Automatic multimodal real-time tracking for image plane alignment in interventional Magnetic Resonance Imaging

Markus Neumann

UNIVERSITÉ DE STRASBOURG

École Doctorale Mathématiques, Sciences de l'Information et de l'Ingénieur

Thèse

présentée pour obtenir le grade de

Docteur de l'Université de Strasbourg
Discipline : Signal, Image, Automatique et Robotique
Spécialité : Robotique

par

Markus Neumann

Automatic multimodal real-time tracking for image plane alignment in interventional Magnetic Resonance Imaging

Soutenue publiquement le 25 février 2014

Membres du jury :

Directeur de Thèse	Michel de Mathelin, Professeur, Université de Strasbourg
Rapporteur Externe	Jocelyne Troccaz, Directeur de recherche CNRS, TIMC, Grenoble
Rapporteur Externe	Bruno Quesson, Directeur de recherche CNRS, CRCTB, Bordeaux
Examinateur	Michael Bock, Professeur, Universität Freiburg
Examinateur	Afshin Gangi, Professeur, Université de Strasbourg
Examinateur	Elodie Breton, Ingénieur de recherche, Université de Strasbourg
Invité	Loïc Cuvillon, Maître de conférences, Université de Strasbourg

Laboratoire des sciences de l'Ingénieur, de l'Informatique et de l'Imagerie - ICube
UMR UdS-CNRS 7357

Meiner Familie
und Andrea

Acknowledgements

First of all, I'd like to thank the thesis committee members and especially the external examiners for having accepted to read and evaluate my work:

- Michael Bock, Professor, Universität Freiburg, Chair of the examining committee
- Jocelyne Troccaz, Directeur de recherche CNRS, TIMC, Grenoble
- Bruno Quesson, Directeur de recherche CNRS, CRCTB, Bordeaux
- Afshin Gangi, Professeur, Université de Strasbourg

Thank you very much for your constructive comments and suggestions, that helped me to improve this work.

Je remercie Michel de Mathelin, mon directeur de thèse, de m'avoir accueilli dans son équipe et de m'avoir confié ce travail de thèse. Je lui suis très reconnaissant pour ses idées, ses conseils ainsi que son enthousiasme et optimisme qui m'ont aidé de mener à bien ce travail.

Je remercie tout particulièrement Elodie Breton et Loïc Cuvillon, mes encadrants, pour leur soutien et leur grande disponibilité au cours de cette thèse, notamment pour les tests à l'IRM, la correction des papiers et de ce rapport. Leur aide, leur investissement ainsi que leurs conseils ont été d'une très grande valeur pour moi. Merci!

Je tiens à remercier le Ministère de l'Enseignement supérieur et de la Recherche de m'avoir attribué une allocation de recherche, ainsi que l'institut Carnot Télécom et Société numérique, l'IHU de Strasbourg et le LABEX CAMI pour leurs financements qui m'ont permis de réaliser ma thèse dans de très bonnes conditions. Je remercie également l'IRCAD de Strasbourg de m'avoir donné la possibilité de travailler dans un cadre exceptionnel.

Merci au service de la radiologie interventionnelle du Nouvel Hôpital Civil de Strasbourg et notamment au Professeur Afshin Gangi et son équipe pour leur soutien et bonne humeur.

Furthermore, I'd like to thank the Siemens team, namely Eva Rothgang, Li Pan, Patrick Gross, Arne Hengerer and Christine Lorenz, for their help and support. Thank you very much for the

invitation to the Siemens Interventional MRI workshop in Boston which has been an enriching and interesting experience.

De même, je veux remercier Wilfried Uhring et Daniel Gounot qui m'ont donné la possibilité de tester avec leur matériel et d'utiliser leurs locaux.

Je tiens à remercier particulièrement Lucille Zorn et Laurent Barbé pour la conception et la mise en service du banc d'essai compatible IRM. Sans eux, je n'aurai pas pu faire la validation expérimentale. Je remercie aussi Laurent Goffin et Nicolas Padoy pour leur aide avec la mise en service du capteur RGB-D. Je tiens également à remercier Florent Nageotte et Philippe Zanne pour leur disponibilité pour mes divers questions au cours de cette thèse.
Je tiens à remercier tous les membres permanents de l'équipe AVR que j'ai pu côtoyer de près ou de loin pour la bonne ambiance de travail.

Je remercie très chaleureusement les anciens et nouveaux doctorants de l'équipe AVR dont j'ai eu la chance de faire connaissance durant ces 3 années: Bérengère, Julien, Norbert, Fadi, Houssem, Salih, Xavier, Lennart, Antonio, Laure, Nitish, Ryad, Arnaud, Laure-Anais, Nadège, Paolo, Xavier, Rahim, Andru, Nicole et Quentin. Je garderai un très bon souvenir de la bonne ambiance de travail, du week-end mémorable à Lyon et des séances de basket hebdomadaires. Merci à vous tous.

Durant les 3 ans de thèse, j'ai également eu la chance de faire connaissance de beaucoup de stagiaires, qui ont enrichi la vie au labo et dont je veux remercier notamment: Florent, Ivan, Jean-Baptiste, Nicolas, Clémentine, David (oui surtout David ;), Jordane, Vincent.

Weiterhin möchte ich meiner Familie danken. Mama, Papa, Anja und Janine, ich bin euch dankbar für eure Liebe und Zuneigung, euer Interesse an meiner Arbeit und für eure bedingungslose Unterstützung. Ausserdem möchte ich mich bei Johannes und der kleinen Ida bedanken. Ida, die Vorfreude auf deine Ankunft (die du bis aufs Letzte ausgereizt hast) war eine grosse Motivation.
Zum Schluss möchte ich meiner Andrea danken, die mich über die ganze Zeit liebevoll unterstützt hat und mir in den schwierigeren Momenten beigestanden und geholfen hat. Danke für alles was wir in den letzten Jahren zusammen erlebt und erreicht haben.

Ohne euch wäre ich nie so weit gekommen.

Abstract

Interventional magnetic resonance imaging (MRI) aims at performing minimally invasive percutaneous interventions, such as tumor ablations and biopsies, under MRI guidance. During such interventions, the acquired MR image planes are typically aligned to the surgical instrument (needle) axis and to surrounding anatomical structures of interest in order to efficiently monitor the advancement in real-time of the instrument inside the patient's body. Object tracking inside the MRI is expected to facilitate and accelerate MR-guided interventions by allowing to automatically align the image planes to the surgical instrument.
In this PhD thesis, an image-based workflow is proposed and refined for automatic image plane alignment. An automatic tracking workflow was developed, performing detection and tracking of a passive marker directly in clinical real-time images. This tracking workflow is designed for fully automated image plane alignment, with minimization of tracking-dedicated time. Its main drawback is its inherent dependence on the slow clinical MRI update rate. First, the addition of motion estimation and prediction with a Kalman filter was investigated and improved the workflow tracking performance. Second, a complementary optical sensor was used for multi-sensor tracking in order to decouple the tracking update rate from the MR image acquisition rate. Performance of the workflow was evaluated with both computer simulations and experiments using an MR compatible testbed. Results show a high robustness of the multi-sensor tracking approach for dynamic image plane alignment, due to the combination of the individual strengths of each sensor.

Resumé

En imagerie par résonance magnétique (IRM) interventionnelle, des interventions percutanées minimalement-invasives (biopsies, ablations de tumeurs,...) sont réalisées sous guidage IRM. Lors de l'intervention, les plans de coupe acquis sont alignés sur l'outil chirurgical et les régions anatomiques d'intérêt afin de surveiller la progression de l'outil dans le corps du patient en temps réel. Le suivi d'objets dans l'IRM facilite et accélère les interventions guidées par IRM en permettant d'aligner automatiquement les plans de coupe avec l'outil chirurgical. Dans cette thèse, un système d'alignement automatique des plans de coupe établi sur une séquence IRM clinique est développé. Celui-ci réalise automatiquement la détection et le suivi d'un marqueur passif directement dans les images IRM tout en minimisant le temps d'imagerie dédié à la détection. L'inconvénient principal de cette approche est sa dépendance au temps d'acquisition de la séquence IRM clinique utilisée. Dans un premier temps, les performances du suivi ont pu être améliorées grâce à l'estimation et la prédiction du mouvement suivi par un filtre de Kalman. Puis un capteur optique complémentaire a été ajouté pour réaliser un suivi multi-capteurs, découplant ainsi la fréquence de rafraichissement du suivi de la fréquence de rafraichissement des images IRM. La performance du système développé a été évaluée par des simulations et des expériences utilisant un banc d'essai compatible IRM. Les résultats montrent une bonne robustesse du suivi multi-capteurs pour l'alignement des plans de coupe grâce à la combinaison des qualités individuelles de chaque capteur.

Summary

Introduction

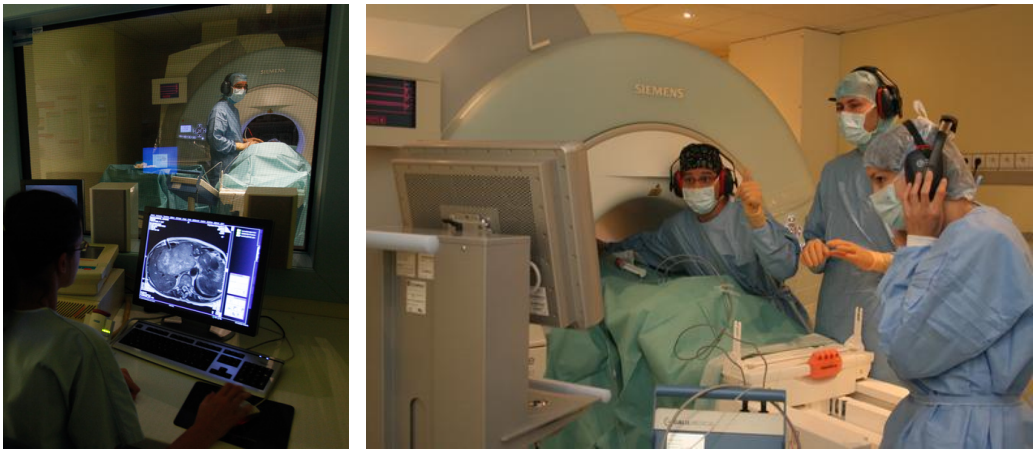
Interventional radiology

Interventional radiology consists in using medical imaging for the guidance of minimally invasive surgical procedures. Typical procedures such as biopsies, infiltrations and ablations of tumors can be performed for diagnostic or therapeutic reasons. They are performed through percutaneous access and are directly monitored through an imaging device. The used imaging modalities include ultrasound (US), computer tomography (CT) or magnetic resonance imaging (MRI). Advantages of MRI-guidance for interventional radiology include absence of ionizing radiation exposure for both patient and medical staff, better soft tissue contrast and free image plane orientation and positioning during acquisitions.

Interventional MRI and thesis motivation

Nevertheless, procedures under MRI guidance can be time consuming. One to three successive image planes are continuously acquired to monitor the real-time procedure. Typically, physicians are interested in the image planes being aligned to the main axis of the interventional instrument to obtain a "3D-like" vision of the instrument, target lesion and the surrounding anatomical structures of interest. As the physician is occupied with the surgical procedure, he/she cannot at the same time align the image planes to the performed instrument movements. In order to monitor the advancement of the instrument inside the patient's body, the image planes are thus manually aligned by a technologist at the MRI console. Due to visual occlusion by the MRI scanner and the surrounding machines (e.g. for monitoring of vital signs of the patient), the technologist cannot observe the movements of the physician's hand in order to alter the image planes accordingly. Vocal communication between physician inside the MRI room and the technologist at the MRI console is usually possible over the standard patient intercom system. In that case, the physician cannot talk to the technologist during image acquisition due to the strong noise inside the MRI room. Thus positioning and orientation of image planes is often performed with the physician using sign language. As a consequence, the quality of the alignment of the image plane to the instrument strongly depends on the experience of the technologist at the MRI console, on his/her communication means with the physician and on the experience of them together as a team.

Hence, we propose throughout this thesis the development of a real-time image plane alignment method, allowing to control the real-time MR image planes automatically from inside the MRI room.



(a) (b)

Figure 1: (a) Interventional MRI environment: physician inside the MRI room performing the intervention and technologist at the MRI console in the MRI control room. (b) Intervention in MRI with the physician lying inside the magnet to advance the needle based on the images displayed in real-time. The physician uses sign language for communication with the technologist in the MRI control room.

Thesis Organization

The first chapter, presents an overview of recent developments, technical constraints and typical procedures in interventional MRI. Assistance for entry point positioning, automatic image plane alignment and automatic detection of the interventional tool are highly interesting features in order to facilitate and eventually improve the accuracy of such procedures. Specifically a literature review of tracking systems for interventional MRI is proposed before the thesis objectives are presented.

In the second chapter, an automatic image plane alignment workflow is presented using a detection and tracking algorithm for a passive marker and resonant micro-coils. This image-based tracking workflow is designed for fully automated image plane alignment, with minimization of tracking-dedicated time. The principles of the workflow as well as the developed detection algorithm are detailed. The implementation of the workflow in a C++ program for communication with the MRI console PC is then described. The tracking performance of the presented workflow is evaluated with simulations and experiments using an MRI compatible testbed.

The presented workflow is then extended with a Kalman filter for improvement of its tracking

performance. Chapter 3 starts with an introduction of the Kalman filter and its benefits. The integration of the filter in the image plane alignment workflow is then described. In order to account for disparities in the measurement confidence for different image planes, a dynamic adaptation of the measurement noise covariance matrix of the Kalman filter has been investigated for the presented workflow. The benefits of the Kalman filter for image plane alignment are then highlighted with the same simulations and experiments as those performed for the initial workflow.

Due to the slow image acquisition rate of the clinical MRI sequence, it is proposed in chapter 4 to decouple the tracking from the MR images alone, through use of an additional sensor. For this reason, an RGB-D sensor is integrated into the initial workflow. The integration of the RGB-D sensor into the existing workflow is presented in its different aspects, i.e. MRI compatibility, 3D position measurement and registration. An online registration step is performed in order to use the RGB-D sensor measurements in the MRI frame. Furthermore, measurements based on the RGB-D sensor and on the MR images are fused using an Information filter. This hybrid workflow, combining active and passive tracking approaches is then experimentally validated.

1 Interventional Magnetic Resonance Imaging (MRI)

1.1 Current status of interventional MRI

1.1.1 Interventional radiology and interventional MRI

In interventional radiology, different imaging modalities such as ultrasound (US), X-ray computer tomography (CT) or magnetic resonance imaging (MRI) are used for guidance of minimally invasive percutaneous procedures including biopsies, infiltrations, ablations and also intravascular interventions. MRI has several benefits over other imaging modalities (CT, US), starting with its excellent soft-tissue contrast and a high image resolution. Both anatomic and physiologic information can be acquired. It is thus possible to assess functional parameters such as flow velocities, thermal imaging, perfusion and diffusion. Compared to CT, MRI is a non-ionizing medical imaging modality, thus preserving both patient and practitioner from radiations during image acquisition. Moreover, MRI allows to image any 3D volume or single slice imaging plane in space as the image plane orientation and position can be freely chosen. Due to these advantages, MRI has become an attractive imaging modality for diagnostic and therapeutic image-guided procedures. However, interventional MRI remains clinically confidential due to strong economical (aspect not covered in this work) and technical constraints, such as the high permanent magnetic field and the radio-frequency (RF) pulses used for imaging, and the restrained spatial access to the patient.

1.1.2 Technical constraints in the MRI environment

Interventional procedures in MRI include thus the use of surgical tools and machines within the MRI environment. As a high permanent magnetic field, transient gradient magnetic fields, and radio frequency (RF) pulses are used for MR imaging, multiple technical constraints have to be respected for the use of additional tools inside the MRI room ([Daa01]).

The permanent magnetic field (B_0) of the MRI scanner is probably the most constraining property of this imaging modality due to its field strength typically between 0.1 T and 3 T

for interventional MRI scanners. Strength of attraction of a material by the MRI scanner magnet depends on the magnetic properties of the material and the scanner field strength. Paramagnetic materials are only weakly attracted to the magnetic field in contrast to ferromagnetic materials, which are subject to high attraction. Furthermore, the use of ferromagnetic materials in the proximity of the permanent magnetic field causes a disturbance of its homogeneity resulting in geometrical distortions of the acquired images. Static disturbances of metallic materials can be compensated through shim coils, whereas dynamic motion of metallic materials causes local disturbances of the field homogeneity, which result in artifacts within the MR image.

During MR imaging, RF pulses between ~8 MHz (0.1 T) and ~123 MHz (3 T) are used for excitation of the hydrogen nucleus. Instrumentation inside the MRI room must not only endure these pulses but must also be resistant against RF induced heating. Furthermore, data transmission inside the MRI room can be compromised through the RF pulses. Thus instrumentation has to be shielded against RF noise. Due to the high sensitivity of the reception antennas of the MRI scanner, shielding is also important for protection of the MR image quality with respect to electromagnetic noise produced through electrical appliances used within the MRI room. Furthermore, disturbances from outside the MRI room must be avoided. Hence, data transmission between the MRI room and the outside must be performed with shielded cables and RF filters or with optical fiber cables in order to sustain the functionality of the Faraday cage surrounding the MRI room.

The typical MRI bore is a small radius long tunnel in which access to the patient is rendered extremely difficult for the interventionalist. Even in recent "open-bore" MR systems, spatial access to the patient remains constrained.

The last typical constraint of the MRI environment is the strong acoustic noise of due to fast current switches in the gradient coils. This noise impedes oral communication between staff members inside and/or outside the MRI room, but also between the staff and awaken patient. Both patient and practitioner must wear noise protecting headsets or earplugs.

Apart from the technical constraints, the patient herself/himself must not present any contraindications to the MRI exam such as metallic implants (metallic cardiac valves, vascular clips,...), pacemakers or other non-MR safe electrical medical implants inside the body.

Despite these constraints, interventional MRI has emerged over the years thanks to the evolution of MRI imaging sequences, hardware and software [MTKB08].

1.1.3 Technical developments in interventional MRI

First interventional MRI scanners were C-arm type magnets where the permanent magnetic field is oriented vertically between two magnet pole shoes. Further developments resulted in double-donut MRI scanners where two solenoid coils with a vertical gap in between are

used for image acquisition (Figure 1.1a). The advantage of such systems is that they allow to access the patient from several directions and are thus well suited for biopsies or ablations. However due to their low permanent magnetic field strength of under 1 T, these systems had a lower image quality and acquisition speed than diagnostic systems with a field strength of 1.5 T or more. A C-arm type with a field strength of 1 T has been developed (Philips Panorama MR, Philips Healthcare, Best, Netherlands), but the pole shoes are much larger than for the initial C-arm type and restrain access to the patient (Figure 1.1b). Interventional solenoid magnets with short (length of ~1.5 m) and wide (diameter of ~0.7 m) bores have then been developed allowing to acquire MRI images of higher quality at a field strength of 1.5 T with the possibility for the physician to access the patient within the MRI scanner from the tunnel opening (Figure 1.1c).

Besides the development of interventional MRI scanners, further efforts have been made to

(a) (b) (c)

Figure 1.1: Interventional MRI scanners. (a): Double donut MRI scanner (Signa SP, GE, Little Chalfont, UK). Reprinted from [SGKU02]. (b): MRI scanner with horizontal gap (Panorama, Philips, Best, Netherlands). Reprinted from [htt13a]. (c): Open bore MRI scanner (MAGNETOM Espree, Siemens AG, Erlangen, Germany) with the physician accessing the patient through the tunnel opening.

furnish the interventional MRI room with compatible hardware such as shielded monitors for MR image display to the physician inside the MRI room. Furthermore, MR-compatible input devices such as table controls, keyboards and foot switches have been developed allowing to change imaging parameters interactively from within the MRI room. Accordingly, software protocols for real-time communication with the MRI console PC have been opened to the research community allowing to develop software interfaces for interactive control of sequence parameters in real-time. As presented in [BG12], several commercial solutions are available for interactive real-time control of MR imaging parameters.

One of the major requirements for interventional MRI was the development of interventional imaging pulse sequences allowing to acquire MR images rapidly with a high soft tissue contrast. For this purpose, great efforts have been made for the development of rapid imaging sequences allowing to acquire different types of contrasts such as T1-weighted, mixed T1- and T2-weighted and $T2^*$-weighted images ([BG12]). Interventional pulse sequences have been developed for interactive alteration of image plane position and orientation, as well as image contrast. Such pulse sequences typically allow to successively acquire and display several

image planes in order to monitor the needle advancement in percutaneous procedures.

Thanks to these developments, interventional MRI is today a powerful tool for therapeutic procedures and part of the clinical routine for cancer treatment.

1.1.4 Clinical workflow of MR-guided percutaneous procedures

Typical therapeutic and diagnostic percutaneous procedures under MR-guidance are biopsies, infiltrations, ablations and intravascular interventions ([Jol98]). Biopsies, infiltrations and ablations are typically performed in the prostate, liver, kidney and breast. Intra-vascular interventions are performed in order to deliver stents, vena cava filters or embolization materials using catheters.
During a biopsy, a sample of tissue or bone is retrieved from a suspicious area for examination. The sample is retrieved through a puncture with a hollow needle extracting a core of the suspicious area. During ablations, one or several needles are placed within a tumor, which is then destroyed either through RF ablation, cryoablation, laser ablation, high-intensity focused ultrasound (HIFU) or brachytherapy. For an RF ablation, a needle is placed within the tumor and an electrical current is applied in order to heat the tissue. In the same manner, laser ablation and HIFU rely on tissue heating for destruction of tumors. During a cryoablation, the tissue is destroyed through freezing. For this purpose a gas (e.g. argon) is led through a hollow needle into the tumor and freezes the surrounding tissues. Thawing cycles can be alternated with freezing until destruction of the tissue ([GTA+12]). During brachytherapy, radioactive seeds are implanted into the tumor in order to destroy the surrounding tissues.

Hence, typical interventional tools for MR-guided procedures are either rigid, such as needles, or flexible, such as catheters and guidewires. During the procedures, it is mandatory to correctly detect the position and orientation of the interventional tool in order to monitor its advancement. For that matter, the image planes are typically aligned to the tool main axis or perpendicular to it (Figure 1.2).
A typical workflow of a percutaneous intervention can be described as follows:

1. Pre-operative planning
   The patient is installed inside the MRI scanner. A 3D volume acquisition of the zone of interest is acquired and the needle insertion point and path are planned in the MR images.
2. Marking of the needle insertion point
   The point of needle insertion is then marked manually on the patient's skin by identifying the insertion point through placement of an MR-visible marker on the patient. The zone around the needle insertion point is sterilized. A local anesthesia is administered to the location of the needle insertion point.
3. Procedure and real-time MRI-guidance
   An incision is performed at the needle insertion point in order to facilitate needle advancement and limit needle bending. The needle is then inserted and its advancement

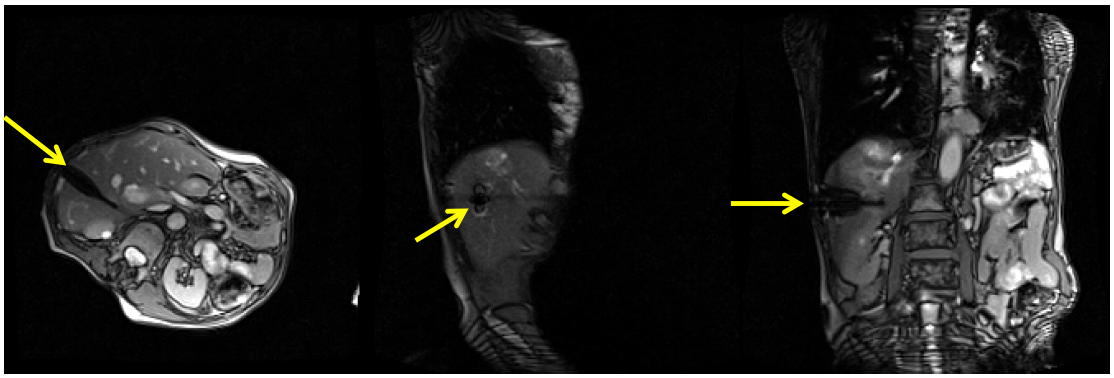


Figure 1.2: Images of a liver biopsy under MRI-guidance. The image planes at the left and right are orthogonal to each other and aligned to the main axis of the needle. The image plane in the middle is oriented perpendicular to the needle axis. The needle artifact is indicated with yellow arrows.

is monitored through the MR images. MR image plane position and orientation are dynamically manually updated for the physician to have a precise view of its needle advancement within the surrounding anatomy. Typically, two image planes are aligned to the needle main axis and orthogonal to each other in order to depict the anatomy surrounding the needle in a 3D-like fashion. Once the needle reaches the targeted region of interest, the treatment or sampling is performed. The needle is eventually retracted.

1.1.5 Current needs in MR-guided percutaneous procedures

Very little dedicated solutions are commercially available to facilitate the described clinical workflow, so that interventional MRI remains clinically limited to highly trained centers. Several difficulties remain for the clinical teams performing percutaneous procedures under MR-guidance.

The exact marking of the needle insertion point on the patient's skin is manual and can therefore be time consuming and prone to error. The insertion point determined in pre-operative planning images has to be located in the MRI coordinates. Two manual methods are used for patient skin marking of the entry point. The first method relies on the MRI laser normally used for isocenter definition relative to the patient. The patient table is manually moved so that the entry point lies underneath the laser plane. Then the right-left position of the entry point is measured with a ruler from the laser cross. For the second method, MR images of the insertion plane are continuously acquired while the physician palpates the patient's body until detecting his finger in the MR image. Then he/she translates his finger to the appropriate right-left position, before marking the entry point on the patient skin. Once the entry point is correctly marked, the practitioner relies on his/her own experience to introduce the needle with the correct angulation in order to reach the targeted lesion. Laser pointers are commercially available in CT but not for MRI so far.

Manual tracking of the needle insertion and real-time image plane alignment is a tedious task as the physician has no means for controlling the image planes. Interfaces allowing to control the MR images planes from within the MRI room exist, but the physician is occupied with the needle insertion and has thus not the possibility to manipulate them. The real-time image plane orientation and positioning is thus performed manually by a technologist at the MRI console and its time efficiency strongly relies on the experience of physician and technologist as a team as no communication is possible during image acquisition in a normal MRI scanner setup. A dedicated interventional MRI headphone system using optical microphones for avoiding interference with the image acquisition is commercially available (IMROC IR from Optoacoustics, Or Yehuda, Israel). Nevertheless, image plane alignment based on oral communication remains tedious. Fast and accurate localization of and automatic image plane alignment to the interventional tool are thus of high interest for interventional MRI. An overview over the research carried out in this domain is given in section 1.2.

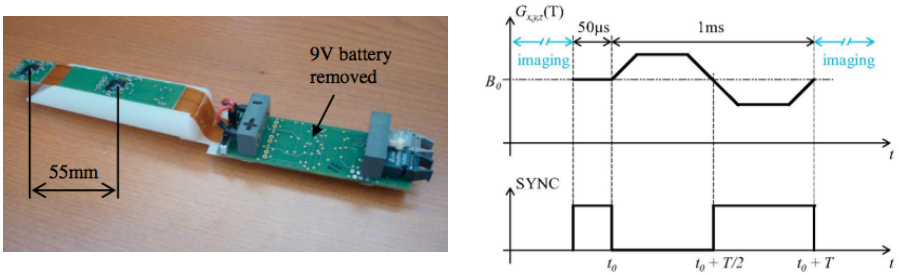
1.2 Literature review of tracking systems for interventional MRI

Numerous efforts have been made for automatic interventional tool detection and image plane alignment. The presented works can be grouped in active and passive strategies ([DLF06], [KKR11], [DWL02]). As this classification is not unique in the literature, classification for this review is realized as follows. Active strategies are defined as relying on additional electrical hardware connected to the MRI scanner or another measurement device in order to track interventional tools and adjust image planes automatically. In contrast, passive strategies rely on detection and tracking directly in the MR images based on contrast enhancement and do not use additional active devices. The following overview over the literature in this domain has been inspired from existing literature ([BW08],[KKR11]).

1.2.1 Active tracking systems

Active tracking systems rely on different principles such as gradient field measurements, tracking RF coils or optical tracking.

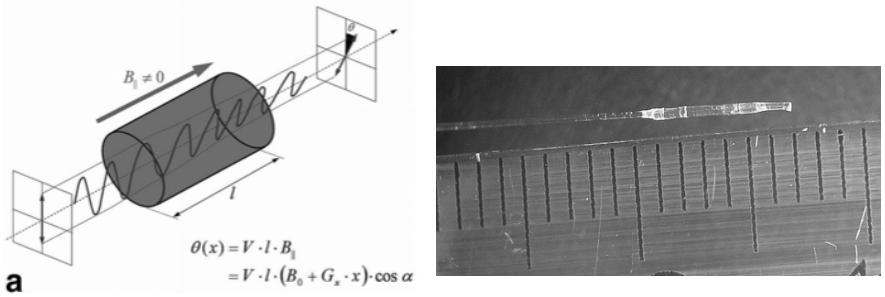
Spatially varying gradients that are applied for imaging or specifically for tracking can be used in order to determine position and orientation of the surgical device to which the sensor is attached. In [SKH+13] a Hall sensor is used for measurement of three specific gradients applied along the three axes of the MRI. The challenge in this approach is to accurately measure the magnetic gradients in the presence of the high permanent magnetic field of the MRI scanner. For this purpose, bipolar pulses are applied sequentially along each axis of MRI and a specific conditioning circuit is used (Figure 1.3). This signal conditioning circuit outputs one analog signal proportional to the magnitude of the applied gradient for every magnetic component. The output signal is transmitted via optical fiber cable outside of the MRI room. Currently, this tracking device could typically be attached to a needle holder and used for percutaneous procedures such as biopsies, ablations and infiltrations.



(a) (b)

Figure 1.3: Gradient measurement using a 3D Hall sensor. (a): The sensor and the conditioning circuit. (b): The dedicated bipolar magnetic gradient pulse and the synchronization signal used for gradient measurements. Reprinted from [SKH+13].

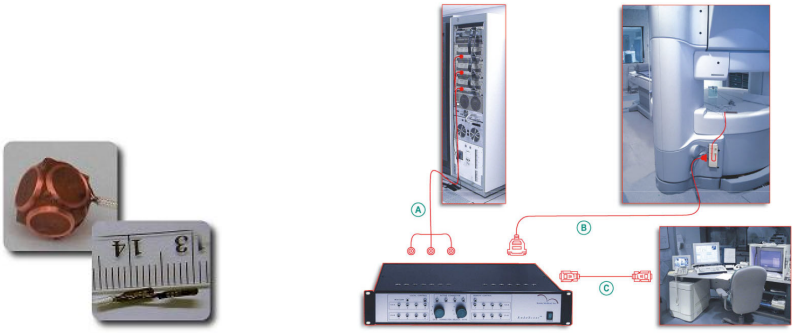
In [BUS+06], the Faraday effect is used for measurement of the gradient fields using a non-magnetic, non-electrically conducting optical sensor. The position of the device can be determined as the polarization plane in an optically active material depends on the local magnetic field. This type of sensor is typically usable for intravascular procedures as it is very small and can be included into a catheter (Figure 1.4).



(a) (b)

Figure 1.4: Gradient measurement using the Faraday effect. (a): The polarization plane in an optically active material is rotated by an angle θ, depending on the strength of the local magnetic field. The position of the sensor can thus be determined according to the transmitted light. (b): Prototype of a Faraday sensor connected to an optical fiber cable. Its small size allows to integrate it into a catheter. Reprinted from [BUS+06].

Another possibility for gradient measurement is the use of pick up coils measuring induced voltages during gradient switching. The EndoScout (Robin Medical, Baltimore, MD) and CathScout (Robin Medical, Baltimore, MD) systems are commercially available solutions using three orthogonal pick up coils for tracking of rigid interventional tools and catheters, respectively (Figure 1.5). The induced voltages are measured and compared to a gradient control signal for determination of position and orientation of the sensor.
Advantages of gradient-based tracking systems are that either no extra-time or only short extra-time for application of tracking gradients is consumed. Furthermore, the Faraday sensor



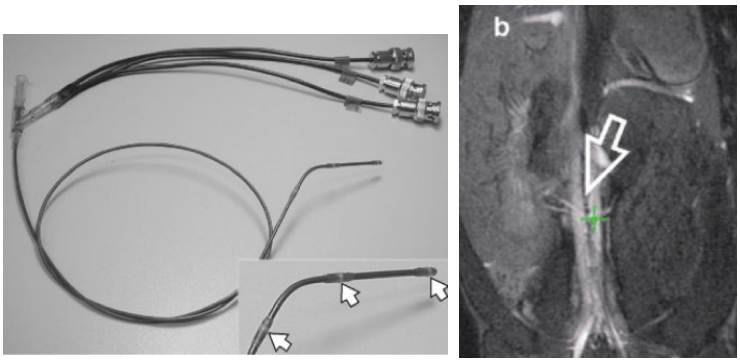
(a) (b)

Figure 1.5: Commercial solution from Robin Medical (Baltimore, MD) using pick up coils for gradient measurement. (a): Three orthogonal coils integrated into a cube sensor or into a microsensor for integration into a catheter. (b): Working principle of the EndoScout system: the system compares the emitted and measured gradient signals, determines position and orientation of the sensor and sends an image plane alignment command to the MRI console PC. Reprinted from [htt13b].

is completely made of non-conducting material, which prevents RF heating. For the Hall sensor and pick up coils, heating can be a problem and are a risk for the patient and the device. Furthermore the physical link between the sensors and a measuring device either through optical fiber or electrical cables restrain the free motion of the device. A further drawback of the Faraday sensor is the fragility of the used optical fiber cable.

Small tracking and profiling RF coils can be integrated in interventional devices and can be used for detection and tracking applications. The coils are directly connected to the receiver system of the MRI scanner. Their detection is realized with dedicated RF pulses that are interleaved with the imaging parts of an MRI sequence. In [BVZ+04] a catheter with three integrated RF coils is presented in combination with a dedicated real-time pulse sequence and a user interface allowing to automatically align the acquired image planes according to the performed catheter motion (Figure 1.6).
The advantage of this approach is the possibility to track a device automatically within a very short time. The coils can be used within catheters when miniaturized and for tracking of rigid interventional tools when attached to them. A possible risk of this technique is potential RF heating due to the conducting cables between the coils and the MRI receiver channels. Furthermore the detection of the coils requires dedicated tracking sequences.

Optical tracking systems use cameras and markers for tracking of holding devices for rigid percutaneous interventional tools used for biopsies, ablations and infiltrations. In [VBR+07] a stereoscopic camera system consisting in four CCD monochromic cameras is used for detection of an infrared LED attached to the distal end of a needle inside a C-arm MRI scanner. The presented system also provides a laser pointer for indication of the needle insertion point on the patient's skin (Figure 1.7). In [BSKO+02] a similar system using only one camera and

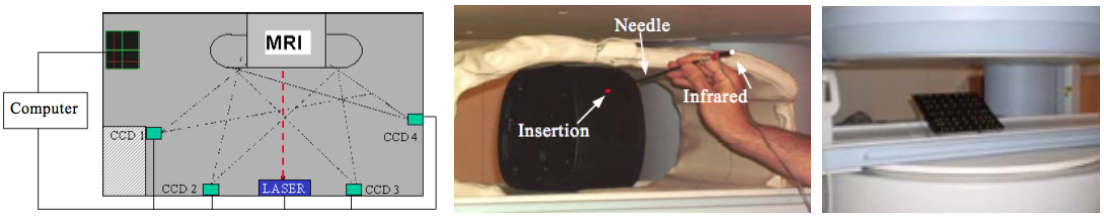


(a) (b)

Figure 1.6: Catheter tracking with RF coils. (a): Catheter with three integrated RF coils connected via coaxial cables to the MRI receiver unit. (b): In-vivo tracking experiment using an active catheter with one coil (green cross). Reprinted from [BVZ+04].

several markers attached to a biopsy tool is presented.
Advantages of optical tracking systems are their high accuracy and the ability to track the interventional tool outside the patient. No pulse sequence alteration is necessary as the optical system is totally independent from the MRI technology. Their drawbacks are the need for a line-of-sight link between marker and camera and thus their limited usability for closed-bore scanners. Furthermore a registration step is necessary between the camera system and the MRI scanner.

Common advantages of active tracking systems are their high accuracy and their independency from MR image contrast. Their drawbacks are the need for additional and often costly hardware that must be made compatible with the specific MRI set-up and pulse sequences.



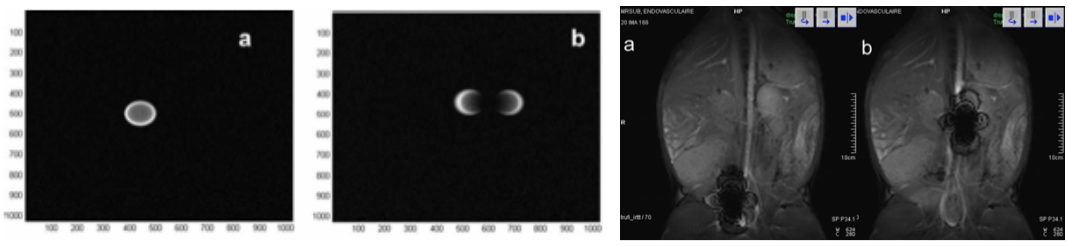
(a) (b) (c)

Figure 1.7: Optical tracking of needle insertion (a): Disposition of the CCD cameras and the laser pointing device in the MRI room. (b): Needle insertion point indicated by a laser pointer. The needle is tracked by the camera with a infrared LED attached to its distal end. (c): A calibration frame, visible for the cameras and for the MRI is used for permanent registration of the camera system with the MRI scanner. Reprinted from [VBR+07].

1.2.2 Passive tracking systems

Passive tracking approaches track interventional tools directly in MR images through signal enhancement or attenuation. The existing systems rely on paramagnetic markers or MRI contrast-agents, inductively coupled coils or application of currents in a wire.

Paramagnetic markers cause local field distortions of the permanent magnetic field and are depicted as dark spots on the MR images resulting from the signal loss in the region surrounding the marker. Due to this negative contrast they can be used for tracking of catheters within surrounding tissue. In [VdWBV01] dysprosium-oxide markers are used for tracking of a guidewire in an intra-vascular intervention.
A positive contrast can be obtained when a dephasing gradient is applied in the slice selection direction during image acquisition ([SVB03]). This dephasing gradient causes a signal loss in the image, and results in a bright marker on a dark image background.
Another possibility of tracking paramagnetic markers within an MR image is to use a selective excitation of paramagnetic equipotential curves caused by a magnetized marker such as a ferromagnetic sphere ([FMBM08]). An RF excitation tuned to the desired equipotential curve is applied and acquisition of three central k-space lines is sufficient for 3D tracking (Figure 1.8).

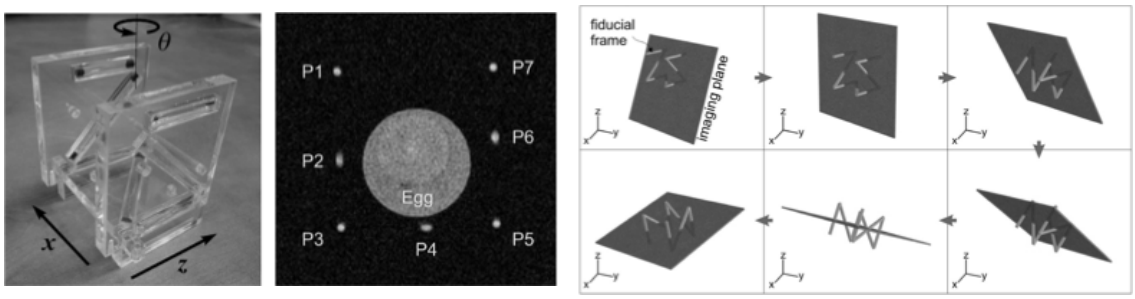


(a) (b)

Figure 1.8: Passive catheter tracking using a magnetized ferromagnetic marker and the selective excitation method. (a): Transversal (left) and sagittal (right) images of a ferromagnetic sphere marker. (b): In-vivo tracking experiment showing the initial (left) and the final (right) position of a catheter with a ferromagnetic marker at its tip. Reprinted from [FMBM08].

The advantage of paramagnetic markers is their small size. They can thus easily be used within catheters. Their drawback is the need for dedicated tracking images.

Contrast agents based on paramagnetic substances (e.g. gadolinium) are often used as passive markers for signal amplification in MR images due to their short T1 relaxation time. In [DSF+07], a stereotactic frame consisting in acrylic plastic with seven embedded contrast agent-filled glass cylinders is used for dynamic scan plane alignment (Figure 1.9). The unique intersection points of the glass cylinders with only one scan plane allow to determine the position and orientation of the frame. This stereotactic frame is presented in the context of a robot-assisted navigation system for MRI-guided prostate biopsy and brachytherapy but could also be used for tracking of a rigid interventional tool when attached to it.

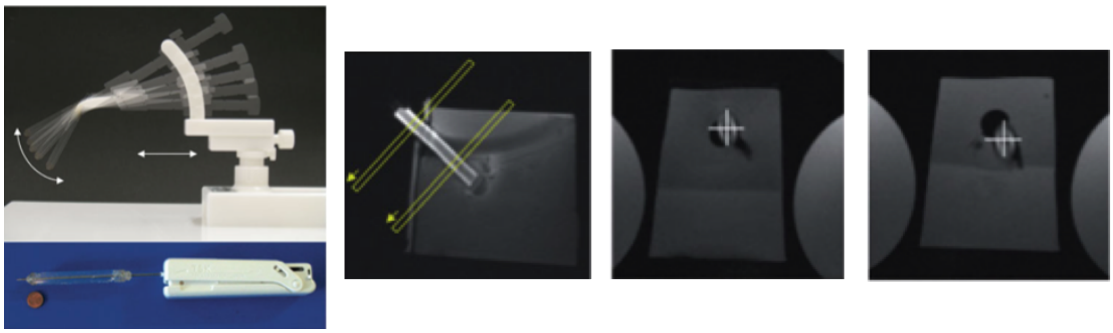


(a) (b)

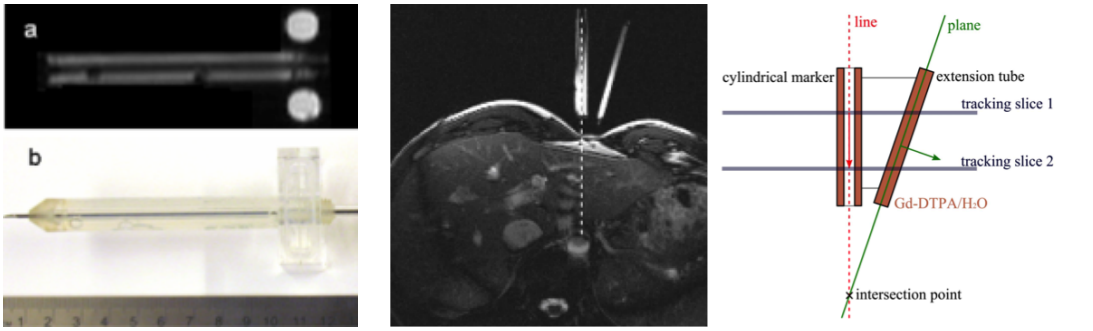
Figure 1.9: Passive tracking using a contrast agent-filled stereotactic frame. (a): Stereotactic frame made of acrylic plastic with seven embedded contrast agent-filled glass cylinders and a corresponding MR image of the cross section of the frame. (b): The image plane follows the motion of the stereotactic frame. Reprinted from [DSF+07].

The advantage of this technique is that the device can be tracked directly in clinical images without the need for dedicated tracking images. Its strong drawback is the bulkiness of the stereotactic frame.
In [dORB+08a], a commercially available prostate biopsy passive marker (PBPM) is used for implementation of an automatic image plane alignment workflow. The PBPM is a contrast agent-filled hollow cylinder through which the needle is passed. Two tracking-dedicated image planes orthogonal to the main axis of the PBPM are acquired. The ring shaped PBPM is then detected through phase-only cross-correlation on both image planes in order to align a third clinical image plane to the needle axis (Figure 1.10a,b). In [dORB+08b] this system is extended with an additional marker in order to track motions of the PBPM along its main axis (Figure 1.10c). In [MKdO+09] this marker is then adapted for percutaneous interventions through addition of a second cylindrical marker that is attached to the hollow cylinder at an angle of 15°. This allows to detect additionally the rotation of the device around its main axis (Figure 1.10d). The initial PBPM is used in [MKS+11] for tracking motions along its main axis without any hardware modifications. This is realized with a projection of the PBPM signal to its main axis in order to determine the device position in needle direction.
The advantage of these techniques is their simplicity as they are based on a commercially available marker. The drawback is that two dedicated tracking images have to be used for tracking and image plane alignment.

Inductively coupled coils are untethered RLC circuits that are tuned to the Larmor frequency of the MRI scanner. They rely on mutual inductance between aligned coils in order to transmit a signal from one coil to another. In this manner the coils couple inductively to the transmit coil of the MRI scanner and locally amplify the flip angle within their windings ([BCY96]). This allows to acquire dedicated low flip angle tracking images where only little signal is measured from tissue or liquids surrounding the coils and where the interior of the coils appear as bright points. A single loop coil is used in [QZK+05b] for visualization of a catheter. In [RME+08] two coils are attached to an endorectal probe for prostate biopsies (Figure 1.11). In [QZK+05a] a



(a) (b)



(c) (d)

Figure 1.10: Passive needle tracking using a commercially available prostate biopsy passive marker. (a): Device for prostate biopsies (top) and detail of passive marker with needle (bottom). (b): Corresponding MR images during phantom experiment with the needle aligned image plane (left). The two tracking planes (yellow) are depicted at the right. Reprinted from [dORB+08a]. (c): Extension of the marker allowing to track motions along its main axis. Reprinted from [dORB+08b]. (d): Extension of the marker allowing to track rotations around its main axis. Reprinted from [MKS+11].

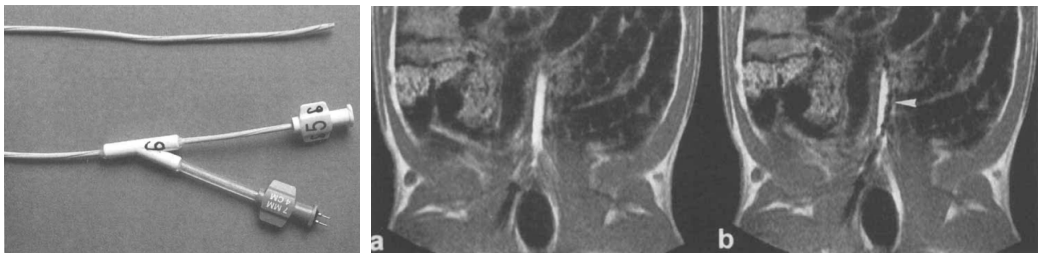
technique for wireless decoupling of the resonant circuit using two crossed diodes is presented in order to avoid image artifacts in the vicinity of the coil during conventional imaging.

(a) (b)

Figure 1.11: Passive tracking using resonant micro-coils. (a): Schematic (left) and photo (right) of a resonant micro-coil. (b): Endorectal device for prostate biopsies equipped with two micro-coils. Reprinted from [RME+08].

The advantages of inductively coupled coils are the simplicity of their detection and their safety against heating as their dimensions are normally chosen smaller than the wavelength of the MRI scanners. Drawbacks of this technique are that the coupling and thus the signal amplification is dependent on the orientation of the coils with respect to the permanent magnetic field of the MRI scanner and that their use is restrained to one MRI scanner field strength.

For visualization of a catheter during an intra-vascular procedure, Glovinski *et al.* ([GAB+97]) presented a system where a current is led through a copper wire loop inside a catheter. The current creates a local magnetic field that disturbs the homogeneity of the permanent magnetic field of the MRI scanner. Due to this disturbance the image zone within this local magnetic field appears dark and is well detectable among the surrounding tissues (Figure 1.12). An image subtraction technique can then be used in order to monitor the advancement of the catheter.



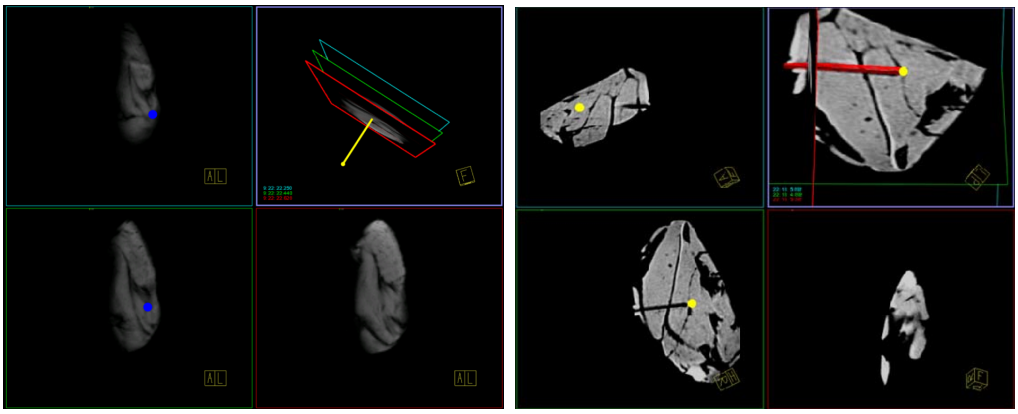
(a) (b)

Figure 1.12: Catheter tracking with a wire loop through which current is led. (a): The conducting wire is wound helically around the catheter. (b): Comparison of the catheter visibility in a pig study without current (left) and with 150 mA (right). Reprinted from [GAB+97].

Advantages of this technique are its simplicity and cost-effectiveness. The drawbacks are potential RF heating due to the conducting wire and the doubling of the image acquisition

time if an image subtraction technique is used.

