POLITECNICO DI MILANO
Master of Science in Automation Engineering
Department of Electronics, Information and Bioengineering

Quaternion-Based Unscented Kalman Filter for Robust Motion Tracking in Neurosurgery

NearLab
Neuro Engineering and medicAl Robotics Laboratory

Supervisor : Prof. Giancarlo Ferrigno

Co-Supervisor : Dr. Elena De Momi

Nima Enayati

Academic year 2012/2013

(This page intentionally left blank.)

To my parents.

Abstract

This dissertation presents a fusion algorithm designed for robust estimation of the pose of a freely moving target in neurosurgery. The filter used for the fusion processes data from an optical tracking system (OTS) and an inertial sensor unit (IMU) containing tri-axial angular rate sensors and accelerometers. While commercial optical tracking systems and inertial measurement units suffer respectively from a low frame frequency and error accumulation, by blending the data from both sensors, the sensor fusion system maintains the advantages of both, i.e. accuracy of OTS and high sampling frequency of IMU, and compensates their drawbacks.

The blending of data from heterogeneous sensors or sensor units is commonly performed via an algorithm derived from Kalman filter. Specifically for systems that involve nonlinear behaviors Extended Kalman Filter (EKF) has been widely implemented. However, due to nontrivial drawbacks of the EKF that can affect the accuracy or even lead to divergence of the system, recently a number of related novel, more accurate and theoretically better motivated algorithmic alternatives to the EKF have surfaced in the literature, with specific application to state estimation for automatic control. The Unscented Kalman Filter (UKF) is a linear estimator which yields performance equivalent to the Kalman filter for linear systems, yet generalizes elegantly to nonlinear systems without requiring the explicit linearization steps required by the EKF. The fundamental component of this filter is the unscented transformation which uses a set of appropriately chosen weighted points to parameterize the means and covariance of probability distributions.

The system represents rotation using quaternions rather than Euler angles or axis/angle pairs. A quaternion representation of the orientation is computationally effective and avoids

problems with singularities. The nonlinearities arising by using the quaternion representation are dealt with by the UKF in an efficient manner.

Beyond improved performance in tracking, the designed system can compensate for brief optical marker occlusions by estimating the pose of the object using only inertial measurements. The accumulated error due to sensor drift is corrected as soon as optical measurements are available. Implementation and testing results of the quaternion-based Unscented Kalman filter are presented. Experimental results validate the filter design, and show the feasibility of using optical/inertial sensor fusion for robust motion tracking satisfying the requirements of neurosurgical computer assisted procedures.

Sommario

Questa tesi presenta un algoritmo di fusione progettato per la stima robusta della posa di un target in movimento libero in neurochirurgia. Il filtro utilizzato per la fusione elabora i dati provenienti da un sistema di localizzazione ottica e da un sensore inerziale composto da sensori di velocità angolare e accelerometri su i tre assi. Mentre i sistemi di localizzazione ottica commerciali e i sensori inerziali soffrono rispettivamente di una bassa frequenza di fotogramma e dell'accumulo di errore, fondendo i dati provenienti da entrambi i sensori, il sistema di fusione dei sensori mantiene i vantaggi di entrambi, ovvero l' accuratezza della localizzazione ottica ed l'alta frequenza di campionamento del sensore inerziale, compensando i loro svantaggi.

La fusione dei dati provenienti da sensori eterogenei viene comunemente effettuata tramite un algoritmo derivato dal filtro di Kalman. In particolare, il filtro di Kalman esteso (EKF) è stato implementato per i sistemi che presentano comportamenti non lineari. Tuttavia, a causa di inconvenienti non banali del EKF che possono influenzare la precisione e anche la divergenza del sistema, di recente sono emersi in letteratura algoritmi alternativi all' EKF, più precisi e teoricamente migliori per applicazioni specifiche nella stima per il controllo automatico. Il filtro di Kalman Unscented (UKF) è uno stimatore lineare che produce prestazioni equivalenti al filtro di Kalman per sistemi lineari, ma generalizza elegantemente per sistemi non lineari, senza richiedere la procedura di linearizzazione richieste dal EKF. La componente fondamentale di questo filtro è la trasformazione Unscented che utilizza un insieme di punti scelti opportunamente ponderati per parametrizzare la media e la covarianza di una distribuzione di probabilità.

Il sistema rappresenta la rotazione utilizzando quaternioni anziché gli angoli di Eulero o la convenzione asse/angolo. Una rappresentazione dell'orientamento in forma di quaternione è computazionalmente efficace ed evita problemi di singolarità. Le non linearità derivanti utilizzando la rappresentazione quaternione sono trattate dalla UKF in modo efficiente.

Oltre a migliorare le prestazioni in termini di tracking, il sistema progettato può compensare brevi occlusioni dei marcatori ottici, stimando la posa dell'oggetto utilizzando solo misure inerziali. L' errore accumulato a causa del drift del sensore viene corretto non appena le misure ottiche sono disponibili. Risultati riguardanti l' implementazione e i test eseguiti sul filtro di "quaternion-baesed Unscented Kalman" vengono presentati. I risultati sperimentali convalidano la progettazione del filtro, e mostrano la possibilità di utilizzare la fusione di sensori ottichi/inerziali per la stima robusta del movimento e soddisfano i requisiti delle procedure neurochirurgiche assistite dal computer.

Acknowledgements

I would like to thank my supervisors and colleagues at Politecnico's NEAR Lab for their intellectual support, and for providing a stimulating academic environment. In particular, I would like to thank my supervisor, Prof. Giancarlo Ferrigno, for his time, support and insight. Special thanks also go to my co-supervisor Dr. Elena De Momi for tirelessly checking on the progress of the project, as well as for taking the time to review this work. Alberto, Mirko, Danilo, Marta and Elisa for their input and suggestions along the way, and Veronica, Luca and all my charming colleagues for support, humor and giving me the opportunity to improve my Italian in the lab!

(This page intentionally left blank.)

Table of Contents

List of Figures

List of Tables

(This page intentionally left blank.)

Chapter 1: Introduction

1.1 Computer-Assisted Surgery

At present, several terms are used for interactive interventions using computers: computer-assisted surgery, computer-integrated surgery, computer-aided surgery, image-guided surgery and navigated surgery that cover a wide range of computer applications during surgery such as robotics, 3D rendering, surface simulations, and intraoperative localization. The general aim is to define operative techniques which make use of advanced computer technology to process data from pre- and intraoperative sources, to present the most relevant information and achieve better performance during some parts of the surgical procedure.

"Surgery is a side-effect of therapy." It can be inferred from this statement that surgical trauma is mainly the result of reaching the target to be treated, rather than the consequence of the treatment itself. This observation has led to the development of systems that approach the affected tissue using a minimally invasive method in which surgical instruments are inserted into the body cavity via one or more small ports, often accompanied by an endoscopic video camera and/or imaging probes such as ultrasound, with externally manipulated instruments performing the intervention.

It is not only trauma that can be improved by Computer-assisted surgery (CAS). As it had been observed about 20 years ago by[1] : "…the safety and quality of the operation will be increased, time consuming procedures can be avoided, and novel procedures in surgery become possible."

Image-guided systems (IGS), augmented reality (AR) and navigation in endoscopic soft tissue surgery are among the main topics of surgical navigation. IGS involve matching the coordinates from medical imaging (preoperative registration) with coordinates from the patient in the operating room (registration and updating images). IGS have become the standard of care in providing navigational assistance during neurosurgery, offering subsurface and functional information to the surgeon.

1.2 Surgical Robotics

"The combination of magnification of the operative field and tool miniaturization has overwhelmed the spatial resolution of the adult human hand. Robots, in contrast, are capable of minute, tremor-filtered movements and are indefatigable." [2]

Frame-based technologies forced neurosurgeons to adapt their microsurgical techniques to the rigid stereotactic systems, which worked with high accuracy but reduced flexibility. Although frame-based stereotactic systems increased their intraoperative flexibility, it was computer-based navigation which, working interactively with images without visible coordinates and without instruments in the operating field, suited the visually oriented neurosurgeons.

Surgical robotics was the obvious next step, enabling minimally invasive procedures to be performed with similar or greater dexterity, as under open surgery. Since then, dozens of research projects have been focusing on brain and spine surgery, differently addressing the challenges of accuracy and effectiveness. The field includes a wide range of procedures: Robot guided biopsy needle insertion for Brain tumor biopsy, Robotic placement of electrodes in deep brain stimulation for Parkinson's disease, Robotic 3-D positioning of depth electrodes to determine focus of epileptic seizure for Epilepsy, Grid pattern tumor treatment, etc.

The main motivation for implementing robotic surgery is clearly to improve the safety and consistency of procedures, as well as the ability to minimize traumatic and often disfiguring incisions to access the target organ. However, robotic surgery offers much more than these. Current research projects are not only dealing with conventional surgical robots, but also new approaches that use robotics-inspired principles toward enhancing the capabilities of conventional surgical approaches; Surgical robotics, therefore, has become a very active area that promises to revolutionize the way many

surgical procedures are performed today. There is no doubt that robotic surgery is transforming the current surgical practice, and not necessarily only those for which robotic technology was originally developed.

1.3 Problem Statement

Image-guided minimally invasive interventions began with Stereotactic neurosurgery, where a target was defined in the brain, and located with respect to a coordinate system defined by a frame attached to the patient's skull with bone screws. While the accuracy of navigation is generally an important factor in computer aided surgery, it is of critical importance in computer aided neurosurgery, due to more restricted requirements.

The error of the navigation can be divided into the technical error of the device calculating its own position in space, the registration error due to inaccuracies in the calculation of the transformation matrix between the navigation and the image space, and the application error caused additionally by anatomical shift of the brain structures during operation.[3]

Current systems allow for sub-millimeter intraoperative spatial positioning, however, certain limitations still remain. Unintended changes are prone to happen in the OR setup, including patient's movement relative to the robot tool frame. This can happen not only due to external sources (forces applied by surgeon, personnel bumping into the operating table, inadequate head fixation etc.) but also due to seizure, that is not uncommon in deep brain stimulation, tumor or epilepsy surgery, in which the patient is kept awake to map functional neural bases via electrophysiological assessment. These disturbances are aperiodic, unpredictable and it is not possible to model them.

Under these circumstances, the positioning inaccuracy in procedures involving robots is a significant danger to patients and accurate and robust tracking of the patient's head and surgical tools is a prime requirement. This research tries to study this need and provide a practical solution that can be implemented in modern operation rooms.

1.4 Tracking and Sensor Fusion

Modern operating rooms often rely on optical tracking systems (OTS) for tracking the patient and surgical tool positions. An optical tracking system consists of stationary cameras and optical markers attached to the object which is being tracked. Commercial optical tracking systems used in computer-assisted surgery generally have state of the art accuracy but they suffer from a low bandwidth of 10–60 Hz and the risk of marker occlusion which make them unsuitable for servo-control of a handheld tool or

continuous head pose tracking. A viable solution would be using a second sensor to compensate for the flaws of the OTS. This process of blending data is often referred to as data fusion or multi-sensor data fusion (see for example [4]).

Sensor fusion is an established technique to improve the quality and reliability of sensor data. Several sensor systems are used to observe a process, and the information from these systems is combined in such a way that it will make use of each individual sensor's strengths and compensate their weaknesses. Beyond improved performance in limited working volumes, future systems will almost certainly involve large-scale hybrid systems, necessitating the blending of data from heterogeneous sensors or sensor units.

The combination of relatively high bandwidth and no line of sight requirement makes inertial measurement units (IMU) a potential candidate. With their small size and low cost IMUs can be integrated into the measurement platforms for both handheld tools and patient's head. The inherent drift problem of the IMUs can be compensated by the accurate OTS measurements.

1.5 Pose estimation by Kalman Filter

Generally the goal in tracking is not finding where the object is at present but knowing where it will be in, for example, two time steps ahead. On the other hand sensors are prone to introduce noise into the measurements and the "real" value needs to be estimated. Within the space of approximate solutions, Kalman Filter (and its extended versions) has become one of the most widely used algorithms with applications in state estimation. The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means of using noisy measurements to estimate the state of a linear system, while minimizing the expected mean-squared estimation error.

On the other hand, fusion of the data from multiple observations is often achieved through some variant of the Kalman filter. The capability of heterogeneous data fusion, combined with the state estimation and predictive applications, makes the Kalman filter (and its extended versions) the optimum method for our problem.

1.6 Research Objectives and Work Overview

This work aims to provide a practical solution for a robust tracking system in the robotic neurosurgery application. As it will be discussed later, similar efforts that have been pursued regarding this or similar problems lack either the required accuracy or appropriate orientation representation. In our approach the pose estimation is achieved

by fusing the data from an optical tracking system and an inertial measurement with a modified Kalman filter, the Unscented Kalman Filter (UKF). The proposed system increases the low acquisition frequency of the optical tracking system and compensates brief marker occlusion, allowing for a robust and accurate pose tracking.

The following paragraph gives the structure of this M.Sc. thesis:

An extensive literature review is done in **Chapter 2**, starting from a brief review of surgical robotics arriving at researches targeting motion tracking and sensor fusion. **Chapter 3** addresses the methods implemented in this work. The used hardware, software and algorithms are described. The quaternion-based unscented Kalman filter is explained and finally the experimentation phase and methods are presented. The results of the experiments and simulations are exhibited in **Chapter 4** and the final discussions and evaluation of the performance are presented in **Chapter 5**.

Chapter 2: Literature review

Several platforms have been presented in the literature for object tracking. Some of which benefit from multi sensor measurements while others simply implement single sensors to estimate the object's pose. This chapter provides a brief literature review, starting with a short summary of efforts in Computer-assisted Neurosurgery and Neurosurgical Robotics and continuing with a review of related motion tracking and sensor fusion researches.

2.1 Computer-assisted Neurosurgery

At the beginning of 20th century, the topic of diagnostic localization of lesions could be understood only by analyzing the neurologic symptoms of the patient, without the possibility of referring to radiologic images. The first imaging technique, introduced in 1918, was visualization of the ventricles by direct injection of air and later of contrast medium into the ventricles. Later, in 1927, angiography was described but direct visualization of the cerebral tissue was not possible before CT was introduced in 1973.

Intraoperative localization technique was based on knowledge of characteristic bony landmarks and the neurosurgeon's skill and 3D knowledge. This anatomic localization method was the gold standard not only before CT and MR imaging but also after their introduction. Micro-neurosurgeons used this detailed anatomic image information for

better intraoperative identification of the anatomic structures and sophisticated planning of approaches.

Parallel to the anatomic localization strategy, there was also the tendency from the beginning of neurosurgery to define the anatomic and pathologic structures in advance by using mechanical devices to define accurate approaches to targets and to provide objective information independent of the individual surgeon's skills. Frame based stereotactic localization techniques were developed based on a rigid coordinate system in which the target and a straight trajectory were determined based on the image information. However, until recently micro-neurosurgeons were more comfortable with intraoperative anatomic identification than with stereotactic coordinates. The reason for this is not only that micro-surgeons like to decide each step separately, depending on the changing anatomic situation in the course of operation, but also because they trust more what they see rather than what computers calculate in abstract coordinates.

Frame-based and robotic technologies forced neurosurgeons to adapt their microsurgical techniques to the rigid stereotactic systems, which worked with high accuracy but reduced flexibility. Although frame-based stereotactic systems increased their intraoperative flexibility, it was computer-assisted methods which, working interactively with images without instruments in the operating field, suited the visually oriented neurosurgeons.

The idea of frameless, interactive, computer-aided surgery consisted in navigation systems able to show in real time the position of the tip of an instrument in the corresponding images and not requiring a stereotactic frame for calculation. This required efficient methods to track the surgical instrument and patient's parts. Because the tracking devices were basically adaptations of robot technology to surgical applications in regard to all the mathematical and technical knowledge necessary for real-time navigation, it seems justified to start with a very brief historical review of surgical robots.

