C-arm-based surgical data visualization and repositioning using augmented reality

by

Tianyu Song

A thesis submitted to The Johns Hopkins University

in conformity with the requirements for the degree of

Master of Science

Baltimore, Maryland

May, 2019

Abstract

Purpose: As the trend towards minimally invasive and percutaneous interventions continues, the importance of appropriate surgical data visualization becomes more evident. Ineffective interventional data display techniques yield poor ergonomics that hinder hand-eye coordination, and therefore, promote frustration which can compromise on-task performance up to adverse outcome. A very common example of ineffective visualization is monitors attached to the base of mobile C-arm X-ray systems.

Methods: We present a spatially- and imaging geometry-aware paradigm for visualization of fluoroscopic images using Interactive Flying Frustums (IFFs) in a mixed reality environment. We exploit the fact that the C-arm imaging geometry can be modeled as a pinhole camera giving rise to an 11 degree of freedom view frustum on which the X-ray image can be translated while remaining valid. Visualizing IFFs to the surgeon in an augmented reality environment intuitively unites the virtual 2D X-ray image plane and the real 3D patient anatomy. To achieve this visualization, the surgeon and C-arm are tracked relative to the same coordinate frame using image-based localization and mapping, with the augmented reality environment being delivered to the surgeon via a state-of-the-art optical see-through head-mounted display.

Results: We present a thorough evaluation of the hand-eye calibration procedure. Results suggest convergence when using 50 pose pairs or more. The mean translation and rotation errors at convergence are 5.7 mm and 0.26°, respectively. The root-mean-squared error of C-arm source tracking after hand-eye calibration was determined as $0.43° \pm 0.34°$ and 4.6 ± 2.7 mm in rotation and translation, respectively. Finally, we demonstrated the application of spatially-aware data visualization for internal fixation of pelvic fractures and percutaneous vertebroplasty.

Conclusion: Our spatially-aware approach to transmission image visualization effectively unites patient anatomy with X-ray images by enabling spatial image manipulation that abides image formation. Our proof-of-principle findings indicate potential applications for surgical tasks that mostly rely on orientational information such as placing the acetabular component in total hip arthroplasty, making us confident that the proposed augmented reality concept can pave the way for improving surgical performance and visuo-motor coordination in fluoroscopy-guided surgery.

Acknowledgments

I would first like to thank my advisor Dr. Nassir Navab for his constant guidance and endless support. His expertise was invaluable in the formulating of the research topic and methodology in particular.

I would also like to thank my lab mates and mentors at the Computer Aided Medical Procedures (CAMP) lab, Javad Fotouhi, Dr. Mathias Unberath for their contribution and constant feedback during my time working at CAMP.

In addition, I want to thank Gerhard Kleinzig and Sebastian Vogt from Siemens Healthineers for their support and making a Siemens ARCADIS Orbic 3-D available.

Finally, I must express my very profound gratitude to my parents and to my girlfriend for providing me with unfailing support and continuous encouragement through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

C-arm fluoroscopy is extensively used to guide minimally invasive surgery in a variety of clinical disciplines including neuro-radiology, orthopedics, and trauma (Hott et al., 2004; Miller et al., 2010; Theocharopoulos et al., 2003; Mason et al., 2014). Fluoroscopy provides real-time X-ray images that enable visualizing and monitoring the progress of surgery on the anatomy level. In fracture care surgery, C-arm imaging is employed to guide the safe placement of implants, wires, and screws. A prominent representative of fracture care surgery is closed reduction and internal fixation of anterior pelvic fractures, *i. e.* fractures of the superior pubic ramus. This procedure exhibits particularly small error margins due to the close proximity to critical structures (Suzuki et al., 2008). To achieve the required surgical accuracy and confidence, C-arm images are acquired from different views to verify acceptable tool trajectories. Yet, geometric interpretation of these interventional images is challenging and requires highly skilled and experienced surgeons that are trained to infer complex 3D spatial relations from 2D X-ray images alone (Starr et al., 2001). This need for ”mental mapping” leads to the acquisition of an excessive amount of

fluoroscopic images, and results in frustration of the surgeon up to compromised surgical efficiency, procedural delays, and radiation hazards (Boszczyk et al., 2006; Synowitz and Kiwit, 2006).

The complexity of interpreting 2D fluoroscopic images to establish spatial connections to the patient anatomy can, at least partly, be attributed to two major shortcomings: 1) Poor surgical ergonomics due to inconvenient off-axis display of image data via external displays, and 2) lack of geometric registration between the image content and the imaged anatomy. There is a wealth of literature on computer-integrated surgical solutions that address *one of the two* aforementioned challenges. In the following we briefly review the relevant state-of-the-art.

1.1 Related Work

First attempts at benefiting the ergonomics of surgery by improving display position placed miniature LCD displays close to the intervention site (Cardin, Wang, and Plewes, 2005; Westwood, 2005), and later, displayed images relative to the surgeon's field of vision using Google Glass (Chimenti and Mitten, 2015; Yoon et al., 2016). More recently, Qian *et al.* (Qian et al., 2017) and Deib *et al.* Deib et al., 2018 described an augmented reality (AR)-based virtual monitor concept delivered via novel optical see-through head-mounted display (OST HMD) devices that use simultaneous localization and mapping (SLAM) to estimate their position within the environment. This knowledge enables rendering of medical images in multiple display configurations, namely: head-, body-, and world-anchored mode. In head-anchored mode, images are

rendered at a fixed pose in relation to the surgeon's field of vision as previously described using Google Glass (Vorraber et al., 2014; Chimenti and Mitten, 2015; Yoon et al., 2016) that can potentially occlude the surgical site. In world-anchored mode, the virtual monitor is static in relation to the OR environment (Chen et al., 2015). Finally, body-anchored mode combines head- and world-anchored concepts such that the image always remains within the field of view, but its global pose in the surgery room is not static. Using this virtual monitor system (Qian et al., 2017; Deib et al., 2018), the surgeon is capable of flexibly controlling the display position, thereby reducing the challenges introduced by off-axis visualization. Another advantage of the virtual monitor system, which distinguishes it from previous hardware-based solutions (Westwood, 2005; Cardin, Wang, and Plewes, 2005; Hanna, Shimi, and Cuschieri, 1998), is the possibility of displaying data with high resolution directly at the surgical site without compromising sterility. Despite the benefits of "in-line" display of images, the disconnect in visuo-motor coordination is not fully reversed since the image content is not spatially registered to the patient nor calibrated to the scanner.

Spatially registering pre- or intra-operative 3D data to the patient interventionally has vastly been considered as it constitutes the basis for navigational approaches (Liebergall, Mosheiff, and Joskowicz, 2006; Joskowicz and Hazan, 2016; Theologis, Burch, and Pekmezci, 2016). In navigated surgery, optical markers are attached to tools, registered to anatomy, and finally tracked in an outside-in setting using active infrared cameras. While highly accurate, these systems are only appropriate for entry point localization since the 3D volume

is not updated. Additionally, navigated surgery suffers from complicated intra-operative calibration routines that increase procedure time, and if not sufficiently robust, foster frustration. Despite improving accuracy (Alambeigi et al., 2017), the aforementioned drawbacks limit clinical acceptance (Joskowicz and Hazan, 2016; Fischer et al., 2016; Hajek et al., 2018). In contrast to explicit navigation and robotic solutions (Sefati et al., 2016; Sefati et al., 2017), scanner-(Fischer et al., 2016; Tucker et al., 2018; Hajek et al., 2018) or user-centric (Unberath et al., 2018; Qian, Deguet, and Kazanzides, 2018; Augmedics, 2018) sensing and visualization have been found effective in relaxing the requirements for markers and tracking by intuitively visualizing 3D spatial relations either on multiple projective images rendered from the 3D scene (Tucker et al., 2018; Fotouhi et al., 2018) or as 3D virtual content in AR environments (Unberath et al., 2018; Augmedics, 2018; Hajek et al., 2018). These approaches work well but require 3D imaging for every patient which is not traditionally available. Image overlays for surgical navigation have also been proposed for fusing multi-modal interventional images such as the EchoNavigator system (Philips Inc., Amsterdam, Netherlands) where the outline of the 3D ultrasound volumes are augmented onto the fluoroscopy images to provide an intuitive geometric mapping between multiple images (Gafoor et al., 2015).

