



Published in final edited form as:

Int J Comput Assist Radiol Surg. 2019 June ; 14(6): 913–922. doi:10.1007/s11548-019-01943-z.

Interactive Flying Frustums (IFFs): spatially aware surgical data visualization

Javad Fotouhi^{1,2}, Mathias Unberath^{1,2}, Tianyu Song¹, Wenhao Gu¹, Alex Johnson³, Greg Osgood³, Mehran Armand^{3,4,5}, Nassir Navab^{1,2,6}

¹Computer Aided Medical Procedures, Johns Hopkins University, Baltimore, MD, USA

²Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA

³Department of Orthopaedic Surgery, Johns Hopkins Hospital, Baltimore, MD, USA

⁴Johns Hopkins University Applied Physics Laboratory, Laurel, MD, USA

⁵Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD, USA

⁶Computer Aided Medical Procedures, Technische Universität München, Munich, Germany

Abstract

Purpose: As the trend toward minimally invasive and percutaneous interventions continues, the importance of appropriate surgical data visualization becomes more evident. Ineffective interventional data display techniques that 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: The root-mean-squared error of C-arm source tracking after hand–eye calibration was determined as $0.43^\circ \pm 0.34^\circ$ and 4.6 ± 2.7 mm in rotation and translation, respectively. Finally, we

B Javad Fotouhi, javad.fotouhi@jhu.edu.

Javad Fotouhi, Mathias Unberath and Tianyu Song are Joint first authors.

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

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors

Informed consent This article does not contain patient data.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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.

Keywords

Augmented reality; Frustum; Fluoroscopy; Surgical data visualization

Introduction

C-arm fluoroscopy is extensively used to guide minimally invasive surgery in a variety of clinical disciplines including neuroradiology, orthopedics, and trauma [21,24,25,35]. 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 [33]. 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 [32]. 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 [7,34].

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.

Related work

First attempts at benefiting the ergonomics of surgery by improving display position placed miniature LCD displays close to the intervention site [8] and later displayed images relative to the surgeon’s field of vision using Google Glass [10,42]. More recently, Qian et al. [28] and Deib et al. [11] have 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 [10,41,42] that can potentially occlude the surgical site. In world-anchored mode, the virtual monitor is static in relation to the OR environment [9]. 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 [11,28], 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 [8,19], 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 [22,23,36]. 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 [1], the aforementioned drawbacks limit clinical acceptance [13, 18,22]. In contrast to explicit navigation and robotic solutions [30,31], scanner-[13,18,39] or user-centric [5,27,40] 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 [14,39] or as 3D virtual content in AR environments [5,18,40]. 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 multimodal interventional images such as the EchoNavigator system (Philips Inc., Amsterdam, Netherlands) where the outline of the 3D ultrasound volumes is augmented onto the fluoroscopy images to provide an intuitive geometric mapping between multiple images [16].

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 [13,39], i.e., an RGB-D camera rigidly attached to the detector of a mobile C-arm, Fotouhi et al. [15] track the position of the C-arm relative to the patient using image-based SLAM. 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 multimodal fiducial markers that can be sensed simultaneously by all involved devices, i.e., the C-arm X-ray system and the OST HMD [2]. 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 theaters. Then, visuo-motor coordination and “mental mapping” are 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 [2] 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.

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) aim 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 [17] combined with improved dynamic inside-out calibration of the C-arm to the AR environment [18] to develop spatially aware visualization. The proposed system, illustrated in Fig. 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.

Methodology

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 [20] 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 [20], then, any image acquired in this fixed C-arm pose can be translated along the camera's z -axis, i.e., along the frustum, while remaining a valid image of the same 3D scene [17]. In transmission imaging, this property of the frustum is convenient because the near and far planes of the frustum can always be held constant at $z = 0$ and $z = D_{SD}$, where D_{SD} is the source-to-

detector distance. In other words, there is no need for adaptive view frustum culling [3] 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.