2.2 Neurosurgical robotics review

In the middle of the 1980s, all the necessary mathematical and technical presuppositions existed for the realization of devices for navigated surgery: Fast computers with appropriate data banks to handle image-based information in real time, hardware and software for manipulators based on industrial robot technology with a high degree of accuracy, a high technical standard of image processing partly due to space research programs. Modified industrial robots with higher positioning accuracy were used in neurosurgery in the beginning of 1990s by Drake [5], Kwoh [6], Benabid [7], Burckhardt [8] and Nakamura [9].

This technology had to be assembled in such a way as to satisfy surgical requirements. Passive manipulators, digitizers, and sensory arms which were able to determine their own position in space and transfer this information into images were introduced into neurosurgery in 1986 by Roberts [10]. He adapted a microscope for navigated surgery, equipped with ultrasound-emitting sources and microphones arranged outside the operating field. The microphone data were transported to a computer, which calculated the position of the microscope in space. The target point chosen on the CT/MR images could be projected into the ocular of the microscope and used for orientation during the operation.

The first reported use of a robot in neurosurgery was in 1985 by Kwoh and colleagues [6], who employed a Programmable Universal Machine for Assembly (PUMA) industrial robot for holding and manipulating biopsy cannulae. Although the robot served only as a holder/guide, the potential value of robotic systems in surgery was evident. In 1991, Drake and coworkers [5] reported the use of a PUMA robot as a retraction device in the surgical management of thalamic astrocytomas. Despite their novel application, both systems lacked the proper safety features needed for widespread acceptance into neurosurgery. Beginning in 1987, Benabid et al [7] experimented with an early precursor to the robot marketed as NeuroMate (Integrated Surgical Systems, Sacramento, CA). NeuroMate uses preoperative image data to assist with surgical planning and a passive robotic arm to perform the procedure. The NeuroMate system has been used in more than 1,000 cases.

In 1987, Watanabe independently introduced an arm-based navigation system for neurosurgical operations [11]. In Switzerland in 1988, Reinhardt [12] was working on an armless navigation system which used a pointer equipped with ultrasound sources. Magnetic sources were also described later by Kato [13], and infrared light-emitting diodes (LEDs) as emitting sources by Zamorano [14].

These first neurosurgical robots relied on preoperative images to determine robotic positioning. As a result, surgeons could not dynamically monitor needle placement under image-guidance and were blind to changes such as brain shift. To satisfy the need for a real-time, image-guided system, Minerva was developed (University of Lausanne, Lausanne, Switzerland). The system consisted of a robotic arm placed inside a Computed Tomography (CT) scanner, thus allowing surgeons to monitor the operation in real-time and make appropriate adjustments to the trajectory as needed[8].

Despite considerable engineering challenges, the design and construction of magnetic resonance (MR)– compatible robotic systems soon followed. MR compatibility ensures that the robot produces minimal MR imaging (MRI) artifact and that the operation of the robot is not disturbed by the electromagnetic field. Proper material selection is critical to avoid adverse effects on image distortion or changes in contrast and signal-to-

noise ratio. The benefits of MR soft-tissue visualization nevertheless prompted investigators from Harvard University (Cambridge, MA)[15], the University of Tokyo (Tokyo, Japan) [9], and the University of Calgary (Calgary, Alberta, Canada) [16] to develop their own MR-compatible robotic systems.

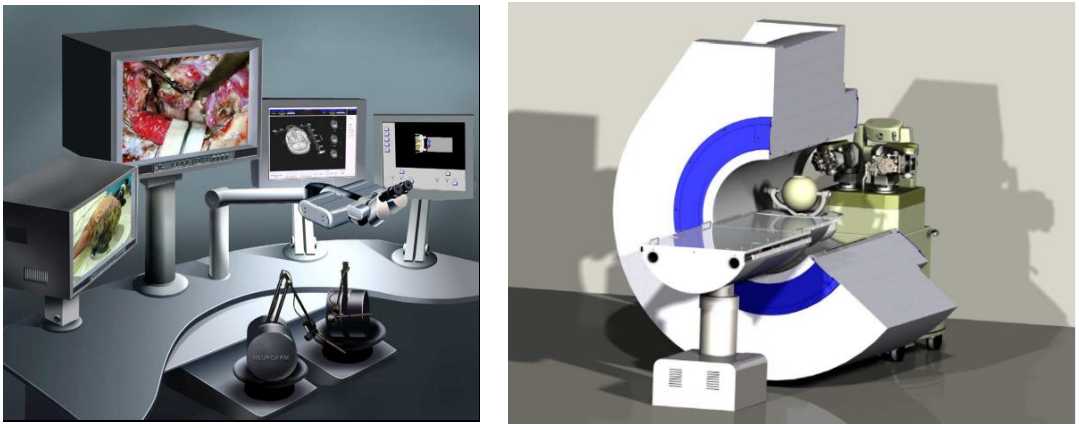

Figure 2.1 NeuroArm (University of Calgary, Calgary, Alberta, Canada) in position for stereotaxy. Figures are from [2]

2.3 Developed Neurosurgical Robots

At the moment only one general surgery robot is commercially available: The daVinci™ [17] system developed by Intuitive Surgical Inc. It has been evaluated mainly in the field of minimally invasive heart surgery and has not been implemented for neurosurgery. The Robot-Assisted Microsurgery System (RAMS) was developed by NASA to enable new procedures for brain, eye, ear, nose and throat but has not been commercialized. Finally, NeuroArm, a research surgical robot is specifically designed for neurosurgery. It is the first image-guided, MR-compatible surgical robot that has the capability to perform both microsurgery and stereotaxy and will be commercialized under the name SYMBIS™ Surgical System. The system is based on master-slave control in which commanded hand-controller movements are replicated by the robot arms (Figure 2.1). In 2008, NeuroArm operated on a on a human patient at the Faculty of Medicine, University of Calgary. This landmark operation was the first time a robot was used to perform image-guided neurosurgery, removing a brain tumor from the 21-year-old patient [18].

2.4 Motion Tracking in Computer Assisted Surgery

As far as motion tracking is considered, two general uses are encountered in computer assisted surgery literature. First, is the human body motion tracking and second is the

surgery tool tracking. Both applications approach another typical problem in medical scenarios, where a robotic system needs to compensate physiological movements of a target region related to the patient.

Many studies using inertial sensors have been performed. Depending on the type, number, and configuration of sensors used, some studies are limited to tracking two degrees of orientation in a plane, while others track 3-D orientation. Full 3-DOF orientation tracking is most commonly performed using nine-axis sensor modules containing three orthogonally mounted triads of angular rate sensors, accelerometers, and magnetometers. Foxlin et al. [19] describe two commercial nine-axis sensing systems designed for head tracking applications. Sensor fusion is performed using a complementary separate-bias Kalman filter. Drift correction is described as only being performed during stationary periods when it is assumed accelerometers are sensing only gravitational acceleration. Thus, the described algorithm requires that all motion stop in order to correct inertial drift errors.

Kraft [20] describes an "unscented," quaternion based Kalman filter for real-time estimation of rigid-body orientation using measurements of acceleration, angular velocity and magnetic field strength. The described filter approximates the Gaussian probability distribution using a set of sample points instead of linearizing nonlinear process model equations. Simulation results demonstrate the general validity of the described filter. Tests of the filter with real measurements are mentioned, but not shown or quantified. In [21], a quaternion-based UKF was used to estimate translation and rotation of an optically tracked surgical tool in cases of short line-of-sight occlusions. The filter showed a maximum error of 2.5 degrees for rotation and 2.36 mm for translation in cases of occlusion lasting for ten samples (20 Hz acquisition rate).

2.5 Multi Sensor Motion Tracking

Tracking systems that employ only one form of sensing all suffer inherent drawbacks. For example, purely inertial trackers suffer from drift, optical trackers require a clear line of sight, and magnetic trackers are affected by ferromagnetic and conductive materials in the environment [22]. To maintain more consistent performance throughout a working environment, across the frequency spectrum, and over a wide range of dynamics, researchers have sought to develop hybrid tracking systems.

A typical example from the field of outdoor navigation is the inertial sensors and GPS combination, with an Extended Kalman filter used as the sensor fusion algorithm. An inertial measurement unit can provide frequent measurements of accelerations and angular velocities of a moving object, while the GPS is used

to correct, from time to time, the position drift accumulated during integration of the inertial measurements [23][24].

Regarding sensor fusion in body motion tracking, [19] has developed a system that is primarily inertial, but aided by angular rate sensors, both [25] and [26] have pursued systems that are primarily optical, but aided by inertial sensors (for prediction), virtual reality researchers at the University of Tokyo have sought to improve the data rate of the Polhemus tracker by augmenting it with rate gyros [27], while researchers at the University of North Carolina have sought to improve the accuracy of the Ascension magnetic tracker by augmenting it with a passive image-based system that observes known fiducial landmarks in the real world [28].

Some efforts have been focused on frequency augmentation or marker occlusion in optical tracking systems. [21] has used optical/electromagnetic UKF fusion for localizing surgical tools. An extensive experimentation has been performed and 9 sample occlusion of 1, 2 and 3 markers are simulated (Both sensors sampling at 10Hz). Authors claim that the fusion drastically improves the estimation during marker occlusion, yielding a RMS error of less than 2.1 mm for translation and 3.6deg for rotation.

Let us focus on approaches that have combined visual tracking and IMU data. In [29], the system tracks a pen-like tool to which is attached an IMU and its tip is detected by four cameras placed on a curved line. The sample rates are 100 Hz for the IMU and 20 Hz for the cameras. The authors of [29] find that with this setup the position error is reduced compared to optical stereo tracking, but the quantity of estimation error is not mentioned. [30] presents a setup for medical applications, consisting of a Vicon optical tracking system with 6 cameras and an optical marker attached to an IMU which is fixed to the object being tracked. With an optical sample rate of 10 Hz and the IMU at 100 Hz, the system accuracy in estimation at 100Hz is found to be as good as the Vicon's, maintaining a RMS of error less than 1mm. however all calculations involving orientation are performed using rotation matrices. We prefer using quaternions for orientation representation which have the advantages:

- Nonsingular representation (compared to Euler angles)
- More compact (and computationally efficient) than matrices

The authors of [31] combine an optical tracking system with three markers with a sample rate of about 55 Hz with an IMU being sampled at 500 Hz which is attached to the optical markers. They find that pose estimation is possible even during short marker occlusions, as long as at least one marker is visible to the cameras. In case of not having any marker visible the estimation error is reported to be 29.1mm during 3 seconds of occlusion. In [32], one video camera is used which tracks the position of a

fiducial marker pattern which is attached to an IMU being sampled at 400 Hz. The camera sample rate is 5–10 Hz. Authors report that this system can track a moving object and compensate short-time marker occlusions, however the position estimation errors are in the order of centimeters which does not fulfill neurosurgery requirements.

Most of the mentioned approaches have used extended Kalman filter as the fusion algorithm and the estimation errors are mostly greater than the accuracy needed for neurosurgery procedures. None of the approaches known to us have implemented unscented Kalman filtering for Optical/IMU fusion. We aim to develop a pose estimation system that integrates optical and inertial data using the quaternion-based UKF. Therefore, our approach combines the benefits of the following key elements:

- Optical/Inertial sensor fusion
- Unscented Kalman filter
- Quaternion representation for orientation

The developed system must perform robust (able to manage occlusions) motion estimation at the higher of the implemented sensors' frequency, while satisfying the requirements for computer-assisted neurosurgical interventions.

Chapter 3: Methods

> *The concept of multisensor data fusion is hardly new. Humans and animals have evolved the capability to use multiple senses to improve their ability to survive. For example, it may not be possible to assess the quality of an edible substance based solely on the sense of vision or touch, but evaluation of edibility may be achieved using a combination of sight, touch, smell, and taste. Similarly, while one is unable to see around comers or through vegetation, the sense of hearing can provide advanced warning of impending dangers. Thus multisensory data fusion is naturally performed by animals and humans to achieve more accurate assessment of the surrounding environment and identification of threats, thereby improving their chances of survival. (D. Hall [33])*

The determination of a rigid body pose from various types of measurements is one of the basic problems of many motion control applications. Yet the solutions to this problem vary widely in terms of sensors, accuracy, stability and computational effectiveness. This chapter describes the hardware and methods that yield a robust pose tracking with the desired bandwidth. This approach combines the benefits of two different key ingredients, Sensor fusion and quaternion based Unscented Kalman filter.

3.1 Sensors

As the principal means of perceiving a physical event, sensors are the first part of a control system that needs to be studied. Several measurement types have been used for object tracking that vary in precision and bandwidth. Cameras, Accelerometers, rate gyroscopes, magnetic field sensors, infrared and ultrasound sensors are among the measurement types used in computer assisted surgery systems. Our approach implements optical tracking system and inertial measurement unit. Both of which are widely used in motion tracking due to their notable characteristics. These characteristics are discussed in the following section.

3.1.1 Optical tracking System

Practical optical tracking systems (OTS) can be separated into two basic categories. Pattern recognition systems sense an artificial pattern of lights and use this information to determine translation and/or orientation. Such systems may be "outside-in" when the sensors are fixed and the emitters are mobile or "inside-out" when sensors are mounted on mobile objects and the emitters are fixed. Image-based systems determine position by using multiple cameras to track mounted points on moving objects within a working volume. The tracked points may be marked actively or passively. The used tracking system in this project is from the latter type and tracks active markers (infrared light emitting diodes).

Optotrak Certus (OC) (Figure 3.1) is a popular tool used in industries, universities and research institutions around the world. OC obtains 3D position utilizing infrared light emitting diodes called optical markers that are sensed by the sensor mounted in a stand. The markers are available in two sizes: 16mm and 7mm. A position sensor consisting of three one-dimensional charge-coupled devices paired with three cylindrical lens cells are mounted on a 1.1 m long stabilized bar and calibrated by the manufacturer. The sensor captures the positions of the markers sequentially with a total sampling speed of 4600 Hz, and maximum frame rate of 400 Hz (marker flashes are time-multiplexed)stati. In addition, OC can track up to 512 markers and the size and weight of OC makes it easy to move between locations.

3.1.2 Inertial Measurement Unit

The term Inertial Measurement Unit (IMU) is widely used to refer to an electronic board containing three accelerometers and three gyroscopes and optionally three magnetometers. The accelerometers are placed such that their measuring axes are orthogonal to each other and measure inertial acceleration. Three gyroscopes are placed

in a similar orthogonal pattern, measuring rotational position in reference to an arbitrarily chosen coordinate system.

Thanks to the MEMS fabrication technology, low cost and small size inertial measurement units are commercially available in various accuracy levels and have been integrated into numerous systems. However, it should be noted that data from low-cost MEMS accelerometers cannot be double-integrated for an extended period of time to determine position, due to a quadratic growth of errors. Any errors in measurement, however small, are accumulated from point to point. This leads to 'drift', or an ever-increasing difference between where the system thinks it is located, and the actual location.

The IMU used in this project is Inertial Two (ATAVRSBIN2) from ATMEL (Atmel, San Jose, USA). It delivers a nine degree-of-freedom sensor platform combining an accelerometer (KXTF9-1026 from Kionix), compass (HMC5883L from Honeywell), gyroscope (IMU-3000 from InvenSense). The sensors provide various sampling ranges (max: 2000 °/s for Gyro and 8 g for accelerometer) and frequencies (max: 2100 Hz for Gyro and 800 Hz for accelerometer). However, the final sampling frequency depends on the performance of the microcontroller board and the quality of communications.

The IMU is mounted on a hardware platform (UC3-A3 Xplained) for evaluating an Atmel AT32UC3A3256 32-bit AVR microcontroller.

3.1.3 Sensor Module Configuration

For the object tracking experimentation, the configuration is a stationary Optical tracking system and a sensor module which is the object being tracked. The IMU and 4 optical markers are rigidly connected and fixed to the sensor module as it can be seen in Figure 3.2.