As a consequence, methods that provide 3D information but only rely on C-arm fluoroscopy are preferred if wide acceptance of the method is desired. Using the same concept as (Fischer et al., 2016; Tucker et al., 2018), *i. e.* an RGB-D camera rigidly attached to the detector of a mobile C-arm, Fotouhi *et*

al. track the position of the C-arm relative to the patient using image-based SLAM (Fotouhi et al., 2017). Doing so enables tracking of successive C-arm poses which implicitly facilitates "mental mapping" as relative image poses can be visualized, however, this visualization is limited to conventional 2D monitors since no *in situ* AR environment is in place. A promising way of achieving calibration between the X-ray images and an AR environment presented to the surgeon is the use of multi-modal fiducial markers that can be sensed simultaneously by all involved devices, *i. e.* the C-arm X-ray system and the OST HMD (Andress et al., 2018). Since the marker geometry is known in 3D, poses of the C-arm source and the HMD can be inferred relative to the marker, and thus, via SLAM also to the AR environment enabling calibration in unprepared operating theatres. Then, visuo-motor coordination and "mental mapping" is improved explicitly by propagating X-ray image domain annotations (*e. g.* a key point) to corresponding lines in 3D that connect C-arm source and detector location, thereby intersecting the patient. This straightforward concept has proved beneficial for distal locking of intramedullary nails (Andress et al., 2018) where availability of 3D down-the-beam information is of obvious interest. Yet and similarly to navigated surgery, the introduction of fiducial markers that must be seen simultaneously by C-arm and HMD is associated with changes to surgical workflow. Consequently, it is unclear whether the provided benefits outweigh the associated challenges in clinical application.

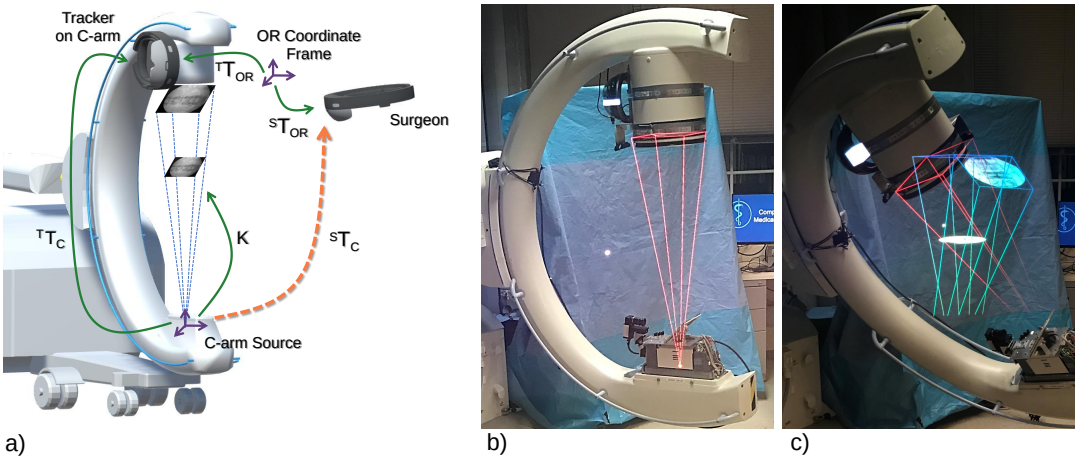

Figure 1.1: a) Schematic illustration of the proposed spatially-aware image visualization of X-ray images on their view frustum. In addition, we show transformation to be estimated dynamically to enable the proposed AR environment. Transformations shown as green arrows are estimated directly while transformations shown in orange are derived. b) Demonstrates the use of a single IFF from the current view, and c) demonstrates the simultaneous visualization of multiple IFFs from the current and previous views.

1.2 Spatially-aware Image Visualization via IFFs

What *if* the surgeon could instantaneously observe all the acquired X-ray images floating at the position of detector at the moment of their acquisition? What *if* the surgeon could interactively move the X-ray image within its geometrical frustum passing through the actual anatomy? What *if* the surgeon could point at the X-ray image that was taken at a given point in the surgery and ask crew to bring the scanner to that X-ray position? What *if* the crew could also observe all the same floating imagery data and the corresponding position of the scanner? What *if* expert and training surgeons could review all acquisitions with the corresponding spatial and temporal acquisition information? Interactive Flying Frustums (IFFs) aims at providing a new augmented reality methodology allowing the realization of all these

'*if*'s. In IFFs paradigm, we leverage the concept of the view frustum (Georgel, Schroeder, and Navab, 2009) combined with improved dynamic inside-out calibration of the C-arm to the AR environment (Hajek et al., 2018) to develop spatially-aware visualization. The proposed system, illustrated in Fig. 1.1, 1) displays medical images at the surgical site overcoming the challenges introduced by off-axis display, and 2) effectively and implicitly calibrates the acquired fluoroscopic images to the patient by allowing the image to slide along the viewing frustum.

Chapter 2

Methodology

2.1 Uniting Patient Anatomy and X-ray Image Using the View Frustum

The C-arm X-ray image formation is geometrically described by the pinhole camera model (Hartley and Zisserman, 2003) with the X-ray source constituting the focal point. While the imaging geometry is largely similar to conventional optical imaging, there are two major differences: First, in contrast to optical imaging where we are interested in reflected light quanta, in X-ray imaging we measure transmitted intensity. Second and as a consequence, the object must be placed between the focal spot (the X-ray source) and the detector plane. Given the 11 degree of freedom (DoF) camera parameters, the view frustum then describes the cone of vision (or pyramid of vision) centered at the focal point with the active area of the X-ray detector plane defining its base. Assuming that the detector plane is normal to the principal ray of the C-arm and using the notational conventions of Hartley and Zisserman (Hartley and Zisserman, 2003), then, any image acquired in this fixed C-arm pose can be

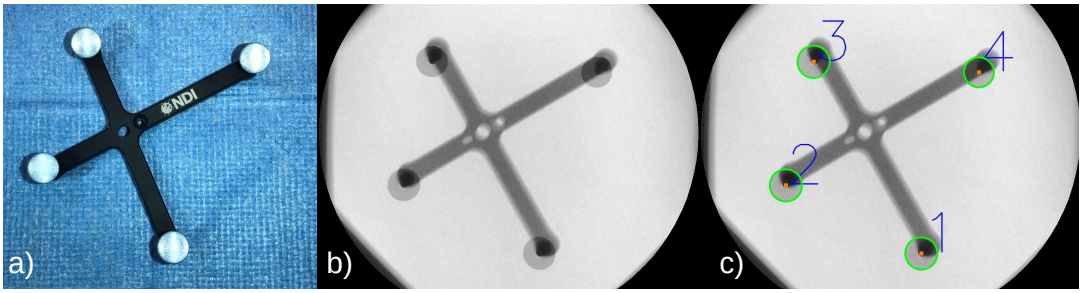

Figure 2.1: a) A photograph of the marker used for offline calibration of the system. Its 3D geometry, and in particular the location of the 4 infrared reflective spheres, is precisely known enabling 3D pose retrieval via outside-in optical tracking. b) An X-ray image of the same marker with c) detected centroids of the spheres. When the marker is stationary, poses extracted from a) and c) enable calibration of the optical tracker to the C-arm source as described in Section 2.2.

translated along the camera's z-axis, *i. e.* along the frustum, while remaining a valid image of the same 3D scene (Georgel, Schroeder, and Navab, 2009). In transmission imaging, this property of the frustum is convenient because the near and far plane of the frustum can always be held constant at $z = 0$ and $z = D_{\mathrm{SD}}$, where D_{SD} is the source-to-detector distance. In other words, there is no need for adaptive view frustum culling (Assarsson and Moller, 2000) since every location on the trajectory of any frustum point will have contributed to the intensity of that point. Consequently, for every structure that is prominent in an X-ray image (*e. g.* a bone contour) there will be a well-defined position z on the frustum, where that image region perfectly coincides with the generating anatomical structure. We will exploit this convenient property to unite and augment the patient with 2D X-ray images acquired in arbitrary geometry. This augmented view onto anatomy is realized using an AR environment that is delivered to the surgeon with a state-of-the-art OST HMD.

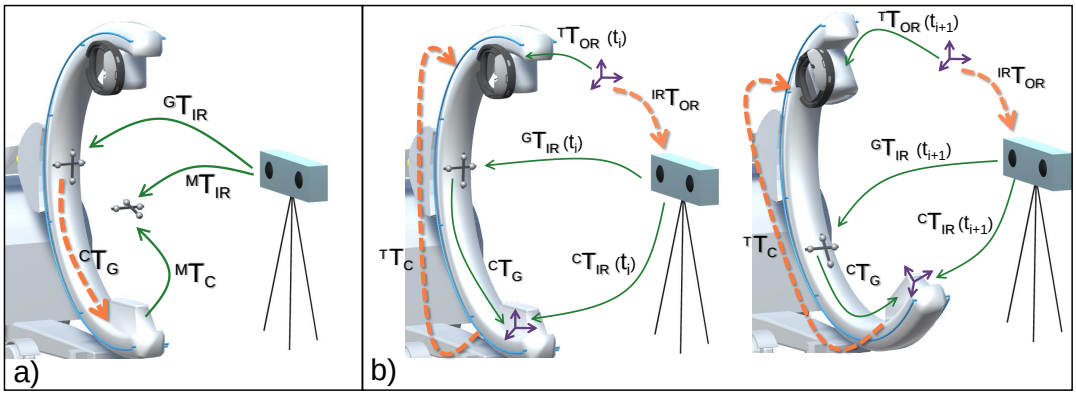

Figure 2.2: Illustrations describing the process of calibrating the tracker to the C-arm X-ray source using hand-eye calibration and an external optical navigation system. a) An infrared reflective marker is attached to the gantry and calibrated to the X-ray source using a second marker that is imaged by the navigation system and the C-arm simultaneously (Fig. 2.1). b) The C-arm gantry, and therefore, the tracker and the optical marker are moved and corresponding pose pairs in the respective frames of reference are collected that are then used for hand-eye calibration following Tsai *et al.* Tsai and Lenz, 1989.

