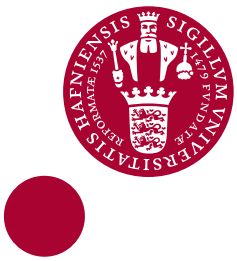

Mobile augmented reality for computer-assisted percutaneous nephrolithotomy

Müller, M.; Rassweiler, M.-C.; Klein, J.; Seitel, A.; Gondan, Matthias; Baumhauer, M.; Teber, D.; Rassweiler, J.J.; Meinzer, H.-P.; Maier-Hein, L.

Published in:
International Journal of Computer Assisted Radiology and Surgery

DOI:
10.1007/s11548-013-0828-4

Publication date:
2013

Document version
Publisher's PDF, also known as Version of record

Int J CARS (2013) 8:663–675
DOI 10.1007/s11548-013-0828-4

ORIGINAL ARTICLE

Mobile augmented reality for computer-assisted percutaneous nephrolithotomy

Michael Müller · Marie-Claire Rassweiler · Jan Klein · Alexander Seitel · Matthias Gondan · Matthias Baumhauer · Dogu Teber · Jens J. Rassweiler · Hans-Peter Meinzer · Lena Maier-Hein

Received: 12 December 2012 / Accepted: 4 March 2013 / Published online: 23 March 2013

Abstract
Purpose Percutaneous nephrolithotomy (PCNL) plays an integral role in treatment of renal stones. Creating percutaneous renal access is the most important and challenging step in the procedure. To facilitate this step, we evaluated our novel mobile augmented reality (AR) system for its feasibility of use for PCNL.
Methods A tablet computer, such as an iPad®, is positioned above the patient with its camera pointing toward the field of intervention. The images of the tablet camera are registered with the CT image by means of fiducial markers. Structures of interest can be superimposed semi-transparently on the video images. We present a systematic evaluation by means of a phantom study. An urological trainee and two experts conducted 53 punctures on kidney phantoms.
Results The trainee performed best with the proposed AR system in terms of puncturing time (mean: 99 s), whereas the experts performed best with fluoroscopy (mean: 59 s). iPad assistance lowered radiation exposure by a factor of 3 for the inexperienced physician and by a factor of 1.8 for the experts in comparison with fluoroscopy usage. We achieve a mean visualization accuracy of 2.5 mm.
Conclusions The proposed tablet computer-based AR system has proven helpful in assisting percutaneous interventions such as PCNL and shows benefits compared to other state-of-the-art assistance systems. A drawback of the system in its current state is the lack of depth information. Despite that, the simple integration into the clinical workflow highlights the potential impact of this approach to such interventions.

Keywords Augmented reality · Mobile · Image-guided surgery · Computer vision · CT

Electronic supplementary material The online version of this article (doi:10.1007/s11548-013-0828-4) contains supplementary material, which is available to authorized users.

M. Müller (✉) · A. Seitel · M. Baumhauer · H.-P. Meinzer · L. Maier-Hein
Department of Medical and Biological Informatics, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
e-mail: michael.mueller@dkfz-heidelberg.de

M.-C. Rassweiler
Department of Urology, Mannheim University Hospital, Theodor-Kutzer-Ufer 1-3, 68167 Mannheim, Germany

J. Klein · J. J. Rassweiler
Department of Urology, SLK-Kliniken Heilbronn, Am Gesundbrunnen 20, 74078 Heilbronn, Germany

M. Gondan
Institute of Medical Biometry and Informatics, University of Heidelberg, Im Neuenheimer Feld 305, 69120 Heidelberg, Germany

D. Teber
Department of Urology, Heidelberg University Hospital, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany

Introduction

Renal stones are a painful disease with a globally increasing incidence independent of sex, race and age [1]. Percutaneous nephrolithotomy (PCNL) plays an integral role in treatment of renal stones. In fact, it is considered the state-of-the-art treatment for large or multiple kidney stones [2]. Due to advances in technical equipment, PCNL reaches stone-free rates of over 90 % while having low complication rates [3]. As a type of minimally invasive procedure, PCNL does not require large incisions in the skin or the kidney, thus lowering the risk of infections and the post-operative stay.