Image-based needle localization without any additional devices has also been investigated in the literature. The needle is detected through the artifact it causes through local disturbance of the permanent magnetic field. Several works have investigated the artifact shape and its dependency on parameters such as the used MRI sequence and the needle orientation with respect to the permanent magnetic field ([DKE+05]) or the simulation and modeling of the needle artifact at the needle tip ([LHMT01]). In [DKE+05] a needle detection algorithm using the Hough transform for detection of linear structures in the MR image is presented. In [KGA04], the needle artifact is modeled as a rectangle in the difference image between a static baseline image and the image acquired during needle insertion. Needle localization is then performed through localization of the corresponding signal of this rectangle in k-space, which is a 2D $sinc$ function. Another technique relying on detection in a difference image in k-space is presented in [SMO+98] where the strongest signal intensity in k-space is identified as the needle tip. In [SRG+10] initial needle detection is performed on several parallel image planes orthogonal to the needle axis (Figure 1.13). The needle is detected using the background subtraction technique and feature analysis for candidate points. The needle tracking is then performed using three imaging planes, either two parallel and one orthogonal or two orthogonal and one parallel to the needle. A cylindrical needle model and a cost function allowing to evaluate the needle position are used for alignment of the imaging planes.



(a) (b)

Figure 1.13: Needle tracking using needle artifact. (a): Initial needle detection in three image planes orthogonal to the needle axis. (b): Needle path tracking (red) towards the target (yellow) with two parallel and one orthogonal image plane to the needle axis. Reprinted from [SRG+10].

The advantage of needle artifact detection is that no additional device is used. Drawbacks of this technique are that the needle artifact depends on the orientation of the needle with respect to the permanent magnetic field, and that the detection depends on the robustness and accuracy of the implemented image processing.

Common advantages of passive tracking methods are their simplicity and cost-effectiveness. Their drawbacks are the inherent need for the tracking device to be located inside the tracking images and the tracking update rate depending on the used MRI sequence. Also, most methods require dedicated MRI acquisition times, slowing down the update rate even more.

1.3 Thesis objectives

The active approaches presented in the literature need additional and often costly hardware whereas passive approaches usually need dedicated tracking images and can be bulky. The objective of this work is to develop a method for dynamic image plane alignment in real-time that would require only little additional instrumentation, be cost effective and minimize dedicated tracking time.
For these reasons, a passive workflow has been developed for automatic image plane alignment without the use of dedicated tracking images.

During interventions, physicians use fast MR imaging sequences acquiring image planes with a typical slice thickness of 4 mm. This thickness has proven to be a good trade-off between spatial resolution and image contrast. Indeed, a decrease of the slice thickness improves the spatial resolution along the direction normal to the image plane but lowers the signal intensity. Based on the slice thickness and the observable needle artifacts in the MR image with a width of up to 10 mm, the objective is to achieve an accuracy of 4 mm.

For dynamic research for the needle insertion point on the patient's skin, the image planes are required to follow the motions of a target over the patient's abdomen. The maximally achievable tracking speed is strongly correlated with the imaging frequency. However, in order to allow the physician to examine the entire width of the patient's abdomen (width assumed ~ 300 mm) in about 5 seconds, a tracking speed of 60 mm/s is desirable.

2 Workflow for automatic image plane alignment in interventional MRI

An automatic workflow for image plane alignment of two real-time image planes to an MRI marker is described in this chapter. First, the principle of the developed workflow is presented. Then, the architecture of the marker and its components is introduced. The MRI scanner, used scanner settings, the developed workflow and the image processing algorithms for marker detection are then presented. The remainder of the chapter describes the calculations for image plane alignment, the developed software architecture and the experimental evaluation of the workflow.

2.1 Principle of the workflow for image plane alignment

The clinical MR imaging sequence that is used during this work allows to acquire sequentially three MR image planes and to change interactively their positions and orientations. For percutaneous interventions, two of these image planes can be oriented orthogonally to each other in order to use them for monitoring of the needle advancement inside the patient's body. This set up allows to depict the needle path and the surrounding tissues in a "3D-like" manner. The idea for the presented workflow was thus to make use of these orthogonal image planes for an automatic image-based tracking technique using an MRI marker.
As the planes are orthogonal to each other and sequentially acquired, it is possible to detect the MRI marker in one image plane and to align the following orthogonal image plane according to this detection. This alignment step can be repeated for every image plane, and it is thus possible to implement a tracking workflow aligning the image planes according to the motion of the MRI marker. No dedicated tracking acquisition is necessary, as the detection of the MRI marker is performed in two clinically used image planes. In order to avoid a manual initialization step where a user aligns one of the two image planes to the MRI marker, one dedicated image is acquired in the beginning for automatic initialization of the tracking. Hence, the objective of the presented tracking workflow is to automatically align two MR image planes, orthogonal to each other, to an MRI marker.

2.2 MRI marker

Two types of MRI markers were used for the presented workflow. In a first approach, an MRI marker, composed of two micro-coils and a passive marker was assembled. Through the progress of work the passive marker only, without the micro-coils was used for the presented workflow.

2.2.1 Micro-coils

Micro-coils are resonant RLC-circuits (resistor, inductor and capacitor) that couple inductively with the radio-frequency (RF) coils of the MRI scanner, when tuned to its Larmor frequency. Due to their resonance, they locally amplify the MR signal within their windings and allow to be detected easily in low flip angle images ([BCY96]).
The micro-coils were built in a similar manner as in [RME+09] and tuned to 63.68 MHz, which is the Larmor frequency of the 1.5 T MRI scanner used during this work. They were wound around Eppendorf plastic tubes (Figure 2.1a) as they have appropriate dimensions (diameter: 7.5 mm, length: 16 mm), and the possibility to fill them with a liquid as signal source. A total of six turns of 0.4 mm diameter enameled copper wire are wound around the plastic tube in order to obtain an inductance of around 300 nH (total coil length 4 mm). A variable capacitor with a range from 8 to 45 pF is then soldered in series to the inductance. The micro-coil is tuned to the Larmor frequency using a loop antenna for inductive coupling and a network analyzer. Tuning is performed through change of the capacitance of the circuit. The two fabricated micro-coils have similar quality factors of about 60. This allows to use them in dedicated low-flip angle acquisitions as the measured MRI signal is amplified within their windings. The plastic tubes are filled with a contrast agent/water solution (Gd-DTPA 5 mM) as signal source for the MRI acquisitions.
Nevertheless, as the micro-coils couple inductively to the RF signals of the MRI scanner, the amplification of the MRI signal depends on their orientation with respect to the RF field.

2.2.2 Passive Marker and Assembled Test Device

The passive marker (Figure 2.1b) is a plastic tube (length: 90 mm, diameter: 35 mm) filled with the same contrast agent/water solution (Gd-DTPA 5 mM) as the micro-coils. The test device is made of two resonant micro-coils rigidly fixed at the endings of the passive marker (Figure 2.1c).

2.3 MRI scanner and scanner settings

All experiments during this work were performed in an open bore (short and wide) 1.5 T MRI scanner (MAGNETOM Aera, Siemens AG, Erlangen, Germany). The RF body coil was used in transmit/receive mode. An interactive, real-time, multi-slice TrueFISP sequence (Beat_IRTTT

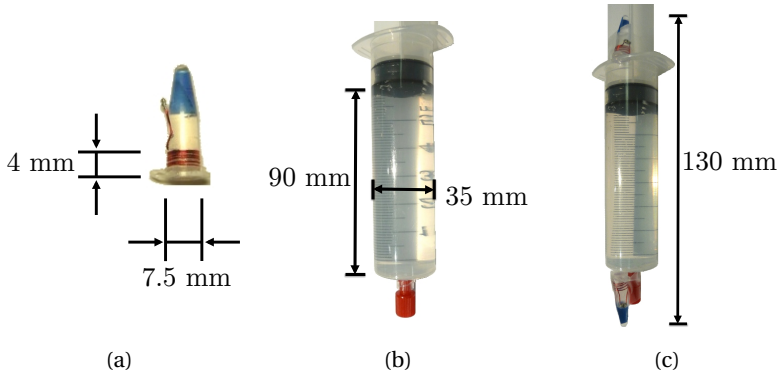


Figure 2.1: (a): Micro-coil wound around a plastic tube that is filled with contrast agent solution. (b): Passive marker consisting of a plastic tube that is filled with the same contrast agent solution. (c): Assembled MRI marker with two micro-coils at both endings of the passive marker.

[PBS+11], Siemens Corporate Research & Technology, USA) was used for image acquisition. The passive marker is used for detection in clinical real-time images. Imaging parameters of the clinical real-time sequence include: matrix 224×202, FOV 450 mm×450 mm, spatial resolution 2 mm×2 mm, slice thickness 4 mm, TE 2.2 ms, TR 5.3 ms, flip angle 50°, bandwidth 260 Hz/Px, temporal resolution 812 ms. Image acquisition time is 812 ms but the image update time is lengthened artificially to 1.2 s in order to allow for image transmission and processing, scan plane calculation and sending of commands to the MRI before the next image acquisition starts.
The two micro-coils are used for detection in dedicated low flip angle volume acquisitions with the same imaging parameters as the clinical real-time images except for their slice thickness (100 mm) and flip angle (1°). Due to the low flip angle, only little signal can be measured from anatomical structures. As a consequence, these volume acquisitions are not clinically usable by the physician and are only used for initial detection. In this work, these images are named "dedicated" in reference to the fact that their acquisition time is lost for the image-guidance of the procedure, and only useful for the automated image plane alignment workflow.

2.4 Image plane alignment workflow

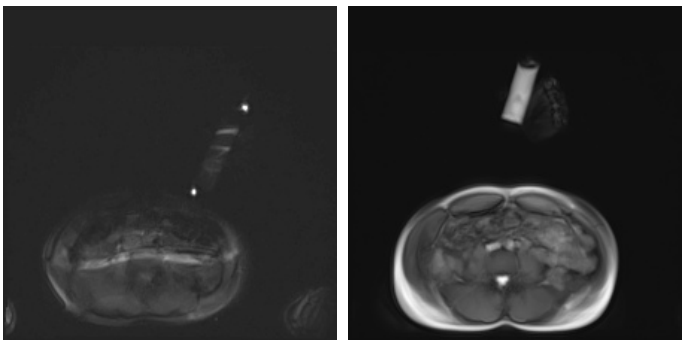
The objective of the presented workflow is to automatically align two image planes, orthogonal to each other, to the main axis of the MRI marker described in section 2.2. The orientation of the two scan planes can either be transversal (oblique) or sagittal (oblique) as the marker is assumed to point towards the MRI table. A purely horizontal marker orientation (aligned to x-axis of MRI scanner) is thus excluded.
The workflow consists in two stages: first an initialization stage where the initial position of the marker is detected in dedicated detection images and second a real-time tracking stage where the passive marker is detected in clinical real-time images.

2.4.1 Initialization

During the initialization stage, the initial position of the marker inside the MRI scanner is detected. It allows thus for automatic initialization and reinitialization of the image plane alignment workflow.
The initialization step can be performed in two different manners using either the marker described in section 2.2 consisting in a passive marker and micro-coils or the passive marker without the micro-coils.
Note that the marker detection for both approaches is realized in acquisitions with a slice thickness of 100 mm. During these acquisitions the MRI signal in the image volume is measured and projected to the image plane. As a consequence a full 3D pose is not reconstructable. The marker position is thus assumed in the middle of the image volume and the main marker axis to be coplanar with the image plane. After the initial detection, image plane parameters of a simple oblique sagittal plane aligned to the marker main axis are calculated and sent to the MRI scanner.

Initialization with micro-coils

A dedicated low flip angle transversal volume image is acquired at the MRI scanner isocenter. As can be seen in Figure 2.2a, these acquisitions contain only little signal from anatomical structures due to their low flip angle. In contrast to this, the signal inside the windings of the micro-coils is locally amplified and very well detectable. The two micro-coils are detected and the position and orientation of the marker are calculated.



(a) (b)

Figure 2.2: Typical initialization images using low-flip angle volume images (ST 100 mm, flip angle 1°) and micro-coils (a) or volume images (ST 100 mm, flip angle 50°) and the passive marker only (b).

Initialization with passive marker

First, a dedicated transversal volume image is acquired (Figure 2.2b) at the MRI scanner isocenter. Note that this volume image has a slice thickness of 100 mm and the same flip angle

as the clinical real-time images. All anatomical structures of the patient are thus depicted. Nevertheless, due to the high slice thickness all measured signal is projected to the image plane and thus superposed. As a consequence, these volume acquisitions are not clinically usable, either. The passive marker is detected in the image and its position and orientation are calculated.
Initialization without the micro-coils was investigated in order to simplify the marker design.

2.4.2 Real-time tracking

Clinically usable real-time image planes are used during the real-time tracking stage of the workflow. It starts typically with the acquisition of a simple oblique sagittal plane ($Image_1$) whose parameters have been sent to the MRI scanner at the end of the initialization stage.
The passive marker is then detected in $Image_1$, its 3D pose ($Pose_1$) is computed and the corresponding pose of a new simple oblique transversal image plane ($Image_2$) aligned with the marker is calculated. After acquisition of $Image_2$, the current pose ($Pose_2$) of the passive marker is again detected in the image in order to update position and orientation of a new sagittal oblique image plane ($Image_3$).
From then on, the acquisition alternates between two orthogonal image planes: a sagittal (odd image numbers) and a transversal plane (even image numbers). These two planes update mutually their position and orientation according to the passive marker detection (Figure 2.3).

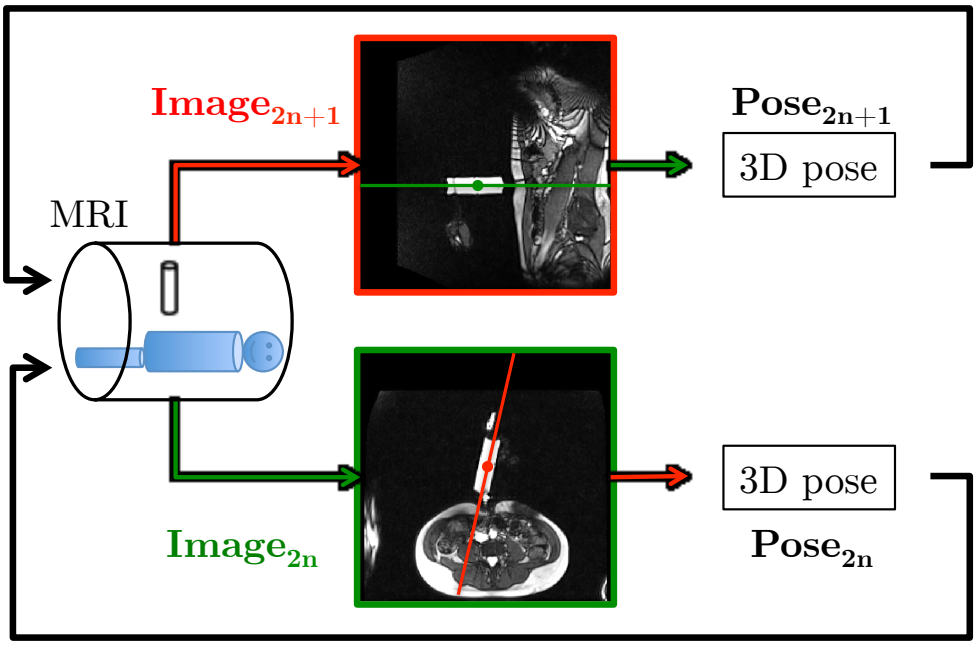


Figure 2.3: Principle of Workflow: two orthogonal image planes, transversal (green border) and sagittal (red border) are alternately acquired. The 3D marker pose is detected on the image plane. A marker aligned orthogonal image plane is calculated and a corresponding command is sent to the MRI console PC for acquisition of the next image plane.

2.4.3 Re-initialization

In case of a non-detection of the marker in two consecutive image planes, the initialization step is repeated. The only difference with the original initialization step being that the projection volume is centered on the latest detected z-coordinate of the marker.

2.5 Marker detection and image segmentation

The marker has to be detected on several types of images, such as low flip angle volume images, volume images and clinical real-time images. Note that all images acquired throughout this work have a size of 224×224 pixels.

2.5.1 Detection in low flip angle volume images

Detection of the marker in dedicated low flip angle volume images is realized through detection of the two attached micro-coils. Detection of the micro-coils is very robust using the phase only cross-correlation (POCC) technique presented in [CDD94] and [dORB+08a].

The phase correlation can be seen as:

- a translation in the Fourier (or frequency) domain of the cross correlation between two images in the spatial (or temporal) domain;
- a clever use of the Fourier Shift theorem, which states that a translation in the spatial domain induces a phase shift in the Fourier domain.

The detection of an object is achieved by a search of the translation (u_0, v_0) that maximizes the similarity (cross-correlation) between the image $i(u,v)$ and a template image of the object $t(u,v)$. The translation (u_0, v_0) is then the most likely location of the object in the image. Performing this search in the Fourier domain has computational cost benefits, if the images are large or readily available in the Fourier domain -as MR Images are. Indeed, the costly cross-correlation between two images becomes through a Fourier transform (DFT) a simple entry-wise multiplication of their Fourier transforms : $DFT(t * i) = DFT(t)^*.DFT(i)$.

If $T(k_u, k_v) = DFT(t(u,v))$ and $I(k_u, k_v) = DFT(i(u,v))$ are then respectively the Discrete Fourier transforms of the template image and the image to process, the phase only correlation is defined by :

$$POCC(k_u, k_v) = \frac{T(k_u, k_v)^*}{|T(k_u, k_v)^*|} \cdot \frac{I(k_u, k_v)}{|I(k_u, k_v)|} \quad , \tag{2.1}$$

with *, the complex conjugate operator.

The most likely position (u_0, v_0) of the object is then assessed by searching the peak of intensity

in the spatial normalized cross-correlation image $pocc(u,v) = DFT^{-1}(POCC(k_u,k_v))$:

$$(u_0,v_0) = \max_{u,v} pocc(u,v) \quad . \tag{2.2}$$

This property is commonly justified and illustrated with the phase correlation of one template image of size $M \times N$ and one image $i(u,v)$ that is an exact translation of the template image:

$$i(u,v) = t(u+u_0, v+v_0) \xrightarrow{DFT} I(k_u,k_v) = T(k_u,k_v)\, e^{i2\pi(\frac{k_u u_0}{M} + \frac{k_v v_0}{N})} \tag{2.3}$$

$$\begin{aligned} POCC(k_u,k_v) &= \frac{T(k_u,k_v)^*}{|T(k_u,k_v)^*|} \cdot \frac{I(k_u,k_v)}{|I(k_u,k_v)|} \\ POCC(k_u,k_v) &= \frac{T(k_u,k_v)^*}{|T(k_u,k_v)^*|} \cdot \frac{T(k_u,k_v)\cdot e^{i2\pi(\frac{k_u u_0}{M} + \frac{k_v v_0}{N})}}{\left|T(k_u,k_v)\cdot e^{i2\pi(\frac{k_u u_0}{M} + \frac{k_v v_0}{N})}\right|} = \frac{|T(k_u,k_v)|^2}{|T(k_u,k_v)|^2} \cdot e^{i2\pi(\frac{k_u u_0}{M} + \frac{k_v v_0}{N})} \\ POCC(k_u,k_v) &= e^{i2\pi(\frac{k_u u_0}{M} + \frac{k_v v_0}{N})} \quad . \end{aligned} \tag{2.4}$$

The translation can then be retrieved from this phase image by computing its inverse Fourier transform that exhibits a unit impulse located at the translation value:

$$pocc(u,v) = DFT^{-1}(POCC(k_u,k_v)) = \delta(u_0,v_0) \tag{2.5}$$

More complex approaches have been proposed to retrieve the translation directly from the phase image and thus avoid an inverse Fast Fourier transform [SOCM01]. Note that the image to proceed is rarely a simple translation of the template image but a complex scene with multiple objects. The image $pocc(u,v)$ -close to a normalized cross-correlation image- has then not only one peak at the object location but multiple peaks whose intensity is a measurement of similarity between the template and this image region.

Detection of the micro-coils is thus performed by detecting the two highest peaks in the signal distribution (see Figure 2.4b). The template image $t(u,v)$ has the same pixel matrix size $M \times N$ than the MR image to process. It is a black background with a white square at the center. The size of the square corresponds to a previously determined mean size of the micro-coils in the images.
Depending on their individual quality factors, the two micro-coils produce different signal intensities. However, the local signal increase is such compared to other structures in the low-flip angle image that the detection algorithm is robust to this signal variation. Finally, in order to determine precisely the position of the micro-coils, a weighted mean is calculated on the original image around the position of the 2 detected peaks.
The distance between the detected micro-coils is compared to their real distance on the marker in order to ensure correctness of the result. The detected micro-coils allow to determine position and orientation of the marker (see Figure 2.4c).

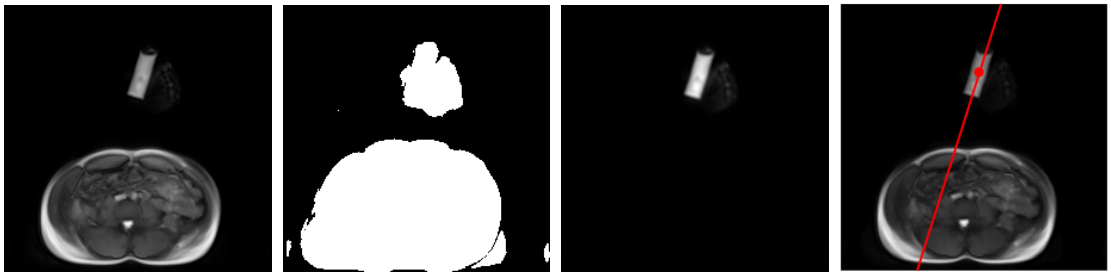
(a) (b) (c)

Figure 2.4: Segmentation of micro-coils in a low-flip angle volume image with (a) - the initial low flip angle volume image, (b) - its phase correlation with the template micro-coil image and (c) - the final detected position of the micro-coils with the derived orientation of the marker.

2.5.2 Detection in volume images

The detection of the passive marker in the volume images (Figure 2.5a) can be divided in two parts: first the segmentation and deletion of the patient's body and second the segmentation of the passive marker.
For the segmentation of the patient's body, a histogram-based thresholding is performed on the image. The threshold is determined in the following manner: the maximum of the histogram is detected (pixel intensity with the highest pixel number). Then, the local minimum after the maximum is detected and defined as the threshold. Note that the pixel intensity "0" (black pixel) is not taken into account for threshold determination as it can appear very often in the image background and varies tremendously depending on the patient. Thus, not considering it, makes the threshold determination more robust against patient size and position on the image.
After thresholding and application of a hole filling algorithm (Figure 2.5b), the biggest segmented object (patient's body) on the thresholded image is determined and the corresponding area is deleted from the original image (Figure 2.5c).
Then the segmentation of the passive marker is performed on the image without the patient's body. A thresholding operation using Otsu's method ([Ots79]) is applied. After thresholding, length and width of the segmented objects are determined. Then the difference between the found values and the real length and width of the marker is calculated for every segmented object. The object whose difference is the smallest is assumed to be the marker (Figure 2.5d). The difference between the detected marker dimensions and its real dimensions must not exceed a fifth of the real marker size.

Currently, initialization using micro-coils is more robust against failing detections due to the higher signal contrast in the dedicated low-flip angle images. Initialization using the passive marker only is prone to detection errors due to small contrast between the marker and the hand holding it in the projection images. Small distances between passive marker and the



(a) (b) (c) (d)

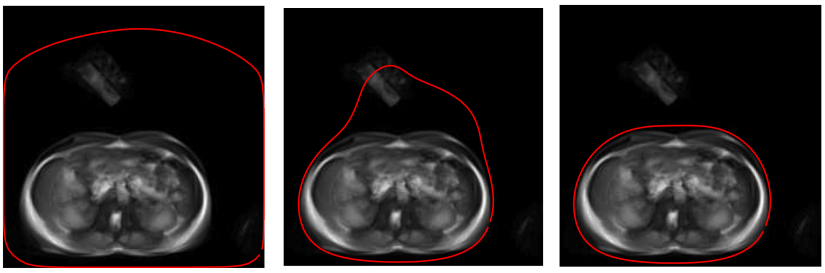
Figure 2.5: Segmentation of the passive marker in a volume image (a). A first threshold is applied (b) and the biggest object is deleted from the original image (c). A second threshold is then applied allowing to separately segment the bright marker from the hand holding it. The marker is then found through a size criterion (d).

patient's body can also cause problems for the proposed algorithm, as the marker and body could be segmented as one object during the first thresholding operation.

Nevertheless, the approach for initialization based on the passive marker gave promising results and will be further developed. A possible solution to the attachment of the marker to the body during thresholding could be the use of an active contour method with a shape constraint. The active contour could e.g. be constrained to keep a rather round shape without formation of corners. In this manner, the patient's body could be fitted without the marker lying within the contour.
A well known active contour approach is the snake algorithm ([KWT88]). This algorithm uses an energy-minimizing spline that is driven by external constrained forces and influenced by image-forces. The image forces are obtained from salient image features such as edges, lines and contours and allow to push the snake towards these features. External constrained forces allow to put the snake in the proximity of a desired local minimum. They can thus be based on inputs from a user-interface or in high-level interpretations. Furthermore, internal forces of a snake allow to impose smoothness constraints of the contour. First tests have been carried out using the snake algorithm with 100 contour points and a maximum of 150 iterations (Figure 2.6). The parameters for the different forces of the algorithm have been obtained experimentally.

2.5.3 Detection in clinical real-time images

The passive marker is used for detection in the clinical real-time images. These images typically depict the marker held by a physician over the patient's body (Figure 2.8a). The detection algorithm has thus to be robust against potentially segmented anatomical structures of the patient or of the physician holding the marker. First step of the detection algorithm is the image segmentation.

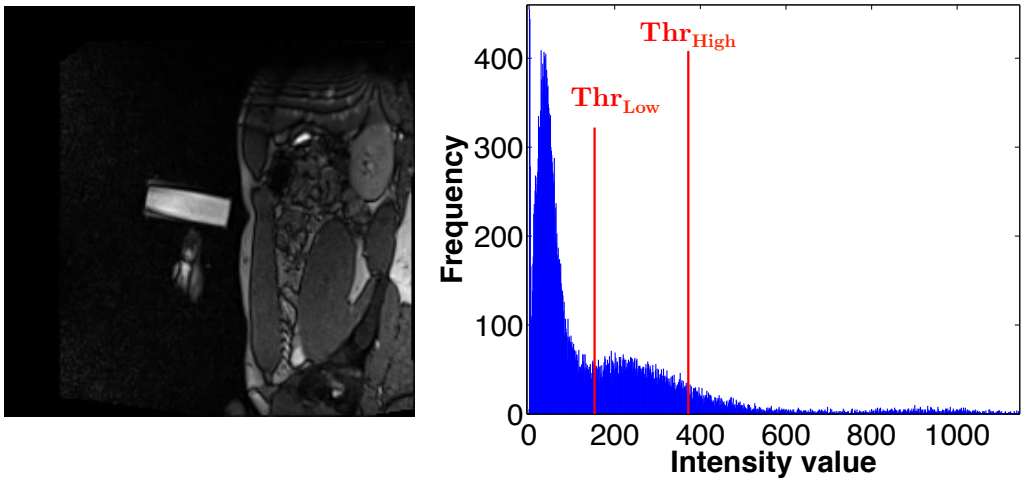


(a) (b) (c)

Figure 2.6: Passive marker detection in volume images using an active contour method. The contour is initialized near the edges of the image (a) and starts then moving over the image. Results after 50 (b) and 150 (c) iterations are presented, where the latter matches well the outer body limits.

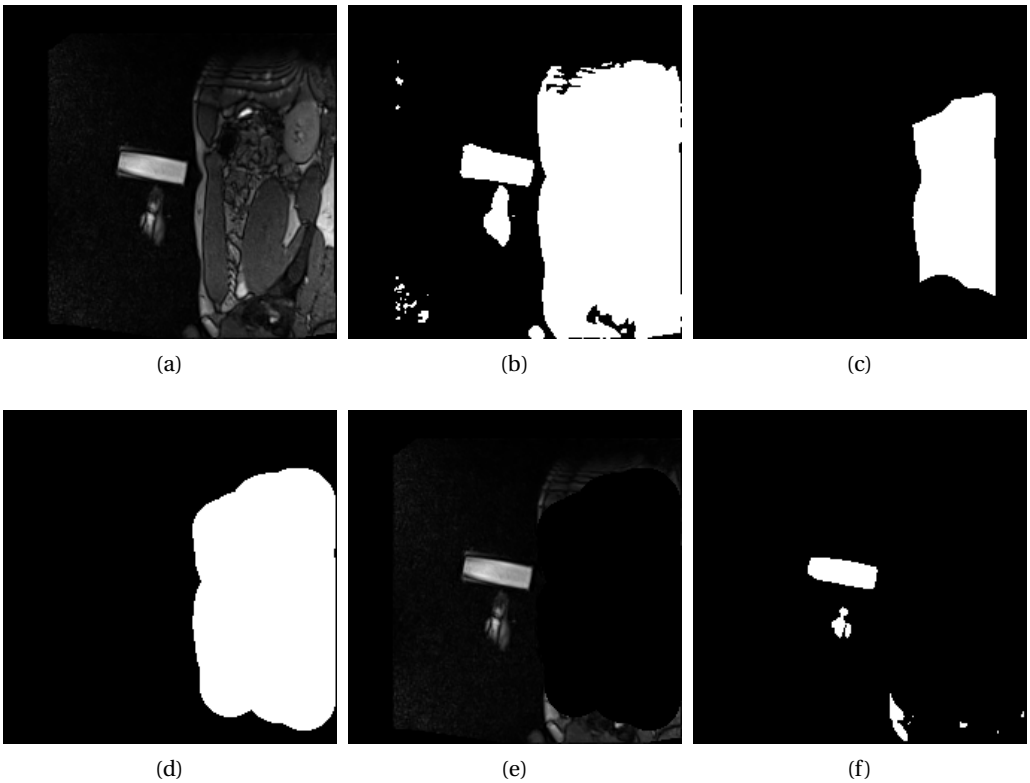
Image segmentation and labeling

Image segmentation starts with a Gaussian blurring step of the image in order to reduce image noise before the following thresholding operations. Thresholding is performed in two consecutive steps in order to avoid over-segmentation of the image and to facilitate the subsequent classification of the segmented objects. A first threshold is determined in order to separate image foreground and background. Foreground of the image consists in the patient's body, the marker and the anatomical structures of the person holding the marker (Figure 2.7a). This first threshold (Thr_{Low}) is established at the 0.65-quantile from the histogram of the image after Gaussian blurring. A second threshold (Thr_{High}) is determined at the 0.9-quantile of the same histogram (Figure 2.7b). Note that these histogram-based thresholds are also computed without considering pixel intensity "0", for the same reason as presented in section 2.5.2.
The first threshold is applied on the Gaussian blurred image. Morphological operations are then performed: first, a hole filling algorithm is applied, followed by a morphological opening (erosion, then dilation). Hole filling allows to close potential gaps inside the segmented body parts (Figure 2.8b). The opening step is realized with a structuring element in the order of 1/10th of the image dimensions. This rather big structuring element allows to delete segmented marker parts from the image while preserving the kernel of the patient's body during the erosion step of the opening (Figure 2.8c). The following dilation allows to rebuild the approximate shape and size of the patient's body (Figure 2.8d). The obtained image contains thus a mask for the segmented patient's body. This mask is inverted and then multiplied with the Gaussian blurred image in order to delete the major part of the patient's body from the scene (Figure 2.8e).
After deletion of the patient's body, the second threshold (Thr_{High}) is applied on the obtained image (Figure 2.8f). As this threshold is a lot higher than the first one, the hand of the person holding the marker and the passive marker are separated. The segmented objects are then labeled.



(a) (b)

Figure 2.7: Histogram-based determination of thresholds for the image depicted in (a). The thresholds are defined at the 0.65-quantile (Thr_{Low}) and the 0.9-quantile (Thr_{High}) of the histogram.



(a) (b) (c)

(d) (e) (f)

Figure 2.8: Image segmentation algorithm: The original image (a) is blurred and thresholded (b). A morphological opening (erosion, then dilation) step is then applied (c,d). The resulting image is inverted and multiplied with the blurred original image in order to mask out the principal body parts (e). A higher threshold (Thr_{High}) is then applied on this image in order to separate the marker clearly from the hand holding it.

Image classification

After labeling, the geometrical properties of the objects, such as position, intensity weighted position and mean intensity are determined. Further geometrical properties are determined using an ellipse fitting algorithm calculating orientation, axes length and eccentricity of the ellipses fitted to the segmented objects.
The classification of the objects is then achieved with a score system that is based on the previously determined geometrical features. According to the score, the objects are classified as marker (only one is expected) or non-marker objects. A score composed of sub-scores for size (A), intensity (B) and eccentricity (C) is calculated for every segmented object i.

The size score is defined by

$$A(i) = \begin{cases} \frac{Size(i)}{typSize} & \text{if} \quad \frac{Size(i)}{typSize} \leq 1 \\ 1 - \left|1 - \frac{Size(i)}{typSize}\right| & \text{if} \quad 1 < \frac{Size(i)}{typSize} \leq 1.8 \\ 0 & \text{if} \quad \frac{Size(i)}{typSize} > 1.8 \end{cases} \tag{2.6}$$

where $Size(i)$ is the size of the object in number of pixels and $typSize$ the expected size of the marker. As the size of the marker in the images depends on the intersection of the image plane with the marker, its typical size in the images is determined experimentally in an optimally marker-aligned image plane. The size score allows to calculate a maximal score for those objects that have the same size as $typSize$. If they are smaller or bigger their score diminishes. The size score is not directly set to zero when the object is bigger than $typSize$, as parts of the hand holding the marker or the micro-coils could be segmented with the object containing the marker.

The intensity score is the result of the division between the summed intensity of pixels ($Int(i)$) inside a segmented object and its size in the image:

$$B(i) = \frac{Int(i)}{Size(i)} \quad . \tag{2.7}$$

Thus this score is independent of the object size and represents the mean intensity of the object. Note that after all intensity scores have been computed with the equation here above, they are normalized to the interval $[0;1]$ through division with the maximal intensity score of all segmented objects. As a consequence the marker, which is normally the object with the biggest intensity-size-ratio in the image obtains the biggest intensity score. This score can be prone to errors, when a part of the hand/arm holding the marker is segmented as an object, as fat tissue is very intense in the used MRI sequence.

Each segmented object is fitted by an ellipse. The used ellipse fitting algorithm ([PM07]) is based on the calculation of raw image moments. A raw image moment (M_{pq}) is a weighted average of the image pixel intensities ($I(u,v)$) that allows to describe image objects after

segmentation:

$$M_{pq} = \sum_u \sum_v u^p \, v^q \, I(u,v) \qquad . \tag{2.8}$$

Typical geometrical properties that can directly be determined from raw moments are the sum of intensity values (M_{00}) and the centroid coordinates ($\overline{u} = M_{10}/M_{00}$, $\overline{v} = M_{01}/M_{00}$) of an object. In case of a binary image, M_{00} indicates thus directly the object size in pixels.
For ellipse fitting, the raw image moments are used for calculation of central moments (μ_{pq}), which describe the distribution of the pixel intensities of an object in a translation-invariant manner:

$$\mu_{pq} = \sum_u \sum_v (u-\overline{u})^p \, (v-\overline{v})^q \, I(u,v) \qquad . \tag{2.9}$$

This distribution is then approximated with an ellipse whose covariance matrix can be computed using the central moments:

$$cov\,[I(u,v)] = \begin{bmatrix} \frac{\mu_{20}}{\mu_{00}} & \frac{\mu_{11}}{\mu_{00}} \\ \frac{\mu_{11}}{\mu_{00}} & \frac{\mu_{02}}{\mu_{00}} \end{bmatrix} \quad . \tag{2.10}$$

Geometrical features of the ellipse (and thus of the object) can then be obtained from the eigenvectors (direction of ellipse axes) and eigenvalues (length of the axes) of the covariance matrix.

The eccentricity score is then calculated based on the object-fitted ellipse:

$$C(i) = \sqrt{1 - \frac{MinAx(i)^2}{MajAx(i)^2}} \qquad , \tag{2.11}$$

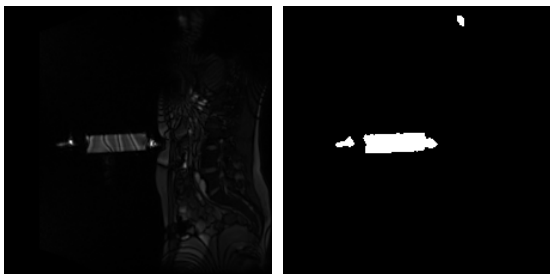
where $MinAx(i)$ and $MajAx(i)$ represent the length of the minor axis and major axis of the ellipse, respectively. This score is maximal if the object is elongated and zero when the segmented object has a circular shape.

The global score is obtained by building the weighted sum of the previously calculated scores:

$$S(i) = 0.6 \cdot A(i) + 0.3 \cdot B(i) + 0.6 \cdot C(i). \tag{2.12}$$

These weights have been determined experimentally. Note that the size score and eccentricity score are weighted with factor 2 compared to the intensity score. This is mainly due to potentially segmented body, arm or hand parts whose fat tissues can obtain big intensity scores due to their brightness with respect to other segmented objects in the image. Moreover, the marker can be subject to artifacts that influence heavily on its intensity (Figure 2.9a).
According to its global score, an object is classified as 'marker' if its score is above a defined threshold. If more than one object is classified as 'marker', only the highest score is retained as marker.

The presented image segmentation and classification methods have been tested for typical scenes with a patient's body and a marker held in hand. The algorithms have been verified with only four volunteers and have thus still to be tested on a larger sample. A marker with and without micro-coils has been used during tests. Due to their brightness in the clinical real-time images, the micro-coils are sometimes segmented as part of the marker, when they are not distant enough from the passive marker (Figure 2.9b). Since the micro-coils are symmetrically positioned at both endings of the marker, and since we ultimately aim at using the workflow without micro-coils, this work did not focus on separating the micro-coils from the passive marker during thresholding in the clinical real-time images.



(a) (b)

Figure 2.9: Challenges of marker segmentation and detection. The passive marker can be subject to artifacts influencing its intensity (a). Moreover, when micro-coils are not distant enough from the passive marker, they can be segmented as one object (b).

In addition, depending on the orientation of the hand/arm holding the marker more or less of these structures are depicted in the image and the algorithm can fail and detect these body parts as the marker. A perspective to solve this issue is the use of a shape-based detection of the marker in the images. Thus, an active contour method using a shape-prior could be used in order to segment the rectangular or elliptical form of the marker in the image. Initial tests of marker shape approximation with a superellipse ([RW95]) have been carried out but are not conclusive at this stage. The advantage of a superellipse parametrization of the contour is its ability to form different shapes such as rectangles, ellipses, diamonds or transitional shapes in between. Mainly the transition between a rectangle and an ellipse is interesting for approximation of the marker shape. Real-time ability of such an approach has to be evaluated.

Another challenge is the insufficient masking of the patient's body inside the images, which results in elongated and bright body parts that stay in the scene after segmentation and are subsequently evaluated through the classification algorithm. A possible solution to this problem is the segmentation of the patient's body with an active contour method as presented in section 2.5.2. The approximation of the patient's body in an initial image plane with a parameterized 3D shape would allow to project this shape on an arbitrarily oriented image plane in order to delete the body in every subsequent image plane.

2.6 Calculations for image plane alignment

In order to define the new position and orientation of an MR image plane, the used Siemens interface requires the following parameters:

- the 3D position of the new image center,
- the rotation between a pure transversal plane and the new image plane (in axis/angle representation).

They must be sent to the MRI console PC in order to align an image plane to a marker pose (position and orientation). As the presented workflow assumes no motion of the marker between two consecutive image planes, the detected 3D pose of the marker in the current image defines the 3D pose of the next image plane. Thus, the following calculation steps have to be performed in order to align the next image plane to the currently detected 3D marker pose:

- conversion of the 2D position of the marker (in image frame) to its 3D position (in MRI frame),
- calculation of the marker-aligned image plane orientation,
- determination of the rotation between a pure transversal image plane and the marker-aligned image plane.

The results of image plane position and orientation computation can then be sent to the Siemens interface for image plane alignment.

2.6.1 Computation of the new image plane position

In order to represent a point from an MR image in 3D, three coordinate systems are defined (Figure 2.10). The first coordinate system is two-dimensional and is linked to the upper left corner of the image ($\mathcal{F}_C : \boldsymbol{u}, \boldsymbol{v}$). The second is three-dimensional and is also linked to the MR image ($\mathcal{F}_I : \boldsymbol{row}, \boldsymbol{col}, \boldsymbol{nor}$) but has its origin in the center of the image. The frame of reference is three-dimensional, linked to the MRI scanner ($\mathcal{F}_M : \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$) and has its origin in the isocenter of the MRI. $\mathcal{F}_C$ is used for clarity, as most image processing libraries define the image origin in an image corner instead of its center. See Appendix A for further illustrations and details about the used coordinate systems. The transformation between a point in the MR image ${}^{C}P = [u, v]^T$ (in pixels) and its 3D correspondence ${}^{M}P = [x, y, z]^T$ (in mm) is given by the following equation in homogenous coordinates

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = {}^{M}\boldsymbol{M}_I \cdot {}^{I}\boldsymbol{M}_C \cdot \boldsymbol{S} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{2.13}$$

where:

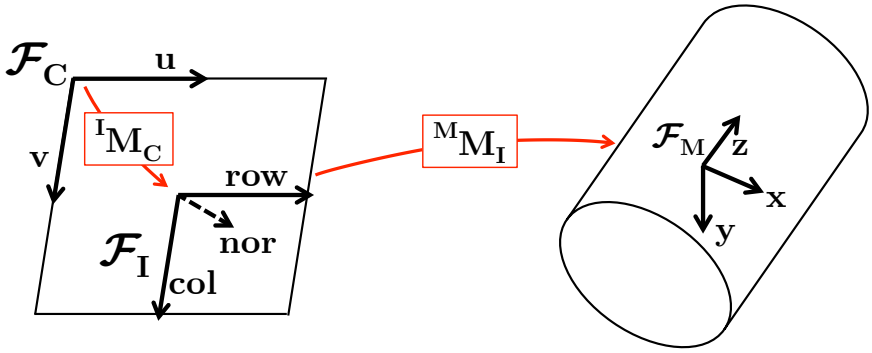


Figure 2.10: Coordinate systems for image plane calculations. Coordinate systems $\mathcal{F}_C$ and $\mathcal{F}_I$ are linked to the upper left corner and the center of the image, respectively. Coordinate system $\mathcal{F}_M$ is linked to the isocenter of the MRI scanner. The transformations between the implied coordinate systems are represented by the homogeneous matrices ${}^I\boldsymbol{M}_C$ and ${}^M\boldsymbol{M}_I$.

- ${}^M\boldsymbol{M}_I$ is the homogenous matrix describing rotation (${}^M\boldsymbol{R}_I$) and translation (${}^M\boldsymbol{t}_I$) between the image frame $\mathcal{F}_I$ and the MRI frame $\mathcal{F}_M$

$$ {}^M\boldsymbol{M}_I = \begin{bmatrix} {}^M\boldsymbol{R}_I & {}^M\boldsymbol{t}_I \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} row_x & col_x & nor_x & t_x \\ row_y & col_y & nor_y & t_y \\ row_z & col_z & nor_z & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} , \tag{2.14} $$

- ${}^I\boldsymbol{M}_C$ is the homogenous matrix describing translation (${}^I\boldsymbol{t}_C$) between the image frames $\mathcal{F}_C$ and $\mathcal{F}_I$

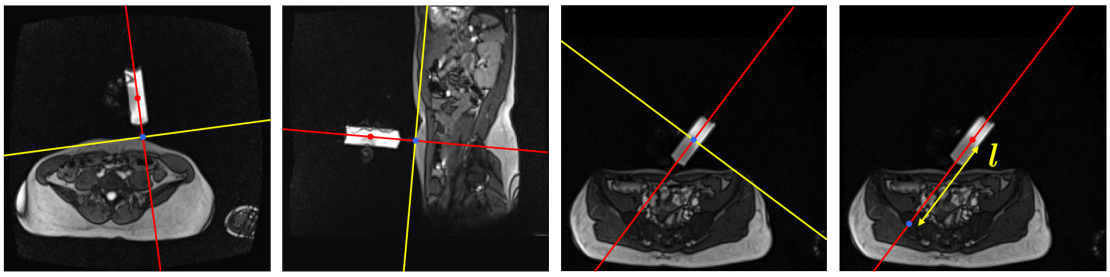
$$ {}^I\boldsymbol{M}_C = \begin{bmatrix} \boldsymbol{I}_{3,3} & {}^I\boldsymbol{t}_C \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -t_u \\ 0 & 1 & 0 & -t_v \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} , \tag{2.15} $$

- and S is the matrix describing the pixel size of the image

$$ \boldsymbol{S} = \begin{bmatrix} PxSize_u & 0 & 0 \\ 0 & PxSize_v & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} . \tag{2.16} $$

The parameters for the construction of matrices ${}^M\boldsymbol{M}_I$ and $\boldsymbol{S}$ are given by the information header that is attached to every MRI image. The parameters $PxSize_u$ and $PxSize_v$ describe the size of the pixels in horizontal and vertical direction respectively and are given in $mm/pixel$. The translation between the two image frames, given by t_u and t_v, is constant and corresponds to the half of image width and image height (in mm) respectively.

For simple scan plane positioning, the next image plane position is defined at the current 3D marker position ${}^{M}P = [x, y, z]^{T}$. As a consequence, the next image plane will depict the marker at its center, resulting in one image half that is empty (typically above the patient) and potentially the patient's body not entirely depicted in the image.
In order to avoid empty image space a first plane positioning strategy has been implemented: first the line corresponding to the marker orientation on the image plane is computed. Then the line normal to the first one and intersecting with the center of the image plane is computed. Their intersection coordinates on the image plane are calculated and defined as the point where the next image plane will be centered (Figures 2.11a and 2.11b). Nevertheless, this method has two flaws: first, when the marker is held near to the image plane center, the position of the next image plane will be defined very near to the center (Figure 2.11c). The second is that the distance between the detected marker position (red points in Figure 2.11) and the new image plane position (blue points) depends on the orientation of the marker on the image plane.
In order to stabilize the distance between the image plane center and the marker, another approach is possible: after computation of the line corresponding to the marker orientation on the image plane, a point on the same line at a pre-defined distance l from the marker position can be calculated and defined as the next image plane position (Figure 2.11d). In this manner the distance between the image plane center and the marker depends only on the marker motion.



(a) (b) (c) (d)

Figure 2.11: Image plane positioning using different strategies. The first strategy (a-c) defines the position of the next image (blue point) plane at the intersection between the marker axis (red) and an orthogonal line (yellow) passing through the image plane center. In (c), marker position and new image plane position coincide and the marker will thus be in the center of the next image plane. A second strategy avoids this by defining the next image plane position at a pre-defined length l from the marker position (d).

Depiction of the marker in the upper image half and stabilization of the distance between marker and the image plane center is of great interest for the workflow as the field of view of the acquired image planes could eventually be reduced to a minimum, depicting only patient's body and the marker. This field of view reduction would result in a reduction of the acquisition time by reducing the acquisition matrix, or in a higher image resolution by keeping the same acquisition matrix.

The first image plane positioning strategy has chosen in the beginning of this work. The second strategy has been developed subsequently and has not been tested yet.

2.6.2 Computation of the image plane orientation

In order to align the used image planes to a detected marker orientation, the planes are rotated around one of the principal axes of the frame linked to the MRI scanner. See Appendix A for further details about plane conventions and used coordinate systems.
As the marker is assumed to point towards the MRI table, a purely horizontal marker orientation (aligned to $x0z$-plane of MRI scanner) is excluded and only two cases have to be distinguished for the calculation of the orientation of the next image plane. When the current image plane is a sagittal oblique image plane, the next image plane is a transversal oblique image plane and vice versa. According to which of the two cases is dealt with, the orientation computation is slightly different.
The orientation of the new image plane is first defined by computing its orientation vectors $\boldsymbol{row}$, $\boldsymbol{col}$ and $\boldsymbol{nor}$. Then the rotation matrix between a pure transversal image plane and the desired image plane is calculated with Roll, Pitch, Yaw angles. Finally, the matrix representation is converted in the axis/angle representation in order to be sent to the Siemens interface.

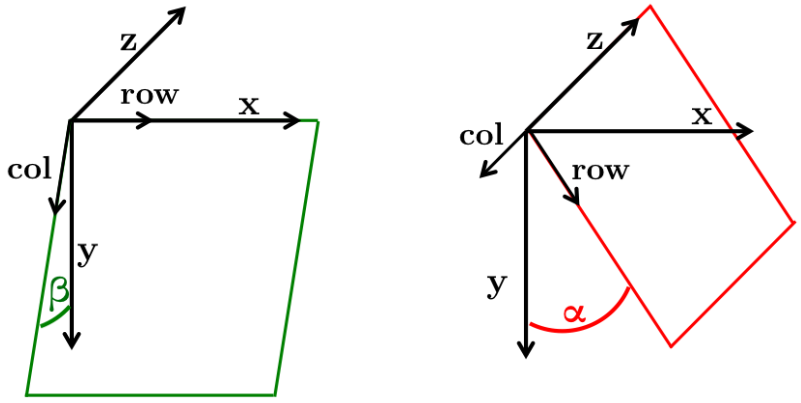
MR image plane conventions

The two image plane orientations used during the following computations are transversal oblique and sagittal oblique image planes (Figure 2.12). A simple oblique transversal image plane is rotated around the x-axis of the MRI scanner by an angle β. Correspondingly, a simple oblique sagittal image plane is rotated around the z-axis of the MRI scanner by an angle α. Their special cases with $\beta = 0$ and $\alpha = 0$ are referred to as pure transversal and pure sagittal image planes, respectively.