2.2 System Calibration

In order to realize the AR visualization of X-ray images in a spatially-aware manner as described in Section 2.1, the pose of the C-arm defining the corresponding view frustum must be known in the coordinate system of the OST HMD delivering the AR experience. To this end, we rely on a recently proposed approach that is marker-less and radiation-free, and uses vision-based inside-out tracking to dynamically close the calibration loop (Hajek et al., 2018). The inside-out tracking paradigm is driven by the observation that both the surgeon and C-arm navigate the same environment, *i. e.* the operating room, which we will refer to as the "OR coordinate system". For

interventional visualization of X-ray images using IFFs, we must recover:

$$ {}^{\mathrm{S}}\mathbf{T}_{\mathrm{C}}(t) = {}^{\mathrm{S}}\mathbf{T}_{\mathrm{OR}}(t) \underbrace{\left({}^{\mathrm{T}}\mathbf{T}_{\mathrm{OR}}^{-1}(t) \, {}^{\mathrm{T}}\mathbf{T}_{\mathrm{C}}(t_0) \right)}_{{}^{\mathrm{OR}}\mathbf{T}_{\mathrm{C}}(t)}, \tag{2.1} $$

the transformation describing the mapping from the C-arm source coordinate to the surgeon's eyes as both the C-arm and the surgeon move within the environment over time t. In Eq. 2.1, t_0 describes the time of offline calibration. Upon acquisition of X-ray image $\boldsymbol{I}_i$ at time t_i, ${}^{\mathrm{OR}}\mathbf{T}_{\mathrm{C}}(t_i)$ will be held constant, since the viewpoint of the corresponding frustum cannot be altered and only translation of the image along the respective z-axis is permitted. The spatial relations that are required to dynamically estimate ${}^{\mathrm{S}}\mathbf{T}_{\mathrm{C}}(t)$ are explained in the remainder of this section and visualized in Fig. 1.1.

2.2.1 Inside-out Tracking of Surgeon and Tracker on C-arm ${}^{OR}\mathbf{T}_{S/T}$:

Vision-based SLAM is used to incrementally build a map of the environment and estimate the camera's pose ${}^{\mathrm{OR}}\mathbf{T}_{\mathrm{S/T}}$ therein (Endres et al., 2012). Using the surgeon as example, SLAM solves:

$$ {}^{\mathrm{OR}}\mathbf{T}_{\mathrm{S}}(t) = \underset{{}^{\mathrm{OR}}\hat{\mathbf{T}}_{\mathrm{S}}}{\arg\min} \, \mathrm{d}\left(\mathbf{f}_{\mathrm{OR}} \left(\mathbf{P} \, {}^{\mathrm{OR}}\hat{\mathbf{T}}_{\mathrm{S}}(t) \mathbf{x}_{\mathrm{S}}(t) \right), \mathbf{f}_{\mathrm{S}}(t) \right), \tag{2.2} $$

where $\mathbf{f}_{\mathrm{S}}(t)$ are features extracted from the image at time t, $\mathbf{x}_{\mathrm{S}}(t)$ are the 3D locations of these feature obtained *e. g.* via multiview stereo, $\mathbf{P}$ is the projection operator, and $d(\cdot, \cdot)$ is the similarity to be optimized. Following (Hajek et al., 2018), the C-arm gantry is also tracked relative to the *exact same* map of the environment by rigidly attaching an additional tracker to the gantry. To this

end, both trackers are of the same make and model, and are operated in a master-slave configuration. The environmental map provided by the master on start-up of the slave must exhibit partial overlap with the current field of view of the slave tracker, ideally a feature rich and temporally stable area of the environment. As a consequence, the cameras of the C-arm tracker are oriented such that they face the operating theater, and not the surgical site.

2.2.2 One-time Offline Calibration of Tracker to C-arm Source ${}^{T}\mathbf{T}_{C}(t_0)$

: Since the fields of view of the visual tracker and the X-ray scanner do not share overlap, it is not feasible to co-register these sensors via a common calibration phantom. Alternatively, we estimate ${}^{\mathrm{T}}\mathbf{T}_{\mathrm{C}}(t_0)$ via hand-eye calibration, *i. e.* the relative pose information from the rigidly connected tracker and the C-arm are used for solving $\mathbf{X} := {}^{\mathrm{T}}\mathbf{T}_{\mathrm{C}}(t_0)$ in $\mathbf{AX} = \mathbf{XB}$ fashion (Tsai and Lenz, 1989). To construct this over-determined system, the C-arm undergoes different motions along its DoFs, and the corresponding relative pose information of the tracker and the C-arm source are stored in $\mathbf{A}$ and $\mathbf{B}$ matrices, respectively.

Since current C-arms do not exhibit encoded joints, we rely on optical infrared tracking to estimate the pose of the C-arm source. To this end, passive markers $\mathbf{M}$ are introduced into the X-ray field of view and another set of reflective markers $\mathbf{G}$ are rigidly attached to the C-arm gantry (Fig. 2.2-a). The spatial link between the gantry and the source is then estimated via the following equation:

$$ {}^{\mathrm{C}}\mathbf{T}_{\mathrm{G}} = {}^{\mathrm{M}}\mathbf{T}_{\mathrm{C}}^{-1}\,{}^{\mathrm{M}}\mathbf{T}_{\mathrm{IR}}\,{}^{\mathrm{G}}\mathbf{T}_{\mathrm{IR}}^{-1}, \tag{2.3} $$

where ${}^{\mathrm{M}}\mathbf{T}_{\mathrm{C}}$ is the rigid extrinsic parameters expressing the source to marker configuration. To estimate this transformation, spherical marker locations are automatically identified in X-ray images via circular Hough transform. Once ${}^{\mathrm{M}}\mathbf{T}_{\mathrm{C}}$ is estimated, marker $\mathbf{M}$ is removed and the C-arm pose is estimated in the frame of the external infrared navigation system ${}^{\mathrm{C}}\mathbf{T}_{\mathrm{IR}} = {}^{\mathrm{C}}\mathbf{T}_{\mathrm{G}}\,{}^{\mathrm{G}}\mathbf{T}_{\mathrm{IR}}$. To solve the calibration problem in a hand-eye configuration, we construct the following chain of transformations:

$$\begin{aligned} &{}^{\mathrm{T}}\mathbf{T}_{\mathrm{OR}}^{-1}(t_i)\,{}^{\mathrm{T}}\mathbf{T}_{\mathrm{C}}(t_0)\,{}^{\mathrm{C}}\mathbf{T}_{\mathrm{IR}}(t_i) = {}^{\mathrm{T}}\mathbf{T}_{\mathrm{OR}}^{-1}(t_{i+1})\,{}^{\mathrm{T}}\mathbf{T}_{\mathrm{C}}(t_0)\,{}^{\mathrm{C}}\mathbf{T}_{\mathrm{IR}}(t_{i+1}), \\ &\underbrace{{}^{\mathrm{T}}\mathbf{T}_{\mathrm{OR}}(t_{i+1})\,{}^{\mathrm{T}}\mathbf{T}_{\mathrm{OR}}^{-1}(t_i)}_{\mathbf{A}}\underbrace{{}^{\mathrm{T}}\mathbf{T}_{\mathrm{C}}(t_0)}_{\mathbf{X}} = \underbrace{{}^{\mathrm{T}}\mathbf{T}_{\mathrm{C}}(t_0)}_{\mathbf{X}}\underbrace{{}^{\mathrm{C}}\mathbf{T}_{\mathrm{IR}}(t_{i+1})\,{}^{\mathrm{C}}\mathbf{T}_{\mathrm{IR}}^{-1}(t_i)}_{\mathbf{B}}. \end{aligned} \tag{2.4}$$

