

Projective Camera Model in Biomedical Navigation Applications

Inauguraldissertation

zur Erlangung der Würde eines

Dr. sc. med.

vorgelegt der Medizinischen Fakultät der Universität Basel

von

Adrian Schneider

aus Gunten, Kanton Bern

Basel, 2016

Original document stored on the publication server of the
University of Basel <http://edoc.unibas.ch/>

Genehmigt von der Medizinischen Fakultät auf Antrag von

Prof. Dr. Philippe C. Cattin
Dissertationsbetreuer, Fakultätsverantwortlicher

Prof. Dr. med. Stephen Wyler
Koreferent

Prof. Dr. med. Rachel Rosenthal
Koreferentin

Prof. Dr. Nicolas Padoy
Externer Gutachter

Basel, den 16. September 2016

Prof. Dr. Thomas Gasser
Dekan

Acknowledgements

I would like to acknowledge my supervisor Prof. Dr. Philippe C. Cattin for giving me the opportunity to make an interesting master thesis and to pursue a PhD in his group. It is a pleasure to be guided by somebody with his passion, scientific background, sense for the essentials and tremendous practical experience. We, his PhD students, use to joke, that it is impossible to find a topic in computer vision Philippe does not know about it yet. I would also like to acknowledge my colleagues Simon Pezold and Mathias Griessen for their remarkable support during the last three years. Further, I would like to thank all the members of our lab for the numerous lively discussions I could participate in.

Finally, I want to thank my family and my friends for their support and inducement during my study time.

Summary

Surgical interventions pose very high demands on the spatial sense of surgeons. Especially in complex surgeries, navigation support is an important factor for the success of an operation. With the introduction of computers in surgical navigation, the field of computer assisted surgery was born. In difference to former mechanical constructions, these systems are highly adaptive and enable the development of versatile medical applications. This thesis examines different aspects of integrating a camera into surgical 3D navigation systems. Depending on the context, a camera can perform a self-localization or has to be tracked. In either case, this information is used to build intuitive visualization interfaces, where the operation planning and the reality are fused in a spatially correct manner. The accuracy and reliability of such augmented reality navigation applications is a major concern in this thesis. However, besides the high level incorporation of existing algorithms, this thesis also investigates the projective transformation at its core. It is shown that the position and behavior of a controllable tilting mirror can be described compactly with the projective camera model.

Zusammenfassung

Operative Eingriffe stellen sehr hohe Anforderungen an den Orientierungssinn von Chirurgen. Besonders bei komplizierten Operationen kann die Verwendung eines Navigationssystems zum Erfolg eines Eingriffes beitragen. Mit der Einführung von computerbasierten Navigationssystemen wurde das Gebiet der computerunterstützten Chirurgie geboren. Im Gegensatz zu den früheren, mechanischen Konstruktionen sind diese neuen Navigationssysteme leicht anpassbar und können vielseitig eingesetzt werden. In dieser Doktorarbeit werden verschiedene Aspekte der Integration einer Kamera in chirurgische 3D-Navigationssysteme untersucht. Je nach Umfeld kann eine Kamera ihre räumliche Position selbst bestimmen oder muss lokalisiert werden. In beiden Fällen lassen sich hilfreiche Visualisierungen erstellen, die einen virtuellen Operationsplan perspektivisch richtig mit der Realität verknüpfen. Die Genauigkeit und Zuverlässigkeit von solchen *Augmented-Reality*-Navigationssystemen ist ein zentrales Thema dieser Doktorarbeit. Ein weiteres wesentliches Thema ist die Untersuchung der Projektionstransformation. Es konnte gezeigt werden, dass die Position und Steuereigenschaften eines kippbaren Spiegels mittels eines projektiven Kameramodells kompakt beschrieben werden kann.

Contents

Contents	iv
Nomenclature and Abbreviations	vii
1 Introduction	1
1.1 Computer Assisted Surgery	2
1.2 Contribution of this Thesis	2
2 Transformations	6
2.1 Rigid Transformation	6
2.1.1 Determining a Rigid Transformation	8
2.2 Projective Transformation	9
2.2.1 Projective Transformation within a Camera	10
2.2.2 Camera Calibration	12
2.2.3 Camera Tracking	13
2.3 Transformation Errors	14
3 Publications	17
4 Augmented Reality Assisted Laparoscopic Partial Nephrectomy	19
5 Augmented Reality Assisted Brain Tumor Extraction in Mice	28
6 Landmark-Based Surgical Navigation	39
7 Direct Calibration of a Laser Ablation System in the Projective Voltage Space	48

CONTENTS

8	Conclusions	57
9	Outlook	60
	References	66

Nomenclature

Mathematical Notation

$\{X\}$	coordinate system labeled as X
${}^Y T_X$	transformation T from $\{X\} \rightarrow \{Y\}$
${}^Y P$	point P expressed in $\{Y\}$
$(T)^{-1}$ and T^{-1}	inverse of matrix T
$\ V\ $	Euclidean length or 2-norm of a vector V
$ N $	absolute value of a number N
$A \times B$	cross product of the vectors A and B

Nomenclature

Acronyms

MRI	magnetic resonance imaging
CT	computed tomography
AR	augmented reality
CS and COS	coordinate system
GPU	graphics processing unit
DLT	direct linear transformation
LPN	laparoscopic partial nephrectomy
PnP	perspective-n-points
TRE	target registration error

Chapter 1

Introduction

Performing surgical interventions demands a skilled team and a dedicated environment. The steadily increasing complexity in surgical interventions poses very high demands to surgeons and mistakes are likely to happen, inasmuch as there are more than one million surgery-related deaths world-wide per year [1]. The circumstances of those incidents are versatile. A commonly known reason are complications caused by infections, which can be reduced by implementing and following a strict sterilization concept. A less popular example are the challenges in anesthesia, such as the weighting of the patient's physical condition against the need of a surgery. However, in either case the equipment in an operation room is of importance. Using a surgical device less prone to contamination or monitoring an increased picture of the patient's vital functionalities enhances the safety of a surgical intervention. Thus, it is important to keep the medical devices in the operation room up to date.

The innovation potential in the medical field is high. Since there is a public interest and an even stronger economical driver, plenty of sophisticated medical devices are released to the market every year. Some of this products incorporate latest technology and might enable the surgical community to perform new treatment methods. Some other devices are less spectacular in terms of surgical leaps, but might just enhance the quality and safety of an operation.

The focus of this thesis are medical devices which support the surgical team in terms of 3D navigation and guarantee a certain accuracy during the intervention. Depending on the type of a surgery, such a surgical navigation system is an

indispensable necessity or it can be applied as a supportive aid. As an example, in neurosurgery it is beyond imagination to perform certain interventions without navigation. But in the field of orthopedics or surgeries on abdominal organs, the spatial sense of a surgeon often meets the required accuracy and navigation system are rarely applied.

1.1 Computer Assisted Surgery

With the introduction of computers in surgical navigation, the field of computer assisted surgery was born. In contrast to the former mechanical constructions, those systems are highly adaptive and can be extended with an arbitrary number of additional functionalities. A typical computer assisted surgery application is the fusion of preoperative anatomical images (CT, MRI), operation planning data and real-time 3D positions of surgical tools. This delivers detailed and accurate information, which can be presented in a comprehensive way to the surgeon, such as a virtual 3D environment as illustrated in Fig. 1.1. It has been shown by several studies [2, 3] that applying intraoperative 3D navigation greatly supports the surgeon and significantly reduces the risk for the patient.

A computer assisted surgical navigation system basically consists of two parts, the tracking device and the processing unit. The tracking device measures the 3D pose of its dedicated markers. This can be achieved by various technologies, such as electromagnetic or stereo optical tracking. The processing unit is a highly surgery specific software application, which receives, processes and visualizes 3D position data delivered from the tracking system. Even though there are versatile applications, the used principles are similar and the applied algorithms are identical.

1.2 Contribution of this Thesis

External Tracking The original purpose of this thesis was targeting the processing side. In particular, a navigation system for laparoscopic partial nephrectomy was built and investigated. As a tracking system, an off-the-shelf electromagnetic tracking device was used. The applied visualization modality is aug-

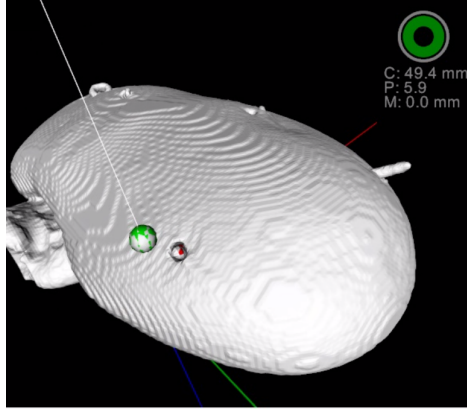


Figure 1.1: Surgical navigation view. The 3D environment shows a pig kidney, a surgical tool and real-time position information.

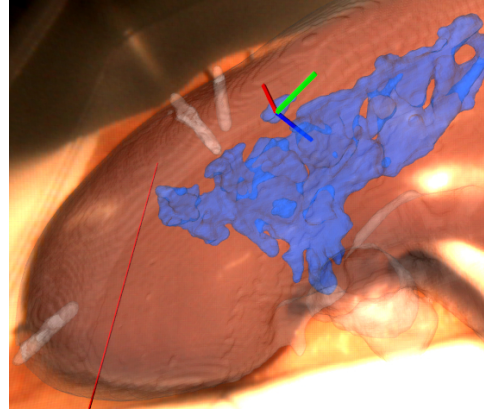


Figure 1.2: Augmented reality assisted navigation on a pig kidney, in which the segmented renal pelvis is overlaid in blue over image.

mented reality, an intuitive technique to present virtual objects embedded in a perspective correct manner in the reality, the video pictures from a camera. The main advantage of this technique is that it can be naturally incorporated into the visual perception. Adopting augmented reality in minimally invasive surgeries is of special interest, inasmuch as laparoscopic or endoscopic cameras are anyway used by the surgeon. Extending those systems with embedded navigation information is a great support. Figure 1.2 illustrates the developed augmented reality assisted surgical navigation system for the partial nephrectomy.

Self-Localization By applying related methods as required for the above augmented reality application, one can localize a camera only based on the content of the camera image itself. As a consequence, there is no need for an expensive tracking system anymore. On the other side, this approach requires the tracked objects to be part of the camera image. This is a major disadvantage and limits the field of applications in medicine. However, two suitable medical scenarios were examined, in which self-localization of a camera is employable.

In a first application, a highly accurate navigation system for the tumor removal in mice brains was developed. This system serves as a sample extraction tool. In a later stage, high resolution imaging is applied to reconstruct the inter

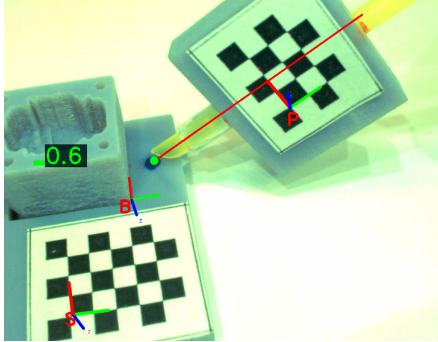


Figure 1.3: Augmented reality assisted mice brain tumor extraction system. Real-time distance information between tool and target is shown.

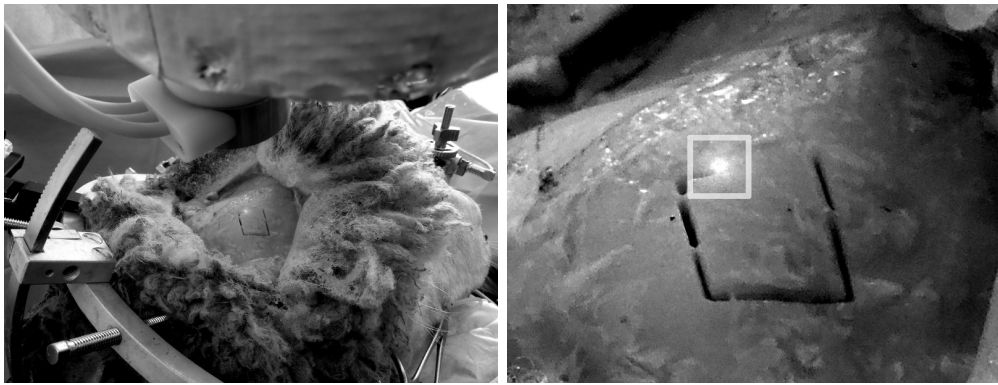


Figure 1.4: Landmark-based navigation system running on a tablet computer. Displays tumor in Mandibulae region and the Alveolaris nerve.

tumor blood vessel system in these samples. The navigation is performed based on a single camera, which tracks two chessboard markers (Fig. 1.3).

A second application demonstrates a low cost navigation system running on an average tablet computer. The intended use lies in the field of *Cranio-Maxillo-Facial* surgery. The navigation is based on tracking artificial landmarks on the target's surface with only the built in camera of the tablet computer. In order to run the application in real-time, an efficient point pattern matching algorithm was invented and heavy computational tasks were outsourced to the GPU. As visualization, augmented reality is used (Fig. 1.4).

Projective Model The thereby gained experience in computer vision and projective geometry enabled us to develop a calibration method for a laser deflecting tilting mirror, which is part of a laser ablation system mounted on a robot (Fig. 1.5). It could be shown, that under certain conditions a deflection mirror can be approximately described by the projective camera model. As a result, a common camera calibration algorithm can be used to calibrate the laser ablation system as well.



(a) Navigated laser cut on a sheep head. (b) Enlarged view of the cutting region. Bright spot is the laser.

Figure 1.5: Robot mounted laser ablation system.

Chapter 2

Transformations

This thesis builds on operations in the 3D space. To align the reader with the necessary knowledge of transformations and familiarize him with the used nomenclature, a brief summary about *rigid transformations*, *projective transformations* and *transformation errors* is presented in the following.

A commonly used knowledge base in computer vision and often cited in this thesis, is the book *Multiple View Geometry in Computer Vision* [4]. Even though the core matter of that book is about projective transformations, it offers a good introduction to arbitrary transformations as well. For this thesis, the relevant content are Part 0 and Part 1.

2.1 Rigid Transformation

A rigid transformation describes a rotational and a translational relationship between two coordinate systems (CS). In this thesis, rigid transformations are often denoted by the less specific term *affine transformation*, which incorporates shearing and scaling in addition to rotation and translation. A rigid transformation T in the 3D space can be compactly represented by a 4×4 matrix

$$T = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

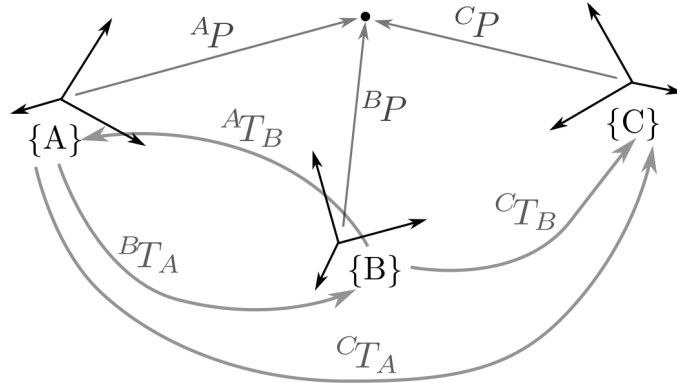


Figure 2.1: Rigid transformations between the coordinate systems $\{A\}$, $\{B\}$ and $\{C\}$. Curved arrows denote rigid transformations. Straight arrows are vectors.

where r_{11}, \dots, r_{33} is the rotation part and the translation is given by t_1, t_2, t_3 . Introducing the additional row $[0, 0, 0, 1]$ lifts the transformation to a 4th dimension, in which also the 3D translation can be applied in form of a linear operation. This is known as *homogeneous transformation matrix*.

The notation ${}^B T_A$ denotes a rigid transformation matrix T that transforms a point expressed in the $\{A\}$ CS ${}^A P$ into a position relative to the $\{B\}$ CS ${}^B P$, as illustrated in Fig. 2.1. The actual transformation from ${}^A P$ to ${}^B P$ is carried out using a matrix–vector multiplication:

$${}^B P = {}^B T_A \cdot {}^A P \quad \Leftrightarrow \quad \begin{bmatrix} B_x \\ B_y \\ B_z \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} A_x \\ A_y \\ A_z \\ 1 \end{bmatrix}$$

Adding a third CS $\{C\}$ with its corresponding transformation ${}^C T_B$ makes the situation interesting. Linear transformations, to which a rigid transformation belongs, can be simply concatenated. However, the concatenation order one has in mind is reverse to the computational flow. One has to think starting from the original vector P at the very right and append each transformation in a sequential order on the left side. Following, the transformation chain to transform a point

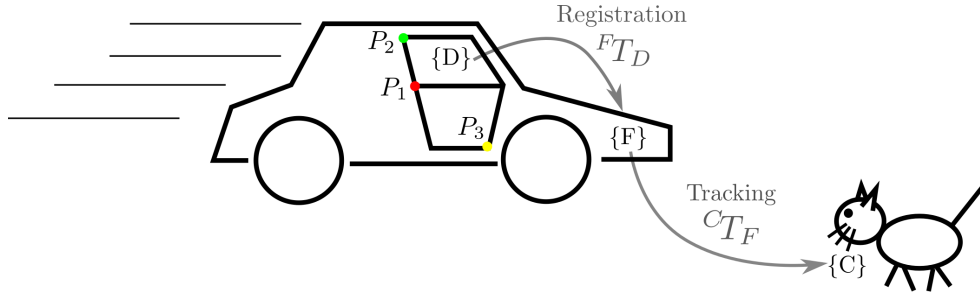


Figure 2.2: Real life example of *tracking* and *registration*.