System calibration

In order to realize the AR visualization of X-ray images in a spatially aware manner as described in “Uniting patient anatomy and X-ray image using the view frustum” section, 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 [18]. 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:

$${}^S\mathbf{T}_C(t) = {}^S\mathbf{T}_{OR}(t) \underbrace{\left(\mathbf{T}_{OR}^{-1}(t) \mathbf{T}_C(t_0) \right)}_{OR\mathbf{T}_C(t)}, \quad (1)$$

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. 1, t_0 describes the time of offline calibration. Upon acquisition of X-ray image I_j at time t_j , ${}^{OR}\mathbf{T}_C(t_j)$ 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 ${}^S\mathbf{T}_C(t)$ are explained in the remainder of this section and visualized in Fig. 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 ${}^{OR}\mathbf{T}_{S/T}$ therein [12]. Using the surgeon as example, SLAM solves:

$${}^{OR}\hat{\mathbf{T}}_S(t) = \arg \min_{{}^{OR}\hat{\mathbf{T}}_S} d\left(\mathbf{f}_{OR}\left(\mathbf{P}^{OR}\hat{\mathbf{T}}_S(t)\mathbf{x}_S(t)\right), \mathbf{f}_S(t)\right), \quad (2)$$

where $\mathbf{f}_S(t)$ are features extracted from the image at time t , $\mathbf{x}_S(t)$ are the 3D locations of these feature obtained, e.g., via multi-view stereo, \mathbf{P} is the projection operator, and $d(\cdot, \cdot)$ is the similarity to be optimized. Following [18], 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.

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 ${}^T\mathbf{T}_C(t_0)$ via hand-eye calibration, i.e., the relative pose information from the rigidly connected tracker and the C-arm is used for solving $\mathbf{X} := {}^T\mathbf{T}_C(t_0)$ in $\mathbf{AX} = \mathbf{XB}$ fashion [38]. 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 is 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. 3a). The spatial link between the gantry and the source is then estimated via the following equation:

$${}^C\mathbf{T}_G = {}^M\mathbf{T}_C^{-1} {}^M\mathbf{T}_{IR} {}^G\mathbf{T}_{IR}^{-1}, \quad (3)$$

where ${}^M\mathbf{T}_C$ 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 ${}^M\mathbf{T}_C$ is estimated, marker \mathbf{M} is removed and the C-arm pose is estimated in the frame of the external infrared navigation system ${}^C\mathbf{T}_{IR} = {}^C\mathbf{T}_G {}^G\mathbf{T}_{IR}$. To solve the calibration problem in a hand-eye configuration, we construct the following chain of transformations:

$$\begin{aligned} {}^T\mathbf{T}_{OR}^{-1}(t_i) {}^T\mathbf{T}_C(t_0) {}^C\mathbf{T}_{IR}(t_i) &= {}^T\mathbf{T}_{OR}^{-1}(t_{i+1}) {}^T\mathbf{T}_C(t_0) {}^C\mathbf{T}_{IR}(t_{i+1}), \\ \underbrace{{}^T\mathbf{T}_{OR}^{-1}(t_{i+1})}^{\mathbf{A}} \underbrace{{}^T\mathbf{T}_{OR}^{-1}(t_i) {}^T\mathbf{T}_C(t_0)}^{\mathbf{X}} &= \underbrace{{}^T\mathbf{T}_C(t_0) {}^C\mathbf{T}_{IR}(t_{i+1})}^{\mathbf{X}} \underbrace{{}^C\mathbf{T}_{IR}^{-1}(t_i)}^{\mathbf{B}}. \end{aligned} \quad (4)$$

Equation 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 :

$$Q_a Q_x = Q_x Q_b. \quad (5)$$

By re-arranging Eq. 5 in the form of $\mathbf{M}Q_x = 0$, we solve for rotation in the following constrained optimization:

$$\min \| \mathbf{M}Q_x \|_2^2, \quad \text{s.t. } \| Q_x \|_2^2 = 1. \quad (6)$$

Finally, the translation component p_x is estimated in a least-squares fashion as expressed in Eq. 7, where R denotes rotation matrix:

$$\begin{aligned} R_a p_x + p_a &\approx R_x p_b + p_x, \\ (R_a - \mathbb{1}) p_x &\approx R_x p_b - p_a. \end{aligned} \quad (7)$$

Generating the frustum \mathbf{K}

The view frustum of the C-arm is modeled via the standard 11 DoF camera parameters. In “System calibration” section, 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.