In order to produce 3-D pose estimations relative to an Earth-fixed reference frame, the filter uses input data from the accelerometers and gyros fixed to the sensor module with their own reference frame and the position and orientation data from the OTS.

Figure 3.1 Left: Inertial two module from ATMEL. Right: Optotrak Certus Optical tracking system

3.1.4 Measurement Coordinate systems

The measurement system comprises sensor groups for acceleration (accelerometer), angular rate (Gyro), 3D position (OTS) and 3D orientation (OTS). Measurements are spatial vectors with respect to sensors' local reference frames. Acceleration and angular rate are measured with respect to the reference frame of the IMU board. Both sensors (Accelerometer and Gyroscope) have their X, Y, and Z axis aligned, and a symbol is provided on the board to indicate their directional alignment. The optical tracking system, however, involves two reference frames. One is attached to the sensor module and is defined as the dynamic reference frame (DRF) of the rigid body that the 4 markers describe. Since the markers are rigidly fixed to the sensor module, the dynamic reference frame moves and rotates with the sensor module. The second reference frame called static reference frame (SRF) is stationary and fixed to the OTS cameras, therefore it can be considered as our world reference frame. The optical tracking system reports the position and orientation of the DRF with respect to the SRF. Figure 3.2 shows the involved reference frames. Naturally, in order to fuse multiple measurements they must be with respect to the same reference frames. This is achieved here in two steps.

First, by using an optical pointer (a pre-calibrated 4-marker object digitizing probe from Northern Digital Inc.) and measuring precise distances of the fixed markers and the IMU reference frame, the dynamic reference frame of the markers is redefined in a way

to coincide with the IMU reference frame. Of course, a more sophisticated way to achieve this would be to design a calibration algorithm and find a homogeneous transformation matrix that transforms the point from one reference point to another. For example refer to [34] and [35].

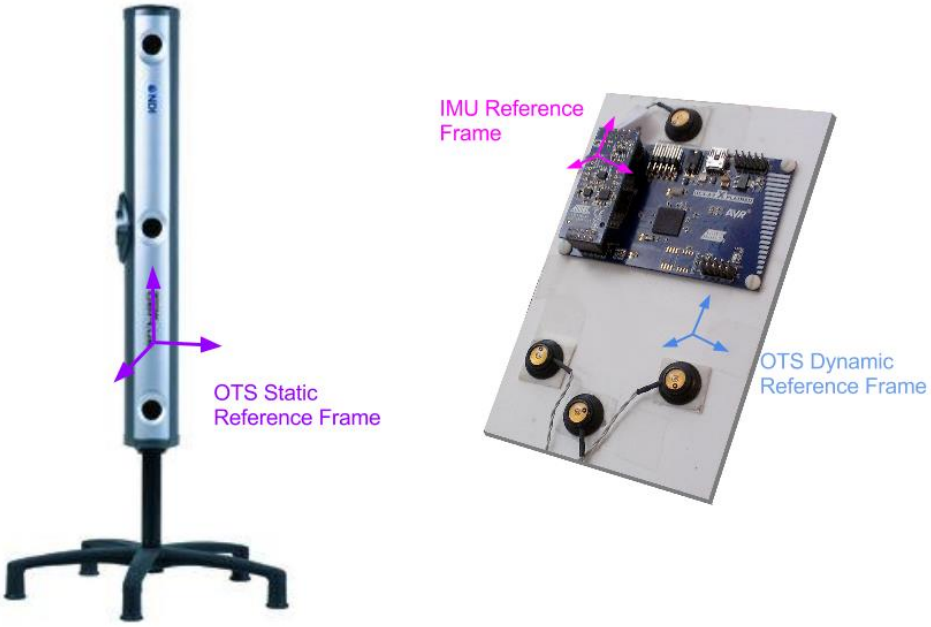

Figure 3.2 Measurement coordinate systems of the implemented sensors

As the second step, the new DRF of the sensor module needs to be rotated to match the direction of the SRF. This conversion is done in each sampling step by using the orientation data provided by the optical tracking system (Or as it will be explained later, the orientation estimation whenever OTS measurement is not available). At this point, all the 4 categories of the measurements are reported with respect to the same reference frame, i.e. Static reference frame, and the fusion can be performed.

3.2 Software

Using multiple sensors means processing data of different type with different time scales that are being communicated in different ways. In order to facilitate the data process, synchronization and communication it is highly beneficial to develop a single platform for sensor management.

For this project software packages for data acquisition from the sensors were developed in Robot Operating System (ROS) framework. ROS is a software framework for robot software development, providing operating system-like functionality on a heterogeneous computer cluster. ROS is released under the terms of the BSD (Berkeley Software Distribution license), and is open source software. It is free for commercial and research use. Available code libraries from the manufacturer made the development of the data acquisition package for the optical tracking system straight-forward. For the inertial measurement unit, however, a firmware had to be written for the microcontroller to make serial data transfer between ROS and the sensor possible. The fusion algorithm was developed in MATLAB, for the purpose of this project, and it was run off-line by having the recorded data acquisitions fed to it.

Regarding temporal issues, the time stamps of data registration by the sensor acquisition software is used and latencies of measurements are not studied in this work. Active synching is not performed and the data from both sensors are arranged off-line based on the ROS time stamps in a way to not have a time difference between the OTS and IMU samples that is greater than half a time interval (5ms for 200Hz).

3.3 Fusion Method

Combining information from multiple sensor types can be achieved with a data fusion algorithm. Kalman filter and its extensions are among the most popular methods of data fusion. Next section describes the reasons for this popularity.

3.3.1 Kalman Filter

There are four typical reasons why Kalman filters are employed in systems where a signal is to be estimated with a sequence of discrete measurements: (1) filtering; (2) data fusion; (3) prediction; and (4) calibration. Here a very brief description is presented. Further details can be found in Kalman filter texts such as [36][37][38].

3.3.1.1 Filtering

"To track" means by definition "to observe or monitor the course of" [39]. Observing the course of any event requires a sensing system to sense physical changes. In practice, all available means of sensing are susceptible to noise, e.g. electrical, mechanical, or

optical noise. Thus an ideal system would be able to somehow perfectly separate the true signal from the inherently noisy observations so that the target could be tracked perfectly.

There is another practical limitation imposed by using digital computers to perform tracking. Because these computers operate in discrete time steps, we are limited to discrete observations of the target. Thus some uncertainty is introduced in terms of the target dynamics in between our discrete observations.

Considering these imperfections, we would like some means to best estimate the target motion, i.e. we would like an optimal estimator.

> *"An optimal estimator is a computational algorithm that processes measurements to deduce a minimum error estimate of the state of a system by utilizing: knowledge of system and measurement dynamics, assumed statistics of system noises and measurement errors, and initial condition information." [37]*

In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discrete-data linear filtering problem [40]. The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means of using noisy measurements to estimate the state of a linear system, while minimizing the expected mean-squared estimation error. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and can do so even when the precise nature of the modeled system is unknown; it is inherently discrete and thus well suited to implementation on a digital computer; it can be extended to model systems with continuous dynamics; and (because it is recursive) it makes use of past information in an efficient manner.

3.3.1.2 Data Fusion

The Kalman filter assumes that the system being estimated has a measurement equation of the form given in equation (3.1) where the matrix $H(t_k)$ relates the state vector $\vec{x}$ to the measurement vector $\vec{z}$, and the vector $\vec{v}$ represents the measurement noise.

$$\vec{z}(t_k) = H(t_k)\,\vec{x}(t_k) + \vec{v}(t_k) \tag{3.1}$$

If the system being estimated has multiple forms of observation, there would be multiple corresponding measurement equations equation (3.1), i.e. multiple instances of the matrix $H(t_k)$, each representing a different relationship. By using the appropriate $H(t_k)$ for each type of measurement, the filter effectively combines, blends, or fuses the information contained in the heterogeneous measurements.

This capability for heterogeneous data fusion, combined with the properties discussed in section 3.3.1.1, has made the Kalman filter a very popular means of data fusion.

3.3.1.3 Prediction

The Kalman filter assumes that the system dynamics can be modeled as in equation (3.2) where matrix $A(t_k)$ relates the state at time step k to the state at step k+1 (in the absence of noise), and $\vec{w}(t_k)$ represents the driving or process noise.

$$\vec{x}(t_{k+1}) = A(t_k)\,\vec{x}(t_k) + \vec{w}(t_k) \tag{3.2}$$

In the classical discrete Kalman filter implementation, the time between filter steps k and k+1 is constant, and thus there is one instance of $A(t_k)$ that is used at each step of the filter to predict the state for that step as in equation (3.2). The best state estimate of the state at time step k+1 given measurements and corresponding estimates through time step k is $A(t_k)$ times the previous estimate.

$$\vec{x}(t_{k+1}|\,t_k) = A(t_k)\,\vec{x}(t_k|\,t_k) \tag{3.3}$$

For obvious reasons, equation (3.3) reflects what is normally called a one-step prediction. To predict more than one step into the future, one can then simply use equation (3.3) with an appropriate change to $A(t_k)$. For nonlinear relationships an extended Kalman filter can be used. For the EKF, $A(t_k)$ becomes the Jacobian of the nonlinear function, and the (nonlinear) predictions are realized by integration.

Even if another prediction scheme is used, the Kalman filter can often offer assistance by estimating states that one cannot directly measure. A common example is that of optimally estimating velocities when one only has measures of position. (Note that the alternative of directly computing derivatives from position measurements is inherently susceptible to noise). With velocities in hand one can better predict where the target will be than if the only information is where the target currently is.

3.3.1.4 Calibration

Knowledge about source and sensor imperfections can be used to improve the accuracy of tracking systems; thus efforts are often undertaken to measure components against a known standard in order to improve the accuracy of their estimated characteristics.

The Kalman filter is generally presented as a way of estimating values of stochastic variables (the states) of linear systems whose associated system parameters (e.g. model dynamics and noise characteristics) have known values. Interestingly enough, the filter

can just as well be turned around and used to estimate values of unknown system parameters when the states are known [38]. In fact, it can even be used to estimate both system states and parameters.

Unlike the first three cases of Kalman Filter applications mentioned earlier, Calibration is not implemented in this research and can be studied in future works.

3.3.2 EKF vs. UKF

Nonlinearities like the nonlinear relationship between estimated orientation and expected measurement in our tracking problem prevent the usage of a classical Kalman filter. It was mentioned earlier that the Extended Kalman Filter (EKF) is probably the most widely used estimator for nonlinear systems. The EKF applies the Kalman filter to nonlinear systems by simply linearizing all the nonlinear models so that the traditional linear Kalman filter equations can be applied. However, in practice, the use of the EKF has two well-known drawbacks:

1. If the assumptions of local linearity are violated, linearization can cause instability in filters

2. In many applications the derivation of the Jacobian matrices is nontrivial and may lead to significant implementation difficulties.

The Unscented Kalman Filter is a linear estimator which yields performance equivalent to the Kalman filter for linear systems, yet generalizes elegantly to nonlinear systems without the linearization steps required by the EKF. The fundamental component of this filter is the unscented transformation which uses a set of appropriately chosen weighted points to parameterize the means and covariance of probability distributions.

It has been argued [41] that the expected performance of the UKF is superior to that of the EKF and, is directly comparable to that of the second order Gauss filter. Further, the nature of the transform is such that the process and observation models can be treated as “black boxes”. It is not necessary to calculate Jacobians and so the algorithm has superior implementation properties to the EKF. The differences in performance in some example application are demonstrated in [42] [43].

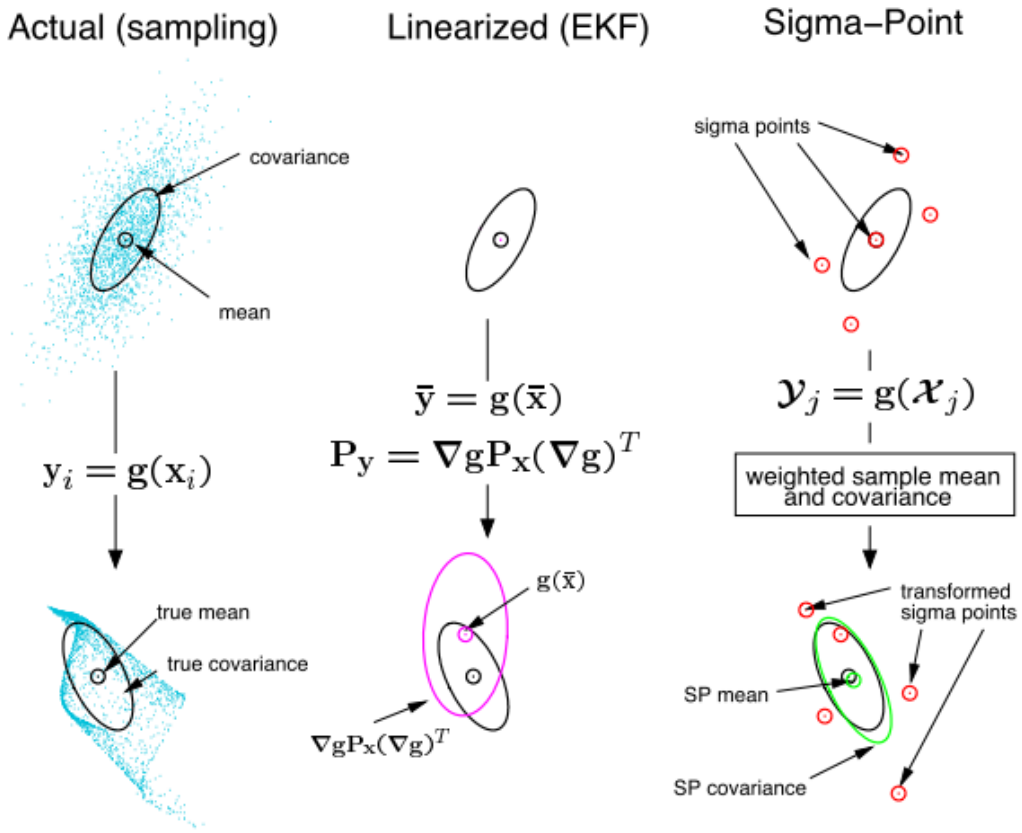

Figure 3.3 An evaluation of the accuracy of the scaled unscented transformation for mean and covariance propagation. From [43]

An example experiment by [43] performs a comparison between EKF and a modified unscented transform called scaled unscented transform (Section 3.4.4). In this experiment a two dimensional Gaussian random variable (GRV) is propagated through an arbitrary highly nonlinear transformation and then the optimally calculated first and second order posterior statistics of the transformed random variable is compared with those calculated through normal linearization (EKF), and by the scaled unscented transformation (SUT). The optimal statistics were calculated with a Monte Carlo approach using 100, 000 samples drawn from the prior distribution and then propagated through the full nonlinear mapping. The result of this experiment is shown in Figure 3.3. On the top-left the prior GRV (indicated by its mean, covariance-ellipse and representative sample distribution) is shown. The posterior sample distribution after going through the nonlinear transformation can be seen at the bottom-left. Its true mean and covariance are as well indicated. The result of propagating the GRV through a first-order linearized system is shown in the center plot. It can clearly be seen that the posterior mean estimate is biased and the posterior covariance estimate is highly inaccurate. The plot on the right shows the results of the estimates calculated by the

scaled unscented transformation. There is almost no bias error in the estimate of the mean and the estimated covariance is also much closer to the true covariance. The superior performance of the SUT approach is clearly evident. Note that solely 5 sigma-points (points sampling the probability distribution 3.4.1)) are used compared to the thousands needed by the pure Monte Carlo method on the left.

3.4 Unscented Kalman Filter

The Unscented Kalman filter (UKF) is an extension of the classical Kalman filter to nonlinear process and measurement models. The main difference to the well-known Extended Kalman Filter (EKF) is that the UKF approximates the Gaussian probability distribution by a set of sample points whereas the EKF linearizes the (nonlinear) model equations. This leads to results which are usually both more accurate (because the original equations are used) and less costly to compute (because no Jacobi matrices need to be calculated).