${}^A P$ into the CS $\{C\}$, ${}^C P$, is given by

$${}^C P = {}^C T_A \cdot {}^A P \quad \text{with} \quad {}^C T_A = {}^C T_B \cdot {}^B T_A.$$

Often, one is in the situation of having a transformation but should actually know its reverse. In respect to Fig. 2.1, the point coordinate ${}^B P$ and the transformation ${}^B T_A$ are assumed to be given. To compute ${}^A P$, the reverse transformation of ${}^B T_A$, correspondingly denoted as ${}^A T_B$, has to be known. The direction of an rigid transformation can be inverted by computing its matrix inverse, as

$${}^A P = {}^A T_B \cdot {}^B P \quad \text{with} \quad {}^A T_B = ({}^B T_A)^{-1}.$$

2.1.1 Determining a Rigid Transformation

So far it was shown how point coordinates are transformed from one CS into another CS, whereat the required rigid transformations were given. But in practice, these transformations are often unknown. Depending on the context of the transformation, different names for the determination process itself are used. Figure 2.2 depicts a simple example, in which a cat with CS $\{C\}$ observes a fast driving car with its CS of the door $\{D\}$ and its front CS $\{F\}$.

Recovering a rather seldom changing transformation is known as *registration* or *calibration*. This is clearly the case for the spatial relationship between the car's door and its front ${}^F T_D$. A commonly applied method to compute such a transformation is based on fitting a set of 3D–3D correspondences [5], denoted in the following as *3dFitting*. The elements of a 3D–3D correspondence are the

coordinates of a particular landmark expressed in both participating CSs. Given the three landmarks P_1 , P_2 and P_3 in Fig. 2.2, ${}^F T_D$ would be computed like

$${}^F T_D = \text{3dFitting}({}^F P_1 \Leftrightarrow {}^D P_1, {}^F P_2 \Leftrightarrow {}^D P_2, {}^F P_3 \Leftrightarrow {}^D P_3).$$

Tracking, on the other side, is the estimation of a rather fast changing spatial relation, also known as *dynamic* transformation. The cat is literally tracking the car by applying a mixture of *motion parallax*, *binocular disparity* and experience to estimate ${}^C T_F$. There exist numerous 3D tracking methods, which are often based on sophisticated technologies. In this thesis, commercial electromagnetic - and stereo-optical devices were applied, which have a tracking frequency up to 30 Hz and a spatial accuracy between 0.1 mm – 1.0 mm.

2.2 Projective Transformation

The projective transformation in the 3D space describes the process of projecting a 3D coordinate to an arbitrary 2D plane, which is equal to a projective camera. In reference to Fig. 2.3, the projective transformation is generally given by

$$u \propto x = K \cdot {}^C T_A \cdot {}^A P$$

and written as components

$$\begin{bmatrix} u_x \\ u_y \\ 1 \end{bmatrix} = \frac{1}{w} \begin{bmatrix} x \\ y \\ w \end{bmatrix} \propto \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \cdot \begin{bmatrix} A_x \\ A_y \\ A_z \\ 1 \end{bmatrix}.$$

In the following, each part of this transformation with its corresponding geometrical meaning depicted in Fig. 2.3 is examined.

Before the actual projection, the 3D point ${}^A P$ has to be transformed from the CS $\{A\}$ into the camera CS $\{C\}$, ${}^C P$. This is done with the rigid transformation ${}^C T_A$. When having a closer look at this transformation, one can notice that the homogeneous line is omitted. This is a handy trick to align the resulting

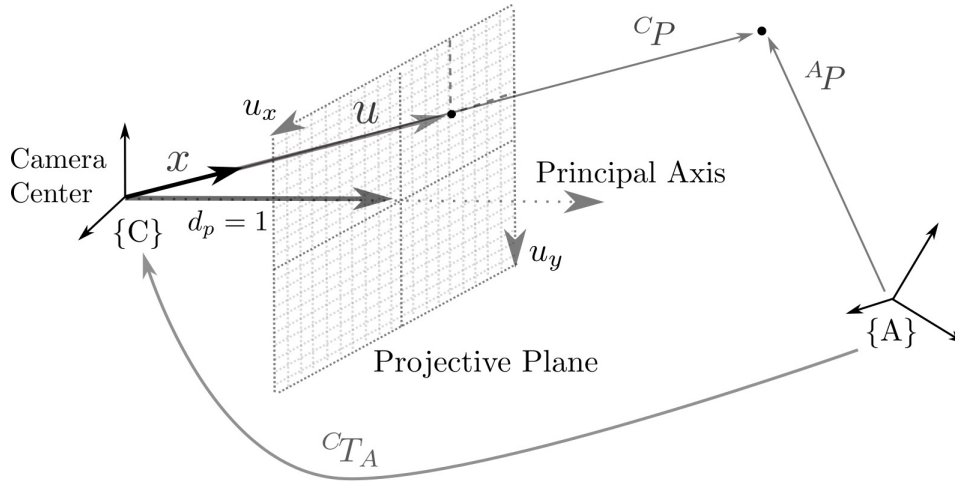


Figure 2.3: The projective camera model, where curved arrows denote rigid transformations and straight arrows are vectors.

vector size of ${}^C P$ with the size of the later used matrix K . In the context of a projective camera, ${}^C T_A$ describes the pose of the camera and is called *extrinsic camera parameters*. Introducing ${}^C P$ to the projective transformation leads to

$$u \propto x = K \cdot {}^C P \quad \text{with} \quad {}^C P = {}^C T_A \cdot A^P.$$

The 3×3 matrix K describes the process of projecting the 3D point ${}^C P$. In terms of a camera, the components of K are known as *intrinsic camera parameters* and represent the optical behavior of the camera. In particular, f_x and f_y are the focal lengths and the principal point coordinate c is given by c_x and c_y . The multiplication of $K \cdot {}^C P$ results in an intermediate vector x . To get the final projection coordinate u , the vector x is scaled based on the distance from the camera center to the projection plane, which is generally chosen to be $d_p = 1$. The two vectors u and x are therefore proportional \propto to each other.

2.2.1 Projective Transformation within a Camera

The general projective transformation represents a *pinhole* camera. For contemporary cameras, however, some remarks have to be done.

The transformation $u \propto K \cdot {}^C P$ is more than a projection itself. It involves a



Figure 2.4: Radial distortion effect bends a straight sliding rail (skatespy.com).

unit change as well, which is incorporated in the intrinsic camera parameters K . In the case of a digital camera, u is a pixel location (u_x, u_y) of the image. But the point coordinate ${}^C P$ is likely to be expressed in a common spatial unit such as millimeter or meter.

Another effect in real cameras are nonlinear distortions introduced by the optical system. Those distortions are not considered by the general projective transformation, even though they are of high importance for the final projection coordinate u . Therefore, an extended projection model is introduced, in which nonlinear corrections are applied directly on u in order to receive a more accurate pixel position u^* .

The most relevant distortion effect is the *radial lens distortion* [4]. This distortion is inevitable and sometimes even reinforced to achieve a so called *fish-eye* effect (Fig. 2.4). It is assumed that the center of the radial lens distortion is equal to the principal point coordinate c and that the distortion is uniformly distributed in all direction over the image. The distortion effect is increasing along with the distance r between a pixel coordinate u and the principal point coordinate c and can be modeled by a polynomial function $L(r)$ of degree n

$$L(r) = 1 + k_1 r + k_2 r^2 + k_3 r^3 + \dots + k_n r^n \quad \text{with} \quad r = \|u - c\| \quad , \text{ where}$$

$k_i (i = 1, \dots, n)$ are radial correction coefficients, which are computed during the camera calibration process. The polynomial degree n is often chosen to be $3 \leq n \leq 5$. Finally, the corrected pixel coordinate u^* is given by

$$u^* = c + L(\|u - c\|) \cdot (u - c).$$

It is important to mention that the radial lens distortion model slightly varies among different implementations. Together with the radial distortion, often also the *tangential distortion* is determined. However, in contemporary cameras this is a minor effect and can be usually neglected.

2.2.2 Camera Calibration

The purpose of camera calibrations is to determine the intrinsic camera parameters. Usually, this is done at an initial step and has to be redone when the optical setup changes. Camera calibration is a subject of ongoing research and methods can be assigned among few general calibration types. The most commonly used approach is the *photogrammetric* calibration, in which a precisely known calibration object is required. Those methods are known to be accurate and robust. However, manufacturing or acquiring such a 3D calibration object is quite an investment. A good trade-off between accuracy and effort is the usage of planar calibration objects. A chessboard or a dot pattern can be precisely created with an off-the-shelf desktop printer. This approach was chosen in this thesis.

Most algorithms applied in the area of photogrammetric camera calibration require so called 2D–3D point correspondences as an input. In reference to Fig. 2.5, the i -th 3D coordinate represented in the calibration object’s CS $\{A\}$, ${}^A P_i$, forms a point correspondence with its projected pixel location u_i in the camera image.

A rudimentary calibration algorithm is the *Direct Linear Transformation* (DLT) [6], inasmuch as it just follows the general projective transformation and finds a proper intrinsic matrix K , but ignores nonlinear distortion effects. In contrast, contemporary calibration algorithms [7, 8] incorporate nonlinear optimization steps to estimate for instance the correction coefficients for radial distortion in addition to the intrinsic camera matrix K .

It is good to know, that solving for the intrinsic camera parameters involves

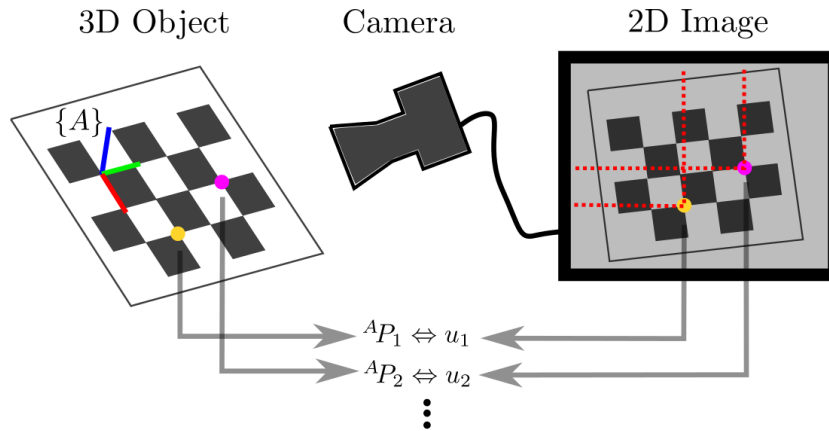


Figure 2.5: 2D–3D point correspondence of a chessboard with its camera picture.

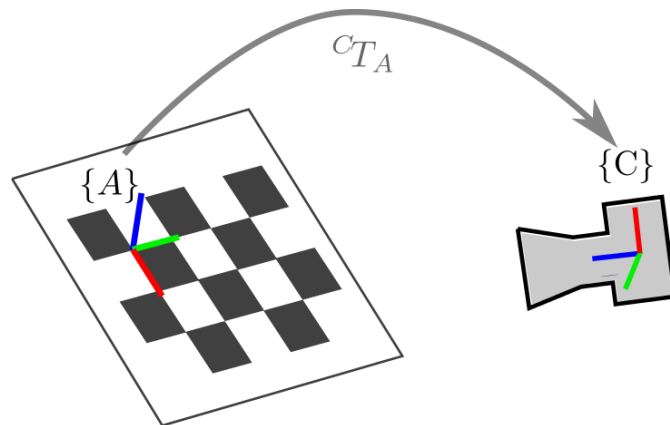


Figure 2.6: Rigid transformation between the CS of the tracked object $\{A\}$ and the CS of the camera $\{C\}$.

also the determination of the extrinsic parameters within this very calibration scene. In particular, the camera pose ${}^C T_A$ in Fig. 2.3 and Fig. 2.6.

2.2.3 Camera Tracking

After the intrinsic camera parameters are known, one can use a camera for estimating the camera pose, such as the transformations ${}^C T_A$ in Fig. 2.6. The methods to compute these kind of problems are known as *perspective- n -point* (PnP) problems. There exist a broad spectrum of sophisticated approaches to solve a PnP problem. A well known type of PnP algorithms is based on 2D–3D

point correspondences as an input. Those correspondences are identical to those used in the photogrammetric camera calibration approach (Fig. 2.5). From that follows, that the geometry of the tracked object has to be known. In this thesis, the PnP implementation of *OpenCV* was used. In particular, the method applies an iterative Levenberg-Marquardt optimization [9] to estimate the camera pose.

There exist also PnP solutions for partially calibrated cameras. For instance, algorithms computing simultaneously the camera pose and the focal length [10]. This is useful if zooming is enabled while tracking.

2.3 Transformation Errors

The quantification of a transformation error is dependent on the applied transformation method and its type of input data, therefore rather specific. In this thesis, however, a recurrent pattern in determining a transformation error occurs. As depicted in Fig. 2.7(a), one assumes that N measured vectors ${}^B\mathbf{V}_i$ and ${}^A\mathbf{V}_i$ ($i = 1, \dots, N$) are given to compute an arbitrary transformation G , so that

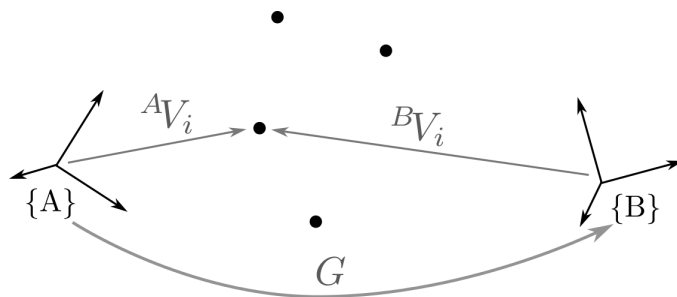
$${}^B\mathbf{V}_i = G \cdot {}^A\mathbf{V}_i,$$

where the corresponding elements ${}^B\mathbf{V}_i$ and ${}^A\mathbf{V}_i$ can be of different units and different dimensions. In general, G does not map ${}^A\mathbf{V}_i$ exactly to its corresponding ${}^B\mathbf{V}_i$. Limited measurement resolution and induced measurement noise results in a measurement error. Processing these data in an algorithm leads inevitably to an error in its resulting transformation G .

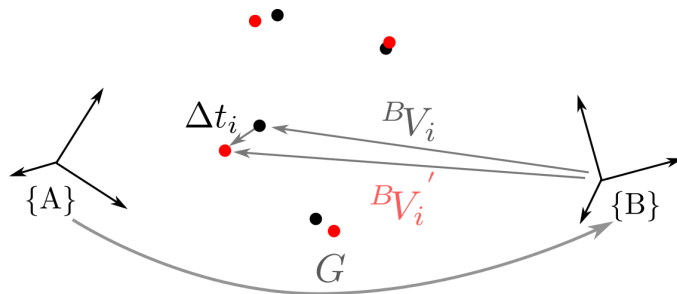
An intuitive method to quantify this error is to transform the measured vectors ${}^A\mathbf{V}_i$ with the computed transformation G , what leads to the projections ${}^B\mathbf{V}_i'$, as illustrated in Fig. 2.7(b). The difference Δt_i between the projected ${}^B\mathbf{V}_i'$ and the measured vector ${}^B\mathbf{V}_i$ is the transformation error for this particular correspondence i , given as

$$\Delta t_i = {}^B\mathbf{V}_i' - {}^B\mathbf{V}_i \quad \text{with} \quad {}^B\mathbf{V}_i' = G \cdot {}^A\mathbf{V}_i.$$

The overall transformation error is based on all N correspondence errors Δt_i . Depending on the application, the standard deviation or the maximum error Δt_{Max} might be of interest. However, the most common error quantification is



(a) Four point correspondences connected by an arbitrary transformation G .



(b) The transformation error Δt_i between the projected vector BV_i' (red dots) and its measured corresponding one BV_i (black dots).

Figure 2.7: Transformation error.

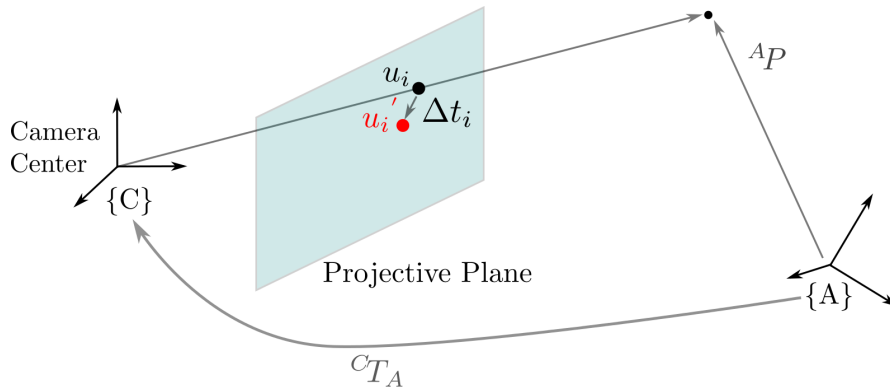


Figure 2.8: The *backprojection error* in the projective plane.

the average of the errors Δt_i , such that

$$E_{\text{Avg}} = \frac{1}{N} \sum_{i=1}^N \|\Delta t_i\| = \frac{1}{N} \sum_{i=1}^N \|{}^B V_i' - {}^B V_i\| = \frac{1}{N} \sum_{i=1}^N \|(G \cdot {}^A V_i) - {}^B V_i\|.$$

It is good to notice that this transformation error quantification can be applied for both mentioned methods, the determination of a rigid transformation (Section 2.1.1) based on 3D–3D correspondences and the camera calibration (Section 2.2.2) based on 2D–3D correspondences. Since the above general illustration (Fig. 2.7(b)) rather shows the error of a rigid transformation, Fig. 2.8 specifically represents the error of a projective transformation, the so called *backprojection error*. The transformation error Δt_i of a correspondence i lies in the projective plane. The corresponding 3D coordinate ${}^A P_i$ is projected with the computed intrinsic parameters K and the particular camera pose ${}^C T_A$ to the projective plane, where this projection is denoted as u_i' . The difference between u_i' and u_i is the transformation error Δt_i , given as

$$\Delta t_i = u_i' - u_i \quad \text{with} \quad u_i' \propto K \cdot {}^C T_A \cdot {}^A P_i.$$

Chapter 3

Publications

The next chapters present the peer-reviewed publications carried out in the context of this thesis. The papers are based on the above described rules of affine and projective transformations in the 3D space and make use of an overlapping set of algorithms and methods. Each publication describes an elegant solution for an interesting clinical problem. However, the medical fields of these papers are versatile, inasmuch as they involve *osteotomy*, *urology* and *cancer research*. All the more it is important to give a rough overview and explain the idea behind the order of the chapters. The first two papers present an innovative clinical 3D tracking approach and an efficient navigation system.