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 green as the image approaches the detector. 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.

Experimental results and evaluation

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 contents was built using the Unity game engine.

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 ${}^T\mathbf{T}_C(t_0)$ between the tracker and the X-ray source, a Polaris Spectra external optical navigation system (Northern Digital, Waterloo, ON) was employed.

Analysis of hand–eye calibration

Offline estimation of the relation between the passive markers \mathbf{G} and the X-ray source constitutes the first and critical step in closing the transformation chain. To this end, we estimated ${}^C\mathbf{T}_G$ via Eq. 3 by acquiring pose information from 7 different poses of marker \mathbf{M} . The translation component of ${}^C\mathbf{T}_G$ was estimated by averaging the translations of each individual measurement in Euclidean space, and the mean rotation was estimated such that the properties of the rotation group $SO(3)$ were preserved.

Next, the hand–eye calibration problem (Eq. 4) was solved by simultaneously acquiring $N=160$ 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 1 were calculated separately for translation and rotation using:

$$p_e = \frac{\sum_{i=1}^N p(\mathbf{A}_i \mathbf{X} - \mathbf{X} \mathbf{B}_i)}{N}, \text{ and} \quad (8)$$

$$R_e = (R_a R_x)^{-1} R_x R_b.$$

C-arm pose estimation via integrated visual tracking

In Table 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. 3. The evaluation was performed over 20 different C-arm angulations.

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 4 radiopaque fiducial markers placed at different heights. We 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 was asked to align the gaze cursor with the fiducial and confirm the selection using the air-tap gesture. The intersection of the gaze cursor ray with the 3D map of the environment created by SLAM was then used as the real 3D position of the fiducial. Finally, the Euclidean distance between the virtual and real 3D locations of a fiducial was measured as TAE. The TAE, averaged over 20 different trials and the 4 fiducials on the phantom, was 13.2 ± 2.89 mm.

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 of pelvic ring fractures [29,37] and (2) percutaneous spine procedures such as percutaneous vertebroplasty [6]. We show exemplary scenes of the aforementioned cases in Fig. 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 theater 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 preoperative 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 [40]. 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, and 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 “Experimental results and evaluation” section. Since IFFs transformation parameters are estimated globally within the operating theater 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 theater. 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 “Experimental results and evaluation” section, 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 1) are similar to the errors observed during tracking (Table 2) suggesting that the remaining error is statistic and the data used for offline calibration were 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 where 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 [14]. We believe that IFFs paradigm is a step toward removing ambiguities present in the projective images acquired intra-operatively. For surgical interventions that require full 3D information, either preoperative CT images need to be registered to the patient, or 3D intra-operative imaging would be employed [4].

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 [26].

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.

Acknowledgements

The authors want to thank Gerhard Kleinzig and Sebastian Vogt from Siemens Healthineers for their support and making a Siemens ARCADIS Orbic 3-D available.

Funding Research in this work was supported in part by the Graduate Student Fellowship from Johns Hopkins Applied Physics Laboratory and NIH R21 EB020113.

Research in this work was supported in part by the Graduate Student Fellowship from Johns Hopkins Applied Physics Laboratory, NIH R01 EB023939, and Johns Hopkins University internal funding sources.

References

1. Alambeigi F, Wang Y, Sefati S, Gao C, Murphy RJ, Iordachita I, Taylor RH, Khanuja H, Armand M (2017) A curved-drilling approach in core decompression of the femoral head osteonecrosis using a continuum manipulator. *IEEE Robot Autom Lett* 2(3):1480–1487
2. Andress S, Johnson A, Unberath M, Winkler AF, Yu K, Fotouhi J, Weidert S, Osgood G, Navab N (2018) On-the-fly augmented reality for orthopedic surgery using a multimodal fiducial. *J Med Imaging* 5(2):021209