3.4.1 Filter Concept

The filter is based on Unscented Transformation which is established on the intuition that it is easier to approximate a Gaussian distribution than *it is to approximate an arbitrary nonlinear function or transformation[44]*. Figure 3.4 illustrates the main concept. A set of points called sigma points are selected so that their mean and covariance are $\bar{x}$ and P_{xx}. The nonlinear function is applied to each point in turn to yield a cloud of transformed points and $\bar{y}$ and P_{yy} are the statistics of the transformed points. Although this method bares a superficial resemblance to Monte Carlo-type methods, there is an extremely important and fundamental difference. The samples are not drawn at random but rather according to a specific, deterministic algorithm. Since the problems of statistical convergence are not an issue, high order information about the distribution can be captured using only a very small number of points.

The transformation processes which occur in an unscented Kalman filter consist of the following steps:

- Predict the new state of the system $\hat{x}_{k+1|k}$ and its associated covariance$P_{k+1|k}$. This prediction must take account of the effects of process noise.
- Predict the expected observation $\hat{z}_{k+1|k}$ and the innovation covariance $P^{vv}_{k+1|k}$. This prediction should include the effects of observation noise.
- Finally, predict the cross-correlation matrix $P^{xz}_{k+1|k}$.

These steps can be easily accommodated by slightly restructuring the state vector, process and observation models. First, the state vector needs to be described.

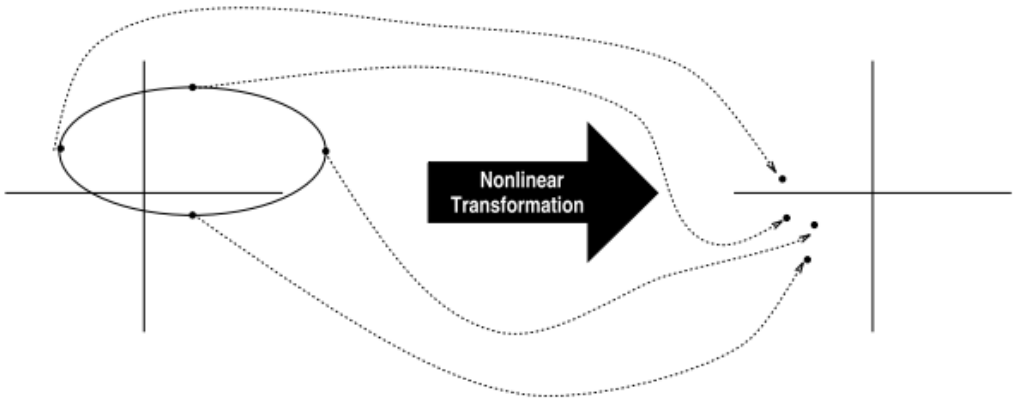

Figure 3.4 The principle of the unscented transform. A set of points (sigma points) are carefully chosen (Left) are transformed by the nonlinear function to yield a cloud of transformed points (Right). From [41]

3.4.2 State Vector

The output of the measurement system is a set of spatial vectors, measured with respect to the static reference frame. Given this data, the filter computes an estimate of the system state vector $\vec{s}$:

$$\vec{s} = [\ \vec{p}\quad \vec{v}\quad \vec{a}\quad q\quad \vec{\omega}\ ] \tag{3.4}$$

Where $\vec{p}, \vec{v}$ and $\vec{a}$ represent 3D vectors of position $\vec{p}(\ p_x\ p_y\ p_z\ )$, velocity $\vec{v}(\ v_x\ v_y\ v_z\ )$and acceleration $\vec{a}(\ a_x\ a_y\ a_z\ )$ along the static reference frame. $q(\ q_0\ q_1\ q_2\ q_3\ )$ is the orientation quaternion and $\vec{\omega}(\ \omega_x\ \omega_y\ \omega_z\ )$ angular rate. The system state vector therefore has 16 components.

The body's orientation is represented by a quaternion , which is a number with four real components. Let i^2=j^2=k^2= -1 then:

$$q \equiv\ q_0 + iq_1 + jq_2 + kq_3 \qquad , q_0, q_1, q_2, q_3\ \in R \tag{3.5}$$

Quaternions express rotation as a rotation angle about a rotation axis. This is a more natural way to perceive rotation than Euler angles. The representation of rotation using quaternions is compact in the sense that it is four dimensional and thereby only contains the four degrees of freedom required according to Euler's theorem. In practical

applications only unit quaternions are used for rotation. Thus, only one constraint on the representation must be upheld during computation compared to the six constraints on rotation matrices.

Unit quaternions are defined as:

$$\|q\| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2} \tag{3.6}$$

$$q \in H, \qquad if\ \|q\| = 1 \quad then \quad q \in H_1 \tag{3.7}$$

H is the set of quaternions, and H_1 the set of unit quaternions. One degree of freedom of the quaternion is lost by applying the constraint on the size. The remaining three degrees of freedom are sufficient to provide a representation for any spatial rotation. However, a side effect of the constraint is that the first four components of the state vector are no longer independent of each other. This causes a conflict with the concept underlying a Kalman filter and the way noise is treated that will be addressed when we define the process and measurement noise.

Another important advantage of using quaternions is that they offer a singularity-free description (as opposed to Euler angles), therefore Gimbal lock does not happen in quaternion representation. Gimbal lock is the loss of one degree of freedom in a three-dimensional space that occurs when the axes of two of the three gimbals are driven into a parallel configuration, "locking" the system into rotation in a degenerate two-dimensional space. Since gimbal lock is innate to the matrix representation of Euler angles, this problem does not appear in the quaternion representation.

3.4.3 Process model

Process model is used to predict the evolution of the state vector. It also describes the influence of the process noise. In the generalization of the classical Kalman filter, the process model equation is given by:

$$\vec{s}(t_{k+1}) = A(\vec{s}(t_k), \vec{w}(t_k)) \tag{3.8}$$

A is an arbitrary function of $\vec{s}(t_k)$ and $\vec{w}(t_k)$. The process noise $\vec{w}(t_k)$ does neither have to be additive, nor does it has to have the same dimension as the state vector. Equation (3.8) is the basis of the process models of the filter described here. Different mathematics governs position and orientation state variables, therefore different models are defined for each.

3.4.3.1 Translation Process Model

A second order equation of motion is employed as the discrete process model for position estimation:

$$p_{k+1} = p_k + v_k.dt + \frac{1}{2}a_k.dt^2 + \vec{w}_p \tag{3.9}$$

$$v_{k+1} = v_k + a_k.dt + \vec{w}_v \tag{3.10}$$

$$a_{k+1} = a_k + \vec{w}_a \tag{3.11}$$

Where vectors $\vec{w}_p, \vec{w}_v$ and $\vec{w}_a$represent process noise and dt is the length of the time interval.

3.4.3.2 Orientation Process Model

For orientation, however, the process model is not as simple as that of translation. This is due to the non-additive nature of rotations expressed in quaternions. Rotation by q_1 followed by rotation by q_2 is equivalent to rotation by $q_2.q_1$ Where (.) represents quaternion multiplication. For a thorough description of quaternion algebra refer to [45][45]. A simple first order model, i.e. constant angular velocity is chosen for orientation estimation:

$$\vec{\omega}_{k+1} = \vec{\omega}_k \tag{3.12}$$

Experimental results prove this choice to be sufficient for a robust prediction. Given $\vec{\omega}_k$ of the previous state estimate and the length dt of the time interval, the differential rotation during this interval is defined with an angle α_d and an axis $\vec{e}_d$:

$$\alpha_d = |\vec{\omega}_k|.dt \tag{3.13}$$

$$\vec{e}_d = \frac{\vec{\omega}_k}{|\vec{\omega}_k|} \tag{3.14}$$

$|\vec{\omega}_k|$ is the total angular velocity in rad/s. The corresponding quaternion is given by:

$$q_d = \left[\cos\frac{\alpha_d}{2} \ , \vec{e}_d \sin\frac{\alpha_d}{2}\right] \tag{3.15}$$

Dividing the quaternion into a scalar part q_0 and a vectorial part(q_1, q_2, q_3), illustrates the similarity between a quaternion and the angle/axis notation of the orientation. The new orientation q_{k+1} is calculated from the previous step quaternion-multiplied by q_d:

$$q_{k+1} = q_k q_d \tag{3.16}$$

The resulting quaternion q_{k+1} is equivalent to the rotation q_d followed by the rotation q_k .The process model for the undisturbed state vector of orientation is defined by Equations (3.12) through (3.16).

3.4.3.3 Process Noise Covariance

Even though the state vector has 16 elements, it has only 15 degrees of freedom. Therefore we choose to describe the influence of the process noise with a 15-dimensional noise vector w_k:

$$w_k = \left(\vec{w}_p \;\; \vec{w}_v \;\; \vec{w}_a \;\; \vec{w}_q \;\; \vec{w}_\omega\right) \tag{3.17}$$

Despite the values of w_k change with every time step, the subscript is omitted in the components for matters of readability. The rate at which the uncertainty of the system state estimate increases with time is represented by the covariance Q of the random variable w_k. Apart from $\vec{w}_q$, the rest of the vector can be assumed to be additive. Based on the method described in [46] the components of matrix Q can be calculated considering the order of the process model with:

$$q_{i,j} = \sigma_w^2 \frac{T^{2m+3-(i+j)}}{(m+1-i)!\,(m+1-j)!\,(2m+3-(i+j))} \tag{3.18}$$

Where T is the time interval and m is the order of the process model. The value of $\sigma_w$is found experimentally.

$\vec{w}_q$ represents an increase of the uncertainty in the orientation (degrees per time interval). Since $\vec{w}_q$ is a three dimensional noise vector, it cannot simply be added to the four component quaternion. It has to be converted into a unit quaternion. Let the random variable $\vec{w}_q$follow a normal distribution with covariance Q_w(a 3*3-matrix) and mean 0. The vector $\vec{w}_q$ can be regarded as a rotation vector. This means that it represents a random rotation with the angle and axis equal to:

$$\alpha_w = |\vec{w}_k| \tag{3.19}$$

$$\vec{e}_w = \frac{\vec{w}_k}{|\vec{w}_k|} \tag{3.20}$$

Therefore the quaternion representation q_w of this rotation is:

$$q_w = \left[\cos\frac{\alpha_d}{2} \;\; , \vec{e}_d \sin\frac{\alpha_d}{2}\right] \tag{3.21}$$

It is possible to add the orientation process noise before the process model. In that case the final process model for orientation estimation is:

$$q_{k+1} = q_k q_w q_d \tag{3.22}$$

$$\vec{\omega}_{k+1} = \vec{\omega}_k + \vec{w}_\omega \tag{3.23}$$

Note that $\vec{w}_\omega$ is assumed to be additive as mentioned before.

3.4.3.4 Measurement Model

The measurement model relates the measurement value to the value of the state vector and describes the influence of a random variable v (measurement noise) on the measured value.

$$\vec{z}(t_{k+1}) = H(\vec{s}(t_k), \vec{w}(t_k)) \tag{3.24}$$

Since the tracking system produces three different types of measurements, three measurement models ($H_{ots}, H_{gyr}, H_{acc}$) are defined. Position and orientation from the Optical tracking system are part of the state vector, leading to the simplest model:

$$H_{ots}: \quad \vec{z}_{p\_ots} = \vec{p}_k + \vec{v}_{ots} \tag{3.25}$$

Angular rate measurements from the gyros need to be transformed to the static reference frame:

$$H_{gyr}: \quad \vec{z}_{gyr} = T_k^{drf/srf}(\vec{\omega}_k + \vec{v}_{gyr}) \tag{3.26}$$

Where $T_k^{drf/srf}(x)$ describes the conversion of vector x from the static reference frame to the dynamic reference frame using the following quaternion preposition [45]:

$$\begin{aligned} &\textit{Let } q \in H_1, \ \ q = [cos\theta, sin\theta n]. \ \ \textit{Let } r = (x, y, z) \in \mathbb{R}^3 \textit{ and } p = [0, r] \in H \\ &\textit{Then } p' = pqp^{-1} \textit{ is } p \textit{ rotated } 2\theta \textit{ about the axis } n. \end{aligned} \tag{3.27}$$

In which q is the orientation quaternion measured by the optical tracking system. Whenever q from OTS is not available (due to marker occlusion or lower sampling rate compared to that of IMU) the orientation quaternion from the estimated state vector is used. Doing so of course introduces an error due to the difference between the estimated orientation of the sensor module and its real orientation. This difference is corrected as soon as data from the OTS is available.

The acceleration measurement function performs one further step that is removing the vector of gravitational field from the measurements. Linear accelerometers measure the vector sum of acceleration $\vec{a}_k$ and gravitational $\vec{g}$ in sensor coordinates. To remove the

component of acceleration due to gravity, the attitude of the accelerometer with respect to vertical must be known. The vector of the gravitational field is constant in the static reference frame, Therefore, by converting the acceleration measurements from DRF to SRF, gravity can be subtracted from the measured accelerations. Thus the measurement function for acceleration becomes:

$$H_{gyr}: \quad \vec{z}_{acc} = T_k^{drf/srf}(\vec{a}_k + \vec{g} + \vec{v}_{acc}) \tag{3.28}$$

Where again $T_k^{drf/srf}(x)$ describes the conversion of vector x from the static reference frame to the dynamic reference frame as described in (**3.27**).

3.4.4 Sigma Points

In the beginning of every UKF recursion the previous estimates of the state vector $\hat{x}_{k-1}$ and its covariance P_{k-1} are known. A set of points (or sigma points) are chosen so that their sample mean and sample covariance are $\hat{x}_{k-1}$ and P_{k-1}.

In the Unscented transform, for an n-dimensional random variable x with mean $\bar{x}$ and covariance P_{xx} , 2n+1 sigma points and their corresponding weights are given by:

$$\begin{aligned} \mathcal{X}_0 &= \bar{x} & \boldsymbol{W_0} &= \boldsymbol{\kappa/(n+\kappa)} \\ \mathcal{X}_i &= \bar{x} + (S)_i & \boldsymbol{W_i} &= \boldsymbol{1/2(n+\kappa)} \\ \mathcal{X}_{i+n} &= \bar{x} - (S)_i & \boldsymbol{W_{i+n}} &= \boldsymbol{\kappa/(n+\kappa)} \end{aligned} \tag{3.29}$$

Where $(S)_i$ is the ith row or column of the matrix S computed as:

$$S = \sqrt{(n+\kappa)P_{xx}} \tag{3.30}$$

Where $\kappa \in \mathbb{R}$ is an arbitrary weighting factor. The effect of κ on the filter performance is summarized in [41]. Since P_{xx} is a covariance matrix, it happens to be symmetric and positive definite. This allows us to use a Cholesky Decomposition to compute S.

The sigma-point selection scheme used in the unscented transform (UT) has the property that as the dimension of the state-space (n) increases, the radius of the sphere that bounds all the sigma points increases as well. Even though the mean and covariance of the prior distribution are still captured correctly, it does so at the cost of possibly sampling non-local effects. If the nonlinearities in question are very severe, this can lead to significant difficulties. In order to address this problem, the sigma points can be scaled towards or away from the mean of the prior distribution by a proper choice of κ. The scaled unscented transformation (SUT) was developed to

address this problem [47]. Again, for an n-dimensional random variable x with mean $\bar{x}$ and covariance P_{xx}, 2n+1 "scaled sigma points" and their corresponding weights are given by:

$$\begin{aligned} \mathcal{X}_0 &= \bar{x} & \boldsymbol{W_0^m} &= \frac{\boldsymbol{\lambda}}{(\boldsymbol{n}+\boldsymbol{\lambda})} \\ \mathcal{X}_i &= \bar{x} + (S)_i & \boldsymbol{W_0^c} &= \frac{\boldsymbol{\lambda}}{(\boldsymbol{n}+\boldsymbol{\lambda})} + (\mathbf{1} - \boldsymbol{\alpha^2} + \boldsymbol{\beta}) \\ \mathcal{X}_{i+n} &= \bar{x} - (S)_i & \boldsymbol{W_i^m} &= \boldsymbol{W_i^c} = \frac{\mathbf{1}}{\mathbf{2}(\boldsymbol{n}+\boldsymbol{\lambda})} \end{aligned} \tag{3.31}$$

Where $\boldsymbol{\lambda}$ is equal to**:**

$$\lambda = \alpha^2(n+\kappa) - n \tag{3.32}$$

Not much research effort has yet been spent on determining if there is a global optimal setting of the SUT scaling parameters (α, β and κ) but some guidelines can be found in [43].