- Development of a computer assisted navigation system for kidney surgeries based on the integration of a commercial electromagnetic tracking device, where the sensors are inserted noninvasively over the urinary passages. *MICCAI 2014*
- Building a cost-effective single camera 3D navigation system for removing farmed tumors in mice brains, in order to study the tumor's blood vessel system. *Salento AVR 2015*

In contrast to the first two publications, the next two papers follow more theoretical aspects, inasmuch as their topics are connected closely with projective geometry and involved challenges.

- Development of a fast point pattern matching algorithm used to create 2D–3D point correspondences. The method is presented in form of a *Cranio-Maxillo-Facial* surgical navigation system running on an *Android* tablet computer. *CLIP MICCAI 2013*
- Applying the projective transformation model to describe an approximated laser deflecting tilting mirror and calibrate this same one. The method is used to calibrate a laser osteotome mounted on a robot. *MICCAI 2015*

Chapter 4

Augmented Reality Assisted Laparoscopic Partial Nephrectomy

The paper was presented at the *MICCAI Main Conference*, September 2014, Boston, USA¹ and won one of the 50 travel awards. It introduces an innovative and progressive surgical navigation system used for kidney operations. The navigation is based on a commercial electromagnetic tracking device. It was shown that the kidney can be tracked noninvasively by inserting an electromagnetic sensor over the urinary passage into the kidney.

Since the proposed navigation system incorporates an electromagnetic tracking system, the extrinsic camera parameters can be determined straight forward. Once the intrinsic camera parameters are known, augmented reality can be performed by aligning the real camera and the virtual camera. In *ex vivo* experiments on pig kidneys, an augmented reality alignment error of $2.1 \text{ mm} \pm 1.2 \text{ mm}$ was achieved.

¹The article is available online at http://link.springer.com/chapter/10.1007/978-3-319-10470-6_45.

Augmented Reality Assisted Laparoscopic Partial Nephrectomy

Adrian Schneider, Simon Pezold, Andreas Sauer, Jan Ebbing,
Stephen Wyler, Rachel Rosenthal, and Philippe C. Cattin

Medical Image Analysis Center, University of Basel, Switzerland

Abstract. Computer assisted navigation is a widely adopted technique in neurosurgery and orthopedics. However, it is rarely used for surgeries on abdominal organs. In this paper, we propose a novel, non-invasive method based on electromagnetic tracking to determine the pose of the kidney. As a clinical use case, we show a complete surgical navigation system for augmented reality assisted laparoscopic partial nephrectomy. Experiments were performed *ex vivo* on pig kidneys and the evaluation showed an excellent augmented reality alignment error of $2.1 \text{ mm} \pm 1.2 \text{ mm}$.

Keywords: Augmented Reality, Electromagnetic Tracking, Navigation.

1 Introduction

Laparoscopic partial nephrectomy (LPN) is considered to be the standard of care for small renal tumors. As opposed to radical nephrectomy where the whole organ is removed, only parts of the kidney get extracted in partial nephrectomy. Although this nephron-sparing surgery is increasingly applied, it is still underused as was found by an investigation among 66 000 patients undergoing radical or partial nephrectomy in the US [2]. In tumors smaller than 4 cm in diameter, LPN has shown to provide equivalent cancer control as compared to radical nephrectomy, but with the advantage of nephron-sparing [8]. The latter results in a higher renal performance and thus in a better quality of life for the patient.

One major challenge in LPN is obtaining optimal surgical margins; that is, removing all cancerous organ parts while keeping as much healthy tissue as possible. Established strategies to decrease the percentage of resections resulting in positive margins make use of intraoperative ultrasound and fresh frozen section analysis [8]. More recently, surgical support by accurate 3D navigation systems started to gain importance.

In this paper, we present a novel, non-invasive method for navigated kidney surgery. In particular, an electromagnetic tracking system is used to determine the kidney pose by localizing a tiny magnetic sensor within a catheter tip that was placed through the urinary passages (urethra, bladder, ureter) at the renal pelvis (Fig. 1, 2). To our knowledge, this transurethral renal access is used for the first time for tracking purposes. It became possible by applying miniaturized electromagnetic sensors that just recently appeared on the market.

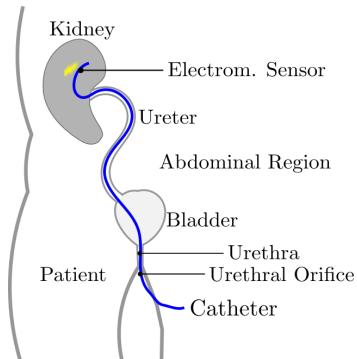


Fig. 1. Transurethral placement of the electromagnetic sensor in the renal pelvis of the right kidney

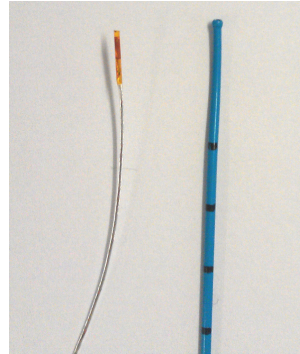


Fig. 2. Electromagnetic sensor, *Model 90, Ascension Corp.* (left). Ureteral catheter, *Angiomed GmbH* (right).

By additionally tracking the laparoscope, we implement a reliable surgical navigation system for LPN. To increase usability and to allow for a seamless integration into the surgical workflow, the classical abstract navigation view is extended with an intuitive augmented reality (AR) visualization technique.

In 2008, Nakamoto et al. [5] proposed a similar LPN guidance method, which uses the *Calypso 4D* to determine the position of implanted wireless magnetic transponders. Additionally, an optical system was used to track the laparoscope. By referencing those two coordinate systems (CS), AR can be performed. An advantage of this approach is the robust tracking of the kidney, as the transponders are implanted and unlikely to shift. The downsides are its extraordinary high price of > \$ 400 000, its large footprint, and the necessity of a second tracker. In our proposed approach, one electromagnetic tracker with wired sensors is sufficient, since the kidney transponder coils and the connecting wire can be packed into a single catheter. This provides us with the opportunity to apply electromagnetic tracking systems that are transportable, much cheaper (< \$ 15 000), and relatively robust to ferromagnetic disturbances. A further advantage of our method is that no transponders have to be implanted into the kidney.

Hughes-Hallett et al. [3] published an excellent review of different AR techniques to perform LPN. Besides the above-mentioned approach, it also describes fundamentally different methods.

2 Materials and Methods

The setup of the proposed navigation system is shown in Fig. 3. In the following, we are going to describe each part of the tracking pipeline in detail, covering deployed materials and algorithms.

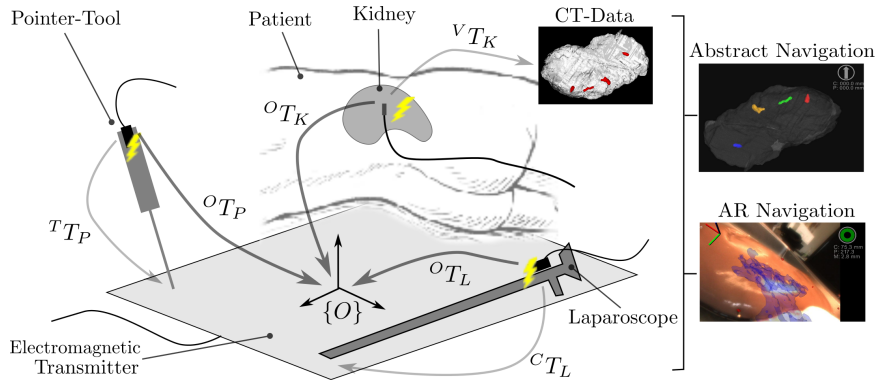


Fig. 3. Navigation system overview. Arrows denote affine transformations.

2.1 Electromagnetic Tracker

As an electromagnetic tracking device, the *trakSTAR 2* together with the *flat* transmitter from *Ascension Technology Corp., Shelburne, USA* is used. The electromagnetic sensors support 6 degrees of freedom. Therefore, the rotation and translation of each sensor can be determined explicitly and are compactly represented by an affine homogeneous 4×4 transformation matrix. A *Model 90* sensor with a diameter of 0.9 mm is built into the catheter. *Model 800* sensors with a diameter of 7.9 mm are attached to the surgical tools. In an OR environment and in the presence of surgical tools (causing ferromagnetic disturbances), a dynamic accuracy of 1.3 mm and a static accuracy of 2.4 mm were determined for a similar setup [10].

As shown in Fig. 3, the electromagnetic tracking system provides the affine transformations O_{TK} (kidney sensor \rightarrow origin), O_{TL} (laparoscope sensor \rightarrow origin) and O_{TP} (pointer-tool sensor \rightarrow origin).

2.2 Laparoscope

We use the 30° *Hopkins2* laparoscope (*Storz GmbH, Tuttlingen, Germany*) together with the *Prosilica GC* laparoscopic camera (*Allied Vision Technologies GmbH, Stadtroda, Germany*). The intrinsic parameters of the laparoscopic optic are determined based on 2D–3D point-correspondences [9] established from a 5×8 chessboard pattern.

As shown in Fig. 3, the transformation C_{TL} links the laparoscopic sensor’s CS and the CS of the camera. The methods to determine C_{TL} are known as *hand-eye calibration*. We implemented the least-squares approach proposed by Park et al. [7].

2.3 Pointer-Tool

A laparoscopic gripper manufactured by *Covidien, Mansfield, USA* is used as a navigated surgical device. The name *pointer-tool* is derived from its application

for determining the 3D position of the tool-tip. The transformation ${}^T T_P$ (Fig. 3) is the translation from the pointer-tool sensor to the pointer-tool tip. We compute it by the method described in [6].

2.4 Kidney Registration

The kidney registration results in the transformation ${}^V T_K$, which maps the CS of the virtual 3D data to the CS of the electromagnetic sensor in the kidney (Fig. 3). In the case of LPN, the 3D data are acquired from a pre-operative diagnostic CT scan.

A well established method to compute the registration is based on 3D–3D point correspondences [1]. If applied to the situation in Fig. 3, one has to select $N \geq 3$ identifiable landmarks ${}^V L_n$ ($n = 1, \dots, N$) from the virtual 3D data. By using the pointer-tool, the corresponding landmarks ${}^T P_n$ are then probed on the real kidney. In order to represent them in the kidney sensor’s CS, denoted as ${}^K P_n$, the following linear transformation is applied:

$${}^K P_n = ({}^O T_K)^{-1} \cdot {}^O T_P \cdot ({}^T T_P)^{-1} \cdot {}^T P_n \quad (n = 1, \dots, N).$$

In a next step, we compute the transformation ${}^V T_K$ by using the two point sets ${}^V L_n$ and ${}^K P_n$ ($n = 1, \dots, N$) as input for the method in [1]. The residual fitting mismatch between the two 3D point sets is the average registration error

$$E_{\text{reg}} = \frac{1}{N} \sum_{n=1}^N (\|{}^V L_n - ({}^V T_K \cdot {}^K P_n)\|).$$

2.5 Navigation

After the kidney registration and device calibration, the remaining task for completing the surgical 3D navigation toolchain is to generate the navigation views.

The Classical Abstract Navigation View can be realized by transforming virtual anatomical 3D data ${}^V D$ and the surgical tool ${}^T P$ into a common CS and by displaying them in a suitable 3D environment. In our implementation, we chose the CS of the virtual 3D data as reference CS. Therefore, ${}^V D$ is already in the correct CS. The surgical tool needs to be transformed from ${}^T P$ to ${}^V P$ by

$${}^V P = {}^V T_K \cdot ({}^O T_K)^{-1} \cdot {}^O T_P \cdot ({}^T T_P)^{-1} \cdot {}^T P.$$

The Augmented Reality Navigation builds upon the classical abstract view. In addition, the position and rotation of the laparoscopic camera, ${}^C T_L$, is mapped to the 3D environment’s virtual camera pose ${}^V C$ by

$${}^V C = {}^V T_K \cdot ({}^O T_K)^{-1} \cdot {}^O T_L \cdot ({}^C T_L)^{-1}.$$

Furthermore, the projective properties of the virtual camera are aligned with those determined during the laparoscopic intrinsic camera calibration. Finally, the undistorted laparoscopic image is put as background into the virtual scene.

3 Experiments and Results

For the following experiments, we used six pig kidneys. Four of them were prepared for the *Sensor Shift* experiment (Sec. 3.1) and two were dedicated to determine the *Overall Navigation Error* (Sec. 3.2). For comparison purposes, a rigid kidney-like mock object was taken into the experiments as well. The *Mock* is constructed using a sponge and holds an artificial tumor made of silicone.

3.1 Sensor Shift

In the presented tracking approach, it is key that the electromagnetic sensor placed in the renal pelvis does not move relative to the kidney while the organ is exposed to external mechanical forces and motion during mobilization. In addition, the error of applying a rigid registration to a soft-tissue structure needs to be evaluated.

In order to observe sensor shifts, we compared an initial 3D–3D registration ${}^V T_{K_0}$ with subsequent registrations ${}^V T_{K_i}$. Between each pair of consecutive registrations, we applied a standardized motion to the kidney, similar to the ones that can be observed during mobilization. The registration differences of the respective 4×4 homogeneous transformation matrices can then be split into a rotational part $\Delta\Theta_i$ and a translational part Δt_i . Since rotations are executed first, Δt_i depends heavily on $\Delta\Theta_i$. At the same time, the distance to the CS origin matters. Therefore, we decided to consider only $\Delta\Theta_i$ as a quantitative measure between the registrations. A good illustration of the effect of $\Delta\Theta_i$ is the resulting point shift ΔP_i in a certain distance d from the sensor. Since the length of an average kidney is about 10 cm, it is reasonable to assume that the sensor can be placed within a range of $d < 2.5$ cm to the region of treatment.

We compute $\Delta\Theta_i$ by the *inner product of unit quaternions* [4] as

$$\Delta\Theta_i = \arccos(|q(r({}^V T_{K_0})) \cdot q(r({}^V T_{K_i}))|),$$

since this rotation metric uses the common unit of radians. Here, $q(\cdot)$ converts a rotation matrix into a 4×1 quaternion and $r(\cdot)$ extracts the 3×3 rotation matrix from a registration. The shift of an arbitrarily chosen point, expressed as Euclidean distance, is calculated as

$$\Delta P_i = \left\| r({}^V T_{K_0}) \cdot \vec{D} - r({}^V T_{K_i}) \cdot \vec{D} \right\| \quad \text{with} \quad \vec{D} = \frac{1}{\sqrt{3}} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot d.$$

The influence of the nonrigid part of the kidney is difficult to isolate. As a reference, the results of the rigid *Mock* can be used. An additional indicator is the registration error E_{reg} . In principle, differences between the pre-operative CT scan and the actual kidney shape lead to an increased registration error.

The following experiment was performed ex vivo with four pig kidneys and the above-described *Mock*. In order to avoid registration errors introduced through landmark correspondence mismatches, five artificial landmarks with precisely

known 3D coordinates from the CT data were used. We chose 20° of rotation and 30 mm of translation relative to the renal hilum (i.e., the entrance of ureter and blood vessels to the kidney) as a reasonable parameterization for simulating the possible mobilization of the kidney during LPN. For every step i , the same motion sequence with the given values was applied to the organ. The repositioning error of the used pointer tool was 0.2 mm.

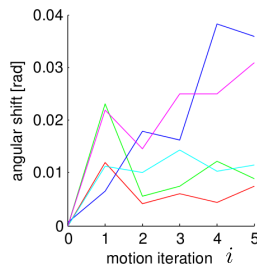


Fig. 4. Registration rotational shift $\Delta\theta$

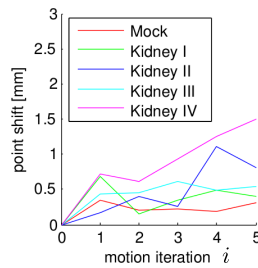


Fig. 5. Point shift ΔP for $d = 25$ mm

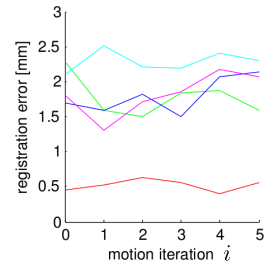


Fig. 6. Registration error E_{reg}

Results: As expected, the *Mock* performed best. Its registration errors (Fig. 6) are about four times smaller than those of the kidneys. The difference can clearly be attributed to the influence of organ deformation or, in general terms, to nonrigidity.

In terms of the sensor shift, the *Mock* and kidneys perform comparably. The rotational shifts (Fig. 4) of *Kidney II* and *IV* stand out, and so do the corresponding point shifts (Fig. 5). In the worst case, an error of 1.5 mm can be expected in a distance of 25 mm from the sensor after applying the motion sequence five times.

In this experiment, the registration error is considerably larger than the point shift. This is the case because the registration is performed over the whole kidney ($d \approx 50\text{--}100$ mm), whereas the point shift is estimated for a distance of $d = 25$ mm from the electromagnetic sensor.