Eq. 2.4 expresses the relations for poses acquired at times t_i and t_{i+1}. We will then decouple the rotation R_x and translation p_x components. The rotation parameters are estimated using unit quaternion representation Q_x:

$$Q_a\,Q_x = Q_x\,Q_b. \tag{2.5}$$

By re-arranging Eq. 2.5 in the form of $\mathbf{M}Q_x = 0$, we solve for rotation in the following constrained optimization:

$$\min ||\mathbf{M}Q_x||_2^2, \quad \text{s.t.} \;\; ||Q_x||_2^2 = 1\,. \tag{2.6}$$

Finally, the translation component p_x is estimated in a least-squares fashion as expressed in Eq. 2.7, where R denotes rotation matrix:

$$\begin{aligned} R_a p_x + p_a &\approx R_x p_b + p_x\,, \\ (R_a - \mathbb{1})p_x &\approx R_x p_b - p_a\,. \end{aligned} \tag{2.7}$$

2.3 Generating the Frustum K

The view frustum of the C-arm is modeled via the standard 11 DoF camera parameters. In Section 2.2, we presented details for computing the 6 DoF extrinsic parameters ${}^{S}\mathbf{T}_{C}(t)$ relative to the surgeon required for visualization. The remaining 5 DoF intrinsic parameters $\mathbf{K}$ are associated with the focal length, pixel spacing, skew, and principle point that are available from internal calibration of the C-arm and usually provided by the manufacturer. Given these 11 parameters, IFFs are rendered in our AR environment.

2.4 User Interaction

The interaction with the virtual frustum of the X-ray image in the augmented surgery environment is built upon the surgeon's gaze, hand gesture, and voice commands. The intersection of the gaze ray and a virtual object is used as the mechanism to select and highlight an X-ray image that, potentially, is minimized to a point in its focal point location. This image can then be manipulated with a single DoF to slide along the z-axis through the frustum following the surgeon's hand gestures that are detected by the gesture-sensing cameras on the OST HMD. The virtual frustum is rendered in red as the image reaches the source, and in blue as the image approaches the detector. The transparency of each image can be changed by moving a slider next to the image. Finally, the voice commands *Lock* and *Unlock* allow the user to lock and unlock the pose of the virtual image, and the use of voice command *Next* highlights the next acquired X-ray image within the corresponding frustum.

Chapter 3

Experimental Results and Evaluation

3.1 System Setup

While the concept described above is generic, we materialized and evaluated a prototype of the described system using hardware components available on site. For intra-operative X-ray imaging we used an ARCADIS Orbic 3D C-arm (Siemens Healthineers, Forchheim, Germany). IFFs and X-ray images were displayed in the AR environment to the surgeon using a Microsoft HoloLens OST HMD (Microsoft, Redmond, WA). The AR environment used to render IFFs and all other virtual content was built as a holographic Universal Windows Platform (UWP) application using the Unity (version 2017.4.17f1 Personal (64bit), Unity Technologies, San Francisco, CA, USA) and Microsoft Visual Studio (version Community 2017, Microsoft Corporation, Redmond, WA, USA).

A second HoloLens device was rigidly connected to the C-arm gantry serving as the inside-out tracker. These two HMDs shared anchors that were

computed from the visual structures of the operating room and communicated over a local wireless network. The interconnection between the HoloToolkit-enabled apps allowed the HMDs to collaborate in a master-slave configuration and remain in sync seamlessly. A sharing service running on a Windows 10 development PC managed the connection of these remote devices and enabled streaming of X-ray images. Transfer of intra-operative X-ray images from the C-arm to the development PC was done via Ethernet. Lastly, for the offline co-calibration of ${}^{\mathrm{T}}\mathbf{T}_{\mathrm{C}}(t_0)$ between the tracker and the X-ray source, a Polaris Spectra external optical navigation system (Northern Digital, Waterloo, ON) was employed.

3.2 Analysis of Hand-Eye Calibration

Offline estimation of the relation between the passive markers **G** and the X-ray source constitutes the first and critical step in closing the transformation chain. To this end, we estimated ${}^{\mathrm{C}}\mathbf{T}_{\mathrm{G}}$ via Eq. 2.3 by acquiring pose information from 7 different poses of marker **M**. From these 7 poses, we seek to compute the mean rotation and translation. The mean rotation matrix $\overline{R}$ is computed on the Special Orthogonal group $SO(3)$ by minimizing:

$$\underset{\overline{R}\in SO(3)}{\arg\min}\sum_{i=1}^{N} d(R_i,\overline{R})^2 \tag{3.1}$$

where $d(.)$ denotes a distance function on the Riemannian manifold. To establish $d(.)$, the rotation matrix is expressed in the Lie algebra (tangent space) of the Lie group as $R = e^{\hat{w}}$. The tangent space $\hat{w}$ is then obtained as $log(R) = \hat{w}$, such that $\hat{w}$ is the skew-symmetric matrix constructed from the

vector w. Consequently, the mean rotation is estimated as (Moakher, 2002)

$$\underset{\overline{R}\in SO(3)}{\arg\min}\sum_{i=1}^{N}\|log(R_i^T\overline{R}\|_F^2 \tag{3.2}$$

where $\|.\|_F^2$ is the Frobenius norm. The mean translation $\bar{t}$ is computed in Euclidean space as:

$$\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t_i \tag{3.3}$$

Next, the hand-eye calibration problem (Eq. 2.4) was solved by simultaneously acquiring $N = 120$ corresponding poses from both the SLAM tracker on the C-arm and the external navigation system as the C-arm gantry underwent different motion. The acquisition of pose data was synchronized by locking the C-arm at each configuration, and recording pose transformations as per the visual tracker and external navigation system using clicker and keyboard commands, respectively. During data acquisition, the C-arm was rotated up to its maximum range for cranial, caudal, and swivel directions. For the left and right anterior oblique views it was orbited up to $\pm 35^{\circ}$.

Residual errors in Table 3.1 were calculated separately for translation and rotation using:

$$\begin{aligned} p_e &= \frac{\sum_{i=1}^{N} p(\mathbf{A}_i\mathbf{X} - \mathbf{X}\mathbf{B}_i)}{N}\,, \text{ and} \\ R_e &= (R_a R_x)^{-1} R_x R_b\,. \end{aligned} \tag{3.4}$$

The calibration accuracy was evaluated for rotation $\overline{e_R}(N)$ and translation $\overline{e_p}(N)$ with respect to the number of pose data N. The evaluation procedure is demonstrated in Alg. 1 (Tsai and Lenz, 1989). The mean and standard deviation for translational and rotational errors are shown in Figs. 3.1 and 3.2.

Table 3.1: Error measures for tracker to C-arm hand-eye calibration of ${}^{T}\mathbf{T}_{C}(t_0)$.

Hand-Eye Calibration	**Residual** $(x, y, z), \|.\|_2$	**RMS**	**Median** (x, y, z)	**SD** $(x, y, z), \|.\|_2$
Rotation (deg)	$(0.77, 1.2, 0.82), 1.7$	0.98	$(0.24, 0.11, 0.24)$	$(0.72, 1.1, 0.75), 1.7$
Translation (mm)	$(7.7, 8.2, 8.7), 14$	8.2	$(3.6, 5.4, 3.4)$	$(7.2, 7.4, 7.6), 8.2$

Figure 3.1: Translational error in mm units with respect to number of pose pairs. The shaded area represents standard deviation of the error.

Results in these two plots indicate that the improvement in accuracy when using more than 50 pose pairs are minimal, suggesting convergence.

3.3 C-arm Pose Estimation via Integrated Visual Tracking

In Table 3.2 we compared the tracking results of the X-ray source using our inside-out visual SLAM system to a baseline approach using outside-in external navigation as in Fig. 2.2. The evaluation was performed over 20 different

Algorithm 1 Assessment of hand-eye calibration

$p(.)$: translation component of a homogeneous rigid transformation
$tr(.)$: matrix trace
$\overline{(.)}$: expected value

procedure HAND-EYE ACCURACY ASSESSMENT (M pairs)
  for `N = 4 to M` **do** ▷ M = 120
    for `i = 1 to T` **do** ▷ T = 800
      `Randomly select N pairs from M`
      `Estimate hand-eye calibration` ${}^{C}\mathbf{T}_{T}$ `with the N pairs`
      for `j = 1 to N` **do**
        ${}^{IR}\mathbf{T}_{W}(j) = {}^{IR}\mathbf{T}_{C}(j)\,{}^{C}\mathbf{T}_{T}\,{}^{T}\mathbf{T}_{W}(j)$
      end for