3. Assarsson U, Moller T (2000) Optimized view frustum culling algorithms for bounding boxes. *J Graph Tools* 5(1):9–22
4. Atria C, Last L, Packard N, Noo F (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
5. Augmedics: Augmedics xvision: Pre-clinical Cadaver Study (2018). <https://www.augmedics.com/cadaver-study>. Accessed 24 Oct 2018
6. Barr JD, Barr MS, Lemley TJ, McCann RM (2000) Percutaneous vertebroplasty for pain relief and spinal stabilization. *Spine* 25(8):923–928 [PubMed: 10767803]
7. Boszczyk BM, Bierschneider M, Panzer S, Panzer W, Harstall R, Schmid K, Jaksche H (2006) Fluoroscopic radiation exposure of the kyphoplasty patient. *Eur Spine J* 15(3):347–355 [PubMed: 15947995]
8. Cardin MA, Wang JX, Plewes DB (2005) A method to evaluate human spatial coordination interfaces for computer-assisted surgery. *Med Image Comput Comput Assist Interv MICCAI* 2005:9–16
9. Chen X, Xu L, Wang Y, Wang H, Wang F, Zeng X, Wang Q, Egger J (2015) Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display. *J Biomed Inform* 55:124–131 [PubMed: 25882923]
10. Chimenti PC, Mitten DJ (2015) Google glass as an alternative to standard fluoroscopic visualization for percutaneous fixation of hand fractures: a pilot study. *Plast Reconstr Surg* 136(2):328–330 [PubMed: 26218380]
11. Deib G, Johnson A, Unberath M, Yu K, Andress S, Qian L, Osgood G, Navab N, Hui F, Gailloud P (2018) Image guided percutaneous spine procedures using an optical see-through head mounted display: proof of concept and rationale. *J Neurointerv Surg* 10:1187–1191 [PubMed: 29848559]
12. Endres F, Hess J, Engelhard N, Sturm J, Cremers D, Burgard W (2012) An evaluation of the RGB-D SLAM system. In: *2012 IEEE international conference on robotics and automation (ICRA)* IEEE, pp 1691–1696
13. Fischer M, Fuerst B, Lee SC, Fotouhi J, Habert S, Weidert S, Euler E, Osgood G, Navab N (2016) Preclinical usability study of multiple augmented reality concepts for K-wire placement. *Int J Comput Assist Radiol Surg* 11(6):1007–1014 [PubMed: 26995603]
14. Fotouhi J, Alexander CP, Unberath M, Taylor G, Lee SC, Fuerst B, Johnson A, Osgood GM, Taylor RH, Khanuja H, Armand M, Navab N (2018) Plan in 2-d, execute in 3-d: an augmented reality solution for cup placement in total hip arthroplasty. *J Med Imaging* 5(2):021205
15. Fotouhi J, Fuerst B, Johnson A, Lee SC, Taylor R, Osgood G, Navab N, Armand M (2017) Pose-aware C-arm for automatic re-initialization of interventional 2d/3d image registration. *Int J Comput Assist Radiol Surg* 12(7):1221–1230 [PubMed: 28527025]
16. Gafoor S, Schulz P, Heuer L, Matic P, Franke J, Bertog S, Reinartz M, Vaskelyte L, Hofmann I, Sievert H (2015) Use of echonavi-gator, a novel echocardiography–fluoroscopy overlay system, for transseptal puncture and left atrial appendage occlusion. *J Interv Cardiol* 28(2):215–217 [PubMed: 25676602]
17. Georgel PF, Schroeder P, Navab N (2009) Navigation tools for viewing augmented cad models. *IEEE Comput Graph Appl* 29(6):65–73 [PubMed: 24806780]
18. Hajek J, Unberath M, Fotouhi J, Bier B, Lee SC, Osgood G, Maier A, Armand M, Navab N (2018) Closing the calibration loop: an inside-out-tracking paradigm for augmented reality in orthopedic surgery. *arXiv preprint arXiv:1803.08610*
19. Hanna GB, Shimi SM, Cuschieri A (1998) Task performance in endoscopic surgery is influenced by location of the image display. *Ann Surg* 227(4):481 [PubMed: 9563533]
20. Hartley R, Zisserman A (2003) *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge
21. Hott JS, Deshmukh VR, Klopfenstein JD, Sonntag VK, Dickman CA, Spetzler RF, Papadopoulos SM (2004) Intraoperative Iso-C C-arm navigation in craniospinal surgery: the first 60 cases. *Neurosurgery* 54(5):1131–1137 [PubMed: 15113467]
22. Joskowicz L, Hazan EJ (2016) Computer aided orthopaedic surgery: incremental shift or paradigm change? *Med Image Anal* 33:84–90 [PubMed: 27407004]