3.2 Navigation Error

In the following experiment, we determined the overall positioning accuracy of our navigation system in both modes: abstract navigation view and AR. The experiment was performed on two ex-vivo pig kidneys and the *Mock*. For each subject, 20 measurements were taken at five known artificial landmarks distributed over the whole kidney. However, in order to be close to the clinical application, the kidney registration was performed using four well identifiable natural landmarks. The repositioning error of the used pointer tool was 0.2 mm,

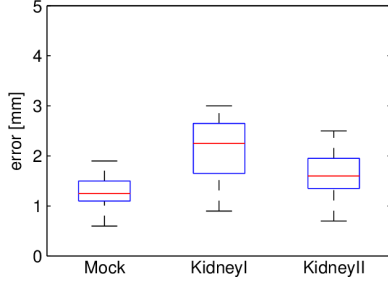


Fig. 7. Navigation error Δ_{Abstract}

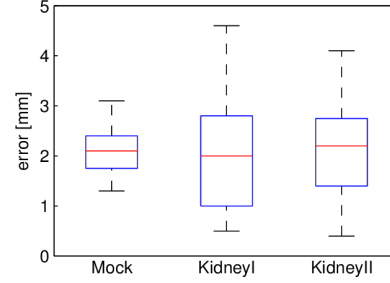


Fig. 8. Navigation error Δ_{AR}

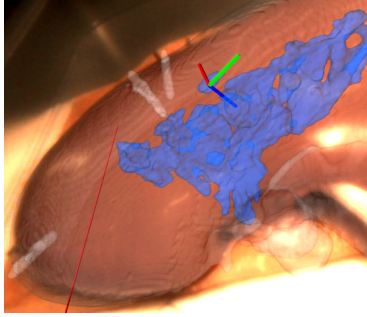


Fig. 9. Real kidney overlaid with virtual renal pelvis (blue)

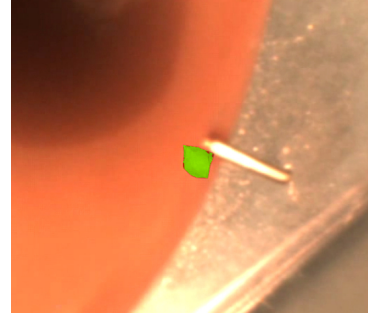


Fig. 10. Virtual marker (green) beside its real corresponding landmark

the error of the intrinsic camera calibration was 0.5 pixel, and the error of the hand-eye calibration of the laparoscopic camera was measured to be 0.4 mm.

In the case of the abstract navigation view, the pointer tool was used to probe a defined landmark a on the kidney and to compare its position ${}^T P_a$ against the known 3D location ${}^V L_a$. The difference is the target registration error $\Delta_{\text{Abstract}_a}$:

$$\Delta_{\text{Abstract}_a} = \| {}^V L_a - ({}^V T_K \cdot ({}^O T_K)^{-1} \cdot {}^O T_P \cdot ({}^T T_P)^{-1} \cdot {}^T P_a) \|.$$

The target registration error of the AR navigation, Δ_{AR_a} , was determined by comparing the visualized location of a landmark on the laparoscopic image stream against its true position (Fig. 10):

$$\Delta_{\text{AR}_a} = \| {}^T P_{R_a} - {}^T P_{V_a} \|.$$

In practice, we used the pointer-tool tip to probe the 3D position of a visualized landmark projected onto the kidney surface ${}^T P_{V_a}$ and to probe the true position ${}^T P_{R_a}$. The distance between the laparoscope and a particular landmark was between 20 mm and 35 mm.

Results: Registration errors are 1.0 mm for the *Mock*, 2.3 mm for *Kidney I*, and 1.8 mm for *Kidney II*. For the kidneys, the determined mean error is 1.9 mm

($std = 0.6$ mm) for abstract navigation (Fig. 7) and 2.1 mm ($std = 1.2$ mm) in AR mode (Fig. 8, 9, 10).

The error of the AR navigation is higher than the one of the abstract navigation, which can be explained by the additional error of the camera transformation. The results also show that the standard deviation of the AR system is much higher. This might be caused by the intrinsic camera parameters. We observed wide-spreading errors of one landmark while changing viewing positions.

4 Conclusion

We showed that our transurethral electromagnetic tracking approach can be applied for LPN, which uses resection margins of 5–7 mm. With an AR error range of 0.9–3.3 mm, our approach performs better than the Calypso based LPN [5] (3–5 mm). However, the experiments also showed that our method is prone to sensor shifts under possible mobilization of the kidney during surgery. For the moment, this issue is tackled by performing re-registrations. In the future, we hope to avoid it by using a dedicated catheter shape.

References

1. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of two 3-d point sets. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (5), 698–700 (1987)
2. Hollenbeck, B.K., Taub, D.A., Miller, D.C., et al.: National utilization trends of partial nephrectomy for renal cell carcinoma: a case of underutilization? *Urology* 254–259 (2006)
3. Hughes-Hallett, A., Mayer, E.K., Marcus, H.J., Cundy, T.P., Pratt, P.J., Darzi, A.W., Vale, J.A.: Augmented reality partial nephrectomy: Examining the current status and future perspectives. *Urology* (2013)
4. Huynh, D.Q.: Metrics for 3d rotations: Comparison and analysis. *Journal of Mathematical Imaging and Vision* 35(2), 155–164 (2009)
5. Nakamoto, M., Ukimura, O., Gill, I., Mahadevan, A., Miki, T., Hashizume, M., Sato, Y.: Realtime organ tracking for endoscopic augmented reality visualization using miniature wireless magnetic tracker. In: Dohi, T., Sakuma, I., Liao, H. (eds.) *MIAR 2008. LNCS*, vol. 5128, pp. 359–366. Springer, Heidelberg (2008)
6. Onprasert, O., S.: A novel method on tool tip calibration for biomedical application. In: *The World Congress on Computer Science and Information Engineering*, pp. 650–653 (2011)
7. Park, F.C., Martin, B.J.: Robot sensor calibration: solving $ax = xb$ on the euclidean group. *IEEE Trans. on Robotics and Automation* 10(5), 717–721 (1994)
8. Weise, E.S., Winfield, H.N.: Laparoscopic partial nephrectomy. *J. Endourol.* 19, 634–642 (2005)
9. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(11), 1330–1334 (2000)
10. Zhou, J., Sebastian, E., Mangona, V., Yan, D.: Real-time catheter tracking for high-dose-rate prostate brachytherapy using an electromagnetic 3d-guidance device: A preliminary performance study. *Medical Physics* 40(2), 021716 (2013)

Chapter 5

Augmented Reality Assisted Brain Tumor Extraction in Mice

In cooperation with the Biomaterials Science Center of the University of Basel, a high accuracy navigation system for the augmented reality assisted tumor extraction in mice brains was developed. The background of this work is the structural analysis of the blood vessel system within tumors. The paper was presented at the *Salento AVR*, September 2015, Lecce, Italy¹.

Instead of using an expensive commercial tracking device, this system works with a single camera. The applied algorithms and image processing routines could be used from the OpenCV project or reused from above publication *Augmented Reality Assisted Laparoscopic Partial Nephrectomy*. Experiments showed a 3D accuracy of $0.48 \text{ mm} \pm 0.25 \text{ mm}$.

¹The paper is available online at http://link.springer.com/chapter/10.1007/978-3-319-22888-4_19.

Augmented Reality Assisted Brain Tumor Extraction in Mice

Adrian Schneider^(*), Peter Thalmann, Simon Pezold, Simone E. Hieber,
and Philippe C. Cattin

Department of Biomedical Engineering, University of Basel, Basel, Switzerland
adrian.schneider@unibas.ch

Abstract. Computer assisted navigation is a widely adopted technique in neurosurgery and orthopedics. In general, the used tracking systems are applicable to multiple situations. However, these general-purpose devices are costly and in case of unusual laboratory applications, a dedicated solution often shows a better performance. In this paper, we propose a cost-effective 3D navigation system for the augmented reality assisted brain tumor extraction in mice, used for cancer research. Off-the-shelf camera 3D reconstruction algorithms are used to individually track a target and a surgical tool. Relative to its costs, the experiments showed an excellent navigation error of $0.48 \text{ mm} \pm 0.25 \text{ mm}$.

Keywords: Augmented reality · 3D Reconstruction · Single camera · Navigation · Micro computed tomography · Cancer

1 Introduction

According to the *World Health Organization*, cancer is a leading cause of death and its prevalence is increasing [12]. Although state-of-the-art oncology is steadily progressing, one of two patients loses the fight against cancer. Current research is tackling the illness at multiple fronts. The main effort targets the design of new drugs, proliferation, the improvement of radiotherapy methods, and the development of sophisticated surgical interventions. To improve the effectiveness of cancer therapies, a better understanding of cancer is highly important. For example, vessel parameters such as diameter and tortuosity are suspected to play a crucial role in the angiogenesis of cancer and therefore also for anti-angiogenic therapies. For the quantification of these vessel parameters, sophisticated 3D imaging techniques are necessary.

In vivo measurements are highly desired, but do not yet provide sufficient resolution. Furthermore, typical absorption-contrast micro-computed tomography (μCT) is not applicable as the contrast for soft tissues is too small for segmentation, and magnetic resonance tomography (MRT) is not yet able to visualize the smallest capillaries (diameter $\sim 5 \mu\text{m}$) due to lack of resolution.

Simone E. Hieber and Philippe C. Cattin shared last authorship.

In the mouse model, one approach to overcome these issues is to use vascular corrosion casting, where the mouse is perfused with Heparin, followed by a polyurethane mixture as described in [6]. The remaining tissue is removed from the polymer specimen with a formic acid solution. Following a standard protocol, synchrotron radiation-based micro-computed tomography (SR μ CT) in absorption-contrast mode [8] or, as shown recently, high-resolution laboratory CT [11] is subsequently used for imaging the specimen. However, this approach is only reliable for tumors at early stages.

A second approach is to use in-line phase-contrast SR μ CT [10], a technique known for much better discrimination of soft tissues compared to standard absorption-contrast μ CT, even without staining. For this technique, however, the spatial resolution of the acquired tomograms highly depends on the specimen size, such that scanning a smaller object enables achieving a higher resolution. The detectors at the synchrotron beamlines typically deliver an image with a fixed size of 2000×2000 pixel. Thus, when scanning the whole mouse brain of about 15 mm size, a spatial resolution of $7.5 \mu\text{m}$ can be reached. This is not enough to visualize the tumor's capillaries. An obvious solution is to measure only the brain part in which the tumor is located. In laboratory mice, the tumors of interest reach a diameter of approximately 2 mm. Extracting them from the brain into specimen of the size of 3 mm, the spatial resolution is increased by a factor of 5 and reaches $1.5 \mu\text{m}$. This enables the visualization of the smallest capillaries, which feature diameters on the level of micrometers.

In this paper, we describe a novel approach to perform such a tumor extraction based on the combination of MRT and computer vision. In particular, a dedicated high resolution MRT device is used to localize the tumor within the mouse brain. Then, a single video camera is used to simultaneously track the brain and a dissection tool. To increase usability and enable a seamless integration into the surgical workflow, an intuitive augmented reality (AR) visualization technique is used. Figure 1 depicts an overview of the whole workflow.

Existing medical 3D navigation systems are highly adaptable and can be used for a multitude of navigation tasks. An accurate and common tracking technology is based on optical stereo frames operating in the infrared spectrum. Such systems achieve submillimeter accuracy in a relatively large measurement volume [2]. Therefore, they are good potential tracking solutions for our application, too. However, these devices are bulky and expensive. By contrast, we are going to present a compact navigation system that achieves comparable results for a fraction of the costs.

2 Materials and Methods

2.1 Specimen Preparation

The tumor samples are gained from mice. At the age of two months, *gliome murine cells* (GL261) are injected into the brain of a mouse. The mouse is sacrificed 12 days after the injection. At this stage, the tumor has reached a

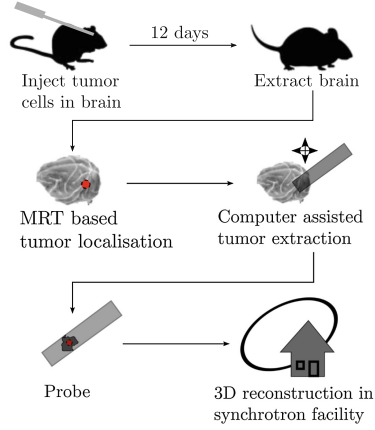


Fig. 1. Workflow from farming the tumor up to 3D reconstruction. This paper is mainly about the tumor localization and extraction. The red dot indicates the tumor (Color figure online).

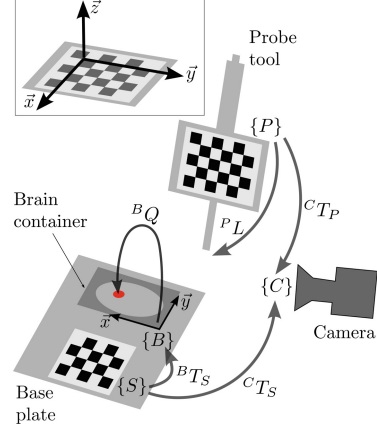


Fig. 2. Setup of the navigation system. Curved arrows denote affine transformations and the red dot indicates the tumor. In the upper left, a chessboard and its spanned coordinate system are shown (Color figure online).

diameter of approximately 2 mm. The target region – cancerous cells – is then extracted from the brain with our newly developed navigation system.

2.2 Navigation System

The setup of the proposed navigation system is shown in Fig. 2. The base plate carries a container for the whole mouse brain. To trepan the tumor from the mouse brain, the probe tool is used. A third part is a camera, which establishes an accurate spatial relationship between the base plate and the probe tool.

In the following, we are going to describe each part of the navigation pipeline in detail, covering deployed materials and algorithms. To familiarize the reader with our nomenclature, we would like to explain briefly the term *affine transformation*. An affine transformation describes a rotational and translational relationship between two coordinate systems (CS) and can be compactly represented by a 4×4 matrix in the 3D case. The notation ${}^X T_Y$ denotes an affine transformation matrix that transforms a point expressed in the $\{Y\}$ CS ${}^Y P$ into a position relative to the $\{X\}$ CS ${}^X P$. The actual transformation from ${}^Y P$ to ${}^X P$ is carried out using matrix–vector multiplication:

$${}^X P = {}^X T_Y \cdot {}^Y P.$$

2.3 Video Camera

The *MQ013CG-E2* digital camera from *XIMEA GmbH* (Münster, Germany) was used, which delivers images of the size of 1280×1024 pixel. As a lens, the *002915* from *Tamron* (Saitama, Japan) with a focal length of 8 mm is used. Using the notation introduced in Fig. 2, the camera image is processed in order to recover the two dynamic affine transformations ${}^C T_P$ and ${}^C T_S$, where ${}^C T_P$ is the transformation from the probe tool to the camera and ${}^C T_S$ is the transformation from the base plate to the camera.

To understand the methods of estimating a 3D transformation from a single camera image, it is necessary to have a closer look at the pinhole model, which is used to describe a projective camera. The projective transformations for ${}^C T_P$ and ${}^C T_S$ are defined as

$${}^P u \propto K \cdot {}^C T_P \cdot {}^P X \quad \text{and} \quad {}^S u \propto K \cdot {}^C T_S \cdot {}^S X,$$

where K are the intrinsic camera parameters, ${}^P u$ is the 2D pixel location of the projected 3D point ${}^P X$, and ${}^S u$ is the respective projection of ${}^S X$. Those pairs are called 2D–3D point correspondences and can be robustly created and recovered by using simple and detectable patterns. The \propto sign indicates proportionality, because the projected pixel coordinates are usually normalized to $[u, v, 1]^T$.

Initially, K has to be determined. These parameters describe the projective behavior of the camera and are determined only once. A common camera calibration method is based on using multiple 2D–3D point correspondences created from planar calibration patterns [13], e. g. chessboard patterns. In addition to K , this calibration algorithm also estimates non-linear lens distortion parameters, such as *tangential distortion* and *radial distortion*.

Tracking an object with a calibrated camera is very similar to the calibration process itself. The 2D–3D point correspondences are created by extracting the 2D coordinates from the camera image and connecting them with the chosen 3D coordinates of the pattern. Based on these correspondences, an iterative method using Levenberg-Marquardt optimization is applied to estimate the extrinsic camera parameters, in particular the affine transformations ${}^C T_P$ and ${}^C T_S$.

Camera Calibration Error and Tracking Error. The camera calibration [13] and the object tracking algorithm [3] both have residual errors. Several error measures can be applied to quantify the quality of the found solution [4]. A commonly used measure is the *backprojection error* E_B , which represents the average pixel error in the image: The computed solution is used to project the i -th 3D point ${}^Y X_i$ of the 2D–3D point correspondences to the image plane. These projections u'_i are then compared with the recorded 2D pixel coordinates u_i taken from point correspondences. If there are N 2D–3D correspondences, then

$$E_B = \frac{1}{N} \sum_{i=1}^N \|u_i - u'_i\| \quad \text{with} \quad u'_i \propto K \cdot {}^C T_Y \cdot {}^Y X_i.$$

In general, the calibration error and tracking error depend heavily on the applied equipment and the scene itself. For further detail, we refer the reader to [7], where a comprehensive error estimation for single camera tracking systems is described.

2.4 Probe Tool

The probe tool is used to trepan the tumor from the mouse brain. The user brings the tool into the right pose by following the guidance instructions of our navigation system (Sect. 2.6).

In reference to Fig. 2, the missing tool tip position ${}^P L$, expressed in the $\{P\}$ CS, can be found by the algorithm described in [9]. However, since the exact orientation of the tool's chessboard CS is known, a better method to determine ${}^P L$ is simply measuring the translation, e. g. by using a sliding calliper.

2.5 Base Plate and Brain Container

The complete brain of the mouse is surgically removed and put into the brain container. The location of the tumor Q is measured using a dedicated high resolution MRT device (*PharmaScan 47/16*, *Bruker BioSpin*, Ettlingen, Germany). Using the container shape, the MRT origin can be aligned with the CS $\{B\}$ easily, at the same time providing the tumor location ${}^B Q$ in $\{B\}$. To enhance the visibility of the container itself on the MRT image, we place it in a water bath during the scan.