      `Estimate mean transformation:` ${}^{IR}\overline{\mathbf{T}}_{W}$
      `Randomly select M pairs from M`
      for `k = 1 to M` **do**
        ${}^{C}\mathbf{T}_{Wa}(k) = {}^{C}\mathbf{T}_{IR}(k)\,{}^{IR}\overline{\mathbf{T}}_{W}$
        ${}^{C}\mathbf{T}_{Wb}(k) = {}^{C}\mathbf{T}_{T}\,{}^{T}\mathbf{T}_{W}(k)$
        $R = {}^{C}\mathbf{R}_{Wa}(k)\,{}^{C}\mathbf{R}^{-1}_{Wb}(k)$
        $e_R(N,i,k) = \cos^{-1}(\frac{tr(R)-1}{2})$
        $e_p(N,i,k) = \left\| p({}^{C}\mathbf{T}_{Wa}(k) - {}^{C}\mathbf{T}_{Wb}(k)) \right\|_2$
      end for

      $\overline{e_R}(N,i) = \frac{1}{M}\sum_{k=1}^{M} e_R(N,i,k)$
      $\overline{e_p}(N,i) = \frac{1}{M}\sum_{k=1}^{M} e_p(N,i,k)$
    end for

    $\overline{e_R}(N) = \frac{1}{T}\sum_{i=1}^{T} e_R(N,i)$
    $\overline{e_p}(N) = \frac{1}{T}\sum_{i=1}^{T} e_p(N,i)$
  end for

end procedure

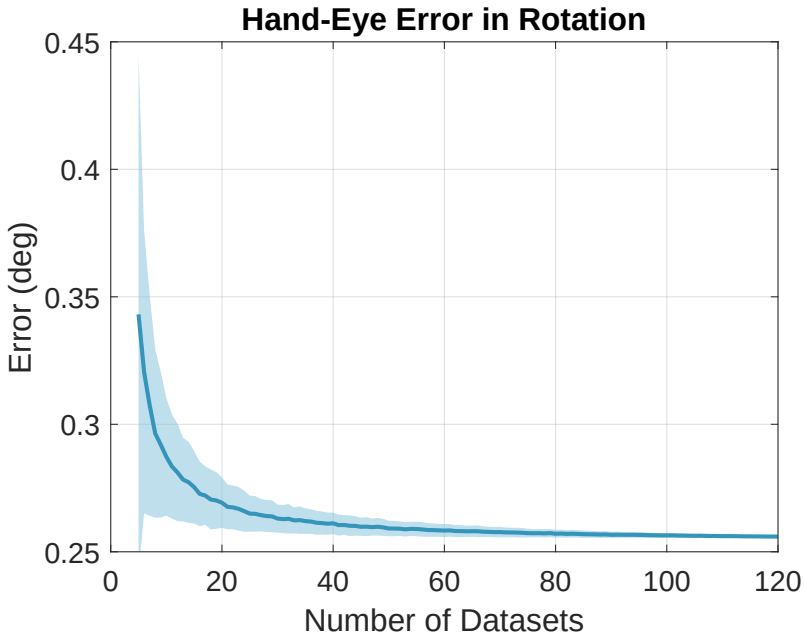

Figure 3.2: Rotational error in degree units with respect to number of pose pairs. The shaded area represents standard deviation of the error.

Table 3.2: Error measures for C-arm extrinsic parameter estimation using SLAM tracking.

C-arm Tracking	**Residual** $(x, y, z), \|.\|_2$	**RMS**	**Median** (x, y, z)	**SD** $(x, y, z), \|.\|_2$
Rotation (deg)	$(0.71, 0.11, 0.74), 0.75$	0.43	$(0.21, 0.12, 0.24)$	$(0.24, 0.09, 0.23), 0.34$
Translation (mm)	$(4.0, 5.0, 4.8), 8.0$	4.6	$(3.3, 3.6, 3.3)$	$(1.3, 1.7, 1.6), 2.7$

C-arm angulations.

3.4 Target Augmentation Error in Localizing Fiducials

The end-to-end error of the augmentation requires a user-in-the-loop design and was evaluated using a multi-planar phantom with $L = 4$ radiopaque fiducial markers placed at different (x, y, z) positions shown in Fig. 3.3. We

Figure 3.3: The multi-level hybrid phantom with X-ray visible fiducials along $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ is used to evaluate the augmentation error.

computed a planar target-augmentation-error (TAE) by manipulating the virtual X-ray image in the frustum for every fiducial separately such that the virtual image plane perfectly intersected the true location of the respective fiducial. Together with a manual annotation of the fiducial in the image plane and the location of the frustum we determined the 3D position of the respective virtual landmark in the AR environment. To retrieve the 3D position of the corresponding real fiducial required for error computation, the user were asked to select the fiducial landmarks at the surface of the phantom using interactive hand gesture or voice commands for $N = 20$ trials. Per selection, a single ray was cast connecting the the HMD and the targeted point on the phantom. To localize the 3D position of the fiducial along the ray, users targeted the same point 4 times from various views around the phantom.

Each ray is defined via two elements: *1)* the 3D position of HMD p_i,

Table 3.3: TAE measurements from 20 trials.

Target Augmentation Error (TAE)	Mean	RMS	Median	Standard Deviation
	13.2 mm	13.5 mm	12.1 mm	2.89 mm

and *2)* u_i as the unit direction vector normal to HMD. After all rays were identified, we estimated the closest point x_l^* to all the rays corresponding to each landmark l via a least-squares minimization strategy as follows:

$$\begin{aligned} x_l^* &= \arg\min_{x \in \mathbb{R}^3} \sum_{i=1}^{N} \| (I - u_i u_i^\top) x - t_i \|^2 \text{, where} \\ t_i &= (I - u_i u_i^\top) p_i \,. \end{aligned} \tag{3.5}$$

Finally, the average TAE error was calculated as the average distance between the corresponding landmarks selected by the user, and the landmarks identified in the X-ray image as follows:

$$\overline{\text{TAE}} = \frac{1}{L \times N} \sum_{l=1}^{l=L\times N} \left\| x_l^* - {}^{\text{S}}\mathbf{T}_{\text{C}} \, {}^{\text{K}}\mathbf{T}_{\text{C}}^{-1} \, x_l^{\text{K}} \right\|_2 , \tag{3.6}$$

where x_l^{K} corresponds to the l^{th} landmark in the virtual frustum coordinate frame, ${}^{\text{K}}\mathbf{T}_{\text{C}}$ is the transformation from each virtual frustum to the C-arm source coordinate. TAE measurements are presented in Table 3.3.

3.5 Spatially-Aware Visualization And Surgical Use Cases

We demonstrate the application of spatially-aware X-ray image visualization on the view frustum using two high volume clinical procedures that are routinely performed under C-arm fluoroscopy guidance: 1) Internal fixation

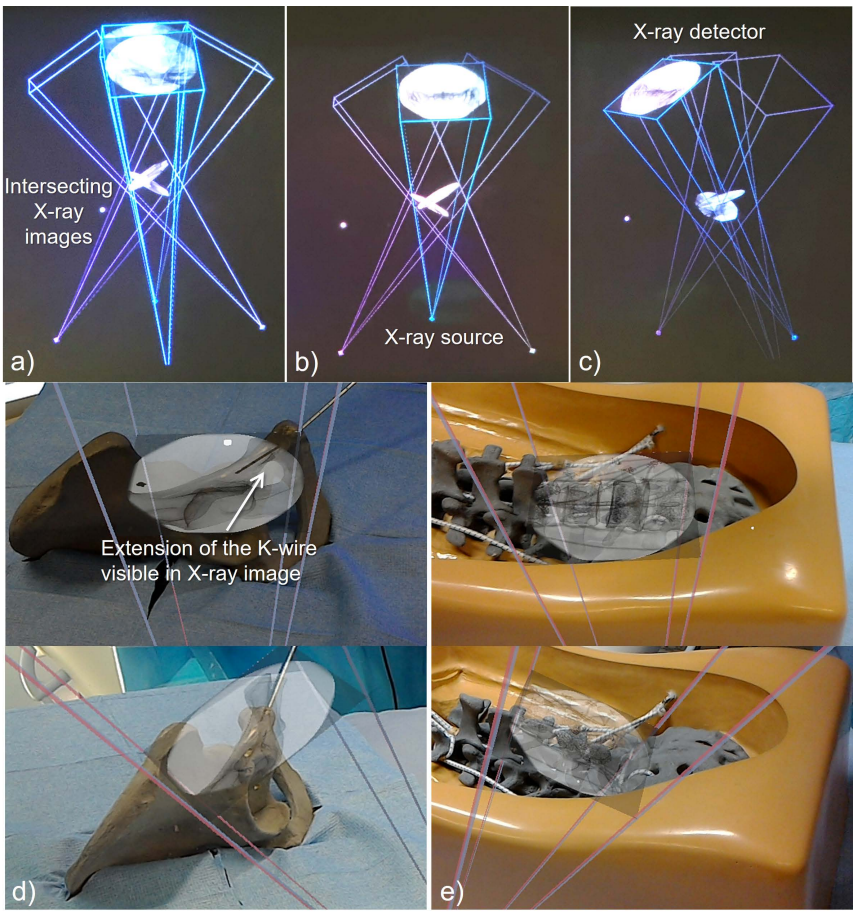

Figure 3.4: Multiple views of IFFs shown on the HMD in a-c). d) and e) show the augmentation of the virtual view frustum and the corresponding C-arm images from two views on a pelvis and a spine phantom.

of pelvic ring fractures (Tile, 1988; Routt Jr, Nork, and Mills, 2000), and 2) percutaneous spine procedures such as percutaneous vertebroplasty (Barr et al., 2000). We show exemplary scenes of the aforementioned cases in Fig. 3.4. In these use cases, the user is able to preview the IFFs from a Windows 10 PC, then send the selected ones to the HMD via a local wireless network (Fig. 3.5). Different features are available to the user to interact with IFFs shown in Fig. 3.6 for better visualization.