Computation of the orientation of a transversal oblique image plane

When the current image plane is sagittal oblique, the next one is a transversal oblique image plane. Thus, the orientation of the marker is detected in the current sagittal oblique image plane and the orientation of the next transversal oblique image plane is then computed accordingly.
As the orientation of a transversal oblique image plane is defined by the orientation vectors

$$\boldsymbol{row} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{col} = \begin{bmatrix} 0 \\ cos(\beta) \\ sin(\beta) \end{bmatrix} \qquad \boldsymbol{nor} = \boldsymbol{row} \times \boldsymbol{col} = \begin{bmatrix} 0 \\ -sin(\beta) \\ cos(\beta) \end{bmatrix} \quad , \tag{2.17}$$



(a) (b)

Figure 2.12: Transversal (a) and sagittal (b) oblique image planes depicted with their ***col*** and ***row*** vectors in the MRI frame.

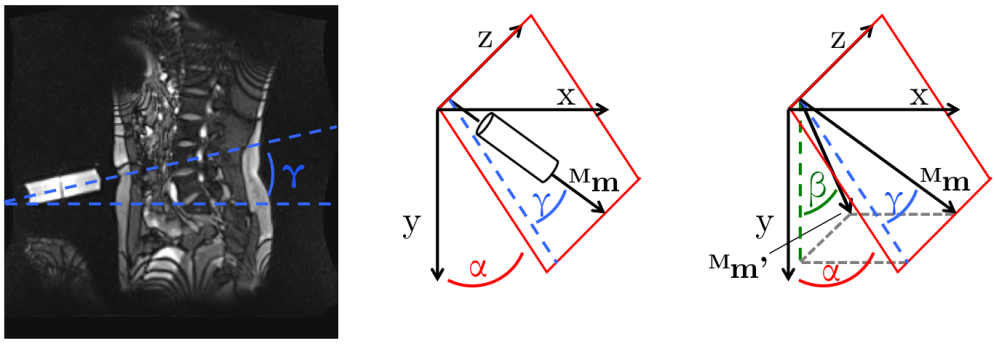
the rotation angle β has to be determined in order to define the plane orientation.
The angle β is defined between the y-axis of the MRI and ${}^{M}\boldsymbol{m}'$, with ${}^{M}\boldsymbol{m}'$ being the orthogonal projection of the marker-axis ${}^{M}\boldsymbol{m}$ on the plane $y0z$ (Figure 2.13). The marker axis vector is defined by the marker orientation in the image (γ) and by the orientation of the sagittal image plane (α). Vector ${}^{M}\boldsymbol{m}$ is thus obtained as

$$
\begin{aligned}
{}^{M}\boldsymbol{m} &= {}^{M}\boldsymbol{R}_{I}\,{}^{I}\boldsymbol{m} \qquad (2.18)\\
&= \begin{bmatrix} -\sin(\alpha) & 0 & -\cos(\alpha) \\ \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} cos(\gamma) \\ sin(\gamma) \\ 0 \end{bmatrix} \\
{}^{M}\boldsymbol{m} &= \begin{bmatrix} -cos(\gamma)\, sin(\alpha) \\ cos(\gamma)\, cos(\alpha) \\ -sin\gamma \end{bmatrix} ,
\end{aligned}
$$

where ${}^{I}\boldsymbol{m}$ is the marker axis represented in the image frame $\mathcal{F}_I$ and ${}^{M}\boldsymbol{R}_{I}$ is the rotation matrix between the image frame and the MRI scanner frame $\mathcal{F}_M$. Accordingly, the orthogonal projection to the $y0z$ plane is obtained with

$$
{}^{M}\boldsymbol{m}' = \begin{bmatrix} 0 \\ cos(\gamma)\, cos(\alpha) \\ -sin\gamma \end{bmatrix} . \qquad (2.19)
$$

Concerning the signs of the components of ${}^{M}\boldsymbol{m}$, note that angle γ is defined between the horizontal image axis and the marker axis in the image plane and rotates around the normal vector of the sagittal oblique image plane.



(a) (b) (c)

Figure 2.13: Projection of the detected marker orientation on a simple oblique sagittal image plane to pure sagittal plane in order to determine the orientation of the next transversal oblique image plane. (a): Detected marker orientation γ on simple oblique sagittal plane. (b): Representation of corresponding orientation in 3D when accounting for orientation α of sagittal oblique plane. (c): Projection ($^M\boldsymbol{m}'$) of the detected marker axis ($^M\boldsymbol{m}$) to the $y0z$-plane in order to determine the orientation β of the next transversal oblique image plane.

The angle β is then given by the angle between the y-axis and vector $^M\boldsymbol{m}'$. The cos and sin of this angle can be calculated based on the dot and cross products of vectors $\boldsymbol{y}$ and $^M\boldsymbol{m}'$. The dot product allows to calculate $cos(\beta)$ with

$$cos(\beta) = \frac{\boldsymbol{y}\ {}^M\boldsymbol{m}'}{|\boldsymbol{y}|\ |{}^M\boldsymbol{m}'|} = \frac{cos(\gamma)\ cos(\alpha)}{\sqrt{\left(cos(\gamma)\ cos(\alpha)\right)^2 + sin(\gamma)^2}} \quad . \tag{2.20}$$

The cross product of both vectors gives a vector that is collinear with the x-axis. In order to obtain $sin(\beta)$ directly, we project thus the cross product onto the x-axis:

$$sin(\beta) = \frac{\boldsymbol{y} \times {}^M\boldsymbol{m}'}{|\boldsymbol{y}|\ |{}^M\boldsymbol{m}'|}\ \boldsymbol{x} = \frac{-sin(\gamma)}{\sqrt{\left(cos(\gamma)\ cos(\alpha)\right)^2 + sin(\gamma)^2}} \quad . \tag{2.21}$$

Hence, the orientation angle of the next transversal image plane is then obtained with

$$\beta = atan\left(\frac{sin(\beta)}{cos(\beta)}\right) = atan\left(\frac{-sin(\gamma)}{cos(\gamma)\ cos(\alpha)}\right) \quad . \tag{2.22}$$

The transformation of a pure transversal image plane to an oblique transversal plane is then obtained by performing a simple rotation around the x-axis by angle β:

$$\boldsymbol{R} = \boldsymbol{R}_{(x,\beta)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & \cos(\beta) \end{bmatrix} \quad . \tag{2.23}$$

Calculations for orientation of a sagittal oblique image plane

When the current image plane is transversal oblique, the next image plane is a sagittal oblique image plane. Thus, the orientation of the marker is detected in the current transversal oblique image plane and the orientation of the next sagittal oblique image plane is then computed accordingly.
As the orientation of a sagittal oblique image plane is defined by the orientation vectors

$$\boldsymbol{row} = \begin{bmatrix} -sin(\alpha) \\ cos(\alpha) \\ 0 \end{bmatrix} \qquad \boldsymbol{col} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \qquad \boldsymbol{nor} = \boldsymbol{row} \times \boldsymbol{col} = \begin{bmatrix} -cos(\alpha) \\ -sin(\alpha) \\ 0 \end{bmatrix} \quad , \tag{2.24}$$

the rotation angle α has to be determined in order to define the plane orientation. The angle α is defined between the y-axis of the MRI and ${}^{M}\boldsymbol{m}'$, with ${}^{M}\boldsymbol{m}'$ being the orthogonal projection of the marker-axis ${}^{M}\boldsymbol{m}$ on the plane $x0y$ (Figure 2.14). The marker axis vector is defined by the marker orientation in the image (γ) and by the orientation of the transversal image plane (β). Vector ${}^{M}\boldsymbol{m}$ is thus obtained as

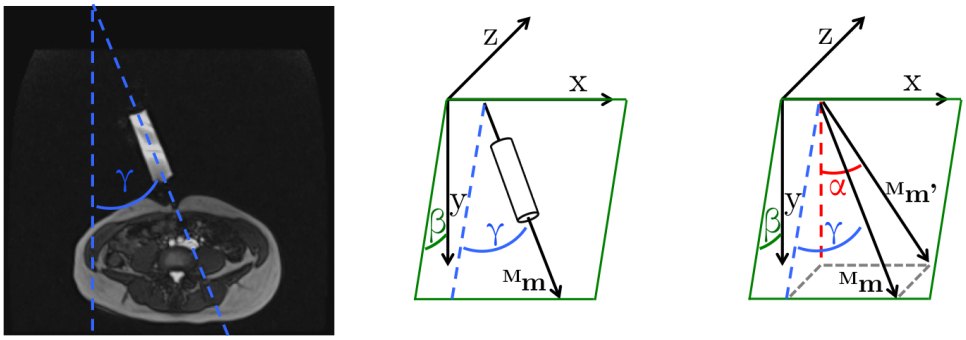
$$\begin{aligned} {}^{M}\boldsymbol{m} &= {}^{M}\boldsymbol{R}_{I}\, {}^{I}\boldsymbol{m} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\beta) & -\sin(\beta) \\ 0 & \sin(\beta) & cos(\beta) \end{bmatrix} \begin{bmatrix} -sin(\gamma) \\ cos(\gamma) \\ 0 \end{bmatrix} \\ {}^{M}\boldsymbol{m} &= \begin{bmatrix} -sin(\gamma) \\ cos(\beta)\, cos(\gamma) \\ sin(\beta)\, cos(\gamma) \end{bmatrix} \quad , \end{aligned} \tag{2.25}$$

where ${}^{I}\boldsymbol{m}$ is the marker axis represented in the image frame $\mathcal{F}_I$ and ${}^{M}\boldsymbol{R}_I$ is the rotation matrix between the image frame and the MRI scanner frame $\mathcal{F}_M$. Accordingly, the orthogonal projection to the $x0y$ plane is obtained with

$${}^{M}\boldsymbol{m} = \begin{bmatrix} -sin(\gamma) \\ cos(\beta)\, cos(\gamma) \\ 0 \end{bmatrix} \quad . \tag{2.26}$$

Concerning the signs of the components of ${}^{M}\boldsymbol{m}$, note that angle γ is defined between the vertical image axis and the marker axis in the image plane and rotates around the normal vector of the transversal oblique image plane.

The angle α is then given by the angle between the y-axis and vector ${}^{M}\boldsymbol{m}'$. As previously, the cos and sin of this angle can be calculated based on the dot and cross products of vectors $\boldsymbol{y}$



(a) (b) (c)

Figure 2.14: Projection of the detected marker orientation on a simple oblique transversal image plane to a pure transversal plane in order to determine the orientation of the next sagittal oblique image plane. (a): Detected marker orientation γ on a transversal oblique plane. (b): Representation of corresponding orientation in 3D when accounting for orientation β of transversal oblique plane. (c): Projection (${}^{M}\boldsymbol{m}'$) of the detected marker axis (${}^{M}\boldsymbol{m}$) to the $x0y$-plane in order to determine the orientation α of the next sagittal oblique image plane.

and ${}^{M}\boldsymbol{m}'$. The dot product allows to calculate $cos(\alpha)$ with

$$cos(\alpha) = \frac{\boldsymbol{y}\ {}^{M}\boldsymbol{m}'}{|\boldsymbol{y}|\ |{}^{M}\boldsymbol{m}'|} = \frac{cos(\gamma)\ cos(\beta)}{\sqrt{\left(cos(\gamma)\ cos(\beta)\right)^2 + sin(\gamma)^2}} \quad . \tag{2.27}$$

The cross product of both vectors gives a vector that is collinear with the z-axis. In order to obtain $sin(\alpha)$ directly, we project thus the cross product onto the z-axis:

$$sin(\alpha) = \frac{\boldsymbol{y} \times {}^{M}\boldsymbol{m}'}{|\boldsymbol{y}|\ |{}^{M}\boldsymbol{m}'|}\ \boldsymbol{z} = \frac{sin(\gamma)}{\sqrt{\left(cos(\gamma)\ cos(\beta)\right)^2 + sin(\gamma)^2}} \quad . \tag{2.28}$$

Hence, the orientation angle of the next transversal image plane is then obtained with

$$\alpha = atan\left(\frac{sin(\alpha)}{cos(\alpha)}\right) = atan\left(\frac{sin(\gamma)}{cos(\gamma)\ cos(\beta)}\right) \quad . \tag{2.29}$$

The transformation of a pure transversal image plane to a sagittal oblique plane is then obtained by performing a rotation around the x-axis by $-\frac{\pi}{2}$

$$\boldsymbol{R}_{(x,-\frac{\pi}{2})} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \quad , \tag{2.30}$$

followed by a rotation around the z-axis by $\frac{\pi}{2}$

$$\boldsymbol{R}_{(z,\frac{\pi}{2})} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.31}$$

and a final rotation around the z-axis by angle α

$$\boldsymbol{R}_{(z,\alpha)} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad . \tag{2.32}$$

Note that two consecutive rotations around the z-axis are applied for clarity. The rotation from a pure transversal image plane to a sagittal oblique plane is thus given by

$$\boldsymbol{R} = \boldsymbol{R}_{(x,-\frac{\pi}{2})}\, \boldsymbol{R}_{(z,\frac{\pi}{2})}\, \boldsymbol{R}_{(z,\alpha)} = \begin{bmatrix} -\sin(\alpha) & 0 & -\cos(\alpha) \\ \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & -1 & 0 \end{bmatrix} \quad . \tag{2.33}$$

The final rotation matrix $\boldsymbol{R}$ is converted to the axis/angle representation in order to be sent to the Siemens interface.

2.7 System architecture

Communication between the MRI console and an external PC is performed via an Ethernet connection and a proprietary Siemens protocol (ReModProt, Siemens Corporate Research & Technology, USA). The MRI console PC sends the acquired images to the external PC where the images are displayed and image processing and image plane calculations are performed. The commands for position and orientation of the image plane are sent to the MRI console PC for an image plane update. The communication interface on the external PC was implemented in C++ using open source libraries Qt, Insight Toolkit (ITK), Visualization Toolkit (VTK), OpenCV and boost.

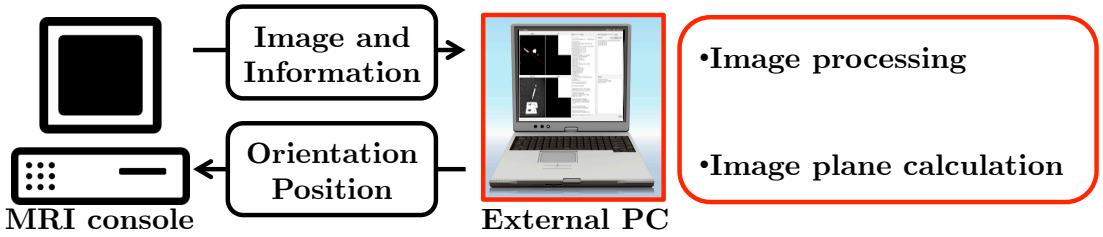


Figure 2.15: System architecture of the implemented workflow: MRI images are sent to an external PC via an Ethernet connection. Image processing and image plane calculations are performed and the created commands are sent to the MRI console in order to align the scan planes.

2.8 Simulations and Experiments

Tracking performance of the presented workflow is evaluated with the simulation program described in section 2.8.1 and the experiments presented in section 2.8.2.

2.8.1 MRI simulator

In order to evaluate the accuracy of the image plane alignment through well known and reproducible motions, a simulation program has been implemented. This program simulates motion of a cylindrical marker in 3D space. In order to simulate the function of the MRI scanner, it calculates the intersection of the currently simulated marker pose with alternately a sagittal oblique or transversal oblique image plane. A computer-generated MR image is created with black pixels as background and white pixels where the image plane intersects with the marker. The orientation of the image planes are controlled by the tracking software presented in section 2.7 with the same protocol as is used for command of the MRI console PC (Figure 2.16). The simulation program sends the images at the same image frequency to the tracking software as the MRI scanner, delayed only by a short time for simulation of the transmission time of the images. Only the image segmentation algorithm of the tracking software has to be adapted in order to detect the marker correctly in the simulated MR images. These changes are necessary as the computer-generated images do not contain the patient's body. The simulator creates images of the marker over a bottle phantom. The advantage of this simulation program is that it allows to simulate the performance of the tracking workflow with reproducible motions and speeds on one PC and with no need for the MRI scanner. One simulated motion was used throughout this work. Its motion range and speed were chosen in typical free-hand displacement. Marker velocity was fixed at 9 mm/s, while rotational speed was 2°/s. First the marker moves along the xz-diagonal in the MRI coordinates for 10s, with a continuous rotation motion along α. Then the motion takes place along the y-axis alone, with a rotation along β, for 6.67s (to ensure the same displacement of 6 cm along the y-axis, similar to the previous displacement along the xy-diagonal, that caused a displacement of

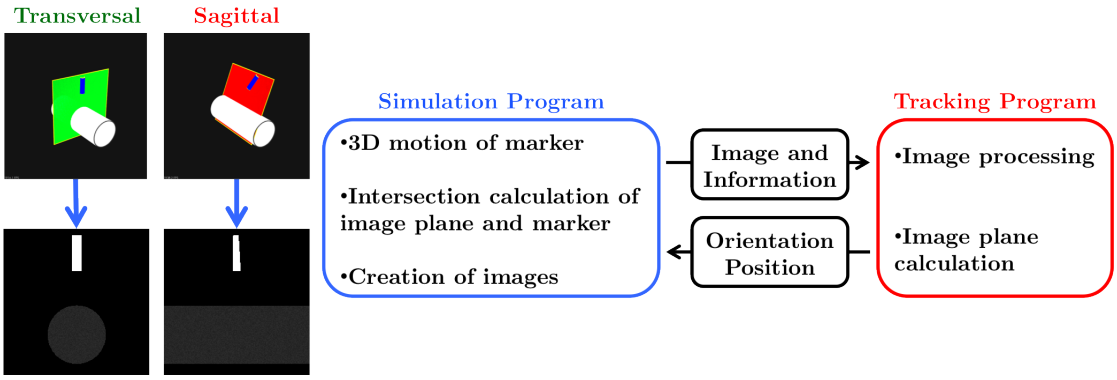
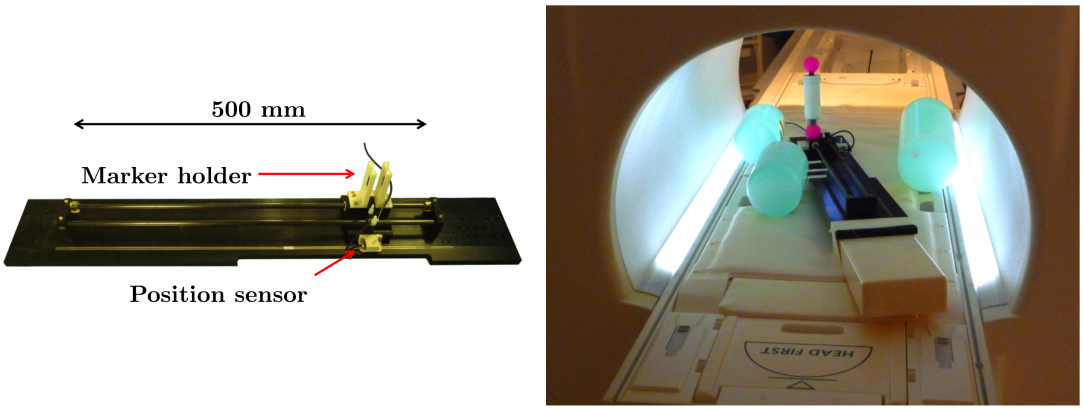


Figure 2.16: The simulation program simulates the MRI scanner and sends alternately transversal and sagittal oblique image planes to the tracking program. A 3D marker motion is simulated and computer-generated MR images are created according to the intersection between the requested plane from the tracking software and the marker pose.

6 cm along x and z). The diagonal xz motion was chosen as it is the worst case scenario for the tracking workflow. A maximum of pose components are rendered unavailable on both orthogonal image planes, through motion perpendicular to the individual image planes.

2.8.2 Experimental testbed

Further experiments were performed using an MRI compatible testbed inside the MRI scanner to assess the tracking effectiveness and accuracy. The testbed provides a ground-truth that can be compared to the measurements from the tracking workflow. This MRI compatible bidirectional translation mechanism allows to manually move a clamp on which the passive marker can be mounted. The position of the clamp is measured along the axis of translation using an optical position sensor with a resolution of 0.5 μm (Figure 2.17). The passive marker can be moved over a distance of 500 mm.
Sending of measurement data from the optical sensor to an external PC inside the MRI control room is realized via a programmable logic controller connected through a fiber optic cable. The programmable logic controller is put within a Faraday cage inside the MRI room in order to shield MR image acquisition from electromagnetic noise. In order to acquire nearly synchronous position measurements, a synchronization signal has been implemented in the used MRI sequence. The signal consists in a 2 ms pulse at the beginning of each MR image acquisition. In this manner, the high measurement acquisition rate of the optical sensor (1 KHz) can be synchronized with the slower MR image acquisition rate (~0.8 Hz). The translation measurement at half of the image acquisition time after the synchronization signal is compared to the corresponding marker detection of the workflow. This allows to obtain a better temporal matching between the measurements. The synchronization signal is output on a trigger output port in the technical room of the MRI. It is then sent to the programmable logic controller via a fiber optic cable and an fMRI Trigger Converter. A recapitulatory connection scheme of the experimental setup is presented in Figure 2.18.



(a) (b)

Figure 2.17: (a): MRI compatible testbed. (b): Testbed in MRI with attached marker.

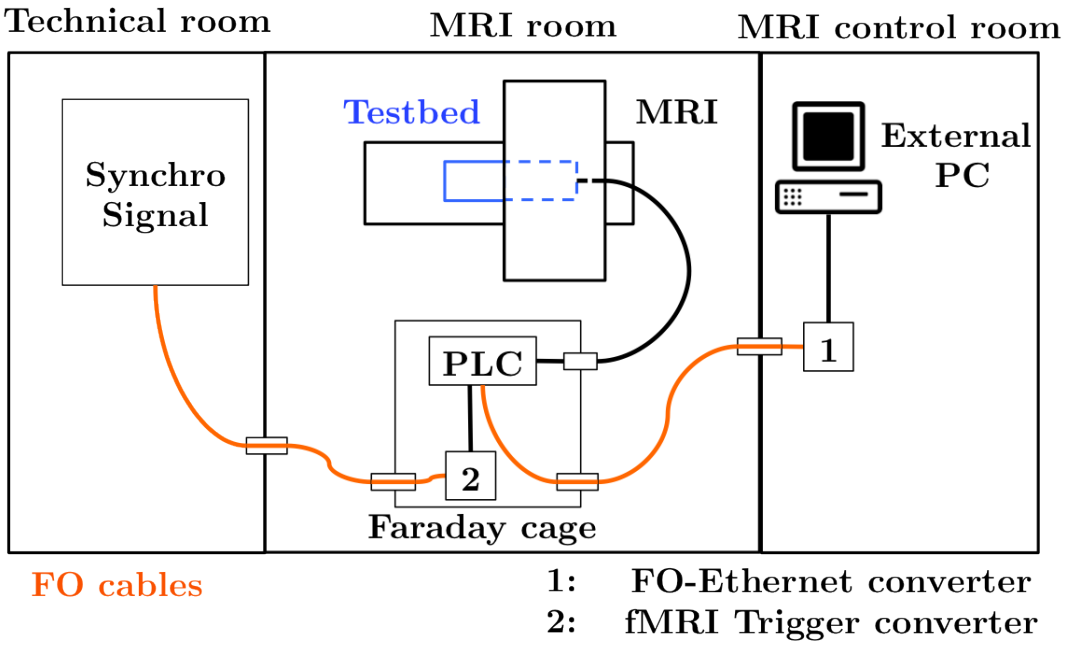


Figure 2.18: Connection scheme of the experimental setup using an MR compatible testbed and a programmable logic controller (PLC) inside a Faraday cage. All orange connections are fiber-optic cables. They pass through the walls of the Faraday cage and of the rooms through waveguides. The optical sensor is connected to the PLC with an RS-232 cable that passes inside the Faraday cage via a built-in filter of the cage.

In order to use the testbed measurements as ground-truth for the marker positions during the experiments, a calibration step is required between the optical measurement of the marker holder and the corresponding holder position in the MRI frame. As only a unidimensional translation is performed, the 3D line equation of the translation parametrized with the value of the optical sensor has to be computed in the MRI frame. For this purpose, the marker is mounted on the testbed clamp and its position is measured accurately in the MR images at different positions on the testbed. Then, a linear regression is computed on the detected marker positions in the MRI frame. From then on, the marker position can be determined in the MRI frame based on the translation measured with the optical sensor.
The testbed is positioned inside the MRI tunnel so that the marker performs a diagonal motion between the x-, y- and z-axis in the MRI frame (Figure 2.17b). All presented experiments have been performed on a single experimental run, hence ensuring identical testbed position. Two motion types were manually performed. First, the operator translated the marker trying to keep the velocity constant. That motion was repeated with increasing speed to try to reach the maximum speed. During motion, the operator looked at the in-room display screen in order to verify that the marker remained visible in the real-time MR images. Second, the operator moved the marker back and forth (oscillatory motion), with increasing speed for each cycle.

2.9 Results

Simulations and experiments as presented in section 2.8 have been carried out in order to evaluate tracking performance in terms of accuracy and speed.
The workflow to evaluate performs marker detection on sagittal oblique and on transversal oblique image planes, that are acquired alternately. Hence, the marker is detected on one plane and the next image plane is then automatically aligned to the detected marker pose.
The MR imaging frequency is set to 0.83 Hz throughout this work, due to the artificial lengthening of the MR image update time (see section 2.9.1). This lengthening is necessary as the workflow algorithms have to be performed on an external PC and the resulting command for image plane alignment has to be sent to the MRI console PC before acquisition of the next image plane.

First, a timing analysis of the real-time image plane alignment is performed and the robustness of the image processing algorithm for marker detection in clinical real-time images is evaluated. Then, the results for evaluation of the maximal tracking speed of the workflow are presented. Finally, the tracking accuracy of the workflow is assessed. Tracking accuracy is evaluated through comparison of the simulated or real pose and the detected 3D marker pose. The 3D marker pose consists in the position (x, y, z) of its center and the orientation of the marker main axis (α, β).

2.9.1 Timing analysis of real-time image plane alignment

In order to perform the image plane alignment of consecutive image planes in real-time, the processing times on the MRI scanner, MRI console PC and the external PC, as well as the communication times between the PCs have to be analyzed:

- Image acquisition time of the MRI scanner (t_{Acq})
- Transmission time of the image from MRI scanner to external PC via the MRI console PC (t_{Tr1})
- Time for image processing and image plane calculation (t_{Proc})
- Transmission and acknowledge time of the image plane alignment command sent from the external PC to the MRI scanner via the MRI console PC (t_{Tr2})

The image update time of an MR image is obtained through calculation of $t_{Update} = t_{Acq} + t_{Tr1} + t_{Proc} + t_{Tr2}$. A recapitulatory timeline of the processing and communication times is presented in Figure 2.19.

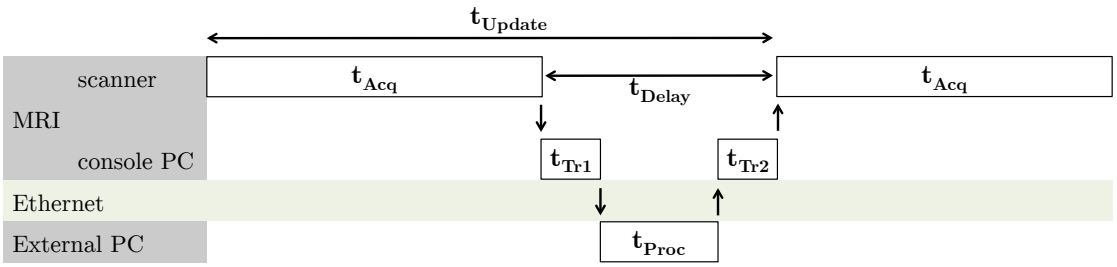


Figure 2.19: Timing schematic for image plane alignment. Image plane acquisition is performed with the scanner (t_{Acq}), image and image information are sent to the MRI console PC, that transmits the data to the external PC (t_{Tr1}). Image processing and image plane calculations are performed (t_{Proc}). The image plane alignment command is transmitted to the MRI scanner via the console PC (t_{Tr2}) before the next image plane acquisition.

In order to account for the image plane alignment command, the MRI scanner PC must obtain the command before the next image plane acquisition. For this reason, it is not possible to acquire the MR images directly one after another. In order to allow for image processing, image plane calculations and data transmission, a delay of $t_{Delay} = t_{Tr1} + t_{Proc} + t_{Tr2}$ has to be introduced between two image plane acquisitions.

A timing analysis of the network communications between external PC, MRI console PC and the scanner allowed to determine the transmission times.
For this purpose 100 image transmissions from MRI scanner to the external PC via the MRI console PC have been evaluated. The time between sending of the first Ethernet packet containing image data from the MRI scanner and reception of the last packet containing image data on the external PC is defined to be the transmission time t_{Tr1}. This transmission time of the MR image and its information (102,036 bytes for a 224 × 224 pixel image) has a mean

value of 7.8 ms with minimal and maximal values of 6.21 ms and 10.8 ms, respectively. Similarly, the transmission of 100 image plane alignment commands from the external PC to the MRI scanner via the MRI console PC have been evaluated in order to determine the transmission and acknowledge time of an image plane alignment command. The time between sending of the first Ethernet packet containing the image plane alignment command from the external PC and the reception of the last packet of the command acknowledge on the external PC is defined as the transmission and acknowledge time t_{Tr2}. This transmission time has a mean value of 48.3 ms with minimal and maximal values of 18.2 ms and 110.3 ms, respectively.

The time for image processing and image plane calculation t_{Proc} is defined between the moment of completed image reception within the workflow program on the external PC and the sending of the image plane alignment command. This time has been determined through evaluation of processing of 200 images on the external PC. For image processing and image plane calculation a mean of 233.2 ms, a maximal value of 284.6 ms and a minimal value of 198.8 ms have been found.

Mean values for transmission and processing times are not significant in this context, as every image plane alignment command should be accounted for. Particularly, the maximal and minimal values of t_{Tr2} and t_{Proc} differ in the order of ~90 ms. Hence, the times t_{Tr1}, t_{Tr2} and t_{Proc} are considered with their maximal values of 10.8 ms, 110.3 ms and 284.6 ms, respectively.

The image plane acquisition time (t_{Acq}) can be obtained from the acquisition software on the MRI console PC. For the MR imaging sequence used in this work, the acquisition time is 812 ms. The MR image reconstruction time could not be determined experimentally and is assumed to be short enough to be included within the acquisition time for the presented timing analysis.

The determined time values are summed up in Table 2.1.

	mean	max
t_{Acq}	812 ms	—
t_{Tr1}	7.8 ms	10.8 ms
t_{Tr2}	48.3 ms	110.3 ms
t_{Proc}	233.2 ms	284.6 ms
t_{Delay}	289.3 ms	405.7 ms

Table 2.1: Results of the timing analysis for real-time image plane alignment

According to the previously determined times for transmission and processing, the delay that must be introduced in order to account for every image plane alignment command would be 405.7 ms. With an acquisition time of 812 ms considering the used MR imaging sequence, this would correspond to a lengthening of the image update time from 812 ms to 1217.7 ms.
Note though, that the accuracy of these timing tests depends heavily on the system charge of the PC's OS during the measurements. Further experiments have thus been carried out

in order to determine experimentally the minimal delay allowing to account for each image plane alignment command. During these experiments, an image update rate of 1200 ms allowed to account for all image plane alignment commands. This corresponds to a delay of 388 ms. As 95% of the values for t_{Proc} are smaller than 270 ms, this delay is confirmed with the previously performed time measurements (270 ms + 10.8 ms +110.3 ms = 391.1 ms). Hence, for the experiments carried out during this work, an image update time (t_{Update}) of 1200 ms, corresponding to a delay (t_{Delay}) of 388 ms, was sufficient for performing real-time image plane alignment.

2.9.2 Robustness assessment of automatic real-time marker detection

The robustness of the image processing algorithm for automatic marker detection in clinical real-time images, presented in section 2.5.3, is evaluated with experimental image data. During experiments, the tracking workflow was performed with a person holding the marker in hand and moving it over the body of two volunteers.
For evaluation of the image processing algorithm, 208 images depicting the volunteer's body and the marker were used. The depicted marker area is required to be bigger than 40% of the marker intersection area in an optimally aligned image plane in order to be considered for the evaluation of image processing. This size criteria allows to avoid the evaluation of the marker detection algorithm on badly aligned images, i.e. when the tracking is not effective and the marker is hardly visible in the images.

First, the automatic detection algorithm is applied to the images. During a post-processing step, the marker detection is then performed manually within the same images. For manual segmentation, a polygon consisting in four points is used in order to define the marker shape. The result of the automatic detection is then evaluated with respect to the manual segmentation. The automatically detected marker position is considered correct, if it lies within the boundaries of the manually segmented marker area.

The success rate of the automatic marker detection is 96.15%, in images where at least 40% of the marker is visible, which corresponds to a successful detection in 200 of 208 images (Table 2.2). In five images (2.4%), the automatic marker detection algorithm could not detect the marker. In three images (1.4%), the marker was detected at wrong positions within the bodies of the volunteers.

	Correct detection	Incorrect detection	No detection
Number of images	200	3	5
Rate	96.15%	1.4%	2.4%

Table 2.2: Statistical performance of the marker detection algorithm when at least 40% of the marker are visible within the images.

Failed detections are mainly due to a failing deletion of the patient's body from the image. If the threshold Thr_{Low} is chosen too low, the mask for deletion of the patient's body is too big and may contain parts of the marker. As a consequence the marker is deleted with the patient's body and the detection fails. Note that, the false positives were detected on images where the marker signal was not very bright or affected by motion artifacts and where its size was just above the lower limit.
In general, it can be concluded that the algorithm can be greatly improved with a better and more appropriate segmentation of the patient's body.

2.9.3 Evaluation of the maximal tracking speed

The maximal tracking speed of the workflow was determined with both, simulations and the experimental testbed.

Simulations

Before determining the maximal tracking speed through simulations, a theoretical value based on the marker and image plane geometry is calculated. In order to theoretically determine the maximal trackable motion speed, this calculation does not account for motion artifacts due to fast marker motion in the created MR images, or command delays due to transmission times between PCs. The maximal trackable motion speed depends on several factors, such as

- image acquisition time of the MRI scanner (t_{Acq})
- transmission time of the image from MRI scanner to external PC via the MRI console PC (t_{Tr1})
- time for image processing and image plane calculation (t_{Proc})
- transmission and acknowledge time of the image plane alignment command sent from the external PC to the MRI scanner via the MRI console PC (t_{Tr2})
- marker diameter (d_{Mark}),
- slice thickness (d_{Slice}),
- the marker motion considered with respect to the image planes.

In order to evaluate the theoretical value, a motion at an angle of 45° between a transversal (Figure 2.20a) and the following sagittal image plane (Figure 2.20b) is considered, which is the worst case scenario for the presented workflow. As previously described, the sagittal plane is aligned to the position of the marker detected in the transversal image plane. In a simplified manner, it is assumed that the marker position cannot be determined on the sagittal image plane when the marker is not intersecting the plane.
It is thus possible to calculate the distance $\left(d_{max} = \frac{d_{Mark}}{2} + \frac{d_{Slice}}{2}\right)$ at which the marker is no longer detectable (Figure 2.20c). The maximal tracking speed can thus be calculated with

$$v_{max} = \left(\frac{d_{max}}{t_{Acqu} + t_{Tr1} + t_{Tr2} + t_{Proc}}\right) \frac{1}{cos(\frac{\pi}{4})} \quad . \tag{2.34}$$

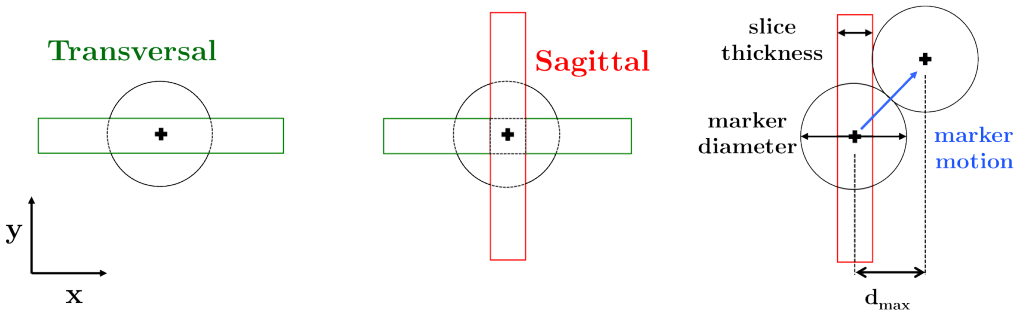


Figure 2.20: Schematic illustration of the calculation of the theoretical maximal tracking speed. (a): Marker (black) and transversal image plane (green) position at time step k, depicted in the x-0-y plane of the MRI frame. (b): Marker and sagittal image plane position (red) at time step $k+1$, when there is no motion of the marker. (c): Marker and sagittal image plane position at time step $k+1$, when the marker moves at 45° between sagittal and transversal image plane.

The marker diameter is 35 mm but the MR-measurable section of the liquid inside is 33 mm (marker wall = 1 mm). The slice thickness of the clinical sequence is 4 mm and the acquisition time is 812 ms. The experimentally determined delay for transmission times (t_{Tr1}, t_{Tr2}) and image processing and plane calculation time (t_{Proc}) is 388 ms (see section 2.9.1), which results in an image update rate of 1200 ms.
The maximal theoretical tracking speed is thus 21.8 mm/s. Note that this theoretical value is calculated for a worst case scenario, i.e. a 45° motion between sagittal and transversal image plane.

The marker motion described here above was then simulated and tracked in order to determine the maximal tracking speed through simulation. The maximal tracking speed was determined by performing simulations at different marker speeds until finding the maximum speed value for which the marker can be tracked. The simulations were carried out with an image update rate of 1200 ms. The found maximal tracking speed is 19.5 mm/s which is slightly lower than the calculated value.

In order to evaluate the tracking performance regardless of current hardware, software and MR imaging constraints, a hypothetical scenario with a a higher image update rate was simulated. This scenario is based on an image update time of 500 ms including communication and image processing times, allowing to evaluate the tracking performance of this approach at an imaging frequency of 2 Hz.
The maximal tracking speed is then determined in the same manner as before, with a diagonal motion between transversal and sagittal image planes. A maximal tracking speed of 48.5 mm/s was found for the hypothetical scenario. This value, which is more than twice higher than for the previous simulations (image update time of 1200 ms), can be explained with the higher imaging frequency of 2 Hz. As the tracked motion is sampled more often, higher motion

speeds can be followed.
Such high image update frequencies are expected through optimization of the used MR imaging sequence (fast parallel imaging, compressed sensing or non-cartesian acquisition schemes), implementation of the workflow on a single scanner PC (no communication times between external and MRI scanner PC) and optimization of the image processing algorithms (reduction of computation times).

Experiments

An experiment for determination of the maximal tracking speed was carried out. For this purpose an oscillatory motion with increasing speed is manually performed on the testbed, in order to determine the speed at which the marker cannot be tracked anymore. The translation curve of the testbed is depicted in Figure 2.21. The determined translation speed on the last edge before the marker cannot be detected anymore is assumed to be the maximal tracking speed. The obtained value for the presented workflow is 8.81 mm/s, which is smaller than the value found during simulations.
The difference can be explained with the fact that the simulator creates computer-generated images, where the signal intensity of the marker is independent of the intersection volume between image plane and marker. In real MR images, the signal intensity varies though and depends on the intersection between marker and plane. As a consequence, the detection in the computer-generated images works even for small intersections of marker and image plane which is not the case for real MR images. In addition, no motion artifacts are simulated within the computer-generated images. Hence, image quality is not degraded with motion in the simulations.

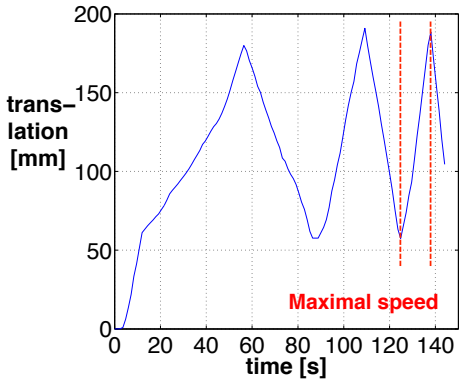


Figure 2.21: Experimental determination of the maximal tracking speed. The oscillatory motion is performed with increasing speed until the marker cannot be tracked anymore. The last edge of the motion before the tracking is interrupted is used for maximal speed computation, which is here 8.81 mm/s.

Nevertheless, the determined maximal tracking speed is approximative as the real maximal speed is difficult to determine experimentally without the possibility of performing a motion

at a repeatable and well known speed.

2.9.4 Accuracy assessment

As for the evaluation of the maximal tracking speed, the tracking accuracy of the workflow was also determined with both, simulations and the experimental testbed.

Simulations

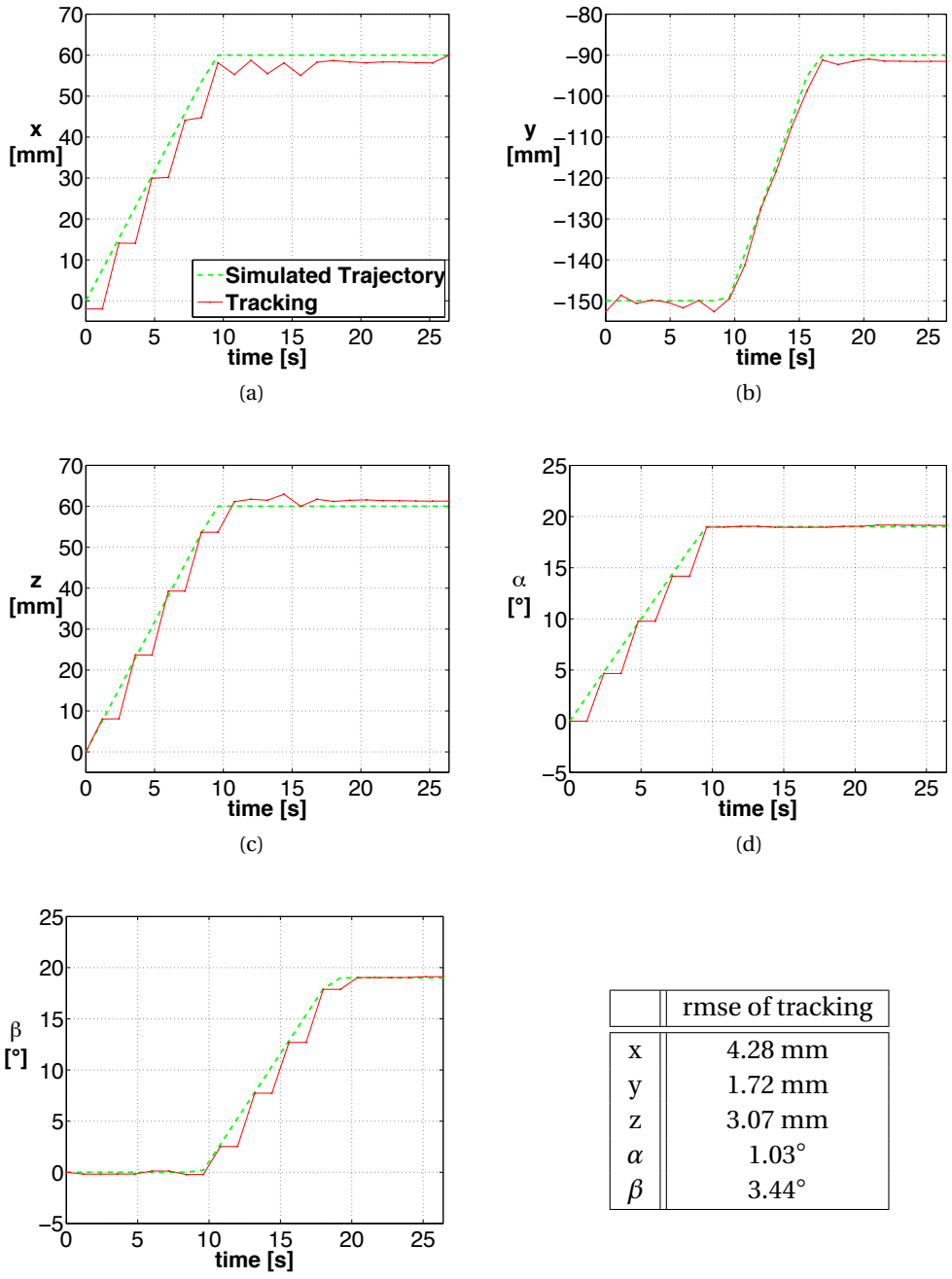
Scan plane alignment accuracy was evaluated with a simulated motion of the marker of 9 mm/s of translational and 2°/s of rotational speed according to the experimentally determined maximal tracking speed (Figure 2.22).

The root mean square error (rmse) between the simulated marker pose and the detected marker pose is presented in Figure 2.22f. The overall rmse of position detection is 5.54 mm. The corresponding graphs of the detected motion with respect to the simulated motion show a step pattern during tracking.
This step pattern can be explained through the fact that it is not possible to measure a 3D position in the image plane with the cylindrical geometry of the marker. This is due to the intersection of the marker and the image plane, providing no information about the marker position in the direction of the image normal. Hence, the marker is detected in the image plane and the unavailable position information is completed with the corresponding information from the image plane in order to form a 3D position measurement. As the image plane is aligned to the previously detected marker position, a part of the marker pose stays equal with respect to the previous image plane. An example for a tracking of a diagonal motion between the x- and z-axis with pure sagittal and transversal image planes is presented in Figure 2.23.

Furthermore, it is also not possible to detect a change of marker orientation for the orientation component whose rotation axis is collinear to the rotation axis of the image plane. As a consequence the angle α of the marker pose cannot be updated with a sagittal image plane and accordingly the angle β cannot be updated with a transversal image plane.

As the image planes in the presented motion in Figure 2.22 are only slightly inclined, the x- and α-components cannot be updated with a sagittal image plane. Accordingly the z- and β-components cannot be updated with a transversal image plane. The y-coordinate can be updated on both sagittal and transversal image planes and is thus not affected by the step pattern.

The scan plane alignment accuracy was also evaluated using the hypothetical high image update frequency (2 Hz) scenario presented in section 2.9.3. In order to compare these results with the previously performed simulation, the same motion type with a translational speed of 9 mm/s and a rotational speed of 2°/s were simulated. The results are presented in Table 2.3. The overall position rmse in the hypothetical scenario is 3 mm. Compared to the values for



	rmse of tracking
x	4.28 mm
y	1.72 mm
z	3.07 mm
α	1.03°
β	3.44°

(f)

Figure 2.22: Simulation results of a marker motion of 9 mm/s translational and 2°/s rotational speed. The simulated marker motion is decomposed to every pose component and depicted with the green graphs (a-e). The tracked position and orientation values are depicted with the red graphs. A step pattern is observable for the x-,z-, α- and β components. They are typically not detectable on both transversal and sagittal image planes when the marker motion is normal to one of both. For the presented motion, the y-component is detectable on both image plane types and is thus steadily updated. The rmse for the pose components is depicted in (f).

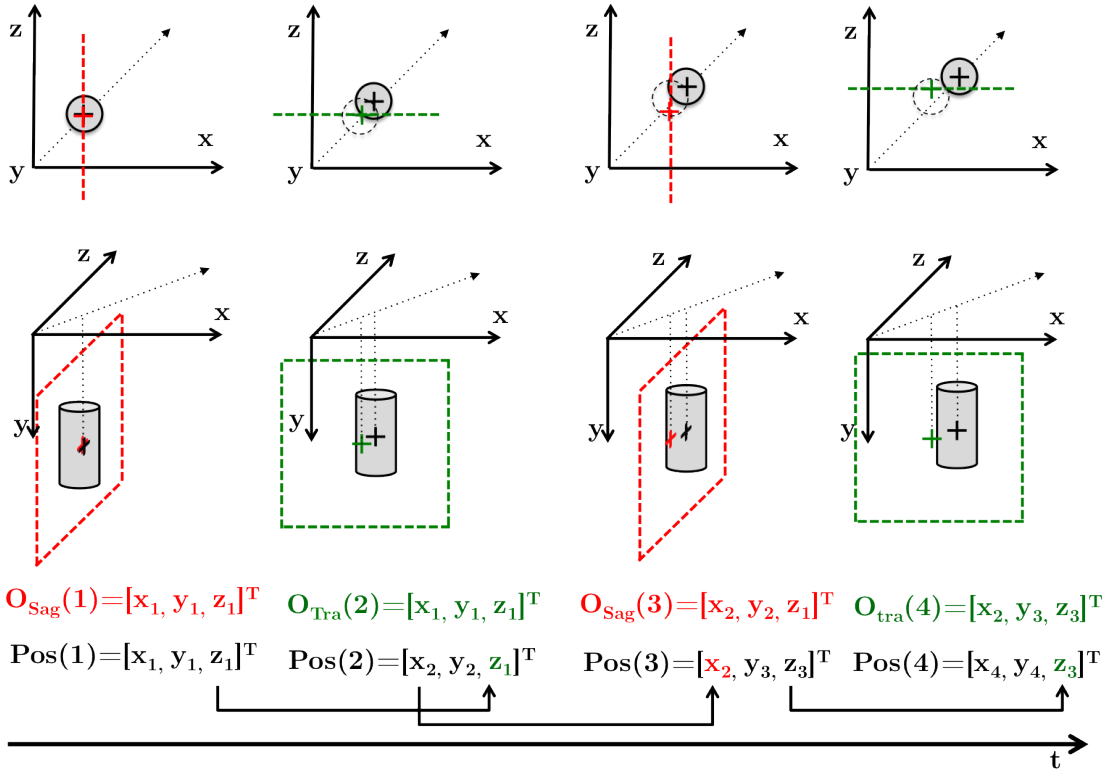


Figure 2.23: Marker motion along diagonal between x-axis and z-axis, tracked with pure sagittal (red) and pure transversal (green) image planes. At every time step t, the marker position ($Pos(t)$) is detected and the following image plane position ($O_{Sag}(t+1)$ or $O_{Tra}(t+1)$) is aligned accordingly. The x- and z-component of the marker pose cannot be updated on the sagittal and transversal image planes, respectively. As an image plane is positioned at the previously detected marker position, the unavailable pose component is replaced by its previous value.