The base plate has a slot to plug in the brain container. Since both geometries and the orientation of the base plate's chessboard are known, the rigid transformation ${}^B T_S$ can be measured very precisely. In a more complex situation, e. g. if the construction data of the base plate is not available, one could apply a 3D–3D registration as described in [1].

2.6 Augmented Reality Assisted Navigation

Once all affine transformations of the navigation system are known, the remaining task is to generate the navigation view. Since we use an AR user interface, it makes sense to choose the camera as the common CS. In reference to Fig. 2, the probe tool tip ${}^P L$ and the tumor location ${}^B Q$ are transformed to $\{C\}$ by

$${}^C L = {}^C T_P \cdot {}^P L \quad \text{and} \quad {}^C Q = {}^C T_S \cdot {}^B T_S^{-1} \cdot {}^B Q.$$

The tumor location ${}^C Q$ is visualized on the camera image using the camera calibration. Furthermore, the user gets updated with the current distance from the tool tip to the tumor, $d_T = \|{}^C L - {}^C Q\|$, by means of a dynamically scaled vertical bar that we overlay on the camera image. Figures 3 and 4 give an impression of the AR assisted navigation.

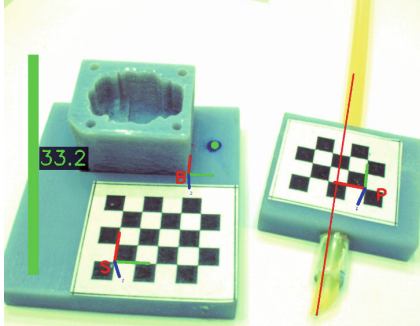


Fig. 3. AR visualization I: The base plate with the brain container is on the left; a vertical green bar displays the distance d_T , the augmented green dot indicates a simulated tumor location. The probe tool is on the right (Color figure online).

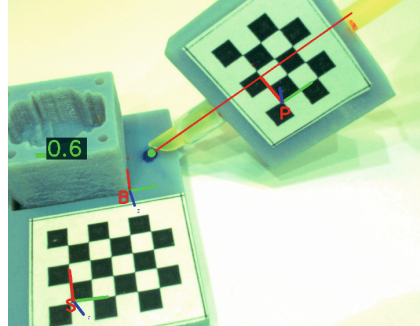


Fig. 4. AR visualization II: The chessboard coordinate system frames $\{B\}$, $\{S\}$, $\{P\}$, and the probe tool's orientation are overlaid. The distance d_T becomes smaller with the tool tip approaching the tumor location.

3 Experiments and Results

In this section, we determine the precision and accuracy of our method under several aspects. The used hardware is as presented in Sect. 2. For the following experiments, a camera calibration with a backprojection error $E_B = 0.2$ pixel was used. The calibration was performed based on 20 different chessboard poses and setting the tangential distortion to zero.

3.1 Transformation Robustness

In this experiment, our goal is to estimate the robustness of the proposed tracking method. In particular, two chessboards with a known, rigid transformation in between them are simultaneously tracked (Fig. 5). The estimated transformation between both chessboards can be quantitatively compared against the true one. This is done from several different camera poses.

The first chessboard has 4×5 fields and spans the CS $\{X\}$. The second board has 4×3 fields and spans the CS $\{Y\}$. Both have square fields, each with a length of 3 mm. The rigid transformation ${}^X T_Y$ between the two chessboard CSs was chosen so that there is no rotation, but only a translation of 30.0 mm. The transformations from each chessboard to the camera CS $\{C\}$, ${}^C T_X$ and ${}^C T_Y$, are recovered from the camera image. Therefore

$${}^X T_Y = {}^C T_X^{-1} \cdot {}^C T_Y.$$

In the following, this transformation is estimated from $N_q = 25$ different camera poses. Each transformation ${}^X T_{Y_q}$ is compared with the true transformation ${}^X T_{Y_t}$.

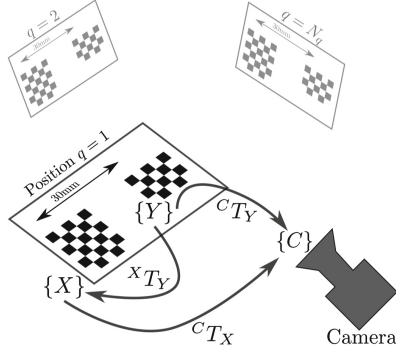


Fig. 5. Transformation robustness experiment setup.

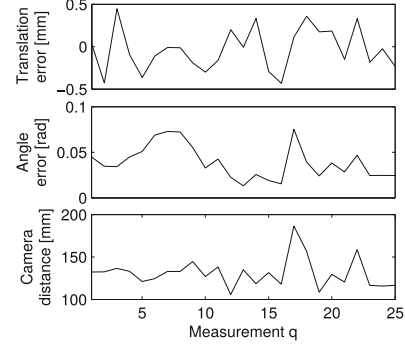


Fig. 6. Transformation robustness experiment results.

The differences can be split into a rotational part $\Delta\Theta_q$ and a translational part Δt_q . We compute $\Delta\Theta_q$ by the *inner product of unit quaternions* [5] and Δt_q by the difference of the CS offsets, as

$$\Delta\Theta_q = \arccos(|q(r({}^X T_{Y_t})) \cdot q(r({}^X T_{Y_q}))|), \quad \Delta t_q = \|t({}^X T_{Y_t})\| - \|t({}^X T_{Y_q})\|.$$

The function $q(\cdot)$ converts a rotation matrix into a 4×1 quaternion, $r(\cdot)$ extracts the 3×3 rotation matrix from the transformation and $t(\cdot)$ returns the 3×1 translation component.

The distance d_q between the camera and the tracked pattern has an influence on the accuracy [7]. In our case, d_q is the average distance from both chessboards to the camera and is computed like

$$d_q = \frac{1}{2} (\|t({}^C T_{X_q})\| + \|t({}^C T_{Y_q})\|).$$

Results The results can be seen in Fig. 6. The mean translation error is 0.20 mm and has a standard deviation of 0.25 mm. The average rotation error is approximately 0.035 rad (2°). The experiments were performed within a camera distance range of 100 mm to 200 mm. However, the supposed correlation between the camera distance and the corresponding errors is not observed. The model for a theoretical error estimation of a single camera navigation system [7] does not consider an equivalent situation, thus a comparison is difficult. But it shows that the determined rotation errors and translation errors are plausible.

3.2 Navigation System Accuracy

In the above experiment, we determined the precision and robustness of tracking two chessboards. Regarding the proposed system (Fig. 2), this only corresponds

to the transformations ${}^C T_S$ and ${}^C T_P$. In the following, the accuracy of the whole navigation system is assessed.

The transformations ${}^B T_S$ and ${}^P L$ are known from the construction data or are determined by physical distance measurements. ${}^C T_S$ and ${}^C T_P$ are estimated based on the camera image. We determine the overall navigation accuracy by defining a known point in the CS $\{B\}$, denoted as ${}^B Q$, and by pointing the tip of the probe tool at it. This correlates with the actual task of a navigated tumor extraction. Figure 4 is a snapshot of this experiment. The green dot is the point ${}^B Q$. By using the complete affine transformation chain of our navigation system, the probe tool tip's location ${}^P L$ is then transformed to ${}^B L$, the tip's position expressed in the CS $\{B\}$. The Euclidean distance of the computed tip position ${}^B L$ from its reference location ${}^B Q$ can be considered as the navigation error E_N . It is given by

$$E_N = \|{}^B L - {}^B Q\| \quad \text{with} \quad {}^B L = {}^B T_S \cdot {}^C T_S^{-1} \cdot {}^C T_P \cdot {}^P L.$$

The marker for the base plate is a 4×5 chessboard that spans the CS denoted as $\{S\}$. The probe tool is tracked via a 4×3 chessboard that spans the CS denoted as $\{P\}$. Both chessboards have square fields, each with a length of 3 mm.

Results The experiment was repeated 25 times within a camera distance range of 100 mm to 200 mm. The average navigation error E_N is 0.48 mm with a standard deviation of 0.25 mm. This corresponds to the accuracy observed with commercial medical navigation systems using a pointing tool [2]. However, our measurement volume is much smaller but adequate for the problem at hand.

3.3 Brain Extraction

The navigated tumor extraction (Fig. 7) was tested on two mice and two mock objects, where the tumor was built from modeling clay surrounded by soft paraffin simulating the brain tissue. An experiment is qualitatively rated following to the amount of the tumor which was trepanned from the brain. We introduce the labels *Complete*, *Partial* and *Off Target* tumor extraction.

Results. The two experiments with the mock objects resulted in *complete* tumor extractions. Both experiments with mice led to *partial* tumor samples (Fig. 8). This difference could be explained with the observation that the real brains slightly moved and got deformed at the moment the probe tool punctured the brain surface, whereas this was not the case with the mock objects. In general, it is important to notice that our navigation system is on the edge of meeting the required accuracy for a *complete* tumor extraction. Sampling a target with a 2 mm diameter using a sampling tool that has a diameter of 3 mm requires an accuracy ≤ 0.5 mm. The experiment above showed that we reach $0.48 \text{ mm} \pm 0.25 \text{ mm}$.

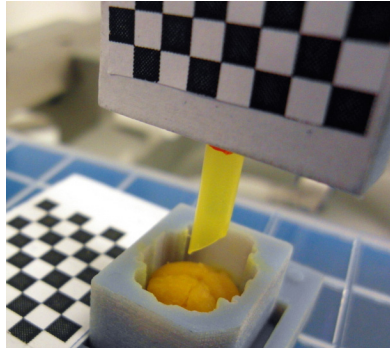


Fig. 7. Tumor extraction from a real mouse brain using the proposed navigation system. The experiment took place at the Animal Imaging Center of ETH Zurich.

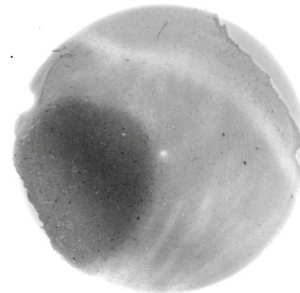


Fig. 8. Reconstructed slice with 3 mm diameter using inline phase contrast SR μ CT (pixel size 1.9 μ m). The dark area represents the partially extracted tumor.

4 Conclusion

In this paper we proposed an AR navigation system applicable for guided brain tumor extraction in mice. The achieved accuracy of 0.48 mm combined with its relatively low cost opens up possibilities for using the system in other fields. However, for the presented task it would be useful to enhance the accuracy. The current system is operated with one camera and the tracked patterns are planar. A better accuracy can be achieved by using two cameras, which together form a stereo frame. The drawback with the latter approach is the reduced field of view, because both cameras need to spot the patterns. Another solution could be the usage of non-planar patterns. In theory, this enhances the 3D reconstruction accuracy [7]. The disadvantage is the complexity of constructing such a pattern. Our current research is focusing on this challenge.

Acknowledgments. We would like to thank Bert Müller for his initiative and helpful discussions, Therese Bormann for her help in the container design and the 3D printing at the FHNW Muttentz, Mathias Griessen and Alex Seiler for their support in the software implementation, Sandra Burgi for preparing the murine brain, and Marco Dominietto for his useful feedback and the MRT measurements at the Animal Imaging Center of ETH Zurich.

References

1. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of two 3-d point sets. *IEEE Trans. Pattern Anal. Mach. Intell.* **5**, 698–700 (1987)
2. Broers, H., Jansing, N.: How precise is navigation for minimally invasive surgery? *Int. Orthop.* **31**(1), 39–42 (2007)

3. Gao, X.S., Hou, X.R., Tang, J., Cheng, H.F.: Complete solution classification for the perspective-three-point problem. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(8), 930–943 (2003)
4. Hartley, R., Zisserman, A.: *Multiple view geometry in computer vision*. Cambridge University Press, New York (2010)
5. Huynh, D.Q.: Metrics for 3D rotations: comparison and analysis. *J. Math. Imag. Vis.* **35**(2), 155–164 (2009)
6. Krucker, T., Lang, A., Meyer, E.P.: New polyurethane-based material for vascular corrosion casting with improved physical and imaging characteristics. *Microsc. Res. Tech.* **69**(2), 138–147 (2006). <http://dx.doi.org/10.1002/jemt.20263>
7. Luhmann, T.: Precision potential of photogrammetric 6DOF pose estimation with a single camera. *ISPRS J. Photogrammetry Remote Sens.* **64**(3), 275–284 (2009)
8. Müller, B., Lang, S., Dominietto, M., Rudin, M., Schulz, G., Deyhle, H., Germann, M., Pfeiffer, F., David, C., Weitkamp, T.: High-resolution tomographic imaging of microvessels. *Proc. SPIE - Int. Soc. Opt. Eng.* **7078**, 70780B-1–70780B-10 (2008). <http://dx.doi.org/10.1117/12.794157>
9. Onprasert, O., Suthakorn, J.: A novel method on tool tip calibration for biomedical application. In: *Proceedings of the World Congress on Computer Science and Information Engineering*, pp. 650–653 (2011)
10. Paganin, D., Mayo, S.C., Gureyev, T.E., Miller, P.R., Wilkins, S.W.: Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object. *J. Microsc.* **206**(1), 33–40 (2002). <http://dx.doi.org/10.1046/j.1365-2818.2002.01010.x>
11. Thalmann, P., Hieber, S.E., Schulz, G., Deyhle, H., Khimchenko, A., Kurtcuoglu, V., Olgac, U., Marmaras, A., Kuo, W., Meyer, E.P., Beckmann, F., Herzen, J., Ehrbar, S., Müller, B.: Three-dimensional registration of synchrotron radiation-based micro-computed tomography images with advanced laboratory micro-computed tomography data from murine kidney casts. *Am. J. Physiol. Heart Circ. Physiol.* **9212**, 92120Y-1–92120Y-9 (2014). <http://dx.doi.org/10.1117/12.2060809>
12. WHO: Research for patient safety (2008). http://www.who.int/patientsafety/information_centre/documents/ps_research_brochure.en.pdf
13. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000)

Chapter 6

Landmark-Based Surgical Navigation

This paper was published at the workshop *Clinical Image-Based Procedures, Translational Research in Medical Imaging - MICCAI*, September 2013, Nagoya, Japan¹ and won the second best paper award. It introduces a computationally fast, semi-rigid point pattern matching algorithm to create 2D–3D point correspondences. This enabled the development of a cost-effective surgical navigation system running in real-time on a tablet computer.

In contrast to above publication, the 2D–3D correspondences cannot be extracted directly from a defined marker structure like the used chessboards. The correspondences are computationally determined by the proposed recursive 2D–2D point pattern matching algorithm. An augmented reality alignment error of $0.8 \text{ mm} \pm 1.0 \text{ mm}$ could be measured.

¹The paper is available online at http://link.springer.com/chapter/10.1007/978-3-319-05666-1_8.

Landmark-Based Surgical Navigation

Adrian Schneider^(*), Christian Baumberger, Mathias Griessen, Simon Pezold,
Jörg Beinemann, Philipp Jürgens, and Philippe C. Cattin

Medical Image Analysis Center, University of Basel, Basel, Switzerland
`adrian.schneider@unibas.ch`

Abstract. Navigational support is a widely adopted technique in surgery that has become a part of the clinical routine. This navigation support either comes in the form of an abstract display that shows for example the distance and direction to the target position or in the form of augmented reality where segmented anatomical structures of interest are overlaid onto a visual image sequence in real-time.

In this paper we propose a cost-effective real-time augmented reality approach using an off-the-shelf tablet PC in combination with a novel 2D/3D point correspondence mapping technique. The proposed point pattern matching algorithm is tailored towards moderate projective distortions and suitable for computational low-power devices. Experiments and comparisons were done on synthetic images and accuracy was measured on real scenes. The excellent performance is demonstrated by an Android 3D guidance application for a relevant medical intervention.

Keywords: Augmented reality · Point pattern matching · Navigation

1 Introduction

According to the *World Health Organization*, there are more than one million surgery-related deaths world-wide per year [12]. One of the reasons is that performing surgical interventions poses very high demands on the spatial sense of the surgeons. As has been documented by several studies [6, 9] intraoperative 3D navigation greatly supports the surgeon in complex interventions and significantly reduces the risk for the patients. However, the currently available systems for 3D navigation are bulky, complex to operate and expensive.

In this paper, we focus on the development of an image-guided 3D navigation system that can be quickly brought in place, requires minimal training and is affordable. We show that an off-the-shelf tablet computer can bridge this gap although certain challenges appear. A common tablet has only one built-in CMOS camera. Therefore, our 3D navigation system is restricted to operate in single view mode. Furthermore, a tablet is a computational low-power environment and thus computationally expensive image processing routines are inapplicable for real-time applications.

^(*) Christian and Adrian contributed equally to this work.

To increase usability and allow a seamless integration into the surgical workflow, an intuitive augmented reality (AR) visualization technique is used. In the proposed navigation system, virtual anatomical objects are rendered and superimposed onto the image stream from the camera and shown on the tablet. This requires the determination of the exact 3D position and orientation of the tablet’s camera with respect to the patient and handing over these parameters to the virtual camera of the renderer. This brings us to the essential challenge and core contribution of this work, namely the development of an accurate but computationally cheap camera pose estimation.

In applications with only one tracking camera, a reliable camera pose estimation method uses 3D/2D point correspondences, where the 2D coordinates represent the pixel locations of detected landmarks and the corresponding 3D coordinates are known with respect to an MR or CT data set of the patient. Having several of these 3D/2D point correspondences, one can compute the camera pose transformation relative to the 3D coordinate system of the landmarks and therefore render anatomical structures from the right perspective.