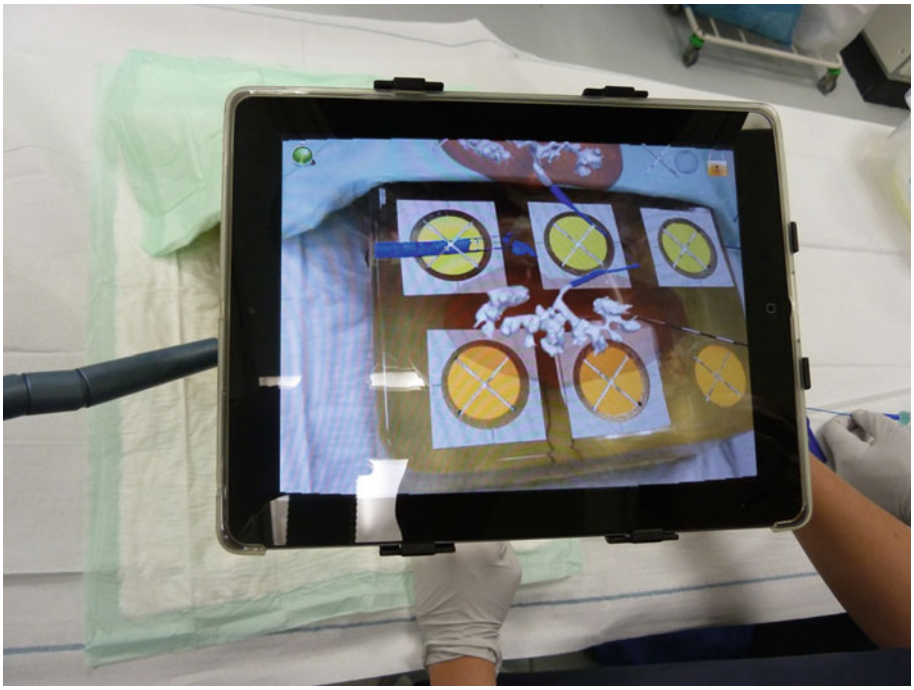

Fig. 1 The AR view of a kidney phantom on the iPad. The collecting system is shown in white, the kidneys in semi-transparent red

Creating percutaneous renal access, i.e., inserting a needle through the skin into the renal collecting system with the patient in a prone position, is the most important step in the procedure and requires a thorough understanding of renal, retroperitoneal and thoracic anatomy so as to minimize the risk of complications. PCNL is typically done with support of imaging devices such as fluoroscopy or ultrasound or a combination of both. While it is still a matter of discussion which modality to choose [4], both modalities have several limitations. Ultrasound as well as fluoroscopy only delivers 2D images for a very small window. In particular, for complex cases, the images provided by these modalities are hard to interpret spatially which results in prolonged puncturing times due to multiple reorientation or reinsertion of the needle. Moreover, it is known that many complications arise from injury of surrounding tissue or structures (colon, spleen, liver, pleura, lung) which are hardly recognizable in the 2D images [3]. Hence, ultrasound and fluoroscopy cannot provide optimal guidance. Besides that, using fluoroscopic imaging, the patient is exposed to a certain amount of radiation depending on the time it takes to access the collecting system.

To facilitate percutaneous needle insertion, numerous computer-assisted approaches, for example, based on external markers [5], internal markers [6], calibrated visualization devices [7], robotic devices [8] and multi-modal image registration [9], have been proposed in the literature (cf. Wood et al. [10] for a detailed review). Specifically, for PCNL, three navigation approaches should be considered: Lazarus et al. [11] present a novel system for PCNL which consists of a mechanical device fixed to the operating table to stabilize the needle during the puncture. The authors report puncturing times of 225 s without and 118 s with the respective device in an in vitro study. Huber et al. utilize electromagnetic tracking [12] to assist PCNL. A small sensor is placed via an ureteral catheter, and a tracked needle is guided to the desired puncture site in a "rendezvous" approach. They recorded a mean puncturing time of 14 s in an ex vivo porcine model after a learning phase of 30 punctures. Very recently, a laser guidance system which belongs to the Uro DynaCT device was evaluated by Ritter et al. [13]. For 10 punctures in a biological model, they report a median time of 4.6 (2–10.2) min. However, thus far, none of these navigation concepts have been widely accepted in clinical routine, because the benefit for the patient did not exceed the additional complexity and higher costs introduced by the required equipment.