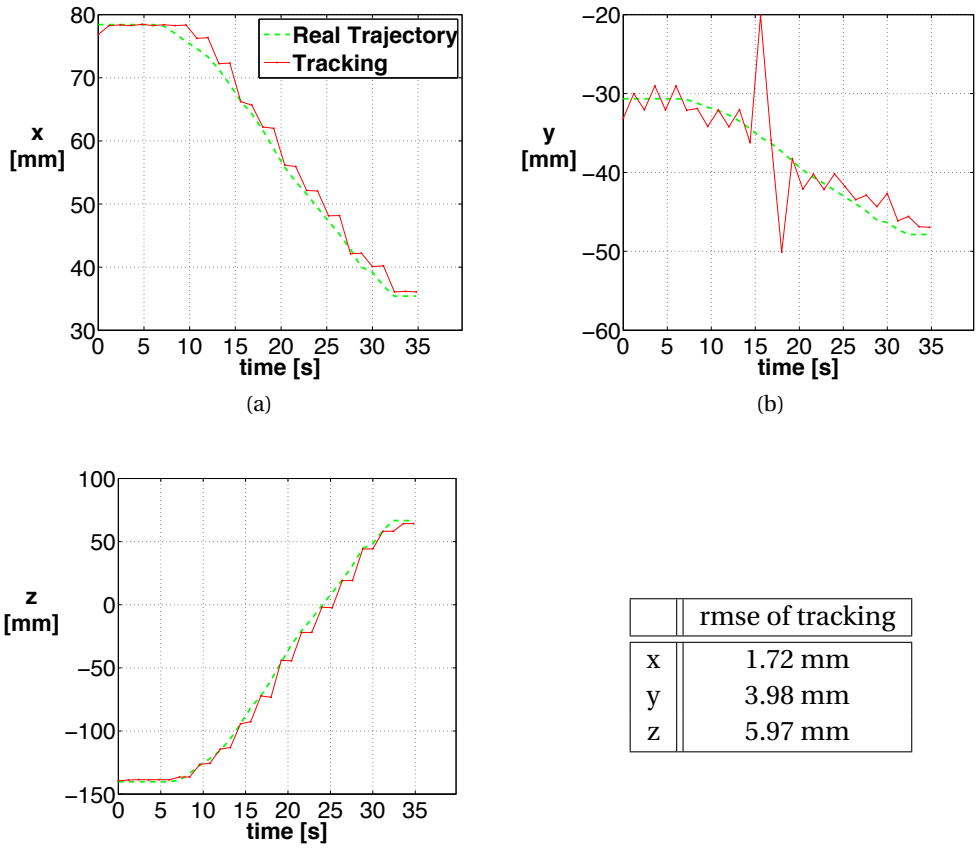
the previously performed simulations (Table 2.22f), the results in terms of accuracy are better due to the higher imaging frequency.
During the previous simulations, the y-component could be updated on both sagittal and transversal image planes due to their small inclination. This corresponds to a duplication of the update frequency for this component compared to the others that could only be updated on one of both image planes. For this reason, the improvement of the tracking error in the hypothetical scenario is less significant for the y-component than for the other pose components.

	rmse of tracking	
	at 2 Hz	at 0.83 Hz
x	2.20 mm	4.28 mm
y	1.15 mm	1.72 mm
z	1.68 mm	3.07 mm
α	0.50°	1.03°
β	2.47°	3.44°

Table 2.3: Tracking accuracy evaluation in simulations with a 2 Hz and 0.83 Hz imaging frequency.

Experiments

Experimental accuracy evaluation was performed through comparison of the value of the ground truth of the testbed and the detected marker pose from the presented workflow. Such a motion is presented in Figure 2.24 for the different position components of the marker pose. Note that the orientation components are not compared during experimental evaluation as they are constant over time.
The measured mean speed of the translation is determined on the translation curve of the testbed and is approximately 8.81 mm/s. The observed step pattern during simulations is also observable for the experimentally determined x- and z-components. The noise on the y-component can be explained through its update on every image plane and the image noise resulting in detection differences of 1 or 2 pixels between the two image plane types. According to the pixel size of the image planes (2 mm), these detection differences quantify as a position error between 2 to 4 mm. The two outliers of the y-coordinate are due to image artifacts which separated the marker in two objects during segmentation and which shifted thus the detected position value. The overall rmse of this motion is 7.38 mm which can also be computed separately for every position component: the rmse is 1.72 mm, 3.98 mm and 5.97 for the x-, y- and z-coordinates, respectively. The higher value for the y-coordinate is due to the two outlier values of marker detection. The high value for the z-coordinate can be explained with the translation direction that is primarily along the z-axis of the MRI scanner. The motion speed as well as the measured alignment errors along this component are thus higher than for the other two position components.



	rmse of tracking
x	1.72 mm
y	3.98 mm
z	5.97 mm

Figure 2.24: Experimental results for tracking of a marker motion of approximately 8.81 mm/s. The marker motion is decomposed to every pose component and depicted with the green curves (a-c). The tracked position values are depicted with the red graphs. As for the simulations results, a step pattern can be observed for the x- and z-component. The outliers on the curve of the y-component are due to segmentation errors of the marker shape. The obtained rmse for the position components is presented in (d).

The motion carried out during the experiment and during the simulation have approximately the same translation speed. The overall rmse for the position component obtained from simulation is slightly better than the experimentally determined rmse. This can be explained with image noise, inconstant marker speed and the two outlier detections during experimental evaluation. The value found during simulation is better, as no image noise or artifacts are simulated and the marker pose components are thus not affected by detection noise. The marker also moves at a constant speed over time during simulations which facilitates its tracking.

2.10 Conclusion

In this chapter, a workflow for MR image plane alignment based on detection and tracking in real-time clinical MR images, without any active tracking device, is presented. A marker, an image-based tracking algorithm, a software interface and a simulation program have been developed. The proposed workflow was designed to minimize dedicated acquisition time. In the ideal case, only one dedicated, non-clinically relevant (low flip angle) volume image is acquired for initialization. Initial scan plane alignment can be performed automatically with either resonant micro-coils in combination with a low flip-angle volume acquisition or with a passive marker in combination with a volume acquisition, hence no manual user initialization step is required. Tracking is then performed directly in the clinical real-time images through detection of the passive marker. The proposed workflow has been included in an existing clinical MRI sequence. Alternating acquisition of transversal and sagittal scan planes allow for a good spatial depiction of the anatomy, necessary for medical procedure.
The developed automatic marker detection algorithm was evaluated with experimental images of two volunteers in which at least 40% of the marker is visible and has achieved a success rate of 96.15%. The failed detections can be explained by an inappropriate patient body segmentation during which parts of the marker are deleted. Hence, the algorithm could be greatly improved with a more efficient segmentation of the patient's body.
Tracking simulations were carried out and revealed an overall rmse of ~ 5 mm for position tracking of the marker and a maximal tracking speed of ~ 19 mm/s. For experimental validation, an MR compatible testbed was used. The tracking accuracy and maximal tracking speed obtained during simulations are better than the experimental values, respectively 7.4 mm and 8.8 mm/s. This can be explained through absence of image noise, image artifacts and transmission times between PCs during simulations. In order to determine the maximal tracking speed experimentally, the possibility to perform a well known repeatable motion would be of great benefit. Initial tests with an MR compatible pneumatic stepper motor have been carried out and will be performed for determination of the maximal tracking speed. For that purpose, the identification and the command of the system are under process.
Furthermore, tracking simulations in a hypothetical scenario assuming a higher MR image update rate of 500 ms were carried out. A maximal tracking speed of 48.5 mm/s and an overall rmse of 3 mm have been found, highlighting the interest of an optimization of the imaging frequency for this tracking approach.
The proposed workflow allows to use the marker above the patient's body as an interventional imaging probe and can thus be used for the search of the needle insertion point on the patient.
The obtained results have also revealed a flaw of the presented tracking method which is the non-availability of certain pose components on the image planes and the resulting step pattern during tracking. These incomplete pose measurements deteriorate the rmse and decrease thus the tracking accuracy of the workflow. Overcoming this effect would be of great benefit for the tracking accuracy. A solution is presented in the following chapter, where a Kalman filter is used for image plane alignment.

3 Kalman Filtering for automatic image plane alignment in interventional MRI

The previously presented workflow for MR image plane alignment is based on tracking in two orthogonal real-time MR images. As pointed out in section 2.10, the main limitation of such an approach is the inability to detect position changes perpendicular to the acquired image plane. Indeed due to the desired compact and simple geometry of the marker, the shape of its cross-section with the image plane does not provide reliable information on its position in the direction perpendicular to the plane.
In order to overcome this limitation and the resulting step pattern on the tracking trajectory, a Kalman filter is proposed for improvement of the tracking. The Kalman filter is chosen as it has proven its benefits as estimation, prediction and data fusion algorithm in numerous applications. Furthermore, its simplicity and effectiveness make it the first choice for the presented workflow. In the presented workflow, the Kalman filter is used in order to estimate and predict the 3D pose of the marker. Hence, the extension of the MR scan plane alignment workflow through prediction of the marker pose is investigated in this chapter.
First, the Kalman filter and its algorithm are introduced. Then, the extension of the workflow presented in chapter 2 through the Kalman filter is described and the expected benefits for the marker detection are explained. The chapter finishes with the obtained simulation and experimental results.

3.1 Introduction to the Kalman filter

An excellent introduction to the Kalman filter ([Kal60]) is given by Welch and Bishop in form of a tutorial in [WB01] and by Maybeck in [May82]. The Kalman filter is there presented as an optimal, recursive data processing algorithm that has found a great variety of applications in numerous domains. It can be considered as optimal and recursive as it incorporates all available information, with no need of storing and reprocessing it every time new information is available.
The Kalman filter has several well known properties. As a filter, it can extract available information from noisy measurements. As a data fusion algorithm, it allows to combine available measurement data from different sources, while accounting for prior knowledge about the

observed systems and about the measuring devices. An estimate of an observed variable can be made in such a manner that the estimation error is statistically minimized.
The Kalman filter can be derived for an observed system by assuming simplifying constraints on the system dynamics: the measurement and process model are assumed to be linear and subject to zero-mean Gaussian noise. Even if these constraints are often not close to reality, the Kalman filter is one of the most established estimation and data fusion method and has been studied in great detail in theoretical and practical applications. This is mainly due to its simplicity, ease of implementation and optimality in terms of minimization of the estimate error covariance ([WB01]).
For non-linear systems, approximations have to be made. The Extended Kalman filter linearizes the system around the current state estimate with the first order terms of a Taylor series ([WB01]). The Unscented Kalman filter is an alternative to the Extended Kalman filter, where mean and covariance are parameterized using a set of discretely sampled points ([JU97]).
The most interesting application of the Kalman filter for the presented tracking workflow is its ability to make predictions of the awaited marker pose based on the process model and past measurements.

3.1.1 Kalman Filter algorithm

We assume a discrete-time, controlled process governed by a linear equation such as

$$\boldsymbol{x}(k) = A(k)\,\boldsymbol{x}(k-1) + \boldsymbol{B}(k)\,\boldsymbol{u}(k) + \boldsymbol{n}(k-1) \quad , \tag{3.1}$$

relating the process state $\boldsymbol{x}$ at time step $k-1$ to k via the transition matrix $\boldsymbol{A}(k)$. The matrix $\boldsymbol{B}(k)$ relates an optional control input $\boldsymbol{u}(k)$ to the state.
The measurement matrix $\boldsymbol{H}(k)$ of the system relates the process state to the measurement $\boldsymbol{z}(k)$ in the following manner:

$$\boldsymbol{z}(k) = \boldsymbol{H}(k)\,\boldsymbol{x}(k) + \boldsymbol{m}(k) \quad . \tag{3.2}$$

Process noise and measurement noise of the system are represented by $\boldsymbol{n}(k-1)$ and $\boldsymbol{m}(k)$, respectively. We assume that both have a Gaussian distribution ($N(mean, variance)$) according to

$$\boldsymbol{n}(k) \sim N(0, \boldsymbol{Q}(k)) \tag{3.3}$$

$$\boldsymbol{m}(k) \sim N(0, \boldsymbol{R}(k)) \quad , \tag{3.4}$$

where $\boldsymbol{Q}(k)$ and $\boldsymbol{R}(k)$ represent the process and measurement noise covariance matrices, respectively.
The internal variables of a Kalman filter are the process state $\boldsymbol{x}$ and the estimate error covariance matrix $\boldsymbol{P}$. The algorithm enables an optimal estimate of the state vector $\boldsymbol{x}(k)$ from its prediction and the available measurement $\boldsymbol{z}(k)$ .It consists thus in a prediction step and an

estimation step :

Prediction Step

$$\hat{\boldsymbol{x}}(k|k-1) = \boldsymbol{A}(k)\,\hat{\boldsymbol{x}}(k-1|k-1) + \boldsymbol{B}(k)\,\boldsymbol{u}(k) \tag{3.5}$$

$$\boldsymbol{P}(k|k-1) = \boldsymbol{A}(k)\,\boldsymbol{P}(k-1|k-1)\,\boldsymbol{A}^T(k) + \boldsymbol{Q}(k) \tag{3.6}$$

Estimation Step

$$\hat{\boldsymbol{x}}(k|k) = [\boldsymbol{I} - \boldsymbol{K}(k)\,\boldsymbol{H}(k)]\,\hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\,\boldsymbol{z}(k) \tag{3.7}$$

$$\boldsymbol{P}(k|k) = \boldsymbol{P}(k|k-1) - \boldsymbol{K}(k)\,\boldsymbol{S}(k)\,\boldsymbol{K}^T(k) \tag{3.8}$$

with $\boldsymbol{K}(k)$ and $\boldsymbol{S}(k)$ known as the gain and innovation covariance matrices, respectively:

$$\boldsymbol{K}(k) = \boldsymbol{P}(k|k-1)\,\boldsymbol{H}^T(k)\,\boldsymbol{S}^{-1}(k)$$

$$\boldsymbol{S}(k) = \boldsymbol{H}(k)\,\boldsymbol{P}(k|k-1)\,\boldsymbol{H}^T(k) + \boldsymbol{R}(k)$$

The notations $\hat{\boldsymbol{x}}(k|k-1)$ and $\hat{\boldsymbol{x}}(k|k)$ define the state prediction for time step k given the measurements up to time step $k-1$ and the state estimate for time step k given the measurements up to time step k, respectively. The same notation is used accordingly for the estimate error covariance matrix.
The algorithm is initialized with initial estimates for $\hat{\boldsymbol{x}}(k-1|k-1)$ and $\boldsymbol{P}(k-1|k-1)$ and starts with the prediction step.

3.2 Kalman extended MR image plane alignment workflow

Due to its ease of use, optimal and well-known characteristics, the Kalman filter is used for improvement of the workflow for automatic image plane alignment presented in section 2.4. The workflow still consists in two stages: an initialization step as presented in section 2.4.1 and then a real-time tracking stage. Only the real-time tracking using two orthogonal MR images changes with respect to the alignment strategy of the initial workflow.

3.2.1 Formalization of workflow

In order to introduce the Kalman filter adapted to the image plane alignment workflow, a formalization step is presented in this section. The introduced formalism is then used throughout the following sections. The image plane alignment workflow is based on the

measurement of the marker pose $\boldsymbol{x}_{Pose}$:

$$\boldsymbol{x}_{Pose} = \begin{bmatrix} \boldsymbol{x}_{Position} \\ \boldsymbol{x}_{Orient} \end{bmatrix} \tag{3.9}$$

with

$$\boldsymbol{x}_{Position} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad \boldsymbol{x}_{Orient} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad . \tag{3.10}$$

Vector $\boldsymbol{x}_{Position}$ consists thus in the position components (x, y, z) of the marker center in the MRI frame. Vector $\boldsymbol{x}_{Orient}$ contains the orientation components (α, β) of the marker main axis in the MRI frame (Figure 3.1). Angle α is the angle between the marker axis projected to a pure transversal plane and the y-axis of the MRI frame. Correspondingly, angle β is the angle between the marker axis projected to a pure sagittal plane and the y-axis of the MRI frame. The angles of the marker pose correspond thus to those presented for image plane alignment in section 2.6.2.
The marker pose can be fully described with only two orientation components as the marker is cylindrical and the rotation around its axis is not considered and free. The pose of a simple

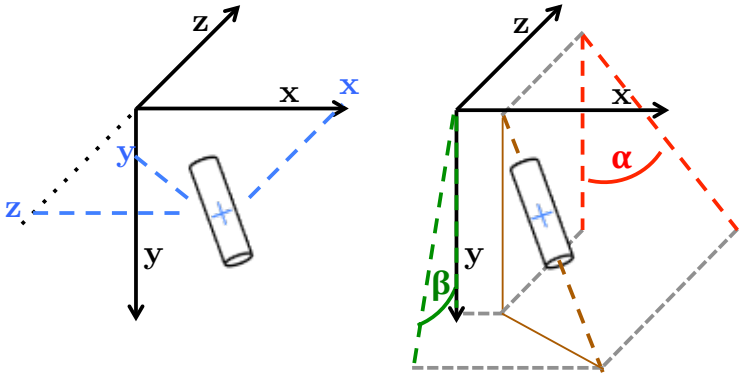


Figure 3.1: 3D pose of the marker consisting in its 3D position given by $[x, y, z]^T$ and its orientation given by $[\alpha, \beta]^T$

oblique image plane can be described with only four components, as the plane undergoes only one rotation around one axis of the MRI frame. For coherence of the notation, the image plane pose is described with the same components than the marker and will be noted $\boldsymbol{x}_{Plane} = [x, y, z, \alpha, \beta]^T$. Depending on which type of plane is described, one of the orientation components is set to zero for $\boldsymbol{x}_{Plane}$:

$$\begin{aligned} \boldsymbol{x}_{SagPlane}(k) &= \boldsymbol{x}_{Plane}(k) \quad \text{with} \quad \beta \to 0 \quad \text{for the sagittal plane} \\ \boldsymbol{x}_{TraPlane}(k) &= \boldsymbol{x}_{Plane}(k) \quad \text{with} \quad \alpha \to 0 \quad \text{for the transversal plane.} \end{aligned}$$

Initial workflow description

With the initial tracking presented in section 2.4, the image plane at time step k is aligned to the previously estimated pose of the marker from time step $k-1$ (Figure 2.23):

$$\boldsymbol{x}_{Plane}(k) = \hat{\boldsymbol{x}}_{Pose}(k-1|k-1) \tag{3.11}$$

So one assumption of this workflow is that motion of the marker is slow enough that at step time k the marker still intersects the image plane acquired at the previous marker pose $\hat{\boldsymbol{x}}_{Pose}(k-1)$.

A second assumption of the workflow is that the position of the marker is given by its intersection of coordinates ${}^{I}\boldsymbol{x}_{Position} = [x_I\, y_I\, 0]^T$ in frame $\mathcal{F}_I$ with the current image plane. As the current image plane position and orientation are chosen equal to the previous position $\hat{\boldsymbol{x}}_{Position}(k-1)$ and orientation ${}^{M}\boldsymbol{R}_I(k-1)$ of the marker, the position of the marker in the scanner frame is:

$$\hat{\boldsymbol{x}}_{Position}(k) = \hat{\boldsymbol{x}}_{Position}(k-1) + {}^{M}\boldsymbol{R}_I(k-1)\, {}^{I}\boldsymbol{x}_{Position} \tag{3.12}$$

$$\hat{\boldsymbol{x}}_{Position}(k) = \hat{\boldsymbol{x}}_{Position}(k-1) + {}^{M}\boldsymbol{R}_I(k-1) \begin{bmatrix} x_I(k) \\ y_I(k) \\ 0 \end{bmatrix} \tag{3.13}$$

with ${}^{M}\boldsymbol{R}_I$, the rotation matrix between $\mathcal{F}_M$ and $\mathcal{F}_I$ (see section 2.6).

From equation 3.13, the estimate of the marker position is the sum of the marker position at the previous time step and the 2D measurement from the image. As one measurement in the normal direction of the image is unavailable and set to zero, the marker position is constrained within a plane defined by ${}^{M}\boldsymbol{R}_I(k-1)$ and including $\hat{\boldsymbol{x}}_{Position}(k-1)$.

As example, with ${}^{M}\boldsymbol{R}_I(k-1) = \boldsymbol{I}_{3,3}$, the current image is a pure transversal plane and the cross section between the plane and the marker does not provide position information in the z-axis direction orthogonal to the plane. From equation 3.13, the marker position along the z-axis is $\hat{z}(k) = \hat{z}(k-1)$ and thus is considered equal to its previous value.

An illustration for a diagonal motion between the x- and z-axis, that is the causes of the step effect depicted in the tracking of the previous chapter, is given in Figure 2.23.

Proposed workflow with prediction

Assuming a model of the motion is known, a prediction of the marker pose for time step k can be used for image plane alignment instead of using the previous pose from time step $k-1$:

$$\boldsymbol{x}_{Plane}(k) = \hat{\boldsymbol{x}}_{Pose}(k|k-1) \quad , \tag{3.14}$$

where $\hat{\boldsymbol{x}}_{Pose}(k|k-1)$ is the prediction of the marker pose for time step k, given measurements up to time step $k-1$.

As the missing pose information is completed with the corresponding information from the image plane, it is no longer kept equal with respect to the previous time step. Instead, as the image plane position and orientation is based on a prediction of the marker pose ($\hat{\boldsymbol{x}}_{Pose}(k|k-1)$) from the previous time step, the missing information is now completed with this prediction. As a consequence, the step pattern on the tracking trajectory is avoided.

However, this workflow is based on the assumption that a prediction can be made using a marker motion model. As the Kalman filter is an optimal estimation and prediction algorithm, it is thus chosen to be integrated in the workflow using a constant velocity model for marker motion approximation.

3.2.2 Application of Kalman filtering to the presented workflow

As freehand motions of the marker are tracked, sudden speed changes are expected, and no process model for the marker motion is known. We thus approximate the process with a constant velocity model.

Adapted to the presented workflow, the state of the process corresponds to the marker pose $\boldsymbol{x}_{Pose} = [x, y, z, \alpha, \beta]^T$ and its derivative:

$$\boldsymbol{x}(k) = \begin{bmatrix} \boldsymbol{x}_{Pose}(k) \\ \dot{\boldsymbol{x}}_{Pose}(k) \end{bmatrix} \quad . \tag{3.15}$$

with a transition matrix of the following form:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_{5\times5} & \delta t \cdot \boldsymbol{I}_{5\times5} \\ \boldsymbol{0}_{5\times5} & \boldsymbol{I}_{5\times5} \end{bmatrix} \quad , \tag{3.16}$$

where δt is the time step between two acquisitions of the MRI scanner, $\boldsymbol{I}_{5\times5}$ is a 5×5 identity matrix and $\boldsymbol{0}_{5\times5}$ is a 5×5 zero matrix.
The marker pose is assumed to be measured directly inside the MR images ($\boldsymbol{z}(k) = \boldsymbol{x}_{Pose}(k)$) resulting in an observation matrix of the form

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{I}_{5\times5} & \boldsymbol{0}_{5\times5} \end{bmatrix} \quad . \tag{3.17}$$

The presented workflow can thus be described with the following equations:

$$\boldsymbol{x}(k) = \boldsymbol{A}(k)\,\boldsymbol{x}(k-1) + \boldsymbol{n}(k-1) \quad , \tag{3.18}$$

$$\boldsymbol{z}(k) = \boldsymbol{H}(k)\,\boldsymbol{x}(k) + \boldsymbol{m}(k) \quad . \tag{3.19}$$

The Kalman filter algorithm is then used in order to predict the pose of the marker for the next time step and eventually calculate an image plane aligned to this prediction.

For this purpose, the filter has to be updated with the 3D pose of the marker at each image acquisition. However, as described in section 2.9, only a subset of the 3D pose is available on each acquired image.
In this workflow, the missing components in the measurement $z(k)$ are replaced by their estimates from $\hat{\boldsymbol{x}}(k|k-1)$ due the acquisition of the image plane at the predicted marker pose.

As the Kalman filter requires the complete 3D pose of the marker in the measurement vector, available and missing (replaced by their estimates) pose components are used as measurements. The measurement vector contains thus pose components with different measurement confidence. In [Wel96], Welch highlights that the measurement covariance matrix is linked to a particular instant in time and can be estimated in real-time during operation of the Kalman filter in order to account for changing measurement noise. A dynamic adaptation of the measurement covariance matrix is thus expected to improve the tracking performance of the Kalman filter, as it allows to account for different measurement confidences in both - the same measurement vector, and in consecutive measurement vectors at different time instants.
Hence, in order to account for disparities in measurement confidence, measurements can be weighted differently for computation of the state estimate and the estimate error covariance matrix according to equations 3.7 and 3.8, respectively. For this purpose the measurement noise covariance matrix $\boldsymbol{R}$ is adapted dynamically to every measurement. As a consequence the gain matrix $\boldsymbol{K}$ changes and the measurement vector's influence on the state estimate and the estimate error covariance is defined through $\boldsymbol{R}$.
The measurement noise covariance matrix can be weighted with different strategies.
One possibility is to calculate the direction of the marker motion and determine at which extent it is normal to the current image plane. In this manner it is possible to admit only little confidence to the normal components of the image plane when the determined motion is majorly normal to the image plane. In the presented workflow the drawback of this approach is the slow image acquisition rate of the MR images and the resulting time delayed marker motion update.
Another strategy is to evaluate the marker shape after image segmentation and marker detection in order to assess measurement confidence. In this manner it would be possible to weight the measurement according to the similarity between the detected marker shape to its typical shape. A drawback of this method is the MR image noise altering the segmented shape of the marker. Furthermore, parameterization of a detected marker shape requires higher computation time.

For the presented workflow, a last strategy is chosen: as marker pose components normal to the image plane cannot be updated, their measurement confidence is adapted accordingly. As pointed out in section 2.9, the spatial components of the normal vector depends on the orientation of the simple oblique image plane. Hence, measurement noise covariance values for position components normal to the image plane vary on an interval $[10^{-5};1]$ and are a function of α and β for sagittal and transversal image planes, respectively. As an orientation component can either be detected or not on an image plane, its measurement noise covariance is not gradually adapted but set to the interval boundaries in a binary manner. Thus, $\boldsymbol{R}$ is obtained as

$$\boldsymbol{R} = \begin{bmatrix} 10^{-a} & 0 & 0 & 0 & 0 \\ 0 & 10^{-b} & 0 & 0 & 0 \\ 0 & 0 & 10^{-c} & 0 & 0 \\ 0 & 0 & 0 & 10^{-d} & 0 \\ 0 & 0 & 0 & 0 & 10^{-e} \end{bmatrix} \tag{3.20}$$

with $a = 5$, $b = 5cos(\beta)$, $c = 5\left|sin(\beta)\right|$, $d = 5$, $e = 0$ for transversal image planes and $a = 5|sin(\alpha)|$, $b = 5cos(\alpha)$, $c = 5$, $d = 0$, $e = 5$ for sagittal image planes.
The process noise covariance matrix $\boldsymbol{Q}$ is constant over time and set experimentally to

$$\boldsymbol{Q} = \boldsymbol{I}_{10\times 10} \cdot 10^{-5} \quad . \tag{3.21}$$

3.2.3 Real-time tracking with Kalman filter

Real-time tracking starts with the acquisition of a simple oblique sagittal plane ($Image_1$) whose parameters have been sent to the MRI scanner at the end of the initialization stage.
The passive marker is then detected, its 3D pose ($Pose_1$) is computed and used for initialization of the Kalman filter. The Kalman filter state is thus initialized with the first measured pose $Pose_1$ and a speed equal to zero. The initial estimate error covariance matrix is set to the identity matrix in order to quickly converge afterwards. A first prediction is then made by the Kalman algorithm, but as this prediction is based on zero speed of the marker, the predicted marker pose ($Prediction_1$) is exactly the same as the first detected one, similar to the original workflow without Kalman filter. Therefore the following simple oblique transversal image plane ($Image_2$) is aligned on the first detected marker pose. After acquisition of $Image_2$, the new marker pose is detected ($Pose_2$) and handed to the Kalman filter as measurement. The Kalman filter then corrects its previous prediction according to this measurement and computes an estimation of the process state. Based on this estimation a prediction ($Prediction_2$) of the marker state is performed and the next sagittal oblique image plane ($Image_3$) is aligned to the predicted marker pose.
From then on, the two orthogonal image planes (sagittal/transversal) alternate and update their positions and orientations mutually according to the predicted marker pose from the Kalman filter (Figure 3.2).
In case of a non-detection of the marker, the measurement for the Kalman filter is unavailable.

Nevertheless, in order to generate a prediction for the next marker pose, the previous prediction of the Kalman filter is handed in as measurement and the measurement covariance is set to a very high value. In this manner, the Kalman filter will ignore this measurement according to eq.3.7 and eq.3.8 and generates a prediction of the next marker pose, based on its internal process model only.

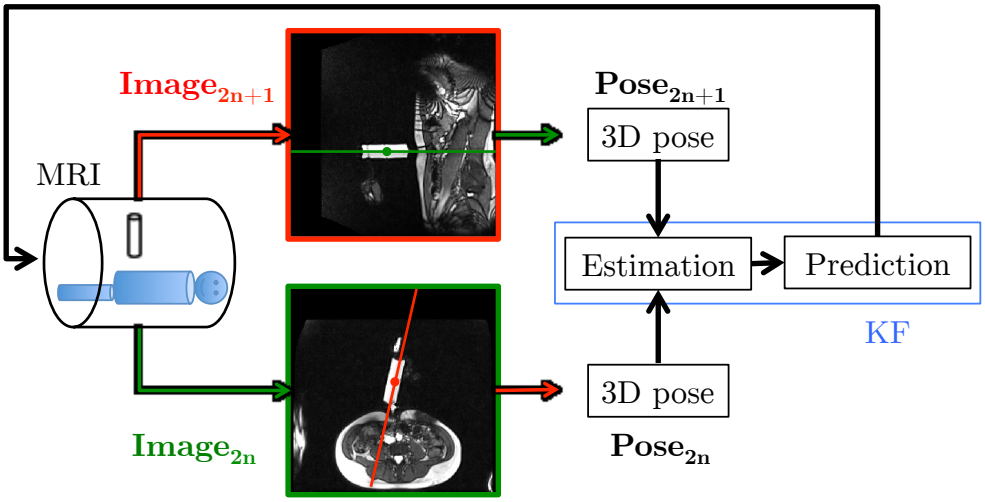


Figure 3.2: Principle of the Kalman extended workflow: two orthogonal image planes, transversal (green border) and sagittal (red border) are alternately acquired. The 3D marker pose is detected on the image plane and fed into the Kalman filter. The Kalman filter computes an estimate and a prediction of the marker pose. An orthogonal image plane aligned to the predicted marker pose is calculated and a corresponding command is sent to the MRI for acquisition of the next image plane.

3.3 Marker detection

Through use of a marker pose prediction from the Kalman filter, additional prior knowledge is available during image segmentation and interpretation and can be of great benefit for these steps. Hence, it is possible to implement either a restrained search window on an image plane or an additional position criterion for the marker detection algorithm.

3.3.1 Region of interest for image segmentation

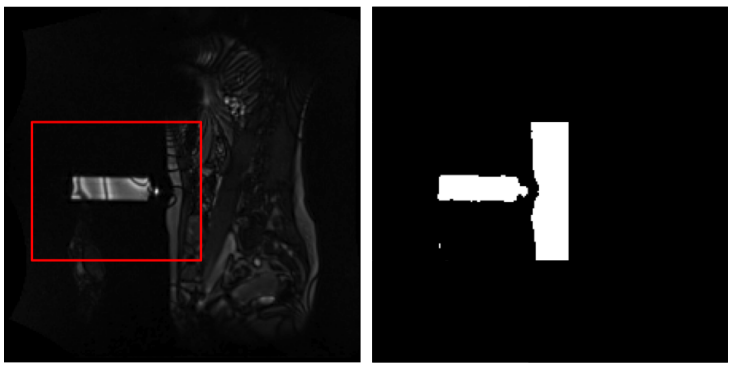
A restrained region of interest can be implemented in order to facilitate the image segmentation step, as the patient's body can almost entirely be deleted from the image to analyze (Figure 3.3a). For this purpose, the predicted 3D position of the marker in the MRI frame is computed as pixel position on the image plane for which it is predicted. A square region of interest of twice the marker length is then computed around the pixel position of the marker.

The region of interest is deliberately chosen of this size, as smaller windows increase the risk of obliterating parts of the marker. The image content outside of the region of interest is then set to zero in order to keep only the image content inside for further analysis. The resulting image contains then typically the passive marker with the hand of the person holding it. The thresholds for the image segmentation are determined in the same manner as described in section 2.5.3 but only the threshold established at the 0.9-quantile is applied to the restrained window in order to segment the passive marker and the hand of the person holding it separately. Segmentation can thus easily be realized, if the patient's body is not depicted inside the region of interest.

3.3.2 Position score for image interpretation

Another approach for application of the prior knowledge is the implementation of a position criterion during image interpretation. For this purpose the predicted 3D position of the marker in the MRI frame is computed as pixel position on the image plane for which it is predicted. Then there are two options: either all segmented objects whose centroid lie outside a predefined region of interest are directly excluded from the image interpretation step, or an additional position score is integrated in the proposed score system in section 2.5.3. The position score can be calculated as a function of the distance between the detected and predicted marker position for an image plane.

Use of the prior knowledge during image interpretation is more useful than its use during image segmentation. This is due to potentially remaining body parts inside the region of interest for image segmentation (Figure 3.3b) that cause problems for the image interpretation method presented in section 2.5.3. The position criterion for the marker detection has no such robustness problems and was thus selected for integration of prior knowledge gained by the Kalman filter.



(a) (b)

Figure 3.3: (a): Region of interest during image segmentation thanks to prior knowledge of the marker position according to the Kalman prediction. (b): Potentially remaining body parts after image segmentation may be problematic for the following image interpretation step.

3.4 Results

In order to assess the performance of the Kalman Filter for image plane alignment, simulations and experiments are performed with the same protocol as presented in section 2.8.1 and section 2.8.2, respectively.
The workflow again is based on marker detection in sagittal oblique and in transversal oblique image planes, that are acquired alternately. The marker is detected on one plane and its current pose is estimated by a Kalman filter. Then, the Kalman filter generates a prediction of the marker pose for the acquisition time of the next image plane. The next image plane is then aligned to the predicted marker pose.
The maximal tracking speed of the workflow and its tracking accuracy are assessed in the same manner as in section 2.9 in order to compare the obtained results.

3.4.1 Timing analysis of real-time image plane alignment

Correspondingly to section 2.9.1, an analysis of the time that is necessary for execution of the Kalman filter algorithm was performed. For this purpose, the execution time of 165 estimation and prediction steps was measured. The additional processing time due to the use of the Kalman filter is less than 250 μs and thus negligible, as expected.
Hence, the experimental image update time remains 1200 ms for the Kalman filter approach for image plane alignment.

3.4.2 Evaluation of the maximal tracking speed

The maximal tracking speed of the Kalman extended workflow is determined with both, simulations and the experimental testbed.

Simulations

Simulations were carried out for determination of the maximal tracking speed for a motion at an angle of 45° between the transversal and sagittal image planes in the same manner as in section 2.9.3. The obtained simulated maximal tracking speed is 21.6 mm/s which is 2.1 mm/s higher than the value for the initial workflow. The maximal tracking speed for the Kalman filter is limited due to its convergence time to a steady state when initialized with a null speed. With a progressive increase of the marker speed or a state initialization with the speed of the marker, the Kalman filter could track even higher speeds for this strictly linear and unidirectional motion type. In return, such simulations are not realistic with regard to the final application of this workflow, which is tracking of freehand motions of the marker.

Correspondingly to section 2.9.3, the maximal tracking speed of the Kalman approach is evaluated with simulations of the hypothetical high update frequency (2 Hz) scenario. For an image update frequency of 2 Hz, the maximal tracking speed is obtained at 52 mm/s.

Compared to the maximal tracking speed in the hypothetical scenario of the initial workflow, the value is 3.5 mm/s higher.

Experiments

The maximal tracking speed is determined experimentally in the same manner as in section 2.9.3, using an oscillatory motion with increasing speeds. The translation curve of the performed oscillatory motion on the testbed is depicted in Figure 3.4. The obtained value for the presented workflow is 21.24 mm/s which is more than twice of the value for the initial workflow. This value is very similar to the maximal tracking speed calculated for the diagonal motion in the previous simulations.

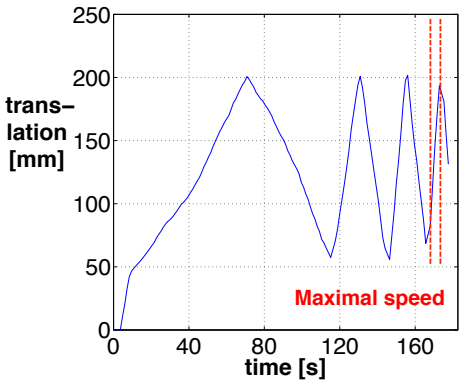


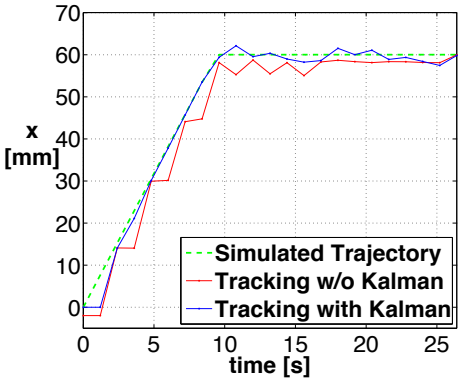
Figure 3.4: Experimental determination of the maximal tracking speed. The oscillatory motion is manually performed along the testbed axis with increasing speed until the marker cannot be tracked anymore. The last edge of the motion before the tracking is interrupted is used for maximal speed computation. Maximal speed in this case is 21.2 mm/s.

3.4.3 Accuracy assessment

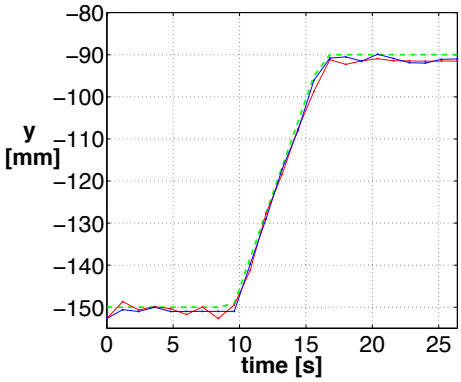
The tracking accuracy of the Kalman extended workflow is also determined with both, simulations and the experimental testbed.

Simulations

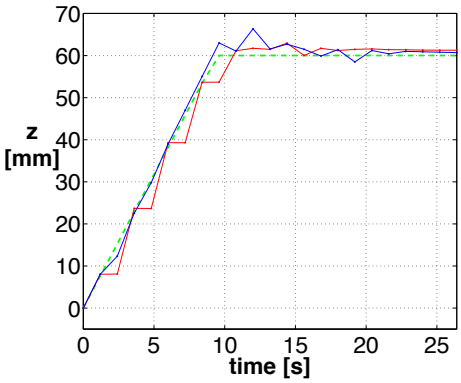
In order to assess the accuracy of the Kalman Filter for image plane alignment, simulations were performed for the same simulated motion as in section 2.9.4. A comparison of tracking performance of the workflow with and without Kalman filtering is presented in Figure 3.5. The step pattern observed for the initial workflow is compensated through the use of the Kalman filter (Figure 3.5a - 3.5e). The overall 3D root mean square error (rmse) of position detection is 3.05 mm, which is an interesting improvement with respect to the initial workflow (5.54 mm). Tracking results in terms of rmse benefit from the Kalman filter for all pose



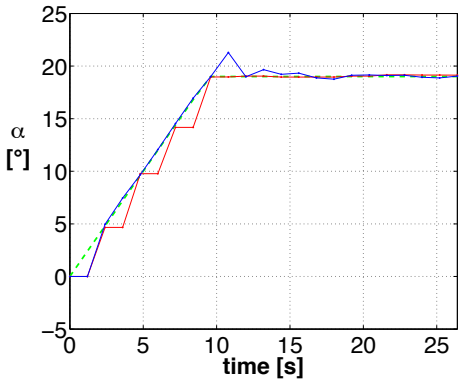
(a)



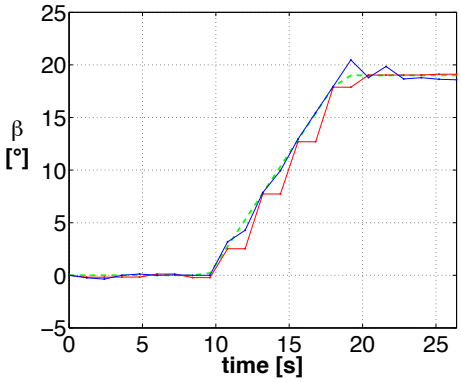
(b)



(c)



(d)



(e)

	rmse of tracking	
	w/o Kalman	with Kalman
x	4.28 mm	1.97 mm
y	1.72 mm	1.32 mm
z	3.07 mm	1.93 mm
α	1.03°	0.72°
β	3.44°	1.58°

(f)

Figure 3.5: Simulation results of a marker motion of 9 mm/s translational and 2°/s rotational speed for tracking without and with Kalman filtering. The simulated marker motion is decomposed to every pose component and depicted with the green graphs (a-e). The tracked position and orientation values are depicted for tracking without (red graphs) and tracking with the Kalman filter (blue graphs). The observed step pattern for different components of the initial tracking workflow is compensated through Kalman filtering. Nevertheless, overshoot can be observed for the Kalman filter, due to its the constant velocity model and the low MR image acquisition frequency. The comparison of the rmse for the pose components is depicted in (f).

components (Figure 3.5f). The rmse for the x and β-components is halved through use of the Kalman filter.
Nevertheless, overshoot is observable for the x-,z-,α- and β-components. Overshoot typically occurs when the motion changes in direction or speed and is due to the implemented constant velocity model and the slow image acquisition rate. The Kalman filter takes two or three images of transition time before adapting to the new motion and coming back to steady state. The transition time for the position components depends on the orientation of the MR image relative to the marker motion. If a changing position component cannot be detected on the next image plane (i.e. component orthogonal to image plane), transition time may be longer than for an immediate detection on the next image plane. In the same manner as the y-component is less affected by the step pattern due to its update on both orthogonal MR images, it is also less affected by the overshoot of the Kalman filter.

In order to evaluate the effect of the dynamic adaptation of the measurement noise covariance matrix ($\boldsymbol{R}$) according to the current image plane (section 3.2.2), further simulations have been carried out again with the same motion. The tracking accuracy is compared for the Kalman filter with dynamic adaptation of the covariance matrix and without adaptation (Table 3.1).

	rmse of tracking using Kalman Filter	
	without adaptation of $\boldsymbol{R}$	with adaptation of $\boldsymbol{R}$
x	2.44 mm	2.26 mm
y	1.11 mm	1.08 mm
z	1.44 mm	1.47 mm
α	0.87°	0.75°
β	2.22°	1.9°

Table 3.1: Comparison of the rmse of the tracking using the Kalman filter with and without dynamic adaptation of the measurement noise covariance matrix $\boldsymbol{R}$.

The dynamic adaptation of the measurement noise covariance matrix with the chosen method results only in slight improvement of the image plane alignment performance in submillimetric order for the position component (overall rmse improvement of ~0.13 mm). The obtained result is below the expectations towards this method and not convincing. As its use makes no significant difference for the presented tracking workflow, dynamic adaptation of the measurement noise covariance has not been used in order to optimize the calculation time of the workflow.

Correspondingly to section 2.9.4, the accuracy assessment of the Kalman approach is also performed in the hypothetical scenario where an image update rate of 2 Hz is simulated. The results in terms of rmse errors are compared to the initial workflow in Table 3.2. The overall position rmse is 2.14 mm, which is almost 1 mm smaller than for the initial workflow in the hypothetical scenario. Thus, a higher imaging frequency is also highly profitable for the Kalman tracking approach.

One can notice that the value for the rmse of the β-component of the Kalman filter approach is slightly higher at a 2 Hz imaging frequency than at 0.83 Hz. This is mainly due to the match between the initial instant of change of the component speed and the acquisition of a sagittal image plane when running at 0.83 Hz. It enables the Kalman filter to readily detect and account for the motion change. On the contrary, at 2 Hz, the initial change of the β-component speed unfortunately matches the acquisition of a transversal image plane, where the component change cannot be detected. Hence, the change can only be detected on the next sagittal image plane.

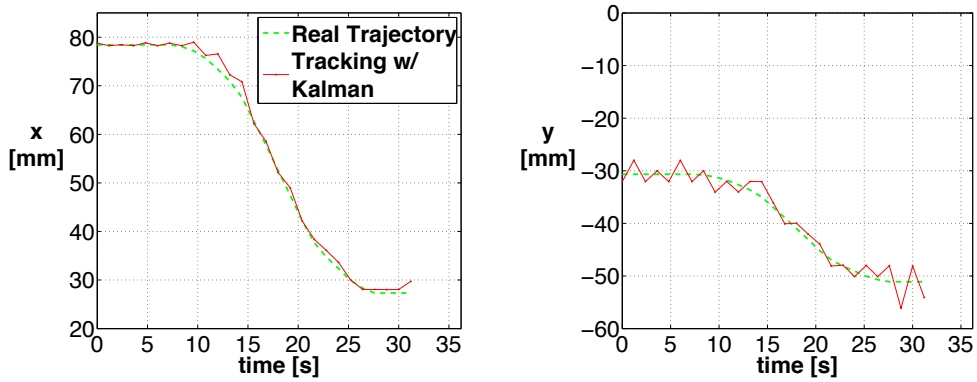
	rmse of tracking at 2 Hz		rmse of tracking at 0.83 Hz	
	w/o Kalman	with Kalman	w/o Kalman	with Kalman
x	2.20 mm	1.27 mm	4.28 mm	1.97 mm
y	1.15 mm	1.13 mm	1.72 mm	1.32 mm
z	1.68 mm	1.30 mm	3.07 mm	1.93 mm
α	0.50°	0.30°	1.03°	0.72°
β	2.47°	1.65°	3.44°	1.58°

Table 3.2: Tracking accuracy evaluation of the initial and Kalman extended workflow in simulations with a 2 Hz and 0.83 Hz imaging frequency.

Experiments

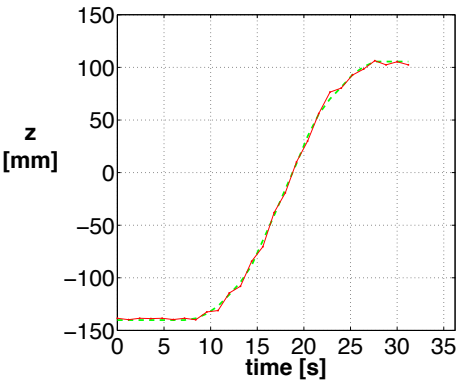
For evaluation of the benefits of the Kalman extended workflow, its performance is assessed with the same test setup using the MR compatible testbed as presented in section 2.8.2. First, the operator manually translates the marker, trying to keep a constant velocity. The operator iteratively tested several increasing speeds, while trying to ensure tracking of the marker through visualization of both real-time MR images on the in-room display screen. The performed motion and tracking are presented in Figure 3.6 for the different position components of the marker pose.
The measured mean speed of the motion is determined on the translation curve of the testbed and is approximately 17.5 mm/s. As for the simulations, it can be observed that the step pattern of the x- and z- components is compensated through the use of the Kalman filter at the cost of slight overshoot during motion changes. The measurement noise on the y-component can be explained in the same manner as for the experiments in section 2.9.4 through its update on every image plane and through image noise resulting in detection differences.
The overall 3D rmse of this motion is 3.7 mm which can also be computed separately for every component: the rmse is 1.2 mm, 2 mm and 2.8 mm for the x-, y- and z-coordinates, respectively. Even though the performed marker motion is more than twice as fast as the marker motion for the initial workflow (8.81 mm/s), the rmse for the Kalman extended workflow is clearly better than without the Kalman filter.
Nevertheless, the improvement of the tracking depends on the compliance of the motion and its model in the Kalman filter, i.e. a motion with constant velocity. As a consequence rmse can

rise if the marker motion exhibits strong acceleration. Smooth free-hand motions are then suggested to the person operating the marker.



(a)

(b)



(c)

	rmse of tracking
x	1.23 mm
y	2.02 mm
z	2.82 mm

(d)

Figure 3.6: Experimental results for tracking of a marker motion of approximately 17.5 mm/s using the Kalman filter. The marker motion is decomposed to every pose component and depicted with the green curves (a-c). The tracked position values are depicted with the red graphs. The step pattern observed for the initial workflow for the x- and z-components are compensated. The obtained rmse (d) is smaller than the one for the initial workflow, even though a much faster marker motion is performed.

3.5 Conclusion

In this chapter, the workflow for MR image plane alignment is extended through a Kalman filter. The Kalman filter is chosen due to its optimal characteristics, its easy-to-implement algorithm and its widespread use in tracking applications. It allows to overcome the tracking inaccuracies inherent to the initial workflow due to the unavailability of the marker pose components perpendicular to the MR image planes. The chosen constant velocity process model

performs well for weak accelerations of the tracked marker motion. The filtering effect of the Kalman filter allows to generate optimal marker pose estimates increasing the robustness of the tracking against measurement noise.
The Kalman filter also shows its benefits for the facilitation of the marker detection algorithm due to prior knowledge of the marker position according to the Kalman prediction.
Furthermore, a weighting technique based on dynamic adaptation of the measurement covariance matrix according to the used image plane and its inherent unavailable pose components is investigated. Its effect on the estimation and prediction is verified and found to be negligible, for the considered tracking with a low update frequency of the Kalman Filter.
The improvement of the image plane alignment is obtained at the cost of slight overshoots due to accelerations of the marker motion. The overshoots are mainly due to the used constant velocity model which is a rather strong simplification of the tracked freehand motion and the low image acquisition rate of the MR images.