Extracting these landmarks from an image has to be efficient and reliable. Fast texture-based methods [4] are prone to perspective distortion [10] and likely to fail on smooth surfaces such as teeth or bone. Therefore we decided to use small, uniformly colored stickers as markers, which can be segmented efficiently by color channel thresholding. In addition, we treat each new image completely separately from the previous frames, i.e. a tracking-by-detection approach was chosen. The advantage of tracking-by-detection is that fast movements do not distract the tracking and there is no accumulation of error or drift over time. The complete segmentation of the landmarks is performed on the GPU of the tablet and results in a binary image. The centers of the circles are then determined by using the blob detector of OpenCV.

Because the proposed landmarks are not accompanied by any unique identifier, such as distinctive colors or unique texture descriptors, there is no direct way to assign the corresponding 3D coordinates to each detected 2D position landmark. What might look like a simple operation at first, turns out to be a challenging task known as *point pattern matching* (PPM) [3]. In our context, a PPM method is required which matches two point patterns (the detected 2D landmarks to the set of known 3D coordinates) related by a projective transformation. Such algorithms exist, but they are computationally too expensive [1] or are restricted to coplanar point patterns [2, 3, 11]. Optimized SLAM methods [5] might be an option, but their structure-from-motion approach does not match our tracking-by-detection requirement.

A computationally feasible approach is to simplify the problem and match the segmented 2D point set v_S with an initially generated reference 2D point pattern v_R , which is the virtual projection of the 3D landmarks from a *reference direction* (Fig. 1). With that step, projective characteristics are banished from the model but return in the form of projective distortions.

A robust method to handle such distortions is to approximate a complex geometric behavior by multiple local low-order transformations [8]. On the basis of

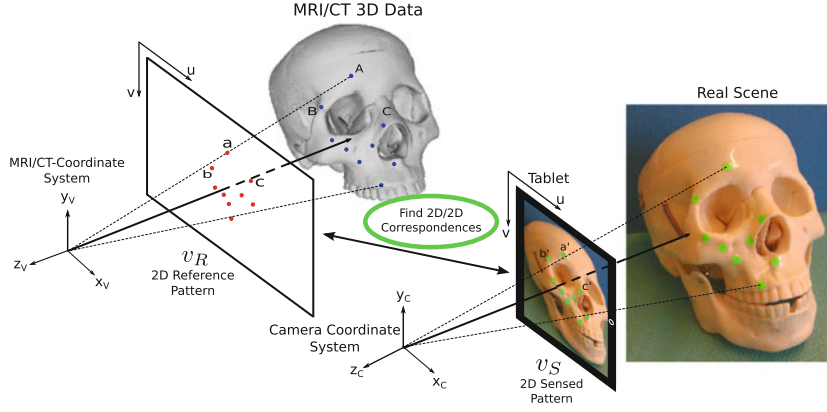


Fig. 1. Simplification of the 3D/2D to a 2D/2D correspondence problem. The blue points denote the 3D object points, the red points the 2D reference pattern projected from the reference direction, and the green points on the tablet the 2D sensed pattern.

this idea, we developed a computationally lightweight algorithm used for rather small point patterns. The newly proposed method is called *Point Recursive Matching* (PRM).

2 Method

The PRM algorithm is based on a recursive structure with an early stopping criterion. Before the algorithm can be applied to find correspondences, point set descriptors Q_R of v_R and Q_S of v_S needs to be computed. The descriptors themselves serve as a look-up table during the recursion and therefore do not need to be recomputed after each iteration. The algorithm is able to handle multiple occlusions since it tries to establish correspondences by locally finding corresponding points and does not try to match the complete pattern at once. By changing the number of nearest point neighbors N_{nb} , among which the solution finding process continues, one can adjust how *locally* the algorithm operates.

Point Set Descriptor. The first step of building the point set descriptor involves the definition of an arbitrarily chosen *base direction*. Subsequently, a descriptor for every point in the point set is computed as follows: compute the connecting vector to every other point in the point set, and compute the angle of this vector with respect to the base direction. The angle of a point to itself is defined as -1 . Figure 2 shows two example point descriptors for point a and b . Finally, all the point descriptors are appended in row-direction into a matrix, which forms the complete point set descriptor Q (Fig. 3). Once the two point set descriptors Q_R of v_R and Q_S of v_S are computed, point correspondences can be established with the recursive approach described in the next paragraph.

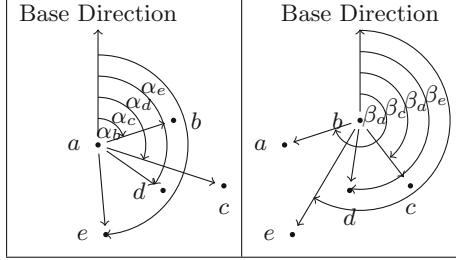


Fig. 2. Two example point descriptors.

Angles					
	a	b	c	d	e
a	-1	α_b	α_c	α_d	α_e
b	β_a	-1	β_c	β_d	β_e
c	γ_a	γ_b	-1	γ_d	γ_e
d	δ_a	δ_b	δ_c	-1	δ_e
e	ϵ_a	ϵ_b	ϵ_c	ϵ_d	-1

Fig. 3. Point set descriptor Q in matrix form.

Matching Algorithm. For practical reasons, we demonstrate the matching algorithm on two example point sets v_R and v_S (Fig. 4, 5, 6 and 7). The set v_S results from the segmentation and therefore has an arbitrary point order. The algorithm starts with the first point of each point set and assumes that these two points correspond to each other. In the example point sets, this corresponds to $a \leftrightarrow a'$ and is shown in Fig. 4.

Next, the algorithm chooses the next point pair among the N_{nb} nearest neighbors of the last assigned correspondence, a respective a' . In this case, the algorithm assigns $b \leftrightarrow b'$ to each other, and reads the angles ϕ_{ab} and $\phi_{a'b'}$ from the pre-computed point set descriptors Q_R and Q_S . In particular these correspond to the entries α_b (Fig. 3) in both Q_R and Q_S . This step is shown in Fig. 5.

The difference of these two angles is kept as the offset between the two point sets $\phi_{offset} = \phi_{ab} - \phi_{a'b'}$. This offset is subtracted in the following from every angle in the sensed point set descriptor Q_S , in order to make it rotation invariant.

The third corresponding point pair is chosen among the not yet assigned N_{nb} nearest neighbor points of b and b' , for example the point pair $c \leftrightarrow c'$ (Fig. 6). With the obtained ϕ_{offset} from the former two correspondences, the algorithm can efficiently validate further point correspondences by comparing the angles between each already assigned point correspondence with the candidate point pair in both descriptors Q_R and Q_S . The assignment is rejected if a predefined angular difference λ_{th} is exceeded.

With regard to the example, the angles ϕ_{ac} and $\phi_{a'c'}$, and ϕ_{bc} and $\phi_{b'c'}$ are compared. If one of the differences $\phi_{diff_\alpha} = |(\phi_{a'c'} - \phi_{offset}) - \phi_{ac}|$ or $\phi_{diff_\beta} = |(\phi_{b'c'} - \phi_{offset}) - \phi_{bc}|$ is larger than the threshold λ_{th} , the assignment $c \leftrightarrow c'$ is rejected and the next correspondence pair $d \leftrightarrow c'$ is tested. Otherwise the algorithm tries to establish further point correspondences by following the same routine.

The PRM algorithm in this way can be formulated very compactly in a recursive manner. With every recursive step, one new point pair is tested and potentially rejected. The current best solution is the one which could determine most point correspondences. If several solutions have an equal number of correspondences, the one with the lowest accumulated angular difference is chosen.

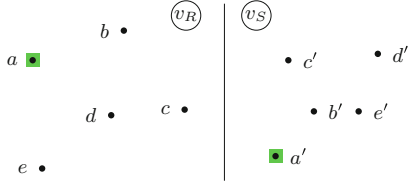


Fig. 4. The first correspondence assumption is marked in green. (Color figure online)

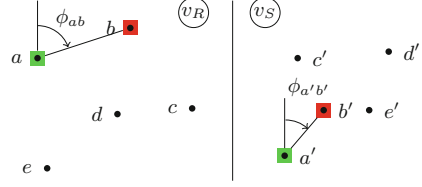


Fig. 5. The second correspondence assumption is marked in red. (Color figure online)

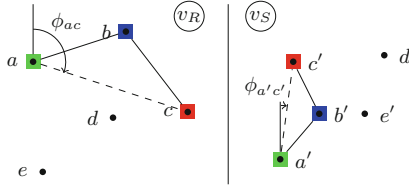


Fig. 6. The next correspondence is added and is marked in red. (Color figure online)

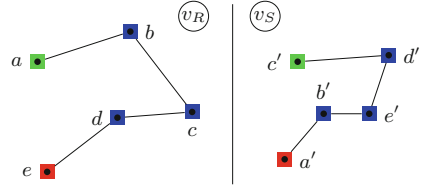


Fig. 7. Final solution to the correspondence problem.

With the angular threshold $\lambda_{th} \geq 360^\circ$ and number of nearest neighbors $N_{nb} = \text{number of all points}$, the PRM algorithm corresponds to an exhaustive brute-force search of the best possible combination. The parameters λ_{th} and N_{nb} help pruning the exhaustive search tree significantly by removing branches with an unreasonably high error early in the recursive search and by limiting the search space. The final assigned solution for the given example can be seen in Fig. 7.

The bottom-up approach of PRM handles occluded points in a natural way. Since the algorithm tries to find corresponding point pairs locally, it will simply skip any occluded points and try to assign the next point in the list. The comparison of angles is facilitated with the use of the point set descriptors Q_R and Q_S as look-up tables. The only arithmetic computations during the recursion process are the subtraction of the angle offset ϕ_{offset} and accumulation of the angular error, the rest are just comparisons.

3 Experiments and Results

In all the experiments, the PRM uses an angular threshold $\lambda_{th} = 30^\circ$ and considers $N_{nb} = 4$ neighbors.

PRM: Varying Camera Viewing Angle. In order to benchmark the PRM algorithm, we evaluated it on synthetic data in a *MATLAB* environment and compared it against a recently published method based on shape context and minimum spanning trees [7], denoted as SC.

The goal of this experiment was to assess the performance of the PPM algorithms with respect to the camera direction, thus varying projective distortion. The camera direction is always measured relative to the reference direction. For every sample camera direction, denoted as θ_{cap} , 5000 randomly selected camera positions within the range from 0° to θ_{cap} (Fig. 8) are generated and for each camera position a random point pattern consisting of nine 3D points is projected and tested. Whenever the sample θ_{cap} produces a correspondence match with at least 6 correct matches (minimum number of points required to compute the camera pose) with no mismatched points, the sample is considered as correct. The random point pattern is constrained to be in a rectangular volume of the same size as the points located on our real objects: 75 mm \times 45 mm and 15 mm in depth relative to the reference direction.

The results are shown in Fig. 9. Between 0° and 5.5° , SC shows better results. For larger viewing angles, PRM performs better.

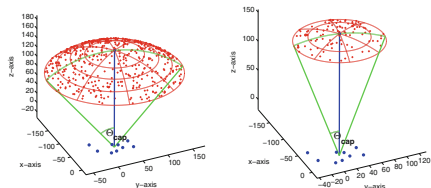


Fig. 8. Two different θ_{cap} . The reference direction is shown in blue, the red crosses on the spherical cap denote the random camera positions, and the blue dots represent a point pattern. (Color figure online)

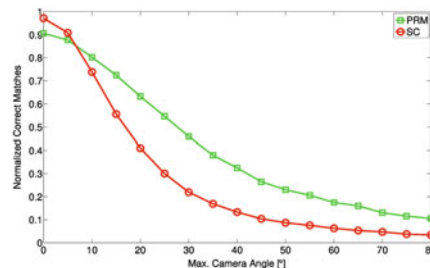


Fig. 9. Success rate with varying camera viewing angle.

Computational Performance of the Android Implementation. The measurements were taken on an *Asus Transformer Prime TF 201, Tegra 3 1400 MHz*. The average times were determined using natural images as seen in the medical application described below. The mean time elapsed for reading an image from the camera required 10 ms whereas the segmentation on the GPU and CPU required 45 ms. Together with an average of 3 ms required for the PRM (C implementation) and some additional overhead to render the 3D object into the image, this resulted in an average framerate of 15 frames per second. A significant amount of time is required to pass data through the *Java Native Interface*.

Accuracy of the AR Navigation System. The accuracy was determined by measuring the distance between a real point and its augmented location. The measurement was performed in the image and converted to metric units. From six different camera positions, one measurement for each was done. A mean error of 0.8 mm between the real and the augmented position could be evaluated. The standard deviation was 1.0 mm with a maximum error of 2.6 mm.

4 Medical Application

Together with the surgeons, the applicability of the proposed technique was shown in-vitro for tumor surgeries in the head and neck area. The surgeon placed the markers in locations well identifiable in the MR/CT scan as well as on the anatomical models. Figures 10 and 11 show our Android navigation solution at execution with the tumor and other critical structures overlaid onto the image.



Fig. 10. A tumor in mandibulae region and the alveolaris nerve.

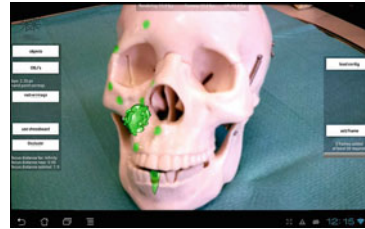


Fig. 11. A tumor in zygomaticum region and a missing tooth.

5 Discussion

The main conclusions that can be drawn from these experiments is that the PRM algorithm is more robust in comparison to the state-of-the-art in finding corresponding point pairs and is suitable for mobile real-time applications. Furthermore, the algorithm is based on a simple concept and can be implemented rather easily. The improved insensitivity, as compared to the state-of-the-art, to perspective changes, PRM lends itself to practical applications for example in clinical navigation tasks.

Although, the monte-carlo simulations showed a small disadvantage of PRM over SC for small perspective distortions, no such effect could be observed for real scenes.

A limitation common to all optical navigation solutions is the line of sight problem. During a surgical intervention landmarks can get covered for example by a surgical instrument or the hand of the surgeon. Once less than six landmarks can be seen no camera pose estimation is possible anymore. A further restriction is the rigid body assumption that is not valid in every case. Putting the landmarks on soft-tissue, for example skin, could thus pose additional challenges.

6 Conclusion

In this paper we proposed a single-view 3D navigation system applicable for general guidance tasks during surgeries. The achieved accuracy combined with its low cost opens a whole new field of easily deployable surgical navigation systems that could also find their application in third-world countries or in remote areas.

Beyond its clinical use, the introduced PPM algorithm performs well under perspective distortion and may contribute in general to AR applications on computational low-powered devices such as tablet computers or smartphones.

References

1. David, D., Duraiswami, S.: SoftPOSIT: simultaneous pose and correspondence determination. *Int. J. Comput. Vision* **59**(3), 259–284 (2004)
2. Denton, J., Beveridge, J.R.: Two dimensional projective point matching. In: *Proceedings of the Fifth IEEE Southwest Symposium on Image Analysis and Interpretation 2002*, pp. 77–81. IEEE (2002)
3. Goshtasby, A.A.: Point pattern matching. In: *Image Registration*, pp. 267–312 (2012)
4. Hofmann, R., Seichter, H., Reitmayr, G.: A gpgpu accelerated descriptor for mobile devices. In: *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 289–290. IEEE (2012)
5. Klein, G., Murray, D.: Parallel tracking and mapping on a camera phone. In: *8th IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2009*, pp. 83–86. IEEE (2009)
6. Kosmopoulos, V., Schizas, C.: Pedicle screw placement accuracy: a meta-analysis. *Spine* **32**(3), E111–E120 (2007)
7. Lian, W., Zhang, L., Zhang, D.: Rotation-invariant nonrigid point set matching in cluttered scenes. *IEEE Trans. Image Process.* **21**(5), 2786–2797 (2013)
8. McIlroy, P., Izadi, S., Fitzgibbon, A.: Kinectrack: agile 6-DoF tracking using a projected dot pattern. In: *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 23–29. IEEE (2012)
9. Oberst, M., Bertsch, C., Wüerstlin, S., Holz, U.: CT analysis of leg alignment after conventional vs. navigated knee prosthesis implantation. Initial results of a controlled, prospective and randomized study. *Der Unfallchirurg* **106**(11), 941 (2003)
10. Schweiger, F., Zeisl, B., Georgel, P., Schroth, G., Steinbach, E., Navab, N.: Maximum detector response markers for SIFT and SURF. In: *Int. Workshop on Vision, Modeling and Visualization (VMV)*, vol. 6 (2009)
11. Voss, K., Suesse, H.: Affine point pattern matching. In: Radig, B., Florczyk, S. (eds.) *DAGM 2001. LNCS*, vol. 2191, pp. 155–162. Springer, Heidelberg (2001)
12. WHO: Research for Patient Safety. http://www.who.int/patientsafety/information_centre/documents/ps_research_brochure.en.pdf (2008)

Chapter 7

Direct Calibration of a Laser Ablation System in the Projective Voltage Space

This paper was published at the *MICCAI Main Conference*, October 2015, München, Germany¹, won one of 50 travel awards and was chosen for an oral presentation. In cooperation with the *Advanced Osteotomy Tools AG*, Basel, Switzerland, a method to calibrate a laser ablation system was invented.

In this paper it was shown, how a laser deflecting tilting mirror can be approximately described by the projective transformation. As a consequence, existing camera calibration algorithms can be applied directly to calibrate the intrinsic and extrinsic parameters of the tilting mirror.

The key idea behind this work was given by Mathias Griessen. Even though his contribution is substantial, Mathias decided not to be listed as an author.

¹The paper is available online at http://link.springer.com/chapter/10.1007/978-3-319-24553-9_34.

