Computer Aided Medical Procedures
Prof. Dr. Nassir Navab

Dissertation

Multimodal Sensing for Autonomous Robotic Ultrasound Imaging

Salvatore Virga

Fakultät für Informatik
Technische Universität München

Technische Universität München
Fakultät für Informatik
Lehrstuhl für Informatikanwendungen in der Medizin

Multimodal Sensing for Autonomous Robotic Ultrasound Imaging

Salvatore Virga

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzende(r): Prof. Dr.-Ing. Alin Albu-Schäffer

Prüfer der Dissertation: 1. Prof. Dr. Nassir Navab

2. Prof. Danail Stoyanov

Die Dissertation wurde am 13.03.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 25.08.2020 angenommen.

Salvatore Virga
Multimodal Sensing for Autonomous Robotic Ultrasound Imaging
Dissertation, Version 1.1

Technische Universität München
Fakultät für Informatik
Lehrstuhl für Informatikanwendungen in der Medizin
Boltzmannstraße 3
85748 and Garching bei München

Abstract

Over the past decades, the advancements in the area of ultrasound imaging have continuously been improving its impact in the medical field. However, one fundamental aspect of this imaging modality has been totally unaffected by these research efforts since its initial conception: ultrasound acquisitions are performed manually by a trained physician or sonographer. Unfortunately, the level of expertise of the operator has a substantial impact on the clinical outcome. An ultrasound transducer has to be maneuvered to the right position, with the right inclination and applying the right force to (re)produce a specific ultrasound image. Often, the lack of an expert operator risks translating into a misdiagnosis. The same characteristics make the acquisition of ultrasound images an incredibly complex task to automate. Human operators rely on visual cues from the live ultrasound images, knowledge about the patient anatomy, and its response to the applied transducer pressure.

Robotics has been tackling the problem since the late 1990s, attempting to provide systems to reduce the required physical interaction of ultrasound operators. The main focus of the research in the field has been on teleoperation for ultrasound acquisitions. Mechanical transducer holders that are remotely controlled by an operator have been designed and validated for many clinical applications, with the common aim to reduce the need for an expert operator in situ. The visual feedback from the ultrasound data stream has been employed in visual servoing tasks, to automatically maintain the visibility of specific anatomical targets or follow the movements of surgical tools. Nonetheless, complete autonomous ultrasound systems have not yet been developed.

This dissertation aims at introducing the design and validation of an autonomous robotic ultrasound system. The same sensorial information that a human operator leverages are evaluated and combined to obtain a complete system that fully benefits from them. Visual information from the scene and pre-operative imaging are employed to obtain a valid robotic trajectory with respect to a selected clinical target. Force sensing enables safe acquisitions and optimizes the overall produced images using patient-specific analysis. The overall system is evaluated in clinical scenarios, targeting both diagnostic and interventional procedures.

Zusammenfassung

In den letzten Jahrzehnten führte der Fortschritt in der Ultraschallbildgebung zu einem stetig zunehmenden Einfluss dieser Technologie im medizinischen Bereich. Ein grundlegender Aspekt dieser Bildgebungsmodalität ist jedoch von diesen Forschungsbemühungen bisher völlig unberührt geblieben: Die Ultraschallaufnahmen werden manuell von einem ausgebildeten Arzt oder Sonographen durchgeführt. Leider hat der Kompetenzgrad des Anwenders einen erheblichen Einfluss auf das klinische Ergebnis der Aufnahmen. Eine Ultraschallsonde muss in die richtige Position, mit der richtigen Neigung und unter Anwendung der richtigen Kraft manövriert werden, um ein spezifisches Ultraschallbild (wieder) zu erzeugen. Oftmals besteht die Gefahr, dass das Fehlen eines sachkundigen Anwenders in einer Fehldiagnose resultiert. Die genannten Eigenschaften machen die Aufnahme von Ultraschallbildern zu einer sehr komplexen Aufgabe, die es zu automatisieren gilt. Der menschliche Anwender ist auf die visuellen Hinweise aus den Live-Ultraschallbildern, die Kenntnis der Anatomie des Patienten und seine Reaktion auf den angewandten Schallkopfdruck angewiesen.

Die Robotik beschäftigt sich seit Ende der 1990er Jahre mit diesem Problem und versucht Systeme bereitzustellen, die die erforderliche physische Interaktion zwichen Anwender und Ultraschallkopf reduziert. Der Forschungsschwerpunkt in diesem Bereich lag in der Teleoperation von Ultraschallakquisitionen. Mechanische Schallkopfhalterungen, die von einem Nutzer ferngesteuert werden, wurden für viele klinische Anwendungen entwickelt und validiert. Ihr gemeinsames Ziel besteht darin, den Bedarf an Expertenanwendern vor Ort zu reduzieren. Das visuelle Feedback aus den Ultraschallbildern wurde bei visuellen Servoaufgaben eingesetzt, um die Sichtbarkeit bestimmter anatomischer Ziele automatisch aufrechtzuerhalten oder den Bewegungen der chirurgischen Werkzeuge zu folgen. Dennoch wurden bisher noch keine vollständig autonomen Ultraschallsysteme entwickelt.

Ziel dieser Dissertation ist es, den Entwurf und die Validierung eines autonomen Ultraschall-Robotersystems vorzustellen. Dieselben sensorischen Informationen, die ein menschlicher Anwender nutzt, werden ausgewertet und kombiniert, um ein komplett integriertes System zu erhalten, das von diesen Informationen profitiert. Visuelle Informationen der Umgebung und die präoperative Bildgebung werden verwendet, um eine gültige Robotertrajektorie in Bezug auf ein ausgewähltes klinisches Ziel zu erhalten. Die Kraftsensorik ermöglicht sichere Aufnahmen und optimiert die insgesamt erzeugten Bilder durch eine patientenspezifische Anpassung. Das Gesamtsystem wird in klinischen Szenarien evaluiert, die sowohl auf diagnostische als auch auf interventionelle Verfahren abzielen.

Acknowledgments

In the past four years, I had the opportunity to be part of a great family from which it is hard to depart.

First of all, I want to express my deepest gratitude to my advisor, Prof. Nassir Navab. His charisma, constant support, and down-to-earth character create a great environment. What he built and achieved over the years is definitely a great example for all his students and the research community in general. The entire CAMP group has been and is composed of great individuals, not only excellent researchers but also wonderful human beings. With the inevitable risk of forgetting someone, I would like to thank people that I want to define as friends rather than colleagues. The senior members of the team: Christoph Hennersperger, Benjamin Frisch, Maximilian Baust, Tobias Lasser, and Federico Tombari, from whom I was lucky to learn a lot. I know I would not have started this path without you as mentors. Thanks to Johanna Wald, Ari Tran, Jakob Weiss, Benjamin Busam, Christian Schulte zu Berge, Nikola Rieke, Helisa Dhamo, Fabian Manhardt, Anca Stefanoiu, David Tan, Keisuke Tateno, Chiara Amat di Sanfilippo, José Gardiazabal and Fausto Milletarí, and in particular to Marco Esposito, Rüdiger Göbl, Julia Rackerseder, Beatrice Demiray, Javier Esteban and Oliver Zettinig with whom I had the pleasure to share the IFL lab as our second home. My everlasting brotherly love goes to Iro Laina and Christian Rupprecht, a constant source of fun, encouragement, and culinary experiments. Lastly, a big thank you goes to Martina Hilla for the non-stop support to the entire group and me.

In Munich, there is one special place that helped me reducing the pain of many long work evenings, so a thank you goes to the entire crew of Pizzesco.

To my friends in Palermo goes my gratitude for their many years of friendship. With many of them, I share the pain of leaving a beautiful land behind and the joy of experiencing it together on our visits back home.

Thanks to my parents Vincenzo and Rosalia, and my sister Rita, for their constant encouragement over all my studies. I hope that even if my parents aren't able to understand the language in which this thesis is written, they will be proud of it. Finally, thanks to Vera for her patience, love, and support during the writing of this dissertation.

Contents

List of Authored and Co-authored Publications

2019

[1] Fernanda Langsch*, **Salvatore Virga***, Javier Esteban, Rüdiger Göbl and Nassir Navab. "Robotic Ultrasound for Catheter Navigation in Endovascular Procedures". *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2019*. (*equal contribution).

2018

[2] Benjamin Busam, Patrick Ruhkamp, **Salvatore Virga**, Beatrice Lentes, Julia Rackerseder, Nassir Navab and Christoph Hennersperger. "Markerless Inside-Out Tracking for 3D Ultrasound Compounding". *Proceedings on Simulation, Image Processing, and Ultrasound Systems for Assisted Diagnosis and Navigation, 2018*.

[3] Javier Esteban, Walter Simson, Sebastian Requena Witzig, Anna Rienmüller, **Salvatore Virga**, Benjamin Frisch, Oliver Zettinig, Drazen Sakara, Yu-Mi Ryang, Nassir Navab and Christoph Hennersperger. "Robotic ultrasound-guided facet joint insertion". *International Journal of Computer Assisted Radiology and Surgery (IPCAI / IJCARS), 2018*.

[4] **Salvatore Virga**, Rüdiger Göbl, Maximilian Baust, Nassir Navab and Christoph Hennersperger. “Use the force: Deformation correction in robotic 3D ultrasound”. *International Journal of Computer Assisted Radiology and Surgery (IPCAI / IJCARS), 2018*.

2017

[5] Oliver Zettinig, Benjamin Frisch, **Salvatore Virga**, Marco Esposito, Anna Rienmüller, Bernhard Meyer, Christoph Hennersperger, Yu-Mi Ryang and Nassir Navab. “3D ultrasound registration-based visual servoing for neurosurgical navigation”. *International Journal of Computer Assisted Radiology and Surgery (IJCARS), 2017*.

[6] Rüdiger Göbl, **Salvatore Virga**, Julia Rackerseder, Benjamin Frisch, Nassir Navab and Christoph Hennersperger. “Acoustic window planning for ultrasound acquisition”. *International Journal of Computer Assisted Radiology and Surgery (IPCAI / IJCARS), 2017*.

2016

[7] Christoph Hennersperger*, Bernhard Fuerst*, **Salvatore Virga***, Oliver Zettinig, Benjamin Frisch, Thomas Neff and Nassir Navab. "Towards MRI-based autonomous robotic US acquisitions: a first feasibility study". *IEEE Transactions on Medical Imaging (TMI), 2016*.[1] (*equal contribution)

[8] **Salvatore Virga***, Oliver Zettinig*, Marco Esposito, Karin Pfister, Benjamin Frisch, Thomas Neff, Nassir Navab and Christoph Hennersperger. "Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms". *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2016*.[1] (*equal contribution)

2015

[9] **Salvatore Virga**, Verena Dogeanu, Pascal Fallavollita, Reza Ghotbi, Nassir Navab and Stefanie Demirci. "Optimal c-arm positioning for aortic interventions". *Proceedings of Workshop Bildverarbeitung für die Medizin (BVM), 2015*.

[1] Part of this work has been developed during the author's Master studies. Findings to this work are here presented and create the foundations on which this dissertation is based.

Part I

Introduction

1 Introduction

Ultrasound imaging has been one of the most used imaging modalities from its introduction in the 1940s, and since then, its development has been rapid and continuous. Its success is mainly due to its advantages over other imaging modalities such as X-ray, Computer Tomography (CT), or Magnetic Resonance Imaging (MRI). Ultrasound imaging equipment, in fact, does not produce any ionizing radiation nor requires a nephrotoxic contrast agent during its use, has a small form factor that allows for its easy transportation and is available at an affordable price compared to the aforementioned modalities. More importantly, ultrasonography produces real-time dynamic images.

However, considering all the mentioned modalities, one can immediately capture the most prominent characteristic that differentiate ultrasound imaging from the others: **it always requires an operator**. In our collective mind, we all have a clear image of a large CT or MRI machine sitting in an otherwise empty room. We are used to be placed on an examination bed and let alone with that machine, and we are confident that the machine itself will be able to produce an image of our anatomy. Operators, in this context, are limited to the simple role of a technician that has to initialize a well-defined protocol from a separated control room. In the case of ultrasound imaging, the practice is very different.

Ultrasound probes are currently designed as hand-held devices, and they require the knowledge and experience of a trained human to maneuver them. Two key elements enable a person to correctly acquire ultrasound images: **vision** and **haptics**. The former, comes in two different forms, on the one hand, the operator uses his visual senses to move the probe along the patient body, targeting specific areas; on the other, she/he receives the visual feedback from the live ultrasound stream and adjusts the probe position based on it. The latter comes in the form of manual palpation and the application of a certain pressure to the probe and onto the patient's body. These characteristics are fundamental to achieve a successful and clinically meaningful ultrasound acquisition, characteristics that are complex to transfer onto a machine.

Robotic systems (i.e., intelligent mechanical structures) have been successfully developed and employed to target complex tasks in various fields, and have also being explored in the field of ultrasound imaging (See Sec. 1.3). This dissertation examines the author's contribution to the field of robotic ultrasound imaging, with a strong focus on providing and utilizing at best vision and force sensing to enable fully autonomous systems.

1.1 Motivation and Main Objective

The development of a robotic system for ultrasound imaging has to have clinical motivations other than the pure engineering interest to solve the task. As already mentioned, other imaging modalities only require the presence of a member of the technical staff, rarely the availability of a radiologist is required. Ultrasound acquisitions, instead, are always performed by a trained physician or by a sonographer. In both cases, the individuals are highly trained, and their experience is fundamental to the clinical outcome of the procedure. As shown by many works in the clinical literature [10, 11, 12, 13], acquisitions by inexperience operators result in less accurate evaluation of the images. That is, missing to capture crucial information from the acquired ultrasound frames, might hinder the overall diagnosis or treatment. This introduces a high variability of the performed acquisitions; a patient might need to undergo further examinations or receive a dangerous late diagnosis. In an interventional setting, it would be often required the presence of an experienced sonographer in the operating room, adding costs to the operation and one more person to an already typically crowded scene. Drawbacks of manual ultrasound acquisitions are not only on the patient side. Over the past decades, it has been unquestionably proven by many analyses that sonographers are often subject of musculoskeletal disorders due to their posture and the repetitive motions during ultrasound acquisitions [14, 15, 16, 17, 18]. In general, it is a manual task that is hard to do well, it is hard to repeat consistently and causes stress to the operators' bodies.

Robotic systems are, overall, optimal solutions for repetitive manual tasks, but in the case of ultrasound acquisition, they could additionally provide a platform to standardize results and reduce operator dependence. However, the complexity of the task requires the availability of crucial information to perform it. To achieve a completely autonomous robotic system for ultrasound acquisitions, a careful design process has to be carried out. This includes the integration of multiple sensing modalities and their coherent use, as well as the implementation of an entirely new workflow. As already mentioned, vision and haptics are the two fundamental input signals for the task. The objective of this thesis is to explore how the two modalities, in the form of ultrasound image feedback, force sensing, and 3D vision, can be effectively employed in a clinically relevant autonomous system for medical ultrasound. The main interest is to demonstrate how these modalities can take advantage of each other strengths to optimize their output and the overall clinical outcome.

1.2 Outline

This thesis is subdivided into three parts. The current one, **Part I**, continues with an overview of the state-of-the-art in the field of robotic ultrasound imaging, with a focus on autonomous systems and their current applications.

In **Part II**, the essential methodology used throughout the work reported in this dissertation is formally introduced, and the specific contributions of this thesis are presented, in details:

- **Chapter 2** introduces fundamental concepts of ultrasound imaging. The main focus lies on the description of ultrasound B-mode image formation, the creation of ultrasound 3D

volumes, and the use of such data with other imaging modalities through image-based registration.

- **Chapter 3** describes the control strategies employed to achieve safe and accurate motions of a robotic ultrasound system. As ultrasound imaging requires contact with the patient's body, modern force control techniques used along this dissertation are presented.

- **Chapter 4** describes the 3D vision technology employed to model the environment with whom a robotic ultrasound system has to interact. Techniques for the alignment of 3D point sets are also presented.

- **Chapter 5** presents a the main ideas, methods, and experiments, developed by the author to achieve accurate, safe and, reproducible robotic ultrasound acquisition in full autonomy. This section is based on the author's original publications.

Part III contains an analysis of the presented contributions:

- In **Chapter 6**, a summary of the findings and improvements concerning the state-of-the-art is presented.

- In **Chapter 7**, an ultimate analysis of the benefits and limitations of the proposed methods is presented here. Further possible directions to improve the current state and enable the use of robotic ultrasound systems in clinical practice are also introduced.

Finally, **Appendix A** presents the abstracts of other publications that are not included in this thesis.

This dissertation is publication-based, text and figures are extracted from the corresponding publications. Note that [7] and [8] partially contain work developed during the author's Master studies, but peer-reviewed, refined and published during his Doctoral studies. The presented findings present the foundations to this dissertation. Further Doctoral work is presented in [4] and [1], composing the complete author's contribution to the field. While all this work was made possible by great contributions of fellow colleagues and co-authors, the personal contribution of this thesis' author is stated for each publication presented in **Chapter 5**.

1.3 Background and Related Work

Robotics has the potential to transform many aspects of the current medical practice. Many fields could benefit from the integration of robotic solutions, from hospital automation to rehabilitation and surgical intervention, so far, the most significant impact of robotics in medicine has undoubtedly been in soft tissue surgery. Based on research from the beginning of the 1990s, taking advantage of concepts like the remote center of motion [19], commercial systems have been able to perform robotic-assisted surgery for a various number of procedures [20, 21], with a focus on urology [22, 23, 24]. Research related to robotic systems for surgery

is nowadays still very active and broad [25]. Robotics is having an impact in other fields such as orthopedics [26, 27], neurosurgery [28], steerable catheters [29, 30], and radiosurgery [31]. For the interest of this dissertation, the focus is on the past and present state-of-the-art research in the field of robotic imaging, and specifically robotic ultrasound. An interested reader can find more information on the overall field of medical robotics in the available literature reviews [32, 33, 34, 35].

The research interest in robotic ultrasound aims at solving some of the limitations of traditional ultrasound imaging presented already in Section 1.1, i.e., reduce or remove the required manual interaction by an expert sonographer. Just like in the case of robotic surgery, one fundamental aspect has been the development of teleoperated systems to allow for ultrasound-based diagnostics in remote areas, driven especially from research in the military. The first concept for a manual tele-ultrasound system has been presented in [36], with a focus on the requirements for image compression and transmission to achieve real-time capabilities.

The conception of robotic ultrasound imaging has to be attributed to a fundamental series of research by Salcudean et al., stated in 1999 [37]. A first mechanical design was proposed to relieve sonographers from the musculoskeletal fatigue of manual ultrasound procedures, taking as first exemplary application the examination of the carotid artery. It is remarkable that already in such an early stage of the research field, the authors had clear what the requirements for such a system should be: able to ensure patient safety at all times, light, limited in the amount of force it could exert onto the patient, with backdrivable joints so that it could be pushed away, and able to cover the range of motions and forces required by the task. The system was then showcased for teleoperation in [38], using image correlation techniques to maintain feature visibility in [39]. The used shared-control approach was then formalized in [40]. Additionally, the authors present some early work on visual servoing control techniques based on the detection and tracking of anatomical structures in ultrasound [41], a field of research more and more explored in the following years. In [42], the complete system usability is demonstrated by employing a shared control schema between the operator, robot controller, and ultrasound image servoing. 3D ultrasound reconstruction and robot control over the internet are also showcased. In the same years, a parallel effort in the development of the first prototypes of mechanical devices for teleoperated ultrasound was also taking place in Japan [43]. Common to all the early-stage work is the development of custom robotic systems, due to the need to fit the essential requirements needed in medical applications. Industrial robots, in fact, did not provide yet the level of safety required for direct contact with the patient and medical staff. Fore control techniques were still limited to the research communities, and sensing capabilities on board of the available industrial manipulators were inadequate. The details provided by [38] influenced the research on robotic platforms for ultrasound imaging that followed, following also the research path for the modernization of industrial manipulators that have been more and more modified also to fit these needs.

The development of systems for teleoperation of ultrasound devices continued with the efforts by European institutions in a series of projects: TeleInVivo [44], TER [45, 46, 47], OTELO [48, 49, 50] and the TERESA project [51]. More recently, [52] focused on the design of an IMU tracked controller that mimics the shape and size of a standard ultrasound probe, to control a remote robotic ultrasound system, while [53] transitioned to the use of an off-the-shelf commercial robotic manipulator for tele-ultrasound. These projects, while proving the growing

interest of the research community for the topic, were still focused on the mechanical design, control characterization, and signal transmission and compression to enable a functional teleoperated system. The integration between multiple sensorial information was mostly limited to ultrasound imaging processing, and the concept of full autonomy was not yet addressed.

Once the basis for such systems has been established, a large wave of research was directed to close the loop between the acquisition of ultrasound images and the robot control. Visual servoing techniques were established for the task, aiming at reaching the desired ultrasound view of the target anatomy [54]. Multiple techniques for the extraction of features in the ultrasound images have been presented: speckle tracking has been employed to compensate for the motion of diagnostic targets [55] and to stabilize the robot motion in real-time [56]. To maintain the visibility of target objects image moments [57], plain intensity-based registration [58], and template-matching algorithms [59] have been explored. Visual servoing techniques have been employed to track not only anatomical targets but also instruments (e.g., needle, catheters, etc.) inserted into the patient's body during an intervention. The use of a robotic system allows, in these cases, to achieve an autonomous tracking of the instrument and maintains its visibility during its insertion. The real-time detection and tracking of rigid and flexible needles have been proposed [60, 61] using small 3D ultrasound volume acquired at a fast rate and a combination of particle and Kalman filters to obtain a robust tracking. A two robot system is guided in [62], with one manipulator acquiring ultrasound images while maintaining the visibility of a needle inserted by the other one, in a fully calibrated system. The tracking of a catheter equipped with an active piezo-element is achieved in [63]. The pulse emitted by the element is captured by the ultrasound transducer equipped onto a manipulator, and that information used to guide the probe position while the catheter has been inserted inside a vascular tree. Another use of visual servoing methods has been targeted at the optimization of the acquired ultrasound image quality, a series of work from Chatelain et al. relies on an ultrasound image quality metric called *confidence maps* and aims at adjusting the transducer pose in order to maximize image quality, avoiding loss of contact with the patient skin [64, 65, 66].

The use of external force sensing has also been leveraged to improve the usability of robotic systems, [67] and [68] explored the use of a collaborative force control approach to limit the force that an operator has to apply during an ultrasound acquisition manually.

Following the release of modern robotic manipulators that enable a safer human-robot interaction, the research in the field has gradually shifted towards the integration of such systems into the clinical routine, rather than focusing on new mechanical designs. Some more recent work on the design of robotic ultrasound systems focus on new kinematic designs adapted to satisfy the needs for a specific clinical application or on the use of recently developed materials. For instance, adopting a specific design for the acquisition of fetal ultrasound using one [69] or two robotic arms [70]. Modern soft robotics, instead, is employed in [71, 72] to achieve safe ultrasound acquisitions on pregnant women by leveraging a novel end-effector design to locally steer the ultrasound probe after it has been placed onto the target area.

Specific clinical applications for robotic ultrasound imaging have also been explored in the literature, in [73, 74] a robotic system for the online monitoring and adjustment of radiation

therapy plans has been presented, making use of ultrasound imaging to track shifts in the patient position and allow to minimize incorrect radiation exposures. A robot for ultrasound guidance during transrectal interventions has been proposed in [75]. Neurosurgical navigation for spinal procedures is enhanced using ultrasound-based robotic guidance, and image-based registration of subsequential acquisitions is explored in [76], the groundwork for the first patient trials performed in [3].

Finally, the interest for robotic ultrasound systems has nowadays permeated also the industry, with tele-echography commercial systems being developed in Europe [77, 78] and China [79]. Research in the field also continues to be supported by European initiatives, such as the MURAB[1] H2020 project [80], targeted at ultrasound-based robotic-guided breast biopsies.

Reviews on the state-of-the-art for robotic ultrasound systems have also been produced in literature over the years. The interested reader can refer to [81, 82, 83, 84, 85].

[1] https://www.murabproject.eu/

Part II

Methodology and Contributions

2 Ultrasound Imaging

The origin of ultrasound research can be traced back to the year 1794, when a fellow Italian scientist, Lazzaro Spallanzani, discovered echolocation in nature. Spallanzani was a biologist, physiologist, and Catholic priest that, as described in [86], performed a long series of experiments on bats with the intent of understanding their ability to navigate in the absence of light. To his surprise, Spallanzani was able to associate this trait to the bats' use of their hearing. He observed that bats were able to avoid obstacles while flying when blinded, only to lose this ability when their ears would be obstructed. He did not believe initially that hearing could be involved, since bats fly very silently, but its extensive and rigorous experiments convinced him; an example that even a man of the Catholic Church has to surrender to the scientific method. The exact scientific principle was to be demonstrated only in 1944 by two American biologists Donald Griffin and Robert Balambos, when the term echolocation was coined. The reason for Spallanzani's surprise was then clear; animal echolocation is based on sound waves at frequencies up to $200\,\mathrm{kHz}$, far above the human hearing limit of $20\,\mathrm{kHz}$. In the same decade, taking inspiration from these observed phenomena in nature, sound waves were becoming a fundamental tool for the navigation of human-made machines with the introduction of the *sonar* (sound navigation and raging).

The use of ultrasonic waves as a diagnostic tool has become a reality since 1942, when Austrian neurologist Karl Theo Dussik attempted to detect brain tumors transmitting an ultrasound beam through a human skull. A first commercial medical device would be released in 1963 by Meyerdirk & Wright, making ultrasound imaging generally available for clinical use. Since then, medical ultrasound imaging has become one of the most available and used imaging modalities, with a plethora of clinical applications. An in-depth presentation of the historical steps that brought ultrasound imaging to its current state is available in [87].

2.1 Ultrasound Physics

Ultrasound Wave Properties

Acoustic waves propagating through a medium can be distinguished into two types: **longitudinal** and **transversal** waves. The latter, also called shear-waves, are less frequently used in ultrasound imaging, although they are gaining more importance as their use is introduced in advance modalities, such as ultrasound elastography [88].

Longitudinal waves can be identified by a back-and-forth particle motion that is parallel to the wave direction. As already mentioned, ultrasound waves exist above the human perceptible limit of $20\,\mathrm{kHz}$. For medical use, ultrasound frequencies are commonly between 1 and $20\,\mathrm{MHz}$, depending on the particular application. Materials characterized by the piezoelectric effect

can produce ultrasound waves. Certain materials, in fact, respond to a mechanical force by generating an electric charge in response, a phenomenon depicted in Fig. 2.1. At the same time, a mechanical deformation can be induced by applying an electric field to the same material. For medical imaging devices, lead zirconate titanate is commonly used as a piezoelectric element. Since a singular piezoelectric element produces a limited amount of energy, ultrasound transducers are composed of a series of elements stacked into layers. That is, an ultrasound transducer is then able to transform electric energy into mechanical pulses and vice versa.

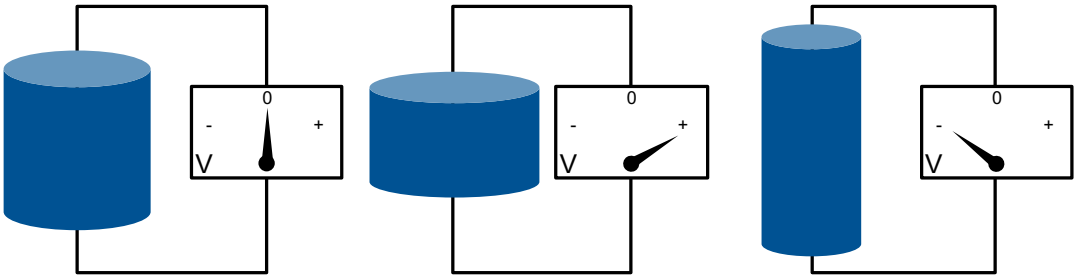

Fig. 2.1. Visual representation of the **piezoelectric effect**, meterials characterized by this effect produce an electric charge in response to a mechanic stress. In this example, a disc produces a certain voltage due to the induced deformations (exaggerated for the visualization).

We can define as *period* the time for an ultrasound wave to conclude one cycle, while its *wavelength* is the space traveled during that time. The wavelength λ can be put in relation to the wave frequency f, i.e., the number of cycles repeated in a second, as

$$\lambda = \frac{c}{f}. \tag{2.1}$$

Where c is the *acoustic velocity*, i.e., the speed at which waves travel through a medium, which depends on the material density and stiffness. In anatomical tissue this velocity ranges between 1400 and $1640\,\mathrm{m/s}$, although in most modern systems this values is approximated to the constant value of $c = 1540\,\mathrm{m/s}$.

Ultrasound waves produced by piezoelectric elements exhibit a *self-focusing* property, which concerns the narrowing of the produced beam at a certain distance. The area near the element where this phenomenon takes place is called the *near field* or *Fresnel zone*, contrary to the *far field* or *Fraunhofer zone* where the beam, instead, diverges. In the near field, the beam diameter is constant, and the field size depends on the signal wavelength and the crystal size d. Thus the near field length can be computed as:

$$\text{Near field length} = \frac{d^2}{4\lambda}. \tag{2.2}$$

The length defines the range best suited for imaging, as focus quality and thus resolution decrease after that. Modern ultrasound transducers, which are composed of multiple piezoelectric elements, can achieve a dynamic focus by introducing a time delay between electrical pulses on the individual elements. This can be modified electronically to achieve an optimal

focusing at multiple depths within the human body. Modern systems also employ a technique called *apodization* to weight the elements signal amplitude across the whole transducer.

Spatial resolution in ultrasound imaging, i.e., the ability to distinguish between two points at a defined depth in tissue, is characterized by *axial* and *lateral* resolution and strongly coupled to the signal wavelength.

The axial resolution, R_a, is the minimum distance that can be distinguished between two reflectors along the scan line direction. It is equal to half the overall pulse length L_p, which, in turn, is the product of the number of excitation pulses ($n \in \{2, 3\}$ in practice) and the wavelength:

$$R_a = \frac{L_p}{2} = \frac{n\lambda}{2} \tag{2.3}$$

The lateral resolution, i.e., across the imaging plane, is primarily determined by the beam width of the ultrasound wave. Intuitively, the lateral resolution is high when the beam's width is narrow, and it is therefore influenced by the aforementioned concepts of near and far-field, and focusing.

Tissue Interaction

A fundamental concept in ultrasound imaging is *acoustic impedance*, a physical property of tissue that describes the resistance an ultrasound wave encounters as it passes through it. It is based on the product of the tissue density ρ and the speed of sound in the medium,

$$Z = \rho c. \tag{2.4}$$

Intuitively, impedance increases as the tissue density does so. Examples of impedance values for bodily tissues are presented in Table 2.1. The effect of acoustic impedance in medical ultrasound becomes visible at interfaces between tissue types with different impedance. A significant difference in tissue impedance would resolve in the reflection of the ultrasound wave. Considering a wave passing orthogonally through two tissue layers, the amount of reflection can be expressed as

$$R = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \tag{2.5}$$

where Z_1 and Z_2 are the impedance values for the first and second tissue, respectively.

Since not all the ultrasound wave would be reflected, the remaining intensity passed to the second tissue layer can be represented with the transmission factor

$$T = \frac{2Z_2}{Z_2 + Z_1}. \tag{2.6}$$

Material Type	ρ [kg/m^3]	c [m/s]	Z [MegaRayls]
Air	0.0012	330	0.0004
Fat	928	1,430	1.327
Water (20°C)	998	1,481	1.482
Brain	1,035	1,562	1.617
Kidney	1,050	1,560	1.638
Muscle	1,041	1,580	1.645
Heart	1,060	1,554	1.647
Liver	1,050	1,578	1.657
Bone	1,990	3,198	6.364

Tab. 2.1. Density (ρ), speed of sound (c) and impedance (Z) values for various anatomical tissue types, from [89].

From the data available in Table 2.1, one can observe that, as an example, only less than 1% of an ultrasound wave would be reflected at an interface between fat and liver tissue, while most of it would be lost passing through lung tissue, due to the presence of air.

In reality, most of the acoustic waves do not hit tissue interfaces orthogonally. Part of the ultrasound is, in fact, reflected according to the Snell's law. Similarly to optical reflection, the reflection angle θ_r is equal to the incidence angle θ_i, the transmission angle θ_t can be obtained from the ratio between the tissue speeds of sound:

$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{c_1}{c_2}. \tag{2.7}$$

Leading to a general formulation of acoustic impedance and transmission:

$$R = \frac{Z_2\cos\theta_i - Z_1\cos\theta_t}{Z_2\cos\theta_i + Z_1\cos\theta_t},$$
$$T = \frac{2Z_2\cos\theta_i}{Z_2\cos\theta_i + Z_1\cos\theta_t}.$$

A directional change of the ultrasound wave across tissue interfaces, i.e., $\theta_i \neq \theta_t \neq 0$, is refereed as refraction.

The interaction with objects that are of a smaller size than the signal wavelength, instead, generates diffuse reflections, a phenomenon called *scattering*. The reflections themselves would then cause further disturbances by expanding over all directions, causing the familiar speckle noise that can be observed in ultrasound images. Finally, part of the wave energy is

dissipated in the form of heat by *absorption*, which, again, is highly correlated to the medium that is traversed.

2.2 2D Image Formation

Ultrasound Image Modes

In medical imaging, different ultrasound modes are used for various applications, mainly: A- (amplitude) mode is mostly used to differentiate tissue types [90, 91]. It represents the echoes produced by a single pulse over time, in a one-dimensional diagram. Doppler mode makes use of the Doppler effect to estimate the blood flow through blood vessels [92]. M- (motion) can visualize the movement of body structures, such as hearth valves [93], in a wave-like shape.

For the scope of this thesis, instead, the focus is only on B-mode ultrasound, which is the most used ultrasound mode to acquire general anatomical information. B-mode (brightness mode), in fact, produces a 2D grayscale image of a body area emitting ultrasound waves simultaneously from an array of piezoelectric elements. The intensity of the grayscale values within an image represents the amplitude of the received echoes. It is used for a large variety of clinical applications, both in diagnostics [87] and interventional [94] scenarios, and it is also the only ultrasound image modality used throughout all the contributions presented in this work.

Ultrasound Transducers

As already mentioned, ultrasound imaging makes use of arrays of piezoelectric elements, rather than individual ones. Ultrasound transducers come in different shapes, sizes, and features, depending on the clinical application they are targeting. The characteristics that can help distinguish different transducers are mainly the arrangement of the piezoelectric elements and their frequency range. The most common types of ultrasound transducers are shown in Fig. 2.2:

- **Linear transducers**, in which the piezoelectric elements are arranged linearly, resulting in a rectangular shape of the generated ultrasound beam. These probes usually have a good near-field resolution, and their frequency ranges between $2.5\,\mathrm{MHz}$ and $12\,\mathrm{MHz}$. Common clinical applications that use linear transducers are thyroid, breast, and tendon imaging.

- **Curvilinear transducers** are characterized by having their elements aligned along a circular arc of a fixed angle. The produced beam is convex and ideal for in-depth examinations, which makes this the probe of choice for abdominal imaging. Their frequency ranges between $2.5\,\mathrm{MHz}$ and $7.5\,\mathrm{MHz}$.

Other, more specialized, transducers are also available for intra-corporeal ultrasound imaging, e.g., *trans-rectal*, *trans-vaginal* or *intra-vascular* probes, or with more specialized elements, e.g. *phased array* transducers. Overall, the contributions presented in this work make use of

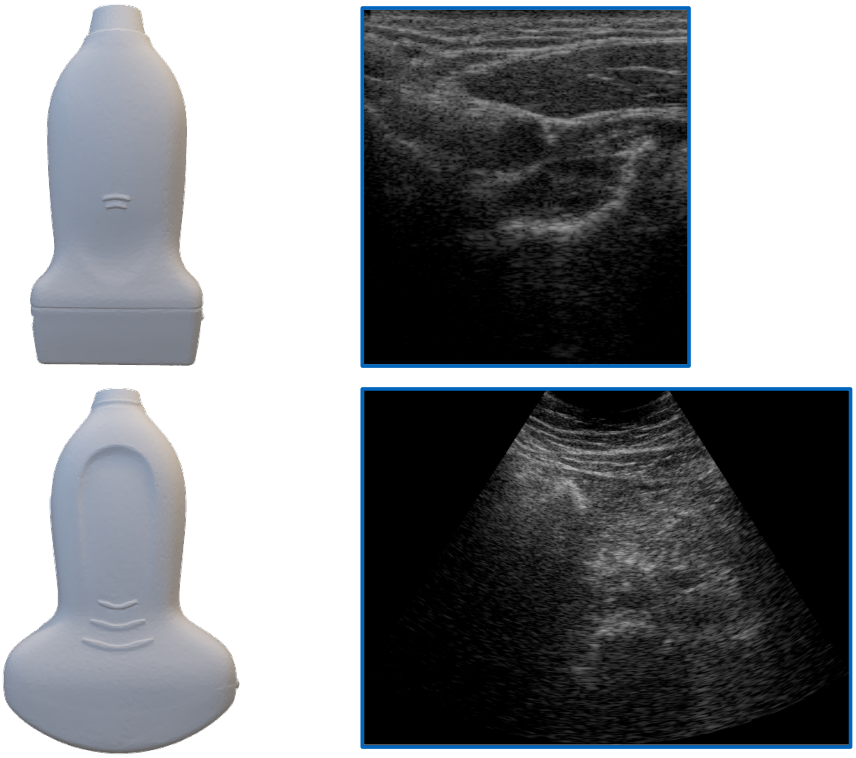

Fig. 2.2. Representation of common ultrasound transducers and examples of ultrasound images obtain using them. **Top row:** A linear transducer. **Bottom row:** a curvilinear transducer. Notice how the acquired images reflect the geometry of the probes' piezoelectric elements.

only linear or curvilinear transducers. Their main characteristics, such as footprint and used frequency, are reported in the respective publication.

B-mode Image Formation

The self-focusing characteristics of piezoelectric elements have already been mentioned. Modern systems make use of a time-delay method to dynamically adjust the focus on the ultrasound beam and optimize the lateral resolution. An extensive discussion about this process, defined as *beamforming*, and more advanced techniques can be found in [87].

After the transmission of the ultrasound pulses, the same elements in the transducer are used to listen to the returning waves, which are affected by the interactions described in 2.1. The readings from multiple elements are then aligned and summed for additional processing.

Given $\tilde{x}(t)$, the acquired raw signal, two scaling factors are applied at the first processing step:

$$x(t) = F_0 \, F_1(t) \, \tilde{x}(t). \tag{2.8}$$

F_0 represents a global gain factor, while $F_1(t)$ is known at *time-gain-compensation* (TGC) and includes multiple scaling factors that are based on the time of arrival of the returning wave to the elements, effectively depending on the depth level to which the produced ultrasound wave initially traveled. In practice, TGC adjustment is available on most ultrasound machines to dynamically reduce the signal loss at greater depths. The instantaneous amplitude of the ultrasound signal is then obtained by *envelope detection*. Considering the Hilbert transform

$$\mathcal{H}[x(t)] = \frac{p}{\pi} \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\tau} d\tau, \tag{2.9}$$

where p is the Cauchy principal value. We can represent $x(t)$ in complex notation as:

$$z(t) = x(t) + i\mathcal{H}[x(t)] = a(t)e^{i\phi(t)}. \tag{2.10}$$

Where $a(t) = \sqrt{[x(t)]^2 + [\mathcal{H}[x(t)]]^2}$ and $\phi(t) = \arctan[\frac{\mathcal{H}[x(t)]}{x(t)}]$ are, respectively, the instantaneous amplitude and phase of the signal. The time derivative of $\phi(t)$

$$\omega(t) = \frac{d\phi(t)}{dt}, \tag{2.11}$$

is instead the instantaneous frequency of $x(t)$. After applying the Hilbert transform, the real party of $z(t)$ represents the original signal:

$$x(t) = a(t)e^{i\int \omega(t)dt}. \tag{2.12}$$

At this stage, the signal is commonly further enhanced, using more advanced filters to reduce noise. The variation in amplitude of the ultrasound signal can be relatively high, causing few high amplitude values to suppress the others. Therefore, the signal cannot be linearly mapped to the gray-scale values. A logarithmic compression is rather used to adjust dynamic range. Considering a gray-scale image with N bit color depth, image intensities are obtained from

$$I(t) = \frac{\ln x(t) - \ln min_t x(t)}{\ln max_t x(t) - \ln min_t x(t)} (2^N - 1), \tag{2.13}$$

with the dynamic range defined by the denominator

$$R = \ln \frac{max_t x(t)}{min_t x(t)}. \tag{2.14}$$

The obtained intensities are composed in the final two-dimensional image. While this is more trivial for linear transducers, due to their simple geometry; curvilinear transducers

require a *scan conversion* step, to interpolate the resulting intensities values in a rectified image according to the specific probe geometry.

2.3 3D Ultrasound Imaging

Although 2D images are the most accepted form to display and analyze ultrasound images, their two-dimensional nature has some limitations. When imaging a complete portion of the anatomy, e.g., an entire kidney or the liver, it is required for physicians to mentally map the acquired 2D images and create a spatial relationship between them. This task is not only time-consuming and mentally demanding but also introduces a high variability, purely depending on the operator skills.

Since the 1980s [95], 3D ultrasound imaging has been developed to address these issues and allow a tomographic reconstruction using ultrasound, similar to what is done with CT or MR imaging. In broad terms, the main idea is to reconstruct a volume from a set of 2D ultrasound images, acquiring knowledge regarding their spatial location. With respect to classic 2D images, 3D ultrasound can improve the accuracy and reduce the variability in the geometrical analysis of anatomical parts. 3D ultrasound acquisitions are also more repeatable, as acquiring an entire 3D volume allows to easily obtain the same information that has already been acquired in previous acquisitions. Over the past three decades, studies using 3D ultrasound have shown its benefits on a large set of clinical applications [95, 96, 97].

Ultrasound Transducer Tracking

3D ultrasound volumes are generated by interpolating 2D ultrasound images based on some knowledge of their spatial relationship. In this sense, several modalities can be used to track the position of the ultrasound transducer during 2D data acquisition.