23. Liebergall M, Mosheiff R, Joskowicz L (2006) Computer-aided orthopaedic surgery in skeletal trauma In: Bucholz RW, Heckman JD, Court-Brown CM (eds) Rockwood and greens fractures in adults, 6th edn Lippincott, Philadelphia, pp 739–767
24. Mason A, Paulsen R, Babuska JM, Rajpal S, Burneikiene S, Nelson EL, Villavicencio AT (2014) The accuracy of pedicle screw placement using intraoperative image guidance systems: a systematic review. *J Neurosurg Spine* 20(2):196–203 [PubMed: 24358998]
25. Miller DL, Vañó E, Bartal G, Balter S, Dixon R, Padovani R, Schueler B, Cardella JF, De Baère T (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. *Cardiovasc Intervent Radiol* 33(2):230–239 [PubMed: 20020300]
26. Navab N, Mitschke M (2001) Method and apparatus using a virtual detector for three-dimensional reconstruction from x-ray images. US Patent 6,236,704
27. Qian L, Deguet A, Kazanzides P (2018) Arssist: augmented reality on a head-mounted display for the first assistant in robotic surgery. *Healthc Technol Lett* 5:194–200 [PubMed: 30800322]
28. Qian L, Unberath M, Yu K, Fuerst B, Johnson A, Navab N, Osgood G (2017) Towards virtual monitors for image guided interventions-real-time streaming to optical see-through head-mounted displays. arXiv preprint arXiv:1710.00808
29. Routt MC Jr, Nork SE, Mills WJ (2000) Percutaneous fixation of pelvic ring disruptions. *Clin Orthop Relat Res* (1976–2007) 375:15–29
30. Sefati S, Alambeigi F, Iordachita I, Armand M, Murphy RJ (2016) FBG-based large deflection shape sensing of a continuum manipulator: manufacturing optimization In: 2016 IEEE SENSORS. IEEE, pp 1–3
31. Sefati S, Pozin M, Alambeigi F, Iordachita I, Taylor RH, Armand M (2017) A highly sensitive fiber Bragg Grating shape sensor for continuum manipulators with large deflections In: 2017 IEEE SENSORS. IEEE, pp 1–3
32. Starr A, Jones A, Reinert C, Borer D (2001) Preliminary results and complications following limited open reduction and percutaneous screw fixation of displaced fractures of the acetabulum. *Injury* 32:SA45–50 [PubMed: 11521706]
33. Suzuki T, Soma K, Shindo M, Minehara H, Itoman M (2008) Anatomic study for pubic medullary screw insertion. *J Orthop Surg* 16(3):321–325
34. Synowitz M, Kiwit J (2006) Surgeons radiation exposure during percutaneous vertebroplasty. *J Neurosurg Spine* 4(2):106–109 [PubMed: 16506476]
35. Theocharopoulos N, Perisinakis K, Damilakis J, Papadokostakis G, Hadjipavlou A, Gourtsoyiannis N (2003) Occupational exposure from common fluoroscopic projections used in orthopaedic surgery. *JBJS* 85(9):1698–1703
36. Theologis A, Burch S, Pekmezci M (2016) Placement of iliosacral screws using 3d image-guided (O-arm) technology and stealth navigation: comparison with traditional fluoroscopy. *Bone Joint J* 98(5):696–702 [PubMed: 27143744]
37. Tile M (1988) Pelvic ring fractures: should they be fixed? *J Bone Joint Surg Br* 70(1):1–12 [PubMed: 3276697]
38. Tsai RY, Lenz RK (1989) A new technique for fully autonomous and efficient 3d robotics hand/eye calibration. *IEEE Trans Robot Autom* 5(3):345–358
39. Tucker E, Fotouhi J, Unberath M, Lee SC, Fuerst B, Johnson A, Armand M, Osgood GM, Navab N (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, p 105790J
40. Unberath M, Fotouhi J, Hajek J, Maier A, Osgood G, Taylor R, Armand M, Navab N (2018) Augmented reality-based feedback for technician-in-the-loop C-arm repositioning. *Healthc Technol Lett* 5:143–147 [PubMed: 30464844]
41. Vorraber W, Voessner S, Stark G, Neubacher D, DeMello S, Bair A (2014) Medical applications of near-eye display devices: an exploratory study. *Int J Surg* 12(12):1266–1272 [PubMed: 25462704]
42. Yoon JW, Chen RE, Han PK, Si P, Freeman WD, Pirris SM (2016) Technical feasibility and safety of an intraoperative head-up display device during spine instrumentation. *Int J Med Robot Comput Assist Surg* 13:e1770