3.4.4.1 Quaternion Sigma Points

In our specific case (unit quaternion in state vector) we already know a simple addition of the random variable to the state vector is not possible because the dimensions do not match. Instead, the sigma points have to be calculated in the way described for the process noise, similar to equation (3.22):

$$\mathcal{X}_i = \begin{pmatrix} q_{k-1} q_S \\ \vec{\omega}_{k-1} + \vec{\omega}_S \end{pmatrix} \tag{3.33}$$

q_S is the quaternion corresponding to three components of S that are related to orientation, $\vec{\omega}_S$ denotes the angular velocity vector built from the respective three components of S, q_{k-1} and $\overrightarrow{\omega}_{k-1}$ are from the previous estate estimation.

As it was mentioned before we have treated the noise according to the number of degrees of freedom of the state vector and not according to the number of elements. This of course affects the dimensionality of the vectors and covariances involved. Vectors of S are 15-dimensional, because they have the same dimension as the process noise vector. Their covariance is hence a [15,15] matrix. On the other hand, the sigma points are state vectors with a quaternion component and are therefore 16-dimensional. The transition between these two sets is performed by (3.33) and is based on the conversion of a rotation vector representation to a quaternion representation.

3.4.5 A posteriori estimate

After generating the sigma points, each point is instantiated through the process model to project them ahead in time resulting in the set of transformed sigma points $\mathcal{Y}_i$:

$$\mathcal{Y}_i = A(\mathcal{X}_i, 0) \tag{3.34}$$

Note that since the influence of the process noise is already included in the sigma points no additional noise vector is being considered in the equation above.

The set $\mathcal{Y}_i$ samples the probability distribution of the *a priori* estimate. $\hat{x}_{k+1}^-$ is defined as the mean value of this distribution $\overline{y}$, and P_{k+1}^- is equal to the covariance P_y. The mean is given by the weighted average of the transformed points:

$$\overline{y} = \sum_{i=0}^{2n} W_i^m \mathcal{Y}_i \tag{3.35}$$

Where W_i^m are the defined weights in (3.31). The covariance P_y is the weighted outer product of the transformed points:

$$P_y = \sum_{i=0}^{2n} W_i^c \, (\mathcal{Y}_i - \overline{y})(\mathcal{Y}_i - \overline{y})^T \tag{3.36}$$

Usually the a priori estimate is computed when a new measurement occurs. A measurement update step requires both the a priori estimate and an estimate z_k^- of the measurement. The set $\mathcal{Y}_i$ is thus transformed further by the measurement model H, resulting in a set $\mathcal{Z}_i$ of projected measurement vectors.

$$\mathcal{Z}_i = H(\mathcal{X}_i, 0) \tag{3.37}$$

Again since the noise is included before there is no need to consider it here. z_k^- is equal to $\bar{z}$ the weighted mean value of $\mathcal{Z}_i$, and the measurement prediction covariance is equal to the covariance P_{zz} :

$$\bar{z} = \sum_{i=0}^{2n} W_i^m \mathcal{Z}_i \tag{3.38}$$

$$P_{zz} = \sum_{i=0}^{2n} W_i^c (\mathcal{Z}_i - \bar{z})(\mathcal{Z}_i - \bar{z})^T \tag{3.39}$$

Note that the UKF uses the mean value of the projected distribution to compute the measurement estimation $\hat{z}_{k+1}^{-}$. Instead In a classical Kalman filter, the measurement estimation is simply equal to $H(\hat{x}_{k+1}^{-})$. Both would be identical if function H is linear.

Innovation ν_k is defined as the difference between the actual measurement z_k and its predicted value z_k^{-}:

$$\nu_k = z_k - z_k^{-} \tag{3.40}$$

Its expected covariance $P_{\nu\nu}$ is the sum of the projected state vector covariance (the uncertainty in the measurement caused by the uncertainty in the state vector prediction) P_{zz} and the measurement noise covariance R (the additional uncertainty induced by the measurement process).

$$P_{\nu\nu} = P_{zz} + R \tag{3.41}$$

The a posteriori estimate $\hat{x}_k$ is finally computed by adding the a priori estimate to the innovation multiplied by the Kalman gain K_k:

$$\hat{x}_k = \hat{x}_k^{-} + K_k \nu_k \tag{3.42}$$

3.4.6 Computing the Mean

In the original Kalman filter (and Unscented Kalman Filter), the state vector is an element of the vector space. The mean value of a vector is equal to its barycentric mean which is simply the sum over all elements of the set divided by the number of elements:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{3.43}$$

The mean value of the state and measurement vector elements can be computed by barycentric mean, except for the orientation components.

Orientations are members of a homogenous Riemannian manifold (the four dimensional unit sphere) but not of a vector space. Their mean value cannot be calculated by a simple barycentric mean because orientations are periodic. Using equation (**3.43**) may not yield correct results as it can be seen from an example with two rotations of -170° and 180° around an axis. The barycentric mean is 5°, while the expected result is -5°. On the other hand, the mean of two unit quaternions has to be a unit quaternion, too. The barycentric mean of the quaternions {(0, 0, 0, 1), (0, 0, 0, 1)} does not obviously fulfill this requirement.

The intrinsic gradient descent algorithm described in [48] can be implemented to compute the mean of quaternion values. The key ingredient is the definition of a new metric which describes the distance between two elements. For orientations, we use θ the angle of the rotation which turns one orientation into the other. Given two quaternions q_1 and q_2, the rotation q_{12} which fulfills:

$$q_2 = q_{12}\, q_1 \tag{3.44}$$

Is given by:

$$q_{12} = q_2\, {q_1}^{-1} \tag{3.45}$$

Rotation angle θ can be computed form the scalar part of q_{12}, see (**3.15**).

$$\theta = 2.\arccos(q_{12}^0) \tag{3.46}$$

In this method the mean orientation $\bar{q}$ is estimated by iteration starting with an arbitrary orientation. Iteration steps are denoted by the index t. In each step, the so called *error vectors* are computed for every set element. An error vector $\vec{e}_i$ is the rotation vector corresponding to the relative rotation between the set element q_i and the estimated mean of the last iteration $\bar{q}_t$. The quaternion representation e_i of $\vec{e}_i$is:

$$e_i = q_i {\bar{q}_t}^{-1} \tag{3.47}$$

So that:

$$q_i = e_i \bar{q}_t \tag{3.48}$$

Which means $e_i$rotates the mean into the set element q_i. The barycentric mean $\vec{e}_i$ of all error vectors:

$$\vec{e} = \frac{1}{n}\sum_{i=1}^{n} \vec{e}_i \tag{3.49}$$

is a measure of the deviation between the estimated mean and the real mean orientation. $\vec{e}$ is a rotation vector that points in the direction of the real mean. The corresponding quaternion *e* can hence be used to calculate a better estimate for the next iteration step:

$$\bar{q}_{t+1} = e\bar{q}_t \tag{3.50}$$

Since $\vec{e}$ is used to adjust the estimated mean from iteration to iteration it is called the adjustment vector. Its value is zero if $\bar{q}_t$ is equal to the real mean of the set of

orientations. The size of the adjustment vector can consequently be used to stop the iteration once a satisfactory precision is achieved.

The starting value of the iterations is arbitrary, but naturally it determines the number of iterations needed to reach the desired precision. It is therefore feasible to use the previous state vector estimate $\hat{x}_{k-1}$and use its quaternion as a start value.

3.4.7 Algorithm Summary

1. Sigma point generation:

$$\mathcal{X}_{k-1} = [\hat{x}_{k-1}^{-} \quad \hat{x}_{k-1}^{-} + \gamma\sqrt{P_{k-1}} \quad \hat{x}_{k-1}^{-} - \gamma\sqrt{P_{k-1}}] \tag{3.51}$$

   (Where γ represents the scaling factors.)
2. Time-Update equation:

$$\mathcal{Y}_{k|k-1} = A(\mathcal{X}_{k-1}) \tag{3.52}$$

$$\hat{x}_{k}^{-} = \sum_{i=0}^{2n} W_i^m \mathcal{Y}_i \tag{3.53}$$

$$P_{y_k}^{-} = \sum_{i=0}^{2n} W_i^c (\mathcal{Y}_{i,k|k-1} - \hat{x}_{k}^{-})(\mathcal{Y}_{i,k|k-1} - \hat{x}_{k}^{-})^T \tag{3.54}$$

3. Measurement-update equation:

$$\mathcal{Z}_{k|k-1} = H(\mathcal{X}_{k-1}) \tag{3.55}$$

$$\hat{z}_{k}^{-} = \sum_{i=0}^{2n} W_i^m \mathcal{Z}_{k|k-1} \tag{3.56}$$

$$P_{z_k} = \sum_{i=0}^{2n} W_i^c (\mathcal{Z}_{k|k-1} - \hat{z}_{k}^{-})(\mathcal{Z}_{k|k-1} - \hat{z}_{k}^{-})^T \tag{3.57}$$

$$P_{x_k z_k} = \sum_{i=0}^{2n} W_i^c (\mathcal{Y}_{i,k|k-1} - \hat{x}_{k}^{-})(\mathcal{Z}_{k|k-1} - \hat{z}_{k}^{-})^T \tag{3.58}$$

$$K_k = P_{x_k z_k} P_{z_k}^{-1} \tag{3.59}$$

4. And finally a posteriori estimates:

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - \hat{z}_k^-) \tag{3.60}$$

$$P_{y_k} = P_{y_k}^- - K_k\, P_{z_k} K_k^T \tag{3.61}$$

3.4.8 Temporal Concerns

According to Nyquist sampling theorem [38] the measurement or sampling frequency should be more than twice the bandwidth of the target's motion, or an estimator may track an alias of the true motion. Nevertheless, the sampling theorem provides a sufficient condition, but not a necessary one, for perfect reconstruction a higher bandwidth is necessary. Given that common arm and head motion bandwidth specifications range from 2 to 20 Hz [19][49], the sampling rate should ideally be greater than 40 Hz. Furthermore, the estimation rate should be as high as possible so that slight (expected and acceptable) estimation error can be discriminated from the unusual error that might be observed during periods of significant target dynamics.

On the other hand tool tracking may be used for tool servo-controlling. Servo-control demands for a fast tracking system which has to be faster than human reaction in order to correct small errors the surgeon might make and compensate for unintended target displacements during the surgery. A bandwidth of 200 Hz seems to satisfy all the requirements needed for a fast object tracking system.

For initialization the first position (and orientation) measurements from the OTS is used and velocities (angular rates) and accelerations are assumed to be zero. The state covariance initial value is a diagonal matrix with the OTS measurement noise on the diagonal.

3.5 Experiment Setup

The performance of the developed fusion algorithm was first simulated by numerically generated trajectories. After confirming that the system is capable of robustly tracking generated trajectories, it was put to test by acquiring real data from the sensors.

The experimental setup was the following:

- An optical tracking system (OTS)
- A sensor module (the board to which the inertial measurement unit and 4 active optical markers were fixed)

- A computer running Linux Ubuntu operating system, in charge of data storage, implementing the packages that we developed in C++ for Robotic Operating System (ROS).

The data was planned to be acquired at frequency of 200 Hz from both sensors, via Ethernet communication from the OTS and via serial communication from the IMU. However due to hardware problems the OTS could not go above 100Hz. Since the data was used off-line, a spline interpolation was performed on OTS data to reach 200 Hz. Doing so, of course, introduces uncertainty into samples of OTS that is used as a ground truth for error computation. To prevent this uncertainty from being taken into account, only the real values of OTS was used for error computation.

In motion estimation, some researchers prefer to move the tracked object by mechanical systems. This would facilitate maintaining a systematic approach by generating repeatable trajectories. However, we decided to move the sensor module by hand in space, replicating the exact case scenario of a hand-held tool. This will include the higher frequency movements introduced by hand shake into the measurement signals.

The performed experiments fall into two categories. First is the frequency augmentation test in which the OTS data are fed to the filter with a lower bandwidth than that of the IMU. The estimation is performed with the higher frequency and the estimation error is computed. In the second category marker occlusion is tested and the performance of the filter in case of temporary loss of optical tracking data is evaluated.

3.5.1 Frequency Augmentation

Existing optical tracking systems used in computer-assisted surgery have a low bandwidth of 10–60 Hz. In our experiment we consider the frame frequency of the OTS to be 20 Hz. Sampling frequency of the IMU is selected to be the maximum that it can provide, which is 200 Hz for both accelerometer and rate gyros. We define frequency augmentation ratio (FAR) as the ratio of sampling frequency of IMU to that of OTS. Having a FAR of 10 means that for each 10 samples from the IMU, 1 sample is acquired from the OTS that "corrects" the drift caused by estimation from the IMU. While, the state vector is estimated at each time instance, we will call the estimated samples when OTS measurement is available "Correction" and the rest that use only IMU measurements "Estimation". Thus, after each correction 9 steps of estimation are performed till the next correction (Figure 3.5)

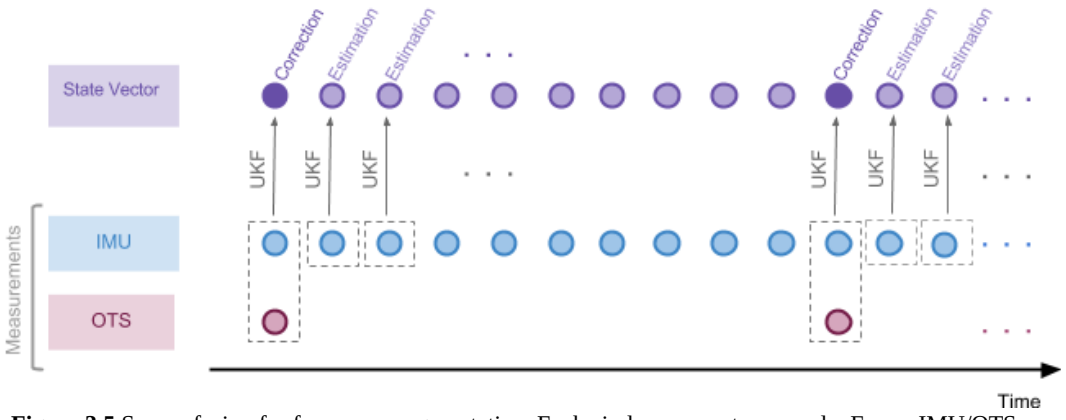

Figure 3.5 Sensor fusion for frequency augmentation. Each circle represents a sample. For an IMU/OTS FAR of 10, 9 samples are estimated per each correction. The estimated state vector has the same frequency as the faster sensor.