The most common technique is the so-called *free-hand* ultrasound [98, 99], which is based on two main technologies that rely on the use of a *marker* attached to the ultrasound probe: *optical* or *electromagnetic* (EM) tracking, also illustrated in Fig. 2.3. An optical marker or EM sensor is attached and calibrated to the ultrasound transducer, while an external sensing system (an infrared camera or an EM field generator) can track their location over time. Optical tracking is quite precise, with a sub-millimeter tracking accuracy, but it commonly suffers from line of sight issues and a limited field of view. EM tracking is less accurate, it is limited to the active area of the EM field and, although it allows for obstructions between the generator and the marker, the readings can be disturbed by the presence of other ferromagnetic items in the vicinity. A variation of the optical tracking paradigm that uses an external camera is the *inside-out* optical tracking technique, which is characterized by the use of an optical sensor attached to the tracked object. In this case, information about the scene are captured by the sensor and used to re-localize the object with respect to it. In [2], the author of this thesis showcased this technology for 3D ultrasound compounding.

Employing free-hand ultrasound, since images are still obtained manually by a user, the acquisition usually presents a non-uniform spacing and might capture the anatomy only

Fig. 2.3. Tracking devices for free-hand 3D ultrasound. **Left:** an EM sensor is attached to an ultrasound linear transducer (white element inserted in the green casing), an EM field is emitted by the generator (cube on the top right) and the sensor position with respect to it obtained. **Right:** An optical marker is attached to the transducer, while an external camera emits infrared signals and captures their reflections from the marker's spheres.

sparsely. This disadvantage is due to the fact that it is hard for a human to optimize the acquisition, in terms of path, the pressure used, etc., using their sensing, and the images themselves are feedback. Another technique for the tracking of ultrasound transducers is attaching them to a mechanical device, e.g., linear stages, and perform their tracking using the information provided by the motors or encoders. *Mechanical* tracking can make use of more advanced mechatronic hardware to not only track the probe but guide it along the patient surface with great precision and repeatability. It is the case of robotic ultrasound systems that can use various sensing information to improve the produced ultrasound volumes and guarantee a constant optimal acquisition, as explored in this thesis. A major drawback of mechanical tracking solutions is their bulkiness and high cost, compared to free-hand techniques.

Exploratory techniques for ultrasound compounding also include *sensor-less* tracking, which attempts to regress the spatial relationship between 2D ultrasound frames using image-based [100, 101] or learning-based techniques [102]. Finally, pyramidal 3D volumes can be directly obtained using a 2D ultrasound array, although they are very expensive, have a small footprint, and their clinical use is currently quite limited.

Compounding Techniques

Upon acquisition of tracked 2D ultrasound data, a reconstruction technique has to be employed to obtain a complete 3D volume. Only in rare cases, trivial methods to stack the 2D images can be used, since they required smooth, uniform acquisitions along a straight line. In all other cases, a method to interpolate between the available images has to be integrated.

Compounding methods can be distinguished into two categories:

- **Forward** or **pixel** methods reconstruct a 3D volume visiting each pixel of the 2D ultrasound images. On a first step, the methods visit each pixel of all the 3D ultrasound images. Based on the available spatial information, the nearest 3D voxel of the recon-

structed volume is filled with the respective value. The values of multiple pixels that correspond to an individual voxel are normally averaged. Since this step can produce holes in the final volume, a hole-filling phase is then executed by using the average, median, maximum, or minimum value across voxel neighbors. While these methods can reconstruct the 3D volume rapidly, they tend to cause blurred results and might lose some information from the 2D images. Modern techniques attempt to recover some of these drawbacks, using a fast marching method [103] or Olympic operation [104] to interpolate the empty voxel accurately.

- **Backward** or **voxel** methods build a regular Cartesian grid from the acquired 2D ultrasound images. In this case, the voxel grid is traversed, and the nearest pixel value for the set of 2D images is inserted in the respective voxel. The voxel grayscale intensity value is computed as a weighted average of the neighbor pixel values. In this case, the original texture is preserved, holes avoided, and alternative views not available in the 2D set can be generated. However, if volume voxels are reconstructed using pixel values too distant from each other, large reconstruction artifacts can be introduced, a fine sampled volume is therefore required.

Extensive reviews of ultrasound reconstruction algorithms are also available in [99, 105, 106, 107].

2.4 Intensity-based registration

Image registration has, in general, the task of aligning two or more images, bringing them to a common reference frame, and enable to match anatomical information available in them. Medical image registration represents a very active field of research on its own [108, 109, 110]. For the scope of this thesis, the focus of this section is on the general framework to achieve intensity-based registration between ultrasound and CT or MRI volumes since it is one of the main building blocks leveraged by many of the contributions presented in this thesis.

Given two images I_f and I_m, intensity-based registration relies directly on the image intensities to align them. Fixing the image I_f, the aim is to find the optimal transformation T^* to apply to the moving image I_m:

$$T^* = \arg\max_T \alpha(I_f, T(I_m)). \tag{2.15}$$

Being $\alpha : (I_f, I_m) \rightarrow \mathcal{R}$ a *similarity metric* between the two images. It provides a measure of how well the two images are aligned.

This process, illustrated in Fig. 2.4, is an iterative optimization: the optimizer takes into account the value of α at each iteration to find the set of parameters for T^* that maximize the similarity between the two images. If a similarity gradient can be obtained, gradient-based methods such as gradient descent, or Gauss-Newton can be employed. Other optimization techniques include BOBYQA [111] and NEWOA [112], in the case the similarity is not

differentiable. The number of parameters to optimize depend on the transformation model. For rigid transformations, there are six degrees of freedom, while affine ones include three additional parameters for scaling. Non-linear deformation can also be address using techniques such as free-form deformations (FFD) [113], thin-plate splines [114] or energy-based methods [115].

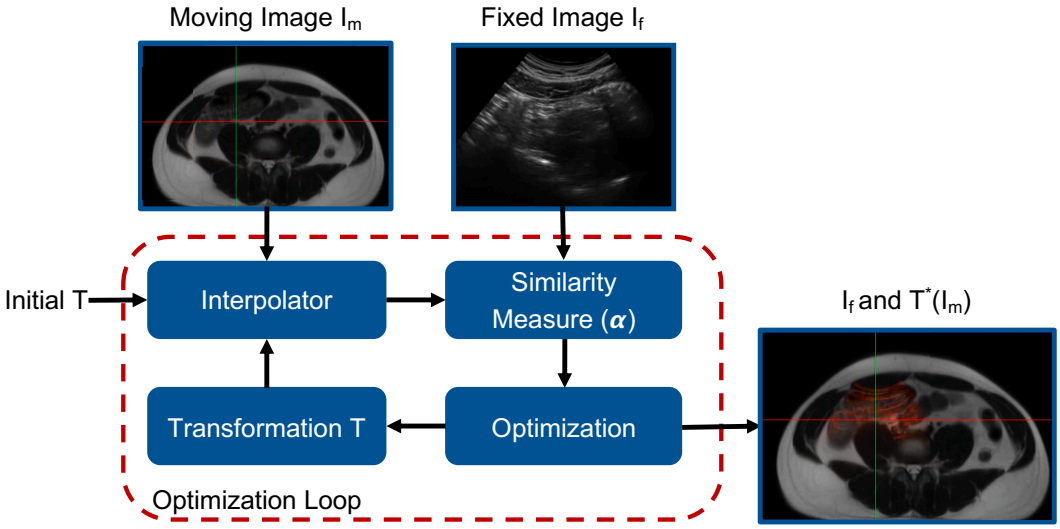

Fig. 2.4. **A generic framework for intensity-based registration methods.** An initial transformation T is applied to the moving image. A similarity metric is used to compute the similarity between the moving image and the fixed one. An optimization technique refines the transformation to maximize the similarity until convergence.

Similarity Metrics

A similarity metric $\alpha(I_f, I_m^{'})$ measures how well the transformed image $I_m^{'} = T(I_m)$ matches the fixed image I_f, with T the transformation obtained by the current iteration of the optimizer. The corresponding pixels or voxels from the overlapping area Ω of the two images are compared according to a chosen function f, and the result is integrated over the area:

$$\alpha(I_m, I_m^{'}) = \frac{1}{\Omega}\int_{\Omega} f(x)dx. \tag{2.16}$$

Classic chosen functions f are the **Sum of Squared Differences** (SSD),

$$SSD(x) = (I_m(x) - I_m^{'}(x))^2, \tag{2.17}$$

or the Sum of Absolute Differences (SAD),

$$SAD(x) = \left|I_m(x) - I_m^{'}(x)\right|. \tag{2.18}$$

For both these metrics, however, holds the assumption that both images include intensity values in the same range, which is often not true even for images of the same modality. To overcome the different illumination level of the two images, **Normalized Cross-Correlation** (NCC) is commonly used:

$$NCC(x) = \frac{(I_m(x) - \mu_f)(I_m^{'}(x) - \mu_m)}{\sigma_f \sigma_m}, \tag{2.19}$$

with μ and σ, the mean and standard deviation of the image intensities, respectively. NCC is a valid metric for intra-modal registration, i.e., registration of two images acquired with the same modality, since it assumes the intensities of both images to have a linear correlation. In fact, in the case of multi-modal registration, such as ultrasound and MRI, bright areas in one modality are not necessarily represented with the same intensity range in the other. To handle this non-linear correlation, various similarity metrics have been proposed, such as **Mutual Information** (MI) [116], based on the Shannon entropy formulation from information theory

$$H(I) = -\sum_i p_i \log p_i. \tag{2.20}$$

$p_i(I)$ is the probability distribution of an image pixel having intensity i. For two images, instead, the joint probability distribution can be taken into account, and the joint entropy

$$H(I_m, I_m^{'}) = -\sum_i \sum_j p_{I_m, I_m'}(i,j) \log p_{I_m, I_m'(i,j)} \tag{2.21}$$

has to be minimized to achieve an optimal alignment of the two images. For robustness, MI is commonly formulated as

$$MI(I_m, I_m^{'}) = H(I_m) + H(I_m^{'}) - H(I_m, I_m^{'}) = \sum_i \sum_j p_{I_m, I_m'}(i,j) \log \frac{p_{I_m, I_m'(i,j)}}{p_{I_m}(i) p_{I_m'}(j)}, \tag{2.22}$$

and maximized instead. MI has been proven a powerful metric for multi-modal registration but has not been very successful when used for the registration of ultrasound imaging with other tomographic modalities.

Defined specifically for registration between ultrasound and MRI acquisition, the **Linear Correlation of Linear Combination** (LC^2) metric [117] has been employed in many of the contributions presented in Chapter 5 of this thesis, such as [7] and [1]. This similarity method

correlates the MRI intensities p and gradient magnitudes $g = |\nabla p|$ to the ultrasound intensity values. For each pixel x_i in an ultrasound image, the intensity function $f(x_i)$ is defined as

$$f(x_i) = \alpha p_i + \beta g_i + \gamma, \tag{2.23}$$

where $y_i = \alpha, \beta, \gamma$ are the unknown parameters that represent the influence of p_i and g_i in a neighborhood of x_i. These coefficients can be approximated by minimizing the difference between y_i and the ultrasound intensities u_i:

$$\left\| M \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} - \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} \right\|^2 \text{ and } M = \begin{pmatrix} p_1 & g_1 & 1 \\ \vdots & \vdots & \vdots \\ p_m & g_m & 1 \end{pmatrix}, \tag{2.24}$$

solving for y_i using ordinary least squares with the pseudo-inverse of M. The final local similarity is then

$$LC^2(I_{US}, M) = 1 - \frac{\sum_{x_i} |I_{US}(x_i) - My|^2}{\sum_{x_i} Var(I_{US}(x_i))}. \tag{2.25}$$

In [118], the LC^2 similarity metric has been shown applicable for rigid, affine, and deformable registration, delivering high robustness and a wide convergence range.

3 Force Control for Robot Manipulators

The ability to physically interact with the surrounding environment is fundamental for robotic systems in many applications. In the medical domain and in particular for robotic ultrasound applications, this is even of greater importance. As seen in Chapter 2, contact with the patient is always required to acquire ultrasound images in real-time. In the case of robotic actuated ultrasound transducers, it is crucial for the system to be able to measure the force that has been applied to the examination area and adapt it as needed.

Pure motion control for robotic manipulators is not adequate in these situations since errors in the control modeling could create severe damages to the patient; force sensing and force control are therefore mandatory. This chapter explores the main techniques that allow robotic manipulators, like the one employed during the development of this thesis, to interact with patients' anatomies safely.

3.1 Fundamental Concepts

Research focusing on the modeling of the interaction with the environment by robotic systems has been very active in the past decades. As discussed in Chapter 4, a common strategy is to employ visual sensing to map the robot's surroundings. However, many applications require finer information on the interaction with objects, especially in telemanipulation, for which force sensing was introduced to help the remote handling of items without damaging them.

A pure position control strategy would result in unwanted force being applied to the robot parts and the contact object in the case of constrained motions since the controller would only attempt to adjust its end-effector position by applying more torque to the motors. Ensuring a *compliant behavior* of the system can reduce this effect.

Passive Compliance

Passive compliance is obtained by limiting the contact forces by means of mechanical compliance of the robot components. A typical solution is employing a *remote center of compliance* (RCC), which is a compliant end-effector mounted onto a robot's rigid body [119]. This approach doesn't use external sensing, so it is faster and less complex than adapting the control algorithm in real-time using feedback from sensors. However, it requires the design of a custom mechanical component that adapts precisely to the task at hand, which makes the concept hard to generalize for tasks where the environment response is not known a priori, e.g., patients' bodies have a large variability.

Active Compliance

Active compliance involves the use of contact forces and torques measurements as feedback signals for the control schema. They are used to dynamically adapt the generated robot trajectory according to the interaction with the environment. Since this approach requires more complex and, therefore, computationally demanding control systems, it is usually paired with some extend of passive compliance, such that damage from the applied forces can be avoided while they are not yet compensated for in the control loop.

Force Sensing

Six components (three translational force elements and three rotational torque ones) are required to obtain optimal contact information. One strategy is to attach a force/torque sensor between the robot end-effector and the tool, simply measuring the contact forces applied by the equipped tool [121, 122]. External torques can also be estimated via joint torque sensors [120, 123, 124], which provide not only information about the force applied by any tool, but also the ones obtained by the contact of other robot links with external objects. The last one is the typical approach for modern robotic manipulators that are certified for human interaction. This technology, in fact, allows to also react to collision with humans that might be operating in the vicinity of the robot and collide with its body. One drawback of this approach is that the computation of forces applied to the tool is limited by the resolution of the joint sensors, which are propagating their errors; an additional force sensor is often added to the robot wrist to combine the two approaches. Both force sensing approaches are displayed in Fig. 3.1 on the robotic manipulator employed throughout all the contributions presented in this thesis.

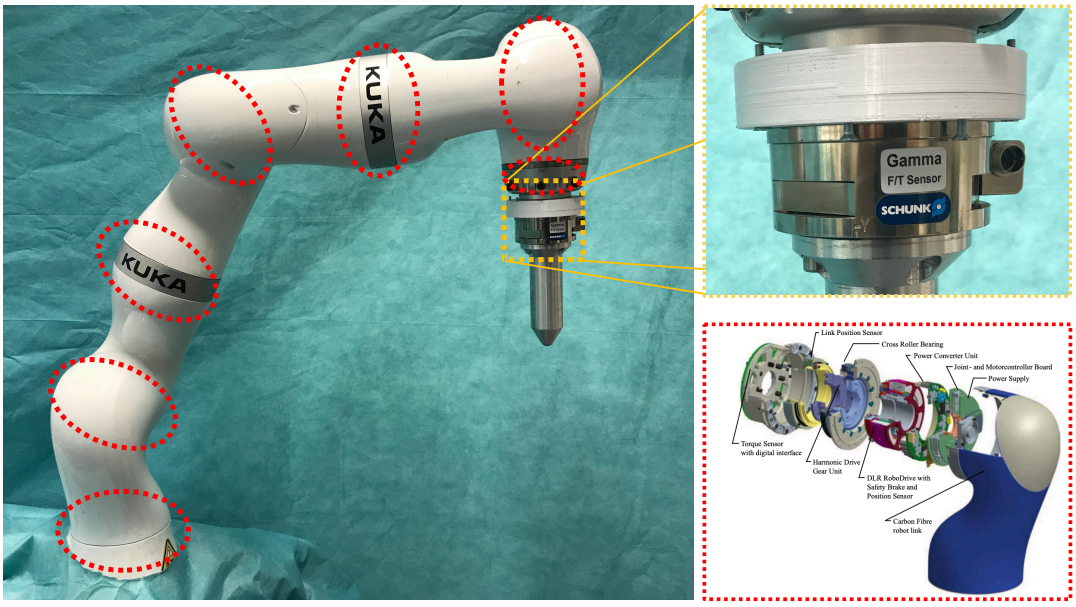

Fig. 3.1. **Hardware for force sensing with an industrial manipulator. Left**: a robotic arm equipped with a 6D force/torque sensor mounted at its flange (marked in yellow) and joint torque sensors for each of its seven joints (marked in red). **Top right:** a particular of the force/torque sensor at the robot's end-effector. **Bottom right:** the design of the robot joints including the torque sensor, reprinted from [120] with kind permission from Emerald Publishing Limited.

3.2 Force Control Strategies

Without aiming at completeness, the main control strategies for systems with active compliance are now explored. The reader should refer to the fundamental literature in the field for a complete and formal introduction to the topic [125, 126].

The two main categories for control schemes are *indirect* and *direct* force control. Strategies in the first category do not directly include force feedback into the control loop, making only use of motion control techniques. The ones in the second category, instead, are able to control the applied forces directly.

3.2.1 Indirect Force Control

A robot manipulator that uses indirect force control strategies can be characterized by a mass-spring-damper system with configurable parameters. In general, using these control systems, the introduction of a contact force translates into a deviation of the final desired end-effector position, to which the control can react to by either applying forces to compensate for it or by enforcing a change of the desired pose.

Considering a robot with n joints, its joint state $q \in \mathbb{R}^n$, the measured end-effector position $p_e \in \mathbb{R}^3$ and rotation $\theta_e \in \mathbb{R}^3$ in Euler angles; the measured end-effector velocity can be computed as

$$v_e = J(q)\dot{q}, \tag{3.1}$$

with $\dot{q} \in \mathbb{R}^n$ the joint velocity and $J \in \mathbb{R}^{6\times n}$ the end-effector geometric Jacobian matrix. The measured end-effector wrench can instead be represented as

$$h_e = (f_e \; m_e)^T, \tag{3.2}$$

with $f_e \in \mathbb{R}^3$ and $m_e \in \mathbb{R}^3$ the linear forces and angular momentum, respectively.

The dynamic model of such manipulator can be described in its operational space as:

$$\Gamma(q)\dot{v}_e + Z(q,\dot{q})v_e + \eta(q) = h_c - h_e, \tag{3.3}$$

with the inertia matrix

$$\Gamma(q) = (JH(q)^{-1}J^T)^{-1} \tag{3.4}$$

and the overall external joint wrench

$$Z(q, \dot{q}) = J^{-1}C(q, \dot{q})J^{-1} - \Gamma(q)\dot{J}J^{-1}, \tag{3.5}$$

including the centrifugal $H(q)$, Coriolis $C(q, \dot{q})$ and gravitational forces $\eta(q) = J^T g(q)$ both in task and joint space. The value $h_c = J^{-1}\tau$ represents the controlled end-effector wrench when applying the joint torques τ.

Stiffness Control

Defining the current and the desired end-effector pose as

$$x_e = (p_e\ \theta_e) \text{ and } x_d = (p_d\ \theta_d), \tag{3.6}$$

respectively, the error between the two poses is

$$\Delta x_{de} = x_d - x_e. \tag{3.7}$$

In the case of a generic proportional-derivative (PD) controller with gravity compensation (See Sec. 8.5.1 of [127]), the motion control schema can be expressed as

$$h_c = A^{-T}(\theta_e)K_p\Delta x_{de} - K_D v_e + \eta(q), \tag{3.8}$$

with $K_P, K_D \in \mathbb{R}^{6\times6}$ symmetric and positive-definite,

$$A(\theta_e) = \begin{pmatrix} I & 0 \\ 0 & T(\theta_e) \end{pmatrix}, \tag{3.9}$$

and $T(\theta_e) \in \mathbb{R}^{3x3}$ mapping rotation and angular velocities, i.e., $\omega_e = T(\theta_e)\dot{\theta}_e$.

In the case of a total of zero external forces (i.e., $h_e = 0$), there would be a correspondence between the desired and effective end-effector pose (i.e., $\Delta x_{de} = 0$). It can be demonstrate that the system is asymptotically stable in that state [126]. Applying a constant value h_e, instead, it results in a non-zero Δx_{de}:

$$\Delta x_{de} = K_p^{-1}A^T(\theta_e)h_e, \tag{3.10}$$

and therefore

$$h_e = A^{-T}(\theta_e) K_p \Delta x_{de}. \tag{3.11}$$

From 3.10, it is clear that the behavior of such a system is equivalent to the reaction of a spring under the application of the wrench h_e, with K_p acting as the stiffness matrix of the spring model. K_p^{-1}, therefore, would act as a compliance component, giving a desired pose p_d an acceptable deviation from this command. Δx_{de} can be achieved exerting a certain force on to the environment; this is the fundamental approach behind **stiffness control**.

The selection of the stiffness matrix parameters is complex and usually depends on the task at hand. Notice that since the end-effector pose is described with six components, it is possible to select different stiffness values along the degrees of freedom, achieving different behavior in the respective (angular) directions. In any case, stiffness control is only based on a static relationship between the desired end-effector position and the contact forces.

Impedance Control

A more dynamic behavior can be achieved via **impedance control**. This time a acceleration-based motion control approach is used, to decouple the robot dynamics using an inverse dynamics control law. In this case, Eq. 3.3 varies into:

$$\Gamma(q)\alpha + Z(q, \dot{q})\dot{q} + \eta(q) = h_c - h_e, \tag{3.12}$$

with α an acceleration control input defined with respect to the robot end-effector.

Setting α as

$$\alpha = \dot{v}_d + K_M^{-1}(K_D \Delta v_{de} + h_\Delta - h_e), \tag{3.13}$$

the closed-loop system can be expressed as

$$K_M \Delta \dot{v}_{de} + K_d \Delta v_{de} + h_\Delta = h_e. \tag{3.14}$$

With $K_m, K_D \in \mathbb{R}^{6\times6}$ symmetric and positive definite, v_d and $\dot{v}_d$ the velocity and acceleration of a desired frame Σ_d, $\Delta \dot{v}_{de} = \dot{v}_d - \dot{v}_e$, $\Delta v_{de} = v_d - v_e$ and h_Δ the total elastic wrench in the case of a displacement between the end-effector and the desired frame.

Eq. 3.14 describes a generic mechanical impedance model. Again, in the case of $h_e = 0$ the system would be in a stable equilibrium and in the presence of a non-zero wrench, a displacement of the desired end-effector pose would take place. In this case, choosing

$$K_m = \begin{bmatrix} mI & 0 \\ 0 & M \end{bmatrix}, \tag{3.15}$$

and K_D a diagonal matrix with 3×3 blocks, Eq. 3.14 would represent a mechanical spring with mass m and inertia $M \in \mathbb{R}^{3\times3}$, having damping K_D.

An impedance controller would use the acceleration input α based on the position and force feedback from the sensors, as defined in Eq. 3.13. The dynamic control law in Eq. 3.12 is then used to compute the joint torques $\tau = J^T h_c$. A compliant behavior is produced according to Eq. 3.14 in the case of external forces from the environment.

3.2.2 Hybrid Force/Motion Control

Under the category of direct force control techniques lays one of the most common modern control strategies, *hybrid force/motion control*. In the case of partially constraint motions, it is possible to split the control task into two, one including free motion in space along the unconstrained degrees of freedom, and one focused on force control along the ones in contact with the environment.

Many control schemes have been proposed in the literature for hybrid force/motion control. For conciseness, this section focuses on the formulation of the typical force/position controller that has been employed throughout the work presented in this thesis. Since direct force control methods require to include a model of the interaction task, the formulation is usually categorized based on the interaction with a rigid or compliant environment. For simplicity, and since in practice modeling of the individual patient body is unrealistic for the task at hand, the following presentation focuses on the rigid case. The interested reader should further refer to classic motion control textbooks [125, 127] and to the current review literature on the topic [128].

Considering a robot's end-effector frame, its force h_e and velocity v_e can be obtained by specifying the vectors $\lambda \in \mathbb{R}^6$ and $\nu \in \mathbb{R}^6$, respectively. Generally, they can be mapped to the measured forces and velocities using

$$h_e = S_f \lambda \tag{3.16}$$

and

$$v_e = S_v \nu. \tag{3.17}$$

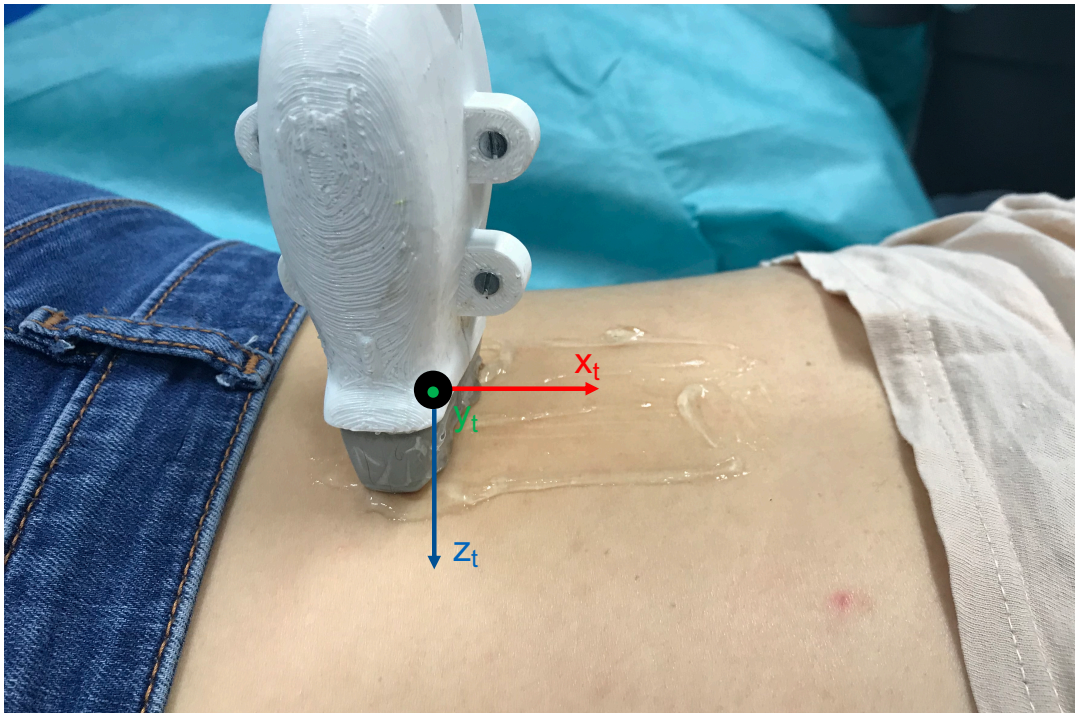

Fig. 3.2. **A ultrasound acquisition performed by a robotic arm.** In this over-simplified example, the task can be described as sliding the ultrasound probe along the x_t axis while exerting a fixed force along z_t.

The matrices $S_f \in \mathbb{R}^{6\times m}$ and $S_v \in \mathbb{R}^{6\times 6-m}$ are commonly known as selection matrices, and they describe the task, i.e., the degrees of freedom controlled by the force or the motion control part. Considering a task reference frame, an interaction can be defined along each of its axes, either using a desired force/torque or linear/angular velocity. Since there are 6 degrees of freedom in total, the dimensionality of S_f and S_v is set once the m positional constraints are set, i.e., if $m = 6$ then the controller aligns with a classical position controller.

As pictured in Fig. 3.2, a simplified ultrasound acquisition along one direction has some resemblance to the typical task of *sliding an object on a planar surface*. We define x_t as the axis along which the movement is executed and z_t the axis along which a force has to be applied. The task would have 3 motions constraints, i.e. non-zero velocity along x_t, zero velocity along y_t and z_t, and 3 force constrains, i.e. non-zero force along z_t, zero torques about x_t and y_t. The resulting matrices S_f and S_v would then be

$$S_f = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } S_v = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.18}$$

The block diagram in Fig. 3.3 provides a simplified description of a typical hybrid force/motion control. It can be proved, and it is represented in the diagram, that a complete decoupling between the force and motion control is indeed possible, leading to the following equalities:

$$\begin{aligned} \dot{\nu} &= \alpha_\nu \\ \lambda &= f_\lambda, \end{aligned} \tag{3.19}$$

with α_ν and f_λ the control inputs of the controller. Providing the desired velocity ν_d, can be obtained using the control law:

$$\alpha_\nu = \dot{\nu_d} + K_{P\nu}(\nu_d - \nu) + K_{I\nu} \int_0^t (\nu_d(\rho) - \nu(\rho))d\rho. \tag{3.20}$$

While the desired force λ_d can be achieve via the control law:

$$f_\lambda = \lambda_d + K_{P\lambda}(\lambda_d - \lambda) \tag{3.21}$$

or

$$f_\lambda = \lambda_d + K_{I\lambda} \int_0^t (\lambda_d(\rho) - \lambda(\rho))d\rho. \tag{3.22}$$

The respective matrices K_P and K_I are positive definite gain matrices. Adding feedback from the end-effector force/torque measurements can be obtained from

$$\lambda = S_f^\dagger h_e, \tag{3.23}$$

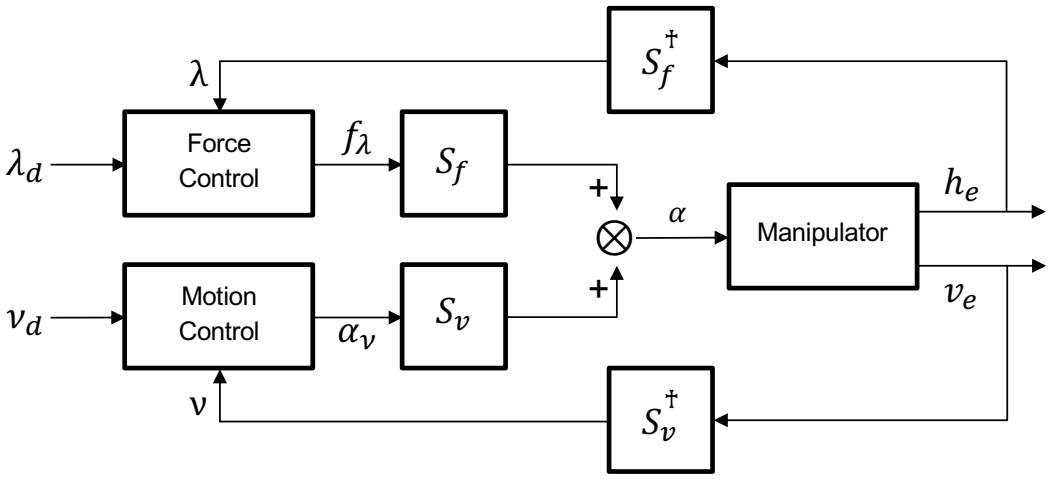

Fig. 3.3. Example of a block diagram of an hybrid force/motion controller.

with $S_f^\dagger$ a weighted pseudo-inverse of S_f. In a similar fashion, an hybrid *force/position* controller can be achieved using instead the control law:

$$\alpha_\mu = \ddot{p_d} + K_D p(\dot{p_d} - \mu) + K_P p(p_d - p), \tag{3.24}$$

with p_d the desired end-effector position.

4 3D Visual Sensing

A complete knowledge about the environment is crucial for robotic tasks. A robotic system has to be able to interact with or navigate in dynamic scenes with high precision. This is important for general robotics tasks such as indoor navigation [129, 130], collision detection and avoidance [131, 132], object grasping [133], and more. It assumes an even more important role in the medical domain, where information regarding the position of the patient position, the medical staff, and other equipment are essential for conducting safe and efficient procedures successfully. Moreover, these environments are often very dynamic, and therefore a real-time update is necessary.

The main principle behind range measurement, triangulation, has been used more than 2000 years ago by Eratosthenes to measure the earth's circumference from shadows created by sunlight. As for many technologies, a great boost came from the military during the two world wars. Later in the 1970s and 1980s, the larger availability of electronic components and optical apparatuses made more affordable the introduction of the first 3D sensor systems for industrial applications. The technologies behind the systems developed in those years were similar to the ones used in more modern systems: laser scanning, pattern projection, and time-of-flight sensors, but the costs and the lack of miniaturization were still the primary limits for the development of commercial systems. A revolution in the field has been the introduction of the Microsoft Kinect sensor in 2010 as the first 3D sensor available for the consumer market [134]. Since then, a large number of companies have introduced sensors to the market, enabling the use of this technology for many applications.

3D information are commonly represented in so-called *depth images*. They are componsed of a matrix, like a 2D image, but it additionally includes a depth value for each image pixel, with respect to the sensor location. A depth information is available for every object located in the field of view of the sensor. A depth estimation for each image pixel results in a dense depth map. This section introduces the main techniques used to obtain that information, focusing on the ones used within the contributions presented in this thesis. Some technologies are omitted since they are not commonly used in medical applications, e.g., laser scanners, due to their large size and the presence of moving parts. An in-depth description of the presented topics is available at [135].

4.0.1 Passive Stereoscopy

Stereoscopic methods, in general, rely on two cameras and achieve the computation of depth information from the relationship between two projection rays from the two cameras, taking inspiration from the human binocular vision system.

Fig. 4.1 represents an exemplary stereo camera system. For a point captured by the two individual sensors, its projection onto the two image planes is characterized by a displacement in position between them. This difference $d = x_l - x_r$ is called disparity, obtaining this information for the the whole image set generates a *disparity map*.

The depth information about each point can be then obtained via triangulation as

$$z = \frac{b \cdot f}{d \cdot pixel_{size}}, \tag{4.1}$$

where b is the *baseline*, i.e., the distance between the two camera centers, f is the cameras focal length and $pixel_{size}$ is the pixel size of the image sensors. This represents a trivial case since it is assumed that the two cameras are identical, i.e., the same focal length, they are horizontally aligned, that the images are acquired at the same instant, and they are rectified, i.e., no lens distortion. In practice, all these possible influences have to be taken into account.

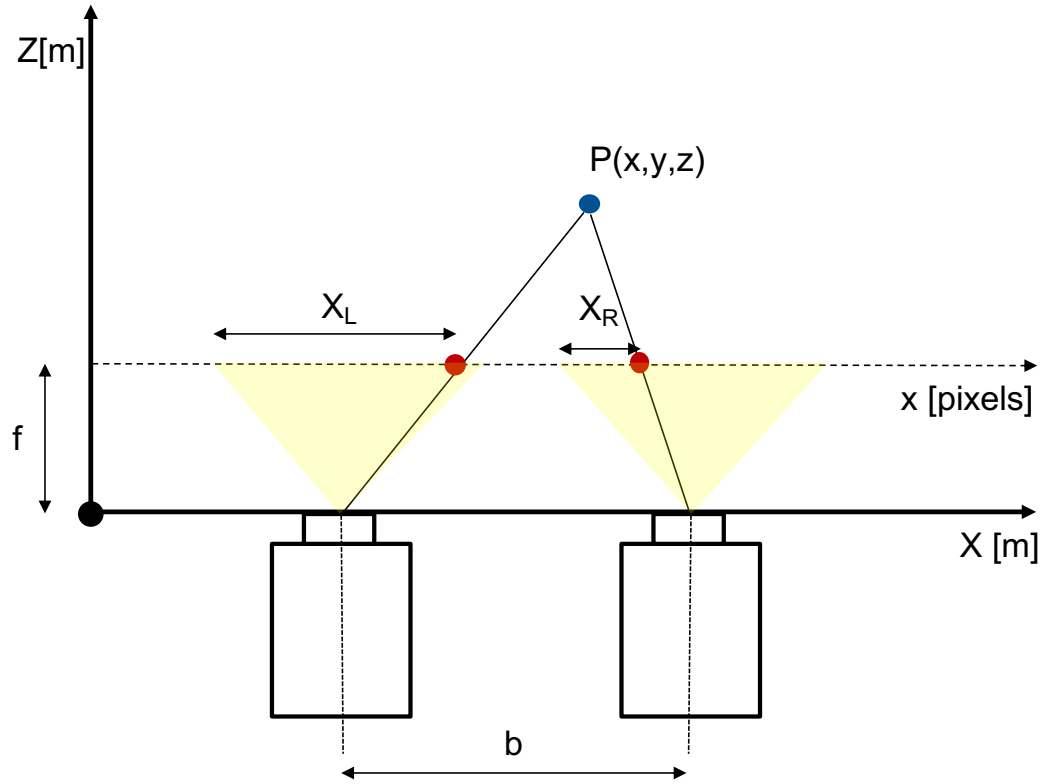

Fig. 4.1. A point in space is captured by the two sensors of a stereo camera system. The distance of the point with respect to the cameras can be estimated using the system baseline (b), its focal length (f) and the pixel coordinates in the two cameras ($x_l$andx_r).

The described approach relies on finding the the corresponding point in the two stereo images, a task defined as *correspondence problem* or *stereo matching*. Establishing these correspondences is itself a fundamental problem that is the subject of a large literature on stereo vision [136, 137]. Classic strategies for this problem perform a comparison of pixel blocks between the two images, using, for instance, similarity metrics such as the one described

in Sec. 2.4. In general, it can be computationally expensive, limiting early-stage depth from passive stereo algorithms to powerful machines of dedicated hardware.

A variation of passive stereoscopy, called *depth-from-motion*, uses multiple frames (typically from videos) to analyze differences between consecutive ones [138], enabling depth extraction also from monocular sensors [139, 140]. *Depth-from-focus*, instead, is particularly popular in microscopic imaging, where a different focus setting is used to acquire multiple images [141, 142].

4.0.2 Structured Light

A solution to the challenges in stereo matching for passive stereoscopic methods is given by active methods based on *structured light* sensors. The needed correspondences are, in fact, obtained using a projected pattern onto the scene. To avoid ambiguities in the correspondence extraction, time- or color-multiplexing can be employed to project multiple patterns [143], or a pseudo-pattern is used, so that the distortion given by its projection onto objects can be analyzed, as the dot pattern employed by the Microsoft Kinect in its first release [134]. Typically, an infrared pattern is projected to avoid interfering with other computer vision tasks, while an infrared camera captures its deformation.

4.0.3 Time-of-Flight Imaging

Depth information can be obtained by using a light source that illuminates the scene at certain intervals. The distance between the sensor and an object is measured based on the time difference between the light emission and the detection of the reflected light.

The *Time-of-flight* principle can be used with various types of signals, including sound. The same principle is used since decades for LIDAR (Light Detection and Ranging) imaging. Only in the recent years solid state LIDAR scanners have been introduced, enabling a broader use of this type of sensors since the presence of mechanical moving parts was limiting it.

Time-of-flight cameras enable the acquisition of very dense depth images, with the rapid computation of distance information for multiple image pixels at the same time. Compared to depth from stereo, the need for additional computations is limited, making these sensors suitable for high dynamic environments.

4.1 Feature-based registration

Regardless of the acquisition method, the scene is represented by the resulting depth map. This contains a set of N points in 3D, a so-called *point cloud*, $P = [\mathbf{p_1}, \mathbf{p_2}, \ldots, \mathbf{p_N}], p \in \mathbb{R}^3$. It is common to obtain 3D information from multiple sensors or sources and require to match them so that they can be referred to a common reference frame. For instance, in [7] and [8], 3D point clouds are obtained from structured light sensors and extracted from available pre-operative medical volumes, Fig. 4.2. Both points sets would contain a representation

of the current patient surface. The ability to match this information would then enable to accurately and safely guide robotic devices to the patient.

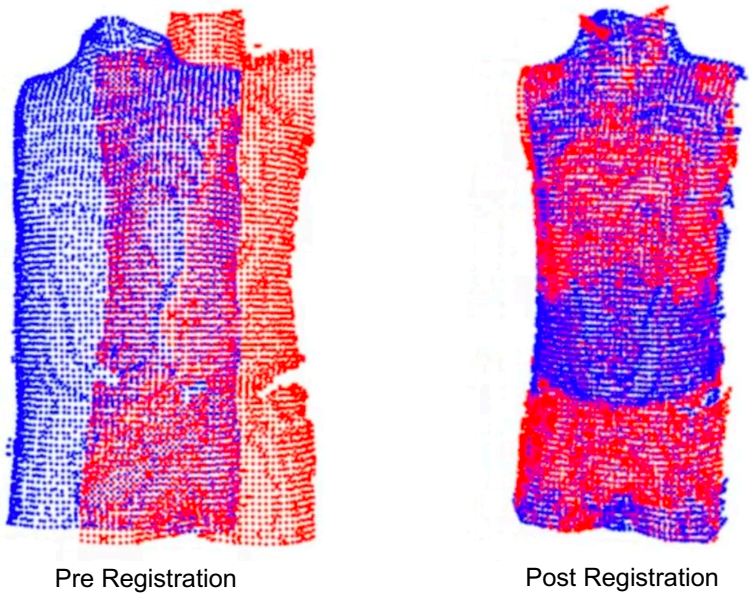

Fig. 4.2. **Feature-based registration is used to match two point sets.** In blue a point cloud extracted from a full-body MRI, in red one obtained from a 3D vision system. The two point sets undergo a deformable registration, its result is shown on the right side.

Image registration, in general, is a problem already described in Chapter 2.4, with a focus on intensity-based approaches for medical images. An alternative is to employ feature-based methods. While the previous ones, as we have seen, are based on the analysis of the image intensities, the later ones use distance measures based on features extracted from the input images to be aligned. Given a source point set S and a moving one M, the general problem, in the rigid case, can be formulated as:

$$M = R * S + t, \tag{4.2}$$

with $R \in \mathbb{R}^{3\times3}$ the rotation matrix, and the translation vector $t \in \mathbb{R}^{3\times1}$, necessary to align the two sets. While the entire point sets can be employed, robust *feature points* are commonly extracted to guide the registration, and a plethora of feature descriptors for 3D point clouds have been presented in literature [144, 145]. Once these are chosen, geometric transformation methods are employed to minimize the distance measure between the two point clouds.

In general, feature-based registration has been extensively discussed in the literature, with an intense use in the medical field [146, 147, 148].