In order to improve the reactivity of the image plane alignment workflow, decoupling of the tracking frequency from the MR image acquisition frequency would be of great use. For this reason the combination of the presented workflow with an additional sensor is presented in the next chapter.

4 Extension of the workflow through an RGB-D sensor

The Kalman extended workflow has shown benefits for the alignment of the image plane to the marker. Nevertheless, the slow acquisition rate of the clinical MRI sequence limits the speed of the motion that can be tracked. In order to overcome this limit, an additional tracking sensor with a higher acquisition frequency should be used. In order to minimize additional hardware cost and avoid modification of the MRI hardware and sequences, an RGB-D sensor has been chosen as it proposes a cheap and easy solution for 3D pose measurements and has been abundantly used in tracking applications over a wide area of research fields.
The purely MR image based approach presented in chapter 2 is thus combined with an RGB-D sensor based approach. In this manner the advantages of a passive and active approach are combined and they mutually compensate for their inherent drawbacks.

This chapter gives first an insight on the technique of RGB-D sensors and their applications. Second, a hybrid image plane alignment workflow combining the passive workflow presented in chapter 2 and an RGB-D sensor based approach is presented. Then, the detection algorithm for the RGB-D sensor is presented, as well as the registration approach between RGB-D and MRI frames and the developed system architecture. Focus is then put on the fusion of the MR image and RGB-D sensor based measurements through use of an Information filter. In this context a general introduction to multi-sensor data fusion is presented. The chapter is ended with the presentation of the achieved results and the conclusion.

4.1 RGB-D sensor

An RGB-D sensor delivers RGB images and a depth image of an observed scene. It combines an RGB camera with an infrared (IR) projector and an IR camera. The IR projector and camera work together as a depth sensor: an IR speckle dot pattern is projected on a scene and the reflected speckles are captured by the IR camera. The spatial relationship between the projector and camera is known thanks to an offline calibration. The projected dot pattern and the camera observed dot pattern are compared. Due to a position shift of the observed dot pattern that is a function of the distance of the object to the RGB-D sensor, it is possible to

determine the depth of the object (Figure 4.1).

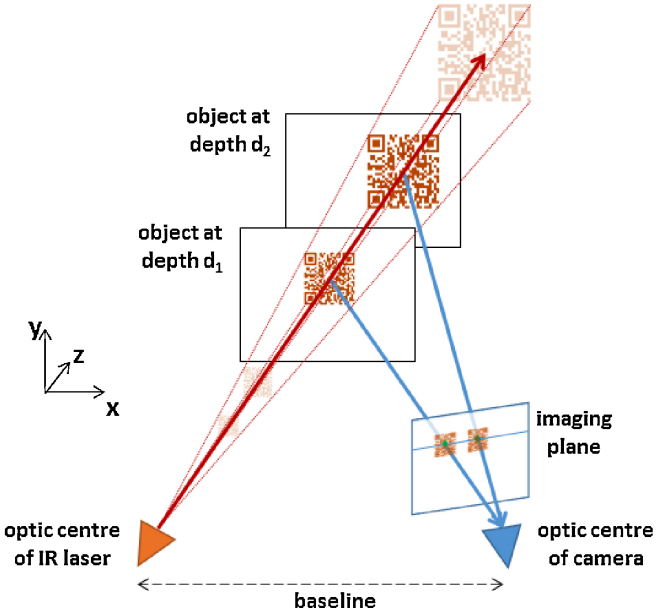


Figure 4.1: Depth measurement with an RGB-D sensor. The left-right shift of the observed dot pattern allows to reconstruct the depth of the reflecting surface (reprinted from [HSXS13]).

There are several RGB-D sensors available on the market. The most popular are the *Microsoft Kinect* and the *Asus Xtion Pro Live*. During this work the *Asus Xtion Pro Live* (Figure 4.2) was chosen as it is directly powered via the USB cable used for data transmission and does not need an additional power supply as the *Microsoft Kinect* does. The sensor delivers RGB and depth images with a resolution of 640×480 pixels at an acquisition frequency of 30 Hz. Depth sensor range lies between 800 mm up to 3500 mm. Detailed specifications of the sensor can be found on the website of the manufacturer ([htt13c]). Calibration of the RGB camera was

Figure 4.2: RGB-D sensor *Asus Xtion Pro Live* with infrared projector, RGB camera, infrared camera and microphones (reprinted from [htt13c]).

performed in order to determine its intrinsic parameters (see section 4.1.4 for details). The RGB camera and the IR camera are horizontally separated by a few centimeters, which makes the determination of the extrinsic parameters between them necessary in order to correlate the depth images to the RGB images. During this work, the manufacturer delivered set of

extrinsic parameters has been used.
As RGB-D sensors provide RGB images and synchronized depth images of a scene at a high acquisition rate and at low material cost, they have been used extensively in a widespread area of research during the last years.

4.1.1 Literature review of RGB-D sensors in research applications

In [HSXS13], Han *et al.* present a review of Kinect-based computer vision algorithms and applications on which is based the following literature review. They highlight that the complementary and synchronized RGB and depth information can be used in order to find solutions to classical problems in computer vision at a lower cost than with traditional 3D cameras such as stereo cameras and time-of-flight (TOF) cameras. The type of problems that can be addressed or whose solution can be improved include object tracking and recognition, human activity analysis , hand-gesture recognition, and indoor 3D mapping (Figure 4.3). In its first

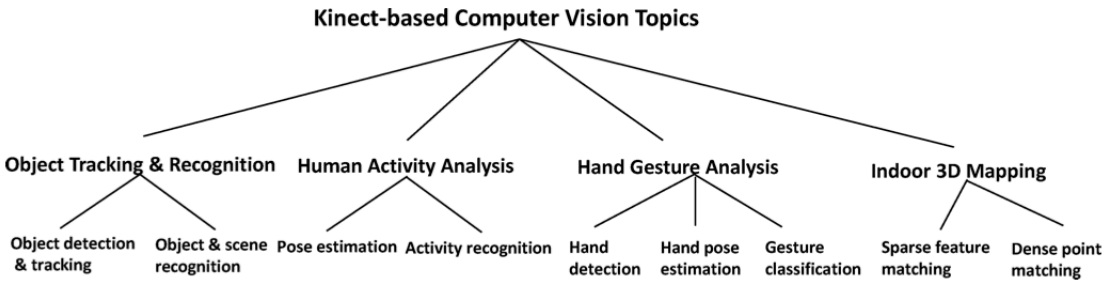
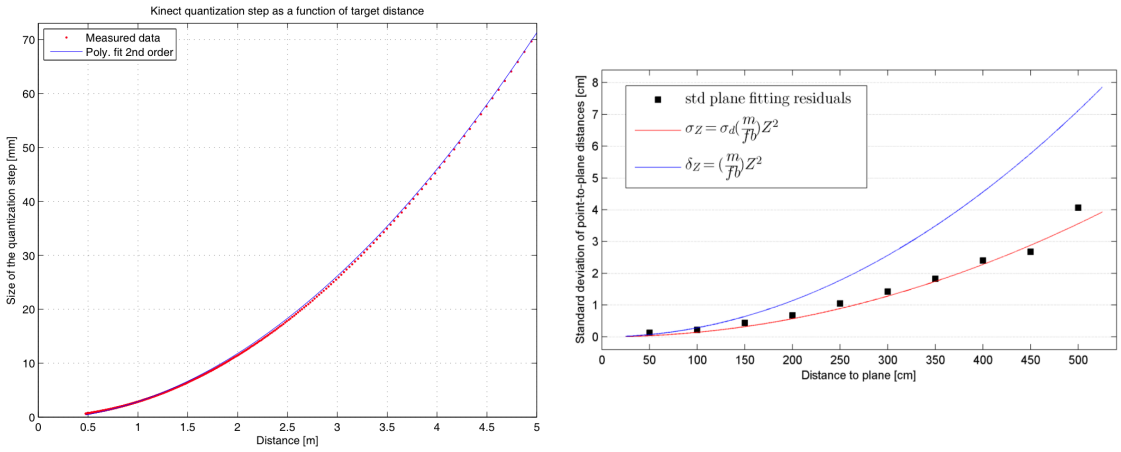


Figure 4.3: Type of problems that are currently under investigation through RGB-D sensor-based research (reprinted from [HSXS13]).

section the review presents research focussed on performance evaluation:
In [SJP11], the resolution and precision of Kinect depth measurement is investigated. In terms of resolution the quantization step q, which is the distance between two consecutive depth values, was evaluated. The function q, describing the quantization step as a function of depth z was found to be non-linear : $q(z) = 2.73z^2 + 0.74z$ (Figure 4.4a). The authors also compared the Kinect camera precision to a stereo camera and a TOF camera. The authors found the Kinect to be superior in accuracy to the TOF camera and close to the performance of the stereo camera.
In [SLAL11], the authors compare the Kinect and two TOF cameras to an actuated laser range sensor. Their results reveal that none of the devices can compete with the laser range sensor in terms of accuracy, when its full range is evaluated. When environments of a constrained size (within a sphere of radius 3.5 m) are considered, the Kinect has an accuracy similar to the one, obtained through the actuated laser range sensor.
Khoshelham $etal$. explain the depth measurement by triangulation in [KE12] and highlight possible error sources and their influence. The three main sources for erroneous depth measurement are related to the sensor, the measurement setup and the properties of the observed



(a) (b)

Figure 4.4: (a): Kinect quantization step as function of target distance, given for a distance interval from 500 mm to 5000 mm (reprinted from [SJP11]). (b): Standard deviation of depth measurements as function of distance between sensor and object (black squares). The theoretical random error and resolution of individual depth measurements at different distances from the sensor are derived and depicted as red and blue graphs, respectively (reprinted from [KE12]).

object surface. Sensor errors refer to an inadequate calibration and inaccurate measurement of disparities for depth reconstruction. For measurement setup errors, the authors mention mainly the lighting conditions (influencing the disparity measurement) and the imaging geometry. Imaging geometry includes the distance between the object and the sensor and occlusions of or shadows on parts of the scene (due to invisible pattern parts for the IR camera). Both occlusions and shadows result in gaps in the depth image. In the same manner, smooth and shiny object surfaces make the disparity measurement impossible and their depth values can neither be reconstructed. Experiments were carried out in order to evaluate the random error in depth measurement as a function of distance between object and sensor. For this purpose, depth values of a planar surface were observed at different distances and the standard deviation of 4,500 samples was calculated for every distance (Figure 4.4b). The results reveal that the random error in depth measurement increases with increasing distance between object and sensor from a few mm at close range to 4 cm at maximum range of 5 m.
Major progress has been made in image segmentation with the facilitation of background subtraction methods. These algorithms could be improved and made more robust against lighting condition changes and lack of contrast, which are difficult problems for algorithms based on traditional images without depth information. Main application of these new algorithms is people detection, either based on depth images only or on combined use of RGB and depth images.
Few medical applications are known up to now. In [WCW13] a semantic map of the environment of a wheelchair is built based on the acquired data by an RGB-D sensor. In [MNB+11], three RGB-D sensors are used in order to track personnel in a surgical environment. In [LR13],

an approach for detection and localization of equipment in an interventional room using RGB-D sensors is presented in the context of an X-ray monitoring system.

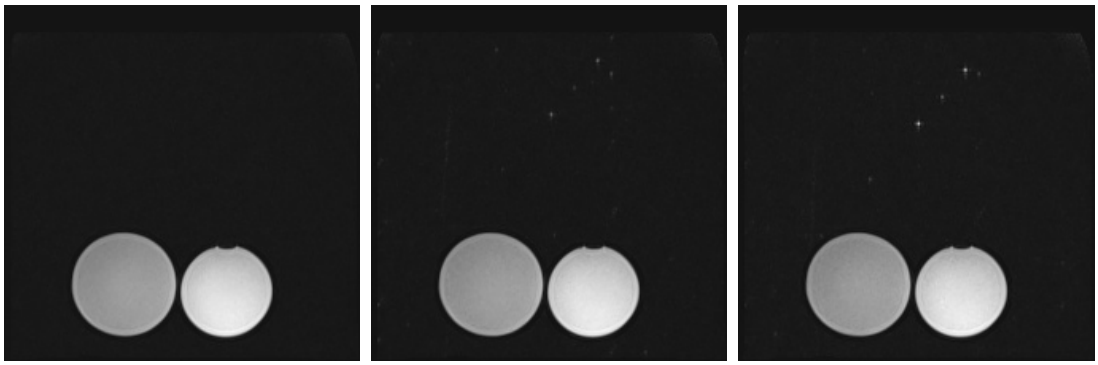
Due to its low cost, good accuracy and ease of use, the RGB-D sensor is a valuable additional sensor for integration in the MR image plane alignment workflow as it allows direct 3D measurements of an observed scene. The results found in terms of precision and depth resolution seem to be convenient for our application as the typical working distance between marker and sensor won't exceed 1.6 m where a maximal quantization step of ~8 mm is expected according to Figure 4.4. As currently an RGB-D sensor has not yet been used within an MRI room, compatibility tests of the sensor within a 1.5 T MRI environment are performed.

4.1.2 Use of an RGB-D sensor in the 1.5 T MRI environment

Compatibility tests are performed in order to evaluate functionality and interference with MR images of the RGB-D sensor inside the MRI room. The RGB-D sensor is positioned at a typical working distance of ~1.5 m on the MRI z-axis to the MRI scanner isocenter. Connection of the RGB-D sensor to an external PC inside the MRI control room is realized via an active USB extension cable that is passed through the waveguide of the MRI room.
The RGB-D sensor works normally during and outside MR image acquisition. Furthermore, only minor image artifacts are visible on MR images acquired during operation of the RGB-D sensor (see Figure 4.5). The real-time MRI sequence for clinical interventions is used for assessing the artifacts of the RGB-D sensor in the MR images. The observed artifacts are mainly caused by the passage of the USB extension cable through the waveguide of the MRI room. When the USB cable is passed through the waveguide but neither connected to the external PC nor to the RGB-D sensor, the acquired MR images are attained by the same type of artifacts but less intense than during operation of the RGB-D sensor (compare Figure 4.5 b and c). When the USB cable is removed from the waveguide no artifacts appear. A possible explanation is that the copper-containing USB extension cable compromises the shielding function of the MRI room and introduces noise to it. The observed artifacts do not interfere with the marker detection in the MR images.
Use of an optical fiber and a transducer to convert the USB2 electric signal in an optical signal for transmission of the RGB-D sensor data outside the MRI room could be assessed. In this manner, passing the USB cable through the waveguide and thus comprising the Faraday shield of the MRI room could be avoided in order to suppress the artifacts from the MR images.

4.1.3 Three-dimensional position measurement with the RGB-D sensor

The RGB camera image is formed according to the camera model presented in Appendix B, where a point in camera coordinates (${}^{C}\boldsymbol{P} = [x, y, z]^{T}$) and its projection in image coordinates



(a) (b) (c)

Figure 4.5: Evaluation of MR image artifacts due to operation of the RGB-D sensor inside the MRI room. MR images for three scenarios are depicted: First, no active device is inside the MRI room (a). Second, the USB extension cable is passed through the waveguide of the MRI room with no devices connected at its ends (b). Third, RGB-D sensor and external PC are connected via USB through the waveguide and turned on (c). The image artifacts in (b) and (c) are similar but slightly more intense for (c) where the RGB-D sensor is active.

(${}^I\boldsymbol{P}' = [u, v]^T$) are related through the following equation:

$$\begin{bmatrix} s\,u \\ s\,v \\ s \end{bmatrix} = \boldsymbol{C} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad , \tag{4.1}$$

where parameter s is an arbitrary scale factor. Matrix $\boldsymbol{C}$ is called the calibration matrix and contains the intrinsic parameters of the camera:

$$\boldsymbol{C} = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad . \tag{4.2}$$

The intrinsic parameters of the RGB camera are obtained during one-time calibration (described in section 4.1.4): parameters α_u and α_v are scale factors along the u and v axes of the image frame, respectively. The coordinates of the image center in pixels in the image frame are given by the parameters u_0 and v_0. The image center, also called principal point, is defined by the intersection of the camera optical axis and the image plane. The following equations are thus obtained for the coordinates u and v of point ${}^I\boldsymbol{P}'$:

$$u = \alpha_u \frac{x}{z} + u_0 \tag{4.3}$$

$$v = \alpha_v \frac{y}{z} + v_0 \quad . \tag{4.4}$$

Note that $[u, v]^T$ is given in pixels and $[x, y, z]^T$ in mm.
Inversely, in order to measure a 3D position based on the RGB image we can change equations 4.3 and 4.4 to

$$\frac{x}{z} = \frac{u - u_0}{\alpha_u} \tag{4.5}$$

$$\frac{y}{z} = \frac{v - v_0}{\alpha_v} \quad , \tag{4.6}$$

where only the relations $\frac{x}{z}$ and $\frac{y}{z}$ can be obtained from the image coordinates. Thus, without further information about the observed scene, it is not possible to reconstruct the 3D position of a point, based only on its 2D position from the image plane.
This is where the depth measurement of the RGB-D sensor can be employed. As the RGB-D sensor measures the depth (z) of an observed scene directly in the camera frame, equations 4.5 and 4.6 can be solved:

$$x = \frac{u - u_0}{\alpha_u} \; z \tag{4.7}$$

$$y = \frac{v - v_0}{\alpha_v} \; z \quad . \tag{4.8}$$

Thus, the RGB-D sensor allows to directly measure a 3D position of an object in the camera frame.
It would also be possible to reconstruct the 3D position of an object with a single camera instead of an RGB-D sensor through the intrinsic parameters of the camera and a prior knowledge of one characteristic dimension of the object (e.g. in [BCG07]). This distinctive dimension can be an external size of the object (width, height, radius) or a distance between two features on the object. Such approach is described below.
Assume the known width (in mm) of an object in the camera frame as $x_d = x_1 - x_2$, where x_1 and x_2 are the points attached to the extremities of the object along the x axis of the camera frame. The corresponding distance $u_d = u_1 - u_2$ (in pixels) can be measured in the image. According to eq.4.7, the points in camera frame and their projections in the image plane can be related:

$$x_1 = \frac{(u_1 - u_0)}{\alpha_u} \; z_1 \tag{4.9}$$

$$x_2 = \frac{(u_2 - u_0)}{\alpha_u} \; z_2 \tag{4.10}$$

Thus, if they are written as distance x_d, the following equation is obtained:

$$x_d = x_1 - x_2$$

$$x_d = \frac{(u_1 - u_0)}{\alpha_u} \; z_1 \quad - \quad \frac{(u_2 - u_0)}{\alpha_u} \; z_2 \tag{4.11}$$

If it is assumed that the object is spherical, it is possible to measure its diameter independently from its orientation with respect to the camera. An equal z coordinate is assumed for both

extremities of the object :

$$z = z_1 = z_2 \quad . \tag{4.12}$$

Accordingly eq.4.11 can be simplified to

$$x_d = \frac{z}{\alpha_u} \left[(u_1 - u_0) - (u_2 - u_0)\right] \tag{4.13}$$

$$x_d = \frac{z}{\alpha_u} (u_1 - u_2) \quad , \tag{4.14}$$

which can be written as

$$z = \frac{x_d \, \alpha_u}{u_d} \quad . \tag{4.15}$$

It is thus possible to calculate the z coordinate (in mm) of the object in the camera frame by measuring its size u_d (in pixels) in the image frame. With the obtained z coordinate, eq.4.7 and eq.4.8 can be applied in order to reconstruct the 3D position of the object.

Nevertheless, the precision of the camera approach depends highly on the accuracy of the object segmentation in the image, the object size and the image resolution. The additional depth measurement from the RGB-D sensor is thus clearly an advantage over this method.

4.1.4 RGB camera calibration

Camera calibration allows to determine the intrinsic parameters of a camera. Zhang's method ([Zha00]) was used for calibration of the RGB camera of the RGB-D sensor.
This technique allows to calibrate a camera through observation of a planar pattern at a few (at least two) different orientations. The feature points of the pattern are detected in the obtained images and the intrinsic and extrinsic parameters of the camera are obtained through an analytical solution. The image deformation due to radial distortion is then estimated. Finally, the determined parameters are refined through a non-linear optimization approach.
At the end of this calibration method, the intrinsic parameters of the camera are obtained, namely the coordinates of the principal point (u_0, v_0), the scale factors α_u and α_v in image coordinates (pixels) and the skewness factor γ of the image axes. The skewness factor defines the angle between the image x-axis and y-axis. The coefficients of the radial distortion k_1, k_2 are also obtained. The radial distortion due to the optical system induces a shift of the pixel position compared to a perfect pinhole model (more details in annex B).
Following this method, the results presented in Table 4.1 are obtained for the RGB camera of the Asus Xtion Pro Live RGB-D sensor. The calibration has been performed for two resolutions of the camera. For each resolution a checkerboard calibration pattern has been observed in 13 different orientations.

Resolution	640 × 480 pixels	1280 × 1024 pixels
Coordinates principal point (u_0, v_0) in pixels	313.25, 241.104	634.549, 513.768
Scale factors (α_u, α_v) in pixels	540.519, 539.709	1082.703, 1081.662
Skewness factor (γ)	0	0
Radial distortion coefficients (k_1, k_2)	0.055, 0.106	0.029, -1.028

Table 4.1: Results for calibration of RGB camera of Asus Xtion Pro Live according to Zhang's method described in [Zha00].

4.2 Hybrid image plane alignment workflow

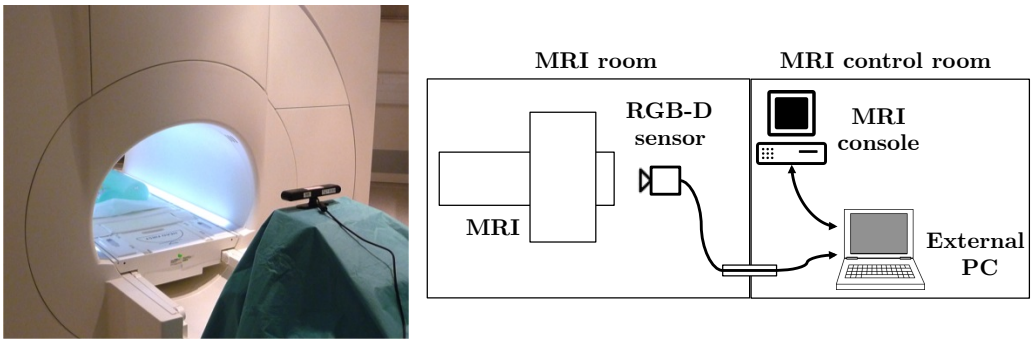
A hybrid workflow for automatic scan plane alignment has been developed combining the passive workflow presented in chapter 2 with an active tracking approach based on measurements of an RGB-D sensor. Main drawback of the purely MR image based approach is the coupling of the tracking speed to the slow image acquisition rate of the clinical MRI sequence. The inherent drawback of an RGB-D sensor based approach is the need for an unobstructed line-of-sight between sensor and tracking marker. Combination of both approaches is expected to compensate for their respective drawbacks and to combine their advantages. The hybrid approach allows to use the MRI or the RGB-D sensor either as stand-alone tracking modalities or to use them simultaneously by fusing their measurements through use of an Information filter.

4.2.1 RGB-D sensor installation

Due to the need for a line-of-sight between the RGB-D sensor and the marker, best position of the RGB-D sensor is in the back of the MRI scanner at a typical distance of about 1.5 m between the sensor and the MRI scanner isocenter along the MRI z-axis (Figure 4.6a). Positioned there, the risk for occlusion of the marker through the person holding it, is minimized. An installation scheme of the RGB-D sensor and its connection to the external PC is presented in Figure 4.6b.

4.2.2 Marker

The marker used for tracking has to be detected in MR images and in RGB-D sensor images. Thus to validate the concept, a cylindrical marker (length: 109 mm, diameter: 40 mm) filled with a contrast agent / water solution (Gd-DTPA 5 mM) and two pink balls (diameter: 40 mm) attached at its distal ends is used (Figure 4.7). The exterior dimensions of the cylindrical marker are slightly bigger than those of the marker presented in section 2.2.2 but the interior dimensions determining the contained liquid volume are exactly the same. The ball positions are symmetric with respect to the center of the marker.
The MR contrast agent solution and the pink balls are well detectable in MR images and RGB images from the RGB-D sensor, respectively.



(a) (b)

Figure 4.6: (a): RGB-D sensor positioned behind MRI scanner inside the MRI room. Typical distance between the RGB-D sensor and the MRI scanner isocenter is about 1.5 m. (b): Schematic of the RGB-D sensor in the MRI environment. An external PC is connected via USB to the RGB-D sensor. The USB connector is passed through the waveguide inside the MRI room.

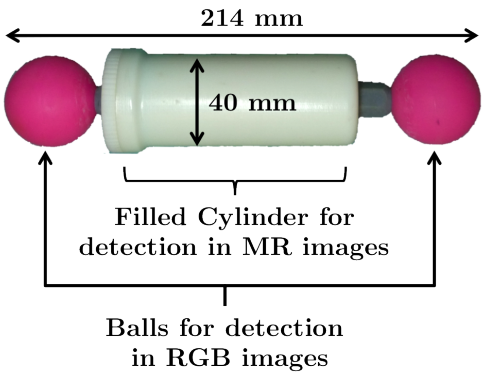


Figure 4.7: Marker containing an MR contrast agent solution for detection in MR images, with two pink balls at its distal ends for detection in RGB images from the RGB-D sensor.

4.2.3 Initialization

The RGB-D sensor is positioned inside the MRI room and connected via an USB cable, passed through the waveguide, to the PC inside the MRI control room. The initial position of the marker inside the MRI tunnel is determined through volume images as presented for the MR image based workflow (section 2.4.1). After alignment of the first real-time scan plane to the marker, the MR image based tracking starts.

4.2.4 Real-time scan plane alignment

An online registration approach is performed during the real-time scan plane alignment stage as described in section 4.4 in order to determine the rigid transformation between the MRI and RGB-D sensor frames. As this transformation is unknown in the beginning, real-time tracking starts based on MR images only: alternating clinical real-time images are aligned according to the detected marker pose in the MR images (section 2.4.2). Simultaneously, the marker pose is detected by the RGB-D sensor. As soon as enough matched marker positions have been acquired in both modalities, the registration can be performed.
After successful registration, the marker pose that is detected with the RGB-D sensor (section 4.3) can be transformed to the MRI frame and is thus available for scan plane alignment, too. The detected 3D poses from the MR image and from the RGB-D sensor are then fed to an Information filter (see section 4.5) that fuses them and predicts the next 3D pose of the marker. Alternately, transversal and sagittal oblique scan planes, aligned to the predicted marker pose, are calculated and corresponding commands are sent to the MRI scanner (Figure 4.8).

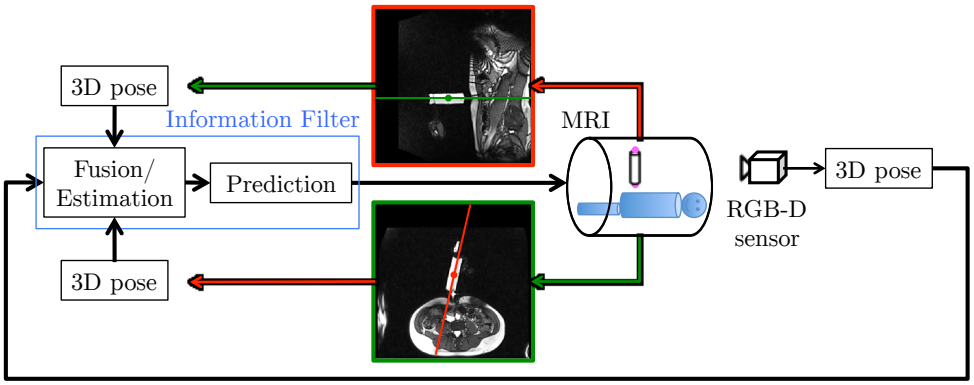


Figure 4.8: Principle of Workflow: two orthogonal image planes, transversal (green) and sagittal (red) are alternately acquired. The 3D marker pose is detected both, in the MR image plane and by the RGB-D sensor. The detected 3D marker pose from the RGB-D sensor is transformed to the MRI frame. Depending on their availability, the 3D poses are then fused in an Information filter and the corresponding command for the position and orientation update of the next image plane is then sent to the MRI.

The working frequency of the Information filter is set to the faster modality which is the RGB-D sensor. Due to this principle and the difference of the acquisition frequencies between the RGB-D sensor and the MRI, the fusion algorithm has to deal with unavailable MRI measurements. Thus, in between two MR image acquisitions the workflow relies on detected 3D poses from the RGB-D sensor, only.
Generally, in case of an unavailable measurement (due to a failed detection, e.g.) from one modality the Information filter will continue working with the measurements of the other modality. When the marker cannot be detected inside both, MR images and RGB-D sensor images, no commands are sent to the MRI console until one of both measurements becomes available again.
As a consequence, the frequency of the sent commands depends on the availability of the measurements. An example of a typical timeline is depicted in Figure 4.9.

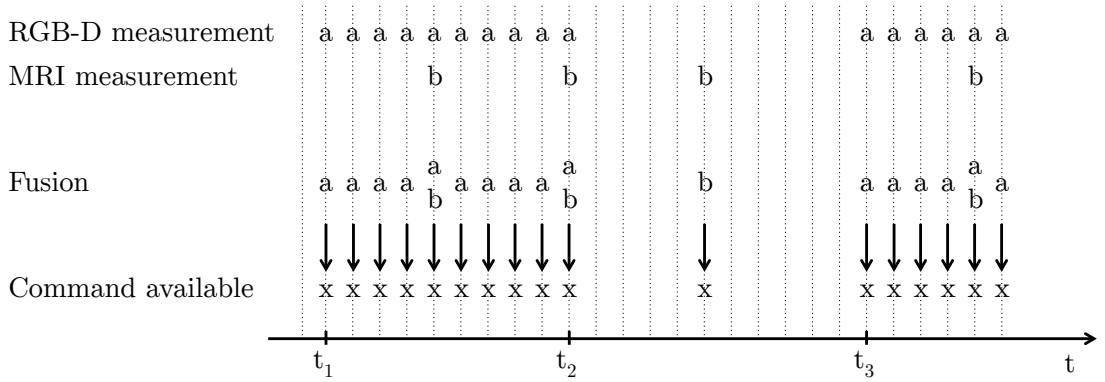


Figure 4.9: Example timeline of commands that are sent to the MRI console PC depending on the available measurements. Between t_1 and t_2 both RGB-D (a) and MRI (b) measurements are available. The measurement fusion algorithm fuses them if an MR image is acquired during an RGB-D image acquisition period. During the intervals where MRI measurements are not available, the measurement fusion algorithm only accounts for the available RGB-D measurements. Between t_2 and t_3 MRI the measurement fusion algorithm only accounts for MRI measurements as RGB-D measurements are unavailable. From t_3 on, all measurements are available again. Commands are only sent to the MRI console PC if at least one modality delivers available measurements.

4.3 Detection algorithm for RGB-D sensor images

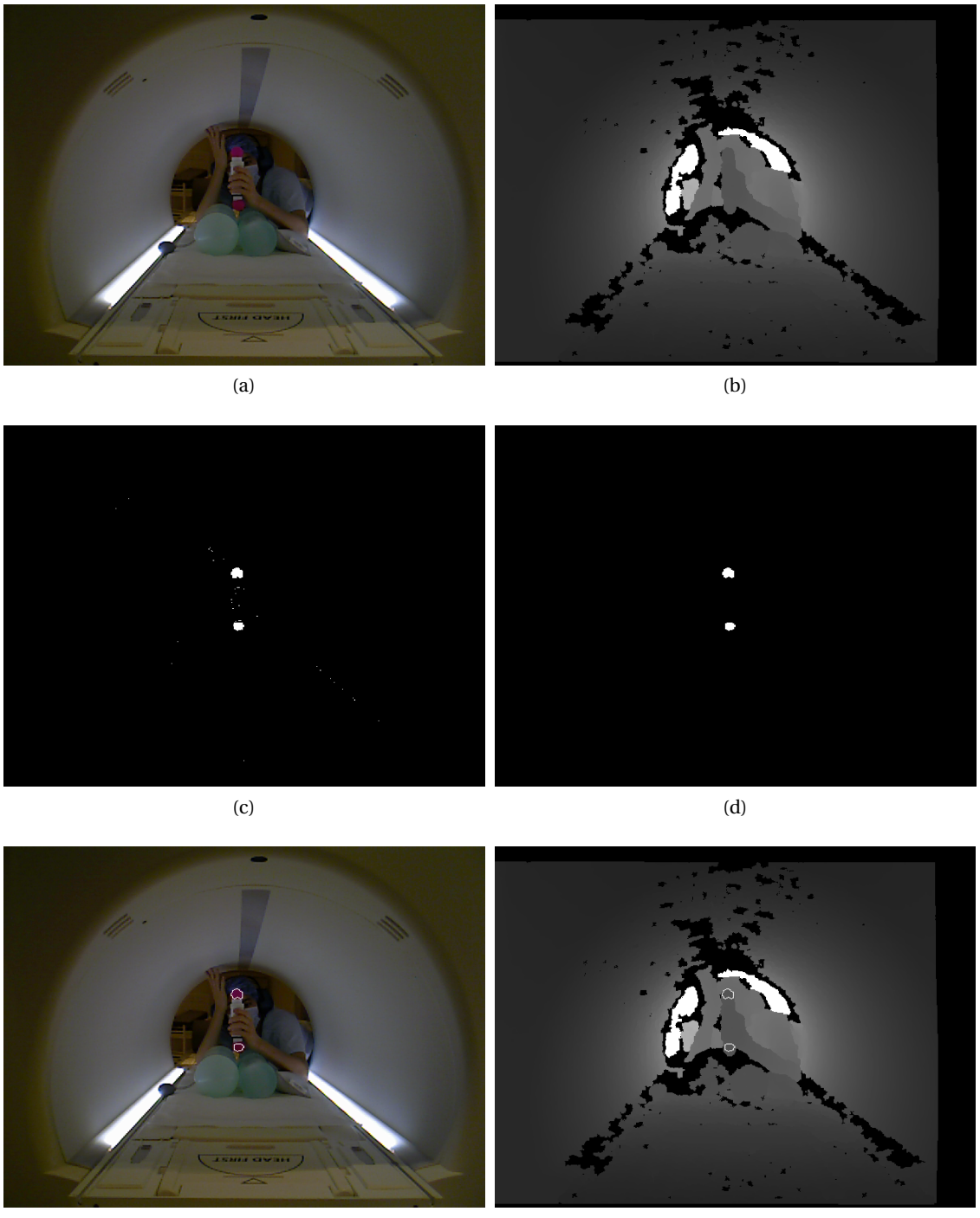
In order to measure the 3D pose of the marker with the RGB-D sensor, positions of the two pink balls are detected in the images and their 3D positions are calculated in the camera frame as described in section 4.1.3. Note that the RGB-D sensor operates with its minor resolution of 640 × 480 pixels for RGB images as this corresponds directly to the maximal resolution of the depth images. In this manner the RGB and depth images are temporally synchronized and acquired with the higher frame rate of 30 frames per second. Using a higher resolution for the

RGB images would decrease their frame rate to 15 frames per second.
Detection of the two pink balls starts with the transformation of the RGB image (Figure 4.10a) to Hue Saturation Value (HSV) color space and the application of an experimentally determined threshold. The HSV space revealed as the most intuitive color representation for the research of a threshold. After thresholding (Figure 4.10c) a morphological opening is performed (Figure 4.10d) using a circular structuring element with a diameter of 5 pixels in order to erase potential noise.
Then a contour detection algorithm is applied (Figure 4.10e) performing a least-square fitting of ellipses to the segmented contours ([FPF99]). The length of the major and minor axes of the ellipses are then verified with respect to an admitted ball diameter interval. This interval is calculated based on the diameter of the ball in the camera frame, on its maximal and minimal allowed distance to the RGB-D sensor and on the scale factors of the image axes. The scale factors are obtained during camera calibration and may be different for the image axes, but are assumed equal for this calculation (for further details about camera calibration see section 4.1.4). The minimal and maximal ball diameter is thus determined for the u axis of the image frame and adopted for the v axis:

$$u_{min} = \frac{x_{ball}\,\alpha_u}{z_{max}} \tag{4.16}$$

$$u_{max} = \frac{x_{ball}\,\alpha_u}{z_{min}} \quad . \tag{4.17}$$

The variables u_{min} and u_{max} are the minimal and maximal ball diameter (in pixels), respectively. The ball diameter (in mm) is given by x_{ball}, the minimal and maximal admitted distance between the camera and the ball by z_{min} (800 mm) and z_{max} (2000 mm), respectively.
If the balls are segmented, the 2D positions inside the image frame of the centers of their fitted ellipses is determined. In order to reconstruct the 3D positions of the balls, as described in section 4.1.3, the depth of the segmented balls must be determined. For this purpose, the depth image (Figure 4.10b) of the scene is accessed and the depth pixels with the same position than those inside the detected contours in the RGB image are verified (Figure 4.10f). First must be verified if they are valid values as the depth map can be incomplete due to a failed reconstruction of the depth by the RGB-D sensor. This typically happens for strongly reflecting surfaces where the emitted IR pattern cannot be captured by the RGB-D sensor. For both balls the depth values inside their detected contours are collected and their median values are determined. The median has revealed as a better depth measurement than the mean value of all collected depth values. This is mainly due to its robustness against depth value outliers at the edges of the balls, where the projected IR pattern cannot be detected on the balls themselves but potentially on a reflecting surface behind the ball (Figure 4.11).
These depth measurements are used for reconstruction of the 3D positions of the balls. Note that with greater distance between the ball and the RGB-D sensor, the number of available depth values for the ball decreases due to the resolution of the structured light pattern used for depth reconstruction. Furthermore, as not the whole surface of an object can be detected on the depth map, but only the side facing the RGB-D sensor, an offset in depth between the



(a) (b)

(c) (d)

(e) (f)

Figure 4.10: RGB (a) and depth image (b) of a typical workflow scene: the marker is held inside the MRI scanner by a person who can control the scan planes through marker motions. The RGB image is transformed to HSV color space and thresholded (c). A morphological opening is performed on the thresholded image in order to reduce noise from thresholding (d). The ball contours are highlighted in the RGB image (e) and the corresponding pixels in the depth image (f) are collected for calculation of their median value.

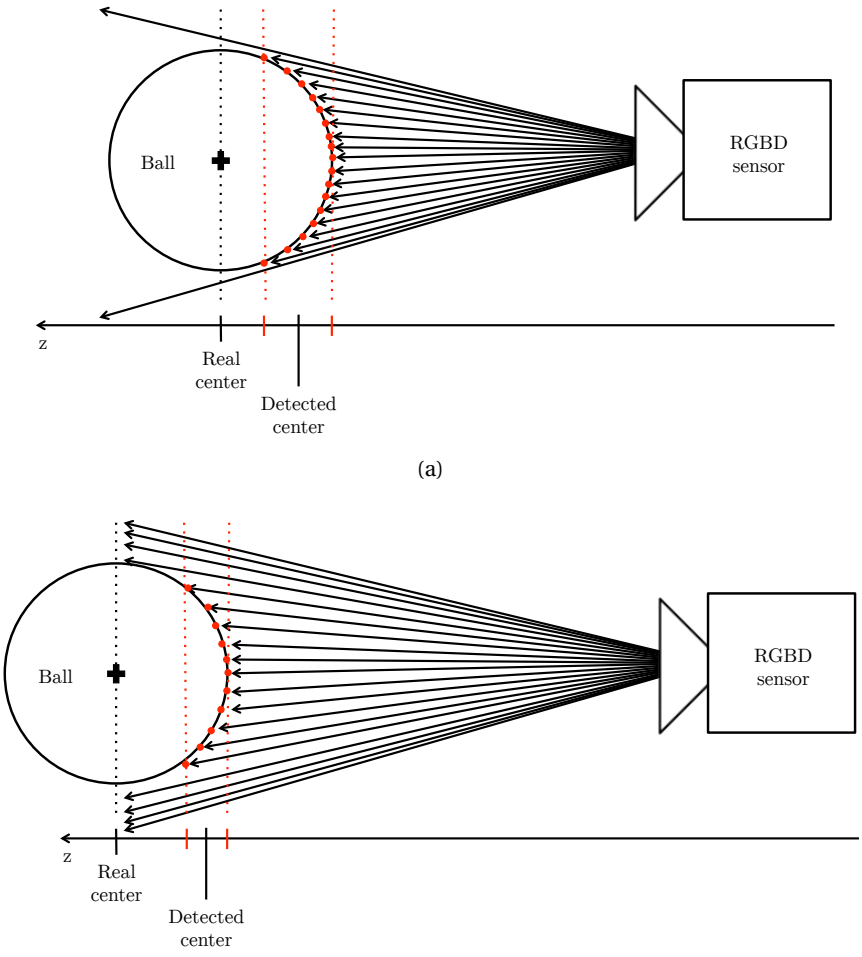


Figure 4.11: Simplified representation in one plane of the projection of the structured light pattern on the ball. The red points can be reconstructed on the ball's surface and correspond thus to its actual depth values.The edges of the ball cannot be reconstructed in this manner due to their lack of reflection of the structured light pattern on the ball surface. The reconstructed depth value of the ball edges on the depth image might thus correspond to a reflecting surface behind the ball. (a): Ball at a small distance from RGB-D sensor. A small shift between the real center and the available values for calculation of the detected center on the z axis can be observed. (b): Ball at a greater distance from RGB-D sensor. The number of available depth points has decreased and the shift between real ball center and the available values for calculation of the detected center has increased.

real center of the ball and its detected center is introduced. The offset is a function of the distance between the ball and the RGB-D sensor (Figure 4.11). During the registration process of the RGB-D sensor with the MRI (section 4.4), a mean value for this offset is identified as part of the translation between the MRI frame and the RGB-D sensor frame. This offset is then assumed to be constant over the depth range of the marker workspace.
The orientation of the marker is computed by defining a vector from the 3D position of the upper ball pointing towards the position of the lower ball. The 3D marker position is defined as the middle position between the two detected ball positions.

4.4 Registration between MRI and RGB-D sensor

As presented in [Tro12], registration consists in the determination of a geometrical relation between two frames using a transformation. Rigid transformations include rotations and translations. The transformation can also be non-rigid, which means that it is non-Euclidean and searched between objects among whose at least one is deformable. In order to register the MRI frame with the RGB-D sensor frame the 3D rigid transformation between both frames has to be determined. This transformation is represented with three parameters for the translation and three parameters for the rotation.
The registration between the MRI frame and the RGB-D sensor frame can be performed offline or online. As presented in [VBR+07] and [Daa01], offline registration between a camera system and an MRI consists in determination of the transformation between them and can be performed with a dedicated registration frame. The transformation is determined once after installation of the camera system and is thus constant over time. The drawback of an offline registration is that in case of a displacement of the camera, the found transformation between the camera and the MRI frame is no longer valid and the registration step has to be repeated. As a permanent installation of the RGB-D sensor behind the MRI machine is currently not compatible with the clinical workflow and thus not feasible, an offline registration would have to be performed before every intervention.
On the contrary, an online registration consists in finding and updating the transformation between the MRI frame and the RGB-D sensor frame online during the intervention. Initially, the intervention is performed using MRI data only. At the beginning of the intervention, a number of matched points (at least three) has to be acquired in both the MRI frame and the RGB-D sensor frame in order to determine the 3D rigid transformation between the point sets. In our case, a pair of matched points is composed of the marker center coordinates in both modalities and detected during the same RGB-D acquisition period. After successful registration, the intervention is performed using data from both modalities. An advantage of the chosen approach is that the point sets in both frames are extended and updated during the entire intervention allowing to improve and refine the registration result over time. Hence, the online registration allows to change the position of the RGB-D sensor at any time during a procedure in order to achieve an optimal line-of-sight between sensor and intervention zone. Due to its flexibility, an online registration approach has been proposed.

The transformation between the MRI frame and the RGB-D sensor frame is given by the homogenous matrix

$$ {}^{MRI}\boldsymbol{M}_{RGB-D} = \begin{bmatrix} {}^{MRI}\boldsymbol{R}_{RGB-D} & {}^{MRI}\boldsymbol{t}_{RGB-D} \\ 0 & 1 \end{bmatrix} \quad , \tag{4.18} $$

where ${}^{MRI}\boldsymbol{R}_{RGB-D}$ is the rotation matrix and ${}^{MRI}\boldsymbol{t}_{RGB-D}$ the translation between the frames (Figure 4.12).

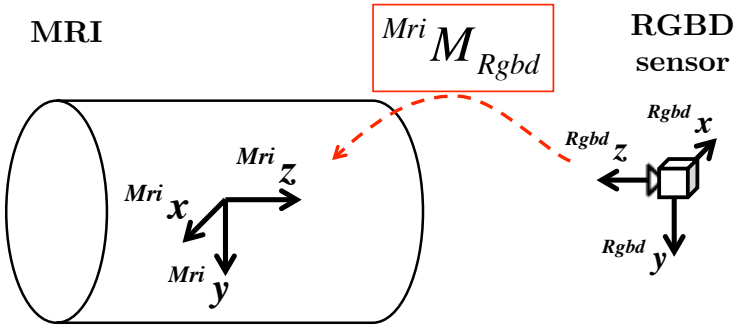


Figure 4.12: Spatial relationship between MRI and RGB-D sensor frame that is determined during registration and represented in the homogeneous matrix ${}^{Mri}\boldsymbol{M}_{Rgbd}$.

In [HM11], Horaud and Monga present the registration problem as the search for an optimal correspondence transformation between two 3D point clouds describing the same observations but acquired in different frames. The ordered lists of 3D points $\boldsymbol{C} = [C_1, C_2, ..., C_n]$ and $\boldsymbol{O} = [O_1, O_2, ..., O_n]$ are represented in the frames of their respective sensors $\mathcal{F}_C$ and $\mathcal{F}_O$. The research for the transformation between the two point clouds is presented as the following optimization problem:

$$ Q^* = \min_{\boldsymbol{R},\boldsymbol{t}} \sum_{i}^{n} |C_i - {}^{C}\boldsymbol{R}_O \, O_i - {}^{C}\boldsymbol{t}_O|^2 \quad , \tag{4.19} $$

where Q^* depends on the rotation ${}^{C}\boldsymbol{R}_O$ and translation ${}^{C}\boldsymbol{t}_O$. When expressing the sets of coordinates, C_i and O_i, with respect to their respective centers of mass, C_{mc} and O_{mc}, this cost function can be split and the optimal rotation and translation are obtained through the following equations

$$ \min_{\boldsymbol{R}} \sum_{i}^{n} |C'_i - {}^{C}\boldsymbol{R}_O \, O'_i|^2 \quad , \tag{4.20} $$

and

$$ \boldsymbol{t} = C_{mc} - {}^{C}\boldsymbol{R}_O \, O_{mc} \tag{4.21} $$

with

$$C_{mc} = \frac{1}{n}\sum_{i}^{n} C_i \tag{4.22}$$

$$O_{mc} = \frac{1}{n}\sum_{i}^{n} O_i \tag{4.23}$$

$$C'_i = C_i - C_{mc} \tag{4.24}$$

$$O'_i = O_i - O_{mc} \quad . \tag{4.25}$$

Thus, in a registration step, we typically start by searching for the optimal rotation between the frames, before computing the translation.

4.4.1 Determination of the optimal rotation between the frames

Based on eq.4.20, Horaud and Monga demonstrate that at least three non aligned matched points are needed for the determination of all rotation parameters between two distinct frames. When this condition is met, matched point sets can be inserted in the following equation.

$$\underbrace{\begin{bmatrix} \vdots & \vdots & \vdots \\ 0 & -(O'_{iz}+C'_{iz}) & (O'_{iy}+C'_{iy}) \\ (O'_{iz}+C'_{iz}) & 0 & -(O'_{ix}+C'_{ix}) \\ \vdots & \vdots & \vdots \end{bmatrix}}_{\boldsymbol{M}} \underbrace{\begin{bmatrix} N_x \\ N_y \\ N_z \end{bmatrix}}_{\boldsymbol{N}} = \underbrace{\begin{bmatrix} \vdots \\ C'_{ix}-O'_{ix} \\ C'_{iy}-O'_{iy} \\ \vdots \end{bmatrix}}_{\boldsymbol{P}} \quad . \tag{4.26}$$

The least-square solution to eq.4.26 is

$$\boldsymbol{N} = \boldsymbol{M}^{+}\boldsymbol{P} \quad , \tag{4.27}$$

where $\boldsymbol{M}^{+}$ is the pseudo-inverse of matrix $\boldsymbol{M}$. This equation allows to find the rotation axis $\boldsymbol{n}^{*}$ and rotation angle θ^{*}, given by

$$\boldsymbol{n}^{*} = \frac{\boldsymbol{N}}{|\boldsymbol{N}|} \tag{4.28}$$

$$\theta^{*} = \pm 2atan(|\boldsymbol{N}|) \quad . \tag{4.29}$$

The residual error can be evaluated through calculation of the rmse between both point clouds, with the point clouds being represented in the same frame thanks to the found transformation. The residual error is used for determination of the sign of the rotation angle. For this purpose, the residual error is computed for both transformations, one with a positive and one with a negative rotation angle. The transformation resulting in a better residual error between the point clouds is the one with the correct rotation angle. The according rotation matrix ${}^{C}\boldsymbol{R}_{O}$ between both frames can then be obtained through the Rodrigues formula.