As it was mentioned before, the OTS measurements (which is claimed to have an accuracy of 0.1mm by the manufacturer) are considered as the ground truth for error computation. Thus, OTS measurement is assumed to be the "real" pose. A down-sampled 20 Hz signal of the acquired OTS data and a 200 Hz IMU measurement are fed to the fusion system that performs pose estimation at 200 Hz. The architecture of the frequency augmentation test is shown in Figure 3.6

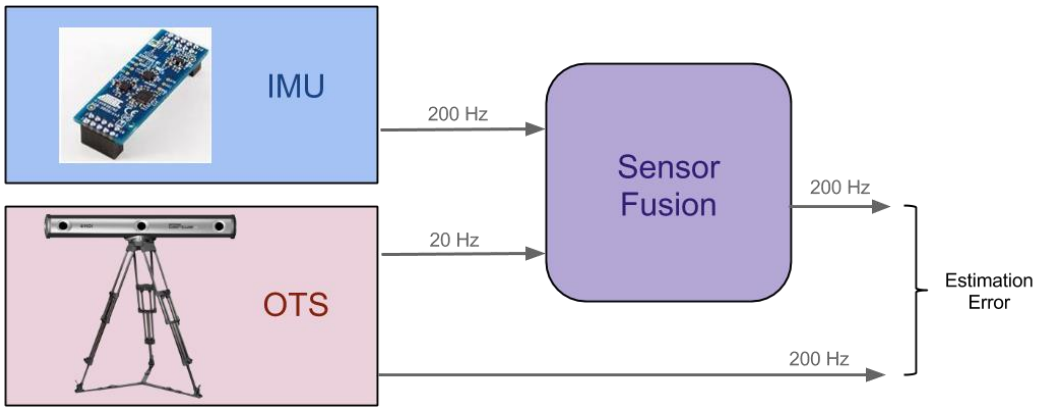

Figure 3.6 Frequency augmentation test. Capability of the sensor fusion in frequency augmentation is tested by comparing the estimated pose from the sensor fusion system with measurements of OTS.

3.5.1.1 Simulation

As the first evaluation step the fusion system was tested by simulated trajectories. 6 random 3d trajectories with average and maximum linear velocities of 150mm/s and 300mm/s are generated. The position trajectory is numerically double-differentiated using central difference approximation to provide acceleration data for the filter. A Gaussian noise with max amplitude of 100mm/s^2 is added to the acceleration to

simulate the IMU measurement noise. Orientation trajectories were periodic rotations around one axis and the gyro measurements were simulated by again numerically differentiating the generated orientation. The orientation trajectories are then converted to quaternions before feeding it to the sensor fusion algorithm. Assigned frequencies are the same as the ones in real data acquisitions (20Hz for OTS and 200Hz for IMU). The predicted pose is compared with the generated path to compute the estimation error.

3.5.1.2 Experiments

In object tracking systems the estimation error depends highly on dynamics of the motion. The higher the acceleration and velocity of the object the greater will be the estimation error. [49] reports the maximum velocity of patient's head movement during awake brain surgery to be 60 mm/s. In case of surgical tool tracking the velocity may be higher. To cover for both cases, we have performed the experiments in two categories. A category of acquisitions having linear velocity with average of 40mm/s and maximum of 100mm/s and second category having linear velocity with average of 100mm/s and maximum of 330mm/s. Orientation experiments are performed in one set of 12 acquisitions with an average angular rate of 20°/s and maximum of 50°/s.

Next the real data acquisition is done by moving the sensor module by hand in square shaped trajectories with a side length of approximately 20 cm for 10 seconds. 12 data acquisitions were performed for each velocity category, thus in total 24 acquisitions were performed for position estimation.

In frequency augmentation high-frequency inertial measurement is fused with low-frequency Optical tracking system measurements. As it was described in Section 3.5.1, having a frequency augmentation ratio (FAR) of 10 means for each OTS correction step, 9 estimation steps are performed using only IMU data. The estimation error is measured at each step by comparing the estimated pose with the real pose (OTS measurement). For example Figure 3.7 shows the estimation error of a 5-second simulated position estimation.

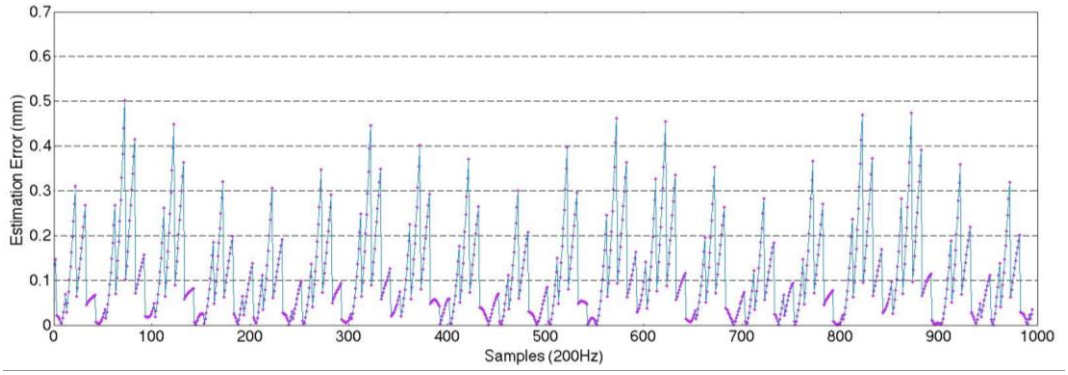

Figure 3.7 Estimation error of all samples during a 5 second position estimation.

Calculating the root mean square, percentiles and other statistical representations of the total error signal shown in Figure 3.7 provides some information about the performance of the filter. However, we would like to improve the representation by one further step.

Since we expect the error to generally increase step by step (due to the drift accumulation) until the correction is performed, a better way to represent the estimation error is to consider the error of each step separately. In other words, after calculating the estimation error for all steps they are categorized in 10 populations called: 1^{st}, 2^{nd},3^{rd},…,10^{th} step estimation errors, depending on how many estimation steps away they are from the last OTS correction. Figure 3.8 describes how the estimation errors of for example 8^{th} and 9^{th} estimation steps are clustered. This classification yields 10 statistical populations that exhibit the behavior of the estimation error in a more informative fashion. The errors do not have a Gaussian distribution. To represent the statistical dispersion, the root mean square error (RMSE) and a 95% confidence interval (CI) are computed from each population and graphically represented as the performance of the fusion algorithm. Equation (**3.62**) shows how the RMSE of n^{th} population for 12 data sets is calculated:

$$RMSE_n = \sqrt{\frac{1}{12N}\sum_{i=1}^{12}\sum_{k=0}^{N}\left(\hat{x}_{10k+n}^{i} - z_{10k+n}^{i}\right)^2} \qquad (3.62)$$

$$n = 1, 2, \dots, 10 \qquad N = \left(\frac{s}{10} - 10\right)$$

Where s represents number of samples in each acquisition.

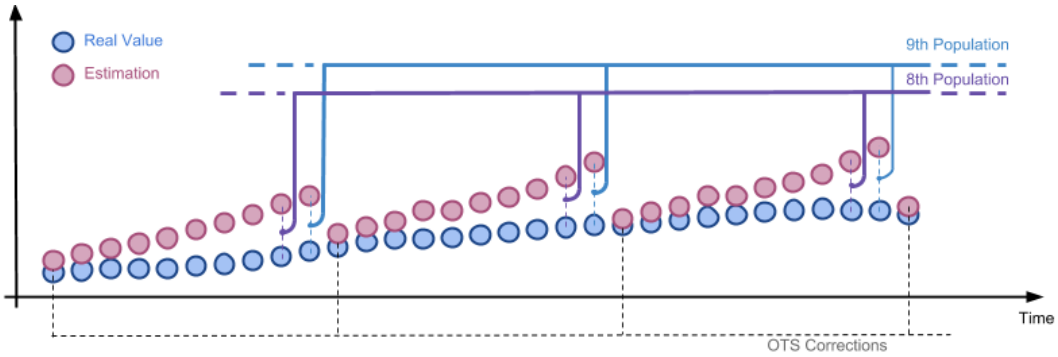

Figure 3.8 Estimation error populations. For each estimated sample the error is computed. Next, errors are categorized in 10 populations depending on their place with respect to OTS corrections.

3.5.2 Marker occlusion

It was mentioned earlier that one of the disadvantages of using optical tracking systems in Computer-assisted surgery is the clear line-of-sight requirement. Sensor fusion can compensate for temporary loss of marker visibility. However, the longer the occlusion time the greater will be the error due to quadratic growth of the drift. In the second category of the experiments we aim to evaluate the performance of our fusion algorithm in marker occlusion.

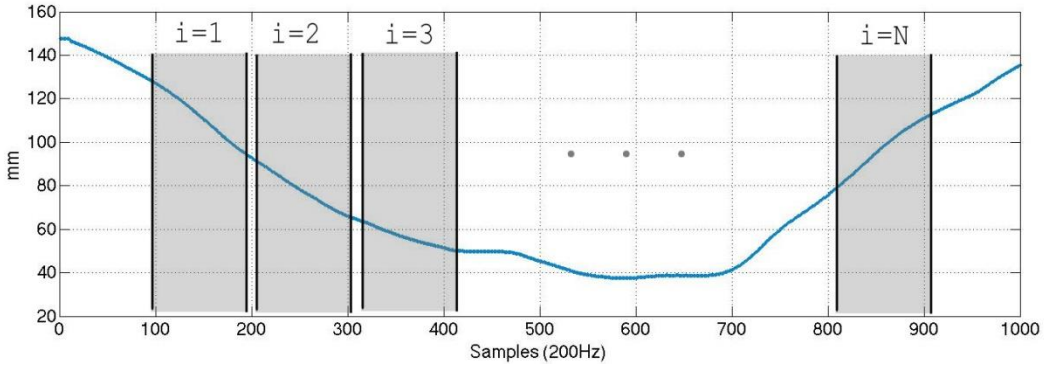

Figure 3.9 Description of moving window of occlusion. For a marker occlusion experiment (1000 samples) an occlusion period of 100 sample is simulated and the signal is used for sensor fusion. Results are recorded and the occlusion window is shifted 100 samples ahead and fed to the fusion system and so on till reaching the end of the signal.

.

The sensor module was moved by hand in space for 5 seconds and data was acquired from both sensors at 200 Hz (feeding a down-sampled 40 Hz OTS signal to the fusion algorithm). The trajectories were random with maximum velocity of 30 mm/s.

To be able to measure the estimation error during the occlusion, the markers were not covered during the acquisitions. Instead, the occlusions were simulated by not feeding the OTS measurements to the fusion algorithm for finite periods. As shown in Figure 3.9, an occlusion window lasting 100 samples was applied N times (N = 9), shifting it over the entire acquisition. 6 acquisitions were performed and 9 occlusions were simulated per acquisition resulting in a total of 54 occlusions. To evaluate the performance of the algorithm as function of the occlusion duration, the errors during each occlusion is recorded and the root mean square error (RMSE) of all the 54 occlusions is calculated as a function of the sample progression from the beginning of the occlusion window. Details of the experimental process are summarized in Table 3.1 and 3.2.

		Sensor Frequency in fusion (Hz)		Max Velocity (mm/s)	Mean Velocity (mm/s)	Number of Acquisitions
		OTS	IMU			
Frequency Augmentation	Simulation	20	200	300	150	6
	Real data	20	200	100	40	12
		20	200	330	100	12
Marker Occlusion	Real data	50	200	100	40	6

Table 3.1 Summary of the performed experiments and simulations for position estimation

		Sensor Frequency in fusion (Hz)		Max Angular rate (°/s)	Mean Angular rate (°/s)	Number of Acquisitions (or Simulations)
		OTS	IMU			
Frequency Augmentation	Simulation	20	200	70	20	6
	Real data	20	200	70	20	12
Marker Occlusion	Real data	50	200	50	10	6

Table 3.2 Summary of the performed experiments and simulations for Orientation estimation

In the next chapter the results of these experiments are presented.

Chapter 4: Results

In this chapter the results of the performed experiments are presented. As it was explained in Chapter 3, the performance of the fusion algorithm was first evaluated by simulated trajectories and then several acquisitions in three categories of experiments were performed (2 for position and 1 for orientation) and fed to the fusion system. The results of each category are presented separately.

Before analyzing the results, it is necessary to have an overview of the performance that is required for our specific application, i.e. robot-assisted neurosurgery. In general, Image-guided surgery (IGS) requires 3–5 mm accuracy, whereas 2 mm is recommended for IG neurosurgery. However, in robot-assisted IGS, even higher accuracy might be necessary. Therefore we have considered 1mm and 2° to be the limits of acceptable estimation error, respectively for position and orientation estimation.

4.1 Frequency Augmentation Results

In 3.5.1 it was explained that for frequency augmentation test a simulation and some experiments were performed. First the result of the simulation is presented and later the experimental results are shown.

4.1.1 Simulated Trajectories

Figure 4.1shows the RMSE of estimated simulated translations. Each bar represents the root mean square of errors corresponding to each estimation population. For example RMS Bar number 2 represents the root mean square of set of the errors of all the 2nd step estimated samples of the simulated trajectory. 95% confidence intervals are calculated to denote the statistical dispersion of each population. The results for the simulated orientation trajectories are shown in Figure 4.2. Errors are reported in degrees. As it can be clearly observed from the figures, both for position and orientation the estimation error stays well below the limits of 1mm and 2°. However, since the linear acceleration and angular rate were calculated numerically the results may not represent the accuracy of the pose estimation precisely. The results of the experimentation on the other hand will provide sufficient proof for capabilities of the fusion system.

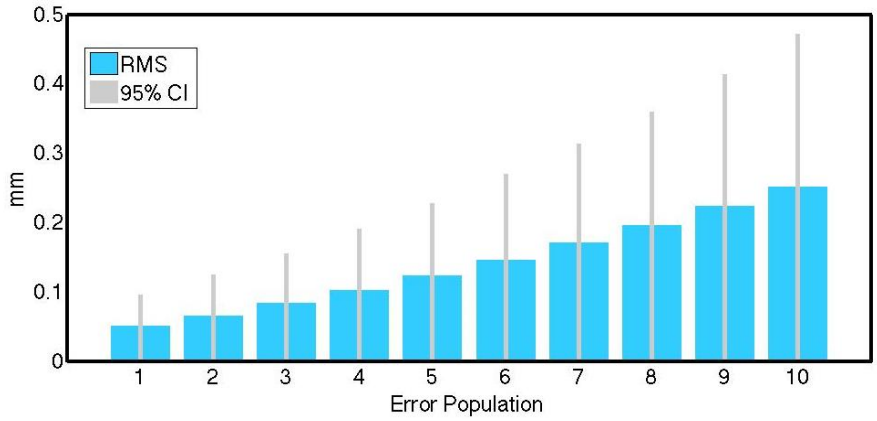

Figure 4.1 Position estimation error along axis y for simulated trajectories. For each estimated sample the error is computed. Next, errors are categorized in 10 populations of "estimation steps after correction".

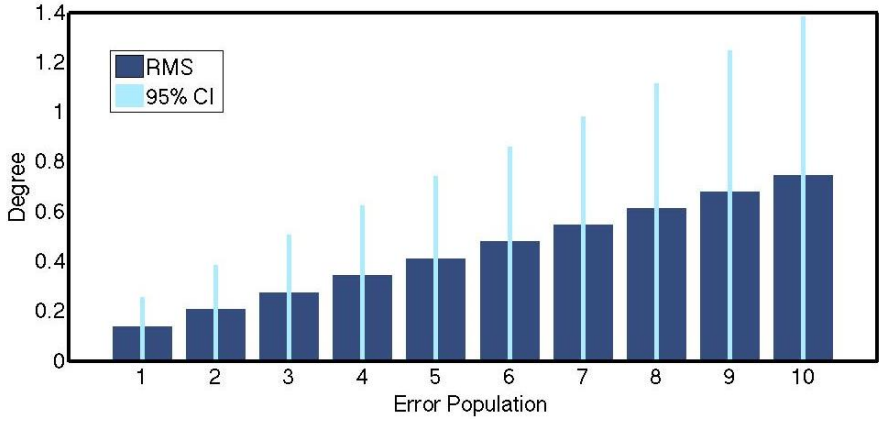

Figure 4.2 Errors of orientation estimation of the simulated trajectories (axis y)

4.1.2 Results of the Frequency Augmentation Experiments

Real data acquisitions were done by moving the sensor module by hand and recording the measurements from the sensors. In 3.5.1 it was explained that the tests were performed in two categories based on velocities. Figure 4.3 shows both the estimated and real trajectories of the object in an acquisition from velocity category 1. Values are all in static coordinate frame. It can be seen from the enlarged part of the drawing that the error in the estimated trajectory increases until the OTS correction is performed. Clearly the sensor fusion system is quite capable of tracking the sensor module and the accuracy of the tracking can be evaluated next.