To address this issue, we developed a new concept for computer-assisted percutaneous needle insertion that can be integrated into the clinical workflow in a straightforward manner. It combines a previously proposed method for camera pose estimation [14] with a recently introduced concept for mobile augmented reality (AR) [15] to provide an AR view of internal anatomical structures on a mobile display, such as an Apple iPad® positioned above the patient (Fig. 1). In an initial study, we demonstrated the clinical feasibility of the proposed system [16] for PCNL. In this paper, however, we aim to systematically assess the benefit of the proposed system in contrast to the established clinical standards using fluoroscopy or ultrasound for guidance during PCNL. We therefore measured the time

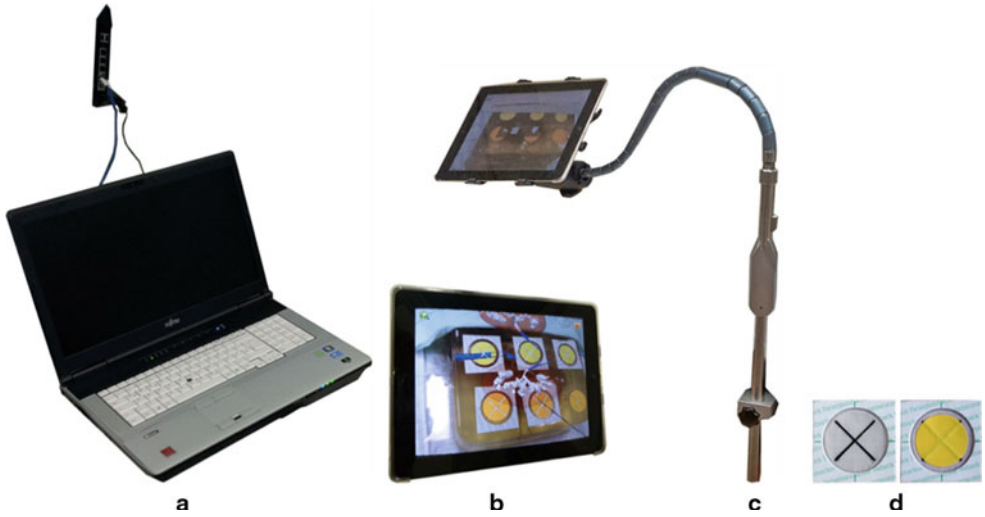

Fig. 2 The components of the navigation system: **a** the navigation server, **b** Apple iPad as mobile display, **c** tablet fixation on a flexible endoscope stand, **d** radio-dense navigation markers, optionally attached with color labels for a more robust localization

and radiation exposure for puncturing the collecting system of a kidney in a specifically designed phantom using the proposed mobile navigation system, fluoroscopy and ultrasound, respectively. Additionally, we examined the accuracy of the augmented reality visualization using different registration algorithms.

Material

System components

All components of the presented navigation system are depicted in Fig. 2. We employed a standard Apple iPad 2 for the mobile visualization. Basically, any other tablet PC equipped with a camera and a wireless network adapter could be used. In the current system design, the tablet is used as a thin-client for showing and recording images, as well as submitting application commands, e.g., for switching between different visualization modes, whereas the processing of the images and the logic of the application run on a server PC. In preliminary trials, the system was implemented as a standalone application on the iPad but this approach was discarded due to limited hardware capabilities of the tablet. On the server side, we use a notebook with adequate equipment for real-time calculations and rendering (Fujitsu TS CELSIUS H910). The server and the tablet computer are connected via a standard Wi-Fi access point. For affixing the tablet during puncturing, we use a commercially available iPad holder[1] mounted on a flexible endoscope stand.[2] Since the stand has a standardized diameter, it can easily be attached to a treatment couch without requiring any alterations.

Low-price, radio-dense skin markers originating from radiotherapy[3] are employed as fiducial markers. These markers are fixed on the patient's skin (respectively the phantom's surface) before imaging and can easily be detected in CT or X-ray images. To improve the robustness of the fiducial segmentation (cf. section *Image Streaming and Processing*), the radio-dense markers are equipped with colored, adhesive labels. The diameter of the fiducial markers is 32 mm.

Phantom design

For evaluation purposes, a phantom model serves as a substitute for the human kidney and its surrounding soft tissue. Therefore, the phantom consists of two porcine kidneys that are covered by a block of ballistic gelatin, which has a consistency similar to human tissue (Fig. 3a, b). The kidneys are catheterized in order to inject contrast agent during the experiments and to be able to verify the needle position. To be able to distinguish the kidney tissue from the surrounding ballistic gelatin during the segmentation of the structures in the planning CT images, a high concentration of contrast agent is inserted into the mixture (cf. Fig. 3c). As shown in Fig. 3b, d, the navigation markers are attached on the surface of the gelatin block and can be easily located in the CT image. Since the ballistic gelatin is slightly transparent and the inner structures can be recognized, the phantom is covered by a surgical drape with notches for the fiducial markers.