The method for determination of the optimal rotation of Horaud and Monga is based on a work for robot hand/eye calibration, presented in [TL89] by Tsai and Lenz. They explain that this method has an exceptional case for a rotation of $\pm\pi$ around the rotation axis. In this case the resulting vectors of the additions of O'_i and C'_i are collinear for all distinct point sets i and eq.4.26 cannot be used. The rotation axis is then given directly by $O'_i + C'_i$ and the rotation angle is π. Nevertheless, in practice it is complicated to verify this collinearity for all distinct point sets, due to the measurement noises of both modalities.
For this reason and in order to use this approach regardless of its exceptional case, an avoidance mechanism is implemented. This mechanism consists in searching for the approximate rotation between the frames $\mathcal{F}_C$ and $\mathcal{F}_O$ before searching for the optimal rotation using eq.4.26 and eq.4.27. If the approximate rotation angle is near $\pm\pi$, the points linked to $\mathcal{F}_O$ are represented in an intermediate frame $\mathcal{F}_I$ allowing to avoid the singularity. The rotation according to the method here above is then calculated between $\mathcal{F}_C$ and $\mathcal{F}_I$. The final rotation between $\mathcal{F}_C$ and $\mathcal{F}_O$ can then be derived.
The approximate solution ([Nag13]) to the registration problem is based on the definition of a frame $\mathcal{F}_{eig}$ that is linked to the point cloud. Three non-aligned matched points are used for definition of this frame with the vectors

$$\begin{aligned} \boldsymbol{u}_O &= \frac{O_1\,O_2}{|O_1\,O_2|}\,, \quad \boldsymbol{w}_O = \frac{O_1\,O_3}{|O_1\,O_3|}\,, \quad \boldsymbol{v}_O = \boldsymbol{w}_O \times \boldsymbol{u}_O\,, \\ \boldsymbol{u}_C &= \frac{C_1\,C_2}{|C_1\,C_2|}\,, \quad \boldsymbol{w}_C = \frac{C_1\,C_3}{|C_1\,C_3|}\,, \quad \boldsymbol{v}_C = \boldsymbol{w}_C \times \boldsymbol{u}_C \qquad . \end{aligned}$$

The rotation matrices ${}^O\boldsymbol{R}_{eig} = [\boldsymbol{u}_O, \boldsymbol{v}_O, \boldsymbol{w}_O]$ and ${}^C\boldsymbol{R}_{eig} = [\boldsymbol{u}_C, \boldsymbol{v}_C, \boldsymbol{w}_C]$ represent the rotation between $\mathcal{F}_{eig}$ and the frames $\mathcal{F}_O$ and $\mathcal{F}_C$, respectively. The approximate rotation matrix between the frames $\mathcal{F}_C$ and $\mathcal{F}_O$ can thus be calculated with

$${}^C\boldsymbol{R}_{O\ approx} = {}^C\boldsymbol{R}_{eig}\left({}^O\boldsymbol{R}_{eig}\right)^T \qquad . \tag{4.30}$$

Note that this method is referred to as approximate as it relies only on three matched points instead of all the available points and is thus highly sensitive to measurement errors on the chosen points. The rotation matrix ${}^C\boldsymbol{R}_{O\ approx}$ is then converted to the axis-angle representation in order to determine the rotation angle between the frames $\mathcal{F}_C$ and $\mathcal{F}_O$. If this angle is within ± 0.2 rad near $\pm\pi$, the aforementioned method is near its exceptional case. In order to avoid the exception, the point cloud $\boldsymbol{O}$ is represented in an intermediate frame $\mathcal{F}_I$, rotated by $-\pi$ around the approximate rotation axis. For this purpose the approximate rotation axis is multiplied with π and then converted to the rotation matrix ${}^I\boldsymbol{R}_O$. The points $\boldsymbol{O}$ are then represented in the frame $\mathcal{F}_I$ with

$${}^I\boldsymbol{O} = {}^I\boldsymbol{R}_O\,O \qquad . \tag{4.31}$$

The rotation matrix ${}^C\boldsymbol{R}_I$ is then searched between the point clouds $\boldsymbol{C}$ and ${}^I\boldsymbol{O}$ according to eq.4.26 ff. Eventually, the final rotation matrix between the point clouds $\boldsymbol{C}$ and $\boldsymbol{O}$ can be

calculated by accounting for the intermediate frame in the following manner:

$$ {}^{C}\boldsymbol{R}_{O} = {}^{C}\boldsymbol{R}_{I}\,{}^{I}\boldsymbol{R}_{O} \quad . \tag{4.32} $$

4.4.2 Determination of translation between the frames

The translation between the two frames is then obtained by computing eq.4.21 with the previously obtained rotation.

4.4.3 Registration point list

The matched points and their spatial distribution heavily influence the quality of the registration. It is therefore important to dispose of point sets that are spatially spread but not too distant to the MRI scanner isocenter as MR image quality (and thus quality of MRI measurements) decreases with increasing distance to the isocenter, due to magnetic field inhomogeneity.
Algorithm 1 shows the chosen approach for management of the point list. The maximal number of points has been set to 20. Arriving points are verified with respect to their distance to the MRI isocenter and the distance to the other points in the list. If the list is already full, it is tested if the arriving point would improve spatial distribution of the point list. If so, an existing point in the list will be replaced by the arriving point.

Algorithm 1: Pointlist update

Data: Point list $\mathcal{L}$ containing points that are used for registration, Arriving matched point $\boldsymbol{P}$ containing RGB-D (${}^{RGB-D}\boldsymbol{P}$) and MRI (${}^{MRI}\boldsymbol{P}$) coordinates , Maximal distance $\boldsymbol{d}_{Isoctr}$ of ${}^{MRI}\boldsymbol{P}$ from MRI isocenter, Minimal distance $\boldsymbol{d}_{min}$ of $\boldsymbol{P}$ to points in $\mathcal{L}$, Current number $numOfPts$ of points in list, Max number $maxListSize$ of points in list

Result: Updated pointlist $\mathcal{L}$

```
begin
    if (abs(MRI P) < d_Isoctr) then
        for P_i in L do
            d_i ←Distance between P_i and P
            if d_i < d_min then
                exit
            end
            d_PtList ← d_PtList ∪ d_i
        end
        if numOfPts < maxListSize then
            Add P to L
            Increment numOfPts
        else
            P_n ← Find nearest neighbor of P in L
            P_n1, P_n2 ← Find 2 nearest neighbors of P_n in L
            distPn_1, distPn_2 ← Distances between P_n and P_n1, P_n2
            distP_1, distP_2 ← Distances between P and P_n1, P_n2
            if (distP_1 > distPn_1) && (distP_2 > distPn_2) then
                Replace P_n by P in L
            end
        end
    end
end
```

4.5 Multi-sensor Data Fusion

In this section, theory of multi-sensor fusion is developed with the aim of combining MRI data with RGB-D sensor data. An improved tracking performance is expected through this fusion.

4.5.1 Introduction to multi-sensor data fusion

Multi-sensor data fusion aims at combining information coming from different sensors in order to improve the estimate of the process that is monitored. Robotics, computer vision, maneuvering of unmanned vehicles, supervision of manufacturing processes and medical applications are fields in which multi-sensor data fusion is commonly used. In [HHL09] an extensive introduction to multi-sensor data fusion, its various techniques, principles and

example applications is given. Khaleghi *etal*. ([KKKR13]) refer to [HHL09] and give a generic overview of contemporary data fusion approaches as well as recent trends in research. The Joint Directors of Laboratories (JDL) model for data fusion with its four layers of abstraction for data fusion is there deemed to be too restrictive and too restrained to military applications. Instead, authors propose a new taxonomy for data fusion methodologies, based on the different data-related challenges in multi-sensor data fusion (Figure 4.13).

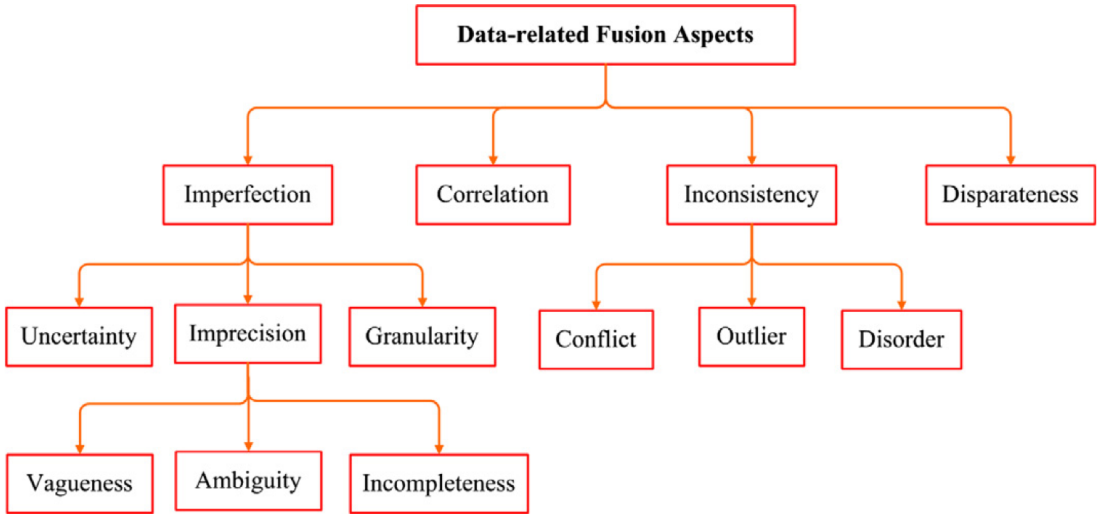


Figure 4.13: Taxonomy for data fusion methods. They are classified according to the four main challenges linked to input data, specified by Khaleghi *et al.* as data imperfection, data correlation, data inconsistency and data disparateness (reprinted from [KKKR13]).

The challenges in data fusion are grouped in imperfection, correlation, inconsistency and disparateness of data. Imperfect data can either be uncertain, imprecise or prone to fusion errors due to different data granularity. Many data fusion algorithms rely either on independence of data or on prior knowledge of its cross-covariances in order to generate consistent estimates. An often encountered problem is that data may be correlated with an unknown covariance, e.g. the observations of different sensors observing different physical variables may be exposed to the same external noise biasing the measurements. If data correlation is not addressed, it can result in biased estimation. Spurious or conflicting data is grouped under the term data inconsistency. Data gathered by different types of sensors (physical sensors, human operators) is referred to as disparate data.
According to Khaleghi *etal*., the most challenging problem is the inherent imperfection of data which is treated in most of the works on data fusion. From this research, different theories have emerged for representation of data imperfection, such as probability theory, fuzzy set theory, possibility theory, rough set theory and Dempster-Shafer evidence theory.
In this introduction, we will restrain the focus on probabilistic fusion methods as they are well adapted for our system and a presentation of all data fusion methods (as realized in [KKKR13]) would exceed the scope of this work. In probabilistic fusion methods, data uncertainty is represented as probability distribution or density functions. These methods are also referred

to as Bayesian Fusion as they rely on the Bayes estimator.
The Kalman filter as a particular case of the Bayes filter as well as the approximations and techniques used for dealing with non-linear systems have been introduced in section 3.1.
Non-linear probability distributions can also be approximated using grid-based methods ([Mah07]), Sequential Monte Carlo (SMC) ([DDFG01]) or Markov Chain Monte Carlo (MCMC) ([Ber04]) algorithms. A widely known recursive implementation of an SMC algorithm is the Particle filter ([CD02]), which approximates the posterior probability of the system state as a weighted sum of random samples. These random samples are predicted through the process model from the prior density and are then converted to predicted measurements through the measurement model. Weights of the random samples are updated according to the likelihood of their corresponding measurements with respect to the actual measurement. Compared to the Kalman filter, a Particle filter may be less performing in terms of computation time, as a large number of random samples may be necessary for a good estimation of the posterior probability.
MCMC algorithms are an alternative to Particle filters for high-dimensional non-linear systems. The probability density of interest is here approximated through samples that are obtained by means of a Markov chain. The Markov chain generates the samples through a transition kernel with Markovian property, meaning that the transition probabilities between samples depends only on their current state. After a certain number of iterations a well-designed Markov chain converges to a stationary density of interest ([Ber04]). Well known algorithms are the Metropolis-Hastings ([Has70]) algorithms and the Gibbs sampling algorithm ([CG92]).

As the treated data fusion problem is neither high-dimensional nor a multi-target tracking application, we decided to choose a simple, easy-to-implement and well studied data fusion algorithm, which is the Kalman filter. Furthermore, as a random free-hand motion of the marker is tracked, we do not possess a reliable process model. Thus we assume a simple linear model for which the Kalman filter is an appropriate choice. The Kalman Filter algorithm and some of its tracking applications are presented in sections 3.1.1 and 3.1, respectively. The Kalman filter has already been used in different variations for multi-sensor fusion ([RDW91], [LOCBR00], [DMC00], [WCD76]).
The Information filter is another form of the Kalman filter but is more adapted to multi-sensor fusion problems than the Kalman filter itself. Mutambara derives the Information filter from the Kalman filter and presents its characteristics in great detail in [Mut98]. Algebraically the Information filter is identical to the Kalman Filter, but instead of estimating a state directly, it estimates the amount of information about a state that is contained in a measurement (*Fisher* information). According to Mutambara, the main advantages of the Information filter over the Kalman filter are the following:

- The estimation equations for the Information filter are simpler than those for the Kalman filter and can be easily partitioned in order to use them for decentralized multi-sensor data fusion.
- Initialization of the Information filter is very simple as the information estimates (matrix and state) can be easily set to zero information in the beginning. In practice though, in

order to make the Information matrix invertible it is initialized with very small non-zero diagonal elements.
- If the state dimension is smaller than the observation dimension (which is often the case in multi-sensor systems), then the matrices that have to be inverted in the Information filter are smaller than the matrices in the Kalman filter.

In [SK08], Durrant-Whyte reveals the simplicity of the update stage of the Information filter as its main advantage over the Kalman filter: "*For a system with n sensors, the fused information state update is exactly the linear sum of information contributions from all sensors*". As a consequence, the Information filter is a highly flexible and robust algorithm that can deal easily with unavailable sensor measurements. This is a crucial point, as the presented workflow fuses information coming from two different modalities with very different acquisition rates (Figure 4.9).
Despite its potential, the Information filter has not found many applications and is not very well covered in the literature. In [BSLK04] and [May82] the Information filter is briefly presented. Grocholsky discussed the filter in more detail in the context of multi-sensor multi-vehicle systems ([Gro02]). Another application of the Information filter presented by Prat in [Pra10] is the fusion of data of a laser scanner and a camera in the context of object and hazard detection in driving context.

4.5.2 Information filter

The Information filter is simply a formulation of the Kalman filter that uses information measurements to represent the states of the system. The implementation of the Information filter proposed in [MDW93] and [Mut98] was selected. Analogically to the Kalman filter (section 3.1.1), the following linear process model (without control input) is assumed

$$\boldsymbol{x}(k) = \boldsymbol{A}(k)\,\boldsymbol{x}(k-1) + \boldsymbol{n}(k-1) \quad , \tag{4.33}$$

where $\boldsymbol{x}(k)$ is the process state and $\boldsymbol{A}$ is the transition matrix (relating process state from time $k-1$ to k), with a measurement equation

$$\boldsymbol{z}(k) = \boldsymbol{H}(k)\,\boldsymbol{x}(k) + \boldsymbol{m}(k) \quad . \tag{4.34}$$

Measurement $\boldsymbol{z}(k)$ is obtained by multiplying the process state with the observation matrix $\boldsymbol{H}(k)$. Variables $\boldsymbol{n}(k)$ and $\boldsymbol{m}(k)$ represent process noise and measurement noise, respectively. Their Gaussian probability distributions are

$$\boldsymbol{n}(k) \sim N(0, \boldsymbol{Q}(k)) \quad . \tag{4.35}$$

$$\boldsymbol{m}(k) \sim N(0, \boldsymbol{R}(k)) \tag{4.36}$$

$\boldsymbol{R}(k)$ and $\boldsymbol{Q}(k)$ are the measurement noise covariance and the process noise covariance matrices, respectively.

Instead of the state estimate $\hat{\boldsymbol{x}}$ and the estimate error covariance matrix $\boldsymbol{P}$, that are the internal variables of the Kalman filter, the Information filter uses the information state vector $\hat{\boldsymbol{y}}$ and the information matrix $\boldsymbol{Y}$. These variables are related through the following equations:

$$\boldsymbol{Y} = \boldsymbol{P}^{-1} \tag{4.37}$$

$$\hat{\boldsymbol{y}} = \boldsymbol{Y}\,\hat{\boldsymbol{x}} \quad . \tag{4.38}$$

The prediction and estimation steps of an Information filter for a system of N sensors can then be represented as follows (from [Mut98]):

Prediction Step

$$\hat{\boldsymbol{y}}(k|k-1) = \boldsymbol{L}(k|k-1)\,\hat{\boldsymbol{y}}(k-1|k-1) \tag{4.39}$$

$$\boldsymbol{Y}(k|k-1) = \left[\boldsymbol{A}(k)\,\boldsymbol{Y}^{-1}(k-1|k-1)\,\boldsymbol{A}^{T}(k) + \boldsymbol{Q}(k)\right]^{-1} \tag{4.40}$$

with local propagation coefficient $\boldsymbol{L}(k|k-1)$ given by:

$$\boldsymbol{L}(k|k-1) = \boldsymbol{Y}(k|k-1)\,\boldsymbol{A}(k)\,\boldsymbol{Y}^{-1}(k-1|k-1)$$

Estimation Step

$$\hat{\boldsymbol{y}}(k|k) = \hat{\boldsymbol{y}}(k|k-1) + \sum_{j=1}^{N} \boldsymbol{i}_j(k) \tag{4.41}$$

$$\boldsymbol{Y}(k|k) = \boldsymbol{Y}(k|k-1) + \sum_{j=1}^{N} \boldsymbol{I}_j(k) \tag{4.42}$$

with local information state contribution $\boldsymbol{i}_j(k)$ and the associated information matrix $\boldsymbol{I}_j(k)$:

$$\begin{aligned} \boldsymbol{i}_j(k) &= \boldsymbol{H}_j^{T}(k)\,\boldsymbol{R}_j^{-1}(k)\,\boldsymbol{z}_j(k) \\ \boldsymbol{I}_j(k) &= \boldsymbol{H}_j^{T}(k)\,\boldsymbol{R}_j^{-1}(k)\,\boldsymbol{H}_j(k) \end{aligned} \tag{4.43}$$

The algorithm is initialized with initial information estimates $\hat{\boldsymbol{y}}(k|k)$ and $\boldsymbol{Y}(k|k)$ and starts with the prediction step. One can observe that the estimation of the information state is an algebraic sum of the information contributions of the different sensors and of the predicted information state (eq.4.41, eq.4.42). This is why the Information filter is so well adapted to multi-sensor fusion problems and why it is used in our case.

4.5.3 Information filter for fusion of tracking data from RGB-D sensor and MR images

Adapted to our application the state of the process corresponds to the marker pose $\boldsymbol{x}_{Pose} = [x, y, z, \alpha, \beta]^T$ and its derivative :

$$\boldsymbol{x}(k) = \begin{bmatrix} \boldsymbol{x}_{Pose}(k) \\ \dot{\boldsymbol{x}}_{Pose}(k) \end{bmatrix} \tag{4.44}$$

A constant velocity model is chosen as the process model. Thus the transition matrix is given by:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_{5\times5} & \delta t \cdot \boldsymbol{I}_{5\times5} \\ \boldsymbol{0}_{5\times5} & \boldsymbol{I}_{5\times5} \end{bmatrix} \quad , \tag{4.45}$$

where δt is the time step between 2 acquisitions of the RGB-D sensor , $\boldsymbol{I}_{5\times5}$ is a 5×5 identity matrix and $\boldsymbol{0}_{5\times5}$ is a 5×5 zero matrix.
The marker pose is directly measured in the MR and RGB-D images. Their observation matrices are therefore equal:

$$\boldsymbol{H} = \boldsymbol{H}_{MRI} = \boldsymbol{H}_{RGB-D} = \begin{bmatrix} \boldsymbol{I}_{5\times5} & \boldsymbol{0}_{5\times5} \end{bmatrix} \tag{4.46}$$

In our two sensor case, the sums of the local information states and information matrices (eq.4.41 and eq.4.42) can be expressed as:

$$\begin{aligned} \sum_{j=1}^{N} \boldsymbol{i}_j(k) &= \boldsymbol{i}_{MRI}(k) + \boldsymbol{i}_{RGB-D}(k) \\ &= \boldsymbol{H}_{MRI}^T(k)\, \boldsymbol{R}_{MRI}^{-1}(k)\, \boldsymbol{z}_{MRI}(k) + \boldsymbol{H}_{RGB-D}^T(k)\, \boldsymbol{R}_{RGB-D}^{-1}(k)\, \boldsymbol{z}_{RGB-D}(k) \end{aligned} \tag{4.47}$$

$$\begin{aligned} \sum_{j=1}^{N} \boldsymbol{I}_j(k) &= \boldsymbol{I}_{MRI}(k) + \boldsymbol{I}_{RGB-D}(k) \\ &= \boldsymbol{H}_{MRI}^T(k)\, \boldsymbol{R}_{MRI}^{-1}(k)\, \boldsymbol{H}_{MRI}(k) + \boldsymbol{H}_{RGB-D}^T(k)\, \boldsymbol{R}_{RGB-D}^{-1}(k)\, \boldsymbol{H}_{RGB-D}(k) \end{aligned} \tag{4.48}$$

In case of an unavailable measurement due to a failed detection for example, the corresponding term of the unavailable sensor measurement is simply retracted from these sums.
The impact of the sensors on the Information filter can be controlled by weighting their measurement noise covariance matrices ($\boldsymbol{R}_{MRI}$ and $\boldsymbol{R}_{RGB-D}$). The optimal weights for both sensors have been determined experimentally.

4.6 System architecture

Communication between the MRI console PC and an external PC is performed in the same manner as described in section 2.7. Additionally, the connection between RGB-D sensor and

external PC is performed via an USB cable, that is passed through the waveguide of the MRI room (Figure 4.6b).
The external PC receives MR images and image information as well as RGB and depth images from the RGB-D sensor. The detection results from both modalities are fused and used for tracking of the marker (Figure 4.14).
The communication interface with the RGB-D sensor was included in the existing C++ program, using open source libraries such as OpenNI, NITE and PCL (Point Cloud Library).

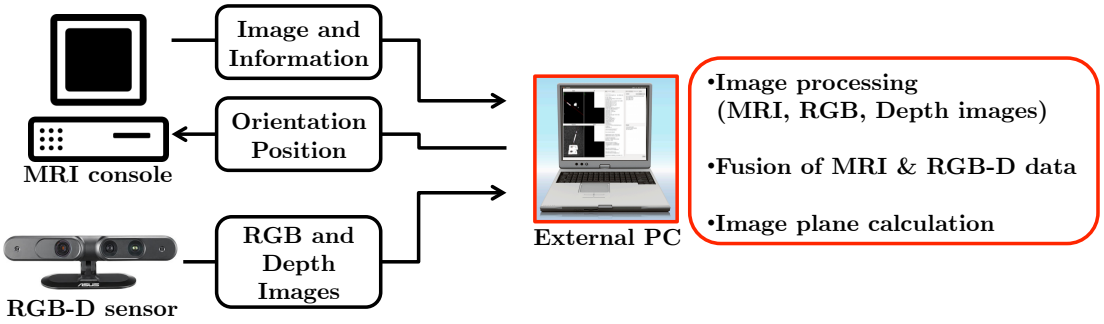


Figure 4.14: System architecture of the implemented hybrid workflow: The external PC receives MR images from the MRI console PC via Ethernet and images from the RGB-D sensor via an USB connection. Image plane control is performed, based on fused detection results from both modalities.

4.7 Results

Results for the registration quality between RGB-D sensor and MRI frame are presented. Furthermore, the maximal tracking speed and the tracking accuracy for the workflow presented in this chapter are experimentally evaluated.
The workflow is based on marker detection in sagittal oblique and in transversal oblique image planes, that are acquired alternately with acquisition frequency of 0.83 Hz. The marker is simultaneously detected on RGB-D images with an acquisition frequency of 30 Hz. The marker pose is estimated by an Information filter, based on measurements from both modalities. The MR image planes are aligned to the marker pose that is predicted by the Information filter.

4.7.1 Timing analysis of real-time image plane alignment

The combined use of MR images and measurements from an RGB-D sensor adds new elements to the timeline presented in section 2.9.1:

- Image acquisition time of the RGB-D sensor (t_A)
- Time for image processing of the RGB-D images (t_I)
- Time for data fusion of the measurements from MR images and from RGB-D images and for marker pose prediction (t_F)

A multi-thread implementation has been chosen, allowing to perform RGB-D image reception and processing in parallel to the MR image reception and processing. The Information filter algorithm is triggered after the image processing of the RGB-D images and fuses the available measurements. Thus, compared to the initial workflow, the addition of the RGB-D sensor does lengthen the overall processing time for calculation of an image plane alignment command only by the data fusion time (t_F).

The analysis of more than 2000 fusion and prediction steps revealed that the average additional processing time due to the use of the Information filter is less than 250 μs and thus negligible. Hence, the image update time for MR images remains the same than for the initial workflow (1200 ms).

A recapitulatory timeline of the processing and communication times is presented in Figure 4.15.

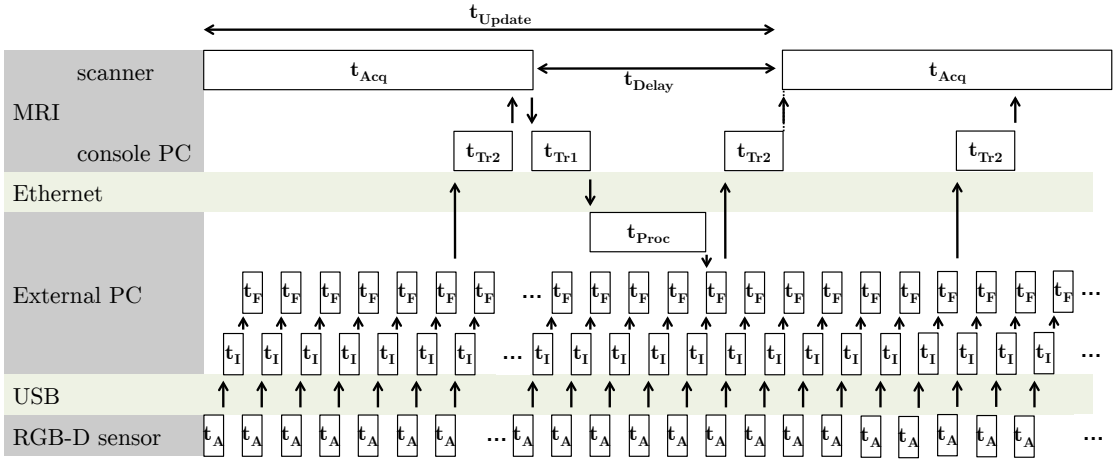


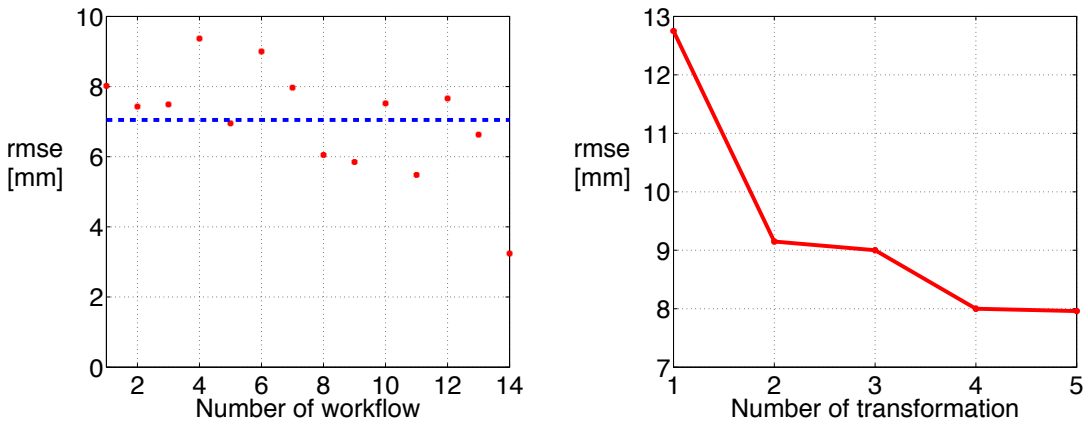
Figure 4.15: Timing schematic for image plane alignment. MR image and RGB-D image acquisitions are performed in parallel. The RGB-D images are acquired and transmitted to the external PC (t_A) and then processed (t_I). The Information filter algorithm is triggered (t_F) after processing of the RGB-D images.

It would be possible to acquire one MR image after another without any delay, as the measurements of the RGB-D sensor allow to control the MR image planes at a higher frequency. Nevertheless, if no delay is introduced, accounting for an MR image based measurement would be impossible, during an occlusion between the marker and the RGB-D sensor.

4.7.2 Online registration evaluation

The online registration presented in section 4.4, with the aim of determining the rigid transformation between the MRI scanner and the RGB-D sensor frames is evaluated. This transformation is mandatory to translate the RGB-D data in the MRI frame of reference where the tracking

is performed. Registration quality is assessed with the rmse between a point set acquired in the MRI frame and the corresponding one in the RGB-D sensor frame back projected to the MRI frame with the rigid transformation to evaluate.
The registration quality of 14 workflows (mean duration 1.5 min) performed by a user displacing the tracking target in a free-hand-motion was evaluated. During each workflow the rigid transformation between MRI and RGB-D sensor frames is refined over time due to updates in the pointlist used for registration. The transformation resulting in the best rmse for each workflow is determined (Figure 4.16a). Hence, a mean registration error of 7.04 mm is obtained for the 14 workflows.
A typical evolution of the root mean square error (rmse) during a workflow is presented in Figure 4.16b. The rigid transformation is updated when an update of the pointlist improves the rmse. Hence, the rmse improves with every new transformation update.
Note that, the computed rmse characterize the registration quality and not the overall tracking performance of the system which is evaluated in section 4.7.3. An initial rigid transformation was typically obtained after approximately 70 s, depending on the performed free-hand-motion of the user.



(a) (b)

Figure 4.16: Evaluation of rmse of registration. (a): The rmse of the best rigid transformation is depicted for every workflow (red). The mean registration error (blue) is 7.04 mm. (b): Typical evolution of the registration error during a workflow with respect to the number of the transformation update between MRI and RGB-D sensor frames.

4.7.3 Experimental evaluation with testbed

Evaluation of the maximal tracking speed

The maximal tracking speed is determined experimentally in the same manner as in section 2.9.3. Note that the maximal tracking speed for the hybrid workflow is only limited by the command acknowledge rate of the MRI interface as the marker can always be tracked in the

RGB-D images. Even if the marker is not depicted in one or several MR images it is either continually tracked with the RGB-D sensor or can be re-detected easily after an occlusion in the RGB-D images. Hence, a maximal tracking speed does not exist in the same manner as for the other workflow versions, where the marker cannot be tracked anymore and a reinitialization step has to be performed. The maximal tracking speed for the hybrid workflow is thus assumed to be the speed at which the MRI console PC cannot align the MR image planes to the marker anymore. The translation curve of the performed oscillatory motion on the testbed is depicted in Figure 4.17. The achieved maximal tracking speed is 45.3 mm/s which is more than twice the value of the initial workflow.

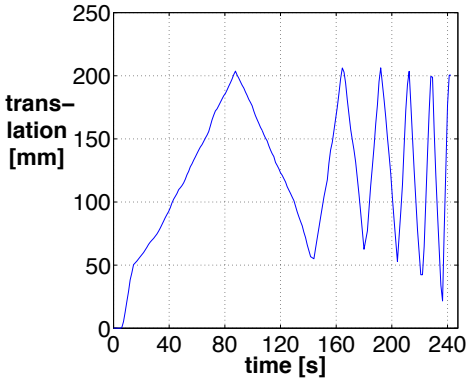


Figure 4.17: Experimental determination of the maximal tracking speed. The oscillatory motion is performed with increasing speed until the marker cannot be tracked in the MRI images anymore. The last edge of the motion before the tracking is interrupted is used for maximal speed computation. Maximal speed in this case is 45.3 mm/s.

Accuracy assessment

In order to evaluate the workflow combining MR image based tracking and tracking based on the RGB-D sensor, the tracking performance is assessed with the same setup as used for the initial and the Kalman extended workflow (section 2.8.2).
Registration between the RGB-D sensor and the MRI frame was achieved through an initial coarse 3D rigid transformation between both frames by comparing RGB-D and MRI position measurements. The transformation was then adjusted interactively until the image planes are well aligned to the marker.
The performed motion and tracking are presented in Figure 4.18 for the different position components of the marker pose. The measured mean speed of the motion is determined on the translation curve of the testbed and is approximately 15.1 mm/s. No step pattern or overshoot is observed due to the image acquisition rate of the RGB-D sensor and the predictions of the Information filter. Also, the curves of the x- and z-coordinate follow the real trajectory closely. The curve of the y-coordinate reveals a systematic error of about 3.5 mm due to an inaccurate manual registration between the frame of the RGB-D sensor and the MRI

frame.

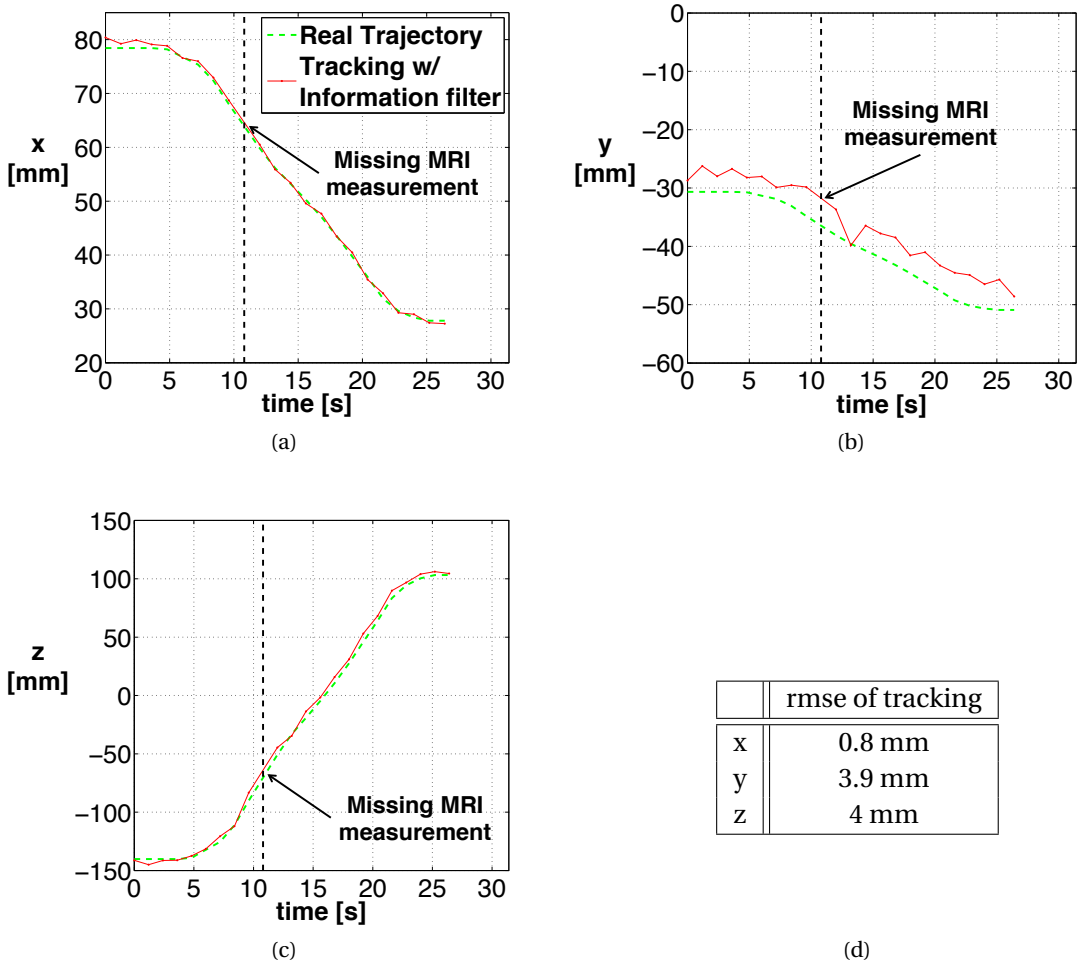


Figure 4.18: Experimental results for tracking of a marker motion of approximately 15.1 mm/s. The marker motion is decomposed to every pose component and depicted with the green curves (a-c). The tracked position values are depicted with the red graphs. One MRI measurement could not be included in the estimations due to an image artifact and the resulting non-detection. A systematic registration error along the y-axis can be observed. The other tracking curves follow the real trajectory closely. The obtained rmse is represented in (d).

The overall rmse of the position tracking is 5.7 mm which can also be computed separately for every component: the rmse is 0.77 mm, 3.91 mm and 4.04 mm for the x-, y- and z-coordinates, respectively. Note that the motion has almost the same speed than the motion for experimental evaluation of the Kalman extended workflow in section 3.4.3. The rmse is higher though. This can also be partly explained with the systematic registration error along the y-axis which can be observed in the rmse for this component. Minor registration errors along the other axes may also affect the result. A failed detection in the MR image after 10.8 s,

due to an artifact separating the marker in half, also influenced the rmse negatively as the MRI part lacks for the estimation of the marker pose. Furthermore, the quantization step of the RGB-D sensor at a distance between 1.3 and 1.5 m (~ 7 mm) can influence the rmse along the z-axis.

Another point is the prediction of the marker pose to which the following image plane is aligned. Currently the prediction of the Information filter is performed for the acquisition of the next RGB-D sensor image (1/30 s in the future). The MR image plane alignment command sent to the MRI console PC is thus also based on a prediction for the time of the next RGB-D sensor image. This results in an MR image plane aligned to a predicted marker pose that has been made for a time before the MR image acquisition. This slight temporal shift of the image plane alignment quantifies in a slight spatial shift of the image plane with respect to the actual marker position. This effect may also contribute to the higher rmse of this workflow version with respect to the Kalman extended workflow. This has been initially implemented in this manner as the time limit after which the MRI console PC does not acknowledge a command anymore is not exactly known (Figure 4.19).

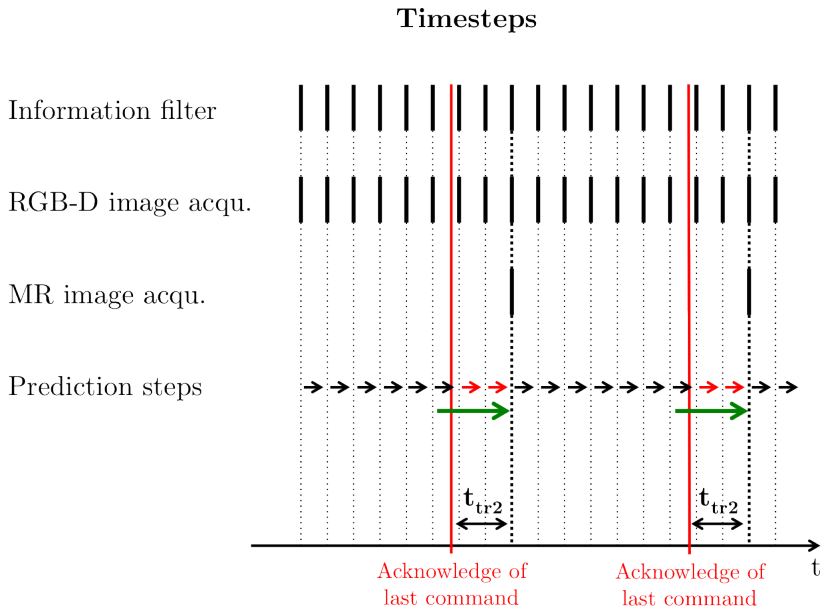


Figure 4.19: Qualitative timeline of the predictions performed by the information filter. Currently the prediction steps are performed according to the frequency of the RGB-D sensor and are sent to the MRI console PC. The time between the last accepted command for image plane alignment and the MR image plane acquisition is denominated t_{tr2}. The commands based on the predictions made after the red line and sent to the MRI console PC are not taken in account anymore (red arrows). The last possible command should thus be based on the prediction made for the next acquisition time of an MR image plane such as the green arrow indicates.

Instead, the image plane alignment command should be based on the prediction of the marker

pose for the time of the next MR image acquisition. For this reason, the transmission and acknowledge time for the image plane alignment command, denominated t_{tr2} in section 2.9.3, must be experimentally determined. A prediction can then be made for the time of acquisition of the next MR image plane and the corresponding command can be sent to the MRI console PC before the acknowledgment limit.

4.8 Conclusion

In this chapter has been presented the use of an RGB-D sensor integrated in an MR image plane alignment workflow. The RGB-D sensor can be used in an MRI room without any functionality problems and without affecting MR image quality. The few image artifacts linked to the use of the RGB-D sensor are caused by the USB connection cable passed through the waveguide of the MRI room but do not interfere with the MR image-based detection of the marker.
A marker, trackable in both MR images and RGB-D images, is used in order to integrate the RGB-D sensor in the MR image-based workflow. The extension of the workflow with the RGB-D sensor allows to decouple the tracked marker motion from the slow image acquisition rate of the clinical MRI sequence. Much higher motion speeds than with the workflow versions presented in chapters 2 and 3 can be tracked. Furthermore, the tracking does not reveal any step pattern or overshoot as could be observed for the previous versions. No reinitialization step using a dedicated volume image is necessary as the marker can nearly always be re-detected with the RGB-D sensor.
The registration of the RGB-D sensor with the MRI is performed online in order to avoid an explicit time-consuming registration step. Moreover, the chosen approach allows to refine the determined transformation over time.
An Information filter is used in order to fuse synchronous data from both modalities. This highly flexible algorithm makes the proposed workflow very robust against unavailable measurements.
The presented tracking workflow is thus highly reactive and can be used for research of the needle insertion point on the patient's skin combining the advantages of an active and passive tracking approach for MR image plane alignment. However it is highly sensitive to the quality of the registration between the MRI and RGB-D sensor frames.

5 Conclusion

5.1 Conclusions

In this work, a method for automatic image plane alignment in real-time is proposed for interventional MRI. Advantages of the proposed method are that it requires only little additional instrumentation, causes little material cost and minimizes the dedicated image acquisition tracking time.

The developed tracking method is based on a device combining a passive marker and resonant micro-coils for automatic image plane alignment in real-time. An image-based tracking algorithm and the software for its implementation with an MRI scanner were developed. Shortly, the method is based on the successive acquisition of two orthogonal image planes aligned to the marker axis, based on the marker detection in the previously acquired image. A simulation program was implemented, allowing to evaluate the tracking performance in controlled and reproducible conditions. An MRI compatible testbed was used for experimental evaluation of the workflow. Currently, the proposed method allows to use the marker as an interventional imaging probe for dynamic alignment of two orthogonal MR image planes. It is thus possible to search interactively for the needle insertion point on the patient.

Due to the use of two orthogonal 2D MR image planes during tracking, only a partial knowledge of the marker pose is measured in each individual successive image plane. Given the current slow MR image acquisition rate, the impossibility to measure the displacement normal to the image plane strongly limits the performance of the object tracking. In order to improve the image plane alignment, the presented approach was thus extended with a Kalman filter in order to benefit from its optimal filtering, estimation and prediction characteristics. For this purpose, a formalization of the existing workflow was realized in order to include the Kalman filter in the control loop. As expected, use of the Kalman filter resulted in an improvement of the tracking performance during simulations and experimental validation. The Kalman filter performed well during weakly accelerated marker motions. Nevertheless, due to the slow MR image acquisition rate and the chosen constant velocity model for the marker motion, the Kalman filter reacts with overshoot when a sudden change in motion occurs. This can result

in a marker loss during tracking.

The shortcoming of this method, common to all passive approaches, is its slow update rate and its dependency on the MR imaging sequence. In order to decouple the tracked motion from the MR image acquisition, an RGB-D sensor was integrated into the proposed tracking workflow. The sensor was chosen as it is a cost effective and easily implementable solution for direct measurements of the 3D pose of the marker. Furthermore, the RGB-D sensor can be used inside the MRI room. An online registration approach was implemented in order to facilitate a potential integration of the proposed system into the clinical workflow. As a consequence, no explicit registration step has to be performed before every intervention and after every position change of the RGB-D sensor. The integration of the RGB-D sensor to the existing workflow has been realized with an Information filter for data fusion of the measurements based on MR and RGB-D sensor images. In this manner, both modalities contribute to the estimation of the marker pose and the corresponding image plane alignment. Furthermore, due to the high acquisition rate of the RGB-D sensor, tracking robustness is greatly improved.

It has to be acknowledged that the purely passive image plane alignment workflow lacks of reactivity for fast and wide marker motions due to the low acquisition rate of the MR images. When the marker is moved at moderate velocities, this approach could be a practicable option for tracking. The combination of a passive and active tracking approach for image plane alignment in interventional MRI overcomes this restrained reactivity and has shown encouraging results. The flexible fusion of the measurements of both modalities allows thus to increase the robustness of the image plane alignment by combining the individual strength of both methods.

5.2 Perspectives

MR image based detection of the marker is still subject to ongoing work. In order to improve the robustness of the object segmentation and image interpretation, the marker and the patient's body can be segmented using an active contour method such as a snake algorithm. Their parametrization with superellipses is expected to result in a higher robustness and accuracy of the image segmentation. Tests for the real-time compatibility of such an approach have to be carried out. Besides, the active contour segmentation is expected to strengthen the marker segmentation in the initialization volume images, based on the passive marker alone without the micro-coils.

The timing analyses of the different workflow versions have shown that the communication times between the PCs cannot be easily estimated and are inconstant, due to the use of no real-time Ethernet protocol and computer systems. The image plane alignment performance could be greatly improved if the marker pose could be predicted for the middle of the next image acquisition period. Hence, the implementation of the developed workflows on the MRI console PC would allow to avoid transmission times. Eventually, this would lead to an acceleration of

the image update rate and to a facilitation of the implementation of the workflows. Therefore, simulations at an hypothetical 2 Hz imaging frequency have been performed and shown the interest of working with the highest possible MR imaging frequency. In this context, an optimization of the MR imaging sequence could be considered. A good compromise between image quality and acquisition time should be found in close collaboration with the physicians.

The proposed system can be used in order to search interactively for the needle insertion point on the patient's skin. A marking device must thus be attached to the tracking device. For this reason, further developments for the optimization of the marker are necessary.

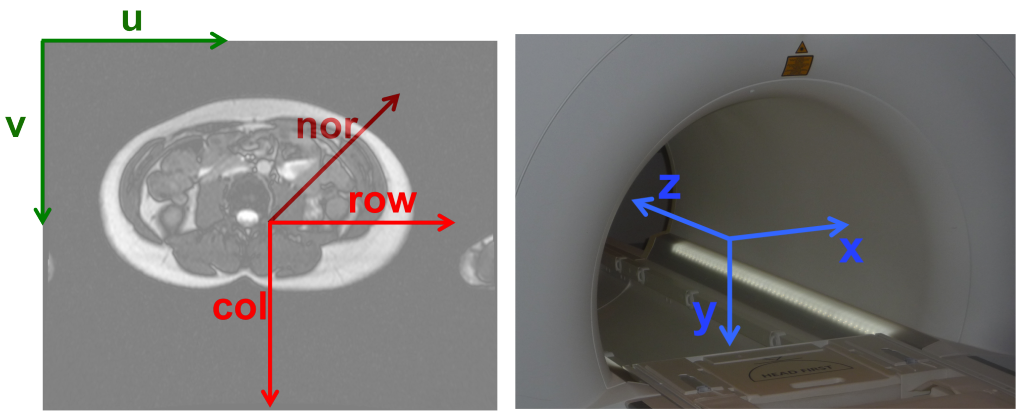
In this work, the passive marker design was chosen so that it corresponds to a simple, already sterile syringe filled with a contrast agent dilution. Further work on an optimal compact marker design is required, especially to propose a marker design compatible with its use for tracking of an interventional needle. Ultimately such needle tracking marker could be combined with a needle artifact detection algorithm in order to precisely monitor the trajectory of the MR-compatible flexible needle.

A Conventions in MRI

A.1 Coordinate systems

The basic coordinate systems in an MRI context are given hereunder. As long as no precision is made, their coordinates are given in mm.

- a coordinate system linked to the MRI scanner ($\mathcal{F}_M : \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$) with its origin at the isocenter of the MRI scanner
- a coordinate system linked to the image plane ($\mathcal{F}_I : \boldsymbol{row}, \boldsymbol{col}, \boldsymbol{nor}$) with its origin at the center of the image plane
- a second coordinate system linked to the image plane ($\mathcal{F}_C : \boldsymbol{u}, \boldsymbol{v}$), which is two-dimensional and has its origin at the upper left corner of the image plane. Its coordinates are represented in Pixels.



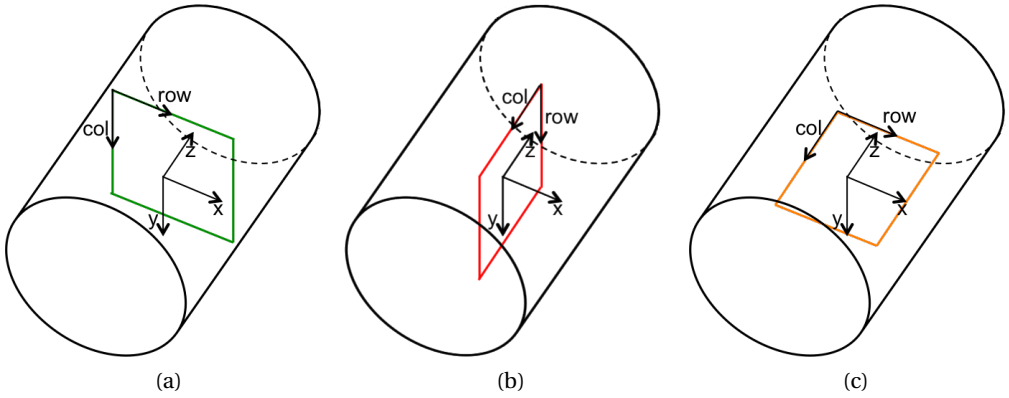
(a) (b)

Figure A.1: Image and Machine coordinate systems used during this work. (a): Coordinate systems $\mathcal{F}_C$ and $\mathcal{F}_I$, both linked to the image plane are represented. (b): Coordinate system linked to the MRI scanner $\mathcal{F}_M$ is represented.