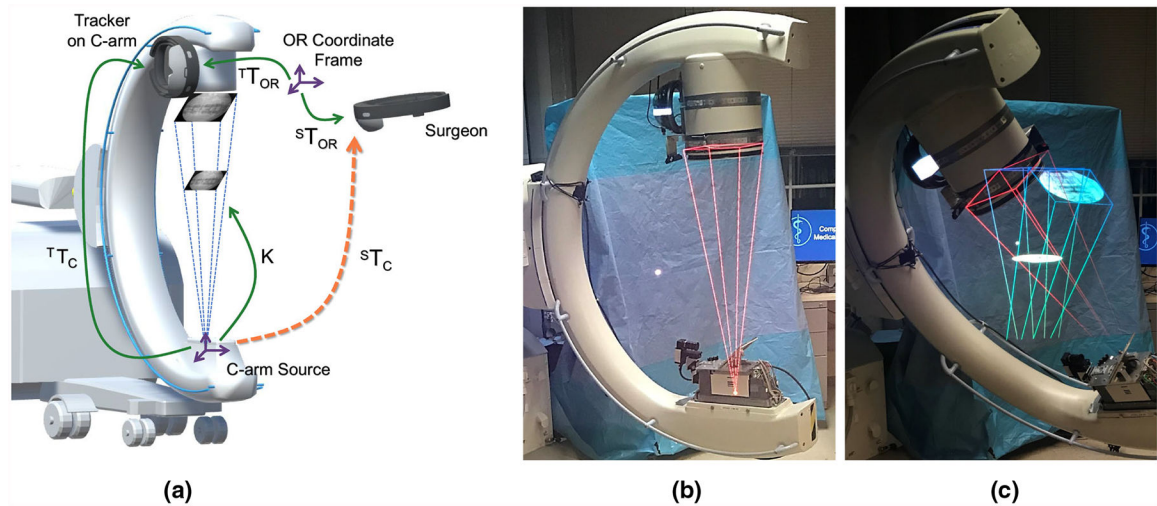


Fig. 1.
a Schematic illustration of the proposed spatially aware image visualization of X-ray images on their view frustum. In addition, we show transformation to be estimated dynamically to enable the proposed AR environment. Transformations shown as green arrows are estimated directly while transformations 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

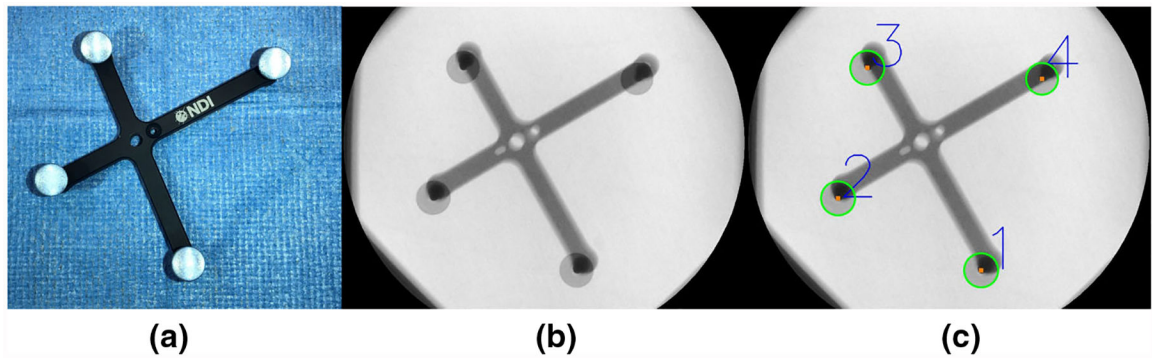


Fig. 2.

