasso, A., et al., *PLUS: open-source toolkit for ultrasound-guided intervention systems.* IEEE Transactions on Biomedical Engineering, 2014. **61**(10): p. 2527-2537.
48. Tokuda, J., et al., *OpenIGTLink: an open network protocol for image-guided therapy environment.* The International Journal of Medical Robotics and Computer Assisted Surgery, 2009. **5**(4): p. 423-434.
49. Kim, T., et al., *The effect of delayed visual feedback on telerobotic surgery.* Surgical Endoscopy and Other Interventional Techniques, 2005. **19**(5): p. 683-686.
50. Hartley, R. and A. Zisserman, *Multiple view geometry in computer vision*. 2003: Cambridge university press.
51. Conrady, A., *Lens-systems, Decentered.* Monthly notices of the royal astronomical society, 1919. **79**: p. 384-390.