Iterative Closest Point (ICP)

Initially presented by Besl and McKay [149], this classic algorithm registers the moving point set M iteratively applying a series of rigid transformations, estimating the points nearest-neighbor correspondences at each step. Formally,

$$dist(R,t) = \frac{1}{N}\sum_{i=1}^{n} ||Rm_i + t - s_i||^2 , \tag{4.3}$$

is minimized at each step until an optimal transformation is obtained. A large number of variations of the ICP algorithm have been proposed throughout the years. To optimize the error minimization step, EM-ICP [150] makes use of the Expectation-Maximization technique [151] rather than the quadratic function built from the Euclidean distance of the point correspondences as in the original formulation; while LM-ICP [152] uses a Levenberg-Marquard optimization. Other works address the need of ICP for an accurate initialization to be able to achieve an optimal solution [153]. Variations of the ICP algorithm to handle the non-rigid registration between point sets are also available in literature [154, 155]. Flavors of the ICP algorithm have been used in [7] and [1], as presented in Chapter 5 of this thesis.

Gaussian Mixture Model (GMM)

This algorithm, [156], describes the two point sets as two Gaussian mixture densities [157] and performs their registration by minimizing their L2 norm:

$$d_{L_2}(S,M,\theta) = \int (gmm(S) - gmm(T(M,\theta)))^2 dx. \tag{4.4}$$

That is, the algorithm attempts to optimize the statistical divergence between the two Gaussian mixtures. For point sets with low dimensionality, closed-form solutions exist for this problem, making the computation efficient. This is, in general, a particular case of the *kernel correlation* approach [158], but in practice, Gaussian kernels are the most used as they can effectively model many real-world problems.

Coherent Point Drift

It is a probabilistic approach similar to the one from GMM, which is agnostic to the used transformation model and can be employed for both rigid and deformable registration [159]. Rather than directly using the points from the set M, this algorithm employs its GMM centroids. For a given data point, its correspondence is the maximum of the GMM posterior probability when the two sets are aligned. The centroid motions in enforced to be coherent during the optimization process, such that the topological structure of the moving points set is preserved. This technique has been employed in the contribution presented in Sec. 5.3.1.

The probability density function for a point $s \in S$ is

$$p(s) = \sum_{i=1}^{M+1} P(i)p(s|i), \tag{4.5}$$

with $p(s|i)$, the Guassian distribution centered on the centroid $m_i \in M$ in D dimensions:

$$p(s|i) = \frac{1}{(2\pi\sigma^2)^{D/2}} \exp\left(-\frac{||s - m_i||^2}{2\sigma^2}\right). \tag{4.6}$$

The posterior probability of the GMM centroid m_i corresponding to the data point s_j is then

$$P(i|s_j) = \frac{P(i)p(s_j|i)}{p(s_j)}. \tag{4.7}$$

Re-parametrizing the GMM centroids location with the parameters θ, one has to maximizing the probability likelihood, or rather minimizing the log-likelihood to find the set of optimal parameters:

$$E(\theta, \sigma^2) = -\sum_{j=1}^{N} \log \sum_{i=1}^{M+1} P(i)p(s|i). \tag{4.8}$$

To find these optimal parameters, the original formulation makes use of the Expectation-Maximization (EM) algorithm [151]. The E-step estimates the values of θ and employs the Bayes' theorem to compute the mixture posterior probability distributions $P^{old}(i, s_j)$. During the M-step, updated values of θ are computed, minimizing the cost function:

$$Q(\theta, \sigma^2) = \frac{1}{2\sigma^2} - \sum_{j=1}^{N} \sum_{i=i}^{M+i} P^{old}(i, s_j)||s_j - T(m_i, \theta)||^2 + \frac{N_p D}{2} \log \sigma^2, \tag{4.9}$$

with

$$N_p = \sum \sum P_{old}(i|s_j) \leq N. \tag{4.10}$$

This function decreases coherently with the negative log-likelihood E, leading to a set of optimal parameters θ.

5 Contributions

With respect to the state-of-the-art presented in Sec. 1.3, the robotic ultrasound systems initially proposed in the literature were mostly used as pure mechanical devices to achieve spatial tracking of 2D transducers, facilitate stable and uniform movements while being manually maneuvered by an operator, or allow for teleoperation. The research focus was initially only on mechanical design and on the definition of proper control strategies to achieve smooth and safe acquisition. More importantly, all systems were considered an aid or remote extension for manual acquisitions, concepts of autonomous robotic acquisition were not explored. Following the advances in computer vision, many works targeted the analysis of the acquired ultrasound images to perform visual servoing tasks. While this initiated the development of functionalities that allow the control of robotic systems based on visual feedback (e.g., to maintain the visibility of anatomical target in the ultrasound field of view), their application in clinical procedures has been very limited.

In the literature prior to the following contributions, robotic ultrasound systems were not integrating all available sensing information to achieve optimal, safe, and robust autonomous acquisition. That is, autonomous systems for ultrasound imaging require more precise knowledge about their surrounding environment and accurate planning to reach their designated anatomical target. The author's contributions aim at the definition of a workflow to integrate multiple sensors and introduce the information they provide in the robot control loop. The final task is to obtain a system where the robotic components (kinematics, force sensing, control) and the imaging ones (ultrasound images, external 3D sensing) behave coherently to acquire the planned imaging target. Rather than focusing on the individual systems, the aim is to let them work in a **symbiotic** way, allowing them to benefit from each other's information and optimize their individual tasks.

The first two presented contributions focus on the definition of the basic concepts for this integration, i.e., using imaging sensing to optimize the behavior of a robotic manipulator **and** using force sensing from the manipulator to optimize the acquired images. The last two contributions, instead, showcase how these techniques can be leveraged to perform specific clinical procedures in the field of vascular surgery, targeting both a diagnostic and an interventional scenario. A summary of each work is presented, directly preciding the respective complete original publication, as requested by the Doctoral regulations.

5.1 Diagnostic Target Definition and Acquisition (TMI 2016)

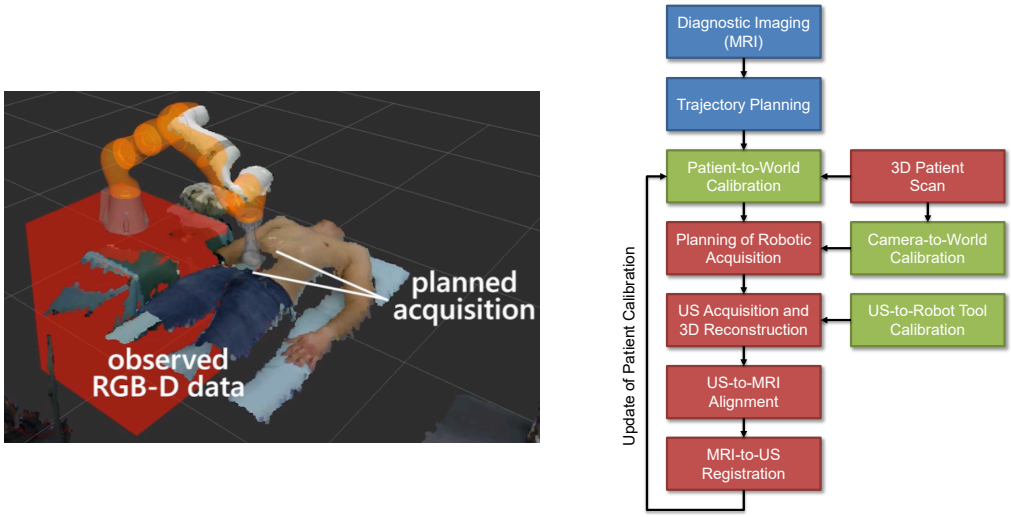

Fig. 5.1. **Left:** Robot kinematics and external 3D sensing information are fused in a common reference frame and visualized in a 3D environment. It is possible to notice the current patient location with respect to the manipulator. The acquisition planned using a pre-operative volume is transfer to the same environment by point-set registration. **Right:** The proposed workflow for autonomous robotic ultrasound acquisitions. It includes planning steps based on pre-operative diagnostic images (blue), a series of calibration between external camera to robot and robot to patient (green), and the actual acquisition and calibration refinement steps (red).

The task of defining the diagnostic target for robotic ultrasound acquisitions had not been targeted yet in literature. To achieve this task, one has to transfer the anatomical information obtained by some other means into the current robotic workspace. This work defines a complete clinical workflow (See Fig. 5.1, right side) to obtain this information for patient-specific acquisitions. The anatomical target can be defined on a previously available tomographic imaging modality (e.g., CT or, in this case, MRI), the proposed method relies on techniques addressed in Sec. 4.1 to align sets of 3D points. That is, a representation of the patient's surface can be extracted from the volumetric image at hand and matched with the current patient position via rigid registration. To achieve so, external 3D sensing is employed in the form of a calibrated RGB-D sensor overlooking the scene. Trajectories that are selected within the MRI volume can then be transferred onto the current patient location and followed completely autonomously by the robotic system. The structured-light 3D scanner, therefore, requires a patient-to-robot and image-to-patient calibration that allows for a proper trajectory initialization in the robot coordinate system. Further refinement of the overall calibration can be achieved via 3D registration of the first acquired ultrasound volume and the initial MRI one, with methods described in Sec. 2.4. This work demonstrates how a distance of only $0.97\,\mathrm{mm}$ between the planned acquisition and the actual one can be achieved using this follow-up adjustment. On the left side of Fig. 5.1, it is shown how the hand-eye calibration between the robotic system and the 3D camera allows to match their information in a common reference frame and transfer the planned acquisition into the scene.

Towards MRI-Based Autonomous Robotic US Acquisitions: A First Feasibility Study

Christoph Hennersperger*[1], Bernhard Fürst*[2], Salvatore Virga*[1,3], Oliver Zettinig[1], Benjamin Frisch[1], Thomas Neff[3], Nassir Navab[1,2]

[1] Chair for Computer Aided Medical Procedures (CAMP), Technische Universität München, Munich, Germany.
[2] Johns Hopkins University, Baltimore, USA.
[3] KUKA Roboter GmbH, Augsburg, Germany.

Contribution. The main design and development of the proposed system, including hardware integration and complete software development, was performed by the author of this thesis. The performed experiments and their evaluation were performed together with Oliver Zettinig. The conception of the initial idea supporting this publication was initially of Christoph Hennersperger. The writing and revision of the manuscript were done jointly by the three main authors, with the support of all co-authors.

Towards MRI-Based Autonomous Robotic US Acquisitions: A First Feasibility Study

Christoph Hennersperger*, Bernhard Fuerst, Salvatore Virga, Oliver Zettinig, Benjamin Frisch, Thomas Neff, and Nassir Navab

***Abstract*— Robotic ultrasound has the potential to assist and guide physicians during interventions. In this work, we present a set of methods and a workflow to enable autonomous MRI-guided ultrasound acquisitions. Our approach uses a structured-light 3D scanner for patient-to-robot and image-to-patient calibration, which in turn is used to plan 3D ultrasound trajectories. These MRI-based trajectories are followed autonomously by the robot and are further refined online using automatic MRI/US registration. Despite the low spatial resolution of structured light scanners, the initial planned acquisition path can be followed with an accuracy of 2.46 ± 0.96 mm. This leads to a good initialization of the MRI/US registration: the 3D-scan-based alignment for planning and acquisition shows an accuracy (distance between planned ultrasound and MRI) of 4.47 mm, and 0.97 mm after an online-update of the calibration based on a closed loop registration.**

***Index Terms*— Automatic imaging, image-guidance, medical robotics, multi-modal registration, robotic ultrasound, ultrasonic imaging.**

I. Introduction

ULTRASOUND (US) has become one of the standard medical imaging techniques and is widely used both within diagnostic and interventional applications. Examples for these areas include e.g. the utilization of ultrasound for vascular imaging of the carotid and abdominal arteries with respect to diagnostic purposes, or needle insertion for liver biopsy and ablation, where ultrasound imaging is used to guide the insertion process throughout the procedure. In general, clinical US is mostly based on 2D-images (except cardiac and obstetric applications), requiring a manual navigation of the probe. The resulting high operator-variability of the manual guidance is not only challenging for the application described above, but impairs a wider clinical acceptance of ultrasound for the extraction of quantifiable parameters from these data [1]. 3D ultrasound can potentially overcome these limitations, and is performed either using native 3D probes or by tracking 2D images in space, using a tracking target attached to the ultrasound probe (tracked freehand 3D ultrasound) [2]. While systems using native 3D probes are still not widely available in clinical practice, tracked ultrasound is easily accessible and can also be interpolated with respect to a regular grid [3]. When comparing native and freehand 3D ultrasound, both techniques have their merits. Native 3D ultrasound allows, on the one hand, for live 3D volume acquisitions in real-time and thus for a direct analysis of volume changes over time; a property which is especially exploited for 3D echocardiography [4]. On the other hand, the systems are expensive and only allow for the imaging of comparably small volumes restricted by the probes' field of view. Freehand 3D ultrasound does not pose limitations with respect to volume sizes, anatomies and trajectories, but can be potentially distorted by varying pressure applied by the operator, or changing anatomy caused by breathing or cardiac pulsation.

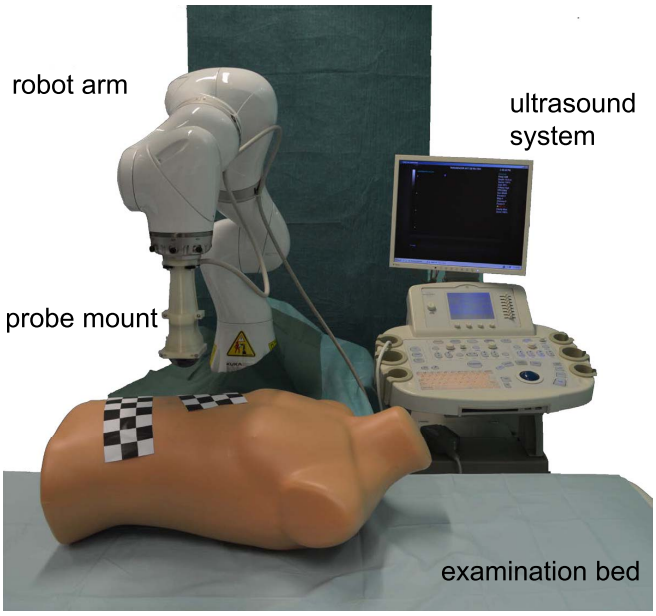

Fig. 1. In the presented system setup, a RGBD-camera is mounted on the ceiling, observing the scene with the patient lying on the examination bed. The robotic arm is approaching the patient from above and is equipped with a curvilinear ultrasound probe.

Manuscript received July 8, 2016; revised October 11, 2016; accepted October 21, 2016. Date of publication October 24, 2016; date of current version February 1, 2017. This work was partially supported by the Bayerische Forschungsstiftung project RoBildOR. C. Hennersperger. B. Fuerst, and S. Virga contributed equally to this work. *Asterisk indicates corresponding author.*

*C. Hennersperger is with the Chair for Computer Aided Medical Procedures, Technische Universität München, 85748 Munich, Germany (e-mail: christoph.hennnersperger@tum.de).

B. Fuerst is with Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA.

S. Virga and T. Neff are with KUKA Roboter GmbH, 86165 Augsburg, Germany.

O. Zettinig and B. Frisch are with the Chair for Computer Aided Medical Procedures, Technische Universität München, 85748 Munich, Germany.

N. Navab is with Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD 21218 USA and also with the Chair for Computer Aided Medical Procedures, Technische Universität München, 85748 Munich, Germany.

Digital Object Identifier 10.1109/TMI.2016.2620723

While ultrasound acquisitions are mostly performed by physicians or medical staff, data collection for both 2D and 3D ultrasound might be improved by automating the acquisition process. This goes in hand with the demands of the healthcare system, showing that current freehand ultrasound scanning techniques suffer from high inter-operator variability, hindering the overall clinical acceptance, especially of 3D applications [1], [5].

In this work, we focus on the improvements robotic technology could provide for interventional ultrasound imaging. While we propose a general system for automatic robotic 3D-US acquisitions (see Fig. 1), the advantages of robotic technology become intuitively apparent by considering a possible clinical workflow, e.g. for the aforementioned application of liver needle placement: To place the tip of a needle at the target site, the needle is advanced carefully under US-guidance to reach a final target position in current practice. Thereby, the procedure is heavily based on pre-interventional X-ray Computed Tompgraphy (CT) or Magnetic Resonance Imaging (MRI) datasets. Throughout the insertion, it is important to maintain steady ultrasound views, providing an overview of all essential structures in the abdomen (tumor, vessels, lung, liver tissue). A robotic ultrasound system could replace the current setup, where the interventionalist has to manually hold the US probe and 'cognitively' fuse the data with the interventional plan. 3D-US reconstructions of the target anatomy performed periodically throughout the procedure can be used to guide the insertion process through 3D visualization, but also to compensate for patient and breathing motion [6]. After each 3D-US acquisition, CT or MRI datasets available from the interventional planning are automatically co-registered to the acquired 3D-US scans. This allows for a careful advancement of the needle based on the US data, where fused data can also be employed to identify the target and potential risk regions.

To this end, a fully automatic system to acquire 3D ultrasound datasets could avoid the high inter-operator variability and resource demands of current clinical systems. Other applications include e.g. the localization of suitable entry-points for port-placement, automatic acquisitions of a pre-defined anatomy, patient-positioning for radiotherapy, or teaching of ultrasound based on robotic acquisitions and guidance. By addressing limitations of today's ultrasound techniques using a robotic framework, this would not only open up the way for new applications in an interventional setting, but also for potential society-wide screening programs using non-invasive ultrasound imaging [1], [7]. In this view, autonomous acquisitions have the potential both to facilitate clinical acceptance of robotic imaging techniques by simplifying the workflow as well as to reduce the scanning time and the necessary amount of manual operator input.

First attempts aiming at automatized ultrasound acquisitions used mechanical stepper motors, moving the ultrasound probe in a controlled fashion [8]. More recently, different robotic systems were developed in the context of clinical ultrasound, including applications to imaging, surgical interventions and needle guidance [9], [10]. While freehand ultrasound scanning enables a fast and dynamic acquisition and screening of several anatomies, modern compact and lightweight robotic arms can further support the physician, i.e. by automatically maneuvering a second imaging probe or tool in order to enable live image fusion [11]. Moreover, such systems eventually incorporate pressure, hand tremor or movement correction [6], or an automatic servoing for tissue biopsy based on a registration to prior US acquisitions [12]. With the goal of fully automatic acquisitions, however, a prior planning of the actual target area of interest is a prerequisite, since whole-body ultrasound scans are impractical and time-consuming. In a clinical interventional setup, planning should be performed by the medical expert based on anatomical data, as given by MRI and CT. Given an appropriate target, (several) 3D ultrasound datasets can then be acquired autonomously during an intervention, without requiring the presence of medical staff guiding the robot. We acknowledge that there have been attempts in the past to incorporate tomographic image information into robotic systems to improve the visualization of ultrasound image information to physicians. However, to our knowledge an integration of these data to enable planning and the *automatic* acquisition of 3D US datasets has not been considered so far.

In this work, we aim at closing this gap in the workflow of current robotic support systems in order to allow for fully automatic 3D ultrasound acquisitions using a robotic imaging system. We present the path towards an autonomous robotic ultrasound system to assist clinicians during interventions by performing multiple and especially reproducible examinations based on pre-interventional planning. To this end, we introduce a first concept for a robotic ultrasound system consisting of a lightweight robotic arm, a clinical ultrasound machine, and a RGB-D camera which is used to observe the scene. This enables the direct planning of a patient-specific trajectory by selecting its start- and endpoints in the patient's MR (or CT) image. A registration of the actual patient position to the MRI allows for the automatic acquisition of 3D ultrasound data. By using intensity-based image registration, we can close the loop and perform an online-update of the patient-to-world registration, accounting for inaccuracies of the RGB-D information as well as patient movement. Thus, the overall goal is to perform fully autonomous image acquisitions within a closed control loop by utilizing 3D surface information, pressure estimations of the robotic system, and image-based servoing, to image regions of interest defined by pre-interventional data.[1] The workflow showing the main steps of the proposed solution is depicted in Fig. 2.

The remainder of this work is organized as follows: Section II will give an overview of related work in the field of autonomous robotic ultrasound acquisitions. We introduce our proposed system and all corresponding and necessary steps in Section III, before experiments and results are presented in Section IV. Finally, opportunities and challenges associated to the system are discussed in Section V, before we draw final conclusions in Section VI.

[1]Supplementary video material to this manuscript is available at http://ieeexplore.ieee.org, provided by the authors. It features a brief overview of the acquisition setup and workflow with the distinct steps.

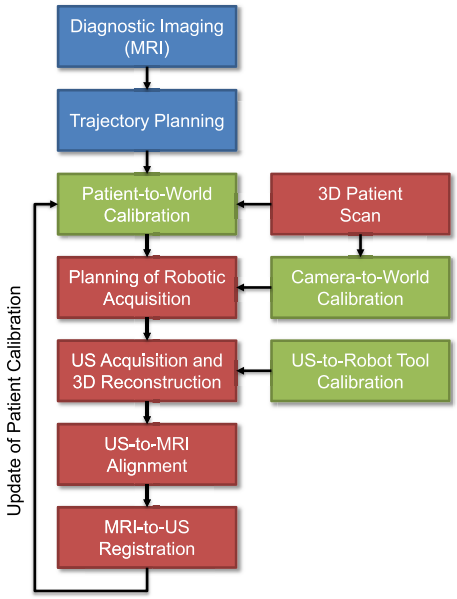

Fig. 2. The proposed workflow for autonomous acquisitions includes a planning based on diagnostic images (blue), calibration of both robot and patient to the world reference (green), and interventional acquisition (red). During acquisition, a 3D patient scan using a structured-light 3D scanner is used for the initial patient to world alignment and robot to world calibration. Following the autonomous robotic ultrasound acquisition, a US to MR registration is conducted to refine the patient to world alignment.

II. Related Work

Over the past two decades, research groups have focused on improving ultrasound image acquisition systems in terms of accuracy, usability, remote controllability as well as by a synchronization with pre-operative planning. For this purpose, (medical) robots have been introduced for different applications. Inspired by [9], we differentiate between situation aware robotic systems, and surgical assistant systems.

Surgical assistant systems primarily serve as advanced tools and are directly controlled by the surgeon, mostly acting in a master-slave fashion. Thereby, these systems are used to augment the operator's ability to perform a specific task, e.g. by allowing precise, jitter-free movements, or the application of constant desired pressure. Systems targeted to specific applications, such as carotid imaging [13], have been explored recently. Furthermore, ultrasound has been added as a tool to the more general surgical da Vinci platform for several applications [14], [15] using specialized drop-in ultrasound probes which can be picked up by the surgeon intra-operatively. Following this concept, a general system providing tele-operated ultrasound using different robotic arms is currently being developed [16]. For a more detailed overview of surgical assistants in ultrasound, the reader is referred to [9].

In contrast to these systems, situation aware robotic system perform at least one task autonomously based on a pre-defined situation-specific model, requiring an awareness of the task and its surrounding environment. In the following, we will focus our review on this area, as it is the primary target area of this work. In order to do so, we differentiate between Automatic Robotic Support Systems (ARSS) providing automatic support for a defined task, as well as Automatic Data Acquisition Systems (ADAS), targeting fully automatic robotic imaging.

A. Automatic Robotic Support Systems

Boctor *et al.* [17] originally proposed a dual robotic arm setup holding a needle in one arm and a US probe in the second, in order to facilitate an accurate placement of the needle to perform liver ablations. In a later work [18], the robotic arm holding the US probe was replaced by a freehand 3D ultrasound setup to improve general usability while still enabling accurate placement. With respect to an integration of tomographic image data into the robotic ultrasound environment, Zhang *et al.* [19] propose a system combining a robotic arm with ultrasound, optical tracking and MRI data manually registered to the system in order to improve the accuracy of needle placement. Targeting at optimal robotic ultrasound acquisitions to cover a predefined volume of interest, a planning-framework is presented in [20] to perform automatic robotic imaging trajectories. Other groups have focused more on a direct integration of the resulting images into the robot control, facilitating visual servoing based on robotic ultrasound. Abolmaesumi *et al.* [21] were among the first exploring the combination of robotic arms and ultrasound imaging based on visual servoing by combining an experimental robotic arm holding an ultrasound probe with vessel feature tracking to follow vessels during the ultrasound scan. Krupa *et al.* [6] further explored visual servoing towards an application to ultrasound by using ultrasonic speckle as a main characteristic feature to compensate for potential patient motion both in-plane and out-of-plane with respect to the ultrasound probe. These concepts were later adapted to fully automatic needle insertion under ultrasound guidance [22] and automatic probe positioning in order to allow for an optimal skin coupling with respect to the current position [23]. Recently, a servoing approach using live ultrasound to ultrasound registration was presented, targeting at screw placement in spine surgery [12], [24]. Another intra-operative application of robotic ultrasound was presented for the mobilization and the resection of the internal mammary artery in the course of coronary artery bypass grafting [25], following a servoing approach using color Doppler ultrasound images. For a comprehensive and detailed background of visual servoing, the reader is further referred to [26], [27], where concepts and basic approaches are explained in detail.

B. Automatic Data Acquisition Systems

While all approaches described above perform ultrasound acquisitions semi-automatically or under direct physician guidance, they either require a manual positioning of the US probe, or a proper definition of the working area. A first step towards fully autonomous scanning of ultrasound trajectories [28] combines a designed pneumatic probe holding case with optical tracking in order to enable scanning of a small volume of interest. While this approach has a high potential for certain applications, it still requires manual placement of the autonomous holding cage on the patient and does not enable fully automatic scanning. Using two robotic arms holding separate ultrasound probes, a system for 3D

robotic tomographic imaging pioneers the idea of localized 3D tomographic ultrasound imaging [29]. With respect to an intra-operative setting, however, the positioning of two robotic arm seems challenging, which is why the major field of applications for this technique might lie in the imaging of smaller structures. Focusing on vascular applications, a targeted scanning system for the automatic acquisition of vascular 3D ultrasound is presented in [30], combining pressure compensation with vessel tracking mechanisms. Finally, another system focusing on automatic data acquisition for liver screening was recently presented [31], combining RGB cameras for an automatic determination of liver scanning regions, followed by an automated acquisition workflow using different scanning directions. While this system operates fully automatically, it is limited to 2D images only and performs acquisition based on classified body features, which are prone to an erroneous feature determination, limiting the practical applicability of the system.

With this work, we focus explicitly on a system which enables fully autonomous robotic 3D ultrasound acquisitions based on an individual planning from tomographic image data (performed before the intervention). Thereby, the system does not rely on an initial probe positioning, manual registration of image data, or feature tracking in ultrasound or RGB-images. Instead, we use an initial patient-to-robot registration to perform 3D US acquisitions, which are then used in order to refine the overall registration and calibration. To do so, an intensity-based image registration of the pre-aligned 3D US and tomographic image data is used to retrieve a refined and updated patient-to-robot registration. Using the robot's built-in force and torque sensors, this further allows ultrasound acquisitions with optimal skin force, providing repeatable 3D ultrasound volumetric datasets.

III. Methods

This section first introduces the main component of the robotic ultrasound hardware setup in III-A. Next, III-B describes the pre-interventional imaging as well as trajectory planning for autonomous acquisitions. Following our proposed application, all required calibration steps and procedures are defined in III-C, before the interventional workflow for robotic acquisitions is explained in detail in III-D, including robotic control strategies, updates, ultrasound acquisitions and refinement of world-calibrations based on US acquisitions.

A. Hardware Setup

Our system consists of a lightweight robot, the ultrasound device, and a structured light RGB-D 3D scanner. Figure 3 shows a schematic overview of the relevant system components with the necessary coordinate transformations employed throughout this work. While the ultrasound transducer is directly mounted onto the end-effector of the robotic arm, the 3D scanner is attached to the ceiling and serves as a vision system allowing for the direct calibration and registration of all system parts.

1) Robotic Arm: Based on developments of the German space center (DLR) [32], KUKA introduced a robotic platform targeted at direct human-machine interaction, referred

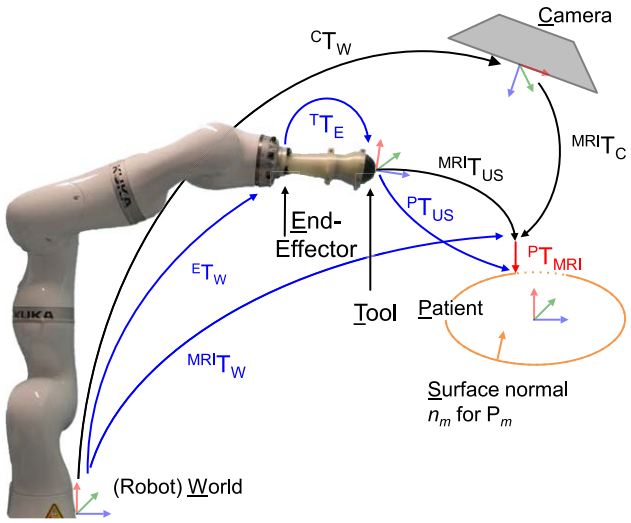

Fig. 3. Using the 3D scan of the patient $^{MRI}\mathbf{T}_{C}$ and the transformation from world to the camera $^{C}\mathbf{T}_{W}$ the world-to-patient transformation $^{MRI}\mathbf{T}_{W}$ is estimated. The tool (transducer apex) reaches the patient's surface (orange) by applying the tool-to-patient transformation under consideration of the surface normals n_m. The discrepancy between the estimated position and the real position of patient is indicated by $^{P}\mathbf{T}_{MRI}$, which is detected by intensity-based registration (red).

to as 'Intelligent industrial work assistant' - iiwa (KUKA Roboter GmbH, Augsburg, Germany). This system consists of a 7 joint robotic arm with corresponding control units and consequently enables one redundant degree of freedom (6+1 in total). As a result of this design, the robot provides dynamic movement and flexible adaption of trajectories to the working environment. With respect to robotic ultrasound, the incorporated high-accuracy torque sensors in each of the seven joints are evenly important, as a robotic ultrasound platform has to be fully compliant to both patient and staff. Based on significant safety measures for collision detection, the robot subsystem is certified for human-machine-collaboration due to the compliance to standards for functional safety. Thus it is considered to be safe for use in direct interaction with humans. Detailed specifications and design choices can be found in [33].

The KUKA native Sunrise.OS and its Sunrise.Connectivity module allow for the full low-level real-time control of the KUKA iiwa via UDP at rates up to 1 kHz, acting similar to the Fast Research Interface (FRI) [34] as proposed for the previous generation of the KUKA LWR robot arms. In this work, a publicly available software module[2] developed in our lab is utilized to enable a direct interaction between Sunrise.OS and the Robot Operating System[3] (ROS) framework. By doing so, low-level functionality and control provided by the robot manufacturer can be integrated with RGD-D information and high-level robotic interaction through ROS, as required for the proposed system.

2) Structured-Light 3D Scanner: The Kinect 3D camera (Microsoft Corporation, Redmond, Washington, USA) allows for the recording of color images and depth data

[2]https://github.com/SalvoVirga/iiwa_stack
[3]http://www.ros.org/

(RGB-D) at 30 frames per second. The system contains one camera for color video, an infrared laser projector for the projection of a structured light pattern and a camera to record said pattern. This device has been used for several medical applications, such as for a touch-less interface [35], for image-to-patient registration [36], or to provide initialization to US/MRI registration [37]. In the proposed system setup, the camera is attached to the ceiling, focusing on the robotic system and examination bed. The RGB-D data is used in order to i) align the camera within the world coordinate system (camera-to-world calibration - Sec. III-C), and to ii) register the patient lying on the examination bed to the world reference frame (patient-to-world calibration - Sec. III-C).

3) Ultrasound System: In general, any ultrasound device could be incorporated within an autonomous system by using the video output of an arbitrary clinical US device and recording those images using hardware frame-grabbers connected to the device. In practice, such systems provide a partially diminished image quality due to compression artifacts and - more importantly for robotic acquisitions - a temporal lag between the acquired image by the US machine and the recorded frame via frame-grabber. To enable direct acquisitions and advanced servoing in real-time, we favor ultrasound devices providing direct interfaces for retrieving acquired ultrasound data with minimal temporal delay. An Ultrasonix RP ultrasound machine (Analogic Corporation, Peabody, Massachusetts, USA) is used in combination with a 4DC7-3/40 4D Convex curvilinear transducer used in single-plane mode. The system provides the ulterius[4] API, enabling both the streaming of ultrasound data to a client as well as the control of US acquisitions parameters remotely through the Ethernet connection.

B. Diagnostic Imaging and Trajectory Planning

The goal of an autonomous 3D ultrasound acquisition is to perform a robotic-ultrasound trajectory based on a planning on pre-interventional images. With respect to a fast and intuitive planning, a physician wants to see either i) the region around a specific target, such as for liver or lymph node biopsy, or ii) scan a defined region, such as an organ, in order to perform a diagnosis based on these images. In this work we focus on the general case, where the operator can select the region of interest (ROI) directly. For our setup, a T2-weighted MRI volume is used as basis to determine the ROI for the ultrasound scan. The physician simply selects the start- and endpoint P_s, P_e of a trajectory in the MRI data, where the acquisition path direction $\vec{d}_t$ is defined by these points:

$$\vec{d}_t = P_e - P_s. \tag{1}$$

To enable both the transfer of the planned acquisition path to the robot world coordinate system and the initialization of the MRI/US registration, the patient's surface is extracted from the MRI by thresholding the image. Based on the patient-to-world calibration (see Section III-C), the trajectory direction and points can be directly transformed into world coordinates. Consequently, the segmented patient surface is exploited to determine the points of interest for the autonomous ultrasound scan.

[4]http://www.ultrasonix.com/wikisonix/index.php/Ulterius

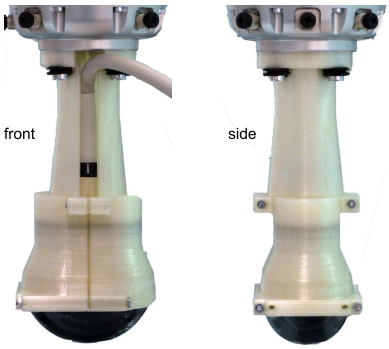

Fig. 4. Ultrasound probe mount designed for lightweight arm end-effector.

C. System Calibration

1) US Tool Configuration and Calibration: The transformation ${}^{\mathrm{T}}\mathbf{T}_{\mathrm{E}}$ from the robot end-effector to the transducer tool tip can be obtained directly from the CAD models used for manufacturing of a customized mount for the rigid attachment of the US probe to the robot end-effector. The probe mount utilized for the robotic system is shown in Fig. 4 and is tailored to the ultrasound transducer for optimal and rigid attachment.

To perform ultrasound acquisitions, a second spatial transformation ${}^{\mathrm{US}}\mathbf{T}_{\mathrm{T}}$, pointing from the probe tool tip to the US image origin, is defined by the image scaling (transformation from px to mm) and size with respect to the US probe apex

$${}^{\mathrm{T}}T_{\mathrm{US}} = \begin{bmatrix} s_x & 0 & 0 & s_x t_x \\ 0 & s_y & 0 & s_y t_y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{2}$$

where s_x, s_y determine the resolution of the ultrasound image, and t_x, t_y the translation from the apex, defined by the center of curvature, to the image origin. If an additional refinement of the transformation from the ultrasound image to the transducer apex is required, a standard ultrasound to tracking calibration technique can be applied [38].

As the used lightweight robot provides force sensors in each joint, an additional calibration with respect to the weight and center of mass of the tool is necessary to allow the force controlled motion. For an accurate load calibration, a proprietary calibration algorithm provided with the robot is used to determine these values in 3D.

2) Camera-to-World Calibration: This calibration step allows for the control of the robot - representing the world reference frame - as observed in camera coordinates. In the case of RGB-D cameras, this transformation relates from the camera to the robot arm, and can be computed either by using the 2D RGB images or the 3D information obtained by the depth sensor. An accurate calibration can be achieved with both techniques, although the 3D case requires more effort, user interaction and processing time [39]. Alternately, in [20], an additional optical tracking system is used to achieve a good

calibration between robot and camera. However, by leveraging the 2D RGB images, it is possible to compute the transformation ${}^{C}\mathbf{T}_{W}$ between the robot base and the RGB-D camera by using a standard hand-eye calibration method [40] in its 'eye-on-base' variant, without requiring an additional optical tracking system. To do so, an Augmented Reality (AR) marker is placed on the robot flange and the transformation between the marker and the camera estimated using the ArUco [41] tracking library. Several tracked poses combining random (non co-linear) movements around several rotation axis are recorded and used as correspondence pairs for calibration. The robot forward kinematics and the presented transformation ${}^{T}\mathbf{T}_{E}$ complete the chain of transformations from the camera reference frame to the US transducer. It should be noted that by using this calibration, the exchange of the US transducer with another one does not require a full camera recalibration, but only recomputing the last transformation of the chain between the end-effector and the US probe ${}^{T}\mathbf{T}_{E}$. Furthermore, the AR marker-based calibration is used for the camera-to-world calibration only, allowing for a fully marker-free data acquisition and update. The transformation between the RGB and the depth sensors of the Kinect can be obtained as proposed by [42]. For the hardware used, this relates to a translation of 2.5 cm along the transverse axis of the camera.

3) Patient-to-World Calibration: With respect to a fully autonomous acquisition, it is necessary to transfer trajectories planned in a tomographic image to the actual position and orientation of a patient lying on the bed. To do so, acquired 3D RGB-D information of the patient (Sec. III-A) provides surface information, which can be extracted simultaneously from the tomographic images. These data can then be used directly for aligning both surfaces and consequently determining their spatial transformation. In the following section, we will describe the necessary steps to enable such a calibration in detail.

Surface Extraction from MRI Data: To extract the patient (or phantom) skin surface from the MRI images, we employ an approach which does not require complex segmentations, as the extracted and matched surfaces will be refined throughout the acquisition process using image-based registration. Therefore, tomographic data (in this work T2-weighted MRI) is thresholded first, neglecting all values $I < \tau$. Throughout all our experiments, $\tau = 100$ provided a sufficient discrimination between the patient and the background (acquisition table and room around patient). As the resulting mask contains holes and is partially irregular, a morphological closing is performed on the thresholded image containing of subsequent dilation and erosion. The surface can then be retrieved from the image data [43], where only the connected surface component covering the highest ratio of total pixels is kept for further processing

$$\Omega_B = \underset{i}{arg\,max} \sum \Omega_i, \quad \Omega_i = \{x_1, x_2, ..., x_m\}, \tag{3}$$

with x_m being the surface positions of the points contained in the component Ω_i in 3D space.

Spatial Change Detection in RGB-D Data: To separate the patient from the background and other objects within the view of the camera, a spatial change detection is performed. Octrees are often used in computer graphics to partition a volume by recursively dividing cubes [44]. First, a background point cloud is recorded and added to an octree. After positioning the patient on the bed, a second point cloud is retrieved and added to another octree. The differences in the tree structure allow the efficient determination of newly added points, which represent the background-subtracted object, in our case the patient. To increase robustness and compensate for noise, a minimum of n pixels (experimentally set to 2) must be contained in each tree node.

Surface Matching: The alignment of surfaces can either be achieved using Iterative Closest Points (ICP), a non-linear derivative of ICP, or by directly aligning features. As the patient surface shape deviates strongly from the shape of the background (e.g. table or ultrasound scanner) a feature alignment process is applicable. On the foundation of the calibration refinement using intensity-based registration (Sec. III-D), our framework automatically accounts for local deformations and inaccuracies. Consequently, we employ ICP as surface matching method, as it provides a robust and especially highly efficient global alignment, which is then used to initialize the intensity-based registration.

Result of the surface-matching between extracted MRI and RGB-D information will be a rigid transformation from MRI-space to RGB-D camera space

$$P_i^C = {}^{C}\mathbf{T}_{MRI}\, P_i^{MRI} = ({}^{MRI}\mathbf{T}_C)^{-1}\, P_i^{MRI}, \tag{4}$$

with P_i^{MRI}, P_i^C being the surface points in MRI and camera-space respectively.

D. Autonomous US Acquisition

Following the workflow as shown in Fig. 2, this section will describe all steps carried out to perform one or multiple autonomous acquisitions.

1) Planning of Robotic Acquisition: Based on the alignment of both camera and patient within the world reference (Sec. III-C), the next step is to transfer the previously planned image trajectory to a robotic control trajectory which can be executed by the robotic arm. To allow accurate and safe trajectories, a proper force control strategy of the robotic manipulator, and the planning of the acquisition are necessary.

Stiffness and Force Control: The manipulator is mainly operated using a stiffness controller, which represents a Cartesian virtual spring between the desired (setpoint) position x_{set} and current (measured) position x_{msr}. The control law for this strategy is defined by the transposed Jacobian matrix J^T

$$\tau_{Cmd} = J^T(k_c(x_{set} - x_{msr}) + D(d_c)) + f_{dyn}(q, \dot{q}, \ddot{q}), \tag{5}$$

where k_c is the Cartesian stiffness of the virtual spring $k_c(x_{set} - x_{msr})$. The damping term $D(d_c)$ is dependent on the normalized damping value, while the dynamic model of the manipulator $f_{dyn}(q, \dot{q}, \ddot{q})$ is predefined by the manufacturer [34]. The resulting joint torque is computed by a Cartesian law using J^T.

In order to allow for compliant force applied in the US probe direction, the force control is modified such that the

force term $D(d_c)$ is set to a high damping in the respective probe direction, allowing for a compliant movement. Thus the high accuracy torque sensors and real-time forward kinematics computation of the system are exploited in this view in order to provide an acoustic force coupling without applying excessive forces (>25 N) on the skin. The stiffness controller allows to achieve an indirect force control [45], considering the relationship between the deviation of the end-effector position and orientation from the planned motion as well as the contact force and moment. The online parametrization of the virtual spring values makes it possible to maintain the contact force constant. The Cartesian stiffness along the direction of the US probe is set in the range [125-500] N/m according to the anatomy of the patient, and the force desired is parametrized as 5 N. The stiffness and forces in the other directions are parametrized to 2000 N/m and 0 N accordingly. If an excessive force > 25 N occurs, the robot's internal subroutine will automatically stop the acquisition process. This feedback loop enables the compliant motion constrained by the patient or other objects.

Planning of Acquisition Path: Based on the selected points (see Section III-B) and corresponding surface normals provided by the RGB-D data, the acquisition path is planned. The selected start and end points of the trajectory P_s,P_e define the direction vector of the trajectory (see Eq. (1)), which is used to define equidistant sampling points $Q^1, Q^2, \ldots, Q^i$ along the line with a distance of 2 cm. For each sampling point, the closest surface point $Q^i \rightarrow P^k$ is retrieved by a nearest-neighbor search. Along with the corresponding surface normal directions n_k, the points can then be used directly as target tool poses for robotic movements. It should be noted that the force controller implicitly commands the robot to adapt the tool position until the surface is reached. As the robot approaches the next trajectory point P^{k+1}, the direction of the transducer is stepwise changed to the subsequent surface normal position defined by P^k, P^{k+1} by the robot control.