The image position of an MR image plane defines the x, y, and z coordinates of the center of the central image voxel. The orientation of an MR image plane is defined through direction cosines given in the row and column vectors of the $\mathcal{F}_M$ frame.

A.2 Plane conventions

In MRI, in contrast to CT scan, not only volume projection can be acquired but thin image planes that are fully orientable in space. The basic orientations of a scan plane are defined as transversal, sagittal and coronal (see Figure A.2).



(a) (b) (c)

Figure A.2: Representation of the different MRI image planes inside the MRI scanner tunnel. Orientation of ***row*** and ***col*** vectors (shifted to image corners for clarity) with respect to the coordinate system linked to the MRI scanner are given for transversal (a), sagittal (b) and coronal (c) image planes.

Every image plane is described by 3 orthogonal vectors, namely ***row***, ***col*** and ***nor***. These vectors describe respectively the row, column and normal direction of an image plane. The ***row*** vector is oriented along the horizontal image axis, the ***col*** vector is oriented along the vertical image axis and the ***nor*** vector is described by their cross product.
An image plane that is rotated around one of the principal axes of the MRI scanner is called a simple oblique image plane (see Figure A.3). In this work are mainly used simple oblique transversal and sagittal image planes. Their orientations are described by the angles $alpha$ and $beta$.
Note, that there are also double oblique image planes which are rotated around two of the principal axes of the MRI scanner. This type of image plane is not used during this work.

Transversal image plane

A pure transversal image plane is described by the following vectors, represented in the MRI scanner frame:

$$\boldsymbol{row} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{col} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \boldsymbol{nor} = \boldsymbol{row} \times \boldsymbol{col} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{A.1}$$

A simple oblique transversal image plane is rotated around the x-axis by an angle β and has thus the following image plane vectors:

$$\boldsymbol{row} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \boldsymbol{col} = \begin{bmatrix} 0 \\ cos(\beta) \\ sin(\beta) \end{bmatrix} \qquad \boldsymbol{nor} = \boldsymbol{row} \times \boldsymbol{col} = \begin{bmatrix} 0 \\ -sin(\beta) \\ cos(\beta) \end{bmatrix} \tag{A.2}$$

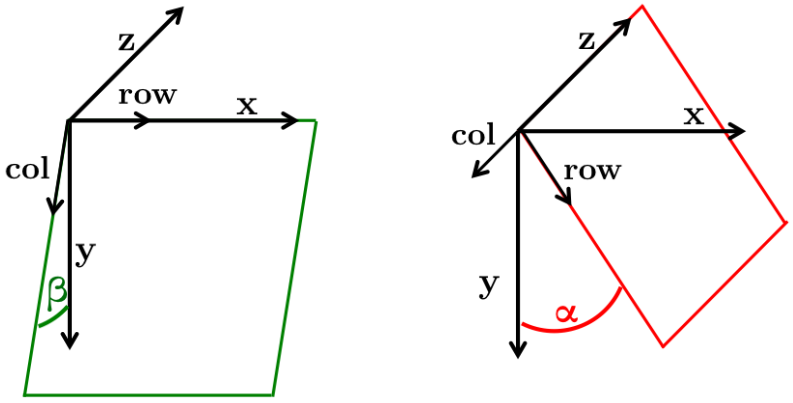
Sagittal image plane

A pure sagittal image plane is described by the following vectors, represented in the MRI scanner frame:

$$\boldsymbol{row} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad \boldsymbol{col} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \qquad \boldsymbol{nor} = \boldsymbol{row} \times \boldsymbol{col} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \tag{A.3}$$

A simple oblique sagittal image plane is rotated around the z-axis by an angle α and has thus the following image plane vectors:

$$\boldsymbol{row} = \begin{bmatrix} -sin(\alpha) \\ cos(\alpha) \\ 0 \end{bmatrix} \qquad \boldsymbol{col} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \qquad \boldsymbol{nor} = \boldsymbol{row} \times \boldsymbol{col} = \begin{bmatrix} -cos(\alpha) \\ -sin(\alpha) \\ 0 \end{bmatrix} \tag{A.4}$$



(a) (b)

Figure A.3: Transversal (a) and sagittal (b) oblique image planes depicted with their ***col*** and ***row*** vectors in the MRI frame.

B Geometrical camera model

A geometrical camera model as presented in [HM11] is described in this Annex and taken as reference for this work.
Two transformations are used in order to describe the geometrical model of the camera.
First, a perspective projection of a 3D point onto the image plane of the camera and second, a transformation from the camera frame to the image frame.
The frame $\mathcal{F}_c = \{x, y, z\}$ is the camera frame with its origin O at the point of aperture of the camera. The image plane of the camera is placed behind its point of aperture at a distance $-f$ along the z axis, corresponding to the focal length of the camera (see Fig.B.1a).

B.1 Perspective projection of a 3D point onto the image plane

A point $A = (x, y, z)^T$ is projected through the aperture of the camera onto its image plane. The image plane is parallel to the plane passing through the x and y axes of the camera frame.
The point $A' = (x', y', z')^T$ is the projection represented in the camera frame. The coordinates of the projected point A' are given by the following equations:

$$x' = -\frac{x}{z} f$$
$$y' = -\frac{y}{z} f$$
$$z' = -f$$

Note, that the negative components of A' correspond to an inversion of the observed scene on the image plane. This transformation can be represented in a matrix $\boldsymbol{P}$, when homogeneous coordinates are used for the points ($\tilde{A}$, $\tilde{A}'$). The transformation can thus be written as

$$\begin{array}{ccc} \tilde{A}' & = & \boldsymbol{P} \quad\quad\quad \tilde{A} \\ \begin{bmatrix} s\,x' \\ s\,y' \\ s\,z' \\ s \end{bmatrix} & = & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{f} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \end{array}$$

where the Cartesian coordinates of the point A' can be obtained through division of the homogeneous coordinates by their scale factor s:

$$A' = \begin{bmatrix} \frac{s\,x'}{s} \\ \frac{s\,y'}{s} \\ \frac{s\,z'}{s} \end{bmatrix}$$

B.2 Transformation from camera to image frame

The image frame $\mathcal{F}_i = \{u, v, w\}$ has its origin in the upper left corner of the image plane. Its coordinates are given in pixels. The image plane is intersected by the z axis of the camera frame in the principal point $P_p = (u_0, v_0, w_0)$. In order to convert the coordinates from the camera frame, given in mm, to coordinates in the image frame, given in pixels, the parameters k_u and k_v are introduced. They correspond to the pixel size of the image sensor ([pixels/mm]) for the u and v axes, respectively.
Thus, the transformation from the camera frame to the image frame can be written as:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} k_u & 0 & 0 \\ 0 & k_v & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \\ w_0 \end{bmatrix} \tag{B.1}$$

This transformation represents a rescaling, rotation and translation. As the w component will always be 0 in the image frame, the third line of this transformation can be discarded. The combined transformation between the 3D camera frame and the 2D image frame corresponds thus to:

$$\begin{bmatrix} s\,u \\ s\,v \\ s \end{bmatrix} = \boldsymbol{K} \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \tag{B.2}$$

with

$$\boldsymbol{K} = \begin{bmatrix} k_u & 0 & 0 & u_0 \\ 0 & -k_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{B.3}$$

B.3 The intrinsic parameters

When the two transformations K and P are multiplied, we obtain the following transformation matrix:

$$\boldsymbol{K}\,\boldsymbol{P} = \begin{bmatrix} k_u & 0 & -\frac{u_0}{f} & 0 \\ 0 & -k_v & -\frac{v_0}{f} & 0 \\ 0 & 0 & -\frac{1}{f} & 0 \end{bmatrix} \tag{B.4}$$

By multiplying the obtained transformation with -f (which doesn't change the transformation due to the use of homogeneous coordinates), the camera matrix $\boldsymbol{I}_c$ is obtained:

$$\boldsymbol{I}_c = \begin{bmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{B.5}$$

where $\alpha_u = -k_u\,f$ and $\alpha_v = k_v\,f$. $\boldsymbol{I}_c$ gives the relationship between the 3D point A and its projection onto the image plane in image coordinates ${}^{i}A' = (u, v)$:

$$\begin{bmatrix} s\,u \\ s\,v \\ s \end{bmatrix} = \boldsymbol{I}_c \begin{bmatrix} x \\ y \\ z \\ s \end{bmatrix} \tag{B.6}$$

If camera coordinates such as

$$x_c = \frac{x}{z} \qquad y_c = \frac{y}{z} \qquad z_c = 1 \tag{B.7}$$

are used, the relationship between these camera coordinates and their image coordinates is given by the equations

$$u = \alpha_u x_c + u_0 \tag{B.8}$$

$$v = \alpha_v y_c + v_0 \tag{B.9}$$

This can be written in matrix form as

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{C} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix} \tag{B.10}$$

with the calibration matrix $\boldsymbol{C}$, containing the intrinsic parameters of the camera:

$$\boldsymbol{C} = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{B.11}$$

$\boldsymbol{I}_c$ can thus be decomposed in a projective transformation and the calibration matrix $\boldsymbol{C}$

$$\boldsymbol{I}_c = \underbrace{\begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{C}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{B.12}$$

Note that in this model, perfect orthogonality between the u and v axes is assumed and thus no skew factor is introduced.
If radial distortion of the camera sensor is accounted for, equations B.8 and B.9 have to be replaced by:

$$u = \alpha_u x_c + u_0 + K_1 r^2 (u - u_0) \tag{B.13}$$

$$v = \alpha_v y_c + v_0 + K_1 r^2 (v - v_0) \tag{B.14}$$

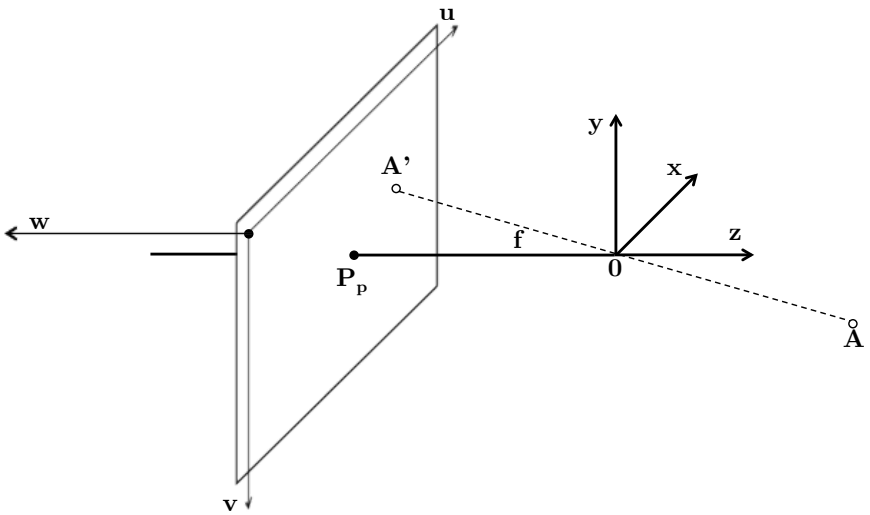
where $r^2 = (u-u_0)^2 + (v-v_0)^2$ is the squared distance of the point ${}^i A' = (u, v)$ from the principal point P_p in the image frame.

B.4 The extrinsic parameters

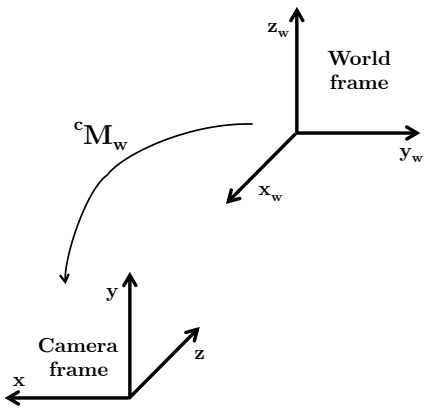
In order to represent an observed point from the surrounding world in the camera frame, the transformation between the world and the camera frame must be known. This transformation corresponds to a rotation ${}^c\boldsymbol{R}_w$ and a translation ${}^c\boldsymbol{T}_w$, which can be represented in a homogeneous matrix (see Fig. B.1b).

$${}^c\boldsymbol{M}_w = \begin{bmatrix} {}^c\boldsymbol{R}_w & {}^c\boldsymbol{T}_w \\ 0 & 1 \end{bmatrix} \tag{B.15}$$

The represented 3 rotations and 3 translations inside the homogeneous matrix are the extrinsic parameters of the camera.



(a)



(b)

Figure B.1: (a): Geometrical model of the camera with the image frame $\{u, v, w\}$ and the camera frame $\{x, y, z\}$. The image plane is situated at a distance corresponding to the focal length f along the z axis, which intersects the image plane in the principal point P_p. The point A is projected through the aperture of the camera (O) onto the image plane where the projection lies on the image plane in the point A'.
(b): The spatial relationship between the camera frame and the world frame is described by the homogeneous matrix ${}^{c}\boldsymbol{M}_{w}$.

C Custom software for image plane alignment in interventional MRI

During this work, a custom C++ program has been developed in order to perform the developed workflow on an external PC connected to the MRI console PC. The final program performs automatic marker detection in MR images and in images from an RGB-D sensor. The detected marker poses are then fused by an Information filter and an image plane alignment command is sent to the MRI.
Hence, the realized functionalities of this program are:

- Reception of MR images from the MRI console PC
- Marker detection in the MR images
- Reception of images from the RGB-D sensor
- Marker detection in the RGB-D images
- Calculation of MR image plane alignment commands according to the marker detection
- Sending of the image plane alignment commands to the MRI console PC

These functionalities are implemented in an automatic workflow, that requires no user interaction.

The program has been developed in a Linux environment using the following open-source cross-platform libraries:

- Qt: framework for creation of a graphical user interface
- Visualization Toolkit (VTK): framework for 3D computer graphics, image processing and visualization
- Insight Segmentation and Registration Toolkit (ITK): framework for image segmentation and registration
- Open Source Computer Vision Library (OpenCV): library of function for real-time computer vision
- Boost: set of C++ libraries for numerous domains such as multi-threading, linear algebra, asynchronous input and output operations and more
- Point Cloud Library (PCL): framework for acquisition and 3D geometry processing of

n-dimensional point clouds
- OpenNI framework: framework providing support for hand gesture recognition, body tracking and voice command recognition in combination with RGB-D sensors

C.1 General description

The graphical user interface (GUI) of the developed program (Figure C.1) has been realized using Qt. The program is executed in a Linux environment on a laptop computer that is connected to the MRI console PC via Ethernet. The MR images are transmitted to the external PC with a proprietary Siemens data header. The MR image plane alignment commands are sent from the external PC to the MRI console PC in an ASCII string, in compliance with a Siemens protocol.

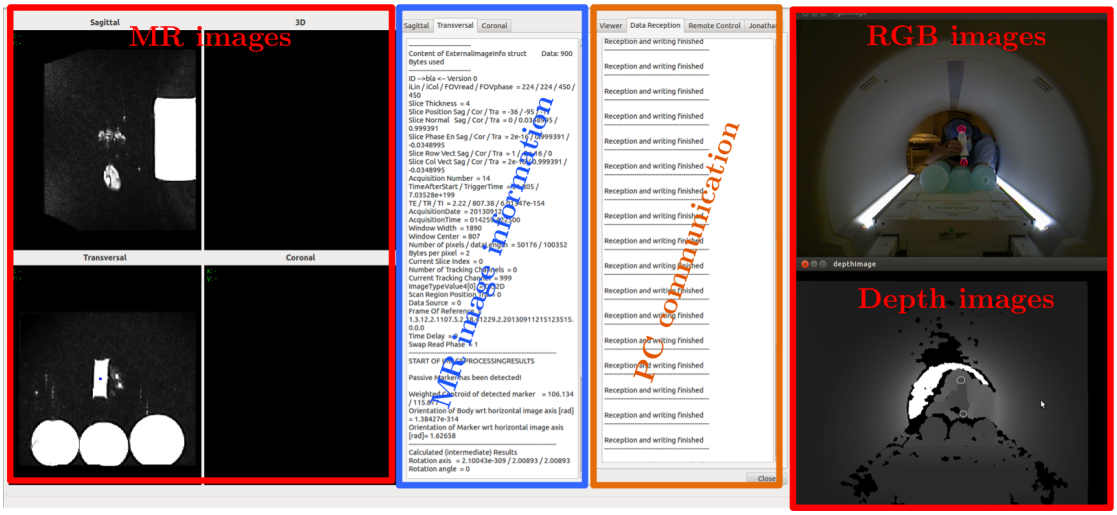


Figure C.1: GUI screenshot of the developed program with the received MR and RGB-D images at the left and right, respectively. Image information and communication protocols between external and MRI console PC are depicted in the middle.

The program was developed and further developed according to the different versions of the workflow presented throughout this work.

C.2 Version 1: MR image plane alignment based on MR images

The first version allows to automatically detect an MRI marker in MR images, calculate its 3D pose and align an image plane according to the detected marker pose. The received MR images and the according information (position, orientation, etc.) are represented in different windows on the GUI. Besides the automatic functionality for image plane alignment, the developed GUI allows the user to interact directly with the MRI console PC. The user can thus sent MR image plane alignment commands entered in an input window on the GUI.

The functionalities of the program are distributed in different threads (Figure C.2).

Thread 1: Creation of GUI and management of event-loop

The first thread creates the graphical interface and executes its inherent event-loop. The event-loop handles user interactions with the program (button clicks, etc.) and communication between other threads. In the context of the presented program it manages the displaying of the MR images, MR image data and sent commands on the GUI.

Thread 2: MR image reception

This thread opens a socket connection as server using the boost library and waits for a connection of the MRI console PC. When an MR image is acquired and reconstructed it is represented on the GUI of the MRI console PC. Synchronously, the MRI console PC connects as client on the socket server and transmits the MR image and its (proprietary) data header as byte stream to the external PC. The byte stream is reassembled on the external PC in order to obtain the MR image and the data header.
After reconstruction of the image and data they are sent to Thread 1 in order to be displayed on the GUI. Furthermore, the image is processed by a marker detection algorithm that is based on functions from the ITK library. If the marker is detected, its 3D pose is reconstructed based on the MR image orientation and position. For vector and matrix calculations, the OpenCV library is used. Finally, the ASCII command string based on the detected marker pose is created and sent to thread 3. The thread returns then to waiting for a connection of the MRI console PC.

Thread 3: Sending of commands

Thread 3 manages sending of image plane alignment commands to the MRI console PC. When it receives an ASCII command string from another thread, a socket connection is established between both PCs and the image plane alignment command is transmitted. Depending on the format of the command string, the MRI console PC answers with an acknowledgement message or not. The acknowledgment message indicates if the command has been executed or refused.

The previously presented threads allow to implement the functionalities of the first version of the workflow, that performs MR image plane alignment, based on MR images, only. Communication between the threads is performed using the signal/slots mechanism proposed by Qt. This mechanism allows to define signals and slots that are linked to objects. During initialization of the program, the connections between signals and slots of different objects/threads are established. During program execution, a signal can then be emitted asynchronously, meaning that the emitting thread continues its work right after emitting its signal without caring by whom it is received.

A slot is a normal member function of an object. It is called when a signal that is connected to it, is emitted. This mechanism is an easy way to transmit data from one thread to another. Note though, when data of a current thread is accessed from another thread, the concerned data structures have to be protected from concurrent access using mutexes. A mutex is a variable that protects two or more threads to access a critical section (e.g. shared data) at the same time.

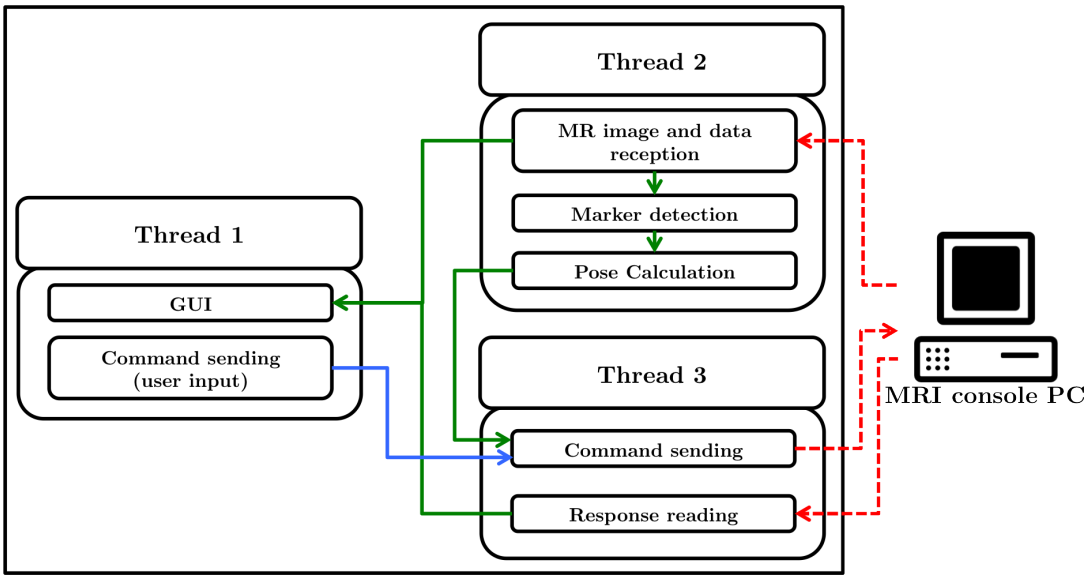


Figure C.2: Global architecture of the program developed for MR image plane alignment based on marker detection in MR images.

C.3 Version 2: MR image plane alignment based on MR images and Kalman filtering

The second version of the workflow performs image plane alignment according to a prediction of the marker pose from a Kalman filter.
For implementation of this approach the Kalman functionalities are added to thread 2 of the first program version. Thus after marker detection, the marker pose is sent as measurement to a Kalman filter algorithm (OpenCV). The Kalman filter algorithm generates then an estimate and a prediction of the marker pose. As in the first program version, the command string is then created and sent to thread 3.

C.4 Version 3: MR image plane alignment based on MR images and RGB-D images

The third version of the program performs MR image plane alignment based on marker detection in the MR images and in images from an RGB-D sensor. The marker pose is determined based on both modalities and the obtained measurement is then fused by an Information filter. The Information filter then generates a prediction of the marker to which the MR image plane is aligned.
Hence, version 1 of the program can be used as baseline and extended through the missing functionalities (Figure C.3). For this purpose, several new threads are created and presented in the following. Moreover, threads 1 and 2 of program version 1 have to be adapted.

Adaptation of thread 1

In order to display the RGB and depth images, two windows are added to the GUI.

Adaptation of thread 2

In order to allow for data fusion, an Information filter is implemented in thread 6. In order to work correctly, the Information filter must receive the measured marker pose from both modalities. For this purpose, the detected marker pose from the MR images has to be sent to the Information filter in thread 6.
Thread 2 keeps thus its essential functions with two changes. First, the detected marker pose has to be sent to the Information filter before thread 2 starts waiting for a new MR image. Second, the image plane alignment command is no longer sent from thread 2.

Thread 4: Reception of images from RGB-D sensor

The OpenNI framework is used in combination with the PCL for reception of the RGB and depth images from the RGB-D sensor. The PCL proposes a grabber interface providing access to different RGB-D sensors and their drivers. Shortly, callback functions of the PCL are connected to the OpenNI framework that is the interface to the RGB-D sensor. In this manner it is possible to receive different data types from the RGB-D sensor such as RGB images, depth images or point clouds. In the case of this program, a callback function providing synchronously acquired RGB and depth images is chosen. Furthermore, an option is set so that the depth image is directly registered to the RGB image. The received images arrive at 30 frames per second with a resolution of 640×480 pixels.

Thread 4 receives thus continuously RGB and depth images and saves to shared variables. When an image couple is received Thread 4 notifies thread 5, the image processing thread, that there is data to process. When notified thread 5 recovers the data from the shared variables and starts to work. The shared data is protected through a mutex. Image processing cannot

directly be executed in thread 4 as the image acquisition frequency is too high and there is not enough time available between the image reception callbacks.

Thread 5: RGB-D image processing

As previously presented, thread 5 is notified by the RGB-D image reception thread that there is image data to process. The image data is then loaded from the shared variables to thread 5. Detection of the marker (pink balls) is then performed in the RGB and depth images using the OpenCV library. If the marker could be detected its 3D pose is reconstructed using the depth images. The detected marker pose is then sent to the Information filter, implemented in thread 6.

Thread 6: Data fusion

The data fusion thread receives the detected marker pose based on MR images (thread 2) and based on RGB-D images (thread 5). The Information filter algorithm is then executed in order to generate an estimation and a prediction of the marker pose. Based on the predicted marker pose, an image plane alignment command is created and sent to thread 3 in order to sent it to the MRI console PC.
The acquisition frequency of the RGB-D sensor (30 Hz) is much higher than the one of the MRI scanner (~0.83 Hz). For this reason thread 6 is executed on every update of the marker pose detected in the RGB-D images. If an MR image is received at the same time, the Information filter fuses their data. Otherwise, the estimation of the marker pose is based on a measurement from an RGB-D image, only.

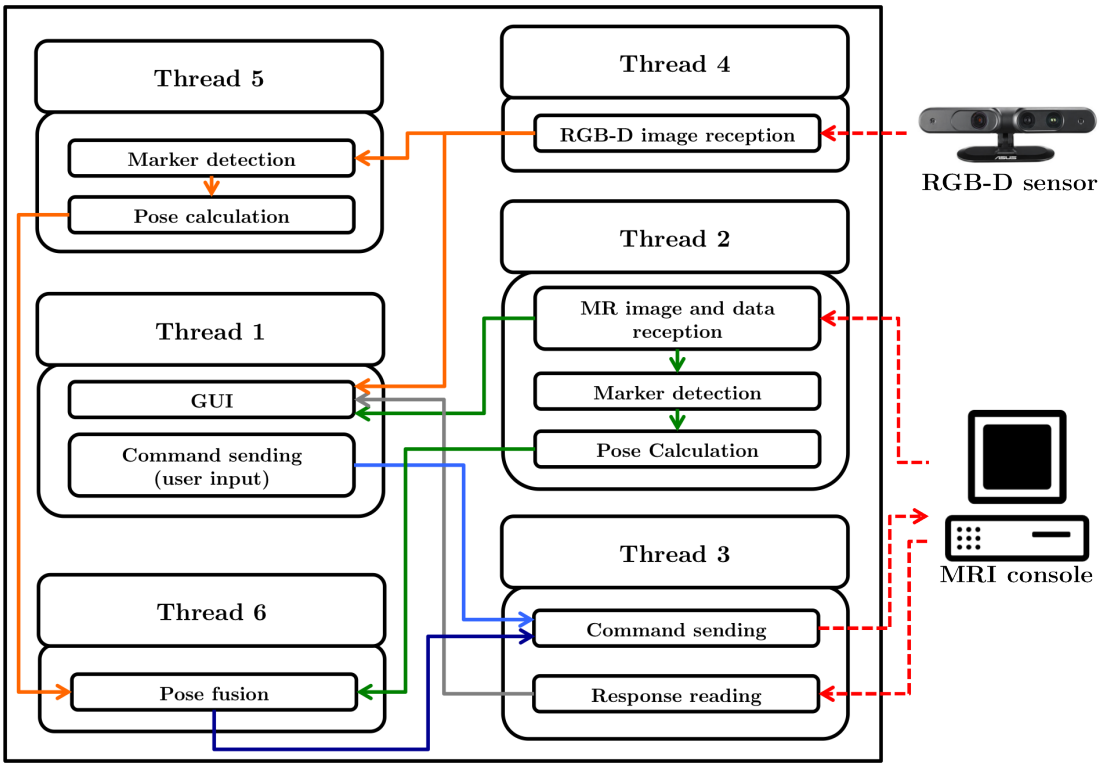
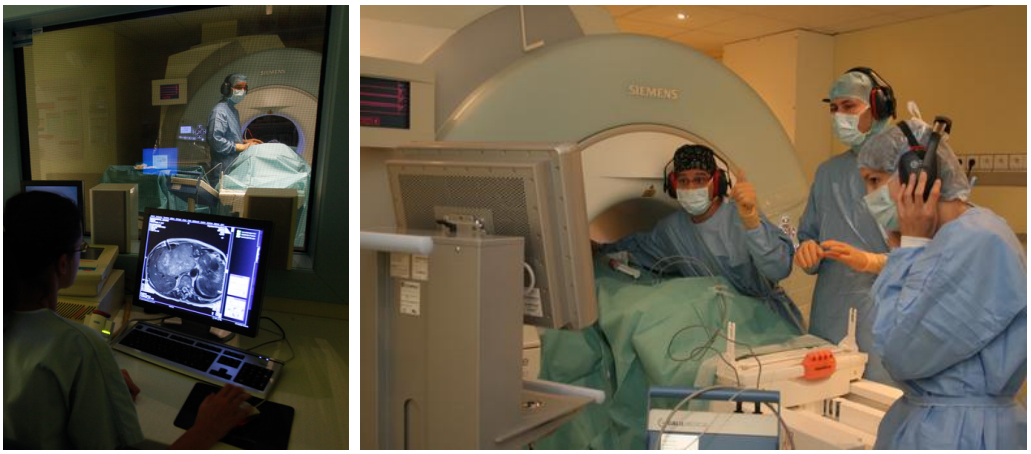


Figure C.3: Global architecture of the program developed for MR image plane alignment based on marker detection in MR and RGB-D images.

D Résumé en français

D.1 Introduction

En radiologie interventionnelle, des procédures chirurgicales minimalement invasives sont réalisées sous guidage d'un imageur médical. Ces gestes sont typiquement réalisés par accès percutané pour des fins diagnostiques ou thérapeutiques incluant des biopsies, infiltrations et ablations. Les modalités d'imagerie utilisées sont l'ultrason (US), la tomodensitométrie (TDM) et l'imagerie par résonance magnétique (IRM). L'IRM a de nombreux atouts face aux autres modalités, notamment le contraste excellent des tissus mous, la haute résolution spatiale et temporelle des images, l'absence de rayonnement ionisant et le positionnement et l'orientation libre des plans image.
Néanmoins, les interventions sous guidage IRM restent des interventions complexes et de longue durée: pendant les interventions 1 à 3 plans d'image successifs sont positionnés manuellement et sont acquis en permanence pour surveiller la procédure en temps-réel. Le radiologue est typiquement intéressé par des plans images orthogonaux alignés à l'axe de l'outil d'intervention, à la cible ou aux structures anatomiques d'intérêt autour. Comme le radiologue est occupé avec la procédure chirurgicale, il n'a pas de moyens d'adapter les plans image selon ses besoins. C'est donc un(e) technicien(ne) dans la salle commande de l'IRM qui aligne manuellement les plans image en temps-réel. Les mouvements de la main du radiologue et de l'aiguille ne sont pas visibles pour le/la technicien(ne) dû au scanner IRM et aux autres machines pour la surveillance des fonctions vitales du patient. L'alignement des images se fait donc par l'intermédiaire d'une communication entre la salle de l'IRM et la salle de commande. Dû au bruit du scanner pendant les acquisitions, le radiologue ne peut pas parler avec le/la technicien/technicienne ce qui fait qu'ils communiquent par langue des signes pour aligner les plans images.
Par conséquence l'alignement des plans images dépends fortement de l'expérience de la/du technicienne/technicien dans la salle de commande de l'IRM, de la communication entre les personnes impliquées et de leur performance en tant qu'équipe.
De ce fait les systèmes de suivi et d'alignement automatique sont d'un grand intérêt pour faciliter interventions chirurgicales mini-invasives sous guidages IRM.



(a) (b)

Figure D.1: (a) Environnement de l'IRM interventionnel: le radiologue dans la salle d'IRM et une technicienne dans la salle de commande. (b) Intervention en IRM avec le radiologue penché dans le tunnel pour avancer l'aiguille dans le corps du patient basé sur les images montrées sur l'écran. Le radiologue utilise la langue des signes pour communiquer avec la technicienne dans la salle de commande.

Etat de l'art des systèmes de suivi pour l'alignement automatique des plans images

Les systèmes de suivi peuvent être classés en approches actives et passives. Les systèmes actifs utilisent du matériel électronique additionnel connecté au scanner IRM ou à un dispositif de mesure pour le tracking des outils interventionnels. Les systèmes passifs se basent sur le suivi et la détection des outils directement dans les images IRM.

Les systèmes actifs sont établis sur la mesure de gradients, de pulsations RF ou sur le suivi par caméra pour la mesure de la pose d'un outil interventionnel. Leurs avantages sont la précision et l'indépendance aux images acquises. Les inconvénients sont le besoin de matériel additionnel, souvent couteux et nécessitant une mise en service ainsi que des séquences IRM spécifiques.

Les systèmes passifs sont établis sur la détection des outils interventionnels dans les images et avec parfois une méthode d'amplification du contraste. Des exemples types sont les marqueurs paramagnétiques, les marqueurs remplis d'une substance de contraste et les micro-bobines. Les avantages des approches passives sont leur simplicité et leurs couts bas. Les inconvénients sont la nécessité pour l'outil de se situer dans le plan image et souvent la nécessité d'acquérir des plans image dédiés à la détection.

Objectifs de la thèse

Dans le contexte des problématiques décrites et de l'état de l'art présenté l'objectif de cette thèse est le suivant: le développement d'un processus d'alignement automatique en temps-réel des plans image, fondé sur l'asservissement visuel d'un marqueur dans les images IRM et permettant de contrôler les plans images depuis l'intérieur du scanner IRM. La méthode développée doit en plus minimiser le temps d'imagerie dédié à la détection et être peu onéreuse.

D.2 Processus pour l'alignement des plans de coupe en IRM interventionnelle

Principe

Le premier processus pour l'alignement automatique des plans image est fondé sur une séquence clinique acquérant en alternance deux plans image orthogonaux. Les plans sagittal et transversal sont utilisés pour suivre un objet composé de deux micro-bobines et/ou d'un marqueur passif rempli d'une substance de contraste.

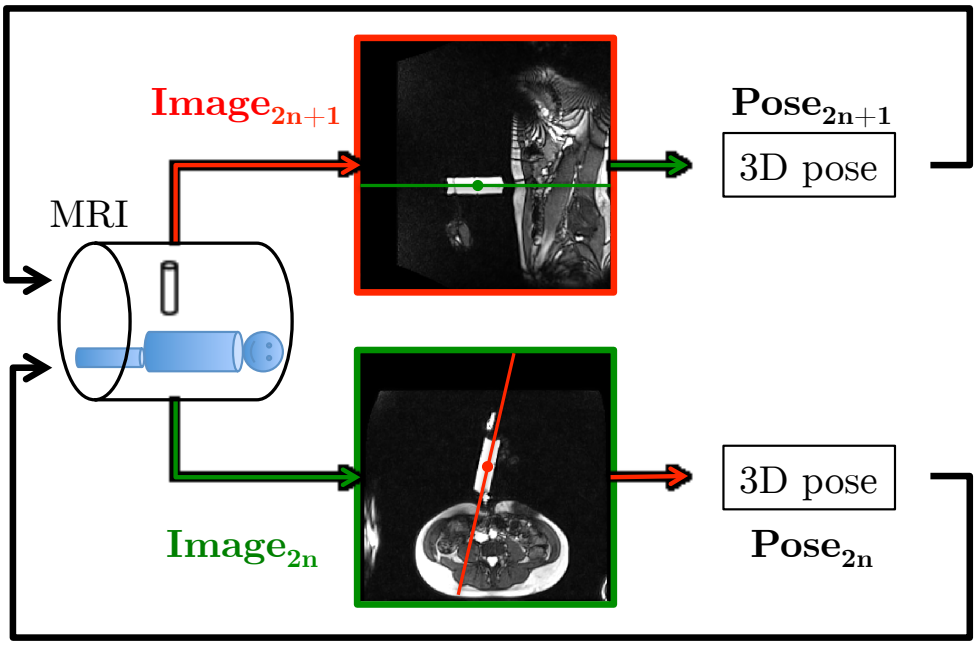


Figure D.2: Principe du processus d'alignement automatique des plans image: 2 plans orthogonaux, transversal (vert) et sagittal (rouge), sont acquis en alternance. La pose 3D du marqueur est détectée dans les images. En fonction de cette pose la position et l'orientation du prochain plan image sont calculées et une commande d'alignement est envoyée à l'IRM.

Seule l'initialisation est réalisée avec une acquisition spécifique: il s'agit soit d'une acquisition d'un volume pour la détection du marqueur passif, soit d'une acquisition à faible angle de bascule d'un volume pour la détection des micro-bobines. L'asservissement des plans images cliniques est réalisé directement par la détection du marqueur passif dans les images.

Détection du marqueur

La détection dans les images d'initialisation peut être réalisée avec les micro-bobines dans les acquisitions de volume à faible angle de bascule ou avec le marqueur passif dans les acquisitions de volume.
La détection des micro-bobines est réalisée avec la méthode POCC ("Phase only cross-correlation") qui permet de détecter les micro-bobines dans une image de corrélation croisée entre l'image IRM et une image modèle des micro-bobines.
La détection du marqueur passif dans l'acquisition de volume débute avec la soustraction du corps du patient après un seuillage. Puis un deuxième seuil déterminé à l'aide de la méthode d'Otsu est utilisé pour segmenter l'image. La détection du marqueur est décidée avec un critère de taille.

Pour la détection du marqueur dans les images temps-réel cliniques, un algorithme de segmentation et de classification a été développé. Un double seuillage établi sur les 0.65 et 0.9 quantiles de l'histogramme permet d'éliminer le corps du patient de l'image et de séparer par la suite le marqueur de la main qui le tient. La classification des objets segmentés est réalisée avec des critères de taille, intensité et excentricité.

Logiciel de suivi

Ce logiciel permet la réception d'images IRM et leurs trames d'informations sur un PC externe connecté au PC de commande de l'IRM. Le logiciel réalise la segmentation du marqueur ainsi que les calculs pour l'alignement du prochain plan image. Une commande correspondante est alors envoyée au PC de commande qui contrôle les plans image de l'IRM.

Evaluation des performances du suivi

Le processus développé est évalué avec des simulations et expériences.

Un simulateur permettant la simulation d'un marqueur a été développé et peut être connecté au logiciel de suivi. En fonction de l'intersection du plan souhaité et de la position actuelle du marqueur, une image IRM artificielle est créée et envoyée au logiciel de suivi.

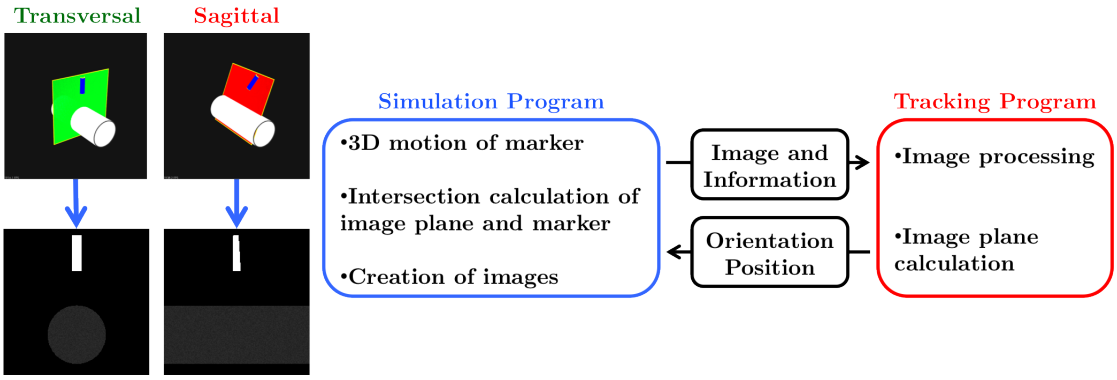
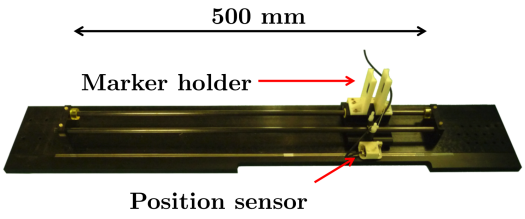
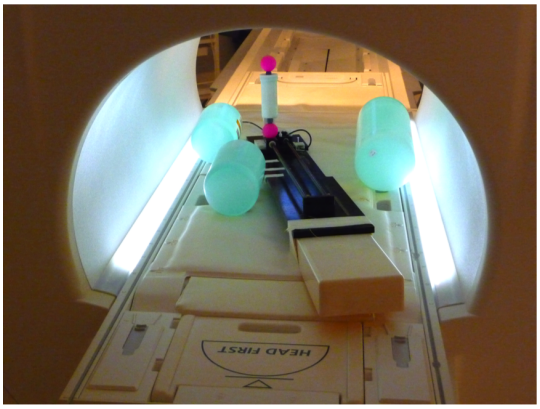


Figure D.3: Principe du simulateur. Il simule le fonctionnement du scanner IRM et envoie en alternance des images sagittales et transversales au logiciel de suivi. Un mouvement 3D du marqueur est défini et des images de synthèse sont créées en fonction de l'intersection entre le plan souhaité et la pose du marqueur.

Pour les expériences, un banc d'essai compatible IRM a été assemblé. Il permet de réaliser manuellement des translations avec le marqueur dont le déplacement est mesuré avec un capteur optique.



(a)



(b)

Figure D.4: (a): Banc d'essai compatible IRM (b): Banc d'essai dans le scanner IRM avec le marqueur monté.

Résultats

Pour évaluer la vitesse maximale du suivi, le cas extrême d'un mouvement du marqueur à 45° entre les plans image a été choisi pour les simulations. La vitesse maximale obtenue est de 19.5 mm/s.

Pour l'évaluation expérimentale, un mouvement oscillatoire de vitesse croissante est généré manuellement. La dernière portion du mouvement qui peut être suivi dans les images IRM

détermine la vitesse maximale. Pour le processus présenté, une vitesse maximale de 8.81 mm/s a été obtenue.

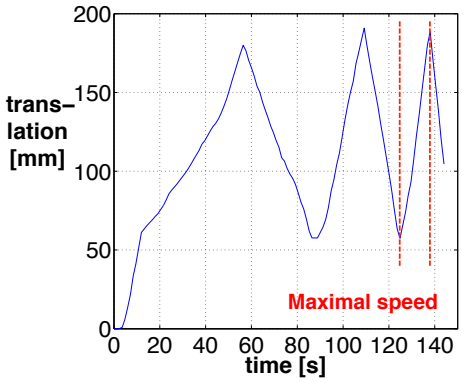


Figure D.5: Détermination expérimentale de la vitesse maximale du suivi. Le mouvement est réalisé avec une vitesse croissante jusqu'à ce que le marqueur ne puisse plus être suivi dans les images. La dernière portion du mouvement avant l'échec du suivi est utilisé pour le calcul de la vitesse maximale: 8.81 mm/s en l'occurrence.

Pour évaluer la précision du suivi, un mouvement de 9 mm/s et 2°/s de vitesse de translation et de rotation a été simulé. La racine carrée de l'erreur quadratique moyenne obtenue pour la position est de 5.54 mm.

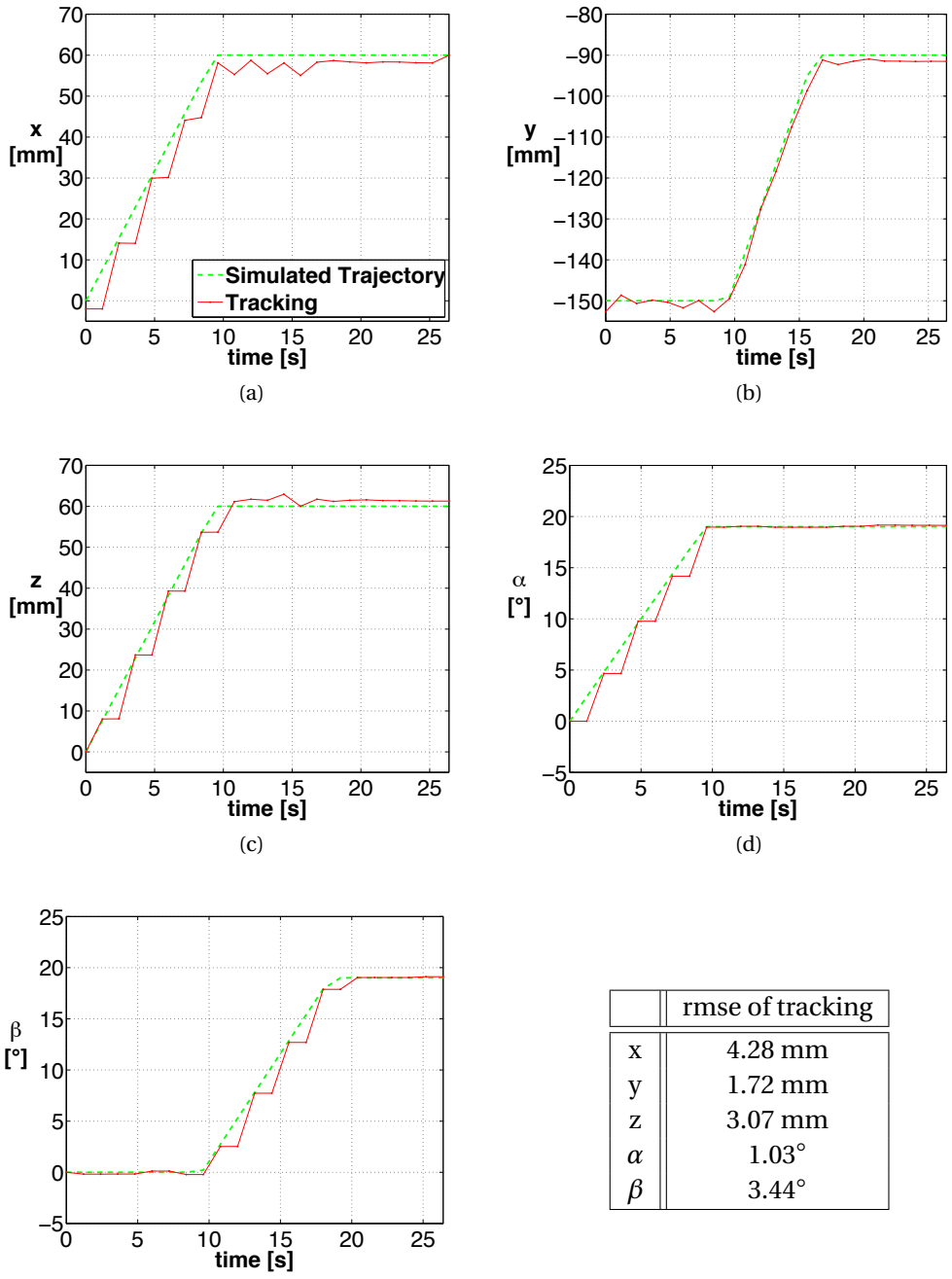
Pour l'évaluation expérimentale de la précision du suivi un mouvement d'une vitesse moyenne de 8.81 mm/s est réalisé sur le banc d'essai. L'orientation du marqueur est constante au cours du temps. L'erreur rmse obtenue pour le suivi est de 7.38 mm.

Conclusion

Un processus de suivi automatique basé sur une séquence IRM a été développé. Seule une image dédiée est utilisée lors de l'initialisation. Des simulations et expériences ont été réalisées pour l'évaluation de la précision et la vitesse maximale. Les résultats obtenus montrent des imprécisions dans le suivi.

D.3 Filtrage de Kalman pour l'alignement des plans image en IRM interventionnelle

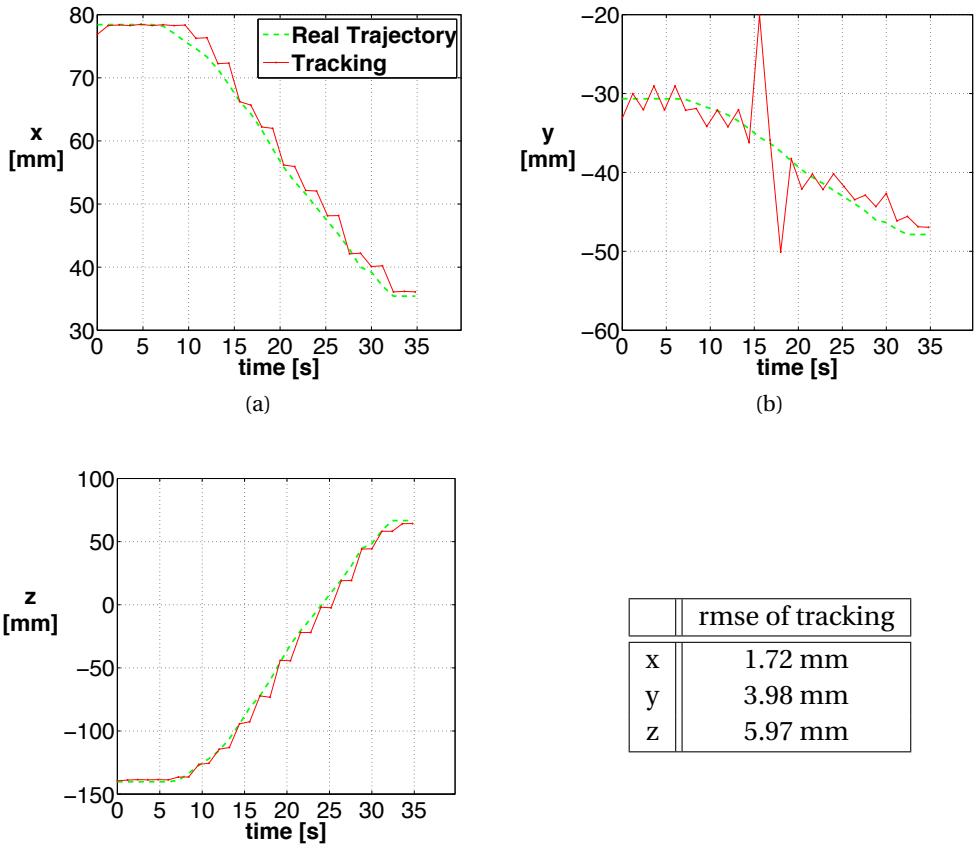
Pour améliorer les performances de suivi du marqueur, un filtre de Kalman est utilisé pour l'alignement des plans image.



	rmse of tracking
x	4.28 mm
y	1.72 mm
z	3.07 mm
α	1.03°
β	3.44°

(f)

Figure D.6: Résultats de simulation d'un mouvement de 9 mm/s de vitesse de translation et 2°/s de vitesse de rotation. Le mouvement simulé est représenté en vert dans les graphes (a-e). Les composantes de position et d'orientation du mouvement détecté sont représentées en rouge. Des imprécisions du suivi peuvent être observées pour les composantes x,z, α and β. Ces composantes ne sont pas détectables sur les deux plans images. La composante y peut être actualisée sur les deux plans dû à leur faible inclinaison. L'erreur rmse pour les composantes de la pose du marqueur est représentée en (f).



	rmse of tracking
x	1.72 mm
y	3.98 mm
z	5.97 mm

Figure D.7: Résultats expérimentaux pour le suivi d'un mouvement de vitesse moyenne de 8.81 mm/s. Comme pour les résultats des simulations, des imprécisions peuvent être observées pour le suivi. Les deux valeurs déviantes de la composante y sont dues à des erreurs de segmentation. L'erreur rmse sur la position est présentée dans le tableau (d).