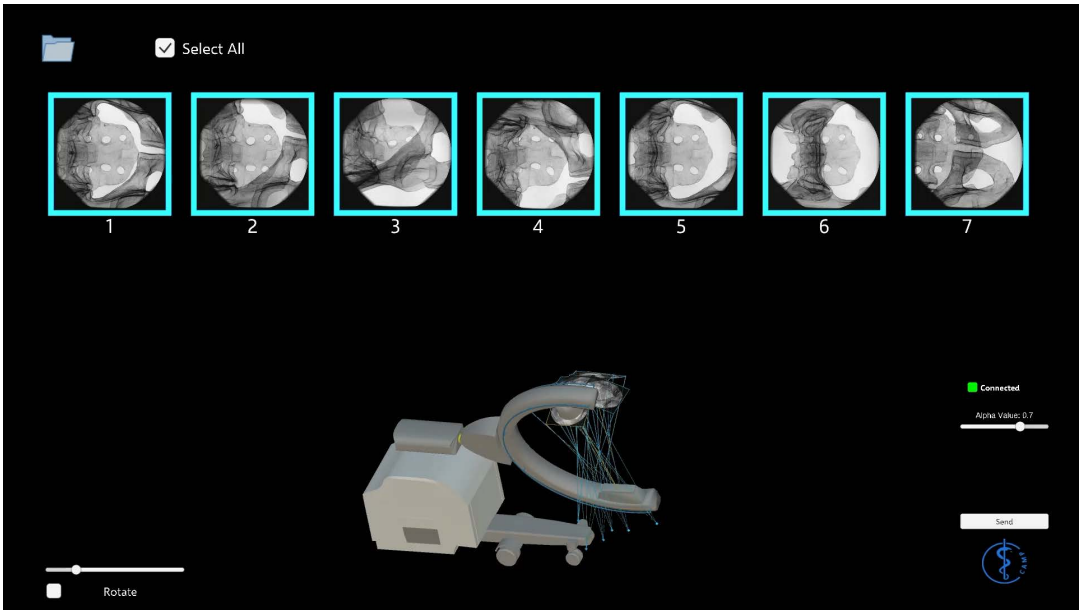

Figure 3.5: multiple views of IFFs shown on the PC running on Windows 10. The user is able to manipulate IFFs on the C-arm to preview the X-ray images from different angles.

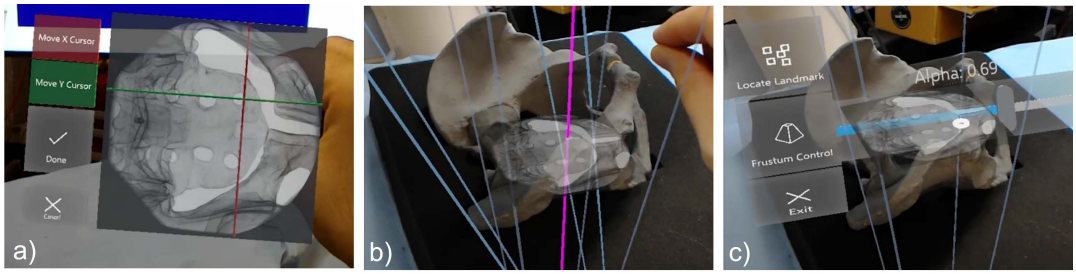

Figure 3.6: a) Annotation on selected X-ray image. b) The resulting ray connecting the C-arm source and the annotation point. c) Adjustable transparency of each X-ray image.

Chapter 4

Discussion and Conclusion

This work presents a spatially-registered AR solution for fluoroscopy-guided surgery. In the proposed visualization strategy interventional images are displayed within the corresponding view frustum of the C-arm and can, therefore, be meaningfully intersected with the imaged objects. This solution introduces several advantages: First, registration between the surgeon and the C-arm X-ray system is real-time, dynamic, and marker-free. Second, the surgeon's HMD and the SLAM tracker on the C-arm are operated in master-slave configuration such that both devices use the same visual structures in the operating theatre as common fiducials. Lastly, exploiting imaging-geometry of projective images for *in situ* augmentation of the patient anatomy eliminates the need for complex and ill-posed image-based 2D/3D registration between X-ray images and pre-operative data.

The concept of spatially-aware AR can also generalize to assist X-ray technicians in reproducing C-arm views by navigating the scanner such that IFFs align (Unberath et al., 2018). Moreover, our system enables storage of a map of the operating theater, the position of the surgeon and the C-arm

including all acquired images in the correct spatial configuration, and the audio footage throughout the procedure, thus virtual replay of surgery is possible and may be an effective training tool for orthopedic surgery residents and fellows.

The visual tracker on the C-arm localizes the scanner in relation to both the surgical environment and the augmented surgeon. Therefore, if the C-arm is displaced, the viewing frustum is dynamically updated in real-time, hence IFFs render with the new alignment. This will, therefore, allow the use of IFFs with mobile C-arm systems as shown in Ch. 3. Since IFFs transformation parameters are estimated globally within the operating theatre coordinate frame, even if the C-arm scanner is moved away, previously acquired images will still render within their spatially-fixed viewing frustum in the operating theatre. Finally, IFFs paradigm enables a flexible AR strategy such that no external setup or additional interventional calibration and registration steps are necessary.

In our quantitative evaluation reported in Chapter 3, we found low orientational errors. On the other hand, the overall translation error for tracking the C-arm using the inside-out tracker was 8.0 mm. The errors persisting after hand-eye calibration (Table 3.1) are similar to the errors observed during tracking (Table 3.2) suggesting that the remaining error is statistic and the data used for offline calibration was acquired with sufficient variation of the C-arm pose. Further reductions in residual error of hand-eye calibration and TAE would be desirable for guiding tools in complex anatomy, but would require improvements in SLAM-based tracking that are intractable given the use of

off-the-shelf OST HMDs that are optimized for entertainment rather than medical application. Results suggest that IFFs is suited for surgical tasks were millimeter accuracy is not required, for instance, C-arm re-positioning and X-ray image re-acquisition from different views for verifying tool placement. We foresee further potential applications in surgical tasks that predominantly require orientational information, *e. g.* adjusting the anteversion and abduction angles for placing acetabular components in total hip arthroplasty in the direct anterior approach (Fotouhi et al., 2018). We believe that IFFs paradigm is a step towards removing ambiguities present in the projective images acquired intra-operatively. For surgical interventions that require full 3D information, either pre-operative CT images need to be registered to the patient, or 3D intra-operative imaging would be employed (Atria et al., 2018).

In C-arm imaging, the X-ray source is typically placed below the patient bed. To ease interpretation of the acquired images, it is common to display the images with left-right flip to provide an impression that the images are acquired from above the surgical bed, since this more closely resembles the surgeon's view onto anatomy. To augment the surgical site with virtual images on the view frustum, the images have to undergo a similar flip such that they align with the patient when observed from the surgeon's viewpoint. Another important note regarding this proof-of-principle work is that we approximated the intrinsic geometry $\mathbf{K}$ of the C-arm to be constant across all poses. However, due to mechanical sag, the relation between the X-ray source and detector, and therefore $\mathbf{K}$, is not perfectly constant but slightly changes at different orientations. In future work, this simplification should be considered, *e. g.* by

using look-up tables or a virtual detector mechanism (Navab and Mitschke, 2001).

The proposed concept and prototypical results are promising and encourage further research that will include user studies on cadaveric specimens to validate the clinical usability of this approach. The future surgeon-centered experiments will evaluate the system performance in real surgical environments under varying conditions for specific surgical tasks. We believe the proposed technique to improve surgical perception can contribute to optimizing surgical performance and pave the way for enhancing visuo-motor coordination.