2) US Acquisition and 3D Reconstruction: The live streams of robot tracking data and ultrasound images are transmitted via Ethernet, therefore the offset between tracking and US data is small compared to the framework overhead. In order to enable 3D registration with the diagnostic image, a volume compounding is performed using a backward normalized-convolution approach following [46], yielding regular-spaced 3D volume data from the arbitrarily sampled ultrasound frames in 3D space.

3) US-to-MRI Alignment: Making use of all previously estimated transformations from camera to MRI ${}^{MRI}\mathbf{T}_C$, world to camera ${}^{C}\mathbf{T}_W$ and ultrasound to tool ${}^{T}\mathbf{T}_{US}$, we can transform both ultrasound and MRI data into the world space using the respective transformations from both image spaces to the world reference frame

$$ {}^{W}\mathbf{T}_{MRI} = ({}^{MRI}\mathbf{T}_W)^{-1} = ({}^{MRI}\mathbf{T}_C \cdot {}^{C}\mathbf{T}_W)^{-1} \quad (6) $$

$$ {}^{W}\mathbf{T}_{US} = {}^{W}\mathbf{T}_E \cdot {}^{E}\mathbf{T}_T \cdot {}^{T}\mathbf{T}_{US}. \quad (7) $$

Both image datasets will be roughly aligned after transformation into the world coordinate frame, relying on prior ultrasound and patient to world calibrations.

4) MRI-to-US Registration: Supposing a rough overlap of MRI and US images as described above, an intensity-based MRI/US registration can directly be initialized in order to obtain an updated transformation between ultrasound and tomographic images.

The LC ^{2} similarity method allows for the registration of MRI and US images by correlating the MRI intensities and gradient magnitudes to the US intensity values. High robustness, wide convergence range and the application for rigid, affine and deformable registration have been shown in [47]. This approach is utilized in a two-step process. Based on the calibration and transformation chain, we can bring recorded ultrasound volume data directly into the MRI reference frame (or both into the world reference frame)

$$ {}^{MRI}\mathbf{T}_{US} = {}^{MRI}\mathbf{T}_W \cdot {}^{W}\mathbf{T}_{US} = ({}^{W}\mathbf{T}_{MRI})^{-1} \cdot {}^{W}\mathbf{T}_{US}. \quad (8) $$

In a second step, LC ^{2} is used to determine the rigid transformation component, aligning the transformed US and MRI images in world space in reference to the actual patient position

$$ {}^{P}\mathbf{T}_W = {}^{P}\mathbf{T}_{MRI}\,{}^{MRI}\mathbf{T}_W, \quad (9) $$

with ${}^{P}\mathbf{T}_{MRI}$ being the updated transformation to compensate for patient movement, deformation, as well as tracking and detection inaccuracies of ${}^{C}\mathbf{T}_{MRI}$ with respect to the world reference frame, *c.f.* Eq. (7). Optionally, a second step consisting of an affine and deformable registration using the same similarity measure can be performed, resulting in a precise and direct correspondence between voxels in the compounded 3D-US and reconstructed MRI volumes, which finally closes the loop of the autonomous ultrasound acquisition system and enabling MRI guided interventions. We evaluate both rigid and affine registrations; the decision whether both are required relies on the specific application in mind. For instance, if bony structures are to be scanned rather than soft tissue, a rigid alignment would be sufficient based on our experience.

5) Update of Patient Calibration: Based on the estimated transformations from the (robotically) acquired 3D ultrasound data and the MRI aligned to the world, the transformation from tomographic imaging to ultrasound space can be refined, in order to enable a more precise initialization of subsequent acquisitions. This becomes especially interesting as such acquisitions can be performed for an on-line refinement of the whole system calibration through an image-based feedback and update. In an interventional setup, an initial acquisition is performed at the beginning to optimize the robot- and patient-to-world calibrations. Subsequent planned trajectories can then be performed automatically based on an updated calibration. Making use of the estimated rigid transformation ${}^{P}\mathbf{T}_{MRI}$ aligning the US and MR volumes in world space as described above, the patient-to-world calibration can be updated accordingly to

$$ {}^{P}\mathbf{T}_{US} = {}^{P}\mathbf{T}_{MRI} \cdot {}^{MRI}\mathbf{T}_C \cdot {}^{C}\mathbf{T}_W \cdot {}^{W}\mathbf{T}_{US}. \quad (10) $$

On the one hand, this reduces processing time, as no rigid alignment is required for those US volumes. On the other hand, the comparability of subsequent records is fully maintained by this approach, as images are provided for the identical anatomy

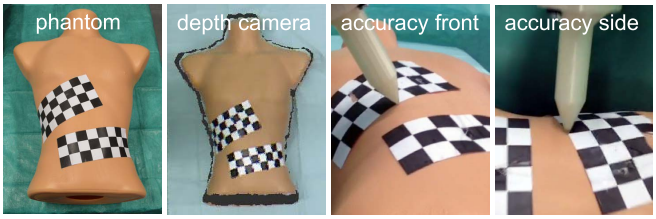

Fig. 5. From the colored point cloud (depth camera), individual checkerboard corner points on the uneven surface are manually selected as target points. The points can be targeted with an accuracy of 2.46 mm in x-y-direction, and 6.43 mm in z-direction.

and planned trajectory. For an exemplary application of US-guided liver biopsy, several acquisitions of the liver could be conducted throughout the procedure, enabling a reliable 3D image-guidance during needle insertion.

IV. Experiments and Results

A. Robot/Camera Calibration and Robot Control

Utilizing the depth images augmented with color information, we evaluate the accuracy of the tool and camera-to-world calibrations by equipping the robot manipulator with a rigid acuminate tool and moving the tool tip onto selected points on a realistic rigid surface. In this view, after tool configuration, the camera-to-robot calibration is performed to obtain the required transformations as described in Sec. III-C, where 13 poses are used in total for the calibration. Next, the RGB-D data of an upper torso phantom surface is recorded and the surface normals computed. The planned acquisition path is defined by manually selecting multiple intersection points on a printed checkerboard attached to the phantom's curved surface, see Fig. 5. The robot is commanded in position control mode to move onto the individual points and the distances to the actual intersection points are measured manually with a caliper. The experiment is performed twice by selecting 10 points in each run and recalibrating the whole system in between the two sessions. For the first run, the results yielded an average accuracy $|x_{set} - x_{msr}|$ of 2.42± 1.00 mm on the x-y plane and 7.20± 3.30 mm along the z-axis (camera's depth axis). For the second one, the accuracy was estimated as 2.5± 0.90 mm (x-y plane) and 5.64± 4.04 mm (z-axis). Overall, accuracies were 2.47 ± 0.96 mm (x-y) and 6.43 ± 3.68 mm (z). During all experiments, the camera was placed at around 1.5 m distance to the phantom. This shows that the calibration yields reproducible results, while the inaccuracies are dependent on the RGB-D camera information. It should be noted that these values are in line with the reported average spatial resolution of the camera, being 3.0 mm (x-y plane) and 10 mm (depth axis) for a camera at 2.0 m from its target [48]. It is also important to notice that while the $x - y$ accuracy directly affects the robot poses, the z errors are effectively compensated by the desired-force control of the robot. In this view, the resulting system layout can compensate for the inaccuracies of the RGB-D camera. Based on these accuracies, the calibration and overall system accuracy is anticipated to be sufficient for the initialization of an image-based registration, and thus also for the full system.

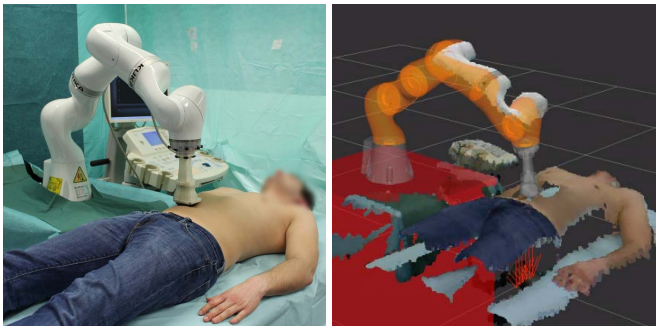

Fig. 6. Scanning setup showing the real system setup scanning a volunteer (left) as well as the visualization of all involved components accordingly with the point cloud given by the RGB-D camera.

B. MRI/US Image Acquisition and Registration

To allow for an evaluation of the overall system, we first use a triple-modality 3D abdominal phantom (Model 057A, Cirs Inc., Norfolk, Virginia, USA), which provides anatomically correct objects such as the liver, cysts, ribs and vertebrae. The phantom is targeted at MRI, CT, and ultrasound imaging and provides realistic image properties for all modalities. We then perform similar automatic robotic US acquisitions on two healthy volunteers, for whom an MRI scan was performed prior to this work.

Following the workflow presented, the MRI images are transformed into the world reference frame by matching the 3D scanned surface with the MRI surface (see Sec. III-C). Based on this global registration, trajectories are planned in the MRI by selecting start- and endpoint. Figure 6 shows the actual scanning setup for a healthy volunteer, where the robot arm performs an automatic acquisition based on a planned trajectory. Given a perfect system accuracy, i.e. a system with perfect calibration and without any tracking or imaging inaccuracies, an automatic acquisition of a 3D ultrasound volume by the robotic system would yield a perfect overlap between reconstructed 3D US data and tomographic imaging. In reality, however, the image alignment will be affected, which is why an intensity-based registration is then performed in order to detect the offset and correct for system inaccuracies. The alignment by intensity-based registration is first performed first rigidly (accounting for tracking, calibration, patient registration, and movement), followed by an affine improvement (accounting for image scaling as well as patient and tissue movement/deformation). Due to the fixed direction of the trajectory and a constant skin-pressure, an affine registration is sufficient for this application. The resulting rigid transformation directly indicates the accuracy of the overall 3D scanning system based on the alignment of planned and real acquisition (distance between desired and actual US). In a second step, we then use the rigid transformation part of the registration to refine the overall system and robot calibration according to Eq. (9). The experiment is repeated with a second planned trajectory, followed by an evaluation of rigid and affine registrations between the calibration-updated scans and the tomographic data. The hypothesis is that by using the rigid part of the intensity-based registration, we can align the pre-interventional MRI with the interventional

TABLE I
QUANTITATIVE RESULTS BASED ON REGISTRATION

	Translation [mm]	
	Phantom	Human
Rigid scan #1	4.69 ± 1.35	4.39 ± 2.47
Rigid scan #2	1.18 ± 0.52	0.89 ± 0.84
Affine scan #1	1.30 ± 1.06	0.92 ± 0.75
	Rotation [°]	
Rigid scan #1	3.66 ± 1.99	4.79 ± 2.42
Rigid scan #2	0.40 ± 0.32	1.30 ± 0.74
Affine scan #1	0.46 ± 0.19	1.37 ± 1.60

US data. This closes the loop between world calibration and robotic-system, where the system calibration can be refined online. In order to also evaluate the errors of the system in different situations, we repeat the overall experiment two times for the phantom, and between each experiment the phantom is moved on the examination bed. For the scans on volunteers, we repeat the overall experiment three times per person, where the person stands up and lies down between each run, evaluating also potential variabilities based on the patient-to-world calibration. Robotic acquisitions are taken at a similar speed as freehand acquisitions, such that a normal acquisition requires <30 s. Ultrasound scans are planned in the the abdominal region and contain organs (e.g. liver, kidney) and vessels. Acquisitions may also contain ribs, where the robot follows the natural curvature of the body surface without exceeding the maximal force applied. Ultrasound parameters are set manually, although modern US systems provide automatic adaption allowing for a direct optimization of settings. Coupling gel is also applied manually, and the volunteers must hold their breath during the scan. The internal subsystems incorporated in the KUKA lightweight system stop actions automatically if the desired force is exceeded, thus allowing for an safe evaluation in the constrained experimental environment. The patient-to-world calibration is performed once before each acquisition, but could be also updated online in future. The global alignment between the patient and MRI surface requires less than 10s, and the intensity-based refinement of the registration requires less than 30s to complete.

For all acquisitions of the phantom and humans combined, the average translation between the initial US-volume and MRI is 4.47 ± 2.15 mm for iteration one (without calibration-update) compared to 0.97 ± 0.76 mm in the second run after the overall patient calibration is updated. Similarly, the rotational part clearly improves after the update of the calibration ($4.50 \pm 2.24°$ before update, versus $1.08 \pm 0.76°$ after), where the rotational error is determined as the Euclidean norm of the three rotation angles. For an additional affine registration after the initial rigid registration, the average translation results in 1.02 ± 0.77mm, which shows that the initial registration successfully accounts for potential tracking and point cloud inaccuracies (rotation errors are 1.43 ± 1.42°). Resulting ultrasound datasets and registrations are depicted in Fig. 7 for a volunteer acquisition, and all results of the different experiments are listed in Table I.

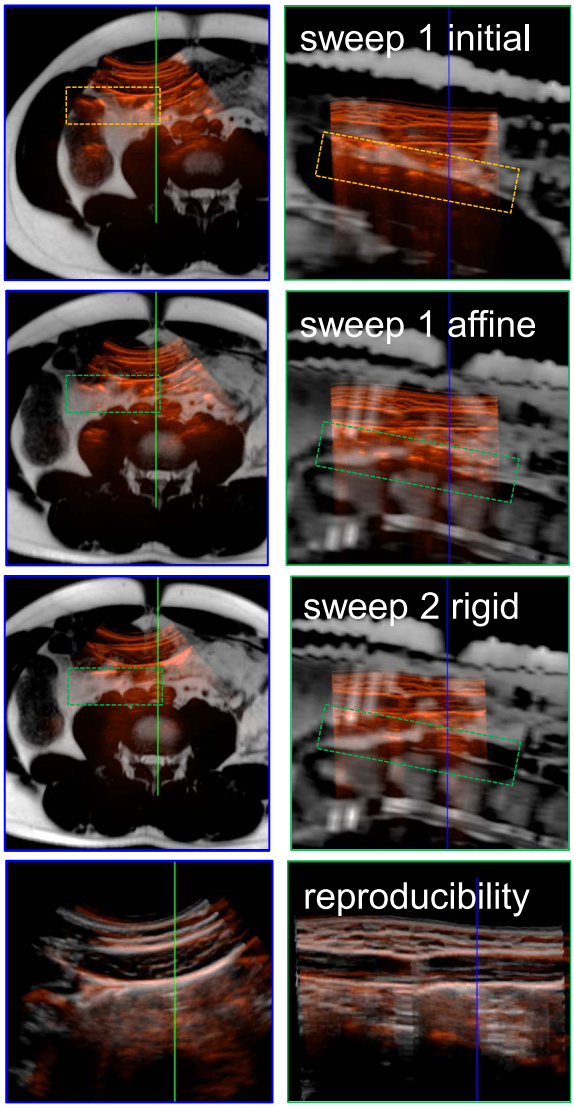

Fig. 7. Left column: axial plane; right column: saggital plane. The initial calibration shows partial misalignment (orange box), while the registrations for sweep #1 and #2 account for this offset. An overlay of both scans also shows the high reproducibility of the sweeps.

Thereby, initial, rigid and affine refer to the initial US scan, the registration based on the calibration update, and the affine registration after rigid registration, respectively.

V. DISCUSSION

Based on the millimeter accuracy achieved on repeated acquisitions with an updated calibration, our results suggest the feasibility of the overall approach as well as a potential path for a clinical integration of the presented system. The initial system-calibration using RGB-D information showed a maximum error below 1 cm, which ensures that the acquired 3D-US datasets will lie within the capture range of the intensity-based registration [47]. Therefore, the system should be able to deal with challenging clinical settings, where a higher deviation might lead to local minima during the registration optimization. As the position and orientation of

the acquired ultrasound images are aligned with the diagnostic image, US data can be correlated directly to the underlying tomographic image data, aiding the medical expert in terms of ultrasound visualization. The results of the intensity-based registration (Table I) show the overall translation/rotation of the US volume with respect to the MRI data, where the affine registration of the second scan appeared to perform worse than the rigid registration after calibration-refinement. While the quantity of data is too little to allow for a statistical meaningful judgement, the local deformation of tissue by the robotically held US probe is the primary reason for this discrepancy in our experience. In view of a specific clinical application in mind, it is thus important to analyze and adapt the utilized registration tools further, especially with respect to a deformable registration of soft tissue.

The integration of ROS and Sunrise.OS as used in this work allows for a full control of the KUKA APIs directly from components being available within ROS. Furthermore, a direct access to the robot state is provided by the KUKA APIs, including native features such as self collision avoidance. In this regard, the utilized approach allows for the best combination of both worlds. This facilitates the rapid development of new approaches, but also ensures safety of the robot environment, i.e. by collision detection and emergency halt features.

In view of a clinical integration, it should be noted that our results provide only a first feasibility evaluation, where future work clearly needs to focus not only on healthy volunteers but also diseased and pathological anatomy. Methods such as patient surface extraction and dynamic registration will require further adaptions, aiming at a clinical integration in the future. Our experiments also showed that the currently used examination bed is not optimal for the robot working space, as the end-effector is moving almost at the base of the robot. With respect to a clinical application, we thus suggest a height-adjustable bed, which could be directly integrated with the imaging system, such that the optimal height would be adjusted based on the planned acquisition. Besides that, acoustic coupling between the US probe and patient surface needs to be explored, facilitating US scans with either automatic gel application, or the exploration of other alternatives.

To this end, our future work will also focus on the possible online-optimization of the trajectory based on the initial planning, as we partially experienced suboptimal US image quality for the selected trajectories in this work. In this view, also an automatic change of the applied coupling force with respect to the target surface (e.g. fat vs. muscle) is analyzed in ongoing efforts.

VI. Conclusion

We have presented a path to an autonomous robotic ultrasound system in order to enable imaging and support during interventions. The set of methods presented shows the basic feasibility of the automatic scanning framework, allowing fast and efficient robotic 3D ultrasound acquisitions based on pre-interventional planned trajectories. On the foundation of an integrated system consisting of a baseline tomographic image with 3D RGB-D information, we automatically register patient data and perform automatic robotic acquisition. We introduced a closed-loop calibration update based on image-based registration to facilitate the acquisition of reproducible 3D US data. Our results show that the overall accuracy of the system is sufficient for clinical applications. Despite challenges which need to be overcome before such systems could be used in daily routine, this work will hopefully facilitate the clinical acceptance of automatic and autonomous robotic ultrasound scanning systems in the future.

Acknowledgment

The authors thank the department of nuclear medicine at Klinikum rechts der Isar for the MRI acquisitions, as well as ImFusion GmbH for the provided registration software.

References

[1] L. Beales, S. Wolstenhulme, J. A. Evans, R. West, and D. J. A. Scott, "Reproducibility of ultrasound measurement of the abdominal aorta," *Br. J. Surg.*, vol. 98, no. 11, pp. 1517–1525, 2011.

[2] A. Gee, R. Prager, G. Treece, and L. Berman, "Engineering a freehand 3D ultrasound system," *Pattern Recognit. Lett.*, vol. 24, nos. 4–5, pp. 757–777, 2003.

[3] O. V. Solberg, F. Lindseth, H. Torp, R. E. Blake, and T. A. N. Hernes, "Freehand 3D ultrasound reconstruction algorithms—A review," *Ultrasound Med. Biol.*, vol. 33, no. 7, pp. 991–1009, 2007.

[4] M. J. Monaghan, "Role of real time 3D echocardiography in evaluating the left ventricle," *Heart*, vol. 92, no. 1, pp. 131–136, 2006.

[5] H. N. Cardinal, J. D. Gill, and A. Fenster, "Analysis of geometrical distortion and statistical variance in length, area, and volume in a linearly scanned 3-D ultrasound image," *IEEE Trans. Med. Imag.*, vol. 19, no. 6, pp. 632–651, Jun. 2000.

[6] A. Krupa, G. Fichtinger, and G. D. Hager, "Real-time motion stabilization with B-mode ultrasound using image speckle information and visual servoing," *Int. J. Robot. Res.*, vol. 28, no. 10, pp. 1334–1354, Oct. 2009.

[7] P. De Rango, "Aneurysm diameter measurement: A challenging and frustrating task," *Eur. J. Vascular Endovascular Surg.*, vol. 43, no. 1, p. 34, 2012.

[8] A. Fenster, D. B. Downey, and H. N. Cardinal, "Three-dimensional ultrasound imaging," *Phys. Med. Biol.*, vol. 46, no. 5, p. R67, 2001.

[9] A. M. Priester, S. Natarajan, and M. O. Culjat, "Robotic ultrasound systems in medicine," *IEEE Trans. Ultrason., Ferroelect., Freq. Control*, vol. 60, no. 3, pp. 507–523, Mar. 2013.

[10] P. Chatelain, A. Krupa, and N. Navab, "3D ultrasound-guided robotic steering of a flexible needle via visual servoing," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 2250–2255.

[11] M. Esposito *et al.*, "Cooperative robotic gamma imaging: Enhancing US-guided needle biopsy," in *Medical Image Computing and Computer-Assisted Intervention*. New York, NY, USA: Springer, 2015, pp. 611–618.

[12] R. Kojcev *et al.*, "Dual-robot ultrasound-guided needle placement: Closing the planning-imaging-action loop," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 11, no. 6, pp. 1173–1181, Jun. 2016.

[13] R. Nakadate *et al.*, "Development of a robot assisted carotid blood flow measurement system," *Mech. Mach. Theory*, vol. 46, no. 8, pp. 1066–1083, 2011.

[14] S. Billings, N. Deshmukh, H. J. Kang, R. Taylor, and E. M. Boctor, "System for robot-assisted real-time laparoscopic ultrasound elastography," *Proc. SPIE*, vol. 8316, p. 83161W, Feb. 2012.

[15] C. Schneider, J. Guerrero, C. Nguan, R. Rohling, and S. Salcudean, "Intra-operative 'pick-up' ultrasound for robot assisted surgery with vessel extraction and registration: A feasibility study," in *Information Processing in Computer-Assisted Interventions*. New York, NY, USA: Springer, 2011, pp. 122–132.

[16] F. Conti, J. Park, and O. Khatib, "Interface design and control strategies for a robot assisted ultrasonic examination system," in *Experimental Robotics*. New York, NY, USA: Springer, 2014, pp. 97–113.

[17] E. M. Boctor, G. Fischer, M. A. Choti, G. Fichtinger, and R. H. Taylor, "A dual-armed robotic system for intraoperative ultrasound guided hepatic ablative therapy: A prospective study," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 3. Apr. 2004, pp. 2517–2522.

[18] E. M. Boctor, M. A. Choti, E. C. Burdette, and R. J. Webster, III, "Three-dimensional ultrasound-guided robotic needle placement: An experimental evaluation," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 4, no. 2, pp. 180–191, 2008.

[19] D. Zhang, Z. Li, K. Chen, J. Xiong, X. Zhang, and L. Wang, "An optical tracker based robot registration and servoing method for ultrasound guided percutaneous renal access," *Biomed. Eng. Online*, vol. 12, no. 1, p. 47, 2013.

[20] C. Graumann, B. Fuerst, C. Hennersperger, F. Bork, and N. Navab, "Robotic ultrasound trajectory planning for volume of interest coverage," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2016, pp. 736–741.

[21] P. Abolmaesumi, S. Salcudean, and W. H. Zhu, "Visual servoing for robot-assisted diagnostic ultrasound," in *Proc. 22nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 4. Jul. 2000, pp. 2532–2535.

[22] A. Krupa, "3D steering of a flexible needle by visual servoing," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. New York, NY, USA: Springer, 2014, pp. 480–487.

[23] P. Chatelain, A. Krupa, and N. Navab, "Optimization of ultrasound image quality via visual servoing," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2015, pp. 5997–6002.

[24] O. Zettinig *et al.*, "Toward real-time 3D ultrasound registration-based visual servoing for interventional navigation," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2016, pp. 945–950.

[25] F. A. Fröhlich, G. Passig, A. Vazquez, and G. Hirzinger, "Robot assisted internal mammary artery detection for coronary revascularisation surgery," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2010, pp. 1849–1855.

[26] F. Chaumette and S. Hutchinson, "Visual servo control. I. Basic approaches," *IEEE Robot. Autom. Mag.*, vol. 13, no. 4, pp. 82–90, Dec. 2006.

[27] F. Chaumette and S. Hutchinson, "Visual servo control. II. Advanced approaches," *IEEE Robot. Autom. Mag.*, vol. 14, no. 1, pp. 109–118, Mar. 2007.

[28] S. Onogi, T. Yoshida, Y. Sugano, T. Mochizuki, and K. Masuda, "Robotic ultrasound guidance by B-scan plane positioning control," in *Proc. CIRP*, vol. 5. Dec. 2013, pp. 100–103.

[29] F. Aalamifar, R. Khurana, A. Cheng, R. H. Taylor, I. Iordachita, and E. M. Boctor, "Enabling technologies for robot assisted ultrasound tomography: System setup and calibration," *Proc. SPIE*, vol. 9040, p. 90401X, Apr. 2014.

[30] S. Merouche, L. Allard, E. Montagnon, G. Soulez, P. Bigras, and G. Cloutier, "A robotic ultrasound scanner for automatic vessel tracking and three-dimensional reconstruction of B-mode images," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 63, no. 1, pp. 35–46, Jan. 2016.

[31] A. S. B. Mustafa *et al.*, "Development of robotic system for autonomous liver screening using ultrasound scanning device," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2013, pp. 804–809.

[32] G. Hirzinger *et al.*, "DLR's torque-controlled light weight robot III—Are we reaching the technological limits now?" in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 2. May 2002, pp. 1710–1716.

[33] R. Bischoff *et al.*, "The KUKA-DLR lightweight robot arm—A new reference platform for robotics research and manufacturing," in *Proc. ISR*, Jun. 2010, pp. 1–8.

[34] G. Schreiber, A. Stemmer, and R. Bischoff, "The fast research interface for the KUKA lightweight robot," in *Proc. IEEE Workshop Innov. Robot Control Archit. Demanding Res. Appl. (ICRA)*, May 2010, pp. 15–21.

[35] L. C. Ebert, G. Hatch, G. Ampanozi, M. J. Thali, and S. Ross, "You can touch this," *Surgical Innov.*, vol. 19, no. 3, pp. 301–307, 2012.

[36] C.-H. Hsieh, C.-H. Huang, and J.-D. Lee, "A non-contact image-to-patient registration method using Kinect sensor and WAP-ICP," in *Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*. New York, NY, USA: Springer, 2013, pp. 95–102.

[37] S. Billings, A. Kapoor, M. Keil, B. J. Wood, and E. Boctor, "A hybrid surface/image-based approach to facilitate ultrasound/CT registration," *Proc. SPIE*, vol. 7968, p. 79680V, Mar. 2011.

[38] L. Mercier, T. Langø, F. Lindseth, and L. D. Collins, "A review of calibration techniques for freehand 3-D ultrasound systems," *Ultrasound Med. Biol.*, vol. 31, no. 2, pp. 143–165, 2005.

[39] S. Kahn, D. Haumann, and V. Willert, "Hand-eye calibration with a depth camera: 2D or 3D?" in *Proc. Int. Conf. Comput. Vis. Theory Appl. (VISAPP)*, vol. 3. Jan. 2014, pp. 481–489.

[40] R. Y. Tsai and R. K. Lenz, "Real time versatile robotics hand/eye calibration using 3D machine vision," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 1988, pp. 554–561.

[41] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognit.*, vol. 47, no. 6, pp. 2280–2292, Jun. 2014.

[42] D. Herrera C, J. Kannala, and J. Heikkilä, "Joint depth and color camera calibration with distortion correction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 10, pp. 2058–2064, Oct. 2012.

[43] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2004.

[44] J. Elseberg, D. Borrmann, and A. Nüchter, "One billion points in the cloud—An octree for efficient processing of 3D laser scans," *ISPRS J. Photogram. Remote Sens.*, vol. 76, pp. 76–88, Feb. 2013.

[45] J. K. Salisbury, "Active stiffness control of a manipulator in Cartesian coordinates," in *Proc. 19th IEEE Conf. Decision Control Including Symp. Adapt. Process.*, Dec. 1980, pp. 95–100.

[46] C. Hennersperger, A. Karamalis, and N. Navab, "Vascular 3D+T freehand ultrasound using correlation of Doppler and pulse-oximetry data," in *Information Processing in Computer-Assisted Interventions*. Springer, 2014, pp. 68–77.

[47] B. Fuerst, W. Wein, M. Müller, and N. Navab, "Automatic ultrasound–MRI registration for neurosurgery using the 2D and 3D LC^2 metric," *Med. Image Anal.*, vol. 18, no. 8, pp. 1312–1319, Dec. 2014.

[48] M. R. Andersen *et al.*, "Kinect depth sensor evaluation for computer vision applications," *Tech. Rep. Electron. Comput. Eng.*, vol. 1, no. 6, Jul. 2015.

5.2 3D Deformation Correction Leveraging Force Sensing (IJCARS 2018)

The need for a tight contact between an ultrasound transducer and the patient has already been discussed in Chapter 2: lack of contact would result in a loss of ultrasound signal since it cannot propagate in air. Therefore, force control techniques (Chapter 3) are required in robotic ultrasound systems to be able to apply the required contact force and, in an optimal implementation, to maintain that force constant. This applied force, therefore, impacts the resulting ultrasound images of any anatomy. That is, the tissue undergoes a certain deformation due to this force, the acquired images reflect this deformation. Using a constant force during a 3D acquisition enables to limit the variability in deformation along the performed trajectory, but an overall deformation is still present. In many cases, the accurate localization and the geometric analysis of anatomical structures can be impaired by these deformations, especially in the case of 3D reconstructed volumes, using methods from Sec. 2.3.

In this work, a method to estimate and partially correct these deformations is proposed, targeting the deformation recovery of an entire 3D ultrasound volume. The 2D tissue displacement (along the image plane) is estimated at sampled locations along an executed trajectory. This process uses an image-based tracking approach that relies on the registration of subsequent images acquired at forces that are increasing over time. Information on how the applied force deforms each pixel is then obtained sparsely along the robot movements, and these sparse measurements are then represented using a 3D graph. An exhaustive information regarding the overall 3D volume is completed expanding the available data using an anisotropic diffusion approach along the graph structure. A high diffusion is allowed along the direction of the performed trajectory, while a small one is enforced along the image plane, to avoid blurring. The output of the method is a full 3D *un*-deformation field, providing information on how each voxel in the volume should move if a different (or no force) would have been applied instead.

An extensive validation of the method is also presented. A public dataset of 30 volumes acquired on human volunteers is created, including ground truth undeformed volumes acquired without any contact to the patient's surface (using the equivalent of a water bath). The proposed technique is able, when using a $5\,\mathrm{N}$ force for the acquisition, to regenerate 3D volumes that have an error of $3.39(186)\,\mathrm{mm}$ with respect to the relative ground truth. Note that the method relies on all the components of the system presented in Sec. 5.1, without requiring the use of additional hardware or complex bio-mechanical models.

Use the Force: Deformation Correction in Robotic 3D Ultrasound

Salvatore Virga[1], Rüdiger Göbl[1], Maximilian Baust[1], Nassir Navab[1,2], Christoph Hennersperger[1]

[1] Chair for Computer Aided Medical Procedures (CAMP), Technische Universität München, Munich, Germany.
[2] Johns Hopkins University, Baltimore, USA.

Contribution. The main contributions of this publications, including the design and implementation of the software used, the acquisition of results from healthy volunteers, their analysis and the writing of the manuscript were done and coordinated by the author of this thesis. Rüdiger Göbl and Maximilian Baust were responsible for the implementation of the image-based tracking pipeline. The revision of the publication was performed jointly together with all the co-authors.

International Journal of Computer Assisted Radiology and Surgery
https://doi.org/10.1007/s11548-018-1716-8

ORIGINAL ARTICLE

Use the force: deformation correction in robotic 3D ultrasound

Salvatore Virga[1] · Rüdiger Göbl[1] · Maximilian Baust[1] · Nassir Navab[1,2] · Christoph Hennersperger[1]

Received: 27 January 2018 / Accepted: 20 February 2018

Abstract

Purpose Ultrasound acquisitions are typically affected by deformations due to the pressure applied onto the contact surface. While a certain amount of pressure is necessary to ensure good acoustic coupling and visibility of the anatomy under examination, the caused deformations hinder accurate localization and geometric analysis of anatomical structures. These complications have even greater impact in case of 3D ultrasound scans as they limit the correct reconstruction of acquired volumes.

Methods In this work, we propose a method to estimate and correct the induced deformation based solely on the tracked ultrasound images and information about the applied force. This is achieved by modeling estimated displacement fields of individual image sequences using the measured force information. By representing the computed displacement fields using a graph-based approach, we are able to recover a deformation-less 3D volume.

Results Validation is performed on 30 in vivo human datasets acquired using a robotic ultrasound framework. Compared to ground truth, the presented deformation correction shows errors of 3.39 ± 1.86 mm for an applied force of 5 N at a penetration depth of 55 mm.

Conclusion The proposed technique allows for the correction of deformations induced by the transducer pressure in entire 3D ultrasound volumes. Our technique does not require biomechanical models, patient-specific assumptions or information about the tissue properties; it can be employed based on the information from readily available robotic ultrasound platforms.

Keywords Robotic ultrasound · Deformation correction · Compounding · Inpainting

Introduction

Ultrasound (US) imaging is extensively used in both diagnostic and interventional scenarios. The high accessibility of US systems and their low costs make US the modality of choice to obtain real-time imaging of various anatomies. In the case of soft tissue applications and subcutaneous pathologies, such as soft tissue sarcomas (STS), US yields morphological information about the target tumor [15], e.g., tumor heterogeneity or vascularization. Being free of ionizing radiation, US is also an ideal imaging modality for follow-up acquisitions after surgical interventions, such as in the case of STS where there is a high risk of recurrence of STS [1].

To achieve optimal acoustic coupling of the US transducer and obtain good visibility of the target anatomy, it is necessary to apply a certain force onto the contact surface. This naturally results in an unavoidable deformation of the imaged tissues. For certain clinical applications, the induced deformation can be exploited to infer mechanical properties of the imaged tissue by means of a static or dynamic excitation—a concept exploited in classical US elastography [12]. On the other hand, this deformation can hamper the exact localization of anatomies of interest, such as STS tumor masses, as well as the precise assessment of their geometry and volume. Additionally, as these deformations do not usually remain constant during free-hand 3D ultrasound acquisitions, different deformations across the individual 2D images of an US sweep eventually impair the volumetric reconstruction. Despite the advantages free-hand 3D-US could provide in several clinical settings, in both diagnostic [13] and interventional scenarios [11], these drawbacks impair a higher acceptance of such approaches in practice. To overcome

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11548-018-1716-8) contains supplementary material, which is available to authorized users.

Salvatore Virga
salvo.virga@tum.de

[1] Technical University of Munich, Boltzmannstr. 3, 85748 Garching bei München, Germany

[2] Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218, USA

Published online: 02 March 2018

these deformation-induced problems, the inhomogeneous deformation created in free-hand 3D US is reconstructed in [18] through a combination of probe tracking information and non-rigid image-based registration, although limited to the deformation in the axial direction. Similarly, in [2,21] nonlinear tissue distortions produced during long 3D US acquisitions are corrected via subsequent elastic registrations of image pairs. All these methods, however, do not aim at removing the tissue deformation, but to make it homogeneous along the entire US volume.

Recently, robotic platforms for US imaging (rUS) have been proposed [5,17]. It is common for these platforms to include force sensing, either via force and torque sensors attached to the robot end-effector, or coupled with elastic joints, such as in [8]. Therefore, contrary to free-hand US, such platforms are able to apply a constant pressure onto the examined anatomy and thus maintain a nearly constant deformation, which significantly alleviates volumetric reconstruction. However, the deformation will still be present due to the necessary contact force. This is of particular interest in an interventional scenario, when registering preoperatively acquired tomographic image data (e.g., MRI or CT) to intraoperatively acquired 3D US data, as the amount of deformation between the two volumes is directly proportional to the required computational effort and thus precious runtime.

Related work

Existing methods for correcting for the induced deformation in US make use of a combination of various types of information such as the US image data, the position of the US transducer (via a tracking system) and the applied force. A group of proposed works additionally employs biomechanical models that reproduce the mechanical properties of the tissue under examination. These include models based on the finite element method (FEM) to predict and correct for the deformation [3]. However, FEM-based methods require a priori knowledge of the tissue properties to obtain an accurate representation and often require specific density, elastic moduli, or similar. In practice, such indices are complex to be retrieved, especially in pathological indications. Besides FEM modeling, 3D mesh models are also used to estimate the deformation [14]. Although these models are not patient-specific itself, the method requires the acquisition of the surface of interest during the procedure, using laser scan technology or similar techniques. Furthermore, both methods assume that deformations are present only along the axial direction of the US transducer. In [4], a proof of concept is proposed to obtain mechanical parameters specific to the examined tissue using mutual information between the US images, avoiding the use of generalized parameters. Finally, a recent method proposed in [16] solely uses tracked US images and the applied force in order to extrapolate the tissue deformation and eventually recover deformations in the axial and lateral direction. However, both methods are presented for 2D images only and validated only on simulations.

Contributions

We present an image-based approach to obtain compression-free 3D US volumes based on robotic acquisitions (see Fig. 1). We do not employ a technique based on biomechanical models, as these require the retrieval of additional parameters that impairs their usability, especially since a patient-specific model is required. It is noticeable, in fact, that these limitations are reflected in the validation of such methods, which is very often performed on synthetic data only—both regarding tissue models and US images—except [14], which is validated on one clinical case.

The robotic platform presented in the following facilitates accurate tracking of the US transducer position and provides control over the applied force via direct force control techniques. The latter ensures that the necessary force to visualize the anatomy of interested is applied as well as that information about the applied force can be obtained along the whole scan trajectory. With regard to the mentioned state of the art, our proposed method features the following aspects:

(a) it makes use of the tracked 2D US images and the force information only, i.e., no additions to a generic robotic US setup are required;
(b) it is able to recover deformations in both axial and lateral direction;
(c) it uses a novel deformation interpolation (3D inpainting) of sparsely measured elastic information to retrieve a full deformation-corrected 3D US volume.

We provide validation for our method on 30 3D acquisitions performed on volunteers and evaluate the corrected volumes against ground truth US data. Table 1 categorizes this work and the related state of art according to their characteristics. Finally, to allow a better comparison of different methods for deformation correction and to improve reproducibility, the human acquisitions that were performed to validate this work are publicly available.

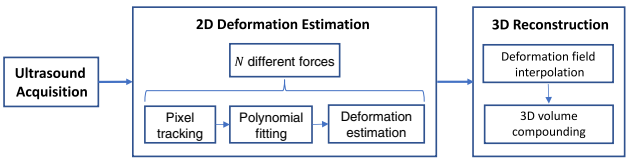

Fig. 1 Proposed workflow for deformation correction of 3D US volumes. For an acquisition of length L, K 2D deformation estimations are performed

Table 1 Comparison with related work

	2D	3D	Axial def.	Lateral def.	Force sensing	No biomech. model	In vivo validation
Burcher et al. [3]	✓		✓	✓	✓		
Pheiffer et al. [14]	✓	✓	✓				✓
Dahmani et al. [4]	✓		✓	✓			
Sun et al. [16]	✓		✓	✓	✓	✓	
Ours	✓	✓	✓	✓	✓	✓	✓

Categorized based on the characteristics of available methods: deformation correction in 2D, in 3D, ability to cope with axial and/or lateral deformations, use of force sensing and/or biomechanical models and in vivo validation

Methods

The aim of the presented method is to obtain a 3D US volume free of the deformation induced by the contact force from the US transducer during the acquisition performed by a rUS platform. Following an image-based approach, this requires to correct for the induced deformation of each individual 2D image that forms the 3D volume, which would imply extensive acquisition and computation time depending on the length of the trajectory. To maintain high clinical usability, we propose to perform the 2D deformation estimation sparsely along the planned trajectory, while providing a complete 3D deformation correction employing a novel 3D-inpainting scheme. A schematic representation of the overall workflow is shown in Fig. 1. Initially, 2D US images are acquired along a linear trajectory of length L. During the whole acquisition, the position of the US transducer is tracked via the forward kinematics of the robotic manipulator, as well as the fixed base force $\mathbf{F_{base}}$ applied onto the surface. At each designated position, we estimate the 2D deformation induced by the transducer pressure as described in "2D deformation estimation" section. This estimation is performed at K equidistant locations (separated by a distance of $L/(K-1)$). Then, the estimated sparse 2D deformation fields are inpainted using a graph-based representation of the whole US sweep in order to obtain a volumetric deformation field. Using this approach allows for filling in the missing information where the direct estimation was not performed. Finally, the undeformed 3D volume is reconstructed as described in "3D reconstruction" section.

2D deformation estimation

While the transducer is moved along the planned trajectory and the base force $\mathbf{F_{base}}$ is applied, for each of the K locations, the tissue deformation due to the applied pressure is examined. Inspired by the approach proposed by Sun et al. [16], 2D deformation fields are generated from a series of 2D images acquired at the same location with N different forces $\mathbf{F_i} \in [\mathbf{F_{start}}, \mathbf{F_{end}}]$ where $i = 1, \ldots, N$. These contact forces are increased by $\mathbf{F_{step}}$ after a small temporal interval t_{step}, s.t.