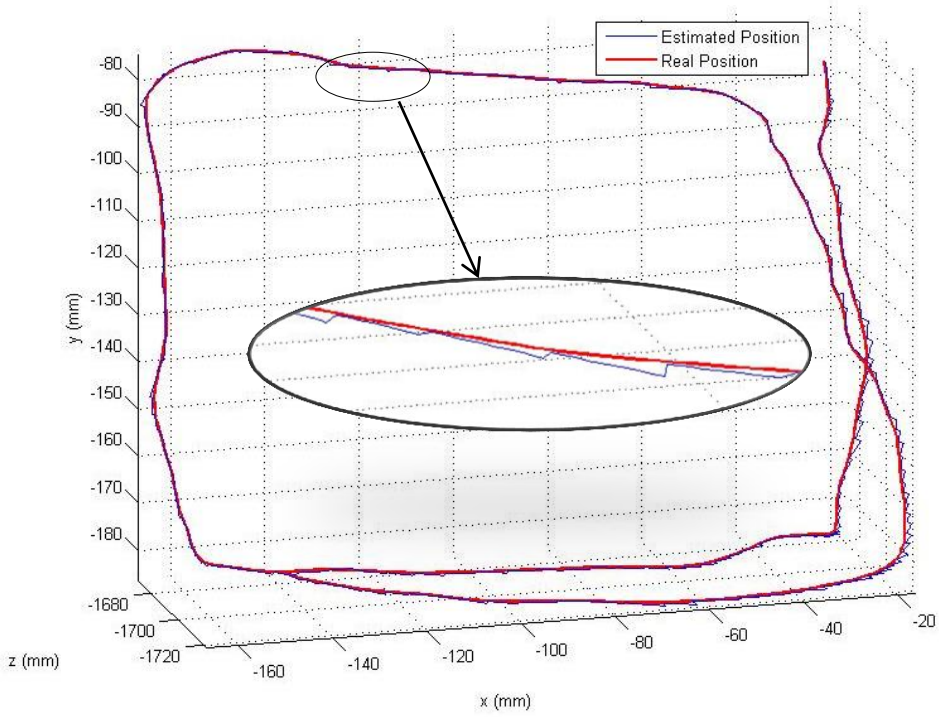

Figure 4.3 Real position and the estimated position vector from the sensor fusion system.

Following the previously described method of visualizing the estimation error, each estimation step is put into the corresponding population and the statistical representations of each population are calculated. Note that each population contains the corresponding estimation step of all the 12 repeated acquisitions in each category. So as an example, 6^{th} population of estimation steps, represents the set of all estimated samples in 12 acquisitions that happen 6 step after an OTS correction. Figure 4.4 shows the result of the position estimation for acquisitions of velocity category 1 with a frequency augmentation ratio (FAR) of 10. As it is expected the error increases as we

go further from the correction step. It can be seen that the RMS of the error populations reaches 0.2mm. In Figure 4.5 the result of the position estimation for velocity category 2 is presented. Naturally due to higher dynamics of movement the errors are higher and the RMSE reaches 0.9mm. Results of position estimation in frequency augmentation test are summarized in Table 4.1.

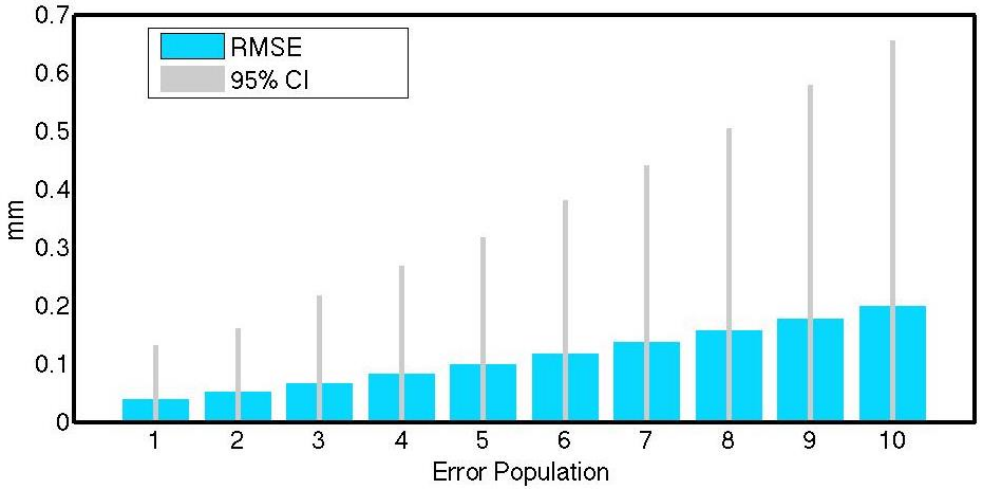

Figure 4.4 Frequency Augmentation results category 1. RMSE of position estimation along axis y for 12 acquisitions of velocity category 1 in case of FAR=10

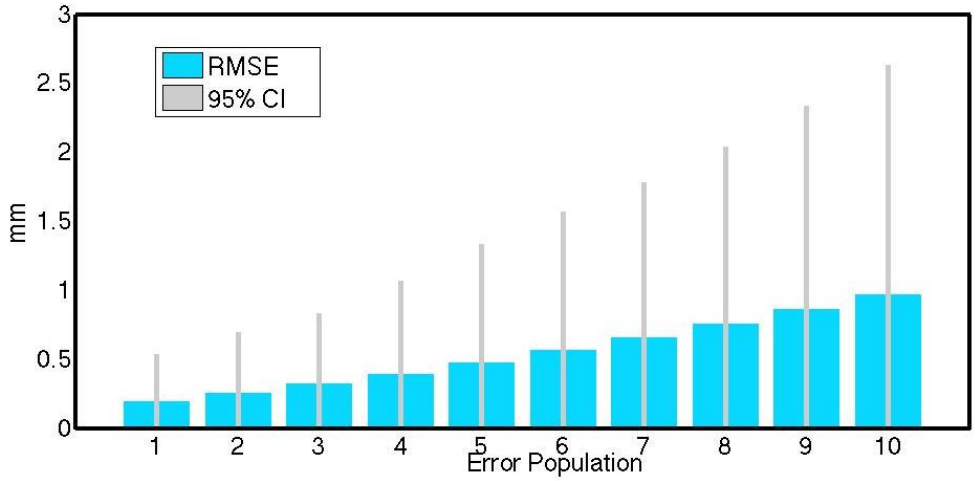

Figure 4.5 Frequency Augmentation results category 2. RMSE of position estimation along axis y for 12 acquisitions of velocity category 2 in case of FAR:10

Orientation experiments were performed in a group of 12 acquisitions. The result of sensor fusion for orientation estimation of the acquired data can be seen in Figure 4.6.

The maximum RMS of the error is 0.7° and the 95% confidence interval reaches 2.3°. It is observed that the values of errors are higher than those achieved in simulated tests. This is due to factors that were not present in simulations like coordinate systems calibration, synchronization and latencies. Nevertheless, the experimentation results suggest that the fusion system is capable of performing pose estimation at the augmented frequency of 200Hz with RMSE below the defined limits of 1mm and 2°.

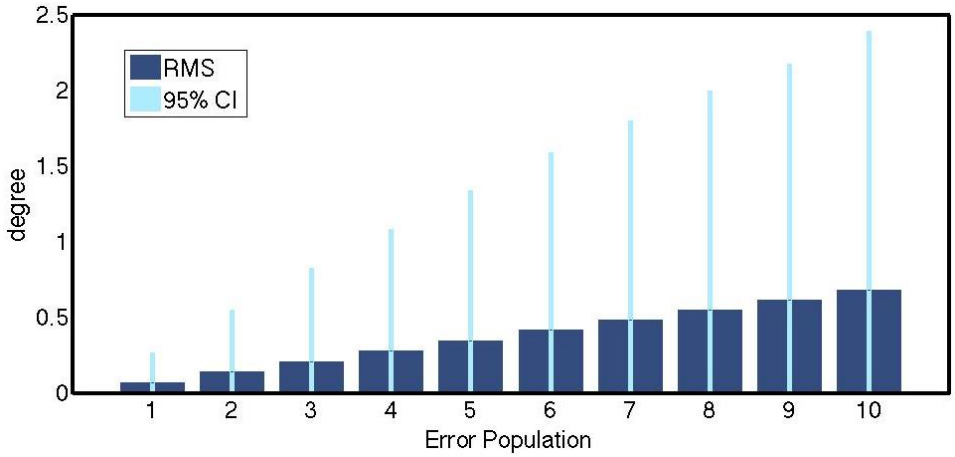

Figure 4.6 Frequency Augmentation results Orientation. RMSE of Orientation estimation around axis y for 12 acquisitions. FAR:10

Note that reducing the ratio of frequency augmentation will of course result in lower estimation errors. As an example Figure 4.7 and Figure 4.10 show the results of pose estimation of the high linear velocity acquisitions of category two with a frequency augmentation ratio of 5 (200Hz IMU/ 40Hz OTS). The RMSE of the position estimation has decreased from 0.9mm in previous case of 10x frequency augmentation to 0.5mm here. For orientation as well a decrease from 0.7 to 0.3 in RMSE can be observed. Table 4.1 and Table 4.2 summarize the results of the frequency augmentation experiments.

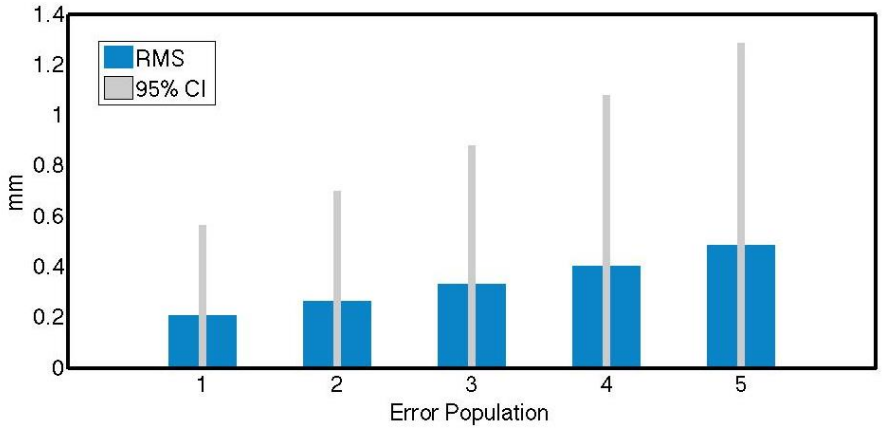

Figure 4.7 RMSE of position estimation along axis y for 12 acquisitions of velocity category 2 in case of FAR: 5

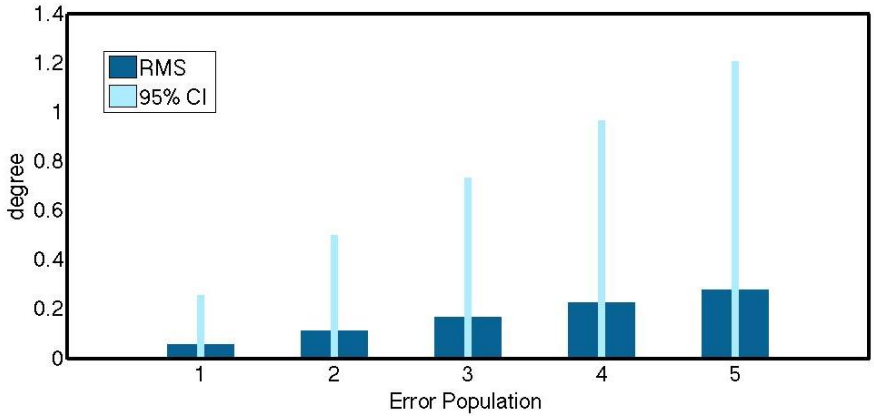

Figure 4.8 RMSE of orientation estimation along axis y for 12 acquisitions around axis y in case of FAR: 5

	FAR	Max RMSE(mm)	95% CI (mm)
Velocity Category 1	10	0.2	0.7
Velocity Category 2	10	0.9	2.6
Velocity Category 2	5	0.5	1.3

Table 4.1 Frequency augmentation results summary. RMSE and 95% confidence interval of position estimation error in experiments.

	FAR	Max RMSE (°)	95% CI (°)
Angular Rate Category 1	10	0.7	2.3
Angular Rate Category 1	5	0.3	1.2

Table 4.2 Frequency augmentation results summary. RMSE and 95% confidence interval of Orientation estimation error in experiments.

4.2 Marker Occlusion Results

Figure 4.9 depicts the performance of the sensor fusion system in presence of marker occlusions. The sensor fusion sampling frequencies are 200 Hz and 40 Hz respectively for IMU and OTS, resulting in an estimation frequency of 200 Hz. In the specific acquisition shown, three occlusions with durations of 50, 70 and 120 samples (Corresponding to 0.25, 0.35 and 0.6 second) happen.

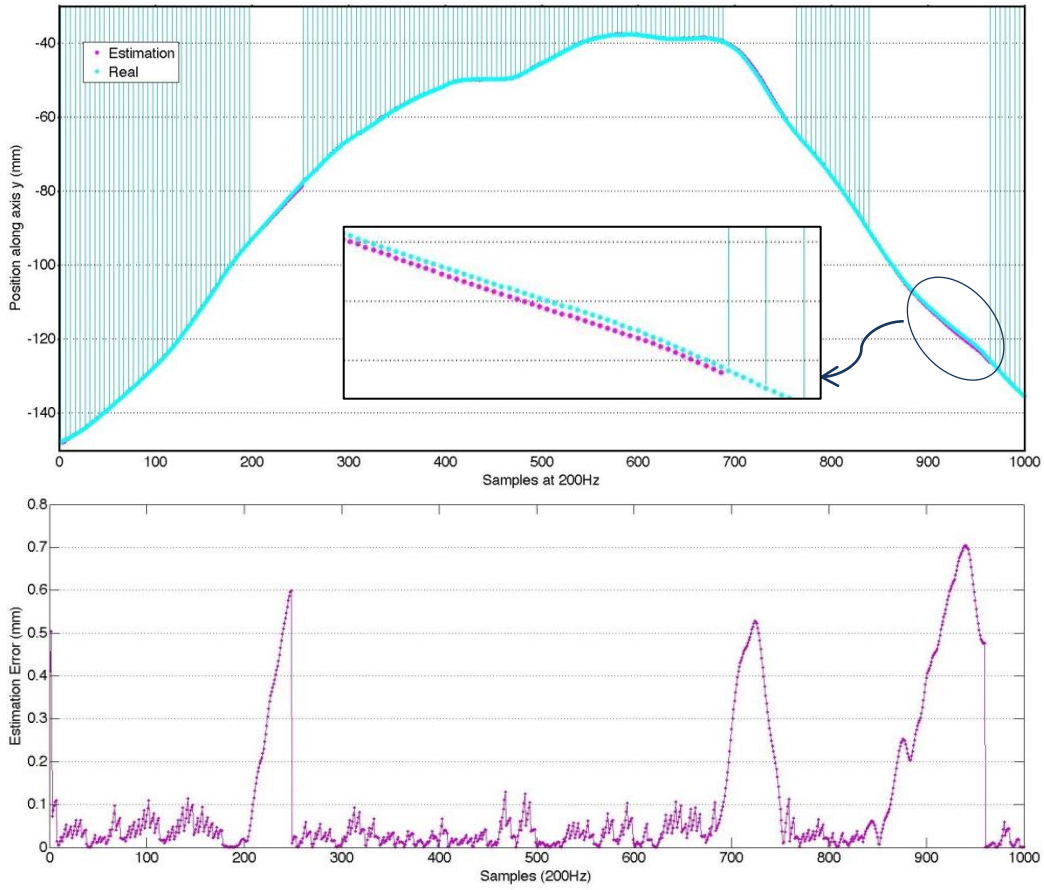

Figure 4.9 Position Estimation In Marker Occlusion Test. Position of the object is estimated using only inertial measurements during 3 occlusions of the optical markers. Vertical bars represent OTS measurements fed to the fusion system. Lowest plot shows the estimation error at each step.