Algorithme du filtre de Kalman

Si on assume un système contrôlé, discret et défini avec l'équation linéaire

$$\boldsymbol{x}(k) = A(k)\,\boldsymbol{x}(k-1) + \boldsymbol{B}(k)\,\boldsymbol{u}(k) + \boldsymbol{n}(k-1) \quad , \tag{D.1}$$

qui met en relation le vecteur état du système $\boldsymbol{x}$ à l'instant $k-1$ et k par la matrice de transition $\boldsymbol{A}(k)$. La matrice $\boldsymbol{B}(k)$ relie une consigne optionnelle et l'état du système.
La matrice de mesure $\boldsymbol{H}(k)$ lie le vecteur état du système et la sortie mesurée $\boldsymbol{z}(k)$:

$$\boldsymbol{z}(k) = \boldsymbol{H}(k)\,\boldsymbol{x}(k) + \boldsymbol{m}(k) \quad . \tag{D.2}$$

Le bruit du processus et bruit de mesure du système sont représentés par $\boldsymbol{n}(k-1)$ and $\boldsymbol{m}(k)$, respectivement. On assume qu'ils ont une distribution Gaussienne ($N(moyenne, variance)$) selon

$$\boldsymbol{n}(k) \sim N(0, \boldsymbol{Q}(k)) \tag{D.3}$$
$$\boldsymbol{m}(k) \sim N(0, \boldsymbol{R}(k)) \quad , \tag{D.4}$$

où $\boldsymbol{Q}(k)$ et $\boldsymbol{R}(k)$ représentent les matrices de covariance des bruits de processus et de mesure, respectivement.
Les variables internes du filtre de Kalman sont le vecteur d'état du système $\boldsymbol{x}$ et la matrice de covariance de l'erreur d'estimation $\boldsymbol{P}$. L'algorithme permet de réaliser une estimation optimale du vecteur d'état, établie sur la prédiction et la sortie mesurée disponible $\boldsymbol{z}(k)$. L'algorithme consiste en une étape de prédiction et une étape d'estimation:

Prédiction

$$\hat{\boldsymbol{x}}(k|k-1) = \boldsymbol{A}(k)\,\hat{\boldsymbol{x}}(k-1|k-1) + \boldsymbol{B}(k)\,\boldsymbol{u}(k) \tag{D.5}$$
$$\boldsymbol{P}(k|k-1) = \boldsymbol{A}(k)\,\boldsymbol{P}(k-1|k-1)\,\boldsymbol{A}^T(k) + \boldsymbol{Q}(k) \tag{D.6}$$

Estimation

$$\hat{\boldsymbol{x}}(k|k) = [\boldsymbol{I} - \boldsymbol{K}(k)\,\boldsymbol{H}(k)]\,\hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\,\boldsymbol{z}(k) \tag{D.7}$$
$$\boldsymbol{P}(k|k) = \boldsymbol{P}(k|k-1) - \boldsymbol{K}(k)\,\boldsymbol{S}(k)\,\boldsymbol{K}^T(k) \tag{D.8}$$

avec $\boldsymbol{K}(k)$ et $\boldsymbol{S}(k)$ dénommé gain de Kalman et matrice de covariance de l'innovation, respectivement:

$$\boldsymbol{K}(k) = \boldsymbol{P}(k|k-1)\,\boldsymbol{H}^T(k)\,\boldsymbol{S}^{-1}(k)$$
$$\boldsymbol{S}(k) = \boldsymbol{H}(k)\,\boldsymbol{P}(k|k-1)\,\boldsymbol{H}^T(k) + \boldsymbol{R}(k)$$

Les notations $\hat{\boldsymbol{x}}(k|k-1)$ et $\hat{\boldsymbol{x}}(k|k)$ définissent l'état prédit pour l'instant k, étant connu les mesures jusqu'à l'instant $k-1$ et l'état estimé à l'instant k, étant connu les mesures jusqu'à l'instant k, respectivement. La même notation est utilisée pour l'estimation de la matrice de covariance de l'erreur.
L'algorithme est initialisé avec les estimations initiales pour $\hat{\boldsymbol{x}}(k-1|k-1)$ et $\boldsymbol{P}(k-1|k-1)$ et démarre avec l'étape de prédiction.

Alignement des plans image avec filtrage Kalman

Le processus de suivi temps-réel avec filtrage Kalman utilise la même séquence IRM clinique pour l'acquisition des images. L'alignement des plans image fonctionne différemment du

processus initial. Au lieu de détecter le marqueur dans une image et aligner le prochain plan image en fonction de la détection, le filtre Kalman intervient pour l'alignement. La pose 3D du marqueur détectée dans l'image est introduite comme mesure dans le filtre de Kalman. Le filtre réalise une estimation et une prédiction de la pose prochaine du marqueur. Le prochain plan image va être aligné en fonction de la prédiction du filtre de Kalman.

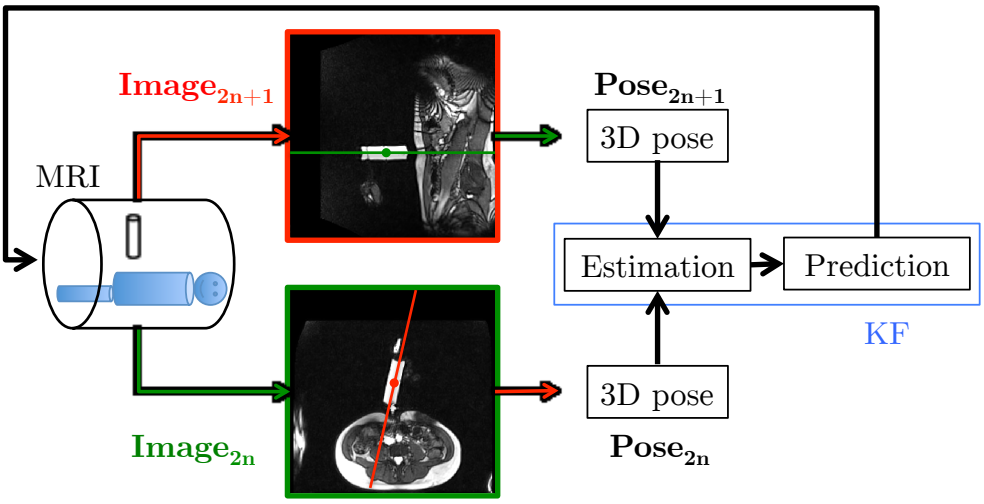


Figure D.8: Principe du processus d'alignement avec filtrage de Kalman: 2 plans orthogonaux, transversal (vert) et sagittal (rouge), sont acquis en alternance. La pose 3D du marqueur est détectée dans les images et introduite dans le filtre de Kalman qui réalise une estimation et une prédiction de la pose du marqueur. En fonction de la prédiction, la position et l'orientation du prochain plan image sont calculées et une commande d'alignement est envoyée à l'IRM.

Résultats

Les performances du processus d'alignement avec filtrage de Kalman sont évaluées de la même façon que pour le processus initial.

Pour évaluer la vitesse maximale pour le suivi, le cas extrême d'un mouvement du marqueur à 45° entre les plans image est simulé. La vitesse maximale obtenue est de 21.6 mm/s.

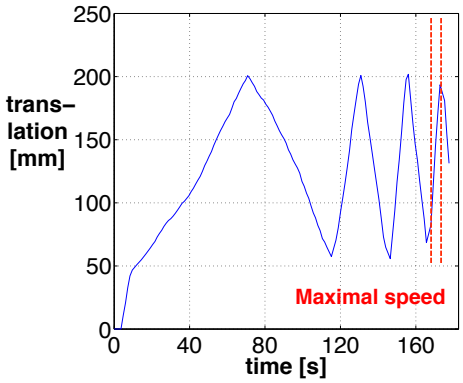


Figure D.9: Détermination expérimentale de la vitesse maximale du suivi. Le mouvement est réalisé avec une vitesse croissante jusqu'à ce que le marqueur ne puisse plus être suivi dans les images. La dernière portion du mouvement avant l'échec du suivi est utilisé pour le calcul de la vitesse maximale: 21.2 mm/s en l'occurrence.

Pour l'évaluation expérimentale, le mouvement oscillatoire est réalisé sur le banc d'essai par un utilisateur comme précédemment. Pour le processus avec filtrage Kalman, une vitesse maximale de 21.24 mm/s a été obtenue.

Pour évaluer la précision du suivi, un mouvement de 9 mm/s et 2°/s de vitesse de translation et de rotation a été simulé. La racine carrée de l'erreur quadratique moyenne est de 3.05 mm pour la position.

Pour l'évaluation expérimentale de la précision du suivi un mouvement d'une vitesse moyenne de 17.5 mm/s est réalisé sur le banc d'essai. L'orientation du marqueur est constante au cours du temps. L'erreur rmse obtenue pour le suivi est de 3.7 mm.

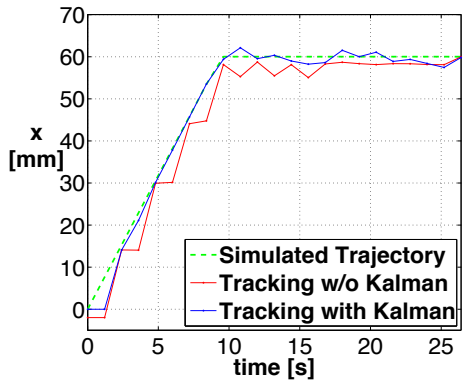
Conclusion

Grâce à l'utilisation du filtre de Kalman, les imprécisions du suivi observées pour le processus initial ont pu être atténuées. De plus la vitesse maximale du suivi est plus de deux fois supérieure grâce au filtre de Kalman.
Néanmoins dû au modèle à vitesse constante utilisé pour le filtre de Kalman et en raison de la fréquence d'imagerie très basse (0.83 Hz), des dépassements peuvent être constatés.

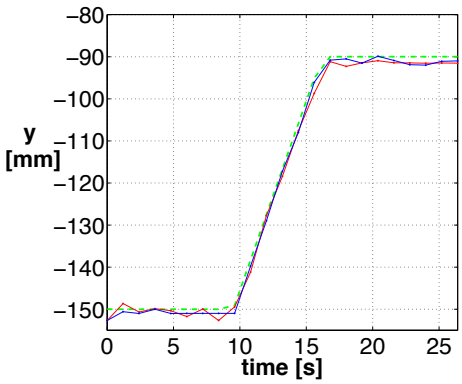
Dû à cette limitation, l'ajout d'un capteur à haute fréquence externe au scanner IRM a été décidé.

D.4 Extension du processus par un capteur RGB-D

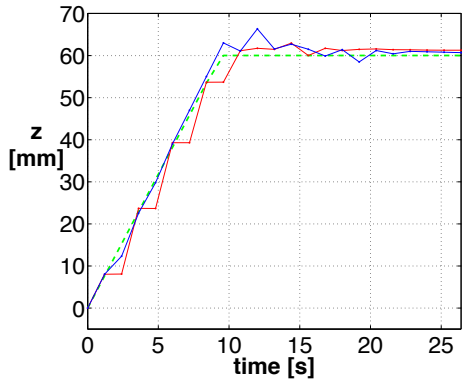
Un capteur RGB-D a été choisi pour améliorer le processus de suivi précédemment présenté. Ce capteur va permettre de mesurer à haute fréquence la pose 3D d'un marqueur dans le



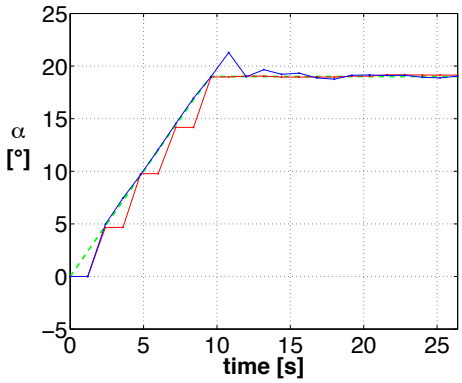
(a)



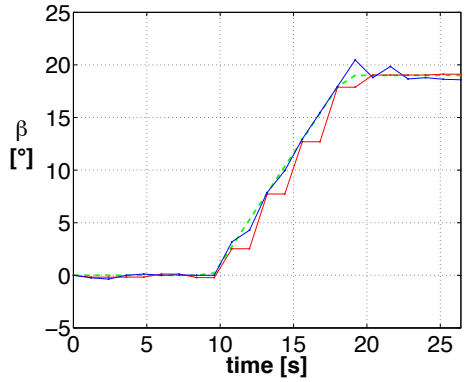
(b)



(c)



(d)



(e)

	rmse of tracking	
	w/o Kalman	with Kalman
x	4.28 mm	1.97 mm
y	1.72 mm	1.32 mm
z	3.07 mm	1.93 mm
α	1.03°	0.72°
β	3.44°	1.58°

(f)

Figure D.10: Résultats de simulation d'un mouvement de 9 mm/s de vitesse de translation et 2°/s de vitesse de rotation pour les processus de suivi sans et avec filtre de Kalman. Le mouvement simulé est représenté en vert dans les graphes (a-e). Les composantes de position et d'orientation du mouvement suivi sans (rouge) et avec Kalman (bleu) sont représentées. Les imprécisions du suivi observées pour le processus de suivi intial peuvent être atténuées avec le filtre de Kalman. Néanmoins, des dépassements peuvent être constatés pour le suivi avec le filtre de Kalman dus au modèle à vitesse constante et à la fréquence d'imagerie très basse (0.83 Hz). La comparaison de l'erreur rmse pour les composantes de la pose du marqueur est représentée en (f).

tunnel du scanner IRM.

Capteur RGB-D

Un capteur RGB-D est équipé d'une caméra RGB et d'un capteur de profondeur. La profondeur est déterminée avec un projecteur infra rouge (IR) projetant un motif qui est détecté en synchrone par une caméra IR dans la scène observée. Basée sur la différence entre motif projeté et détecté, les valeurs de profondeur de la scène sont déterminées. Le pas de quantification du capteur de profondeur est de 8 mm à 1.6m de distance, distance maximale de suivi dans le tunnel du scanner IRM.

Des tests de compatibilité du capteur RGB-D avec un environnement IRM de 1.5 T ont été réalisés pour vérifier son bon fonctionnement et l'absence de perturbations pour le scanner IRM. Des artéfacts mineurs ont pu être constaté du à l'utilisation du capteur RGB-D dans la salle d'IRM. La cause principale des ces artéfacts est l'utilisation d'un câble USB passée à travers le guide d'onde de la salle d'IRM pour connecter le capteur RGB-D à un PC externe dans la salle de commande de l'IRM.

Pour la détection du marqueur dans les images du capteur RGB-D, celui-ci est équipé de deux balles à ses extrémités. La détection du marqueur avec le capteur RGB-D consiste à seuiller l'image RGB et à y segmenter les balles pour ensuite déterminer leur position 3D grâce à l'image de profondeur correspondante.

Alignement des plans image hybride utilisant scanner IRM et capteur RGB-D

Le processus de suivi temps-réel hybride utilise la même séquence IRM clinique pour l'acquisition des images. Le principe est de combiner les mesures de la caméra et les détections dans les images IRM. Dans une première phase, le suivi du marqueur est réalisé en utilisant uniquement l'imagerie IRM. Pendant cette étape, la position du marqueur est détectée en synchrone dans les images IRM et par le capteur RGB-D afin de réaliser le recalage entre les deux modalités. Par la suite, les mesures des deux modalité sont fusionnées pour repousser les limites de chacun des dispositifs.

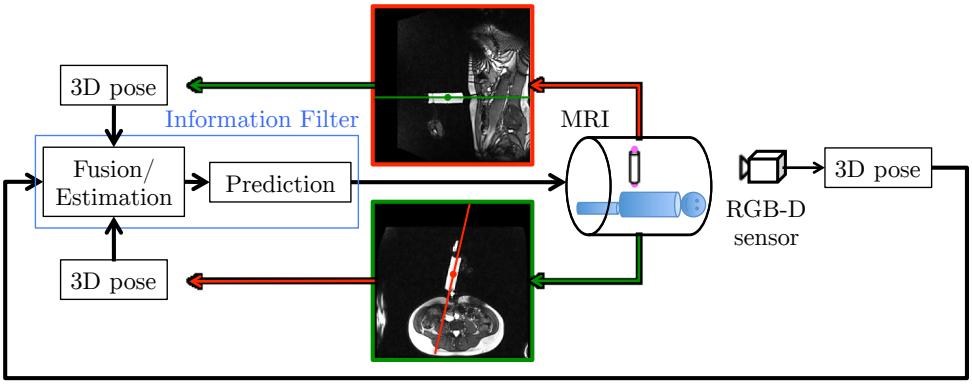


Figure D.11: Principe du processus d'alignement hybride: 2 plans orthogonaux, transversal (vert) et sagittal (rouge), sont acquis en alternance. La pose 3D du marqueur est détectée dans les images IRM et par le capteur RGB-D. La pose détectée par le capteur RGB-D est exprimée dans le repère de l'IRM. Un filtre d'Information fusionne les poses 3D mesurées et réalise une prédiction de la pose du marqueur. En fonction de cette prédiction, la position et l'orientation du prochain plan image sont calculées et une commande d'alignement est envoyée à l'IRM.

Comme la qualité du recalage entre les deux modalités (capteur RGB-D, IRM) dépend fortement des points utilisés, une approche de recalage en ligne a été implémentée. Celle-ci permet d'actualiser la liste des points utilisés pour le recalage au cours du temps. En fonction des actualisations de cette liste, la transformation entre les systèmes de coordonnées des deux modalités est recalculée et améliorée au cours du temps pendant le suivi.

Algorithme du filtre d'Information

Pour décrire l'algorithme du filtre d'Information, le même système contrôlé, discret pour le filtre de Kalman est supposé:

$$\boldsymbol{x}(k) = A(k)\,\boldsymbol{x}(k-1) + \boldsymbol{B}(k)\,\boldsymbol{u}(k) + \boldsymbol{n}(k-1) \quad , \tag{D.9}$$

$$\boldsymbol{z}(k) = \boldsymbol{H}(k)\,\boldsymbol{x}(k) + \boldsymbol{m}(k) \quad . \tag{D.10}$$

Le vecteur d'état du système $\boldsymbol{x}$ et la matrice de covariance de l'erreur d'estimation $\boldsymbol{P}$, variables internes du filtre de Kalman, sont remplacés pour le filtre d'Information par la matrice d'information $\boldsymbol{Y}$ et le vecteur d'information $\hat{\boldsymbol{y}}$:

$$\boldsymbol{Y} = \boldsymbol{P}^{-1} \tag{D.11}$$

$$\hat{\boldsymbol{y}} = \boldsymbol{Y}\,\hat{\boldsymbol{x}} \quad . \tag{D.12}$$

L'algorithme du filtre d'Information consiste comme celui du filtre de Kalman en une étape d'estimation et une étape de prédiction.

Prédiction

$$\hat{\boldsymbol{y}}(k|k-1) = \boldsymbol{L}(k|k-1)\ \hat{\boldsymbol{y}}(k-1|k-1) \tag{D.13}$$

$$\boldsymbol{Y}(k|k-1) = \left[\boldsymbol{A}(k)\ \boldsymbol{Y}^{-1}(k-1|k-1)\ \boldsymbol{A}^{T}(k) + \boldsymbol{Q}(k)\right]^{-1} \tag{D.14}$$

avec le coefficient de propagation $\boldsymbol{L}(k|k-1)$ défini par:

$$\boldsymbol{L}(k|k-1) = \boldsymbol{Y}(k|k-1)\ \boldsymbol{A}(k)\ \boldsymbol{Y}^{-1}(k-1|k-1)$$

Estimation

$$\hat{\boldsymbol{y}}(k|k) = \hat{\boldsymbol{y}}(k|k-1) + \sum_{j=1}^{N} \boldsymbol{i}_j(k) \tag{D.15}$$

$$\boldsymbol{Y}(k|k) = \boldsymbol{Y}(k|k-1) + \sum_{j=1}^{N} \boldsymbol{I}_j(k) \tag{D.16}$$

avec la contribution locale au vecteur d'information $\boldsymbol{i}_j(k)$ et la matrice d'information associée $\boldsymbol{I}_j(k)$:

$$\begin{aligned} \boldsymbol{i}_j(k) &= \boldsymbol{H}_j^{T}(k)\ \boldsymbol{R}_j^{-1}(k)\ \boldsymbol{z}_j(k) \\ \boldsymbol{I}_j(k) &= \boldsymbol{H}_j^{T}(k)\ \boldsymbol{R}_j^{-1}(k)\ \boldsymbol{H}_j(k) \end{aligned} \tag{D.17}$$

Adaptée à l'utilisation hybride des images IRM et du capteur RGB-D, l'étape d'estimation est définie par:

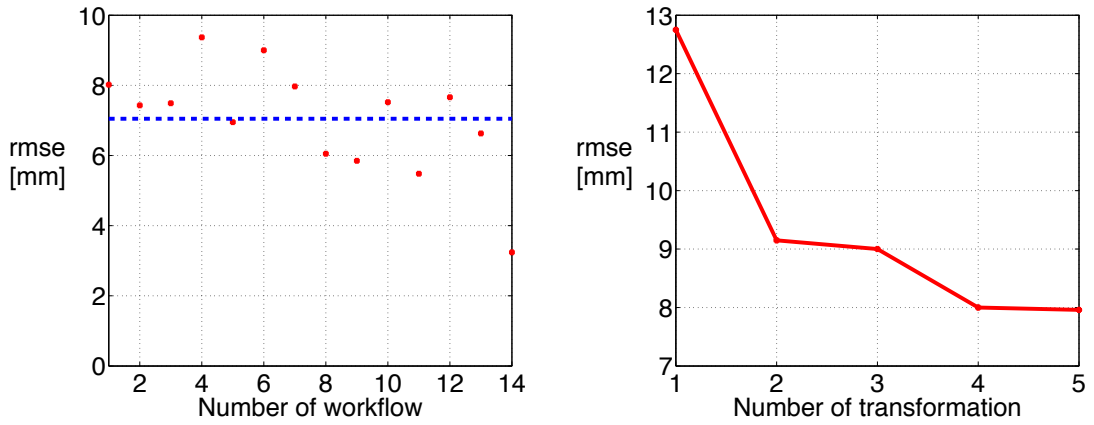
$$\begin{aligned} \sum_{j=1}^{N} \boldsymbol{i}_j(k) &= \boldsymbol{i}_{MRI}(k) + \boldsymbol{i}_{RGB-D}(k) \\ &= \boldsymbol{H}_{MRI}^{T}(k)\ \boldsymbol{R}_{MRI}^{-1}(k)\ \boldsymbol{z}_{MRI}(k) + \boldsymbol{H}_{RGB-D}^{T}(k)\ \boldsymbol{R}_{RGB-D}^{-1}(k)\ \boldsymbol{z}_{RGB-D}(k) \end{aligned} \tag{D.18}$$

$$\begin{aligned} \sum_{j=1}^{N} \boldsymbol{I}_j(k) &= \boldsymbol{I}_{MRI}(k) + \boldsymbol{I}_{RGB-D}(k) \\ &= \boldsymbol{H}_{MRI}^{T}(k)\ \boldsymbol{R}_{MRI}^{-1}(k)\ \boldsymbol{H}_{MRI}(k) + \boldsymbol{H}_{RGB-D}^{T}(k)\ \boldsymbol{R}_{RGB-D}^{-1}(k)\ \boldsymbol{H}_{RGB-D}(k) \end{aligned} \tag{D.19}$$

Dans le cas d'une mesure non disponible d'un des 2 capteurs (dû à une erreur de détection, p.ex.), le terme correspondant est retiré de la somme. L'impact de l'information des 2 capteurs peut être pondéré par les matrices de covariance du bruit de mesure ($\boldsymbol{R}_{MRI}$ et $\boldsymbol{R}_{RGB-D}$).

Résultats

L'évaluation du recalage est effectué effectué sur 14 séquences de suivi (durée moyenne de 1.5 min) où un utilisateur réalise un mouvement libre. La qualité du recalage est évaluée avec la racine carré de l'erreur quadratique moyenne (rmse) entre les points acquis dans le système de coordonnées de l'IRM et les points acquis dans le système de coordonnées du capteur RGB-D reprojetés dans le système de coordonnées de l'IRM par la transformation à évaluer. Pour l'approche de recalage en ligne implémentée, l'erreur rmse s'améliore au cours du temps. Les meilleures erreurs obtenues à la fin de chaque processus ont été comparées. L'erreur moyenne de recalage est de 7.04 mm.
En moyenne, la première transformation entre les deux systèmes de coordonnées a pu être obtenue après 70 s, dépendant du mouvement réalisé par l'utilisateur.



(a) (b)

Figure D.12: Evaluation de l'erreur rmse du recalage. (a): L'erreur rmse de la meilleur transformation trouvée pour chaque processus est représentée (rouge). L'erreur moyenne de recalage est de 7.04 mm. (b): L'évolution type de l'erreur de recalage pendant un suivi en fonction du numéro de transformation identifiée successivement entre IRM et capteur RGB-D.

Pour l'évaluation expérimentale de la vitesse maximale du suivi, le mouvement oscillatoire est réalisé sur le banc d'essai par un utilisateur comme précédemment. Pour le processus hybride, une vitesse maximale de 45.3 mm/s a été obtenue.

Pour l'évaluation expérimentale de la précision du suivi, un mouvement d'une vitesse moyenne de 15.1 mm/s est réalisé sur le banc d'essai. L'orientation du marqueur est constante au cours du temps. L'erreur rmse obtenue pour le suivi est de 5.7 mm.

Conclusion

L'utilisation hybride du capteur RGB-D et des images IRM permet de découpler la fréquence du suivi de la fréquence d'imagerie de la séquence clinique IRM. Or, la vitesse maximale du suivi est plus que 2 fois plus élevée que pour les processus de suivi précédents.

D.5 Conclusion

Conclusions

Dans cette thèse est présentée une approche pour l'alignement automatique des plans image IRM en temps-réel, basée sur la détection d'un marqueur dans les images IRM. Les avantages de cette méthode sont que:

- le temps d'imagerie dédiée à la détection est minimisé, s
- peu d'instrumentation additionnelle est nécessaire et ainsi le cout de matériel est minimisé.

L'approche développée est fondée sur un dispositif consistant en un marqueur passif et des micro-bobines. Un algorithme de suivi établi sur les images IRM a été proposé et implémenté dans une interface logicielle. La méthode consiste en l'acquisition successive de deux plans image orthogonaux. Chacun des plans est aligné à l'axe du marqueur grâce à la détection de celui-ci dans l'image précédente.
Un logiciel de simulation a été développé permettant d'évaluer les performances du suivi dans des conditions contrôlées et reproductibles. De plus un banc d'essai compatible IRM a été utilisé pour la validation expérimentale des performances du suivi.
Actuellement, cette approche permet d'utiliser le marqueur en tant que sonde interventionnelle pour la recherche du point d'insertion de l'aiguille, par exemple.

Dû aux imprécisions observées pendant le suivi, l'approche initiale a été étendue par un filtre de Kalman. Comme attendu, l'utilisation du filtre de Kalman a permis d'améliorer les performances du suivi au niveau de la vitesse et de la précision. Le filtre de Kalman permet d'améliorer le suivi pour les mouvements subissant des accélérations faibles. Pour de fortes accélérations, l'amélioration des performances due au filtre de Kalman disparaît, à cause de l'utilisation d'un modèle à vitesse constante et à la basse fréquence d'acquisition de la séquence IRM utilisée.

La basse fréquence d'acquisition et la dépendance de la vitesse du suivi à la séquence d'imagerie IRM utilisée sont des inconvénients communs des méthodes passives. Afin de découpler la vitesse du suivi de la fréquence d'imagerie IRM, un capteur RGB-D a été intégré au processus de suivi passif. Le capteur RGB-D permet de mesurer la pose 3D d'un objet directement et s'est révélé compatible avec l'environnement IRM. Une méthode de recalage en ligne a été implémentée permettant d'intégrer ce processus hybride facilement dans la

routine clinique. L'intégration du capteur RGB-D dans le processus de suivi passif est réalisé avec un filtre d'information fusionnant les mesures du capteur RGB-D et celles issues des images IRM. Ainsi, les deux modalités contribuent à l'estimation de la pose du marqueur et à l'alignement des plans image. De plus, la robustesse du suivi est améliorée de manière significative grâce à la haute fréquence d'acquisition du capteur RGB-D.

En général, on peut conclure que l'approche purement passive manque de réactivité face à des mouvements rapides et de grande amplitude. Néanmoins, cette approche est praticable pour le suivi de mouvements modérés.
La combinaison de l'approche passive avec une approche active s'affranchit de cette réactivité restreinte et a montré des résultats prometteurs. Ainsi la robustesse du suivi est nettement améliorée grâce à la combinaison des points forts de chaque modalité.

Perspectives

Afin d'améliorer la robustesse des algorithmes de segmentation et de classification, le marqueur ainsi que le corps du patient pourraient être segmentés en utilisant des méthodes de contours actifs (p.ex. un algorithme snake). Une plus grande robustesse de la segmentation des images devrait également résulter d'une paramétrisation des objets segmentés avec des superellipses. Des tests pour la compatibilité temps-réel de telles approches doivent être réalisés. De plus, l'utilisation de contours actifs est censée améliorer la segmentation du marqueur sans micro-bobines dans les images d'initialisation.

Les analyses temporelles des différentes versions du processus ont montré que les temps de communication entre les divers ordinateurs sont difficile à évaluer et sont variables dû à l'utilisation du protocole Ethernet. L'alignement des plans image pourrait être amélioré si la pose du marqueur est prédite pour l'instant d'acquisition de la prochaine image. Or, l'implémentation du processus directement sur la console à architecture PC de l'IRM permettrait donc une accélération du temps de rafraichissement des images et une facilitation de l'implémentation. Des simulations ont permis de montrer l'intérêt de travailler avec la fréquence d'imagerie la plus haute possible. Dans ce contexte, une optimisation de la séquence IRM est envisageable en trouvant un compromis cliniquement praticable entre qualité et temps d'acquisition des images IRM.

L'utilisation d'un câble USB passé à travers le guide d'onde de la salle d'IRM affecte la qualité des images IRM. Pour cette raison l'utilisation d'une pair de transducteurs permettant de transmettre les données RGB-D par fibre optique est à investiguer.

Dans ce travail, une seringue stérile remplie d'un produit de contraste et prêt-à-utiliser dans un environnement chirurgical a été choisie. Un but est de concevoir un marqueur compacte et compatible avec l'utilisation d'une aiguille. Finalement, le processus de suivi développé pourrait être combiné avec un algorithme de détection d'aiguille afin de suivre précisément l'avancement de l'aiguille compatible IRM.

References

[BCG07] G. Boracchi, V. Caglioti, and A. Giusti. Ball position and motion reconstruction from blur in a single perspective image. In *14th International Conference on Image Analysis and Processing, 2007. ICIAP 2007*, pages 87–92, 2007.

[BCY96] M. Burl, G. A. Coutts, and I. R. Young. Tuned fiducial markers to identify body locations with minimal perturbation of tissue magnetization. *Magnetic Resonance in Medicine*, 36(3):491–493, 1996.

[Ber04] B. A. Berg. *Markov Chain Monte Carlo Simulations and Their Statistical Analysis: With Web-based Fortran Code*. World Scientific, January 2004.

[BG12] W. Block and B. Grabow. Pulse sequences for interventional MRI. In *Interventional Magnetic Resonance Imaging*. Springer, August 2012.

[BSKO+02] R. Blanco Sequeiros, R. Klemola, R. Ojala, L. Jyrkinen, E. Lappi-Blanco, Y. Soini, and O. Tervonen. MRI-guided trephine biopsy and fine-needle aspiration in the diagnosis of bone lesions in low-field (0.23 t) MRI system using optical instrument tracking. *European Radiology*, 12(4):830–835, April 2002.

[BSLK04] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan. *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. John Wiley & Sons, April 2004.

[BUS+06] M. Bock, R. Umathum, J. Sikora, S. Brenner, E. N. Aguor, and W. Semmler. A faraday effect position sensor for interventional magnetic resonance imaging. *Physics in medicine and biology*, 51(4):999–1009, February 2006. PMID: 16467592.

[BVZ+04] M. Bock, S. Volz, S. Zühlsdorff, R. Umathum, C. Fink, P. Hallscheidt, and W. Semmler. MR-guided intravascular procedures: Real-time parameter control and automated slice positioning with active tracking coils. *Journal of Magnetic Resonance Imaging*, 19(5):580–589, 2004.

[BW08] M. Bock and F. K. Wacker. MR-guided intravascular interventions: Techniques and applications. *Journal of Magnetic Resonance Imaging*, 27(2):326–338, 2008.

[CD02] D. Crisan and A. Doucet. A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50(3):736–746, 2002.

[CDD94] Q. S. Chen, M. Defrise, and F. Deconinck. Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(12):1156–1168, 1994.

[CG92] G. Casella and E. I. George. Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174, August 1992. ArticleType: research-article / Full publication date: Aug., 1992 / Copyright © 1992 American Statistical Association.

[Daa01] V. Daanen. *Suivi 3D de gestes chirurgicaux. Application à l'IRM interventionnelle*. PhD thesis, Université des sciences et technologies de Lille, Lille, France, May 2001.

[DDFG01] A. Doucet, N. De Freitas, and N. Gordon. *Sequential Monte Carlo methods in practice*. Springer, 2001.

[DKE+05] S. P. DiMaio, D. F. Kacher, R. E. Ellis, G. Fichtinger, N. Hata, G. P. Zientara, L. P. Panych, R. Kikinis, and F. A. Jolesz. Needle artifact localization in 3T MR images. In *Medicine Meets Virtual Reality 14: Accelerating Change in Healthcare: Next Medical Toolkit*, pages 120–125. IOS Press, 2005.

[DLF06] J. N. Draper, M. L. Lauzon, and R. Frayne. Passive catheter visualization in magnetic resonance-guided endovascular therapy using multicycle projection dephasers. *Journal of magnetic resonance imaging: JMRI*, 24(1):160–167, July 2006. PMID: 16758467.

[DMC00] L. Drolet, F. Michaud, and J. Cote. Adaptable sensor fusion using multiple kalman filters. In *2000 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2000. (IROS 2000). Proceedings*, volume 2, pages 1434–1439 vol.2, 2000.

[dORB+08a] A. de Oliveira, J. Rauschenberg, D. Beyersdorff, W. Semmler, and M. Bock. Automatic passive tracking of an endorectal prostate biopsy device using phase-only cross-correlation. *Magnetic Resonance in Medicine*, 59(5):1043–1050, 2008.

[dORB+08b] A. de Oliveira, J. Rauschenberg, D. Beyersdorff, W. Semmler, and M. Bock. A new system for passive tracking of a prostate biopsy device with automatic 3D needle position estimation. In *Proceedings of the 16th Annual Meeting of ISMRM*, page 3003, 2008.

[DSF+07] S. P. DiMaio, E. Samset, G. Fischer, I. Iordachita, G. Fichtinger, F. Jolesz, and C. M Tempany. Dynamic MRI scan plane control for passive tracking of instruments

and devices. In *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention*, MICCAI'07, pages 50–58, Berlin, Heidelberg, 2007. Springer-Verlag.

[DWL02] J. L Duerk, Eddy Y Wong, and Jonathan S Lewin. A brief review of hardware for catheter tracking in magnetic resonance imaging. *Magma (New York, N.Y.)*, 13(3):199–208, January 2002. PMID: 11755097.

[FMBM08] O. Felfoul, J. B. Mathieu, G. Beaudoin, and S. Martel. In vivo MR-tracking based on magnetic signature selective excitation. *IEEE Transactions on Medical Imaging*, 27(1):28–35, January 2008. PMID: 18270059.

[FPF99] A. Fitzgibbon, M. Pilu, and R.B. Fisher. Direct least square fitting of ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):476–480, 1999.

[GAB+97] A. Glowinski, G. Adam, A. Bücker, J. Neuerburg, J. J. van Vaals, and R. W. Günther. Catheter visualization using locally induced, actively controlled field inhomogeneities. *Magnetic Resonance in Medicine*, 38(2):253–258, 1997.

[Gro02] B. Grocholsky. *Information-Theoretic Control of Multiple Sensor Platforms*. PhD thesis, The University of Sydney, Sydney, 2002.

[GTA+12] A. Gangi, G. Tsoumakidou, O. Abdelli, X. Buy, M. de Mathelin, D. Jacqmin, and H. Lang. Percutaneous MR-guided cryoablation of prostate cancer: initial experience. *European Radiology*, 22(8):1829–1835, August 2012.

[Has70] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, April 1970. ArticleType: research-article / Full publication date: Apr., 1970 / Copyright © 1970 Biometrika Trust.

[HHL09] M. Higgins, D. Hall, and J. Llinas. *Handbook of Multisensor Data Fusion*. CRC Press Inc, 2009.

[HM11] R. Horaud and O. Monga. *Vision par ordinateur: outils fondamentaux*. Hermès, 2 edition, May 2011.

[HSXS13] J. Han, L. Shao, D. Xu, and J. Shotton. Enhanced computer vision with microsoft kinect sensor: A review. *IEEE Transactions on Cybernetics*, 43(5):1318–1334, 2013.

[htt13a] http://www.healthcare.philips.com, 2013.

[htt13b] http://www.robinmedical.com/endoscout.html, 2013.

[htt13c] http://www.asus.com/Multimedia/Xtion_PRO_LIVE/. Xtion PRO LIVE - multimedia - ASUS, 2013.

[Jol98] F. A. Jolesz. Invited. interventional and intraoperative MRI: a general overview of the field. *Journal of Magnetic Resonance Imaging*, 8(1):3–7, 1998.

[JU97] S. J. Julier and J. K. Uhlmann. A new extension of the kalman filter to nonlinear systems. pages 182–193, 1997.

[Kal60] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82:35—45, 1960.

[KE12] K. Khoshelham and S. O. Elberink. Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors*, 12(12):1437–1454, February 2012.

[KGA04] E. Kochavi, D. Goldsher, and H. Azhari. Method for rapid MRI needle tracking. *Magnetic Resonance in Medicine*, 51(5):1083–1087, 2004.

[KKKR13] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi. Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14(1):28–44, January 2013.

[KKR11] M. Kaiser, J. Krug, and G. H. Rose. Interventional MRI: minimal-invasive surgery under MR guidance. In *Microwave Symposium Digest (MTT), 2011 IEEE MTT-S International*, page 1, June 2011.

[KWT88] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, January 1988.

[LHMT01] H. Liu, W. A. Hall, A. J. Martin, and C. L. Truwit. Biopsy needle tip artifact in MR-guided neurosurgery. *Journal of Magnetic Resonance Imaging*, 13(1):16–22, 2001.

[LOCBR00] J. A. Lopez-Orozco, J. M. de la Cruz, E. Besada, and P. Ruipérez. An asynchronous, robust, and distributed multisensor fusion system for mobile robots. *The International Journal of Robotics Research*, 19(10):914–932, January 2000.

[LR13] N. Loy Rodas. *Object Detection in the Interventional Room using RGB-D Cameras*. PhD thesis, INSA of Strasbourg, Strasbourg, 2013.

[Mah07] R. P. S. Mahler. *Statistical Multisource-Multitarget Information Fusion*. Artech House, Inc., Norwood, MA, USA, 2007.

[May82] P. S. Maybeck. *Stochastic Models, Estimation and Control*. Academic Press, 1982.

[MDW93] A. G. O. Mutambara and H. F. Durrant-Whyte. A formally verified modular decentralized robot control system. In *Proceedings of the 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems '93, IROS '93*, volume 3, pages 2023–2030 vol.3, 1993.

[MKdO+09] F. Maier, A. J. Krafft, R. J. de Oliveira, W. Semmler, and M. Bock. MRPen – 3D marker tracking for percutaneous interventions. In *Proceedings of the 17th Annual Meeting of ISMRM*, page 4421, 2009.

[MKS+11] F. Maier, A. J. Krafft, R. J. Stafford, J. P. Yung, R. Dillmann, W. Semmler, and M. Bock. 3D passive marker tracking for MR-Guided interventions. In *Proceedings of the 19th Annual Meeting of ISMRM*, page 3749, 2011.

[MNB+11] H. Mönnich, P. Nicolai, T. Beyl, J. Raczkowsky, and H. Worn. A supervision system for the intuitive usage of a telemanipulated surgical robotic setup. In *2011 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 449–454, 2011.

[MTKB08] M. Moche, R. Trampel, T. Kahn, and H. Busse. Navigation concepts for MR image-guided interventions. *Journal of Magnetic Resonance Imaging*, 27(2):276–291, 2008.

[Mut98] A. G. O. Mutambara. *Decentralized Estimation and Control for Multisensor Systems*. CRC Press Inc, January 1998.

[Nag13] F. Nageotte. Registration and image-based control for robot-assisted medical interventions (class notes). 2013.

[Ots79] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, January 1979.

[PBS+11] L. Pan, J. Barbot, S. Shea, S. Patil, K. Kirchberg, G. Meredith, T. Meng, E. Kholmovski, S. Vijayakumar, K. Vij, M. Guttmann, P. Piferi, K. Jenkins, and C. Lorenz. An integrated system for catheter tracking and visualization in MR-Guided cardiovascular interventions. In *Proceedings of the 19th Annual Meeting of ISMRM*, page 195, 2011.

[PM07] D. Padfield and J. Miller. A label geometry image filter for multiple object measurement. *The Insight Journal*, 2007.

[Pra10] A. Prat. *Sensordatenfusion und Bildverarbeitung zur Objekt- und Gefahrenerkennung*. PhD thesis, Technische Universität Braunschweig, Braunschweig, June 2010.

[QZK+05a] H. H. Quick, M. O. Zenge, H. Kuehl, G. Kaiser, S. Aker, H. Eggebrecht, S. Massing, and M. E. Ladd. Wireless active catheter visualization: Passive decoupling methods and their impact on catheter visibility. In *Proceedings of the 13th Annual Meeting of ISMRM*, page 2164, 2005.

[QZK+05b] H. H. Quick, M. O. Zenge, H. Kuehl, G. Kaiser, S. Aker, S. Massing, S. Bosk, and M. E. Ladd. Interventional magnetic resonance angiography with no strings attached: Wireless active catheter visualization. *Magnetic Resonance in Medicine*, 53(2):446–455, 2005.

[RDW91] B. S. Rao and H. F. Durrant-Whyte. Fully decentralised algorithm for multisensor kalman filtering. *Control Theory and Applications, IEE Proceedings D*, 138(5):413–420, 1991.

[RME+08] M. Rea, D. McRobbie, H. Elhawary, Z. Tse, M. Lamperth, and I. Young. System for 3-d real-time tracking of MRI-Compatible devices by image processing. *Mechatronics, IEEE/ASME Transactions on*, 13(3):379–382, June 2008.

[RME+09] M. Rea, D. McRobbie, H. Elhawary, Z. Tse, M. Lamperth, and I. Young. Subpixel localisation of passive micro-coil fiducial markers in interventional MRI. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 22:71–76, 2009. 10.1007/s10334-008-0143-1.

[RW95] P. L. Rosin and G. A. W. West. Curve segmentation and representation by superellipses. *Vision, Image and Signal Processing, IEE Proceedings* -, 142(5):280–288, 1995.

[SGKU02] T. Schirmer, Y. Geoffroy, S. J. Koran, and F. Ulrich. Signa SP/2 – a MRI system for image guided surgery. *Medical Laser Application*, 17(2):105–116, 2002.

[SJP11] J. Smisek, M. Jancosek, and T. Pajdla. 3D with kinect. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 1154–1160, 2011.

[SK08] B. Siciliano and O. Khatib. *Springer Handbook of Robotics*. Springer, 2008.

[SKH+13] J.-B. Schell, J.-B. Kammerer, L. Hebrard, E. Breton, D. Gounot, L. Cuvillon, and M. de Mathelin. Towards a hall effect magnetic tracking device for MRI. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2964–2967, 2013.

[SLAL11] T. Stoyanov, A. Louloudi, H. Andreasson, and A. J. Lilienthal. Comparative evaluation of range sensor accuracy in indoor environments. pages 19–24, 2011. 3D range sensing is one of the important topics in robotics, as it is often a component in vital autonomous subsystems like collision avoidance, mapping and semantic perception. The development of ...

[SMO+98] K. Shimizu, R. V. Mulkern, K. Oshio, L. P. Panych, S. S. Yoo, R. Kikinis, and F. A. Jolesz. Rapid tip tracking with MRI by a limited projection reconstruction technique. *J Magn Reson Imaging*, 8(1):262–264, 1998.

[SOCM01] H.S. Stone, M.T. Orchard, E.-C. Chang, and S.A. Martucci. A fast direct fourier-based algorithm for subpixel registration of images. *IEEE Transactions on Geoscience and Remote Sensing*, 39(10):2235–2243, 2001.

[SRG+10] W. Strehl, E. Rothgang, W. Gilson, K. Kirchberg, J. Hornegger, and C. Lorenz. A passive, image-based navigation tool for real-time MR-guided percutaneous interventional procedures. In *Proceedings of the 18th Annual Meeting of ISMRM*, page 1840, 2010.

[SVB03] J.-H. Seppenwoolde, M. A. Viergever, and C. J. G. Bakker. Passive tracking exploiting local signal conservation: The white marker phenomenon. *Magnetic Resonance in Medicine*, 50(4):784–790, 2003.

[TL89] R. Y. Tsai and R. K. Lenz. A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Transactions on Robotics and Automation*, 5(3):345–358, 1989.

[Tro12] J. Troccaz. Capteurs et recalages per-opératoires. In *Robotique médicale*. Hermès science publications-Lavoisier, Cachan, 2012.

[VBR+07] R. Viard, N. Betrouni, J. Rousseau, S. Mordon, O. Ernst, and S. Maouche. Needle positioning in interventional MRI procedure: real time optical localisation and accordance with the roadmap. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 2748–2751, 2007.

[VdWBV01] R. Van der Weide, C. J. G. Bakker, and M. A. Viergever. Localization of intravascular devices with paramagnetic markers in MR images. *IEEE Transactions on Medical Imaging*, 20(10):1061–1071, 2001.

[WB01] G. Welch and G. Bishop. An introduction to the kalman filter. Los Angeles, CA, 2001.

[WCD76] D. Willner, C. B. Chang, and K. P. Dunn. Kalman filter algorithms for a multisensor system. In *1976 IEEE Conference on Decision and Control including the 15th Symposium on Adaptive Processes*, volume 15, pages 570–574, 1976.

[WCW13] Z. Wei, W. Chen, and J. Wang. Semantic mapping for smart wheelchairs using RGB-D camera. *Journal of Medical Imaging and Health Informatics*, 3(1):94–100, March 2013.

[Wel96] G. Welch. *SCAAT: Incremental Tracking with Incomplete Information*. PhD thesis, University of North Carolina, 1996.

[Zha00] Z. Zhang. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(11):1330–1334, November 2000.

List of publications

International

- M. Neumann, E. Breton, L. Cuvillon, M. de Mathelin. Evaluation of a multi-modal passive and active tracking approach for real-time automatic scan plane alignment in interventional MRI, in 22th Annual meeting of the International Society for Magnetic Resonance in Medicine (ISMRM 2014), Milan, Italy, May 2014.

- M. Neumann, L. Cuvillon, E. Breton, M. de Mathelin. Kalman filter for real-time scan plane control using a passive marker, in 27th International Congress of Computer Assisted Radiology and Surgery (CARS 2013), Heidelberg, Germany, June 2013.

- M. Neumann, L. Cuvillon, E. Breton, M. de Mathelin. Evaluation of an Image-based Tracking Workflow with Kalman Filtering for Automatic Image Plane Alignment in Interventional MRI, in 35th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC 2013), Osaka, Japan, April 2013.

- M. Neumann, E. Breton, L. Cuvillon, L. Pan, C. Lorenz, M. de Mathelin. Evaluation of an Image-Based Tracking Workflow Using a Passive Marker and Resonant Micro-Coil Fiducials for Automatic Image Plane Alignment in Interventional MRI, in 34th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC 2012), San Diego, USA, August 2012.

- M. Neumann, E. Breton, L. Cuvillon, L. Pan, C. Lorenz, M. de Mathelin. Real-time scan plane control using a passive marker and resonant micro-coils, in 9th International Interventional MRI Symposium, Boston, MA, USA, September 2012.

National

- M. Neumann, E. Breton, L. Cuvillon, M. de Mathelin. Asservissement du plan d'acquisition par approche mixte marqueurs visuels et micro-bobines sans-fil en IRM interventionnelle : étude préliminaire, in 1er Congrès de la Société Française de Résonance Magnétique en Biologie et Médecine, Marseille, France, February 2012.

Invited

- M. Neumann, E. Breton, L. Cuvillon, M. de Mathelin. Interventional MRI using robotics vision, in France Life Imaging, Strasbourg, France, November 2013.