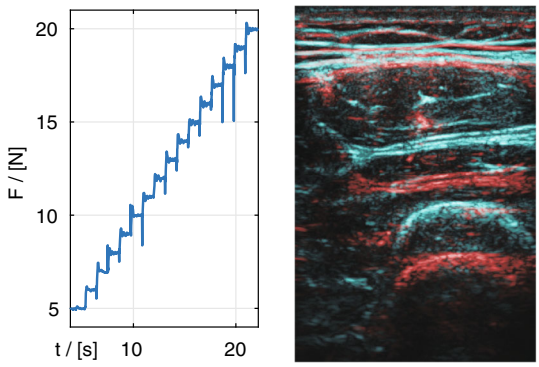

Fig. 2 Exemplary force profile and induced deformation. Forces applied during 2D deformation estimation (left), resulting force-dependent deformed US images for 2 N in red, and 8 N in blue (right)

$$\mathbf{F_i} = \mathbf{F_{start}} + (i-1)\mathbf{F_{step}}, \tag{1}$$

yielding N force-dependent 2D images I_i $(i = 1, \ldots, N)$ as illustrated in Fig. 2. We retrieve the resulting deformation between I_i and I_{i+1} at each pixel using a preconditioned fluid-elastic diffeomorphic demons as described in [19,20] with the following parameter settings: standard deviation for elastic demons $\sigma_e = 1$, standard deviation for fluid demons $\sigma_f = 1$ and step size $\tau = 0.05$. The parameters for the diffeomorphic demons algorithm have been chosen to reach a meaningful compromise between speckle size and degree of desired regularity: while increasing both standard deviations decreases the contributions of individual speckle patterns, too small standard deviations do not achieve satisfactory regularization results. In addition to this, it needs to be considered that by adjusting σ_e the regularity of the entire solution, i.e., the deformation field, can be changed and by adjusting σ_f one can influence the regularity of the individual updates, i.e., how susceptible the algorithm is to spurious and noisy local deformations.

To finally allow for a continuous modeling of the expected tissue displacements as a function of the applied force, a regression function is fitted to the sampled points. In our case, we model the force-dependent deformations using fourth-order polynomial functions in order to regress the deformation field corresponding to the force-free configu-

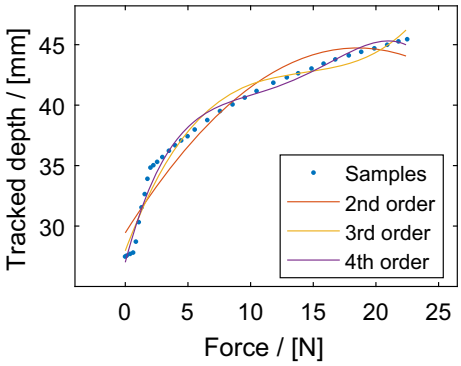

Fig. 3 Qualitative comparison of model quality. The evaluation shows the position of a single tracked sample for all measured incremental force steps in reference to evaluated models based on first-, second- and third-order polynomial

ration, i.e., at zero force. It is important to note that while the method in [16] utilizes median filtering on the obtained deformation fields to guarantee spatial consistency, the proposed reconstruction in "3D reconstruction" section ensures this property in an implicit manner. In conclusion, we obtain a model of the tissue deformation with respect to the applied force. Since each 3D acquisition is performed using F_{base} (e.g., 5 N)—with the exception of the location where the 2D estimation is computed—we will use the regression model to estimate the position at zero force for each volume voxel. Examples of such model obtained by pixel tracking are shown in Figs. 3 and 4.

3D reconstruction

As the 2D deformation estimation takes most of the time during a 3D US acquisition (see "Acquisitions" section) and acquiring them in a dense sampling is impractical, it is performed at K equidistant sampling positions along the scanning trajectory in order to facilitate the proposed method to be applied in a clinical scenario. Thus, the obtained 2D deformation fields have to be propagated or inpainted at the image position where no sampling has been performed. Inpainting can be described as the process of augmenting a set containing missing values with ones based on the surrounding known samples. There are several mathematical approaches to model inpainting, one of which is the solution of a set of partial differential equations (PDEs), where the known values represent the Dirichlet boundary conditions.

It is possible to solve discrete PDE-like problem with a graph-based method as the one proposed by Hennersperger et al. [9], which was introduced for the solution of a quadratic optimization problem. Differently from the original formulation, our inpainting problem requires the generation of new graph nodes to interpolate the missing information. We choose anisotropic diffusion of the deformation fields over the graph structure, such that we can better define the required diffusivity properties for our problem, i.e., high diffusion in the elevational direction and low speed of information propagation in axial and lateral direction. That is, we want the deformations to propagate more over the sweep direction than within the image plane. This can be efficiently implemented over the irregular graph that is built by the method, i.e., the optimization over the graph is performed as local operation. The irregular nature of the graph also allows us to employ non-parallel acquisitions.

For our problem, the boundary conditions (the known values) are the sparse 2D deformation fields. Inpainting is performed over the deformation fields component-wise using a preconditioned conjugate gradient (PCG) with a Jacobi pre-

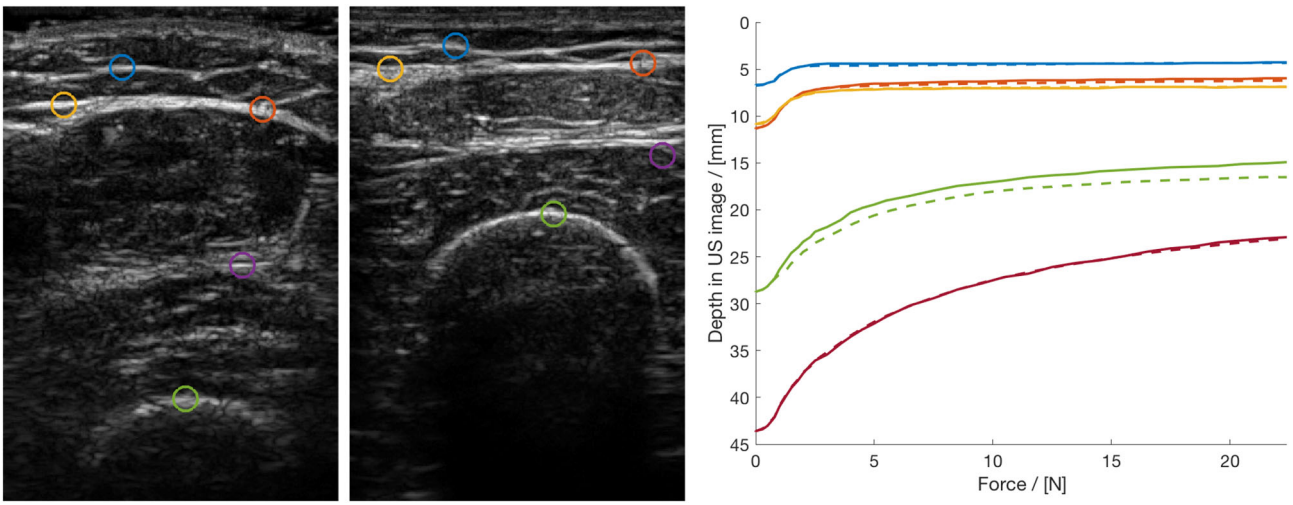

Fig. 4 2D deformation tracking: 5 pixels are tracked—manually and using the demons-based approach—while forces from F_{start} to F_{end} are applied. The points are marked on the first (left) and last (center) frame of the sequence. On the right, their position in axial direction is displayed: manually tracked trajectory (full line) and tracked via demons (dashed line)

conditioner as optimizer. Weights inside the images were fixed to e^{-3} to inhibit blurring within them, while weights between different images were set as described by the original authors [9], applied to the axial and lateral deformation components.

This graph-based inpainting results in a final deformation field that describes how each pixel along the acquired volume is affected by the applied force (forward deformation field). We invert this field to correct for the deformation induced by $\mathbf{F_{base}}$ during the acquisition. Inversion of an explicit field is not possible in the general case, but due to the employed tracking scheme, the resulting field is diffeomorphic, i.e., invertible. It is inverted numerically using an iterative method based on an open-source implementation.[1] After applying the inverse deformation field to the individual images from a full acquisition, a 3D volume is created using an US reconstruction method. We employed a voxel-based interpolation method, in which for each voxel a scalar is computed as weighted interpolation of the relevant US image samples. This approach is described in [7] for advanced reconstruction using tensors instead of scalar values. The resulting undeformed 3D volume thus appears as acquired using no force.

Hardware setup and experiments

Hardware setup

We make use of a robotic platform for autonomous US acquisitions: the system is composed of a manipulator, a KUKA LBR iiwa R800 (KUKA Roboter GmbH, Augsburg, Germany), controlled using the ROS[2] framework. The robotic arm is equipped with torque sensors at its joints, which allow for the estimation of the forces and torques applied by (or to) the robot's end-effector. B-mode US images are acquired using an Ultrasonix RP US machine (BK Ultrasound, Peabody, MA, USA) and a linear transducer (frequency: 3.3 MHz, depth: 55 mm, gain: 50%) which is attached to the robot's flange. Images are transferred using the OpenIGTLink communication protocol[3] to a workstation (Intel Core i7, NVIDIA GTX 1080), where they are synchronized to the transducer tracking and force stream.

Acquisitions

3D US volumes were acquired on the thighs of five healthy volunteers (age 26–30, 4 males, 1 female). We selected the volunteers' thighs as the area for our experimental acquisitions, since extremities are prone to be affected by soft tissue pathologies, such as STS [10], and therefore often subject to US examination. For each volunteer, six volumetric acquisitions of length $L = 70$ mm were performed, for an average of 26 s each and a total of 30 US volumes. Different base forces, applied orthogonally to the contact surface, were employed during the different US sweeps: $\mathbf{F_{base}} \in \{2, 5, 8, 10, 12, 15\}$ Newton. A 2D deformation field, as described in "2D deformation estimation" section, was estimated for $K = 15$ positions along the planned trajectory, with $\mathbf{F_{start}} = 0$ N and $\mathbf{F_{end}} = 20$ N. $\mathbf{F_{step}}$ was varying between 0.25 and 1 N during a single estimation, since smaller variation steps were found beneficial to better capture deformations of the initial tissue layers at lower forces. That is, those layers tend to undergo high deformations with low forces already, using small steps at the beginning of the evaluation allows to better track pixel movements. The required computation time to compute a 2D deformation field from the images acquired at one location was in average of 186 s, with a maximum memory usage of approximately 21 GB for the computations of a full acquisition.

A ground truth volume, free from any compression from the US transducer, was also acquired for every individual by maintaining the probe at about 5 mm from the contact surface while applying a thick layer of US gel to guarantee acoustic coupling. The precise and stable movement required for the ground truth volume acquisition is made possible by the use of such robotic system. It is important to notice that previous works on the topic do not validate their result using real undeformed US images or volumes, but rather using synthetic ones or via registration with other image modalities. On the other hand, due to the missing contact to the patient surface, while the ground truth volumes do not present deformations, the visibility of the underlying anatomy is strongly impaired.

Validation

To validate the framework for deformation-compensation, we assess the quality of the individual components as well as the overall system:

- **2D deformation estimation** We validate the pixel tracking and deformation regression on 5 deformation sequences. For each, 5 points are manually selected on the first frame and their displacement tracked over the successive ones. The position of the same points is also manually annotated, such that the absolute accuracy of the automatically tracked trajectory can be compared.
- **Deformation field interpolation** To evaluate our proposed sparse sampling scheme, we validate how sparsely a 2D deformation estimation can be performed while still obtaining a valuable volumetric deformation correction, as a trade-off between quality and acquisition time is needed. Based on 15 deformation fields estimated per

[1] https://itk.org.

[2] http://www.ros.org.

[3] http://openigtlink.org/.

volume, we first compare different subsamplings (leave-one-out) and compute the Euclidean norm of the resulting difference in deformation for 18 volumes. Second, for a given volume, we exclude the computed central deformation field and interpolate the remaining ones to obtain it. We perform this 7 times, incrementally removing more neighboring fields, until only the first and last sample are used for interpolation. The resulting field in the central location of the volume is then compared to the one directly computed by the (ground truth) 2D estimation.

– **3D volume undeformation** We validate the quality of the overall method using the target registration error (TRE) between the compounded US ground truth volume and the final undeformed volumes acquired applying different forces (2, 5, 8 and 15 N).

Results

2D Deformation Estimation

The 2D deformation estimation, as performed along the planned US trajectory ("2D deformation estimation" section), allows to track the displacement induced by the applied force of the individual image pixels. As can be seen in Fig. 3, the computed displacement is characterized by a nonlinear behavior, with a stronger deformation at lower forces due to the compression of the superficial and more elastic tissue layers. Therefore, while in [16] the authors propose to model this displacement using quadratic functions, we instead propose to use fourth-order polynomials to better capture the high flexibility of the subcutaneous and other relevant tissue layers.

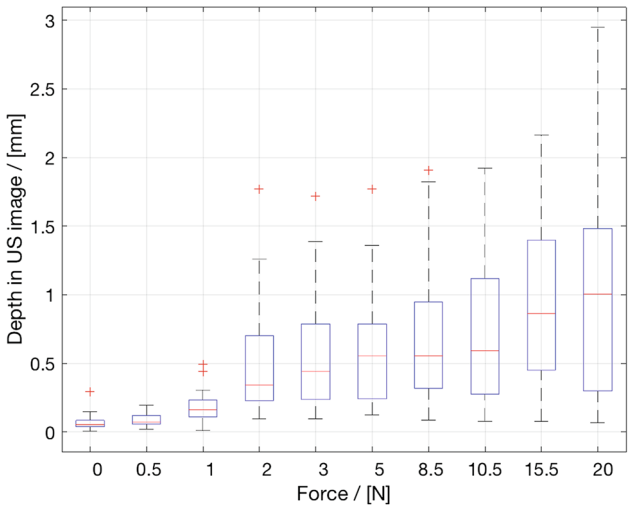

Fig. 5 2D tacking error: Mean and standard deviation of the Euclidean error [mm] between demons-based and manual tracking. 5 points over 5 sequences (25 points in total) of 35 force steps were evaluated, a subsample of the results (10 force values of 35) is shown

In Fig. 4, the 5 points selected for validation of one specific sequence are shown at the beginning and at the end of the deformation sequence together with the resulting models obtained from the demon-based tracking and a ground truth. For the 25 points selected over 5 different sequences, the error between their computed trajectory and the ground truth was found to be 0.64 ± 0.57 mm. Note that information on the tissue state at 0 N is already sampled in our model, so that we do not need to extrapolate to reach the undeformed state. In Fig. 5, the distribution of modeling errors is depicted for the evaluated force steps. The error from our tracking approach tends to accumulate over multiple force steps, with some sharp increases at instants where tissue layers yield to the increasing pressure.

Deformation field inpainting

The results of the validation of our inpainting strategy using different subsampling of the available deformation fields are presented in Fig. 7. For displacement fields sampled at a distance of 35 mm, an average error of 1 mm is obtained. The possibility to sample tissue deformation so sparsely also helps reducing the computational costs of a full acquisition and the clinical feasibility of the proposed method. We also validate the accuracy of the inpainting method computing a known displacement field that is not used during interpolation, together with a subset of its neighbors. In Fig. 6, the original deformation field, computed with the 2D estimation, is shown alongside the magnitudes of the Euclidean error between the inpainted fields and the baseline. As also shown in Table 2, the mean error increases with the number of neighboring samples that are removed from the inpainting process, as expected. The error obtained is comparable to that in Fig. 7.

3D volume undeformation

We validate the performance of the proposed deformation correction method using target registration error (TRE) between the compounded ground truth and the undeformed volumes acquired with different forces (2, 5, 8 and 15 N—20 volumes in total). Anatomical landmarks were manually

Table 2 Quantitative inpainting accuracy: error in *mm* (average and SD) of an interpolated deformation field with respect to respective estimated one

2 Neighbors	4 Neighbors	6 Neighbors	8 Neighbors	10 Neighbors	12 Neighbors
0.94 ± 0.83	0.95 ± 0.84	1.11 ± 0.87	1.20 ± 0.87	1.40 ± 0.97	1.62 ± 1.00

The deformation field for the same location is interpolated incrementally removing neighboring fields. Errors are shown for 2, 4, 6, 8, 10, and 12 fields being removed symmetrically around the selected one

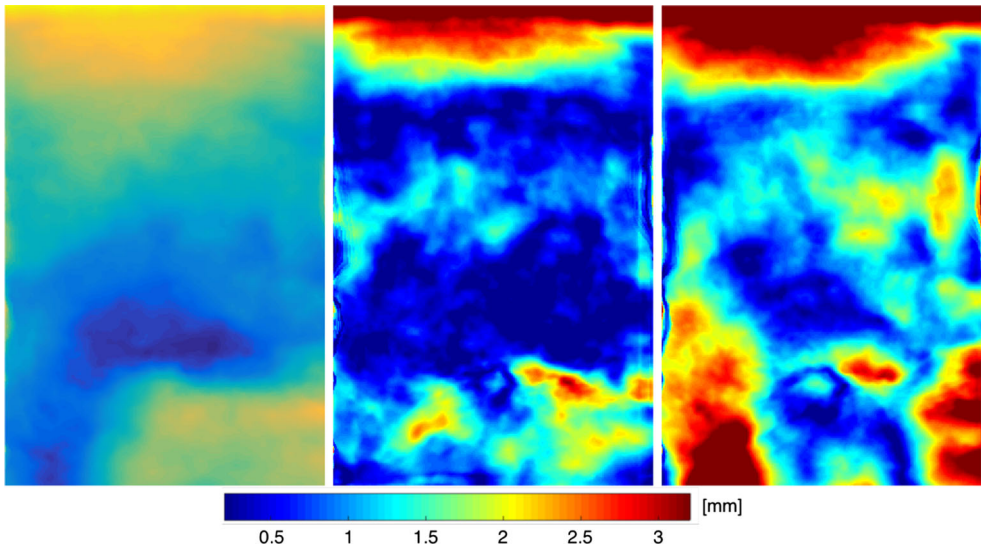

Fig. 6 Qualitative inpainting accuracy: on the left, an exemplary deformation field computed by the 2D deformation estimation is shown. Inpainting is performed with the attempt of reproducing the computed deformation as accurately as possible, while incrementally removing the surrounding neighboring fields to assess how sparse sampling affects output quality. The error magnitudes of the interpolated deformation field with respect to the one on the left are shown in the center (removing 2 neighbors) and on the right (removing 12 neighbors)

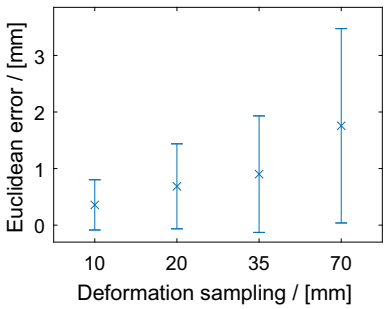

Fig. 7 Deformation subsampling: mean and SD of the Euclidean norm of difference between dense and sparsely sampled deformation estimation

selected with an average of 142 ± 14 fiducial points per volume pair. Table 3 summarizes the measured distances between the chosen points, an increase from 3.11 ± 1.55 to 6.27 ± 2.59 mm can be observed with increasing force. The achieved correction is also shown in Fig. 8.

Discussion and conclusion

The obtained results show that the proposed method is able to capture the deformation induced by the US transducer during a 3D robotic acquisition and effectively correct for it. While a direct comparison to the current state of art is difficult due to variations in the acquisition protocols, the reported deformation correction for a clinical case [14] is 3.5 ± 0.4 mm, which is comparable to our findings for 5 N in Table 3. Such an error would be clinically acceptable for the target application, since diagnosed STS have an average size of 10 cm [6].

Table 3 Target registration error

2 N	5 N	8 N	15 N
3.11 ± 1.55	3.39 ± 1.86	4.65 ± 1.92	6.27 ± 2.59

Average and SD [mm] (3 subjects). The force values used for acquisitions allow the visualization of subcutaneous masses in a large set of body type

To improve the reproducibility of this work and allow future comparative evaluations, we release the acquisitions acquired on volunteers.[4] The dataset contains synchronized US images, tracking data and force information.

It is clear that the overall deformation estimation was able to better correct for low forces, since the estimated deformations from our fitted model are inherently subject to noise due their local nature. Resulting undeformed volumes may contain artifacts at the interface of different tissue layers, as noticeable in Fig. 8, as the tracking of the deformation is more difficult due to the diverse response to the applied force. This effect could be reduced by a better regularization of the obtained deformation fields. It is valuable to note that we do not aim to compensate for all the possible sources of deformations that might be present during an US acquisition, e.g., breathing motion, vascular pulsation, but—similarly to the state of the art—we tackle the deformation caused by

[4] http://campar.in.tum.de/files/virga/dataset.zip.

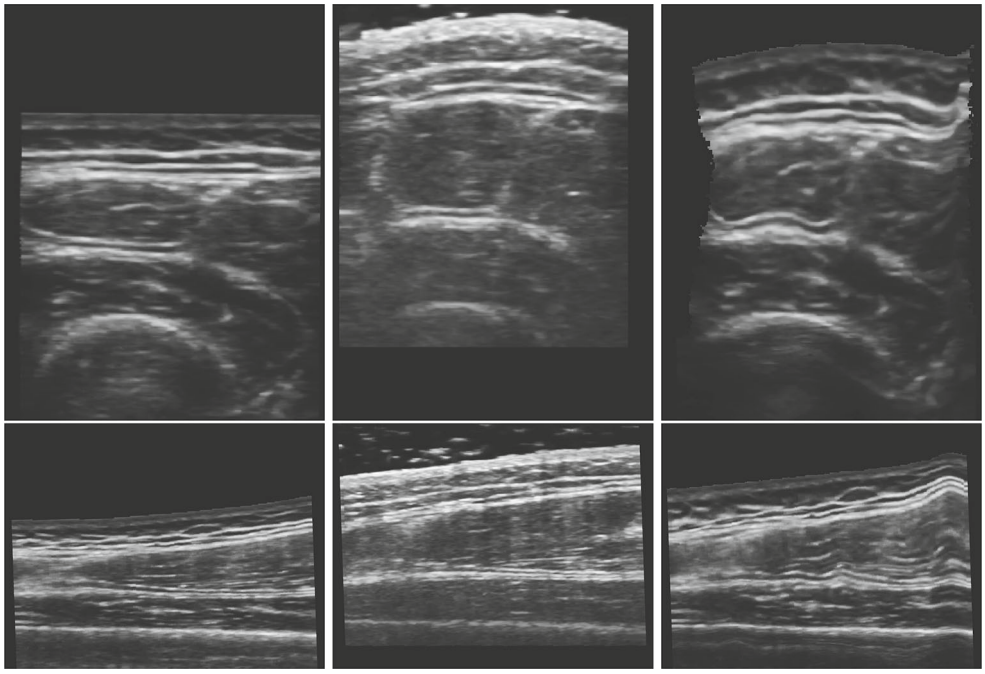

Fig. 8 Comparison of deformed and undeformed volumes. Axial and lateral views of deformed (left), ground truth (center) and undeformed (right) volumes

the probe pressure only. Future work will include improvements in the model to incorporate constrains on the resulting deformation fields and integrate information on elastic tissue behavior. Additionally, a validation of the method for multimodal volumetric registration would be beneficial to assess its potential in additional clinical settings. Also, a prospective validation on clinical patients would be beneficial to evaluate the deformation on pathological tissue, opening the way to clinical impact of robotic US imaging.

Compliance with ethical standards

Conflict of interest The authors declare to have no conflict of interest.

Human and animals rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. No animal experiments were performed in this study.

Informed consent Informed consent was obtained from all participants.

References

1. Arya S, Nagarkatti DG, Dudhat SB, Nadkarni KS, Joshi MS, Shinde SR (2000) Soft tissue sarcomas: ultrasonographic evaluation of local recurrences. Clin Radiol 55(3):193–197
2. Boehler T, Peitgen HO (2008) Reducing motion artifacts in 3-D breast ultrasound using non-linear registration. In: International conference on medical image computing and computer-assisted intervention, pp 998–1005
3. Burcher MR, Han L, Noble JA (2001) Deformation correction in ultrasound images using contact force measurements. In: IEEE workshop on mathematical methods in biomedical image analysis, pp 63–70
4. Dahmani J, Petit Y, Laporte C (2017) Model-based correction of ultrasound image deformations due to probe pressure. In: Medical imaging 2017: image processing, vol 10133, p 101331D
5. Elek R, Nagy TD, Nagy D, Takcs B, Galambos P, Rudas I, Haidegger T (2017) Robotic platforms for ultrasound diagnostics and treatment. In: IEEE international conference on systems, man, and cybernetics (SMC), pp 1752–1757
6. Grimer RJ (2006) Size matters for sarcomas!. Ann R Coll Surg Engl 88(6):519–524
7. Hennersperger C, Baust M, Mateus D, Navab N (2015) Computational sonography. In: International conference on medical image computing and computer-assisted intervention, pp 459–466
8. Hennersperger C, Fuerst B, Virga S, Zettinig O, Frisch B, Neff T, Navab N (2017) Towards MRI-based autonomous robotic us acquisitions: a first feasibility study. IEEE Trans Med Imaging 36(2):538–548
9. Hennersperger C, Mateus D, Baust M, Navab N (2014) A quadratic energy minimization framework for signal loss estimation from arbitrarily sampled ultrasound data. In: International conference on medical image computing and computer-assisted intervention, pp 373–380
10. Morrison BA (2003) Soft tissue sarcomas of the extremities. In: Baylor University Medical Center. Proceedings, vol 16, p 285
11. Natarajan S, Marks LS, Margolis DJ, Huang J, Macairan ML, Lieu P, Fenster A (2011) Clinical application of a 3D ultrasound-guided prostate biopsy system. In: Urologic oncology: seminars and original investigations, vol 29, pp 334–342
12. Ophir J, Cespedes I, Ponnekanti H, Yazdi Y, Li X (1991) Elastography: a quantitative method for imaging the elasticity of biological tissues. Ultrason Imaging 13(2):111–134

13. Pfister K, Schierling W, Jung EM, Apfelbeck H, Hennersperger C, Kasprzak PM (2016) Standardized 2D ultrasound versus 3D/4D ultrasound and image fusion for measurement of aortic aneurysm diameter in follow-up after evar. Clin Hemorheol Microcircul 62(3):249–260
14. Pheiffer TS, Miga MI (2015) Toward a generic real-time compression correction framework for tracked ultrasound. Int J Comput Assist Radiol Surg 10(11):1777–1792
15. Rastrelli M, Tropea S, Basso U, Roma A, Maruzzo M, Rossi CR (2014) Soft tissue limb and trunk sarcomas: diagnosis, treatment and follow-up. Anticancer Res 34(10):5251–5262
16. Sun SY, Anthony BW, Gilbertson MW (2010) Trajectory-based deformation correction in ultrasound images. In: Medical imaging 2010: ultrasonic imaging, tomography, and therapy, vol 7629, p 76290A
17. Swerdlow DR, Cleary K, Wilson E, Azizi-Koutenaei B, Monfaredi R (2017) Robotic arm-assisted sonography: Review of technical developments and potential clinical applications. Am J Roentgenol 208(4):733–738
18. Treece GM, Prager RW, Gee AH, Berman L (2002) Correction of probe pressure artifacts in freehand 3D ultrasound. Med Image Anal 6(3):199–214
19. Vercauteren T, Pennec X, Perchant A, Ayache N (2009) Diffeomorphic demons: efficient non-parametric image registration. NeuroImage 45(1):S61–S72
20. Zikic D, Baust M, Kamen A, Navab NA (2011) general preconditioning scheme for difference measures in deformable registration. In: 2011 IEEE international conference on computer vision (ICCV), pp 49–56
21. Schulte zu Berge C, Kapoor A, Navab N (2014) Orientation-driven ultrasound compounding using uncertainty information. In: International conference on information processing in computer-assisted interventions, pp 236–245

5.3 Applications in Vascular Surgery

The contribution described in Sec. 5.1 included the description of a whole system for autonomous robotic ultrasound acquisitions and analyzed the integration with external 3D sensing for an accurate trajectory planning. On the other hand, in Sec. 5.2, the focus has been on leveraging the force sensing information to enhance the ultrasound acquisitions. The following contributions aim at showcasing the use of all the presented methodologies in specific clinical cases. In Sec. 1.3, some of the clinical applications that robotic ultrasound addresses in literature have already been presented, in this case, we target applications in the vascular surgery domain, specifically, for both diagnostic and treatment of an Abdominal Aortic Aneurysm (AAA).

In brief, AAA is characterized by the enlargement of the main artery in the human body, the aorta, in its abdominal section. Although it is commonly an asymptomatic condition, it causes the weakness of the vessel's wall and increases the chances of its rupture, which would cause a rapid patient's death.

5.3.1 Diagnosis of Abdominal Aortic Aneurysm (IROS2016)

The diagnosis of AAA is commonly reached via unrelated analyses carried out while investigating other pathologies. Typically, an abdominal CT or MRI, acquired for other indications, leads to the discovery of the pathology by chance. Ultrasound imaging is a valid alternative for the diagnosis and staging of AAA: it is a non-invasive imaging modality, and it is highly available. However, as already discussed in Sec. 1.1, the need for a manual acquisition of the images limits its use on a large scale. Implementing a screening program for AAA, which is suggested for all males above 60 years of age, would require a large workforce and, so far, has been therefore unsuccessful.

This work proposes an autonomous robotic system that can acquire a full 3D ultrasound volume of the abdominal aortic tract without any manual interaction with the patient. The aim is to allow a seamless analysis of the patient's aortic diameters and plan for further acquisitions (for AAA staging) or for an intervention to mitigate the risk of a vessel rupture. Based on the system presented in Sec. 5.1, the anatomical area of interest is here selected from a statistical MRI atlas. That is, since it is not possible to assume the availability of a pre-operative volume of a patient undergoing screening, the closest available individual in the atlas (in terms of body size, age, BMI, etc.) is used and the abdominal region of the aorta selected. In this case, a deformable registration (See Sec. 4.1) is employed to match that trajectory to the current patient location.

Showing the potential of integrating multiple sensing in such a system, the proposed solution takes also advantage of the live ultrasound image information to optimize the force control strategy. During the acquisition, in fact, the system attempts to find the best patient-specific force value to be applied such that the image quality is maximal with respect to the aortic area, Fig. 5.2. The concept of *ultrasound confidence maps* [160] is used to assess the visibility of the aortic tract at the current force level, so that the force control scheme can be optimized

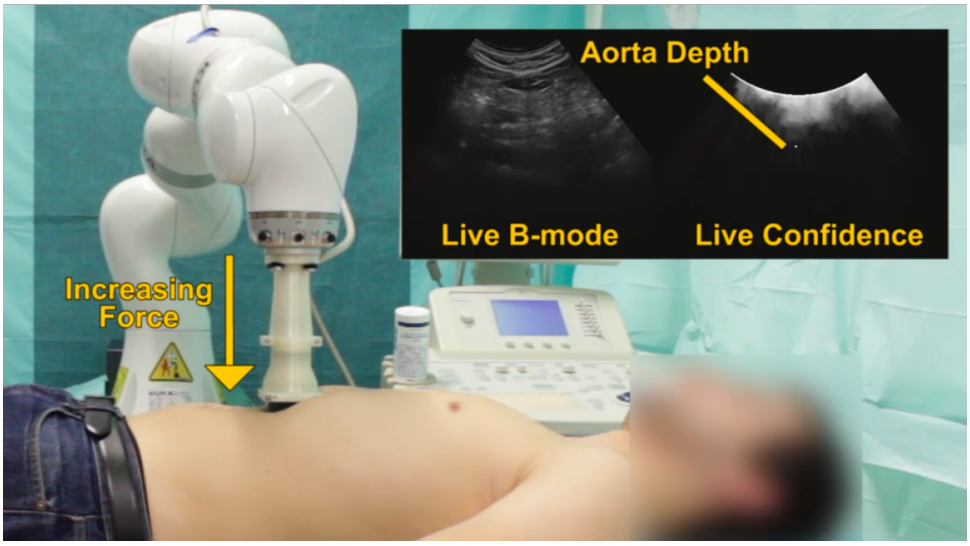

Fig. 5.2. The presented system autonomously selects the best suitable force to apply onto the patient surface to achieve an optimal visibility of the target anatomy. On the top left the live ultrasound image acquired by the robot is shown, together with its respective confidence map. The average confidence value around the aortic region is used as input for the used force control strategy.

for the task at run-time. Additionally, the out-of-plane orientation of the ultrasound probe is also optimized during the trajectory execution, again, to maximize the aorta visibility.

Performing a validation on human volunteers, the contribution shows that, with such a system, it is possible to achieve identical measurements of the aortic diameters with respect to the ones obtained by an expert vascular surgeon using manual ultrasound acquisitions.

Automatic Force-Compliant Robotic Ultrasound Screening of Abdominal Aortic Aneurysms

Salvatore Virga[*1,2], Oliver Zettinig[*1], Marco Esposito[1], Karin Pfister[3], Benjamin Frisch[1], Thomas Neff[1], Nassir Navab[1,4], Christoph Hennersperger[1]

[1] Chair for Computer Aided Medical Procedures (CAMP), Technische Universität München, Munich, Germany.
[2] KUKA Roboter GmbH, Augsburg, Germany.
[3] Division of Vascular and Endovascular Surgery, University Medical Center Regensburg, Regensburg, Germany.
[4] Johns Hopkins University, Baltimore, USA.

Contribution. The main contributions of this publication, including the design and implementation of the force and position optimization algorithms, were performed by the author of this thesis. Oliver Zettinig was responsible for analyzing the results of the performed experiments. Karin Pfister performed and analyzed the manual ultrasound acquisitions on healthy volunteers. The writing and revision of the publication were performed jointly together with all the co-authors.

Automatic Force-Compliant Robotic Ultrasound Screening of Abdominal Aortic Aneurysms

Salvatore Virga*[1,2], Oliver Zettinig*[1], Marco Esposito[1], Karin Pfister[3], Benjamin Frisch[1], Thomas Neff[2], Nassir Navab[1,4] and Christoph Hennersperger[1]

***Abstract*— Ultrasound (US) imaging is commonly employed for the diagnosis and staging of abdominal aortic aneurysms (AAA), mainly due to its non-invasiveness and high availability. High inter-operator variability and a lack of repeatability of current US image acquisition impair the implementation of extensive screening programs for affected patient populations. However, this opens the way to a possible automation of the procedure, and recent works have exploited the use of robotic platforms for US applications, both in diagnostic and interventional scenarios. In this work, we propose a system for autonomous robotic US acquisitions aimed at the quantitative assessment of patients' vessel diameter for abdominal aortic aneurysm screening. Using a probabilistic measure of the US quality, we introduce an automatic estimation of the optimal pressure to be applied during the acquisition, and an online optimization of the out-of-plane rotation of the US probe to maximize the visibility of the aorta. We evaluate our method on healthy volunteers and compare the results to manual acquisitions performed by a clinical expert, demonstrating the feasibility of the presented system for AAA screening.**

I. INTRODUCTION

Ultrasound (US) imaging has become the first-line imaging modality for multiple medical indications, including the focused assessment with sonography for trauma (FAST) as a routine emergency workflow or general vascular conditions [1], [2]. Due to its non-invasiveness and swift imaging capabilities, ultrasound is well suited for screening applications. One target area with a high associated benefit from routine screening would be the abdominal aortic aneurysm (AAA), a dilation (ballooning) of one of the major vessels in the human body. The major risk of an AAA is the rupture of the aneurysm, which is associated with high mortality rates up to 50%. The probability for rupture depends on the size, shape and stress of the aneurysm, with a substantially increased risk for diameters above 6 cm [3]. Ultrasound is already employed as a standard diagnostic tool for the imaging of the aorta, and US-based staging is widely accepted in clinical practice [4]. However, challenges with respect to inter-operator variability and standardized measurement approaches still impair the implementation of national or international sonography-based screening programs [4], [5], [6]. In contrast to conventional clinical 2D ultrasound, 3D US showed an improved localization of aneurysm-shape and endoleak after Endovascular Aortic Repair (EVAR) compared to 2D US, using contrast-enhanced imaging [7]. Based on the extraction of quantitative values from these data, a 3D ultrasound scanning system with reproducible and constant acquisition quality could facilitate the establishment of screening frameworks. While the patient benefit from an early detection and regular check for aortic aneurysms is obvious in such a program, discussions about the overall cost-effectiveness of potential screening activities remain [8]. The full automation of the acquisition using robotic technology would reduce personnel costs a major factor for the data acquisition while achieving full comparability of acquired data. In addition, work-related musculoskeletal discomfort of US examiners could be reduced [9].

In the past decades, attempts for generating operator-independent 3D ultrasound data focused on motorized probes [10], freehand 3D approaches [11], partially also combined with ECG-gating [12] and more recently pulse-oximetry [13] to remove artifacts due to vessel pulsation. While today's methods achieve higher repeatability of the acquired data, they lack an automation of the acquisition itself. In view of servoing approaches, US probes are used and guided by robotic systems for needle placement [14], as well for showing a constant anatomical position by breathing compensation [15], an automatic optimization of the US probe direction for optimal acoustic coupling [16] and the automatic servoing based on live image registration [17]. Finally, first attempts for fully automated ultrasound acquisitions used RGB skin feature detection following a rigid acquisition protocol for liver ultrasound [18], and performed acquisitions based on previously planned trajectories on tomographic image data [19].

In this work, we present a fully autonomous framework to acquire abdominal 3D US images to facilitate AAA screening in clinical routine. Designed to cope with a high anatomical variety in the general population, our system aims to adapt the performed US trajectory to the individual patient. To this end, we elastically register the patient to a generic MRI-based atlas, and autonomously perform a force-optimized robotic US scan of the abdominal aortic region, allowing for manual diameter measurements. In particular, our contributions are as follows:

a) In contrast to [19], a deformable registration in combination with a generic patient atlas is employed to account for various body sizes and shapes. In this way, patients for whom no tomographic imaging data is available can

*These authors contributed equally to this work.

[1]Computer Aided Medical Procedurs, Technische Universität München, Munich, Germany. salvo.virga@tum.de

[2]KUKA Roboter GmbH, Augsburg, Germany

[3]Division of Vascular and Endovascular Surgery, University Medical Center Regensburg, Regensburg, Germany

[4] Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD, USA

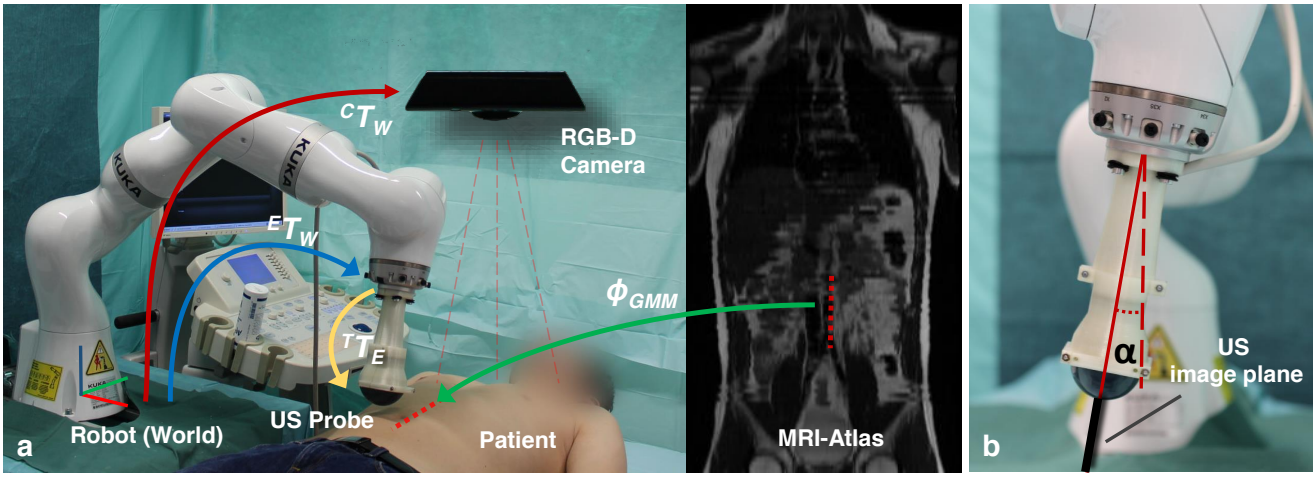

Fig. 1. **a)** System setup showing the robot, the US transducer mounted to its end-effector, the RGB-D camera on the ceiling, the patient, a coronal slice of the MRI atlas, and all required transformations, including the camera-to-robot (${}^{C}T_{W}$) and extrinsic US probe calibrations (${}^{T}T_{E}$). The red-dotted line indicates the scan trajectory. **b)** Close-up of the US transducer, showing the US imaging plane (*thick black line*) and the out-of-plane rotation angle α.

undergo the proposed screening protocol.

b) Similarly, due to the great variety of possible patient conditions, there is a need for an adaptive parametrization of the employed contact force of the US transducer onto the skin [4]. In this work, we propose to utilize confidence maps [20] to automatically determine the optimal force for the US scan.

c) Finally, we propose a control law based on confidence maps as an adapted version of [16] targeted at our application. In this regard, we initially estimate and continuously adjust the out-of-plane rotation of the US transducer during the acquisition. This enables the optimization of the image quality at specified penetration depths (e.g. aorta), compensating for shadowing artifacts due to the presence of bowel gas.

II. METHODS

The presented autonomous screening system consists of a robotic manipulator, allowing for force estimation at its end-effector by means of internal joint torque sensors, and a rigidly attached US transducer. In addition, a structured-light RGB-D camera is mounted at the ceiling of the examination room, as illustrated in Fig. 1a. In this section, we first describe the registration between the MR-atlas and the patient using RGB-D data in Sec. II-A. Then, we summarize in Sec. II-B how the signal loss in US images is estimated in the form of confidence maps. While in Sec. II-C, the overall robot control scheme is introduced, specific details concerning the force estimation and out-of-plane control are described in Sec. II-D and II-E, respectively.

A. Patient Registration and Trajectory Transfer

The aim of a global patient-to-world registration is to gather knowledge about the current patient position with respect to the system setup, so that accurate and safe motions of the robotic arm can be achieved. We exploit the sensing information of a RBG-D camera to transfer a generic trajectory aimed at the aorta of an atlas onto the current patient.

For camera-to-robot calibration, we employ a technique based on augmented reality markers similar to the approach in [19] to determine the transformation ${}^{C}T_{W}$ (cf. Fig. 1a). Note that the computer vision notation for transformations, i.e. 4×4 homogeneous matrices, are used in this work.

Given the wide target population of a screening program, it is not feasible to assume the availability of individual tomographic images for each patient. Hence, we propose to leverage a statistical MRI atlas based on physical and anatomical characteristics, such that the anatomy of each patient can be taken into account for the trajectory planning [21]. A surface point cloud P_{MRI} extracted from a selected MRI atlas image can be elastically registered to the live point cloud P_{RGBD} obtained from the RGB-D camera. We make use of an implementation of the Coherent Point Drift algorithm [22], a probabilistic non-rigid registration method that fits a Gaussian Mixture Model (GMM) to the moving point set. The GMM is initialized using the target points P_{RGBD} and a coherent velocity is enforced to its motion so that a smooth non-linear transformation $\Phi_{GMM} : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ can be computed using spline interpolation. Both point sets are subsampled by a factor f_s for this process, allowing for an optimal trade-off between the fitting accuracy and the computational performance.