References

Hott, Jonathan S, Vivek R Deshmukh, Jeffrey D Klopfenstein, Volker KH Sonntag, Curtis A Dickman, Robert F Spetzler, and Stephen M Papadopoulos (2004). "Intraoperative Iso-C C-arm navigation in craniospinal surgery: the first 60 cases". In: *Neurosurgery* 54.5, pp. 1131–1137.

Miller, Donald L, Eliseo Vañó, Gabriel Bartal, Stephen Balter, Robert Dixon, Renato Padovani, Beth Schueler, John F Cardella, and Thierry De Baère (2010). "Occupational radiation protection in interventional radiology: a joint guideline of the Cardiovascular and Interventional Radiology Society of Europe and the Society of Interventional Radiology". In: *Cardiovascular and interventional radiology* 33.2, pp. 230–239.

Theocharopoulos, Nicholas, Kostas Perisinakis, John Damilakis, George Papadokostakis, Alexander Hadjipavlou, and Nicholas Gourtsoyiannis (2003). "Occupational exposure from common fluoroscopic projections used in orthopaedic surgery". In: *JBJS* 85.9, pp. 1698–1703.

Mason, Alexander, Renee Paulsen, Jason M Babuska, Sharad Rajpal, Sigita Burneikiene, E Lee Nelson, and Alan T Villavicencio (2014). "The accuracy of pedicle screw placement using intraoperative image guidance systems: A systematic review". In: *Journal of Neurosurgery: Spine* 20.2, pp. 196–203.

Suzuki, T, K Soma, M Shindo, H Minehara, and M Itoman (2008). "Anatomic study for pubic medullary screw insertion". In: *Journal of Orthopaedic Surgery* 16.3, pp. 321–325.

Starr, AJ, AL Jones, CM Reinert, and DS Borer (2001). "Preliminary results and complications following limited open reduction and percutaneous screw fixation of displaced fractures of the acetabulum." In: *Injury* 32, SA45–50.

Boszczyk, Bronek M, Michael Bierschneider, Stephanie Panzer, Werner Panzer, Roger Harstall, Katharina Schmid, and Hans Jaksche (2006). "Fluoroscopic radiation exposure of the kyphoplasty patient". In: *European Spine Journal* 15.3, pp. 347–355.

Synowitz, Michael and Juergen Kiwit (2006). "SurgeonâĂŹs radiation exposure during percutaneous vertebroplasty". In: *Journal of Neurosurgery: Spine* 4.2, pp. 106–109.

Cardin, M-A, J-X Wang, and D-B Plewes (2005). "A method to evaluate human spatial coordination interfaces for computer-assisted surgery". In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005*, pp. 9–16.

Westwood, James D et al. (2005). "The mini-screen: an innovative device for computer assisted surgery systems". In: *Medicine Meets Virtual Reality 13: The Magical Next Becomes the Medical Now* 111, p. 314.

Chimenti, Peter C and David J Mitten (2015). "Google Glass as an alternative to standard fluoroscopic visualization for percutaneous fixation of hand fractures: a pilot study". In: *Plastic and reconstructive surgery* 136.2, pp. 328–330.

Yoon, Jang W, Robert E Chen, Phillip K Han, Phong Si, William D Freeman, and Stephen M Pirris (2016). "Technical feasibility and safety of an intraoperative head-up display device during spine instrumentation". In: *The International Journal of Medical Robotics and Computer Assisted Surgery*.

Qian, Long, Mathias Unberath, Kevin Yu, Bernhard Fuerst, Alex Johnson, Nassir Navab, and Greg Osgood (2017). "Towards virtual monitors for image guided interventions-real-time streaming to optical see-through head-mounted displays". In: *arXiv preprint arXiv:1710.00808*.

Deib, Gerard, Alex Johnson, Mathias Unberath, Kevin Yu, Sebastian Andress, Long Qian, Gregory Osgood, Nassir Navab, Ferdinand Hui, and Philippe Gailloud (2018). "Image guided percutaneous spine procedures using an optical see-through head mounted display: proof of concept and rationale". In: *Journal of neurointerventional surgery*, neurintsurg–2017.

Vorraber, Wolfgang, Siegfried Voessner, Gerhard Stark, Dietmar Neubacher, Steven DeMello, and Aaron Bair (2014). "Medical applications of near-eye display devices: an exploratory study". In: *International Journal of Surgery* 12.12, pp. 1266–1272.

Chen, Xiaojun, Lu Xu, Yiping Wang, Huixiang Wang, Fang Wang, Xiangsen Zeng, Qiugen Wang, and Jan Egger (2015). "Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display". In: *Journal of biomedical informatics* 55, pp. 124–131.

Hanna, George B, Sami M Shimi, and Alfred Cuschieri (1998). "Task performance in endoscopic surgery is influenced by location of the image display." In: *Annals of surgery* 227.4, p. 481.

Liebergall, Meir, Rami Mosheiff, and Leo Joskowicz (2006). "Computer-aided orthopaedic surgery in skeletal trauma". In: *Bucholz RW, Heckman JD, Court-Brown CM, editors. Rockwood & greenâĂŹs fractures in adults. 6th ed., Philadelphia: Lippincott*, pp. 739–67.

Joskowicz, Leo and Eric J. Hazan (2016). "Computer Aided Orthopaedic Surgery: Incremental shift or paradigm change?" In: *Medical Image Analysis* 33, pp. 84 –90.

Theologis, AA, S Burch, and M Pekmezci (2016). "Placement of iliosacral screws using 3D image-guided (O-Arm) technology and Stealth Navigation: comparison with traditional fluoroscopy". In: *The bone & joint journal* 98.5, pp. 696–702.

Alambeigi, Farshid, Yu Wang, Shahriar Sefati, Cong Gao, Ryan J Murphy, Iulian Iordachita, Russell H Taylor, Harpal Khanuja, and Mehran Armand (2017). "A curved-drilling approach in core decompression of the femoral head osteonecrosis using a continuum manipulator". In: *IEEE Robotics and Automation Letters* 2.3, pp. 1480–1487.

Fischer, Marius, Bernhard Fuerst, Sing Chun Lee, Javad Fotouhi, Severine Habert, Simon Weidert, Ekkehard Euler, Greg Osgood, and Nassir Navab (2016). "Preclinical usability study of multiple augmented reality concepts for K-wire placement". In: *International journal of computer assisted radiology and surgery* 11.6, pp. 1007–1014.

Hajek, Jonas, Mathias Unberath, Javad Fotouhi, Bastian Bier, Sing Chun Lee, Greg Osgood, Andreas Maier, Mehran Armand, and Nassir Navab (2018). "Closing the Calibration Loop: An Inside-out-tracking Paradigm for Augmented Reality in Orthopedic Surgery". In: *arXiv preprint arXiv:1803.08610*.

Sefati, Shahriar, Farshid Alambeigi, Iulian Iordachita, Mehran Armand, and Ryan J Murphy (2016). "FBG-based large deflection shape sensing of a continuum manipulator: Manufacturing optimization". In: *SENSORS, 2016 IEEE*. IEEE, pp. 1–3.

Sefati, Shahriar, Michael Pozin, Farshid Alambeigi, Iulian Iordachita, Russell H Taylor, and Mehran Armand (2017). "A highly sensitive fiber Bragg Grating shape sensor for continuum manipulators with large deflections". In: *SENSORS, 2017 IEEE*. IEEE, pp. 1–3.

Tucker, Emerson, Javad Fotouhi, Mathias Unberath, Sing Chun Lee, Bernhard Fuerst, Alex Johnson, Mehran Armand, Greg M Osgood, and Nassir Navab (2018). "Towards clinical translation of augmented orthopedic surgery: from pre-op CT to intra-op x-ray via RGBD sensing". In: *Medical Imaging*

2018: Imaging Informatics for Healthcare, Research, and Applications. Vol. 10579. International Society for Optics and Photonics, 105790J.

Unberath, Mathias, Javad Fotouhi, Jonas Hajek, Andreas Maier, Greg Osgood, Russell Taylor, Mehran Armand, and Nassir Navab (2018). "Augmented reality-based feedback for technician-in-the-loop C-arm repositioning". In: *Healthcare Technology Letters*.

Qian, Long, Anton Deguet, and Peter Kazanzides (2018). "ARssist: augmented reality on a head-mounted display for the first assistant in robotic surgery". In: *Healthcare Technology Letters*.

Augmedics (2018). *Augmedics xvision: Pre-clinical Cadaver Study*. https://www.augmedics.com/cadaver-study.

Fotouhi, Javad, Clayton P Alexander, Mathias Unberath, Giacomo Taylor, Sing Chun Lee, Bernhard Fuerst, Alex Johnson, Greg M Osgood, Russell H Taylor, Harpal Khanuja, et al. (2018). "Plan in 2-D, execute in 3-D: an augmented reality solution for cup placement in total hip arthroplasty". In: *Journal of Medical Imaging* 5.2, p. 021205.