Direct Calibration of a Laser Ablation System in the Projective Voltage Space

Adrian Schneider^{1,2}, Simon Pezold¹, Kyung-won Baek¹, Dilyan Marinov²,
and Philippe C. Cattin^{1,2}

¹ Department of Biomedical Engineering, University of Basel, Switzerland

² Advanced Osteotomy Tools AG, Basel, Switzerland

Abstract. Laser ablation is a widely adopted technique in many contemporary medical applications. However, it is new to use a laser to cut bone and perform general osteotomy surgical tasks with it. In this paper, we propose to apply the *direct linear transformation* algorithm to calibrate and integrate a laser deflecting tilting mirror into the affine transformation chain of a sophisticated surgical navigation system, involving next generation robots and optical tracking. Experiments were performed on synthetic input and real data. The evaluation showed a *target registration error* of $0.3 \text{ mm} \pm 0.2 \text{ mm}$ in a working distance of 150 mm.

Keywords: Robotics, Navigation, Laser, Mirror, DLT.

1 Introduction

Laser ablation is a tissue cutting technique that is widely adopted in ophthalmology and dentistry. Although such a *contact-free* cutting method would also be beneficial when cutting bones, i.e. in osteotomy, only little research has been invested in this area so far. One major reason for this was the lack of a compact laser source able to efficiently cut bone without carbonizing it.

With the proposed laser osteotome, see Fig. 1, we try to bridge this gap. To guarantee a high cutting precision, the laser source is directly mounted on a robot's end effector and is optically tracked using a stereo optical tracking device. A reflective mirror mounted on a 2-axes tilting mirror stage was introduced to deflect the laser beam. This tilting mirror permits quickly changing the direction of the laser beam. Large displacements are covered by the robot arm, whereas the small changes in the target location are handled by the tilting mirror. The question remains on how to align a voltage controlled mirror with the coordinate systems (CS) of the optical tracker, patient, and robot.

In this paper, we present a robust method to calibrate the tilting mirror and integrate it into the affine transformation chain. The essence of our approach is to apply a projective camera model to the given situation. In reference to Fig. 2, the camera center C corresponds to the mirror's laser deflecting spot D , from where light rays are received or emitted in a conical manner. In the case of an

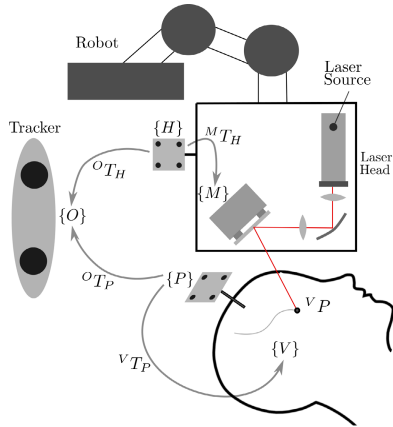


Fig. 1. Navigated laser system. Arrows denote affine transformations.

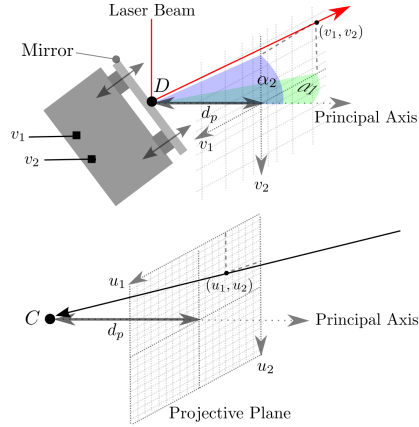


Fig. 2. Analogy of a laser tilting mirror (top) and a camera (below).

actual digital camera, the projective plane is an equidistant grid of photosensors u_1, u_2 , whereas the mirror operates in a *virtual* voltage space v_1, v_2 .

In the field of computer vision, well established camera calibration methods exist. In 1971, Abdel-Aziz and Karara introduced the *Direct Linear Transformation* (DLT) [1], which was a commonly used but rudimentary calibration technique. The main difference to contemporary calibration algorithms is that DLT does not consider nonlinear lens distortion effects, such as radial or tangential distortion. However, this is not required in the case of reflective optics as only mirrors and no lenses are involved. An interesting property of the DLT is that the extrinsic and intrinsic mirror parameters can be determined simultaneously.

In robotics, comparable work has been done on laser rangefinders, which perform depth measurements by triangulation of a moving laser beam and a camera [5]. Unfortunately, they focus only on extrinsic parameters and take the laser's steering properties (intrinsic) as given. The same pattern can be observed in many other applications. Often, the optical path and the mechanics, such as the steering mirror, are known very accurately. It makes sense to use this information and lock as many degrees of freedom as possible. In our system, the intrinsic parameters of the tilting mirror are known as well. But for two reasons we cannot use them: First, the optical setup of our prototype laser head is changing frequently. Second and more importantly, the regulatory authorities require the system to be calibrated on a regular basis when used in clinical practice, which certainly involves the determination of the mirror's overall properties.

Interesting related publications can be found in the field of catadioptric systems. From there we learned that it is common to apply a pinhole model to describe a moving mirror with a fixed center of reflection. A comparable situation is described in [4]. However, the setup there involves a hyperbolic mirror and a camera, which has to be calibrated with a non-linear approach. Our situation is comparably simple and we show that applying the DLT is appropriate.

2 Methods

2.1 Tilting Mirror Calibration with the DLT

As an input, the DLT algorithm requires several 2D–3D point correspondences. Using the pinhole camera model (Fig. 2), the projection of the i -th point from 3D spatial coordinates ${}^H X_i = [x, y, z, 1]_i^T$ to 2D pixel coordinates $u_i = [u_1, u_2, 1]_i^T$ on the projection plane is generally given by

$$\kappa u_i = \tilde{K} \cdot {}^M T_H \cdot {}^H X_i \Leftrightarrow \kappa \begin{bmatrix} u_1 \\ u_2 \\ 1 \end{bmatrix}_i = \begin{bmatrix} \tilde{f}_1 & 0 & \tilde{c}_1 \\ 0 & \tilde{f}_2 & \tilde{c}_2 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_i, \quad (1)$$

with the intrinsic parameter matrix \tilde{K} and the affine transformation ${}^M T_H$, where \tilde{K} holds the focal distances \tilde{f}_j and the principal point coordinates \tilde{c}_j ($j = 1, 2$), ${}^M T_H$ consists of a rotation part r_{11}, \dots, r_{33} and a translation part t_1, t_2, t_3 , and κ is a normalization constant. In our model, ${}^M T_H$ represents the rigid transformation of the mirror's CS $\{M\}$ with respect to the optical marker's CS $\{H\}$ (Fig. 1). The corresponding spatial coordinate ${}^H X_i$ is expressed in the $\{H\}$ CS and denotes the 3D point where the laser beam impacts.

Choosing the distance from the projective center to the projective plane $d_p = 1$, u_j ($j = 1, 2$) can be rewritten as $u_j = \tan(\alpha_j)$, where α_j denotes the angles of the two axes in which the deflection mirror is tilted. The angles are unknown, but they are linear to the known applied voltages v_j , enabling us to rewrite them as $\alpha_j = a_j v_j + b_j$ with the linearity parameters a_j, b_j . Putting these reformulations together, we can rewrite the projection as

$$\kappa \begin{bmatrix} \tan(a_1 v_1 + b_1) \\ \tan(a_2 v_2 + b_2) \\ 1 \end{bmatrix}_i = \tilde{K} \cdot {}^M T_H \cdot {}^H X_i. \quad (2)$$

The small-angle approximation enables us to simplify $\tan(\alpha_j) \approx \alpha_j$ ($\alpha_j < 10^\circ$), thus $\tan(a_j v_j + b_j) \approx a_j v_j + b_j$, which we use to simplify the projection as

$$\kappa \hat{C} \begin{bmatrix} v_1 \\ v_2 \\ 1 \end{bmatrix}_i \approx \tilde{K} \cdot {}^M T_H \cdot {}^H X_i \quad \text{with} \quad \hat{C} = \begin{bmatrix} a_1 & 0 & b_1 \\ 0 & a_2 & b_2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3)$$

Combining \hat{C} and \tilde{K} leads to the final approximative projection from spatial coordinates to voltage space:

$$\kappa v_i = \kappa \begin{bmatrix} v_1 \\ v_2 \\ 1 \end{bmatrix}_i \approx K \cdot {}^M T_H \cdot {}^H X_i \quad \text{with} \quad K = \hat{C}^{-1} \tilde{K} = \begin{bmatrix} f_1 & 0 & c_1 \\ 0 & f_2 & c_2 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4)$$

where K holds the new intrinsic parameters f_j ($= \frac{\tilde{f}_j}{a_j}$) and c_j ($= \frac{\tilde{c}_j}{a_j} - \frac{b_j}{a_j}$).

We would like to point out two important aspects of the final model. First, it is feasible for small tilting angles only. This holds in our case, as our mirror is operated in a range of $\pm 6^\circ$, resulting in a relative approximation error of 0.4%. Second, the actual angle–voltage relation need not be known, in fact: K and ${}^M T_H$ are calculated solely from correspondences between voltage pairs v_i and 3D points ${}^H X_i$, both of which are known. Showing the relationship between v_i, K and u_i, \tilde{K} was only necessary to establish the model.

The vector v_i is proportional to the vector $[K \cdot {}^M T_H \cdot {}^H X_i]$, see Eq. (4). As a consequence, their crossproduct (\times) is 0. Applying the DLT, we find K and ${}^M T_H$ by solving for their product $P = K \cdot {}^M T_H$, which leads to

$$v_i \times (P \cdot {}^H X_i) = 0 \Leftrightarrow \begin{bmatrix} v_1 \\ v_2 \\ 1 \end{bmatrix}_i \times \left(\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_i \right) = 0, \quad (5)$$

with $p_{11} \dots p_{34}$ as unknowns and $[v_1, v_2]_i^T, [x, y, z]_i^T$ given by the point correspondences. This can be converted into a linear system of equations

$$A \cdot \begin{bmatrix} p_{11} \\ \vdots \\ p_{34} \end{bmatrix} = 0 \quad \text{with} \quad A = \begin{bmatrix} 0 & 0 & 0 & 0 & -x_0 & -y_0 & -z_0 & 1 & v_{2_0} x_0 & v_{2_0} y_0 & v_{2_0} z_0 & v_{2_0} \\ x_0 & y_0 & z_0 & 1 & 0 & 0 & 0 & 0 & -v_{1_0} x_0 & -v_{1_0} y_0 & -v_{1_0} z_0 & -v_{1_0} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & -x_i & -y_i & -z_i & 1 & v_{2_i} x_i & v_{2_i} y_i & v_{2_i} z_i & v_{2_i} \\ x_i & y_i & z_i & 1 & 0 & 0 & 0 & 0 & -v_{1_i} x_i & -v_{1_i} y_i & -v_{1_i} z_i & -v_{1_i} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & -x_N & -y_N & -z_N & 1 & v_{2_N} x_N & v_{2_N} y_N & v_{2_N} z_N & v_{2_N} \\ x_N & y_N & z_N & 1 & 0 & 0 & 0 & 0 & -v_{1_N} x_N & -v_{1_N} y_N & -v_{1_N} z_N & -v_{1_N} \end{bmatrix}. \quad (6)$$

Each correspondence results in three equations, one of which is redundant due to linear dependence. Thus, a number of $N \geq 6$ correspondences is required to solve for the 12 unknowns. The resulting product P can then be decomposed into K and ${}^M T_H$ as described in [3].

Input Data Normalization: The algorithm described above is the basic DLT. To enhance the numerical stability, we first transform the 2D–3D point correspondences in order to reach certain spatial properties. In [3], this can be found as *Normalized DLT*. In the presence of measurement noise, it is highly recommended to normalize the input data.

2.2 Calibration Errors

Due to the small-angle approximation and measurement noise, the computed solution $K, {}^M T_H$ will not map the given N input 2D–3D correspondences perfectly. Several error measures can be applied to quantify the quality of the calibration procedure. The *algebraic error* is the residual of the underlying least squares problem in Eq. (6). The *backprojection error* E is a geometric error quantity in the voltage plane. With the computed calibration result, the acquired 3D points ${}^H X_i$ are virtually projected as v'_i and compared with corresponding voltage pairs v_i by computing their Euclidean distances, as

$$E_i = \left\| v_i - v'_i \right\| \quad \text{with} \quad v'_i \propto K \cdot {}^M T_H \cdot {}^H X_i \quad \text{for} \quad i = 1, \dots, N. \quad (7)$$

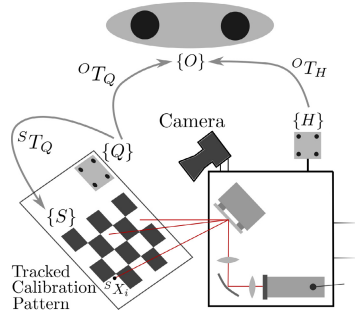


Fig. 3. Acquiring correspondences with a tracked calibration pattern.

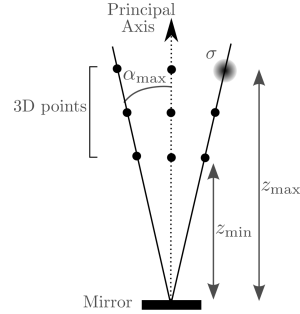


Fig. 4. Synthetic input data generation for the calibration.

The most important error measure for surgeons is the deviation from the planned location on the target site, i.e. on the patient herself and in millimetre, the so called *target registration error* (TRE). An error in the voltage plane E_i can be extrapolated with the distance between tilting mirror and target d_i , which is easy to determine after transforming the 3D points into the mirror CS. The TRE T_i is then computed for all N point correspondences as

$$T_i = d_i \left\| K^{-1} \cdot [E_i, 0, 0]^T \right\| \quad \text{with} \quad d_i = \left\| {}^M T_H \cdot {}^H X_i \right\| \quad \text{for} \quad i = 1, \dots, N. \quad (8)$$

2.3 Acquiring 2D–3D Correspondences

The accuracy of the mirror calibration depends strongly on the quality of its input data. Figure 3 illustrates our acquisition setup. The robot is driven into an appropriate position. Then $N_v \geq 1$ predefined voltage pairs v_i ($i = 1, \dots, N_v$) are applied to the tilting mirror. Their laser impact on a chessboard is recorded with a camera, and the standard blob-detector of *OpenCV* is used to recover their pixel positions, which can be easily transformed into the given chessboard CS ${}^S X_i$. It is important to notice that recovering ${}^S X_i$ with the camera is an independent process. Neither the relative position of the camera to the tilting mirror nor the rest of the system matters. However, a focused image preferably orthogonal to the chessboard enhances the accuracy of acquired positions.

In order to transform the laser position from the chessboard CS $\{S\}$ into the laser head CS $\{H\}$, first one has to resolve the transformation from the chessboard to its tracked marker ${}^S T_Q$. A common method based on fitting two 3D point sets [2] was applied for that purpose. The final transformation is

$${}^H X_i = ({}^O T_H)^{-1} \cdot {}^O T_Q \cdot ({}^S T_Q)^{-1} \cdot {}^S X_i, \quad (9)$$

where ${}^O T_Q$ and ${}^O T_H$ are given by the optical tracking system.

These steps are repeated from N_p different robot positions. Therefore, the total amount of collected 2D–3D correspondences is $N = N_v N_p$. A simple robot trajectory is orthogonal to the chessboard surface.

2.4 Integration of the Tilting Mirror

Using the notation introduced in Fig. 1, the transformations \mathcal{O}_{T_H} and \mathcal{O}_{T_P} are given by the tracking system. The 3D-3D registration ${}^V T_P$ from the patient marker to the operation planning data (CT, MR) can be performed with the method described in [2]. Given a cutting position ${}^V P$ on the patient, the two voltages (v_1, v_2) for the mirror can be computed by

$$[v_1, v_2, 1]^T \propto K \cdot M_{T_H} \cdot (\mathcal{O}_{T_H})^{-1} \cdot \mathcal{O}_{T_P} \cdot ({}^V T_P)^{-1} \cdot {}^V P, \quad (10)$$

which finally forms the complete transformation chain.

3 Experiments and Results

In this section, the performance of the described calibration approach is examined in detail based on synthetic and real input data. These experiments were performed by applying the normalized DLT approach.

Error Analyses with Synthetic Input Data: In these experiments, synthetic data was produced to analyze the presented method in terms of error behavior. As illustrated in Fig. 4, the calibration data generation can be configured by the four parameters α_{\max} (maximum angle for both mirror axes), z_{\min} and z_{\max} (distance in principal direction between the mirror and the 3D points), N_v (number of different voltage pairs applied in each robot position), N_p (number of different robot positions), and σ (standard deviation of zero-mean Gaussian noise applied to the 3D points). The point correspondences are generated in a deterministic way. For our simulated tilting mirror, v_1, v_2 are chosen to be equal to α_1, α_2 ($1^\circ/\text{voltage}$). Based on a given α_{\max} , an equidistant voltage array of size N_v is generated, where values of both axes α_1, α_2 are in the range of $-\alpha_{\max} \leq \alpha_{1,2} \leq \alpha_{\max}$. These are the 2D points. Their corresponding 3D coordinates are generated by applying this voltage array to the mirror and projecting to N_p different orthogonal planes with distance z_p , so that $z_0 = z_{\min}, \dots, z_{N_p} = z_{\max}$. As already mentioned, this leads to a total number of $N = N_v N_p$ point correspondences. To simulate the presence of noise, zero-mean Gaussian noise σ is added in all three dimensions to each 3D coordinate.

Maximum Deflection Angle Influence: Since the proposed method is based on the small-angle approximation, calibrations with increasing maximum deflection angle were performed. In particular, synthetic data sets with varying $\alpha_{\max} = 2^\circ, \dots, 8^\circ$ were generated, whereas the other parameters were kept constant at $z_{\min} = 140$ mm and $z_{\max} = 160$ mm, $\sigma = 0$, $N_v = 25$, and $N_p = 5$. Of each of these data sets, the mirror calibration was computed and the TRE T_i was determined and presented as a box plot, where the central mark is the median, the edges of the box are the 25th and 75th percentiles and the whiskers extend to the minimum and maximum errors. Figure 5 shows the results.