Figure 4.9.top. shows both the estimated position (along axis y) and the real path (OTS measurement). AS it can be seen the object has been moved about 100mm and returned to the first position with an average velocity of 40mm/s. The vertical lines represent OTS measurement. Since the sensor fusion algorithm is using OTS with 5 times sampling frequency of IMU, for each two estimated sample, one OTS correction

happens. Figure 4.9.down, shows how the estimation error changes during the occlusion periods. As it is expected the error increases due to the drift and it reaches a maximum of 0.7mm for this specific acquisition.

As it was explained in 3.5.2 a set of 6 acquisitions were performed for testing marker occlusion compensation capabilities of the system, using a moving 100-sample window of simulated markers occlusion. In Figure 4.10 the results of the all marker occlusion intervals are reported. The graph represents the RMSE of all the 54 marker occlusion as a function of duration of the occlusion. As it is expected the error increases by extending the duration of markers occlusion reaching a maximum of 2.7mm at the end of 100 samples.

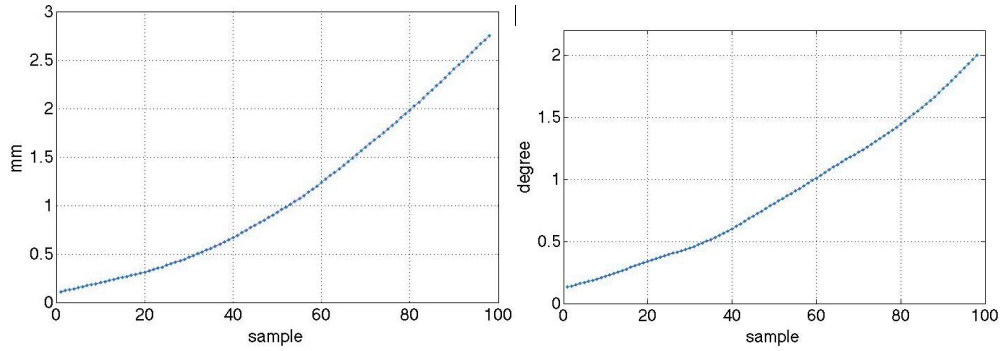

Figure 4.10 RMSE of position estimation error along axis y (Left) and orientation estimation error around axis y for 54 simulated occlusions

Chapter 5: Discussion

This chapter provides a concise, retrospective synopsis of the work presented in this dissertation and emphasizes the notable results. From these results conclusions are drawn and directions for possible future research are proposed.

In this work a quaternion based unscented Kalman filter was developed to be used as a fusion algorithm for Optical and Inertial measurements. Despite that the problems of marker occlusion and frequency augmentation for optical tracking systems have been previously addressed by researches, the accuracy of those efforts were either insufficient [31][32] or not evaluated [29]. The presented fusion algorithm performs robust pose tracking at the higher of the sampling frequencies of the sensors with accuracy in the required range for neurosurgery applications, while it is robust to brief optical marker occlusions. The developed sensor fusion system is based on quaternion representation for orientation, an advantage compared to using the common Euler angles ([30] for example) that suffer from kinematic singularities. In the next sections the achieved results are briefly reviewed and discussed. Last section describes possible future works.

5.1 Frequency Augmentation

The goal of frequency augmentation in sensor fusion is to compensate for low sampling frequency of a sensor by using a second sensor of a different type and estimating at a higher frequency, while keeping the measurement accuracy in an acceptable range. Here we have fused data from an Optical Tracking system at 20 Hz with spatial accuracy of 0.1 mm and an Inertial Measurement Unit sampling at 200 Hz. Results of

several real data acquisitions shown in chapter 4 exhibit that the fusion algorithm maintains the desired accuracy of 1mm (2° for orientation) at velocities as high as 300mm/s while the maximum velocities during awake brain surgery is reported to be 60mm/s [49]. Results also show that, by implementing a higher frequency OTS (40Hz) and therefore having a lower frequency augmentation ratio, the accuracy of the estimation will increase considerably.

5.2 Optical Marker Occlusion

Loosing line of sight is one of the main disadvantages of using optical tracking systems in operating room. Our proposed approach benefits from sensor fusion by estimating the pose of the tracked object using inertial measurements during marker occlusion. Presented results in 4.2 exhibit an acceptable performance from the sensor fusion system. The RMSE remains under our defined limits of 1mm and 1° for up to 55 samples in position and up to 95 samples for orientation (in 200hz sampling). However, it must be considered that the movements during the occlusions had a relatively high velocity (up to 100mm/s) compared to those expected in neurosurgery applications. In case of facing lower velocities of movement the occlusion period can of course be extended.

5.3 Limitations

In this work calibration of the IMU coordinate system and the dynamic reference frame of the OTS has not been performed. Due to hardware problems the OTS signal could not be acquired at 200Hz and the ground truth signal was interpolated from 100 Hz OTS measurements. The object tracking was performed off-line and on-line implementations were not pursued.

5.4 Future Works

Marker occlusion is addressed in this work as the total loss of visibility of all markers. In other words losing visibility of 1 marker will stop the optical tracking data. A more extensive approach is considering the information of individual markers. This will provide partial information during partial occlusions that can be used to improve the accuracy of the sensor fusion system.

Regarding the accuracy improvement, performing a detailed calibration algorithm to calculate a precise transformation matrix from IMU coordinate system to the OTS dynamic reference frame will definitely improve the accuracy of the pose estimation.

On the matter of implementation, the algorithm in this project was used off-line on recorded data from sensors. A definitive next step for this work is on-line implementation. The developed software packages in this project are capable to be implemented for On-line use. However, a study of sensor latencies and synchronization will be necessary first. Real time implementation of the sensor fusion algorithm is another valuable future work.

Bibliography

[1] L. Adams, W. Krybus, D. Meyer-Ebrecht, R. Rger, JM. Gilsbach, R. Msges, G. Schlndorff, **"Computer Assisted Surgery**", IEEE Computer Graphics Application, 1990, pp.10-43.

[2] Paul B. McBeth, Deon F. Louw, Peter R. Rizun, Garnette R. Sutherland, **"Robotics in neurosurgery"**, The American Journal of Surgery 188 (Suppl to October 2004). pp. 68S-75S.

[3] P. Grunert, K. Darabi, J. Espinosa, R. Filippi. **"Computer-aided navigation in neurosurgery"**, Neurosurgical Review May 2003, Volume 26, Issue 2, pp. 73-99

[4] J. L. Crowley and Y. Demazeau, **"Principles and Techniques for Sensor Data Fusion,"** Signal Processing (EURASIP), 1993, Vol. 32. pp. 5-27.

[5] JM. Drake, M. Joy, A. Goldenberg, D. Kreindler, **"Computer and robot assisted resection of thalamic astrocytomas in children"**, Neurosurgery 1991, pp. 27-31.

[6] YS. Kwoh, J. Hou, EA. Jonckheere, S. Hayati, **"A robot with improved absolute positioning accuracy for CT guided stereotactic brain surgery"**. IEEE Trans Biomed Eng 1988, pp. 35-153.

[7] AL. Benabid, P. Cinquin, S. Lavalle , JF. Le Bas, J. Demongeot, J. de Rougemont, **"Computer-driven robot for stereotactic surgery con nected to CT scan and magnetic resonance imaging: technological design and preliminary results".** Appl Neurophysiol 1987, pp. 50:153-4.

[8] CW. Burckhardt, P. Flury, D. Glauser, **"Stereotactic brain surgery".** IEEE Eng Med Biol Mag, 1995, pp.14:314-7.

[9] R. Nakamura, K. Masamune, Y. Nishikawa, **"Development of an MRI-compatible needle insertion manipulator for stereotactic neurosurgery",** J Image Guid Surg 1995, pp.1-242.

[10] DW. Roberts, A. Hartov, FE. Kennedy, MI. Miga, KD. Paulsen, **"Intraoperative brain shift and deformation: a quantitative analysis of cortical displacement in 28 cases"**, Neurosurgery, 1998, pp. 749-760.

[11] E. Watanabe, Y. Mayanagi, Y.Kosugi, S. Manaka, K. Takakura, **"Open surgery assisted by the neuronavigator, a stereotactic, articulated, sensitive arm"**, Neurosurgery , 1991, pp. 79-800.

[12] HF. Reinhardt, GA. Horstmann, O. Gratzl, **"Sonic stereometry in microsurgical procedures for deep-seated brain tumors and vascular malformations"**, Neurosurgery, 1993, pp. 51–57.

[13] A. Kato, T. Yoshimine, T. Hayakawa, Y. Tomita, T. Ikeda, M. Mitomo, K. Harada, H. Mogami, **"A frameless, armless navigation system for computer-assisted neurosurgery"**, Neurosurg, 1991, pp. 845-849.

[14] L. Zamorano, LP. Nolte, AM. Kadi, **"Interactive intraoperative localization using infrared-based system"**. Neurol Res 15, pp. 290-298

[15] K. Chinzei K, K. Miller , **"Towards MRI guided surgical manipulator",** Med Sci Monit, 2001, 7, pp.153- 63.

[16] DF. Louw, T. Fielding, PB. McBeth, D. Gregoris, P. Newhook, GR. Sutherland, **"Surgical robotics: A review and neurosurgical prototype development"**, Neurosurgery 2004, 54, pp.525-37.

[17] G. Guthart and J. Salisbury, **"The intuitive telesurgery system: Overview and application",** In Proceedings of the 2000 IEEE International Conference on Robotics and Automation, 2000

[18] GR. Sutherland, S. Wolfsberger, S. Lama S, K. Zarei-nia, ***"The Evolution of neuroArm"***. Neurosurgery 72, 2013, pp. A27-A32.

[19] E. Foxlin, **"Inertial Head Tracking"** Master's Thesis, Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 1993.

[20] E. Kraft, **"A Quaternion-based Unscented Kalman Filter for Orientation Tracking",** Proceedings of the Sixth International Conference of : Information Fusion, 2003, Volume: 1, pp. 47-54

[21] A. Vaccarella, E. De Momi, M. Valenti, G. Ferrigno, and A. Enquobahrie, **"Application of unscented kalman filter for robust pose estimation in image-guided surgery,"** in Proc. SPIE, vol. 8316. 2012, pp. 1–6.

[22] F. Raab, E. B. Blood, T. O. Steiner, and H. R. Jones, **"Magnetic Position and Orientation Tracking System"** IEEE Transactions on Aerospace and Electronic Systems, 1979, Vol. AES-15, pp. 709-718.

[23] S. Sukkarieh, M, Eduardo, Nebot and F. Hugh, **"Achieving Integrity in an INS/GPS Navigation Loop for Autonomous Land Vehicle Applications"**, Proceedings of IEEE Conference on Robotics and Automation, May 1998.

[24] M. George and S. Sukkarieh, **"Tightly Coupled INS/GPS with Bias Estimation for UAV Applications"**, In Proceedings of Australiasian Conference on Robotics and Automation (ACRA), December2005.

[25] NRC National Research Council, **"Virtual Reality, Scientific and Technological Challenges,"** National Academy Press, 1994, Washington, DC.

[26] K. Meyer, H. Applewhite and F. Biocca, **"A Survey of Position Trackers. Presence"**, a publication of the Center for Research in Journalism and Mass Communication, The University of North Carolina at Chapel Hill,1992.

[27] S. Emura and S. Tachi, **"Sensor Fusion based Measurement of Human Head Motion",** Proceedings 3rd IEEE International Workshop on Robot and Human Communication, 1994, Nagoya (Nagoya University, Nagoya, Japan).

[28] A. State, G. Hirota, D. Chen, B. Garrett, M. Livingston, "**Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking"**, Proceedings of ACM SIGGRAPH '96 (New Orleans, LA, August 1996), Computer Graphics, Annual Conference Series.

[29] N. Parnian, S.P. Won, F. Golnaraghi, **"Position Sensing Using Integration of a Vision System and Inertial Sensors"**, 34th Annual Conference of the IEEE Industrial Electronics Society (2008), pp. 3011-3015

[30] D. Roetenberg, H. J. Luinge, T. M. Baten, and P. H. Veltink, "**Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation",** IEEE Trans. Neural Syst. Rehab. Eng., Sep. 2005, vol. 13, no. 3, pp. 395-405.

[31] A. Tobergte, M. Pomarlan, G. Hirzinger, **"Robust Multi Sensor Pose Estimation for Medical Applications"**, IEEE/RSJ International Conference on Intelligent Robots and Systems (2009), pp. 492-497

[32] B. Hartmann, N. Link, G.F. Trommer, **"Indoor 3D Position Estimation Using Low-Cost Inertial Sensors and Marker-Based Video-Tracking"** IEEE/ION Position Location and Navigation Symposium (2010), pp. 319–326

[33] D. Hall and J. Llinas. **"An introduction to multisensor data fusion".** IEEE Proceedings, 85(1), January 1997.

[34] A. Kim and M. F. Golnaraghi **"Initial calibration of an inertial measurement unit using an optical position tracking system"**, *Proc. IEEE Posit. Locat. Navigat. Symp.*, 2004, pp. 96 -101

[35] P. Lang and A. Pinz, **"Calibration of hybrid vision/inertial tracking systems."** 2nd InverVis 2005: Workshop on Integration of Vision and Inertial Systems, Barcelona, Spain.

[36] R. G. Brown and P. Y. C. Hwang, **"Introduction to Random Signals and Applied Kalman Filtering"**, 1992, 2nd Edition, John Wiley & Sons, Inc.

[37] A. Gelb, **"Applied Optimal Estimation"**, 1974, MIT Press, Cambridge, MA.

[38] O. L. R. Jacobs, **"Introduction to Control Theory"**, 1993, 2nd Edition. Oxford University Press.

[39] The American Heritage Dictionary, Houghton Mifflin Company, 1981, p. 980.

[40] R.E. Kalman, **"A New Approach to Linear Filtering and Prediction Problems,"** Transaction of the ASME—Journal of Basic Engineering, 1960, pp. 35-45.

[41] S.J. Julier and J.K. Uhlmann, **"A New Extension of the Kalman Filter to Nonlinear Systems",** In Proc. SPIE - Int. Soc. Opt. Eng. (USA) (Orlando, FL, April 1997), vol. 3068, pp. 182-193.

[42] S.J. Julier and J.K. Uhlmann, **"Unscented Filtering and Nonlinear Estimation"** Proceedings of the IEEE, Volume: 92 , Issue: 3 (2004) , pp): 401 – 422

[43] R. van der Merwe, **"Sigma-point Kalman filters for probabilistic inference in dynamic state-space models",** Ph.D. thesis, 2004, Oregon Graduate Institute School of Science and Engineering, Beaverton, Oregon. A

[44] J. K. Uhlmann. **"Simultaneous map building and localization for real time applications."** Technical report, University of Oxford, 1994.

[45] E. Dam, M. Koch, M. Lillholm, **"Quaternions, Interpolation and Animation",** Technical Report DIKU-TR-98/5, Department of Computer Science, Univ. Copenhagen, Denmark, 1998.

[46] S. Fioretti, **"Signal processing in movement analysis (a state-space approach Human Movement Science**", 1996, 15, pp. 389-410.

[47] S.J. Julier, **"The Scaled Unscented Transformation",** In Proceedings of the American Control Conference (May 2002), vol. 6, pp. 4555-4559.

[48] X. Pennec, **"Computing the mean of geometric features- Application to the mean rotation",** Institut National de Recherche en Informatique et en Automatique (IN-RIA), Le Chesnay, France, 1998.

[49] D. De Lorenzo, et al. **"Intraoperative forces and moments analysis on patient head clamp during awake brain surgery."** Medical & biological engineering & computing (2013): pp.1-11.