By a projection of the start- and endpoint ($\mathbf{p}_s$,$\mathbf{p}_e \in \mathbb{R}^4$) of the aortic region of interest from the atlas to its surface, the robotic trajectory on the patient surface ($\mathbf{p}'_s$,$\mathbf{p}'_e$) is obtained by transfering these points to the world coordinate system

$$\mathbf{p}'_s = {}^{C}T_{W} \cdot \Phi_{GMM}\left[f_{NN}(\mathbf{p}_s, \mathbf{e}_z)\right], \tag{1}$$

where $\mathbf{e}_z = (0, 0, 1, 0)$ is the vertical unit vector, and $f_{NN}(\mathbf{p}, \mathbf{n})$ computes the nearest element of the point cloud P_{MRI} to the ray $\mathbf{p} + \lambda\mathbf{n}, \lambda \in \mathbb{R}_0$ (analogous for $\mathbf{p}'_e$).

B. US Acquisition and Confidence Map Computation

In our framework, a series of 2D B-mode frames $I_i \in \mathbb{R}^2$ are acquired using a convex transducer suited for abdominal scans. For an estimation of the ultrasound quality, we employ confidence maps as introduced by [20]. In short, a graph is constructed between the pixels of the B-mode image, with source (1) and sink (0) nodes at the transducer elements and the bottom of the image, respectively, and edge weights based on the US intensities between pixels. The confidence map $C_i \in \mathbb{R}^2 \rightarrow [0, 1]$ is then defined as the equilibrium diffusion solution, i.e. the probability of a random walk starting from a particular pixel to rather reach the transducer than the bottom (see Fig. 3a). We refer the reader to [23] for further details. For any given US frame, we denote with the feature

$$\zeta = \frac{1}{|R|} \sum_{(x,y)\in R} C(x, y), \tag{2}$$

the average confidence in the rectangular region R, centered around the estimated world-coordinate location of the aorta in the corresponding B-mode frame I. Since the spine and the aorta are almost incompressible, it can be estimated from the atlas as in Eq. 1 and considered constant regardless of the current force onto the patient. The averaging copes with the inherent level of noise in confidence maps [24].

C. Robot Control Scheme

The overall goal of the robot control scheme is threefold: First, the US sweep acquisition requires following a predefined trajectory on the patient. Second, the force exerted by the US transducer onto the tissue is intended to not only be sufficient but also optimal over time, achieving good acoustic coupling throughout the sweep. Third, the out-of-plane rotation needs to be adjusted on-line to maintain high image quality even in the presence of bowel gas, shadowing, and other artifacts. Considering the desired tool-tip pose

$$\mathrm{P}_d = \begin{bmatrix} \mathrm{R}(\alpha, \beta, \gamma) & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{3}$$

with translation $\mathbf{t} = (x, y, z)$ and rotation R using the Euler angles (α, β, γ), it is possible to decouple the three tasks. A standard position controller is used to command translations (x, y) in the horizontal plane, guided by the planned trajectory.

D. Adaptive Force Estimation

The downward translational component z is controlled by a force controller as already demonstrated in prior art [16], [17], [25]. Constant force control for manipulators with elastic joints is typically achieved by balancing external Cartesian forces $\mathbf{F_{ext}}$ acting on the end-effector with a desired force $\mathbf{F_d}$ so that $\mathbf{F_{ext}} - \mathbf{F_d} = 0$, as historically described in [26]. For real-time behavior, torque sensors in all robotic joints are utilized to compute the external forces using both the Jacobian and the known inverse dynamics system of the manipulator.

Although the area of general robotic force control has been extensively discussed throughout the last decades, the choice of an appropriate force $\mathbf{F_d}$ for a particular medical scenario depended on a manual parametrization up to now. While too little pressure will compromise good acoustic coupling and sufficient image quality during US acquisitions, excessive force might overly deform the anatomy or even harm the patient. In the view of a fully autonomous robotic system for US screening, an optimal force value cannot be known a priori but has to be estimated online to cope with a variety of patients constitution and tissue density. Therefore, we propose an online adaptive force estimation based on confidence values presented in II-B. During initialization, we vertically approach the start pose $\mathbf{p}'_s$ until skin contact ($\mathbf{F_{d}}_0 = 0$). Next, the desired force exerted onto the tissue is increased iteratively by $\mathbf{F}_{step}$ until a mean confidence threshold Θ is reached (H is the Heaviside step function):

$$\mathbf{F_{d}}_{i+1} = \mathbf{F_{d}}_i + \mathbf{F}_{step} \cdot H\left(\Theta - \zeta_i\right). \tag{4}$$

E. Optimization of Out-of-Plane Rotation

The Euler angles (α, β, γ) of the pose P_d can be interpreted as out-of-plane rotation, in-plane rotation, and rotation around the transducer axis, respectively. Contrary to the target anatomies investigated in [16], [17], axial aortic scans benefit only marginally from in-plane or transducer axis rotation: On the one hand, turning the US probe around its axis does not avoid acoustic obstacles between the transducer and the aorta. On the other hand, an in-plane rotation during axial scans constitutes lateral tilting, quickly translates the aorta away from the image center, and is not considered helpful in clinical routine [7]. Thus, we define $\beta = \gamma = 0$ for all experiments and concentrate on the more challenging out-of-plane rotation α for image quality optimization.

Initially, a sweep is acquired with angles between $[\alpha_{min}; \alpha_{max}]$, where the confidence feature ζ_α is recorded for each rotatory pose. The optimal out-of-plane rotation α_0 for the start point of the sweep is then defined as the angle that maximizes confidence at the aorta: $\alpha_0 = \arg\max_\alpha \zeta_\alpha$.

Throughout the acquisition, we aim to maintain an optimal echoing pose. As the direction to tilt the probe out of its plane cannot be directly inferred from 2D frames, we propose to compute the following parameters for each frame. First, the binary parameter κ indicates a drop in confidence below the average of the preceding M frames

$$\kappa = H\left(\zeta_i - \frac{1}{M} \sum_{j=i-M-1}^{i-1} \zeta_j\right). \tag{5}$$

It is used to determine whether the current probe orientation provides sufficient image quality. Second, the parameter $s \in \{-1, 1\}$, $s_0 = 1$, states the direction the probe should tilt to:

$$s_i = \begin{cases} -s_{i-1} & \text{if } \zeta_i < \zeta_{i-M} \ \wedge \left| \sum_{j=i-M-1}^{i-1} s_j \right| = M \\ s_{i-1} & \text{else .} \end{cases} \tag{6}$$

As a result, s will change sign only if there has not been a change in the previous M iterations, and the current

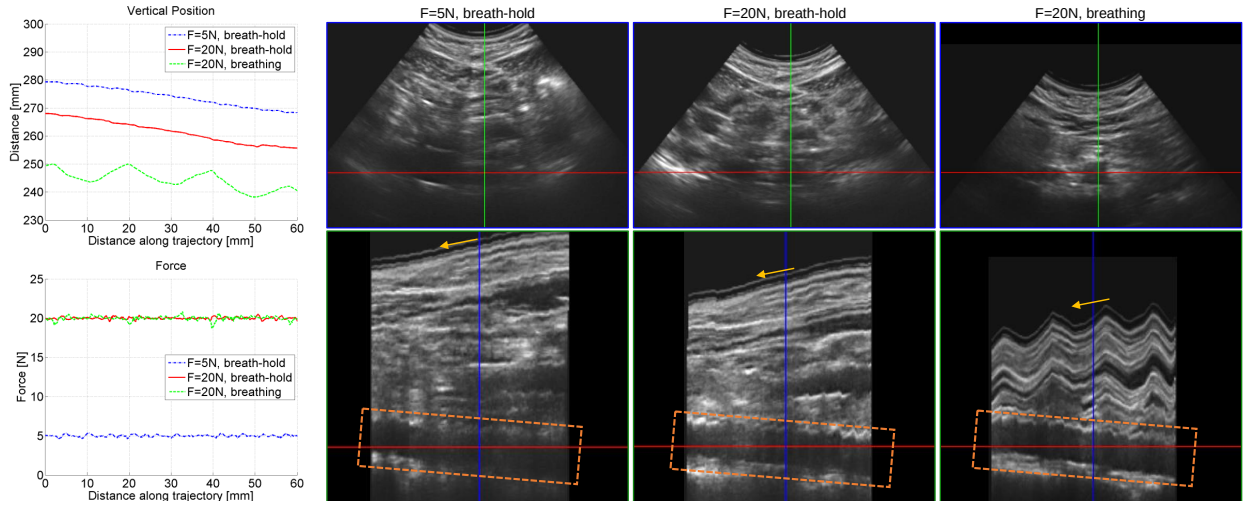

Fig. 2. Results of force adaptation experiments for one subject, including the measured vertical position of the transducer and the corresponding force (*figures on the left*), and axial (*top*) and sagittal (*bottom*) slices of the compounded US volumes. Without proper force onto the tissue (5 N experiment, *left image column*), the US quality at the aorta, in particular its posterior wall, is poor (*orange box*). Quality improves with the adaptive force estimation (20 N experiments, *center column*). Results also show that the force controller can successfully account for breathing motion and allows for steady aortic acquisitions, even if the US image of tissue directly underneath the skin becomes unusable (*right image column*). Yellow arrows indicate the scan direction.

confidence dropped below the one M iterations ago. Altogether, the desired out-of-plane rotation can be computed by combining these factors $\alpha_i = \alpha_{i-1} + \kappa \cdot s \cdot \alpha_{step}$. Updates of α as thus not continuous but are handled by the position controller in a smooth fashion as in [19].

III. Experiments and Results

A. Material and Experimental Setup

The robotic manipulator used in this work is a KUKA LWR iiwa R800 (KUKA Roboter GmbH, Augsburg, Germany) with the KUKA Sunrise.Connectivity software package. A software module[1] developed by the authors allows a direct control of the robot functionalities via the Robot Operating System[2] (ROS) framework. The methods presented in Sec.II are implemented in custom ROS modules, which forward their control output to the KUKA robot controller. For US acquisition, an Ultrasonix® Sonix RP US system equipped with a 4DC7-3/40 curvilinear transducer (Analogic Corporation, Peabody, MA, USA) is used (frequency: 3.3 MHz, depth: 140 mm, gain: 50%). The US probe is attached to the robot flange using a custom designed holder (cf. Fig. 1b). Spatial and depth information are acquired using a Kinect camera (Microsoft Corporation, Redmond, WA, USA) placed above the patient. Experiments ran on a workstation (Intel Core i5, NVIDIA GTX 970) connected to the aforementioned systems. Medical image processing is performed within the ImFusion Suite 1.2.16 (ImFusion GmbH, Munich, Germany). The performed camera-to-robot calibration led to an average error of 2.46 ± 0.96 mm on the x-y plane and 6.42 ± 3.67 mm along the camera's depth axis.

We evaluated our method on five different healthy volunteers (age 24-27, 2 female, 3 male), scanning the aorta from slightly inferior of the rib cage in downward direction roughly until the navel (scanning time approx. 2 minutes per patient). The atlas consisted of a T2-weighted MR image (resolution $1.2 \times 1.2 \times 6$ mm) of one healthy individual (age 26, male), which was deemed sufficient for this volunteer study due to the similar anatomical condition. Similar to [19], the Hausdorff distance between P_{RGBD} and the warped atlas surface mesh was on average 3.7 mm (maximum 9.8 mm), robustly allowing the visualization of the aorta in the US frame without further compensation. For all experiments, the following set of parameters was used: $f_s = 0.01$, $\mathbf{F}_{step} = 2$ N, $\mathbf{F}_{max} = 25$ N, $\alpha_{step} = 2°$, $M = 4$ and the region R comprised an area of 10×10 px. The robot moved with $v = 5$ $^{\mathrm{mm}}/_{s}$ during sweeps, and Θ was empirically set to 0.2.

B. Validation of Force Estimation

For all subjects, patient registration and several US acquisitions were performed. In a first sweep, a minimum force for US screening (5 N) was applied statically. For a second sweep, the force estimation as described in Sec. II-D was used, greatly improving the image quality as illustrated in Fig. 2. As expected [16], the confidence is dependent on the exerted force on the tissue, as visualized in Fig. 3b. In total, the estimated force $\mathbf{F_d}$ was 14.8 ± 6.4 N for all volunteers. In a final sweep (only one volunteer), we tested the capabilities of the robot controller to compensate for motion in real-time to maintain a constant force. Therefore, we asked the volunteer to perform one deep chest inhalation and then breath abdominally throughout the acquisition. Results show that the force controller maintained the desired force with an average error of 0.17 ± 0.24 N, and that a steady acquisition of the (almost incompressible) aorta is possible while breathing.

[1] https://github.com/SalvoVirga/iiwa_stack
[2] http://www.ros.org/

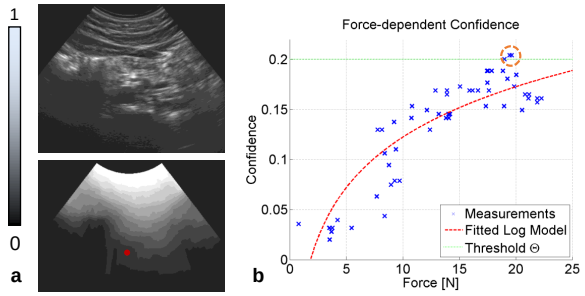

Fig. 3. **a)** Exemplary B-mode frame of the aorta, corresponding confidence map, and expected center of the aorta (*red*). **b)** Confidence values at the depth of the aorta during initial force estimation for one representative subject, showing a degressive dependency on the exerted force. The final force of 20 N was reached after the confidence exceeded $\Theta = 0.2$ (*circled*).

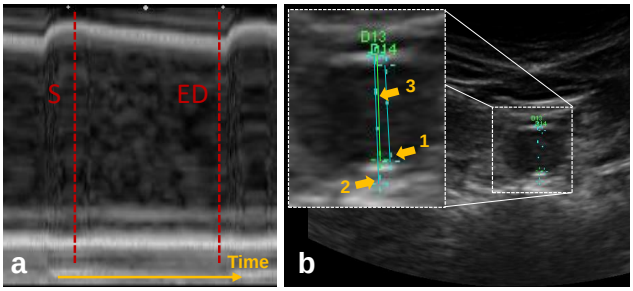

Fig. 5. Result of manual measurement as in clinical routine. **a)** M-mode frame showing cross-section of the aorta over time, allowing for the selection of B-mode frames corresponding to systole (*S*) and end-diastole (*ED*). **b)** B-mode frame with zoomed window of the aorta, showing how inner (*1*), outer (*2*), and leading-edge diameters (*3*) were determined.

C. Optimization of Out-of-plane Rotation

In a next set of experiments, we evaluated the impact of the proposed out-of-plane rotation estimation. A first US sweep was performed with static rotation, i.e. $\alpha = 0$. A representative case is depicted in Fig. 4. In particular in the beginning of the sweep, shadowing artifacts made the aorta not detectable. In contrast, the initial rotation estimation for a second sweep determined an optimal angle of $\alpha = 9.2°$, which led to a significantly increased confidence within the first 30 mm of the sweep. For the remaining trajectory, our controller gradually lowered the out-of-plane rotation and maintained high confidence. For all subjects, the optimal initial rotation was found to be $\alpha_0 = 3.2 \pm 8.0°$.

D. Aortic Diameter Measurements

In all five volunteers, the aortic diameter was measured by a medical expert in the compounded US volumes in sagittal and axial slices according to the guidelines in [4]. For comparison, a medical expert blind to formerly mentioned measurements performed a standard US scan on each volunteer according to the clinical protocol as in [7] (see Fig. 5). In particular, inner diameters d_{in} (without walls), outer diameters d_{out} (with walls) and leading-edge diameters d_{le} (with closer wall only) were measured. Table I reports the results obtained in both scenarios. On average, the error between manual US scan measurement and the ones performed in the robotically acquired volumes was 0.5±0.3 mm.

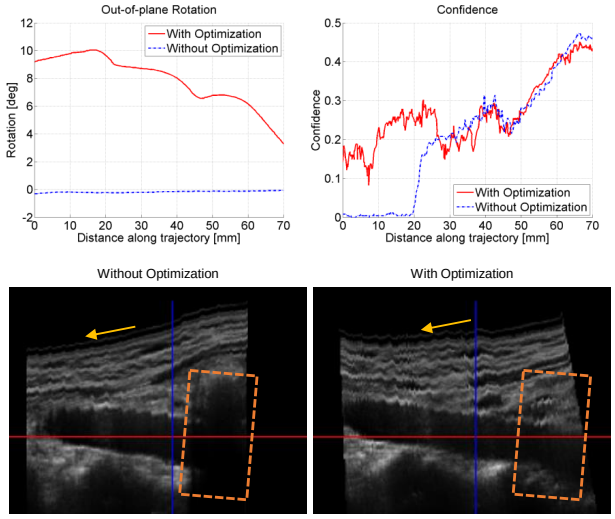

Fig. 4. Comparison between static out-of-plane rotation and optimization based on confidence for one representative subject. The proposed method avoids the shadowing in the beginning of the trajectory (*orange box*). Throughout the acquisition, optimization of the rotation (*top left*) maintains high confidence (*top right*). Yellow arrows indicate the scan direction.

IV. DISCUSSION

With regards to the aortic diameter measurement, our results indicate that the proposed framework allows for the same quantitative measurements as obtained in current clinical practice (cf. Tab. I). The results also validate the adaptive choice of the optimal contact force for the procedure, as the values estimated by our method resulted in a superior image quality. A low force would have led to an overall less visible aorta, while a too high force could have compressed it, thus affecting the diameter measurements. Additionally, the high standard deviation of our automatic force estimation shows the importance of a patient-specific pressure estimation. It should be noted that in this work, all volunteers were within a close range of age and physique, such that much higher variations of various tissues layers can potentially impact the overall ultrasound image quality and contact force for more diverse patient populations. It is further interesting that a good imaging of the aorta was also possible under normal patient breathing (cf. Fig. 2). This might open up the way for optimizations of the current acquisition protocol, in the prospect of a system exploiting normal breathing. With respect to the out-of-plane rotation of the transducer,

TABLE I
AVERAGE DIAMETER MEASUREMENT RESULTS [MM] (5 SUBJECTS)

	Robotic Scan		Manual Scan		Error	
	S	ED	S	ED	S	ED
d_{in}	13.50	11.68	13.38	11.58	0.32±0.13	0.54±0.22
d_{out}	16.68	15.00	16.68	14.84	0.48±0.28	0.56±0.34
d_{le}	15.08	13.34	15.00	13.08	0.52±0.38	0.46±0.15

obtained results demonstrated that the choice of an optimal angle at the beginning, coupled with an ongoing optimization throughout the sweep, leads to an overall improved image quality (cf. Fig. 4). This is in line with clinical practice, where the transducer is regularly tilted to avoid bowel gas and to follow a potential curvature of the aorta [7]. It needs to be noted, however, that reaching and maintaining the globally optimal orientation is not guaranteed by the control scheme. A more in-depth analysis of control stability and confidence convergence is suggested for future studies.

Finally, the presented system allows for a further extension to other applications, such as more complex diagnostic procedures in the vascular domain. Beyond that, the system could facilitate a broader implementation of robotic systems for repetitive and already standardized medical practices.

V. CONCLUSION

In this work, we introduced a fully autonomous robotic system aimed at ultrasound screening for abdominal aortic aneurysms. We have demonstrated a generalized approach to cope with the large variety of anatomies involved in a screening program. Our results show improved quality of 3D US acquisitions, and the clinical validation shows comparable diameter measurements to the ones obtained following current standard of care. This work demonstrates the potential impact of robotic systems on the medical domain, especially in contexts demanding flexibility and adaptation to individual patients. We believe that the generality of the system, based on the foundations of prior work, enables its extension to additional clinical applications and further helps to promote the use of robotic systems in standard medical care.

ACKNOWLEDGMENT

This work was partially supported by KUKA Roboter GmbH, Augsburg, Germany, and Bayerische Forschungsstiftung award number AZ-1072-13 (RoBildOR). We thank ImFusion GmbH, Munich, Germany, for providing their image processing framework and their continuous support.

REFERENCES

[1] T. H. Wong, K. H. Tay, M. G. Sebastian, and S. G. Tan, "Duplex ultrasonography arteriography as first-line investigation for peripheral vascular disease." *Singap Med J*, vol. 54, no. 5, pp. 271–274, 2013.

[2] A. Nchimi, J. Biquet, D. Brisbois, P. Reginster, K. Bouali, C. Saive, and P. Magotteaux, "Duplex ultrasound as first-line screening test for patients suspected of renal artery stenosis: prospective evaluation in high-risk group," *Eur Radiol*, vol. 13, no. 6, pp. 1413–1419, 2003.

[3] D. C. Brewstera, J. L. Cronenwett, J. W. Hallett, K. W. Johnston, W. C. Krupski, and J. S. Matsumura, "Guidelines for the treatment of abdominal aortic aneurysms: report of a subcommittee of the joint council of the american association for vascular surgery and society for vascular surgery," *Vascular Surg*, vol. 37, no. 5, pp. 1106–1117, 2003.

[4] W. Schäberle, L. Leyerer, W. Schierling, and K. Pfister, "Ultrasound diagnostics of the abdominal aorta," *Gefässchirurgie*, vol. 20, no. 1, pp. 22–27, 2015.

[5] L. Beales, S. Wolstenhulme, J. Evans, R. West, and D. Scott, "Reproducibility of ultrasound measurement of the abdominal aorta," *Brit J Surg*, vol. 98, no. 11, pp. 1517–1525, 2011.

[6] P. De Rango, "Aneurysm diameter measurement: A challenging and frustrating task," *Eur J Vasc Endovasc*, vol. 43, no. 1, p. 34, 2012.

[7] K. Pfister, W. Schierling, E. M. Jung, H. Apfelbeck, C. Hennersperger, and P. M. Kasprzak, "Standardized 2D ultrasound versus 3D/4D ultrasound and image fusion for measurement of aortic aneurysmdiameter in follow-up after EVAR," *Clin Hemorheol Micro*, pp. 1–13, 2015.

[8] T. Schmidt, N. Mühlberger, I. Chemelli-Steingruber, A. Strasak, B. Kofler, A. Chemelli, and U. Siebert, "Benefit, risks and cost-effectiveness of screening for abdominal aortic aneurysm." *Fortschr Geb Rontgenstrahlen Nuklearmed Erganzungsbd*, vol. 182, no. 7, pp. 573–580, 2010.

[9] K. Evans, S. Roll, and J. Baker, "Work-related musculoskeletal disorders (WRMSD) among registered diagnostic medical sonographers and vascular technologists: a representative sample," *J Diagn Med Sonogr*, 2009.

[10] A. Fenster, D. B. Downey, and H. N. Cardinal, "Three-dimensional ultrasound imaging," *Phys Med Biol*, vol. 46, no. 5, p. R67, 2001.

[11] A. Gee, R. Prager, G. Treece, and L. Berman, "Engineering a freehand 3D ultrasound system," *Pattern Recognit Lett*, vol. 24, no. 4, pp. 757–777, 2003.

[12] K. Bredahl, N. Eldrup, C. Meyer, J. Eiberg, and H. Sillesen, "Reproducibility of ECG-gated ultrasound diameter assessment of small abdominal aortic aneurysms," *Eur J Vasc Endovasc Surg*, vol. 45, no. 3, pp. 235–240, 2013.

[13] C. Hennersperger, A. Karamalis, and N. Navab, "Vascular 3D+T freehand ultrasound using correlation of doppler and pulse-oximetry data," in *Information Processing in Computer-Assisted Interventions*, 2014, pp. 68–77.

[14] E. M. Boctor, M. A. Choti, E. C. Burdette, and R. J. Webster Iii, "Three-dimensional ultrasound-guided robotic needle placement: an experimental evaluation," *Int J Med Robot*, vol. 4, no. 2, pp. 180–191, 2008.

[15] C. Nadeau and A. Krupa, "Intensity-based ultrasound visual servoing: Modeling and validation with 2-d and 3-d probes," *Robotics, IEEE Transactions on*, vol. 29, no. 4, pp. 1003–1015, 2013.

[16] P. Chatelain, A. Krupa, and N. Navab, "Confidence-driven control of an ultrasound probe: Target-specific acoustic window optimization," in *Robotics and Automation, 2016 IEEE Int. Conference on*, 2016.

[17] O. Zettinig, B. Fuerst, R. Kojcev, M. Esposito, M. Salehi, W. Wein, J. Rackerseder, B. Frisch, and N. Navab, "Toward real-time 3D ultrasound registration-based visual servoing for interventional navigation," in *Robotics and Automation, 2016 IEEE Int. Conference on*, 2016.

[18] B. Mustafa, A. Safwan, T. Ishii, Y. Matsunaga, R. Nakadate, H. Ishii, K. Ogawa, A. Saito, M. Sugawara, K. Niki *et al.*, "Development of robotic system for autonomous liver screening using ultrasound scanning device," in *Robotics and Biomimetics, 2013 IEEE International Conference on*, 2013, pp. 804–809.

[19] C. Hennersperger, B. Fuerst, S. Virga, O. Zettinig, B. Frisch, T. Neff, and N. Navab, "Towards MRI-based autonomous robotic US acquisitions: A first feasibility study," *arXiv preprint arXiv:1607.08371 [cs.RO]*, 2016.

[20] A. Karamalis, W. Wein, T. Klein, and N. Navab, "Ultrasound confidence maps using random walks," *Med Image Anal*, vol. 16, no. 6, pp. 1101–1112, 2012.

[21] M. Reyes, M. A. G. Ballester, Z. Li, N. Kozic, R. M. Summers, and M. G. Linguraru, "Anatomical variability of organs via principal factor analysis from the construction of an abdominal probabilistic atlas," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2009, pp. 682–685.

[22] A. Myronenko, X. Song, and M. A. Carreira-Perpinán, "Non-rigid point set registration: Coherent point drift," in *Adv Neural Inf Process Syst*, 2006, pp. 1009–1016.

[23] C. Schulte zu Berge, D. Declara, C. Hennersperger, M. Baust, and N. Navab, "Real-time uncertainty visualization for B-mode ultrasound," in *Scientific Visualization, Proceedings on*, 2015.

[24] C. Hennersperger, D. Mateus, M. Baust, and N. Navab, "A quadratic energy minimization framework for signal loss estimation from arbitrarily sampled ultrasound data," in *Medical Image Computing and Computer-Assisted Intervention*, 2014, pp. 373–380.

[25] F. A. Fröhlich, G. Passig, A. Vazquez, and G. Hirzinger, "Robot assisted internal mammary artery detection for coronary revascularisation surgery," in *Intelligent Robots and Systems, IEEE/RSJ International Conference on*, 2010, pp. 1849–1855.

[26] M. W. Spong, "Modeling and control of elastic joint robots," *J Dyn Syst Meas Control*, vol. 109, no. 4, pp. 310–318, 1987.

5.3.2 Treatment of Abdominal Aortic Aneurysm (IROS2019)

To reduce the mortality rate in case of aortic rupture, endovascular procedures (e.g., Endovascular Aneurysm Repair) are commonly carried out to insert a catheter into the vasculature and deploy a stent-graft. Visual feedback during the stent insertion is fundamental to ensure its optimal placement. The aim is to cover the entire aneurysm area with the stent sheet, such that the stent would channel the blood flow in case of rupture, but also avoid to obstruct any secondary vessel that branches out from the main aortic tree. The use of live X-ray fluoroscopy imaging is the current gold standard in clinical practice, with the unavoidable risk linked to ionizing radiation exposure. Ultrasound imaging is a potential alternative to fluoroscopy, but a manual acquisition would have to be performed directly by the operating vascular surgeon or an assistant. More importantly, the visual analysis of the ultrasound images to observe and follow the catheter insertion can be a difficult task, as only expert sonographers would be able to do so. Once again, robotic ultrasound can remove the need for a physical manipulation of the imaging device.

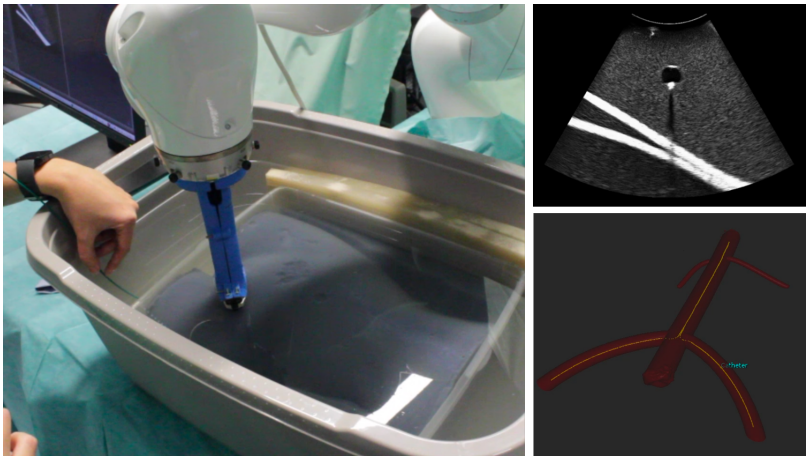

Fig. 5.3. The presented system in action during a mock-up endovascular intervention on a phantom. **Left:**. The system tracks and follows the inserted catheter, the tracking output is fed to the control loop to adjust the transducer position accordingly. **Top right:** The live ultrasound image acquired by the robotic system, the catheter is tracked based on this visual information using a template-matching algorithm. **Bottom right:** The current catheter location is displayed to the used within a 3D model of the vasculature tree.

The presented contribution takes leverage of the discussed robotic ultrasound system for autonomous catheter tracking and navigation in this type of procedure. Spatial information about the vasculature is first extracted pre-operatively patient volumetric imaging data, commonly acquired in this for this procedure, and then registered to the current patient location using an initial 3D ultrasound volume, a combination of point set and image-based registration is used in this case. During the procedure, the catheter is localized in the live 2D ultrasound images, and its position referred to the global map obtained by the initial registration. Using a hybrid force/motion control technique, a constant force is applied onto the patient body while the ultrasound transducer follows at all times the tracked catheter tip (Sec. 3.2.2). Experiments are performed both on healthy volunteers as well as a custom-built phantom to assess the precision of the catheter tracking system, showing an average tracking error of the moving catheter tip of $1.78(102)\,\mathrm{mm}$. Fig. 5.3 displays one of the experiments performed on a phantom and the visualization of the final catheter location within the geometrical structure of the vasculature.

Robotic Ultrasound for Catheter Navigation in Endovascular Procedures

Fernanda Langsch*[1], Salvatore Virga*[1], Javier Esteban[1], Rüdiger Göbl[1], Nassir Navab[1,2]

[1] Chair for Computer Aided Medical Procedures (CAMP), Technische Universität München, Munich, Germany.
[2] Johns Hopkins University, Baltimore, USA.

Contribution. The author of this thesis conceived the main idea behind this publication and coordinated the scientific work. The software to perform the evaluations and the phantom used during experiments were developed by Fernanda Langsch. Tests were performed and evaluated by the two main authors. The writing and revision of the publication were performed jointly together with all the co-authors.

Robotic Ultrasound for Catheter Navigation in Endovascular Procedures

Fernanda Langsch*,[1], Salvatore Virga*,[1], Javier Esteban[1], Rüdiger Göbl[1] and Nassir Navab[1,2]

***Abstract*— Endovascular procedures require real time visual feedback on the location of inserted catheters. This is currently achieved by means of X-ray fluoroscopy, causing exposure to radiation. This work presents a robotic ultrasound system for catheter tracking and navigation in endovascular interventions, focusing on Endovascular Aneurysm Repair. It relies on registration of pre-operative images to provide both a tracking trajectory and visual feedback of the real-time catheter position. The solution is validated on healthy volunteers as well as on a phantom including a realistic vessel structure, showing an average tracking error of the moving catheter tip of $1.78 \pm 1.02\,\mathrm{mm}$.**

I. INTRODUCTION

An Abdominal Aortic Aneurysm (AAA) is a vascular condition that is characterized by the enlargement of a portion of the aorta, leading to the weakening of its walls and a possible vessel rupture. With a worldwide incidence rate ranging between 4 and 11% per year and a mortality rate of 80% in case of rupture [1], AAA presents a significant health risk for a vast part of the population, particularly for males over 65 years. AAA is typically treated via Endovascular Aneurysm Repair (EVAR), a minimally invasive procedure in which a catheter is guided to the site of the aneurysm and used to deploy a stent graft that captures the blood flow, reducing the mechanical stress onto the vessel's walls and preventing rupture [2]. Pre-surgical planning is commonly based on pre-operative images, typically Computed Tomography Angiography (CTA), to determine the appropriate stent graft's dimensions and manufacture it accordingly. During the procedure, the stent graft has to be precisely guided to a location that guarantees full coverage of the aneurysm. At the same time, occlusion of other vessel branches, such as the renal arteries, has to be avoided. For EVAR procedures, fluoroscopy imaging is the standard modality for catheter localization and navigation; causing exposure to ionizing radiation for the patient and the medical staff and requiring the use of contrast agent.. Additionally, pre-operative imaging is typically not integrated into the intra-operative navigation, forcing the surgeon to perform a mental mapping between the two modalities, fluoroscopy and CTA.

In contrast to fluoroscopy, ultrasound (US) is a widely available non-ionizing imaging modality that allows for a clear view of the aorta, and it is already employed in screening programs for AAA diagnosing in some countries [3]. It has, however, the limitation of being a highly user-dependent modality yielding poor reproducibility. As a solution to this problem, robotic systems allow for precise and repeatable acquisitions, and they naturally provide the tracking information needed to obtain 3D US volumes from conventional 2D scanners. This work aims to investigate the potential of robotic US imaging as a replacement for fluoroscopy in endovascular procedures.

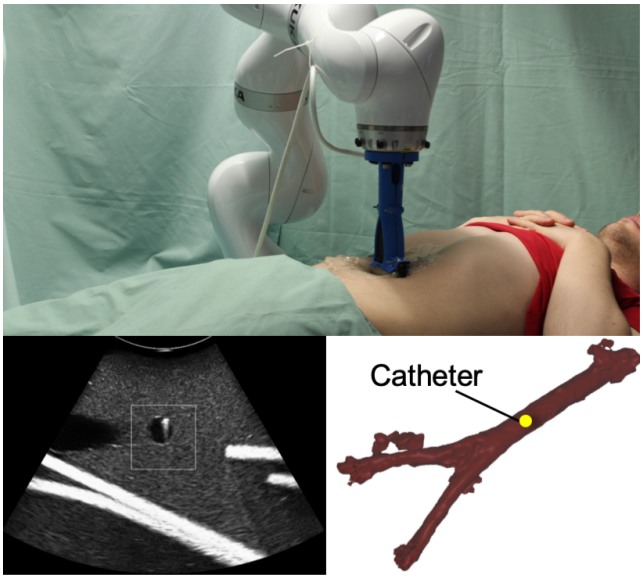

Fig. 1. System overview. Top: autonomous robotic US on volunteer; bottom left: catheter view in US; bottom right: navigation view of catheter location. We refer the reader to the provided supplementary material for an additional view of the proposed solution in action.

The introduction of mechatronic platforms aimed at automatizing US acquisitions and allowing telemanipulation of US transducers has been the focus of research for almost two decades [4]. The research efforts in the field extend from the design of specific end-effectors for the steering of US probes [5] to the introduction of complete systems for the imaging of arbitrary human anatomies using MRI-based trajectories [6]. Specifically for AAA, a robotic solution for autonomous diagnosis using US has also been proposed [7]. An overview of the state-of-the-art in the field of robotic US research is available in [8].

When considering US-based guidance for endovascular procedures, the main goal is to provide the surgeon with the live position of the inserted catheter at all times based on its detection in the available images. Approaches for visual servoing of medical robots based on live tracking and detection of surgical devices have been proposed for various

*These authors contributed equally to this work.
[1]Computer Aided Medical Procedures, Technische Universität München, Munich, Germany fernanda.langsch@tum.de
[2]Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD, USA

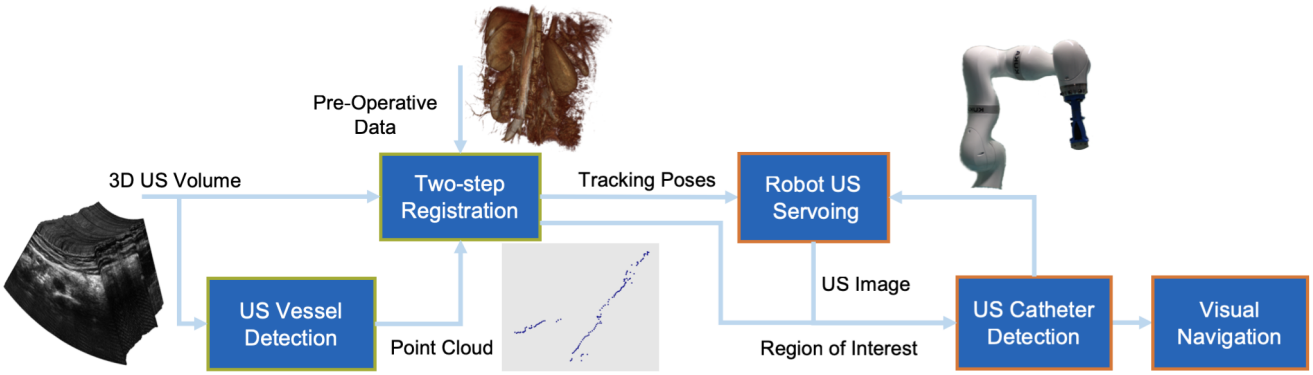

Fig. 2. Diagram of system workflow. A 3D US volume acquired with robotic US and its detected vessel centers (US Vessel Detection), together with the pre-operative image data go through registration (intensity-based and feature-based), yielding the location of vessel centers in the robot frame (Two-step Registration). The robot performs catheter tracking by moving over the computed vessel locations (Robot US Servoing). Image-based catheter detection is facilitated by reducing the search space with knowledge of the vessel location (US Catheter Detection). Finally, the catheter position is displayed over segmented vessel for navigation (Visual Navigation).

scenarios. The automatic alignment of an US probe to the insertion path of a surgical needle has been presented in [9], while in [10] the tip of a flexible needle is followed by a robotic system for minimally invasive procedures. Other approaches, instead, focus on the direct servoing of steerable catheters based on visual information obtained from 3D US [11]. A review of visual servoing techniques in medical robotics is also available [12]. Most notably, a robotic system for US-based tracking on manually inserted catheters has been presented in [13]. The system comprises a robotic manipulator equipped with a linear transducer and a custom catheter with a built-in US active element on its tip. That is, the tracking of the catheter is not performed by means of visual processing, but by detecting the US signal received by the single piezo-electric element installed onto the catheter. While this technique simplifies the tracking problem and might out-perform vision-based tracking, it requires modifications to standard clinical catheters and the consequent integration of additional hardware. Additionally, the proposed method, as presented in [13], does not allow for a seamless bi-directional tracking of the catheter position, and it does not combine the tracking information into any form of visual navigation for the user.

In this work, a complete robotic US system targeted at integrating US-based catheter tracking and visualization to enable radiation-free navigation for EVAR procedures is presented. To achieve this, standard pre-operative volumes are processed to segment the vessel structure of interest and obtain its centerlines. 3D US volumes of the target anatomy are registered to the pre-operative data to create a correspondence of the known vasculature to the intra-operative setting. The computed vessel path is then used to guide the US probe during the manual insertion of a catheter. Image-based catheter detection within the US frames is employed for real-time tracking, allowing to retrieve the position of the catheter along the known path. Finally, the location of the catheter tip is visualized within the segmented vessel structure, providing intuitive navigation feedback to the end-user. The proposed registration scheme is validated on simulated catheter motions for four healthy volunteers, and the complete system is demonstrated on a custom phantom, realistically mimicking both the tissue structure and contours of a human body. In this way, it is shown that the system is able to follow bi-directional motion of the catheter, and visualize its current position precisely.

II. Methods

The following section describes the individual components that enable the system to perform catheter navigation for EVAR procedures using a robotic US platform. The required preprocessing of the pre-operative data is described in Sec. II-A while Sec. II-B provides details on how the computed path along the vessel centerlines is matched to the current patient position. Sec. II-C presents the image-based catheter detection employed and the correspondent visual servoing strategy. An overview of the system components and their interaction is also shown in Figs. 1 and 2.

A. Data Preprocessing

The abdominal vessel tree is segmented from an input pre-operative volume and its centerlines are computed. This includes the abdominal tract of the aorta, a short portion of the iliac arteries after the aorta bifurcation and the main renal arteries. Segmentation and centerline extraction are performed using the methods provided by the Vascular Modelling Tool Kit (VMTK)[1], which are based on active contour level set segmentation and on a weighted geodesic search over a Voronoi diagram [14], respectively. The output of this process is a set of centerline points $\mathcal{C} \subset \mathbb{R}^3$ that will be used to guide the robot motions during the catheter tracking phase. That is, as the catheter position will be naturally constrained by the vessel geometry, we can limit the robot motions and the catheter tracking to the computed locations.

[1] http://www.vmtk.org

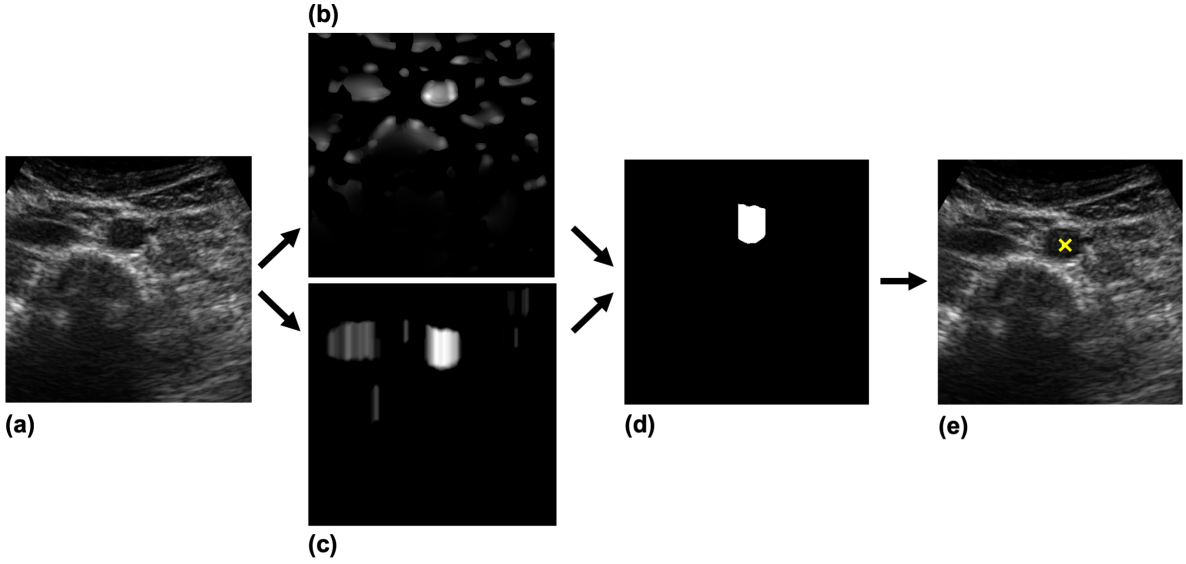

Fig. 3. Vessel detection steps. a) Original US image, b) vesselness response using the modified Frangi filter, c) intensity dip mask, d) binary mask applied to the vesselness response, e) estimated vessel centerline by ellipse fitting.