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 “System calibration” section

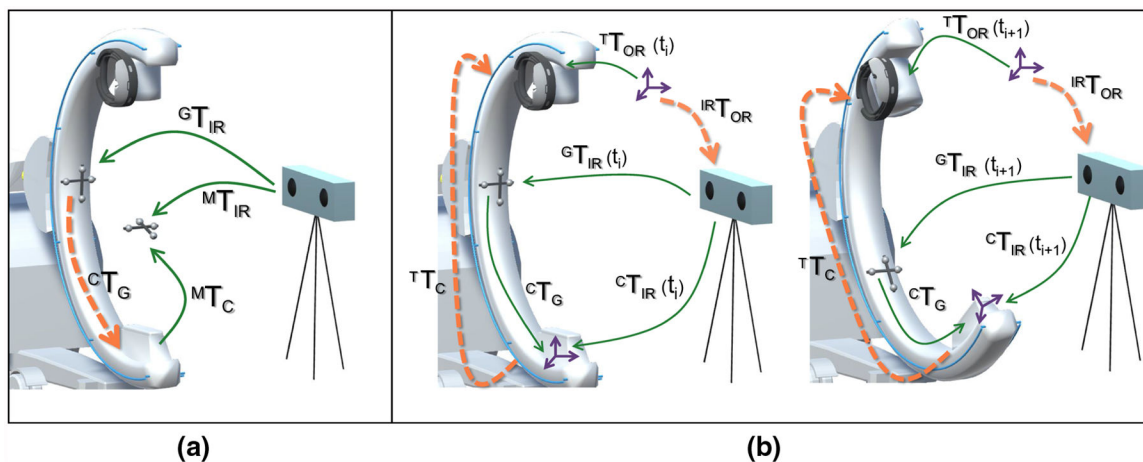


Fig. 3. Illustrations describing the process of calibrating the tracker to the C-arm X-ray source using hand-eye calibration and an external optical navigation system. **a** An infrared reflective marker is attached to the gantry and calibrated to the X-ray source using a second marker that is imaged by the navigation system and the C-arm simultaneously (Fig. 2). **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. [38]

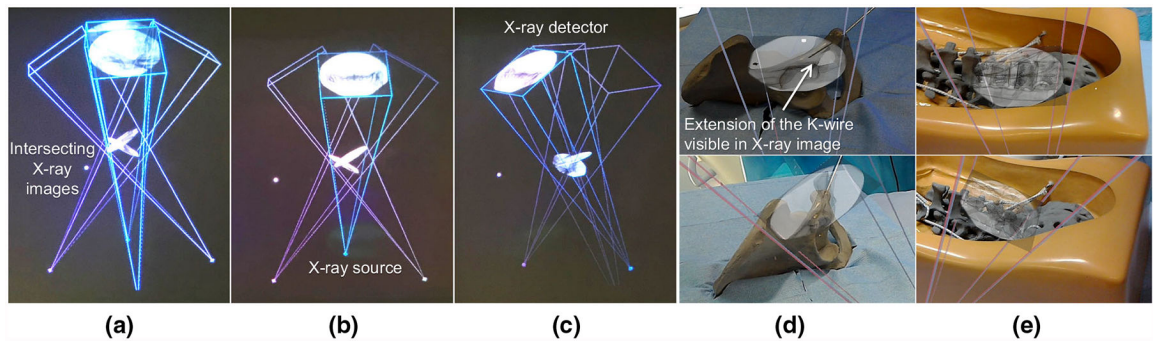


Fig. 4. Multiple views of IFFs are shown in **a–c**. **d, e** show the augmentation of the virtual view frustum and the corresponding C-arm images from two views on a pelvis and a spine phantom

Table 1

Error measures for tracker to C-arm hand-eye calibration of $T_C(t_0)$

Hand-eye calibration	Residual (x, y, z), \cdot ₂	RMS	Median (x, y, z)	SD (x, y, z), \cdot ₂
Rotation (°)	(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

Table 2

Error measures for C-arm extrinsic parameter estimation using SLAM tracking

C-arm tracking	Residual $(x, y, z), \cdot _2$	RMS	Median (x, y, z)	SD $(x, y, z), \cdot _2$
Rotation ($^{\circ}$)	(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