Gafoor, Sameer, Philipp Schulz, Luisa Heuer, Predrag Matic, Jennifer Franke, Stefan Bertog, Markus Reinartz, Laura Vaskelyte, Ilona Hofmann, and Horst Sievert (2015). "Use of EchoNavigator, a Novel Echocardiography-Fluoroscopy Overlay System, for Transseptal Puncture and Left Atrial Appendage Occlusion". In: *Journal of interventional cardiology* 28.2, pp. 215–217.

Fotouhi, Javad, Bernhard Fuerst, Alex Johnson, Sing Chun Lee, Russell Taylor, Greg Osgood, Nassir Navab, and Mehran Armand (2017). "Pose-aware C-arm for automatic re-initialization of interventional 2D/3D image registration". In: *International journal of computer assisted radiology and surgery* 12.7, pp. 1221–1230.

Andress, Sebastian, Alex Johnson, Mathias Unberath, Alexander Felix Winkler, Kevin Yu, Javad Fotouhi, Simon Weidert, Greg Osgood, and Nassir Navab (2018). "On-the-fly augmented reality for orthopedic surgery using a multimodal fiducial". In: *Journal of Medical Imaging* 5.2, p. 021209.

Georgel, Pierre Fite, Pierre Schroeder, and Nassir Navab (2009). "Navigation tools for viewing augmented cad models". In: *IEEE Computer graphics and applications* 29.6, pp. 65–73.

Hartley, Richard and Andrew Zisserman (2003). *Multiple view geometry in computer vision*. Cambridge university press.

Assarsson, Ulf and Tomas Moller (2000). "Optimized view frustum culling algorithms for bounding boxes". In: *Journal of graphics tools* 5.1, pp. 9–22.

Tsai, Roger Y and Reimar K Lenz (1989). "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration". In: *IEEE Transactions on robotics and automation* 5.3, pp. 345–358.

Endres, Felix, Jürgen Hess, Nikolas Engelhard, Jürgen Sturm, Daniel Cremers, and Wolfram Burgard (2012). "An evaluation of the RGB-D SLAM system". In: *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, pp. 1691–1696.

Moakher, Maher (2002). "Means and averaging in the group of rotations". In: *SIAM journal on matrix analysis and applications* 24.1, pp. 1–16.

Tile, Marvin (1988). "Pelvic ring fractures: should they be fixed?" In: *The Journal of bone and joint surgery. British volume* 70.1, pp. 1–12.

Routt Jr, ML Chip, Sean E Nork, and William J Mills (2000). "Percutaneous fixation of pelvic ring disruptions." In: *Clinical Orthopaedics and Related Research (1976-2007)* 375, pp. 15–29.

Barr, John D, Michelle S Barr, Thomas J Lemley, and Richard M McCann (2000). "Percutaneous vertebroplasty for pain relief and spinal stabilization". In: *Spine* 25.8, pp. 923–928.

Atria, Cristian, Lisa Last, Nathan Packard, and Frederic Noo (2018). "Cone beam tomosynthesis fluoroscopy: a new approach to 3D image guidance". In: *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*. Vol. 10576. International Society for Optics and Photonics, p. 105762V.

Navab, Nassir and Matthias Mitschke (2001). *Method and apparatus using a virtual detector for three-dimensional reconstruction from x-ray images*.

Tianyu Song

• Baltimore, MD • 410-800-8663 • tsong11@jhu.edu • https://github.com/stytim

EDUCATION

Whiting School of Engineering, Johns Hopkins University August 2017 – Present
Master of Science in Robotics Expected Graduation Date: May 2019
Core Courses: •Computer Vision •Augmented Reality •Computer-Integrated Surgery •Machine Learning •ROS Programming •Algorithm for Sensor-Based Robotics

Purdue School of Engineering and Technology, Indiana University – Purdue University Indianapolis
Bachelor of Science in Mechanical Engineering August 2015 – May 2017
Core Courses: •Feedback Control of Dynamic System •CAD/CAM •Finite Element Analysis •Machine Design

School of Engineering, Sun Yat-sen University
Bachelor of Science in Theoretical and Applied Mechanics August 2013 – June 2017

SKILLS

- **Programming:** Python, C#, C++, Matlab
- **Tools & Libraries:** Unity, Vuforia, ROS, OpenCV, Pytorch
- **Languages:** English (fluent), Mandarin Chinese (native), Cantonese Chinese (native)

PUBLICATIONS

J. Fotouhi*, M. Unberath*, **T. Song***, W. Gu, A. Johnson, M. Armand, G. Osgood, N. Navab. "Interactive Flying Frustums (IFFs): Spatially-aware Surgical Data Visualization", *International Journal of Computer Assisted Radiology and Surgery,* 2019, pp. 1-9.
* Joint first authors
Special Issue: *10th International Conference on Information Processing in Computer-Assisted Interventions (IPCAI), 2019*

T. Song*, C. Yang*, O. Dianat, E. Azimi. "Endodontic Guided Treatment Using Augmented Reality on a Head-Mounted Display System", *IET Healthcare Technology Letters,* vol. 5, no. 2, 2018, pp. 201-207
* Joint first authors
Received the **Outstanding Paper Award** at the *12th MICCAI Workshop on Augmented Environments for Computer Assisted Interventions (AE-CAI)* September 2018

WORK

Course Assistant – Augmented Reality January 2019 – Present
(Professor: Dr. Nassir Navab)

Course Assistant – Computer Vision September 2018 – December 2018
(Professor: Dr. Haider Ali)

Summer Research Intern
Computer Aided Medical Procedures, Johns Hopkins University May 2018 – August 2018

- Studied and designed various Medical Augmented Reality solutions to improve medical procedures' quality, efficiency, and safety.

Product Engineer Intern
Remote Sensing Park of Digital China - Peking University June 2017 – August 2017

- Designed and 3D-Printed various parts for industrial UAVs and FPV racing drones.

PROJECTS

Augmented Reality Magnifying Loupe for Surgery
(Mentor: Dr. Mathias Unberath, Dr. Peter Kazanzides, Long Qian) February 2018 – Present

- Objectives: Design a surgical loupe mount for optical see-through head-mounted display (HMD) and develop a calibration method to associate the field-of-magnified-vision, the HMD screen space and the task workspace.

Surgical Data Visualization in AR for Fluoroscopy-guided Surgery
(Mentor: Dr. Nassir Navab, Dr. Mathias Unberath, Javad Fotouhi) September 2018 – November 2018

- Implemented a prototype using HoloLens to visualize interventional X-ray images within the corresponding view frustum of the C-arm to the surgeon in an AR environment.
- Conducted 2D/3D registration, camera calibration and hand-eye calibration.
- Paper entitled ***Interactive Flying Frustums (IFFs): Spatially-aware Surgical Data Visualization*** has been accepted for presentation at the 10th International Conference on Information Processing in Computer-Assisted Interventions (IPCAI), 2019.

AR for Endodontic (Root canal treatment) Planning and Implementation
(Mentor: Dr. Nassir Navab, Ehsan Azimi) March 2018 – July 2018

- Implemented a prototype using HoloLens with Unity and Vuforia to provide Augmented Reality guidance in the medical procedure of creating access cavity.
- Paper entitled ***Endodontic Guided Treatment Using Augmented Reality on a Head - Mounted Display System*** was accepted for long oral presentation at 12th MICCAI Workshop on Augmented Environments for Computer Assisted Interventions (AE-CAI) in Granada, Spain.

3D Visual SLAM and Path Planning with drone
(Mentor: Dr. Louis Whitcomb) March 2018 – May 2018

- Implemented LSD-SLAM in ROS Kinetic using monocular vision on the drone to robustly track its pose and reconstructed the surrounding 3D environment simultaneously.

3D Reconstructions from Multiple Images Using SFM
(Mentor: Dr. Austin Reiter) October 2017 – December 2017

- Implemented algorithms in OpenCV for camera calibration, feature detection and matching.
- Implemented the SFM algorithm to reconstruct 3D scene from a sequence of views.

Collision-free Auto Pick-and-place Using UR5 Robot Arm
(Mentor: Dr. Noah J. Cowan) November 2017 – December 2017

- Implemented control algorithms in Matlab and ROS to drive UR5 robot arm in a desired trajectory.
- Implemented computer vision algorithms to drive UR5 to draw pictures based on their contour.

Mobile Device App Development for Systems Engineering Learning
(Mentor: Dr. Shuning Li) September 2016 – March 2017

- Developed an interactive iOS App using Swift 3.0 to introduce Systems Engineering to high school students.
- Submitted a paper abstract, *Developing a Mobile App to Introduce Systems Engineering to High School Students*, to American Society for Engineering Education (ASEE) 2017 conference.