B. Patient Registration

1) US Compounding: A regular-spaced 3D US volume is obtained by combining the live tracking information supplied by the robot and the 2D US images. A backward normalized-convolution technique is employed, as in [15]. The initial intra-operative acquisition can be performed by moving the robotic arm equipped with a US probe manually or executing a planned trajectory autonomously as in [7].

2) Image-based Registration: To spatially align the two available volumes (pre- and intra-operative) an image-based registration based on the LC^2 similarity measure [16] is used. This is beneficial to cope with the two pre-operative modalities used for the proposed validation, i.e. CT for a phantom and Magnetic Resonance Imaging (MRI) for human volunteers. In fact, in the US/CT case, US intensities are correlated to the output of an US simulation from the CT volume; while in the US/MRI case, the MRI intensities and gradient magnitudes are associated to the US intensity values. In contrast to [16], we only perform a rigid registration. A deformable registration using the same similarity method is feasible, but difficult due to the high deformation induced to the soft tissue in the case of human US acquisitions. LC^2-based registration is employed to provide an initialization for the feature-based method described in II-B.4, which focuses on the optimal alignment of the vessel structures visible in both modalities.

3) Centerline Extraction from US: To extract the vessel centerline from the 3D US volume acquired as in Sec. II-B.1, a slice-wise approach is pursued. A typical method for the detection of tubular structures in medical images is to apply a Frangi filter [17], to enhance the vessel structures. In the case of 2D images, the original formulation of the Frangi vesselness filter is dedicated to highlight longitudinal vessel structures. In detail, the vesselness response for bright vessels at a certain scale s can be defined as

$$V(s) = \begin{cases} 0, & \text{if } \lambda_2 > 0 \\ \exp\left(\frac{-R_B^2}{2\beta^2}\right)\left(1 - \exp\left(\frac{-S^2}{2c^2}\right)\right) & \text{otherwise} \end{cases} \quad (1)$$

with

$$R_B = \frac{|\lambda_1|}{|\lambda_2|} \quad \text{and} \quad S = \sqrt{\lambda_1^2 + \lambda_2^2}\,, \quad (2)$$

where $|\lambda_1| \leq |\lambda_2|$ are the eigenvalues of the Hessian matrix of the US image and β and c are constants.

Since our 2D images provide cross-sectional views of the vessels, the vesselness response needs to be altered so that it provides strong responses to circular structures instead of tubular ones. A modification to the definition of R_B is hereby proposed as

$$R_B = |\lambda_2 - \lambda_1|\,. \quad (3)$$

That is, the original formulation of R_B would lead to a high response for elliptical structures (i.e., $|\lambda_1| \ll |\lambda_2|$), while our variation does so for circular ones (i.e., $\lambda_1 \approx \lambda_2$). Additionally, enforcing $V(s) = 0, \text{if } \lambda_1, \lambda_2 < 0$ filters out undesired bright circular structures and saddle points.

As remarked in [18], the Frangi filter yields poor responses for US images. For this reason, the results obtained using the formulation above are further refined by an intensity dip mask [19], that relies on the computation of average intensities in the image along its columns, given a certain vessel diameter. Intuitively, the average intensity along a column for a region of a given size will be the lowest if the region is centered around dark pixels (i.e., the center of a vessel) rather than bright ones (i.e., including other tissue or the vessel walls). The obtained binary mask is then applied to the vesselness response, such that the center of ellipses fitted to the resulting regions can be considered as

part of the vessel centerline. Fig. 3 outlines the computation of the vesselness response, dip mask and subsequent vessel detection for a given 2D US image. A similar approach for the segmentation of liver vasculature has also been presented in [20].

4) Point-based Registration: A feature-based registration between the two point sets is performed using the classic Iterative Closest Point (ICP) [21] with a RANSAC initialization. Since the centerline point set extracted in Sec. II-B.3 could contain false positives points (e.g., other vessels present in the US images could be detected), possible outliers are preventively excluded using a statistical approach [22].

C. Catheter Tracking and Robot Servoing

Given our aim to provide a complete system for the navigation of catheters in EVAR procedures based on robotic US imaging, the robotic arm equipped with the US transducer has to autonomously explore the patient anatomy, seeking to maintain the inserted catheter in view at all times. The registration process presented in Sec. II-B provides a knowledge of the position of the centerline points extracted from a pre-operative volume in the current robot coordinate frame. Therefore, assuming that during the actual procedure the catheter will be located inside the vessel tree, the controlled tool-tip position can be limited to the ones of the computed centerlines. Only a subset of equidistant centerline points is used during the catheter tracking phase; these points are used both to autonomously guide the US transducer and to perform the catheter detection step on the respective US images. Their distance can be adjusted at runtime, so that a more sparse or dense set can be used during the procedure. This allows, for instance, to have larger movements (e.g., 10 mm) of the transducer when a large movement of the catheter takes place; or, on the other hand, to have very fine motions (e.g., 2 mm) that allow for a better catheter tip detection.

1) Robot Control Scheme: Given a centerline point in the current subset $\mathbf{c_i} \in \widetilde{\mathcal{C}} \subset \mathcal{C} \subset \mathbb{R}^3$, the desired transducer pose is determined as follows. The distance vector between subsequent centerlines points is computed as

$$\mathbf{d} = \mathbf{c_{i+1}} - \mathbf{c_i}\,. \tag{4}$$

Fixing the Z-axis to

$$\mathbf{z} = (0, 0, -1)^T\,, \tag{5}$$

the X-axis is computed as the vector rejection of d on z

$$\mathbf{x} = \frac{(\mathbf{d}_x, \mathbf{d}_y, 0)^T}{|(\mathbf{d}_x, \mathbf{d}_y, 0)^T|} \in \mathbb{R}^3\,, \tag{6}$$

while the Y-axis is simply

$$\mathbf{y} = \mathbf{z} \times \mathbf{x}\,. \tag{7}$$

The final transducer pose is composed as

$$P = \begin{bmatrix}\mathbf{x} & \mathbf{y} & \mathbf{z} & \mathbf{c_i}\end{bmatrix}\,. \tag{8}$$

That is, the US probe is always kept above the current point in an downright position, oriented to follow the centerline direction. A hybrid position/force controller [23] is used throughout the robot motion, such that a chosen fixed force is applied onto the patient surface to maintain acoustic contact for US imaging and prevent harm. The Z component of the commanded pose will be effectively ignored as the controller will adjust it to fulfill its primary task and apply a constant force.

2) Catheter Detection: At the current position $\mathbf{c_i}$, an attempt to detect the catheter in the current 2D US image is performed. A template matching algorithm using normalized cross correlation [24] is employed. A synthetic template that simulates the appearance of a catheter inside a vessel in US is created, as shown in Fig. 4. To minimize the search space, we can assume that the catheter will lay inside the vessel and can be located around the computed centerline point. Consequently, an image patch surrounding the vessel center will be considered for the detection algorithm. Formally, given a 2D US image I, a Region of Interest (ROI) $\widetilde{I} \in \mathbb{R}^{k\times k}$ is chosen. We define the template matching function as

$$\mathrm{TM}\colon \mathbb{R}^{k\times k} \to \mathbb{R} \tag{9}$$

and consider a point to match when

$$\mathrm{TM}(\widetilde{I}) > \tau\,, \tag{10}$$

with $k = 150$ and $\tau = 5 \times 10^6$ in our case. According to the detection response, the index of the centerline point guiding the robot motion is updated as

$$i_{t+1} = \begin{cases} i_t + 1 & \text{if match in } \widetilde{I} \\ i_t - 1 & \text{else}\,. \end{cases} \tag{11}$$

An additional indexing is used for the tracking of the catheter tip location,

$$j_{t+1} = \begin{cases} i_t & \text{if match in } \widetilde{I} \\ i_t - 1 & \text{else}\,. \end{cases} \tag{12}$$

That is, while the robot will continuously be in motion to capture any catheter movements in both possible directions, the location of the catheter is updated to the centerline point $\mathbf{c}_j$ that is closest to the location of its last detection. As a remark, since the distance between the used centerline

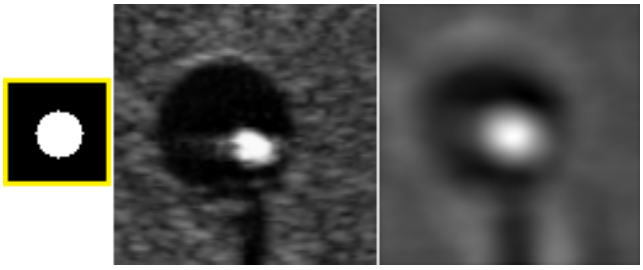

Fig. 4. Catheter detection. Synthetic template (left), selected ROI (center) and the response from template matching (right).

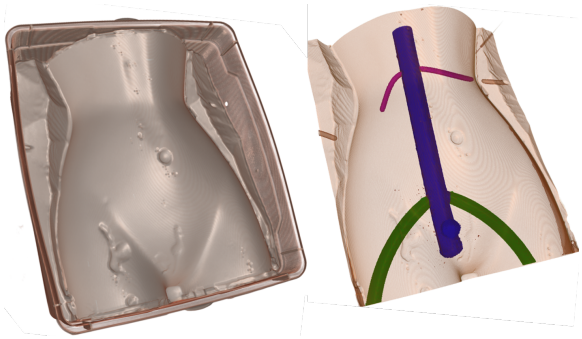

Fig. 5. Experiments are performed on a phantom mimicking a human figure (left) and containing a realistic vessel tree (right). Vessel branches marked in green represent the femoral arteries, in blue the abdominal aortic tract and in red the renal arteries.

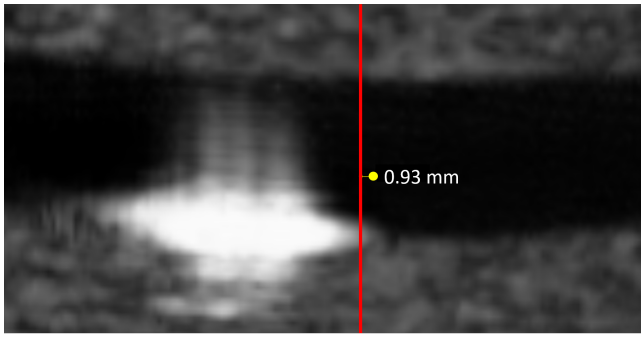

Fig. 6. Tracking validation. The catheter tip location computed by the system (yellow dot) is compared to a manual annotated tip location (red line). Distances are computed along the centerline.

points can be chosen at run-time, the accuracy of the catheter location can be increased or decreased according to the procedure's phase. As a feedback to the user, the tracked catheter location is displayed in 3D, overlayed over the vessel model extracted from the pre-operative volume, as shown in Fig. 8. This enables the user to easily identify the current position with respect to the patient anatomy and use this as guidance information.

III. Results

A. Materials and Experimental Setup

We make use of a KUKA LBR iiwa R800 manipulator (KUKA Roboter GmbH, Augsburg, Germany) controlled via the Robotic Operating System[2] (ROS) using a custom software interface[3]. The robot's end-effector is equipped with a curvilinear US transducer C5-2/60 connected to an Ultrasonix® Sonix RP US system (Analogic Corporation, Peabody, MA, USA). Image processing and robot control is performed on a workstation (Intel i5, NVIDIA GeForce GTX 970) and implemented by custom software plugins integrated to the visualization framework provided by the ImFusion Suite[4] platform (ImFusion GmbH, Munich, Germany). The transmission of US images to the processing station is achieved using a custom messaging library[5].

B. Evaluation

To assess the quality of the proposed system, we individually validate the involved components.

1) Patient Registration: To evaluate system performance on a phantom, as well as on human volunteers, multiple modalities are used for the acquisition of pre-operative data. 3D Turbo Field Echo MRI (TFE, a gradient echo pulse sequence for contrast enhancement of the vasculature) volumes are acquired for the volunteers; while a CT volume is obtained for the experiments performed on the phantom. For trials on volunteers, the methodology described in Sec. II-A (Data Preprocessing) and Sec. II-B (Patient Registration) was performed in all its phases to achieve an optimal registration between the vessel centerlines extracted from the MRI and US volumes. Autonomous robotic US servoing is then performed, using simulated catheter motions. That is, a small subset of points is extracted from the computed vessel centerline and the robot is commanded to their positions, simulating the possible path of an inserted catheter.

At certain intervals, the simulated servoing is paused to acquire small 3D volumes around the current robot location. Assuming a correct registration, this validation volume should already be spatially aligned to the pre-operative volume. In Fig. 7, examples of validation volumes and their spatial relation with the correspondent registered pre-operative volume are shown. The strong agreement between the vessel structures in the US and MRI volumes suggests that the employed registration pipeline allows for a fine guidance based on pre-operative data. In all these autonomous US acquisitions, the hybrid position/force controller was commanded to apply a force of $5\,\mathrm{N}$ along the end-effector Z-axis. For the US, following parameters were used: frequency: $3.3\,\mathrm{MHz}$, depth: $100\,\mathrm{mm}$, gain: 62%. 3D US volumes were compounded using a $0.196\,\mathrm{mm}$ voxel spacing.

2) Catheter Tracking and Robot Servoing: To evaluate the accuracy of the catheter tracking and subsequent robot servoing methodology, a full set of experiments is performed on a US-compatible phantom, shown in Fig. 5. It is built using a 1.7% agar-water mix. Graphite is added to generate stronger US responses, inducing more contrast and speckle noise in the acquired images. The complete pipeline shown in Fig. 2 is executed during these experiments, using a pre-operative CT of the phantom and a clinical vascular catheter (Abbott FlexAbility™, Chicago, IL, USA).

Once the registration steps are performed, the catheter is inserted from either one of the two vessels simulating the human iliac arteries (marked in blue in Fig. 5). Once inserted, the catheter is navigated through the vessel tree while the robot follows its path according to the output of the catheter detection algorithm. Forward and backward motions were performed with the catheter, showing the bi-directional capability of the solution. During large movements of the catheter, points had a distance of $5\,\mathrm{mm}$ from each other,

[2] http://www.ros.org/
[3] https://github.com/IFL-CAMP/iiwa_stack
[4] https://www.imfusion.de/
[5] https://github.com/IFL-CAMP/simple

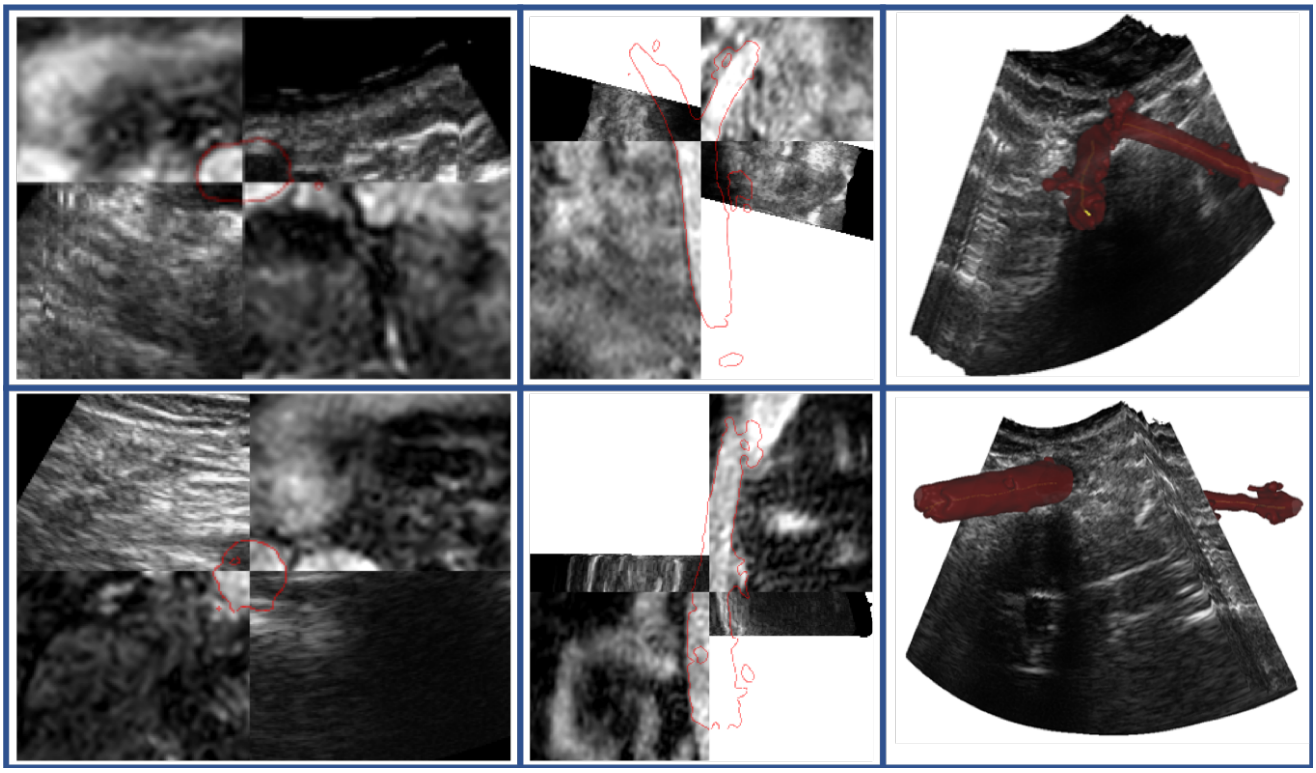

Fig. 7. Registration validation on volunteers. US sweep overlapped with MRI (transverse view - left column, coronal view - center column), and 3D view of a validation sweep with the respective vessel segmented from MRI (right column).

while during a fine search of the catheter tip, we adjust their distance to be $1\,\mathrm{mm}$.

At 25 random locations, the catheter is stopped and a small 3D US volume is acquired around its last detected location. For each of these volumes, we compute the distance between the estimated tip location and a manually annotated ground truth position. This distance is computed only along the longitudinal direction of the vessel, as that represents the error of interest. In Fig. 6 an exemplary measurement of the tracking error is presented. A mean error of $1.78 \pm 1.02\,\mathrm{mm}$ for the localization of the catheter tip has been found. To provide a feedback on the current tracked catheter location to the user, a 3D visualization is provided, an example is given in Fig. 8. The phantom was submerged in water for all experiments, and the following US settings were used: frequency: $5\,\mathrm{MHz}$, depth: $140\,\mathrm{mm}$, gain: 60%.

IV. Discussion

With the performed experiments, we assessed all parts of the proposed workflow. In all simulated catheter insertions performed on human volunteers, the targeted vessel structure was consistently centered in the 2D US images, demonstrating that the proposed methods are able to successfully follow target vessels in real procedures. This is, in part, due to the high quality of the achieved registration between the pre-operative data and the intra-operative 3D US volumes. The examples in Fig. 7 exhibit the resulting image alignment.

During the catheter tracking, we observed strong reflections in US close to the metallic catheter tip, as shown in Fig. 6, where a weaker but still relevant response can be observed to the right of the red line. Those artifacts are a result of the physics underlying US imaging. It is likely they influenced the manual selection of the catheter tip during the validation phase, as well as the used catheter detection method. Nevertheless, the observed tracking error for the detection of the inserted catheter is below $2\,\mathrm{mm}$ on average, which is in line with other presented systems for the tracking and navigation of catheters in endovascular procedures [25].

It is also worth mentioning that any component of the

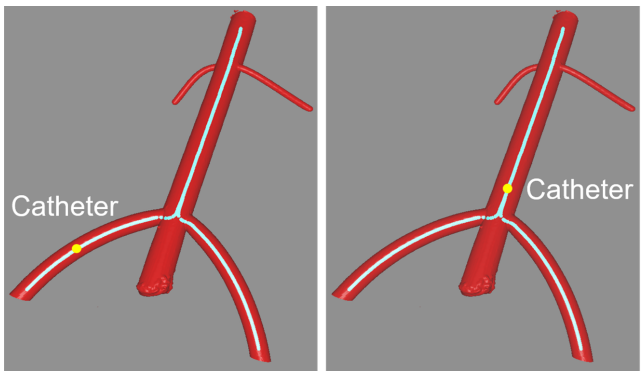

Fig. 8. The vasculature "road map" as displayed to the end-user for visual navigation during endovascular catheter insertions.

presented system could be exchanged for other techniques that provide a similar output without the need to incur in considerable modifications of the system; displaying the high modularity and flexibility of our solution.

While US imaging of the aorta on volunteers has been proven feasible, there is a high variability in the quality of the resulting images due to various factors (e.g., intestinal gas, lack of gel coupling). These are challenges future intelligent robotic US systems will need to tackle to achieve the required robustness for their introduction into clinical routine. For instance, methods to quantify the overall visibility of the target anatomy could be employed to detect possible failure cases. Yet, even without the use of those measures we could show the feasibility of the proposed system.

V. Conclusions

As of now, endovascular procedures heavily rely on fluoroscopic X-ray imaging for catheter guidance within patient vasculature. This work explores how robotic US can be integrated into the intra-operative workflow to provide catheter visualization and intuitive navigation. By combining knowledge of the vascular structure obtained from pre-operative data with interventional 3D US, the system proposed in this work enables for planning of robot trajectories over the targeted vessels, in order to provide live catheter tracking and navigation without exposure to ionizing radiation.

Autonomous robot servoing over the target anatomy was demonstrated by *simulating* catheter insertions on healthy volunteers and *performing* them on a custom phantom. An inter-modality two-step registration pipeline was validated during human experiments while the catheter detection and tracking performance were assessed through experiments performed on a phantom. In conclusion, this work demonstrates the potential of the integration between robotic systems and medical imaging to provide full solutions for interventional applications. It establishes ground concepts that cover the existing medical requirements, exemplifying the envisioned future role of robotics in endovascular procedures.

ACKNOWLEDGMENT

We thank ImFusion GmbH, Munich, Germany, for their image processing framework and their continuous support.

References

[1] X. Li, G. Zhao, J. Zhang, Z. Duan, and S. Xin, "Prevalence and trends of the abdominal aortic aneurysms epidemic in general population–a meta-analysis," *PLOS One*, vol. 8, no. 12, 2013.

[2] E. L. Chaikof, J. D. Blankensteijn, P. L. Harris, G. H. White, C. K. Zarins, V. M. Bernhard, J. S. Matsumura, J. May, F. J. Veith, M. F. Fillinger *et al.*, "Reporting standards for endovascular aortic aneurysm repair," *J VASC SURG*, vol. 35, no. 5, pp. 1048–1060, 2002.

[3] M. Davis, M. Harris, and J. J. Earnshaw, "Implementation of the national health service abdominal aortic aneurysm screening program in england," *J VASC SURG*, vol. 57, no. 5, pp. 1440–1445, 2013.

[4] S. E. Salcudean, W. H. Zhu, P. Abolmaesumi, S. Bachmann, and P. D. Lawrence, "A robot system for medical ultrasound," in *Robotics Research*, 2000, pp. 195–202.

[5] L. Lindenroth, A. Soor, J. Hutchinson, A. Shafi, J. Back, K. Rhode, and H. Liu, "Design of a soft, parallel end-effector applied to robot-guided ultrasound interventions," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2017.

[6] C. Hennersperger, B. Fuerst, S. Virga, O. Zettinig, B. Frisch, T. Neff, and N. Navab, "Towards mri-based autonomous robotic us acquisitions: A first feasibility study," *IEEE T MED IMAGING*, vol. 36, no. 2, pp. 538–548, 2017.

[7] S. Virga, O. Zettinig, M. Esposito, K. Pfister, B. Frisch, T. Neff, N. Navab, and C. Hennersperger, "Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2016, pp. 508–513.

[8] R. Elek, T. D. Nagy, D. Á. Nagy, B. Takács, P. Galambos, I. Rudas, and T. Haidegger, "Robotic platforms for ultrasound diagnostics and treatment," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1752–1757.

[9] P. Chatelain, A. Krupa, and M. Marchal, "Real-time needle detection and tracking using a visually servoed 3d ultrasound probe," in *IEEE International Conference on Robotics and Automation*, 2013, pp. 1676–1681.

[10] G. J. Vrooijink, M. Abayazid, and S. Misra, "Real-time three-dimensional flexible needle tracking using two-dimensional ultrasound," in *IEEE International Conference on Robotics and Automation*, 2013, pp. 1688–1693.

[11] S. B. Kesner and R. D. Howe, "Robotic catheter cardiac ablation combining ultrasound guidance and force control," *INT J ROBOT RES*, vol. 33, no. 4, pp. 631–644, 2014.

[12] M. Azizian, N. Najmaei, M. Khoshnam, and R. Patel, "Visual servoing in medical robotics: a survey. part II: tomographic imaging modalities–techniques and applications," *INT J MED ROBOT COMP*, vol. 11, no. 1, pp. 67–79, 2015.

[13] Q. Ma, J. D. Davis, A. Cheng, Y. Kim, G. S. Chirikjian, and E. M. Boctor, "A new robotic ultrasound system for tracking a catheter with an active piezoelectric element," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2016, pp. 2321–2328.

[14] L. Antiga, "Patient-specific modeling of geometry and blood flow in large arteries," *Politecnico di Milano*, 2002.

[15] C. Hennersperger, A. Karamalis, and N. Navab, "Vascular 3d+ t freehand ultrasound using correlation of doppler and pulse-oximetry data," in *International Conference on Information Processing in Computer-Assisted Interventions*, 2014, pp. 68–77.

[16] B. Fuerst, W. Wein, M. Müller, and N. Navab, "Automatic ultrasound–mri registration for neurosurgery using the 2d and 3d lc2 metric," *MED IMAGE ANAL*, vol. 18, no. 8, pp. 1312–1319, 2014.

[17] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 1998, pp. 130–137.

[18] C. Hennersperger, M. Baust, P. Waelkens, A. Karamalis, S.-A. Ahmadi, and N. Navab, "Multi-scale tubular structure detection in ultrasound imaging," *IEEE T MED IMAGING*, vol. 34, no. 1, pp. 13–26, 2015.

[19] G. P. Penney, J. M. Blackall, M. Hamady, T. Sabharwal, A. Adam, and D. J. Hawkes, "Registration of freehand 3d ultrasound and magnetic resonance liver images," *MED IMAGE ANAL*, vol. 8, no. 1, pp. 81–91, 2004.

[20] Y. Song, J. Totz, S. Thompson, S. Johnsen, D. Barratt, C. Schneider, K. Gurusamy, B. Davidson, S. Ourselin, D. Hawkes *et al.*, "Locally rigid, vessel-based registration for laparoscopic liver surgery," *INT J COMPUT ASS RAD*, vol. 10, no. 12, pp. 1951–1961, 2015.

[21] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor Fusion IV: Control Paradigms and Data Structures*, vol. 1611, 1992, pp. 586–607.

[22] R. B. Rusu, "Semantic 3d object maps for everyday manipulation in human living environments," *KI-Künstliche Intelligenz*, vol. 24, no. 4, pp. 345–348, 2010.

[23] M. H. Raibert and J. J. Craig, "Hybrid position/force control of manipulators," *J DYN SYST MEAS CONTROL*, vol. 103, no. 2, pp. 126–133, 1981.

[24] J. P. Lewis, "Fast template matching," in *Vision interface*, vol. 95, no. 120123, 1995, pp. 15–19.

[25] Q. M. de Ruiter, F. L. Moll, and J. A. Van Herwaarden, "Current state in tracking and robotic navigation systems for application in endovascular aortic aneurysm repair," *J VASC SURG*, vol. 61, no. 1, pp. 256–264, 2015.

Part III

Conclusions and Outlook

6 Summary of Findings

The presented contributions base their strength on the three technological pillars already discussed in Chapters 2 to 4. Ultrasound imaging is the core technique used to obtain live information about the observed anatomy. Force sensing is employed to ensure a contact force that is safe for the patient and clinically valid for the required acquisition. 3D visual sensing provides information about the system surroundings and the current patient position, enabling planned information to be accurately transferred onto the targeted anatomical area.

Autonomy in Robotic Ultrasound With respect to the presented literature in the field of robotic ultrasound, the overall work demonstrates various improvements and brings novelty. As an initial groundwork, in [7], a complete system for autonomous robotic ultrasound is presented. To benefit the research community, a rigorous definition of the required tasks (e.g., calibration steps among the various components) to reproduce such a system is proposed. As a first in literature, to the best of the author's knowledge, this work introduces the definition of an autonomous robotic task to be performed on a human based on pre-operative information. A complete workflow to achieve that in a safe and robust manner is presented. It is leveraging all the technical components described in this thesis that such a system is able to achieve precise acquisitions comparable to the ones acquired by a professional figures trained in ultrasound imaging. While most of the previous state-of-the-art focused on teleoperation of a robotic system or on individual tasks to solve specific problems (e.g., anatomy or tool tracking), this work introduces the concept of autonomy for this type of imaging procedures, end to end. In this case, vision sensing controls the manipulator actions and decision-making, while force sensing satisfies the clinical requirements of patient and staff safety (i.e., controlling the force applied onto the patient and external force applied onto the robot in case of collision).

Force Sensing to Improve Image Quality Control strategies can leverage visual information, as shown by [7] and many other works focused on visual servoing. The aspect of using force sensing to improve the imaging output of such a system, instead, was rarely investigated. That is, in [4] the key to achieve better clinical results is the *haptic* sensing of a robotic system. The lack of a human operator in robotic ultrasound tasks can often be seen as a limiting factor. The use of haptic sensing, in fact, is natural for a human and can provide significant information in the case of ultrasound acquisition, i.e., palpation of the examined area is often common. Feeling a change in tissue stiffness by using hand contact can lead to a more precise diagnostic outcome. In this sense, leveraging this sensing information, not only to perform safe and compliant acquisitions but also to add value to the system's output is an absolute need. In [4], we demonstrate this concept by producing 3D ultrasound volumes that are not strongly deformed by the applied contact force. Force and visual feedback are here employed together to achieve the result: vision-based techniques are used to characterize the tissue movements with respect to changes in force sensing. While other techniques are also shown to be able to recover the induced deformation (e.g., using biomechanical models or additional

sensing), we demonstrate with this contribution how the system introduced in [7] has all the sensing capabilities necessary to achieve so, effectively adding an additional benefit from the use of force information. Not only a mere input parameter for the control loop strategy but an integral part of the imaging pipeline.

From Concept to Clinical Application The development and validation of the technology behind a fully autonomous robotic ultrasound system would be a pure scientific exercise without demonstrating its abilities in targeted clinical applications. Many applications would benefit from such a system, and literature has already explored some of them. As a contribution to the field, this thesis aims at establishing the role of such robotic platform both in diagnostic and interventional scenarios. Vascular surgery applications have been presented in [8] and [1]. The former extends the concepts from [7] to streamline a repetitive operation to be performed on a large scale during AAA screening programs. The system shows capabilities in the assessment of the current patient state that are equivalent to the ones of an expert vascular surgeon. Potentially, AAA screening programs that are not implemented for the lack of personnel could be transformed by the introduction of a completely automatic system. From a technical standpoint, this work introduces the online optimization of the used hybrid force/motion control algorithm based on live ultrasound data in a complete patient-specific way. The ultrasound acquisition is, therefore, not anymore based on the static transfer of a planned trajectory and its blind execution. Intelligent adaptation of the probe position and orientation, as well as the applied force, during the whole movement, allows for a more dynamic and less error-prone execution. The latter work, instead, focuses on the treatment of the diagnosed pathology. The presented task is a pure visual servoing one: to track and follow a catheter inserted into the vasculature. The system validation demonstrates that the achieved tracking accuracy can support the use of such a system in clinical practice, while the modular approach of the used methods allows to dynamically exchange any of the components with different techniques. An essential aspect of the introduction of autonomous ultrasound systems in interventional settings would be reducing the dependency from expert operators being available on-site during the operation and reduce the use of X-ray radiation. All the techniques discussed, implemented, and validated by the presented contributions move the technology towards the application of such systems into real clinical scenarios.

An important aspect to note is that throughout all the contributions of this dissertation, the author chose to validate, in complete safety and following ethical standards, his proposed solutions on human volunteers. In fact, to the best of the author's knowledge, extremely few robotic ultrasound systems have been assessed on humans in the available literature prior to [7]. The use of synthetic validation often minimizes the impact of some technological solutions, it is important, instead, for roboticists in the medical field to be aware of the needs and complexities that real clinical applications have and strive for the integration of their solution in real practice. Noteworthy is that an extension of the system used in the presented contributions has been successfully applied in a neurosurgical application on actual human patients [3].

7 Future Directions

With the presented contributions, the author hopes to have exposed the technological potential for ultrasound imaging to become a more reproducible and less user-dependent imaging modality. A system similar to the one presented could potentially be perfected to become a truly clinical device for autonomous acquisition, filling the gap with other modalities (e.g., CT and MRI) that already reached that autonomy.

Many improvements and future work could be beneficial to this final goal. From a complete system design perspective, current robotic ultrasound systems often translate into a pure integration task. It is, in fact, a complex task to assemble the various system components (i.e., ultrasound machine, robotic manipulator, external sensor) into a coherent system, with the constant need of various calibration steps. Looking back to the previous literature in Sec. 1.3, after a first large wave of research was oriented on the design of custom robotic systems for ultrasound imaging. Since the advent of modern robotics design (i.e., robots capable of human-robot interaction) in commercial systems, new designs are less and less frequent. The author believes that the current research has made full use of the current hardware characteristics, while new custom designs that specifically target ultrasound imaging could be introduced. Modern designs should, in fact, target at the direct assembly of all system components into an individual one, by construction. A focus on specialized end-effector design could provide a seamless solution to manipulate and exchange multiple ultrasound transducers. After all, the current design of ultrasound probes is strictly linked to their hand-held use, transducers designed for robotic systems could have very different features. Additionally, the inclusion on data transmission capabilities within the manipulator body is also a possibility since modern ultrasound image generation pipelines can be integrated into custom embedded platforms; this solution would result in a cable-free ultrasound connection, reducing control strategy complexity, since the external torques applied by a floating cable are hard to model. Finally, an acceptance of the system into the current clinical practice would also be easier with smaller form factors and integration of the kinematic structure into the medical beds used for examinations.

Regarding the technology used in this work, especially in the area of Computer Vision, many techniques do not take leverage of modern data-driven approached. The strong focus of the research community in the definition and application of machine learning algorithms has led to significant advancements in the field and generated a plethora of available solutions. Various image analysis tasks have been targeted by solutions using learning techniques, including classification, segmentation, detection, registration and quality assessment, and focusing of various anatomical structures, such as prostate, breast, liver, heart, spine, etc. Reviews of the current state-of-the-art in the field are available [161, 162, 163]. Due to the high accuracy and success rate of such methods, their integration to the sensing pipeline of a robotic ultrasound system could spark its use in further applications or enable more sophisticated control strategies. An exciting learning approach, already largely employed in the robotic

research field, is reinforcement learning. This rewards-based learning technique has shown promising results in the definition of robotic tasks based on simulated extensive trials [164]. In ultrasound imaging, a reinforcement-learning technique to achieve a desired anatomical view in manual transducer guidance has been proposed in [165]. The integration of such a technique into a robotic motion control loop would be a natural extension, providing to an autonomous system the ability to perfect its position and guarantee the desired acquisition from a rough initial trajectory. Other learning-based techniques could be employed to estimate the tissue deformation computed in [4], strongly reducing the computation costs of the current solution; or to introduce real-time motion compensation techniques (e.g., for respiratory motion) to obtain artifact-free ultrasound volumes. In general, a modular approach in the design of the system software architecture allows for the integration of new algorithms with ease and should always be preferred.

The autonomous manipulation of an ultrasound device opens to the possibility of combining multiple imaging modalities imaging during the same diagnostic or intervention procedure, an opportunity that could target either combined robotic imaging systems or approaches based human-robot collaboration. In the first case, robotic systems for other modalities (e.g., cone-beam CT using robotic C-Arm systems) and a robotic ultrasound system could be co-calibrated to perform their acquisition simultaneously, effectively augmenting the imaging output with more information. Moreover, the two imaging modalities could support each other in optimizing their acquisition, e.g., the live imaging provided from the ultrasound system could be used to compensate for patient movements. Multi-modal collaborative approaches have already been explored in literature [166], although without taking into account possible physical interactions between humans and the system.

Finally, an essential component, often neglected in the literature, is the social and ethical impact of such technology and its acceptance by the clinical staff and the patients. The social aspect of human-robot interaction is, in general, a field of study on its own [167, 168, 169], although with limited coverage of medical robotic systems. Studies in this field would allow to better design hardware and software that can be welcomed into the clinical practice, for instance, taking into account possible patient discomfort and act in real-time to minimize it or adjust the mechanical design of the current system to accommodate the everyday needs of interventional rooms. Often, pure engineering approaches strive for technical correctness without including any feedback from the end-users and recipients of the final system. The contributions presented in this dissertation have always been developed in close collaboration with medical experts, and the author strongly encourages the research community to leverage such a valuable asset, following the example of excellence research groups that are more and more directly integrated into the medical facilities and faculties.

Part IV

Appendix

A Abstracts of Publications not Discussed in this Dissertation

Markerless Inside-Out Tracking for 3D Ultrasound Compounding [2]

Benjamin Busam, Patrick Ruhkamp, Salvatore Virga, Beatrice Lentes, Julia Rackerseder, Nassir Navab and Christoph Hennersperger

Tracking of rotation and translation of medical instruments plays a substantial role in many modern interventions and is essential for 3D ultrasound compounding. Traditional external optical tracking systems are often subject to line-of-sight issues, in particular when the region of interest is difficult to access. The introduction of inside-out tracking systems aims to overcome these issues. We propose a marker-less tracking system based on visual SLAM to enable tracking of ultrasound probes in an interventional scenario. To achieve this goal, we mount a miniature multi-modal (mono, stereo, active depth) vision system on the object of interest and relocalize its pose within an adaptive map of the operating room. We compare state-of-the-art algorithmic pipelines and apply the idea to transrectal 3D ultrasound (TRUS). Obtained volumes are compared to reconstruction using a commercial optical tracking system as well as a robotic manipulator. Feature-based binocular SLAM is identified as the most promising method and is tested extensively in challenging clinical environments and for the use case of prostate US biopsies.

Point-Of-Care Ultrasound Workshop Held in Conjunction with MICCAI 2018, Granada, Spain (2018)

Robotic Ultrasound-Guided Facet Joint Insertion [3]

Javier Esteban, Walter Simson, Sebastian Requena Witzig, Anna Rienmüller, Salvatore Virga, Benjamin Frisch, Oliver Zettinig, Drazen Sakara, Yu-Mi Ryang, Nassir Navab and Christoph Hennersperger

Purpose. *Facet joint insertion is a common treatment of chronic pain in the back and spine. This procedure is often performed under fluoroscopic guidance, where the staff's repetitive radiation exposure remains an unsolved problem. Robotic ultrasound (rUS) has the potential to reduce or even eliminate the use of radiation by using ultrasound with a robotic-guided needle insertion. This work presents first clinical data of rUS-based needle insertions extending previous work of our group.*

Methods. *Our system implements an automatic US acquisition protocol combined with a*

calibrated needle targeting system. This approach assists the physician by positioning the needle holder on a trajectory selected in a 3D US volume of the spine.
***Results**. By the time of submission, nine facets were treated with our approach as first data from an ongoing clinical study. The insertion success rate was shown to be comparable to current clinical practice. Furthermore, US imaging offers additional anatomical context for needle trajectory planning.*
***Conclusion**. This work shows first clinical data for robotic ultrasound-assisted facet joint insertion as a promising solution that can easily be incorporated into the clinical workflow. Presented results show the clinical value of such a system.*

International Journal of Computer Assisted Radiology and Surgery (2018)

3D Ultrasound Registration-Based Visual Servoing for Neurosurgical Navigation [5]

Oliver Zettinig, Benjamin Frisch, Salvatore Virga, Marco Esposito, Anna Rienmüller, Bernhard Meyer, Christoph Hennersperger, Yu-Mi Ryang and Nassir Navab

***Purpose**. We present a fully image-based visual servoing framework for neurosurgical navigation and needle guidance. The proposed servo-control scheme allows for compensation of target anatomy movements, maintaining high navigational accuracy over time, and automatic needle guide alignment for accurate manual insertions.*
***Method**. Our system comprises a motorized 3D ultrasound (US) transducer mounted on a robotic arm and equipped with a needle guide. It continuously registers US sweeps in real-time with a pre-interventional plan based on CT or MR images and annotations. While a visual control law maintains anatomy visibility and alignment of the needle guide, a force controller is employed for acoustic coupling and tissue pressure. We validate the servoing capabilities of our method on a geometric gel phantom and real human anatomy, and the needle targeting accuracy using CT images on a lumbar spine gel phantom under neurosurgery conditions.*
***Results**. Despite the varying resolution of the acquired 3D sweeps, we achieved direction-independent positioning errors of 0.35 ± 0.19 mm and 0.61◦ ± 0.45◦, respectively. Our method is capable of compensating movements of around 25 mm/s and works reliably on human anatomy with errors of 1.45 ± 0.78 mm. In all four manual insertions by an expert surgeon, a needle could be successfully inserted into the facet joint, with an estimated targeting accuracy of 1.33 ± 0.33 mm, superior to the gold standard.*
***Conclusion**. The experiments demonstrated the feasibility of robotic ultrasound based navigation and needle guidance for neurosurgical applications such as lumbar spine injections.*

International Journal of Computer Assisted Radiology and Surgery (2017)

Acoustic Window Planning for Ultrasound Acquisition [7]

Rüdiger Göbl, Salvatore Virga, Julia Rackerseder, Benjamin Frisch, Nassir Navab and Christoph Hennersperger

Autonomous robotic ultrasound has recently gained considerable interest, especially for collaborative applications. Existing methods for acquisition trajectory planning are solely based on geometrical considerations, such as the pose of the transducer with respect to the patient surface.
***Purpose**. This work aims at establishing acoustic window planning to enable autonomous ultrasound acquisitions of anatomies with restricted acoustic windows, such as the liver or the heart.*
***Methods**. We propose a fully automatic approach for the planning of acquisition trajectories, which only requires information about the target region as well as existing tomographic imaging data, such as X-ray computed tomography. The framework integrates both geometrical and physics-based constraints to estimate the best ultrasound acquisition trajectories with respect to the available acoustic windows. We evaluate the developed method using virtual planning scenarios based on real patient data as well as for real robotic ultrasound acquisitions on a tissue-mimicking phantom.*
***Results**. The proposed method yields superior image quality in comparison with a naive planning approach, while maintaining the necessary coverage of the target.*
***Conclusion**. We demonstrate that by taking image formation properties into account acquisition planning methods can outperform naive plannings. Furthermore, we show the need for such planning techniques, since naive approaches are not sufficient as they do not take the expected image quality into account.*

International Journal of Computer Assisted Radiology and Surgery (2017)

Optimal C-Arm Positioning for Aortic Interventions [9]

Salvatore Virga, Verena Dogeanu, Pascal Fallavollita, Reza Ghotbi, Nassir Navab and Stefanie Demirci.