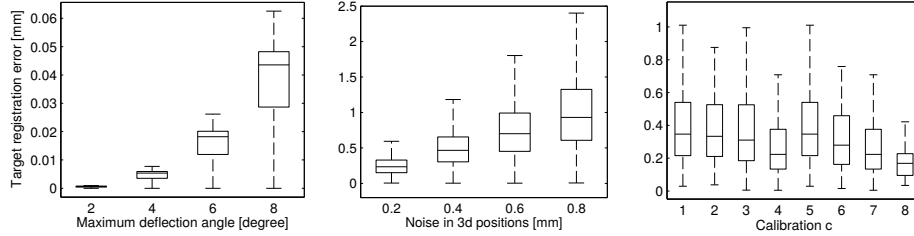


Fig. 5. TRE with increasing maximum mirror angle α_{\max} . **Fig. 6.** TRE with increasing Gaussian noise σ . **Fig. 7.** TRE for real acquired correspondences.

Impact of Noise: In this experiment, calibrations with increasing Gaussian noise $\sigma = 0.2\text{ mm}, \dots, 0.8\text{ mm}$ were made. Other parameters remained constant at $\alpha_{\max} = 6^\circ$, $z_{\min} = 140\text{ mm}$, $z_{\max} = 160\text{ mm}$, $N_v = 25$, and $N_p = 5$. For each noise level, $N_c = 100$ calibrations were made. Figure 6 shows the resulting TRE.

Results: One can clearly see that the tangent approximation for small angles is unproblematic. In our system, the maximum deflection angle α_{\max} is 6° , which corresponds to a maximum TRE of about $30\ \mu\text{m}$. Measurement noise, however, is an issue. The reported accuracy of our used tracking system is 0.25 mm . Based on the simulation, this corresponds to a TRE of about 0.5 mm .

Calibration with Real Data: In this experiment, the proposed calibration method was tested within the actual laser ablation system. As a tilting mirror, the OIM5001 (Optics In Motion) was used. The used optical tracking system was the CamBar B2 (AXIOS 3D Services) and the robot was the iiwa (KUKA Laboratories). The distance z between the chessboard and the tilting mirror was around 150 mm . The maximum deflection angle α_{\max} was about 6° . However, since the extrinsic and intrinsic mirror properties ${}^M T_H$, K are unknown at this time, the exact values of z and α can be determined only afterwards.

In the following, $N_c = 8$ independent tilting mirror calibrations were done. For each calibration c , $N_p = 5$ different robot positions along the chessboard normal were used. In each position, $N_v = 25$ voltage pairs were applied. Therefore, the maximum number of 2D–3D correspondences for each calibration c was $N = 125$. But due the regularly failure of the visual blob-detection within black chessboard fields, N showed to easily drop to 80.

Results: The results of the 8 calibrations can be seen in Fig. 7. The average TRE is 0.3 mm , with a standard deviation of 0.2 mm . The maximum error is 1.0 mm . The average distance z between the mirror and all involved correspondences was 141 mm and the average α_{\max} was 6.8° . When comparing the measured error with the results of noisy synthetic input data, this meets the expected error when using an optical tracking device with a spatial accuracy of 0.25 mm .



Fig. 8. The laser system performing a navigated cut on a sheep head.

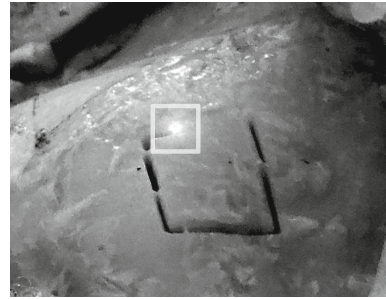


Fig. 9. Enlarged view of the cutting region. The bright spot is the laser.

4 Conclusion

We showed that a voltage controlled tilting mirror can be accurately calibrated by using the pinhole camera model and the *direct linear transformation* approach to solve it. With a *target registration error* of $0.3\text{ mm} \pm 0.2\text{ mm}$ at a working distance of 150 mm, our laser ablation system not only exceeds general osteotomy requirements, but also opens up new surgical possibilities in terms of cutting shapes. Although a maximum cutting error of 1.0 mm is tolerable, it does not meet our own demands. Currently we are designing non-planar optical markers to increase the tracking accuracy. Preliminary results are promising.

To conclude, we would like to give the reader an impression of the presented laser system in action. Figures 8 and 9 show a navigated cut on a sheep skull.

References

1. Abdel-Aziz, Y., Karara, H.: Direct Linear Transformation from Comparator Coordinates Into Object Space Coordinates in Close-range Photogrammetry (1971)
2. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of two 3-D point sets. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (5), 698–700 (1987)
3. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge Univ Press (2010)
4. Orghidan, R., Salvi, J., Mouaddib, E.M.: Calibration of a structured light-based stereo catadioptric sensor. In: Conference on Computer Vision and Pattern Recognition Workshop, CVPRW 2003, vol. 7, p. 70. IEEE (2003)
5. Zhang, Q., Pless, R.: Extrinsic calibration of a camera and laser range finder. In: Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004), vol. 3, pp. 2301–2306. IEEE (2004)

Chapter 8

Conclusions

In this thesis, four medical 3D navigation systems were invented, developed and examined. Different approaches of incorporating a projective camera into the field of computer assisted surgery was shown, such as fusing virtual data with a camera image or extract spatial information from this very same image as well.

In cooperation with surgeons, the medical applicability of augmented reality as visualization technique was examined in terms of usage and accuracy. Generally, this visualization concept of overlaying a camera image with virtual information in a perspective correct way feels intuitive and does not demand any instruction. The feedbacks from surgeons were overall positive. Especially in laparoscopic interventions, augmented reality is of great support. Following to urological surgeons, the operation time could be significantly shortened by finding the place of intervention faster and by receiving detailed planning information. Together with the same surgeons, an effort was done to bring this technology into the operation rooms and a trial with three human cadavers was prepared. Unfortunately, the cadaver study failed as the cadaver fixation altered the tissue properties such that surgeons were not able to install the navigation system described in Chapter 4. In particular, *Thiel-fixated* bodies were too soft, inasmuch as the ureter got pierced during the catheter insertion. On the other side, *formaldehyde-fixated* cadavers were too stiff, so that it was not possible to insert the catheter into female or male genitals. Possible solutions would have been to organize experiments with living animals or to enter directly human trials. However, for the former we do not have the facilities in Basel and the latter would be connected with an immense

effort in order to receive the clinical approval. Therefore, we decided to cancel the experiment, even though the trial would have given important answers of how augmented reality actually performs under real circumstances.

The achieved accuracy in those four navigation systems are not directly comparable, since the setups vary and different factors contribute to the navigation errors. Both single camera navigation systems described in Chapter 5, *Augmented Reality Assisted Brain Tumor Extraction in Mice*, and Chapter 6, *Landmark-Based Surgical Navigation*, achieved high accuracies relative to their costs. The first application showed a 3D accuracy of $0.48 \text{ mm} \pm 0.25 \text{ mm}$ and in the second system an augmented reality alignment error of $0.8 \text{ mm} \pm 1.0 \text{ mm}$ could be measured. It is important to notice that in the first application the effective 3D error was determined. In the second application, the setup allows only to assess the augmented reality error in the 2D image, though its actual 3D accuracy tends to be worse. There are three reasons which explain the navigation error difference between both systems. First, the marker accuracy of the chessboard is much higher than the accuracy of the landmarks attached to the target. Second, the distances between the camera and the chessboards are shorter than the distances in the application with the landmarks. And last, the usage of a high-class camera and a professional lens leads to a better accuracy than when using a tablet computer's optic.

The application described in Chapter 4, *Augmented Reality Assisted Laparoscopic Partial Nephrectomy*, uses a commercial electromagnetic tracking device to perform surgical navigation. Even though the measured augmented reality alignment error of $2.1 \text{ mm} \pm 1.2 \text{ mm}$ is sufficient for this particular surgical intervention, it is less accurate than the cost-effective single camera systems described in the Chapters 5 and 6. One major reason is the complexity of the setup itself. The system is operated under a large volume and the kinematic chain of transformations incorporates several registration and calibration steps. This induces significant additional spatial errors. Further on, electromagnetic 3D tracking is strongly affected by ferromagnetic disturbances. In the presence of surgical tools, a solely tracking error above 1 mm can be expected.

The surgical instrument published in Chapter 7, *Direct Calibration of a Laser Ablation System in the Projective Voltage Space*, can be compared against the

three other applications to greatest possible extent in context of a camera calibration. However, this camera, respectively tilting mirror calibration error, is given in voltages and not in pixel. Thus, a comparison is meaningless. From a pure navigation error point of view, the achieved laser ablation accuracy on the target of $0.3 \text{ mm} \pm 0.2 \text{ mm}$ in a distance of 150 mm is more than sufficient.

Chapter 9

Outlook

The four presented projects are at a different state. Some are in progress, some are completed and others are stopped. In the following, a short outlook for each project is given.

- From a technical point of view, the project described in the publication of Chapter 4, *Augmented Reality Assisted Laparoscopic Partial Nephrectomy*, is mainly completed. A possible extension could be the integration of a stereoscopic laparoscope. That would enable the navigation system to offer depth impressions to the surgeons. Certainly, such a stereoscopic laparoscope could be used for many other interesting tasks as well, like surface reconstruction and registration.

At the clinical side, the next step would be to enter clinical trials. Personally, I think that it is unlikely that the University is providing the required resources to proceed with this project. It might be necessary to cooperate with an industrial partner. Further, one should also try to find other clinical applications for this type of navigation system.

- The high accuracy navigation system described in Chapter 5, *Augmented Reality Assisted Brain Tumor Extraction in Mice*, is finished and was several times successfully used to extract brain tumors in mice. Currently, there are discussions to apply this method also to extract farmed tumors in mice kidneys.

The tracking accuracy could be increased by integrating a second camera. However, that would increase also the complexity and the cost of the system, which is in contrary to its key features. A better solution could be the usage of nonplanar visual markers, which are accurately measured by using a *Micro-CT*. In the same run, even the two registrations from the chessboards to their corresponding CSs could be determined. A first test showed an amazing registration error of 21 μm .

- The navigation solution introduced in Chapter 6, *Landmark-Based Surgical Navigation*, is currently extended with a second motion model. In particular, the optical camera pose computation is combined with the output of the built in *inertial measurement unit*, a pose estimation sensor, so that the navigation is able to work for a short time also without markers. This could be interesting to create the 2D–3D correspondences too. The presented 2D–2D point correspondence algorithm could be significantly enhanced and accelerated by including this prior pose information.
- From a scientific point of view, the calibration method presented in Chapter 7, *Direct Calibration of a Laser Ablation System in the Projective Voltage Space*, should be further examined in terms of stability and correctness. So far, the calibration error was computed by projecting and comparing the point correspondences by using the determined mirror properties, and this led to low errors. However, we noticed that a low calibration error does not necessarily comes from a good calibration. We observed that an inaccurate mirror pose can be compensated by wrong intrinsic parameters, but resulting in a good calibration error. This is harmless for many applications. But our laser focuses in a certain distance from the mirror. Thus, it is crucial to operate with accurate extrinsic parameters.

Since early 2015, we are performing animal trials on a regular basis. At the same time, tedious regulatory work has to be done to get the system approved by the regulatory body. So, there is as well a lot of safety related issues to be tackled.

In general, I hope that future research projects in the field of surgical navi-

gation and augmented reality focus a little bit more on the clinical translation. It is an important task on its own, inasmuch as I strongly believe that these technologies could enrich the daily medical work.

CURRICULUM VITAE

ADRIAN SCHNEIDER

PARTICULARS

Address: Seestrasse 102, CH-3654 Gunten
Mobile: +41795924965
Email: adrian.schneider@eidelen.ch
Birthday: 21.02.1981

SKILLS

C++ / C / Matlab Programming
Computer Vision, Robotics and Image Processing
Experienced Smartphone Developer

EDUCATION

- | | |
|------------|---|
| since 2012 | PhD thesis <i>Projective Camera Model in Biomedical Navigation Applications</i> at the University of Basel |
| 2011–2012 | Master thesis <i>Augmented Reality in CMF Surgery</i> at the University of Basel |
| 2009–2011 | Master in Biomedical Engineering at the University of Bern <ul style="list-style-type: none">• Image processing, pattern recognition, numerics• Biomechanics, human physiology & anatomy |
| 2004–2005 | Graduation work <i>BlueSox</i> at Swisscom Innovations |
| 2000–2005 | Bachelor in Computer Science at the University of Applied Sciences Biel <ul style="list-style-type: none">• Telematics, software engineering |
| 1997–2000 | Maturity and education as chemical laboratory worker at Nitrochemie Wimmis |

PROFESSIONAL EXPERIENCES

- | | |
|------------|--|
| since 2013 | Research at Advanced Osteotomy Tools AG (AOT), Basel <ul style="list-style-type: none">• Research in computer assisted surgery and robotics• Responsible for the optical 3D navigation• Medical certified software development• Scientific publications |
|------------|--|

- 2009–2013 Senior SW Engineer at SwissQual, Zuchwil
- Android system and Linux kernel engineering
 - Symbian application development
- 2008–2009 Senior SW Engineer & Architect at Nokia, Helsinki
- Symbian firmware development and maintenance
- 2006–2007 Senior SW Engineer at Flander OY, Helsinki
- Implementation of Symbian *voice over IP* components
- 2002–2006 SW Engineer at SwissQual, Zuchwil
- Symbian C++ / C / Assembly programming
- 2006 Lecturer at the University of Applied Sciences Biel for the course *Symbian Mobile Programming*
-

LANGUAGES

German	mother tongue
English	fluent
Finnish	basics

PROGRAMMING

C++, C, CMake	→	QT, OpenCV, OpenGL
Matlab/Octave, Python	→	Signal processing, image processing, simulations
Java	→	Android, web applications
PHP, SQL	→	Web pages
L ^A T _E X	→	Documents

OPERATING SYSTEMS

Linux, Windows	→	Application development (10 years)
Symbian OS	→	Application and system development (8 years)
Android	→	Application, system and kernel development (3 years)
OS X	→	Application development (2 years)

REFERENCES

University of Basel & AOT AG PhD Supervisor and CTO of AOT AG
Prof. Dr. Philippe C. Cattin; philippe.cattin@unibas.ch

Publications

- 2015 | Adrian Schneider, Simon Pezold, Kyung-won Baek, Dilyan Marinov, and Philippe C Cattin. Direct calibration of a laser ablation system in the projective voltage space. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*. Springer, 2015
- 2015 | Adrian Schneider, Peter Thalmann, Simon Pezold, Simone E. Hieber, and Philippe C Cattin. Augmented reality assisted brain tumor extraction in mice. In *Salento AVR*. Springer, 2015
- 2015 | Waldemar Deibel, Adrian Schneider, Marcello Augello, Alfredo E Bruno, Philipp Juergens, and Philippe Cattin. A compact, efficient, and lightweight laser head for carlo®: integration, performance, and benefits. In *SPIE Optical Engineering+ Applications*, pages 957905–957905. International Society for Optics and Photonics, 2015
- 2014 | Adrian Schneider, Simon Pezold, Andreas Sauer, Jan Ebbing, Stephen Wyler, Rachel Rosenthal, and Philippe C Cattin. Augmented reality assisted laparoscopic partial nephrectomy. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 357–364. Springer, 2014
- 2013 | Adrian Schneider, Christian Baumberger, Mathias Griessen, Simon Pezold, Jörg Beinemann, Philipp Jürgens, and Philippe C Cattin. Landmark-based surgical navigation. In *Clinical Image-Based Procedures. Translational Research in Medical Imaging*, pages 57–64. Springer, 2014

Talks

- 2015 | Paper presentation at the *MICCAI*, October 2015, München, Germany
- 2015 | Paper presentation at the *Salento AVR*, September 2015, Lecce, Italy
- 2015 | Paper presentation at the *DBE Research Day*, August 2015, Basel, Switzerland
- 2013 | Paper presentation at the *Workshop on Clinical Image-Based Procedures - Translational Research in Medical Imaging*, September 2013, Nagoya, Japan

References

- [1] WHO, “Research for Patient Safety,” 2008. [1](#)
- [2] V. Kosmopoulos and C. Schizas, “Pedicule screw placement accuracy: a meta-analysis,” *Spine*, vol. 32, no. 3, pp. E111–E120, 2007. [2](#)
- [3] M. Oberst, C. Bertsch, S. Wüerstlin, and U. Holz, “CT analysis of leg alignment after conventional vs. navigated knee prosthesis implantation. Initial results of a controlled, prospective and randomized study.,” *Der Unfallchirurg*, vol. 106, no. 11, p. 941, 2003. [2](#)
- [4] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge Univ Press, 2010. [6](#), [11](#)
- [5] K. S. Arun, T. S. Huang, and S. D. Blostein, “Least-squares fitting of two 3-d point sets,” *Pattern Analysis and Machine Intelligence, IEEE Trans. on*, no. 5, pp. 698–700, 1987. [8](#)
- [6] Y. Abdel-Aziz and H. Karara, *Direct Linear Transformation from Comparator Coordinates Into Object Space Coordinates in Close-range Photogrammetry*. 1971. [12](#)
- [7] J. Heikkila and O. Silvén, “A four-step camera calibration procedure with implicit image correction,” in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pp. 1106–1112, IEEE, 1997. [12](#)
- [8] Z. Zhang, “A flexible new technique for camera calibration,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, pp. 1330–1334, 2000. [12](#)

REFERENCES

- [9] K. Levenberg, “A method for the solution of certain non-linear problems in least squares,” 1944. [14](#)
- [10] Y. Zheng, S. Sugimoto, I. Sato, and M. Okutomi, “A general and simple method for camera pose and focal length determination,” in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pp. 430–437, IEEE, 2014. [14](#)