Due to the continuous integration of innovative interventional imaging modalities into vascular surgery rooms, there is an urgent need for computer assisted interaction and visualization solutions that support the smooth integration of technological solutions within the surgical workflow. In this paper, we introduce a new paradigm for optimal-view controlled maneuvering of Angiographic C-arms during thoracic endovascular aortic repair (TEVAR). This allows the semi-automatic pre-computation of well-defined anatomy-related optimal views based on pre-operative 3D image data and automatic interventional positioning of the imaging device relative to the patient's anatomy through inverse kinematics and CT to patient registration. Together with our clinical partners, we have evaluated the new technique using 5 patient datasets and are able to show promising results.

Proceedings of Workshop Bildverarbeitung für die Medizin (2015)

Bibliography

[1] F. Langsch, S. Virga, J. Esteban, R. Göbl, and N. Navab. “Robotic Ultrasound for Catheter Navigation in Endovascular Procedures”. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2019, pp. 508–513 (cit. on pp. 1, 7, 24, 41, 90).

[2] B. Busam, P. Ruhkamp, S. Virga, et al. “Markerless Inside-Out Tracking for 3D Ultrasound Compounding”. In: *POCUS 2018, BIVPCS 2018, CuRIOUS 2018, CPM 2018. Lecture Notes in Computer Science, vol 11042*. Springer, 2018, pp. 56–64 (cit. on pp. 1, 20, 95).

[3] J. Esteban, W. Simson, S. R. Witzig, et al. “Robotic ultrasound-guided facet joint insertion”. In: *International journal of computer assisted radiology and surgery* 13.6 (2018), pp. 895–904 (cit. on pp. 1, 10, 90, 95).

[4] S. Virga, R. Göbl, M. Baust, N. Navab, and C. Hennersperger. “Use the force: Deformation correction in robotic 3D ultrasound”. In: *International journal of computer assisted radiology and surgery* 13.5 (2018), pp. 619–627 (cit. on pp. 1, 7, 89, 92).

[5] O. Zettinig, B. Frisch, S. Virga, et al. “3D ultrasound registration-based visual servoing for neurosurgical navigation”. In: *International journal of computer assisted radiology and surgery* 12.9 (2017), pp. 1607–1619 (cit. on pp. 1, 96).

[6] R. Göbl, S. Virga, J. Rackerseder, B. Frisch, N. Navab, and C. Hennersperger. “Acoustic window planning for ultrasound acquisition”. In: *International journal of computer assisted radiology and surgery* 12.6 (2017), pp. 993–1001 (cit. on p. 1).

[7] C. Hennersperger, B. Fuerst, S. Virga, et al. “Towards MRI-based autonomous robotic US acquisitions: a first feasibility study”. In: *IEEE transactions on medical imaging* 36.2 (2016), pp. 538–548 (cit. on pp. 2, 7, 24, 39, 41, 89, 90, 97).

[8] S. Virga, O. Zettinig, M. Esposito, et al. “Automatic force-compliant robotic ultrasound screening of abdominal aortic aneurysms”. In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2016, pp. 508–513 (cit. on pp. 2, 7, 39, 90).

[9] S. Virga, V. Dogeanu, P. Fallavollita, R. Ghotbi, N. Navab, and S. Demirci. “Optimal c-arm positioning for aortic interventions”. In: *Bildverarbeitung für die Medizin 2015*. Springer, 2015, pp. 53–58 (cit. on pp. 2, 97).

[10] M Röthlin, R Näf, M Amgwerd, D Candinas, and O Trentz. “How much experience is required for ultrasound diagnosis of blunt abdominal trauma?” In: *Langenbecks Archiv fur Chirurgie* 377.4 (1992), pp. 211–215 (cit. on p. 6).

[11] M. L. Chuang, R. A. Parker, M. F. Riley, et al. “Three-dimensional echocardiography improves accuracy and compensates for sonographer inexperience in assessment of left ventricular ejection fraction”. In: *Journal of the American Society of Echocardiography* 12.5 (1999), pp. 290–299 (cit. on p. 6).

[12] D. Ehler, D. K. Carney, A. L. Dempsey, et al. “Guidelines for cardiac sonographer education: recommendations of the American Society of Echocardiography Sonographer Training and Education Committee”. In: *Journal of the American society of echocardiography* 14.1 (2001), pp. 77–84 (cit. on p. 6).

[13] L Brulhart, H. Ziswiler, G Tamborrini, and P Zufferey. “The importance of sonographer experience and machine quality with regards to the role of musculoskeletal ultrasound in routine care of rheumatoid arthritis patients.” In: *Clinical and experimental rheumatology* 33.1 (2015), pp. 98–101 (cit. on p. 6).

[14] M. Necas. “Musculoskeletal symptomnatology and repetitive strain injuries in diagnostic medical sonographers: A pilot study in Washington and Oregon”. In: *Journal of Diagnostic Medical Sonography* 12.6 (1996), pp. 266–273 (cit. on p. 6).

[15] M. Val Gregory and M Dmu. “Musculoskeletal injuries: an occupational health and safety issue in sonography”. In: *Educational Supplement* (1998), pp. 1–3 (cit. on p. 6).

[16] A. Russo, C. Murphy, V. Lessoway, and J. Berkowitz. “The prevalence of musculoskeletal symptoms among British Columbia sonographers”. In: *Applied Ergonomics* 33.5 (2002), pp. 385–393 (cit. on p. 6).

[17] G. Brown. “Work related musculoskeletal disorders in sonographers”. In: *BMUS Bulletin* 11.3 (2003), pp. 6–13 (cit. on p. 6).

[18] K. Evans, S. Roll, and J. Baker. “Work-related musculoskeletal disorders (WRMSD) among registered diagnostic medical sonographers and vascular technologists: a representative sample”. In: *Journal of Diagnostic Medical Sonography* 25.6 (2009), pp. 287–299 (cit. on p. 6).

[19] R. H. Taylor, J. Funda, D. D. Grossman, J. P. Karidis, and D. A. LaRose. *Remote center-of-motion robot for surgery*. US Patent 5,397,323. 1995 (cit. on p. 7).

[20] E. J. Hanly and M. A. Talamini. “Robotic abdominal surgery”. In: *The American journal of surgery* 188.4 (2004), pp. 19–26 (cit. on p. 7).

[21] G. S. Weinstein, B. W. O’Malley, W. Snyder, E. Sherman, and H. Quon. “Transoral robotic surgery: radical tonsillectomy”. In: *Archives of otolaryngology–head & neck surgery* 133.12 (2007), pp. 1220–1226 (cit. on p. 7).

[22] C.-C. Abbou, A. Hoznek, L. Salomon, et al. “Laparoscopic radical prostatectomy with a remote controlled robot”. In: *The Journal of urology* 165.6 Part 1 (2001), pp. 1964–1966 (cit. on p. 7).

[23] A. W. Partin, J. B. Adams, R. G. Moore, and L. R. Kavoussi. “Complete robot-assisted laparoscopic urologic surgery: a preliminary report.” In: *Journal of the American College of Surgeons* 181.6 (1995), pp. 552–557 (cit. on p. 7).

[24] D. D. Thiel and H. N. Winfield. “Robotics in urology: past, present, and future”. In: *Journal of endourology* 22.4 (2008), pp. 825–830 (cit. on p. 7).

[25] R. H. Taylor, A. Menciassi, G. Fichtinger, P. Fiorini, and P. Dario. “Medical robotics and computer-integrated surgery”. In: *Springer handbook of robotics*. Springer, 2016, pp. 1657–1684 (cit. on p. 8).

[26] A. D. Pearle, D. Kendoff, V. Stueber, V. Musahl, and J. A. Repicci. “Perioperative management of unicompartmental knee arthroplasty using the MAKO robotic arm system (MAKOplasty)”. In: *American Journal of Orthopedics* 38.2 (2009), pp. 16–19 (cit. on p. 8).

[27] J. Lang, S Mannava, A. Floyd, et al. “Robotic systems in orthopaedic surgery”. In: *The Journal of bone and joint surgery. British volume* 93.10 (2011), pp. 1296–1299 (cit. on p. 8).

[28] C. Faria, W. Erlhagen, M. Rito, E. De Momi, G. Ferrigno, and E. Bicho. “Review of robotic technology for stereotactic neurosurgery”. In: *IEEE reviews in biomedical engineering* 8 (2015), pp. 125–137 (cit. on p. 8).

[29] K. J. Chun, B. Schmidt, B. Köktürk, et al. “Catheter ablation–new developments in robotics”. In: *Herz Kardiovaskuläre Erkrankungen* 33.8 (2008), pp. 586–589 (cit. on p. 8).

[30] W. Saliba, V. Y. Reddy, O. Wazni, et al. “Atrial fibrillation ablation using a robotic catheter remote control system: initial human experience and long-term follow-up results”. In: *Journal of the American College of Cardiology* 51.25 (2008), pp. 2407–2411 (cit. on p. 8).

[31] J. R. Adler Jr, S. D. Chang, M. J. Murphy, J. Doty, P. Geis, and S. L. Hancock. “The Cyberknife: a frameless robotic system for radiosurgery”. In: *Stereotactic and functional neurosurgery* 69.1-4 (1997), pp. 124–128 (cit. on p. 8).

[32] R. D. Howe and Y. Matsuoka. “Robotics for surgery”. In: *Annual review of biomedical engineering* 1.1 (1999), pp. 211–240 (cit. on p. 8).

[33] B. Davies. “A review of robotics in surgery”. In: *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* 214.1 (2000), pp. 129–140 (cit. on p. 8).

[34] R. H. Taylor. “A perspective on medical robotics”. In: *Proceedings of the IEEE* 94.9 (2006), pp. 1652–1664 (cit. on p. 8).

[35] R. A. Beasley. “Medical robots: current systems and research directions”. In: *Journal of Robotics* 2012 (2012) (cit. on p. 8).

[36] J. W. Sublett, B. J. Dempsey, and A. C. Weaver. “Design and implementation of a digital teleultrasound system for real-time remote diagnosis”. In: *Proceedings Eighth IEEE Symposium on Computer-Based Medical Systems*. IEEE. 1995, pp. 292–298 (cit. on p. 8).

[37] S. E. Salcudean, G. Bell, S. Bachmann, W.-H. Zhu, P. Abolmaesumi, and P. D. Lawrence. “Robot-assisted diagnostic ultrasound–design and feasibility experiments”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 1999, pp. 1062–1071 (cit. on p. 8).

[38] S. E. Salcudean, W. H. Zhu, P Abolmaesumi, S. Bachmann, and P. D. Lawrence. “A robot system for medical ultrasound”. In: *Robotics Research*. Springer, 2000, pp. 195–202 (cit. on p. 8).

[39] W.-H. Zhu, S. Salcudean, S. Bachmann, and P. Abolmaesumi. “Motion/force/image control of a diagnostic ultrasound robot”. In: *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*. Vol. 2. IEEE. 2000, pp. 1580–1585 (cit. on p. 8).

[40] P. Abolmaesumi, S. Salcudean, W.-H. Zhu, S. P. DiMaio, and M. R. Sirouspour. “A user interface for robot-assisted diagnostic ultrasound”. In: *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*. Vol. 2. IEEE. 2001, pp. 1549–1554 (cit. on p. 8).

[41] P Abolmaesumi, M. Sirouspour, S. Salcudean, and W. Zhu. “Adaptive image servo controller for robot-assisted diagnostic ultrasound”. In: *2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Proceedings (Cat. No. 01TH8556)*. Vol. 2. IEEE. 2001, pp. 1199–1204 (cit. on p. 8).

[42] P. Abolmaesumi, S. E. Salcudean, W.-H. Zhu, M. R. Sirouspour, and S. P. DiMaio. “Image-guided control of a robot for medical ultrasound”. In: *IEEE Transactions on Robotics and Automation* 18.1 (2002), pp. 11–23 (cit. on p. 8).

[43] M. Mitsuishi, S. Warisawa, T. Tsuda, et al. “Remote ultrasound diagnostic system”. In: *Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164)*. Vol. 2. IEEE. 2001, pp. 1567–1574 (cit. on p. 8).

[44] G Kontaxakis, S Walter, and G Sakas. “An integrated portable telemedicine workstation featuring acquisition processing and transmission over low-bandwidth lines of 3D ultrasound volume images”. In: *Information Technology Applications in Biomedicine. ITAB 2000*. 2000 (cit. on p. 8).

[45] F Courreges, N Smith, G Poisson, et al. “Real-time exhibition of a simulated space tele-echography using an ultra-light robot”. In: *Int. Symp. on Artif. Int., Rob. and Auto. in Space, i-SAIRAS*. Vol. 1. 2001, pp. 15–16 (cit. on p. 8).

[46] A. Vilchis, J. Troccaz, P. Cinquin, K. Masuda, and F. Pellissier. “A new robot architecture for tele-echography”. In: *IEEE Transactions on Robotics and Automation* 19.5 (2003), pp. 922–926 (cit. on p. 8).

[47] T. Martinelli, J.-L. Bosson, L. Bressollette, et al. “Robot-Based Tele-Echography: Clinical Evaluation of the TER System in Abdominal Aortic Exploration”. In: *Journal of Ultrasound in Medicine* 26.11 (2007), pp. 1611–1616 (cit. on p. 8).

[48] C Delgorge, L Al Bassit, C Novales, et al. “OTELO project: mObile Tele-Echography using an ultra-Light rObot”. In: *Telemed’02* (2002) (cit. on p. 8).

[49] C. Delgorge, F. Courrèges, L. A. Bassit, et al. “A tele-operated mobile ultrasound scanner using a light-weight robot”. In: *IEEE transactions on information technology in biomedicine* 9.1 (2005), pp. 50–58 (cit. on p. 8).

[50] P. Vieyres, G. Poisson, F. Courrèges, N. Smith-Guerin, C. Novales, and P. Arbeille. “A tele-operated robotic system for mobile tele-echography: The OTELO project”. In: *M-Health*. Springer, 2006, pp. 461–473 (cit. on p. 8).

[51] P. Vieyres, G. Poisson, F. Courrèges, O. Mérigeaux, and P. Arbeille. “The TERESA project: from space research to ground tele-echography”. In: *Industrial robot: an international journal* (2003) (cit. on p. 8).

[52] X. Guan, H. Wu, X. Hou, et al. “Study of a 6DOF robot assisted ultrasound scanning system and its simulated control handle”. In: *Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), 2017 IEEE International Conference on*. IEEE. 2017, pp. 469–474 (cit. on p. 8).

[53] K. Mathiassen, J. E. Fjellin, K. Glette, P. K. Hol, and O. J. Elle. “An ultrasound robotic system using the commercial robot UR5”. In: *Frontiers in Robotics and AI* 3 (2016), p. 1 (cit. on p. 8).

[54] W. Bachta and A. Krupa. “Towards ultrasound image-based visual servoing”. In: *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* IEEE. 2006, pp. 4112–4117 (cit. on p. 9).

[55] A. Krupa, G. Fichtinger, and G. D. Hager. “Real-time tissue tracking with B-mode ultrasound using speckle and visual servoing”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2007, pp. 1–8 (cit. on p. 9).

[56] A. Krupa, G. Fichtinger, and G. D. Hager. “Real-time motion stabilization with B-mode ultrasound using image speckle information and visual servoing”. In: *The International Journal of Robotics Research* 28.10 (2009), pp. 1334–1354 (cit. on p. 9).

[57] R. Mebarki, A. Krupa, and F. Chaumette. “2-d ultrasound probe complete guidance by visual servoing using image moments”. In: *IEEE Transactions on Robotics* 26.2 (2010), pp. 296–306 (cit. on p. 9).

[58] C. Nadeau and A. Krupa. “Intensity-based direct visual servoing of an ultrasound probe”. In: *2011 IEEE International Conference on Robotics and Automation*. IEEE. 2011, pp. 5677–5682 (cit. on p. 9).

[59] R. Nakadate, J. Solis, A. Takanishi, E. Minagawa, M. Sugawara, and K. Niki. “Out-of-plane visual servoing method for tracking the carotid artery with a robot-assisted ultrasound diagnostic system”. In: *2011 IEEE International Conference on Robotics and Automation*. IEEE. 2011, pp. 5267–5272 (cit. on p. 9).

[60] P. Chatelain, A. Krupa, and M. Marchal. “Real-time needle detection and tracking using a visually servoed 3D ultrasound probe”. In: *2013 IEEE International Conference on Robotics and Automation*. IEEE. 2013, pp. 1676–1681 (cit. on p. 9).

[61] P. Chatelain, A. Krupa, and N. Navab. “3D ultrasound-guided robotic steering of a flexible needle via visual servoing”. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2015, pp. 2250–2255 (cit. on p. 9).

[62] R. Kojcev, B. Fuerst, O. Zettinig, et al. “Dual-robot ultrasound-guided needle placement: closing the planning-imaging-action loop”. In: *International journal of computer assisted radiology and surgery* 11.6 (2016), pp. 1173–1181 (cit. on p. 9).

[63] X. Guo, B. Tavakoli, H.-J. Kang, J. U. Kang, R. Etienne-Cummings, and E. M. Boctor. “Photoacoustic active ultrasound element for catheter tracking”. In: *Photons Plus Ultrasound: Imaging and Sensing 2014*. Vol. 8943. International Society for Optics and Photonics. 2014, p. 89435M (cit. on p. 9).

[64] P. Chatelain, A. Krupa, and N. Navab. “Optimization of ultrasound image quality via visual servoing”. In: *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE. 2015, pp. 5997–6002 (cit. on p. 9).

[65] P. Chatelain, A. Krupa, and N. Navab. “Confidence-driven control of an ultrasound probe: Target-specific acoustic window optimization”. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2016, pp. 3441–3446 (cit. on p. 9).

[66] P. Chatelain, A. Krupa, and N. Navab. “Confidence-driven control of an ultrasound probe”. In: *IEEE Transactions on Robotics* 33.6 (2017), pp. 1410–1424 (cit. on p. 9).

[67] T.-Y. Fang, H. K. Zhang, R. Finocchi, R. H. Taylor, and E. M. Boctor. “Force-assisted ultrasound imaging system through dual force sensing and admittance robot control”. In: *International journal of computer assisted radiology and surgery* 12.6 (2017), pp. 983–991 (cit. on p. 9).

[68] R. Finocchi, F. Aalamifar, T. Y. Fang, R. H. Taylor, and E. M. Boctor. “Co-robotic ultrasound imaging: a cooperative force control approach”. In: *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*. Vol. 10135. International Society for Optics and Photonics. 2017, p. 1013510 (cit. on p. 9).

[69] S. Wang, J. Housden, Y. Noh, et al. “Design and implementation of a bespoke robotic manipulator for extra-corporeal ultrasound”. In: *JoVE (Journal of Visualized Experiments)* 143 (2019), e58811 (cit. on p. 9).

[70] S. Wang, J. Housden, Y. Noh, et al. “Robotic-assisted Ultrasound for Fetal Imaging: Evolution from Single-arm to Dual-arm System”. In: *Annual Conference Towards Autonomous Robotic Systems*. Springer. 2019, pp. 27–38 (cit. on p. 9).

[71] L. Lindenroth, A. Soor, J. Hutchinson, et al. “Design of a soft, parallel end-effector applied to robot-guided ultrasound interventions”. In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2017, pp. 3716–3721 (cit. on p. 9).

[72] L. Lindenroth, R. J. Housden, S. Wang, J. Back, K. Rhode, and H. Liu. “Design and integration of a parallel, soft robotic end-effector for extracorporeal ultrasound”. In: *IEEE Transactions on Biomedical Engineering* (2019) (cit. on p. 9).

[73] H. T. Sen, M. A. L. Bell, I. Iordachita, J. Wong, and P. Kazanzides. “A cooperatively controlled robot for ultrasound monitoring of radiation therapy”. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2013, pp. 3071–3076 (cit. on p. 9).

[74] H. T. Şen, M. A. L. Bell, Y. Zhang, et al. “System integration and preliminary in-vivo experiments of a robot for ultrasound guidance and monitoring during radiotherapy”. In: *2015 International Conference on Advanced Robotics (ICAR)*. IEEE. 2015, pp. 53–59 (cit. on p. 9).

[75] C. Kim, F. Schäfer, D. Chang, D. Petrisor, M. Han, and D. Stoianovici. “Robot for ultrasound-guided prostate imaging and intervention”. In: *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2011, pp. 943–948 (cit. on p. 10).

[76] O. Zettinig, B. Fuerst, R. Kojcev, et al. “Toward real-time 3D ultrasound registration-based visual servoing for interventional navigation”. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2016, pp. 945–950 (cit. on p. 10).

[77] J. Jakubiak, M. Drwięga, and B. Stańczyk. “Control and perception system for ReMeDi robot mobile platform”. In: *2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR)*. IEEE. 2015, pp. 750–755 (cit. on p. 10).

[78] G. Stollnberger, C. Moser, M. Giuliani, et al. “User requirements for a medical robotic system: enabling doctors to remotely conduct ultrasonography and physical examination”. In: *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2016, pp. 1156–1161 (cit. on p. 10).

[79] MGIUS-R3. `https://en.mgitech.cn/products/instruments_info/11/`. [Online; accessed 20.02.2020] (cit. on p. 10).

[80] R Spoor, M. Abayazid, F. J. Siepel, V. Groenhuis, and S. Stramigioli. “Design and evaluation of a robotic needle steering manipulator for image-guided breast biopsy”. In: *6th Dutch Bio-Medical Engineering Conference 2017*. 2017 (cit. on p. 10).

[81] A. M. Priester, S. Natarajan, and M. O. Culjat. “Robotic ultrasound systems in medicine”. In: *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 60.3 (2013), pp. 507–523 (cit. on p. 10).

[82] A. A. Moshaii and F. Najafi. “A review of robotic mechanisms for ultrasound examinations”. In: *Industrial Robot: An International Journal* (2014) (cit. on p. 10).

[83] R. Elek, T. D. Nagy, D. A. Nagy, et al. “Robotic platforms for ultrasound diagnostics and treatment”. In: *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE. 2017, pp. 1752–1757 (cit. on p. 10).

[84] D. R. Swerdlow, K. Cleary, E. Wilson, B. Azizi-Koutenaei, and R. Monfaredi. “Robotic arm–assisted sonography: Review of technical developments and potential clinical applications”. In: *American Journal of Roentgenology* 208.4 (2017), pp. 733–738 (cit. on p. 10).

[85] L. Nouaille, M. A. Laribi, C. A. Nelson, S Zeghloul, and G. Poisson. “Review of kinematics for minimally invasive surgery and tele-echography robots”. In: *Journal of Medical Devices* 11.4 (2017) (cit. on p. 10).

[86] D Kane, W Grassi, R Sturrock, and P. Balint. “A brief history of musculoskeletal ultrasound:‘From bats and ships to babies and hips’”. In: *Rheumatology* 43.7 (2004), pp. 931–933 (cit. on p. 13).

[87] T. L. Szabo. *Diagnostic ultrasound imaging: inside out*. Academic Press, 2004 (cit. on pp. 13, 17, 18).

[88] J.-L. Gennisson, T. Deffieux, M. Fink, and M. Tanter. “Ultrasound elastography: principles and techniques”. In: *Diagnostic and interventional imaging* 94.5 (2013), pp. 487–495 (cit. on p. 13).

[89] F. A. Duck. *Physical properties of tissues: a comprehensive reference book*. Academic press, 2013 (cit. on p. 16).

[90] B. J. Ostrum, B. B. Goldberg, and H. J. Isard. “A-mode ultrasound differentiation of soft-tissue masses”. In: *Radiology* 88.4 (1967), pp. 745–749 (cit. on p. 17).

[91] B. B. Goldberg and H. M. Pollack. “Differentiation of renal masses using A-mode ultrasound”. In: *The Journal of urology* 105.6 (1971), pp. 765–771 (cit. on p. 17).

[92] R. Aaslid, T.-M. Markwalder, and H. Nornes. “Noninvasive transcranial Doppler ultrasound recording of flow velocity in basal cerebral arteries”. In: *Journal of neurosurgery* 57.6 (1982), pp. 769–774 (cit. on p. 17).

[93] G. R. DeVore, R. L. Donnerstein, C. S. Kleinman, L. D. Platt, and J. C. Hobbins. “Fetal echocardiography: I. Normal anatomy as determined by real-time—directed M-mode ultrasound”. In: *American Journal of Obstetrics & Gynecology* 144.3 (1982), pp. 249–260 (cit. on p. 17).

[94] H. H. Holm and B. Skjoldbye. “Interventional ultrasound”. In: *Ultrasound in medicine & biology* 22.7 (1996), pp. 773–789 (cit. on p. 17).

[95] E. A. Geiser, L. G. Christie Jr, D. A. Conetta, C. R. Conti, and G. S. Gossman. “A mechanical arm for spatial registration of two-dimensional echocardiographic sections”. In: *Catheterization and cardiovascular diagnosis* 8.1 (1982), pp. 89–101 (cit. on p. 20).

[96] R. N. Rankin, A Fenster, D. Downey, P. Munk, M. Levin, and A. Vellet. “Three-dimensional sonographic reconstruction: techniques and diagnostic applications.” In: *AJR. American journal of roentgenology* 161.4 (1993), pp. 695–702 (cit. on p. 20).

[97] Z. Luo, J. Cai, S. Wang, Q. Zhao, T. M. Peters, and L. Gu. “Magnetic navigation for thoracic aortic stent-graft deployment using ultrasound image guidance”. In: *IEEE Transactions on Biomedical Engineering* 60.3 (2012), pp. 862–871 (cit. on p. 20).

[98] A. Gee, R. Prager, G. Treece, and L. Berman. “Engineering a freehand 3D ultrasound system”. In: *Pattern Recognition Letters* 24.4-5 (2003), pp. 757–777 (cit. on p. 20).

[99] O. V. Solberg, F. Lindseth, H. Torp, R. E. Blake, and T. A. N. Hernes. “Freehand 3D ultrasound reconstruction algorithms—a review”. In: *Ultrasound in medicine & biology* 33.7 (2007), pp. 991–1009 (cit. on pp. 20, 22).

[100] R. J. Housden, A. H. Gee, G. M. Treece, and R. W. Prager. “Sensorless reconstruction of freehand 3d ultrasound data”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2006, pp. 356–363 (cit. on p. 21).

[101] H. Gao, Q. Huang, X. Xu, and X. Li. “Wireless and sensorless 3D ultrasound imaging”. In: *Neurocomputing* 195 (2016), pp. 159–171 (cit. on p. 21).

[102] R. Prevost, M. Salehi, S. Jagoda, et al. “3D freehand ultrasound without external tracking using deep learning”. In: *Medical image analysis* 48 (2018), pp. 187–202 (cit. on p. 21).

[103] T. Wen, Q. Zhu, W. Qin, et al. “An accurate and effective FMM-based approach for freehand 3D ultrasound reconstruction”. In: *Biomedical Signal Processing and Control* 8.6 (2013), pp. 645–656 (cit. on p. 22).

[104] D. Dewi, T. L. Mengko, I. K. E. Purnama, A. G. Veldhuizen, and M. H. Wilkinson. “An improved olympic hole-filling method for ultrasound volume reconstruction of human spine”. In: *International Journal of E-Health and Medical Communications (IJEHMC)* 1.3 (2010), pp. 28–40 (cit. on p. 22).

[105] R. Rohling, A. Gee, and L. Berman. “A comparison of freehand three-dimensional ultrasound reconstruction techniques”. In: *Medical image analysis* 3.4 (1999), pp. 339–359 (cit. on p. 22).

[106] A. Fenster, D. B. Downey, and H. N. Cardinal. “Three-dimensional ultrasound imaging”. In: *Physics in medicine & biology* 46.5 (2001), R67 (cit. on p. 22).

[107] F. Mohamed and C. V. Siang. “A Survey on 3D Ultrasound Reconstruction Techniques”. In: *Artificial Intelligence-Applications in Medicine and Biology*. IntechOpen, 2019 (cit. on p. 22).

[108] J. A. Maintz and M. A. Viergever. “A survey of medical image registration”. In: *Medical image analysis* 2.1 (1998), pp. 1–36 (cit. on p. 22).

[109] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. “Medical image registration”. In: *Physics in medicine & biology* 46.3 (2001), R1 (cit. on p. 22).

[110] F. P. Oliveira and J. M. R. Tavares. “Medical image registration: a review”. In: *Computer methods in biomechanics and biomedical engineering* 17.2 (2014), pp. 73–93 (cit. on p. 22).

[111] M. J. Powell. “The BOBYQA algorithm for bound constrained optimization without derivatives”. In: *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge* (2009), pp. 26–46 (cit. on p. 22).

[112] L. M. Rios and N. V. Sahinidis. “Derivative-free optimization: a review of algorithms and comparison of software implementations”. In: *Journal of Global Optimization* 56.3 (2013), pp. 1247–1293 (cit. on p. 22).

[113] A. Sotiras, C. Davatzikos, and N. Paragios. “Deformable medical image registration: A survey”. In: *IEEE transactions on medical imaging* 32.7 (2013), pp. 1153–1190 (cit. on p. 23).

[114] F. L. Bookstein. “Principal warps: Thin-plate splines and the decomposition of deformations”. In: *IEEE Transactions on pattern analysis and machine intelligence* 11.6 (1989), pp. 567–585 (cit. on p. 23).

[115] U. Clarenz, M. Droske, S. Henn, M. Rumpf, and K. Witsch. “Computational methods for nonlinear image registration”. In: *Mathematical Models for Registration and Applications to Medical Imaging*. Springer, 2006, pp. 81–101 (cit. on p. 23).

[116] P. Viola and W. M. Wells III. “Alignment by maximization of mutual information”. In: *International journal of computer vision* 24.2 (1997), pp. 137–154 (cit. on p. 24).

[117] W. Wein, A. Ladikos, B. Fuerst, A. Shah, K. Sharma, and N. Navab. “Global registration of ultrasound to MRI using the LC 2 metric for enabling neurosurgical guidance”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2013, pp. 34–41 (cit. on p. 24).

[118] B. Fuerst, W. Wein, M. Müller, and N. Navab. “Automatic ultrasound–MRI registration for neurosurgery using the 2D and 3D LC2 Metric”. In: *Medical image analysis* 18.8 (2014), pp. 1312–1319 (cit. on p. 25).

[119] N. Ciblak and H. Lipkin. “Design and analysis of remote center of compliance structures”. In: *Journal of robotic systems* 20.8 (2003), pp. 415–427 (cit. on p. 27).

[120] C. Loughlin, A Albu-Schäffer, S Haddadin, et al. “The DLR lightweight robot: design and control concepts for robots in human environments”. In: *Industrial Robot: an international journal* (2007) (cit. on p. 28).

[121] J. Lee. “Apply force/torque sensors to robotic applications”. In: *Robotics* 3.2 (1987), pp. 189–194 (cit. on p. 28).

[122] H Qiao, B. Dalay, and R. Parkin. “Robotic peg-hole insertion operations using a six-component force sensor”. In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 207.5 (1993), pp. 289–306 (cit. on p. 28).

[123] J Luh, W. Fisher, and R Paul. “Joint torque control by a direct feedback for industrial robots”. In: *IEEE Transactions on Automatic Control* 28.2 (1983), pp. 153–161 (cit. on p. 28).

[124] G. Hirzinger, N. Sporer, A. Albu-Schaffer, et al. “DLR’s torque-controlled light weight robot III-are we reaching the technological limits now?” In: *Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No. 02CH37292)*. Vol. 2. IEEE. 2002, pp. 1710–1716 (cit. on p. 28).

[125] L. Sciavicco and B. Siciliano. *Modelling and control of robot manipulators*. Springer Science & Business Media, 2012 (cit. on pp. 29, 32).

[126] L. Villani and J. De Schutter. “Force control”. In: *Springer handbook of robotics*. Springer, 2016, pp. 195–220 (cit. on pp. 29, 30).

[127] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo. *Robotics: modelling, planning and control*. Springer Science & Business Media, 2010 (cit. on pp. 30, 32).

[128] V. Ortenzi, R. Stolkin, J Kuo, and M. Mistry. “Hybrid motion/force control: a review”. In: *Advanced Robotics* 31.19-20 (2017), pp. 1102–1113 (cit. on p. 32).

[129] K. Celik, S.-J. Chung, and A. Somani. “Mono-vision corner SLAM for indoor navigation”. In: *2008 IEEE International Conference on Electro/Information Technology*. IEEE. 2008, pp. 343–348 (cit. on p. 37).

[130] Y. Zheng, G. Shen, L. Li, C. Zhao, M. Li, and F. Zhao. “Travi-navi: Self-deployable indoor navigation system”. In: *IEEE/ACM transactions on networking* 25.5 (2017), pp. 2655–2669 (cit. on p. 37).

[131] H. Baltzakis, A. Argyros, and P. Trahanias. “Fusion of laser and visual data for robot motion planning and collision avoidance”. In: *Machine Vision and Applications* 15.2 (2003), pp. 92–100 (cit. on p. 37).

[132] Y. Watanabe, A. Calise, and E. Johnson. “Vision-based obstacle avoidance for UAVs”. In: *AIAA guidance, navigation and control conference and exhibit*. 2007, p. 6829 (cit. on p. 37).

[133] A. Saxena, J. Driemeyer, and A. Y. Ng. “Robotic grasping of novel objects using vision”. In: *The International Journal of Robotics Research* 27.2 (2008), pp. 157–173 (cit. on p. 37).

[134] Z. Zhang. “Microsoft kinect sensor and its effect”. In: *IEEE multimedia* 19.2 (2012), pp. 4–10 (cit. on pp. 37, 39).

[135] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge university press, 2003 (cit. on p. 37).

[136] L. Nalpantidis, G. C. Sirakoulis, and A. Gasteratos. “Review of stereo matching algorithms for 3D vision”. In: *16th International Symposium on Measurement and Control in Robotics 21-23 June 2007-Warsaw, POLAND*. 2007 (cit. on p. 38).

[137] B. Tippetts, D. J. Lee, K. Lillywhite, and J. Archibald. “Review of stereo vision algorithms and their suitability for resource-limited systems”. In: *Journal of Real-Time Image Processing* 11.1 (2016), pp. 5–25 (cit. on p. 38).

[138] S. Knorr, M. Kunter, and T. Sikora. “Stereoscopic 3D from 2D video with super-resolution capability”. In: *Signal processing: Image communication* 23.9 (2008), pp. 665–676 (cit. on p. 39).

[139] B. Ummenhofer, H. Zhou, J. Uhrig, et al. “Demon: Depth and motion network for learning monocular stereo”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 5038–5047 (cit. on p. 39).

[140] R. Mahjourian, M. Wicke, and A. Angelova. “Unsupervised Learning of Depth and Ego-Motion from Monocular Video Using 3D Geometric Constraints”. In: *CVPR*. 2018 (cit. on p. 39).

[141] T. Darrell and K. Wohn. “Pyramid based depth from focus”. In: *Proceedings CVPR’88: The Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE. 1988, pp. 504–509 (cit. on p. 39).

[142] Y. Xiong and S. A. Shafer. “Depth from focusing and defocusing”. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 1993, pp. 68–73 (cit. on p. 39).

[143] A. Wiegmann, H. Wagner, and R. Kowarschik. “Human face measurement by projecting bandlimited random patterns”. In: *Optics express* 14.17 (2006), pp. 7692–7698 (cit. on p. 39).

[144] X.-F. Hana, J. S. Jin, J. Xie, M.-J. Wang, and W. Jiang. “A comprehensive review of 3d point cloud descriptors”. In: *arXiv preprint arXiv:1802.02297* (2018) (cit. on p. 40).

[145] Y. Guo, M. Bennamoun, F. Sohel, M. Lu, J. Wan, and N. M. Kwok. “A comprehensive performance evaluation of 3D local feature descriptors”. In: *International Journal of Computer Vision* 116.1 (2016), pp. 66–89 (cit. on p. 40).

[146] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš. “A review of 3D/2D registration methods for image-guided interventions”. In: *Medical image analysis* 16.3 (2012), pp. 642–661 (cit. on p. 40).

[147] B. Maiseli, Y. Gu, and H. Gao. “Recent developments and trends in point set registration methods”. In: *Journal of Visual Communication and Image Representation* 46 (2017), pp. 95–106 (cit. on p. 40).

[148] S.-Y. Guan, T.-M. Wang, C. Meng, and J.-C. Wang. “A review of point feature based medical image registration”. In: *Chinese Journal of Mechanical Engineering* 31.1 (2018), p. 76 (cit. on p. 40).

[149] P. J. Besl and N. D. McKay. “Method for registration of 3-D shapes”. In: *Sensor fusion IV: control paradigms and data structures*. Vol. 1611. International Society for Optics and Photonics. 1992, pp. 586–606 (cit. on p. 41).

[150] S. Granger and X. Pennec. “Multi-scale EM-ICP: A fast and robust approach for surface registration”. In: *European Conference on Computer Vision*. Springer. 2002, pp. 418–432 (cit. on p. 41).

[151] T. K. Moon. “The expectation-maximization algorithm”. In: *IEEE Signal processing magazine* 13.6 (1996), pp. 47–60 (cit. on pp. 41, 42).

[152] A. W. Fitzgibbon. “Robust registration of 2D and 3D point sets”. In: *Image and vision computing* 21.13-14 (2003), pp. 1145–1153 (cit. on p. 41).

[153] J. Yang, H. Li, and Y. Jia. “Go-icp: Solving 3d registration efficiently and globally optimally”. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2013, pp. 1457–1464 (cit. on p. 41).

[154] B. Amberg, S. Romdhani, and T. Vetter. “Optimal step nonrigid ICP algorithms for surface registration”. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE. 2007, pp. 1–8 (cit. on p. 41).

[155] S. Cheng, I. Marras, S. Zafeiriou, and M. Pantic. “Statistical non-rigid ICP algorithm and its application to 3D face alignment”. In: *Image and Vision Computing* 58 (2017), pp. 3–12 (cit. on p. 41).

[156] B. Jian and B. C. Vemuri. “A robust algorithm for point set registration using mixture of Gaussians”. In: *Tenth IEEE International Conference on Computer Vision (ICCV’05) Volume 1*. Vol. 2. IEEE. 2005, pp. 1246–1251 (cit. on p. 41).

[157] D. A. Reynolds. “Gaussian Mixture Models.” In: *Encyclopedia of biometrics* 741 (2009) (cit. on p. 41).

[158] Y. Tsin and T. Kanade. “A correlation-based approach to robust point set registration”. In: *European conference on computer vision*. Springer. 2004, pp. 558–569 (cit. on p. 41).

[159] A. Myronenko and X. Song. “Point set registration: Coherent point drift”. In: *IEEE transactions on pattern analysis and machine intelligence* 32.12 (2010), pp. 2262–2275 (cit. on p. 41).

[160] A. Karamalis, W. Wein, T. Klein, and N. Navab. “Ultrasound confidence maps using random walks”. In: *Medical image analysis* 16.6 (2012), pp. 1101–1112 (cit. on p. 68).

[161] D. Shen, G. Wu, D. Zhang, K. Suzuki, F. Wang, and P. Yan. “Machine learning in medical imaging.” In: *Comp. Med. Imag. and Graph.* 41 (2015), pp. 1–2 (cit. on p. 91).

[162] Q. Huang, F. Zhang, and X. Li. “Machine learning in ultrasound computer-aided diagnostic systems: a survey”. In: *BioMed research international* 2018 (2018) (cit. on p. 91).

[163] S. Liu, Y. Wang, X. Yang, et al. “Deep learning in medical ultrasound analysis: a review”. In: *Engineering* (2019) (cit. on p. 91).

[164] J. Kober, J. A. Bagnell, and J. Peters. “Reinforcement learning in robotics: A survey”. In: *The International Journal of Robotics Research* 32.11 (2013), pp. 1238–1274 (cit. on p. 92).

[165] F. Milletari, V. Birodkar, and M. Sofka. “Straight to the point: reinforcement learning for user guidance in ultrasound”. In: *Smart Ultrasound Imaging and Perinatal, Preterm and Paediatric Image Analysis*. Springer, 2019, pp. 3–10 (cit. on p. 92).

[166] M. Esposito, B. Busam, C. Hennersperger, et al. “Cooperative robotic gamma imaging: Enhancing us-guided needle biopsy”. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2015, pp. 611–618 (cit. on p. 92).

[167] C. Bartneck and J. Forlizzi. “A design-centred framework for social human-robot interaction”. In: *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759)*. IEEE. 2004, pp. 591–594 (cit. on p. 92).

[168] S. Lemaignan, M. Warnier, E. A. Sisbot, A. Clodic, and R. Alami. “Artificial cognition for social human–robot interaction: An implementation”. In: *Artificial Intelligence* 247 (2017), pp. 45–69 (cit. on p. 92).

[169] B. Alenljung, J. Lindblom, R. Andreasson, and T. Ziemke. “User experience in social human-robot interaction”. In: *Rapid Automation: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2019, pp. 1468–1490 (cit. on p. 92).

!

List of Figures

List of Tables