[1] iKlip stand adapter, IK Multimedia Production, http://www.ikmultimedia.com/products/iklip/.

[2] Braun Melsungen AG, http://www.bbraun.de/.

[3] Rebeck patient markers, Fobeck GbR, http://fobeck.com/cms2/en.

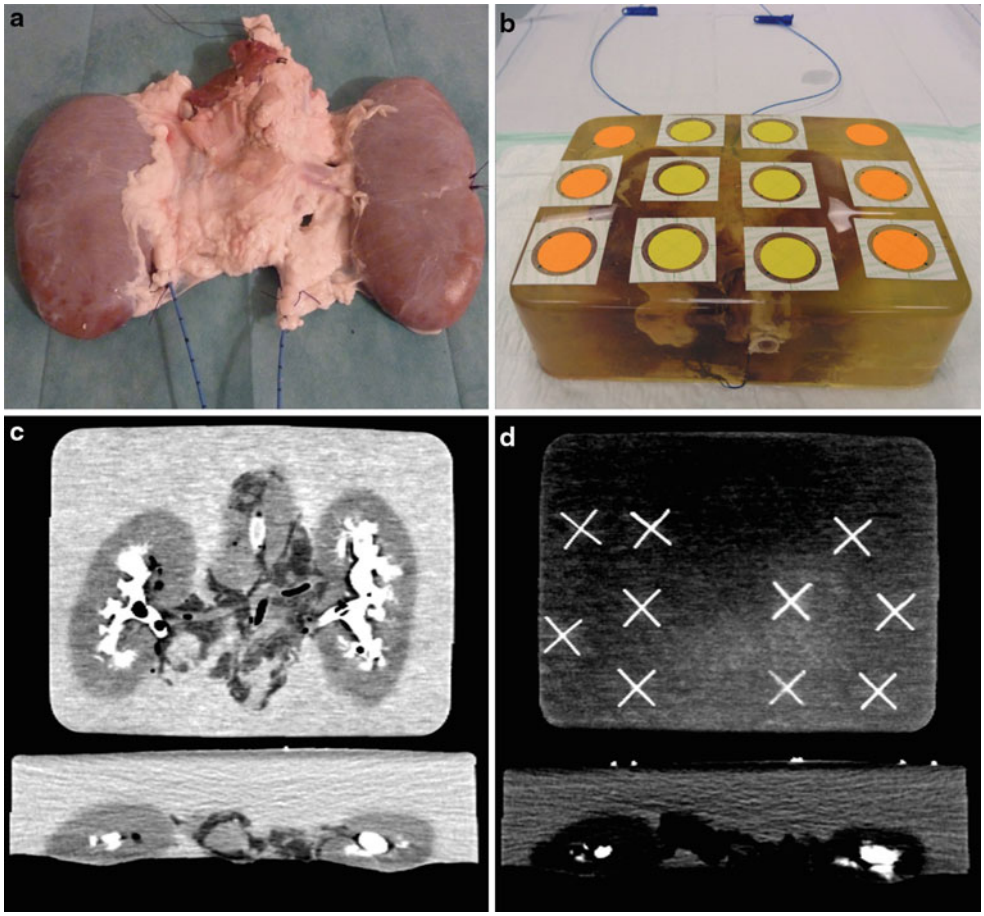

Fig. 3 Images of the constructed phantom: **a** the prepared porcine kidneys that were embedded into the gelatin mixture. **b** The gelatin body with the markers attached. **c** Coronal and transversal slices of the CT image showing the phantom's interior and **d** slices depicting the appearance of the fiducial markers in the CT image

Imaging modalities

Two types of imaging modalities were utilized throughout the experiments: a 3D imaging device and 2D imaging (fluoroscopy and ultrasound). The former was used preinterventionally to create a virtual model which is then superimposed on the video images by the mobile navigation system. The latter was used as in clinical routine during the intervention, i.e., either fluoroscopy alone or in combination with ultrasound. Both routinely applied variants were then compared to the proposed navigation system.

For a part of the experimental series, a Siemens syngo DynaCT device (see footnote 3) was available (Fig. 4). The DynaCT device is a rotatable C-Arm scanner for acquiring CT-like image volumes within 6–10 s. The dimensions of the volume that is acquired are about 250 mm × 250 mm × 194 mm. Therefore, it is perfectly suited for our phantom or the human kidney with its surrounding anatomy. The advantage is the possibility to derive low-dose 3D scans of the anatomy intraoperatively. To show the feasibility of our navigation system in combination with DynaCT imaging, we incorporated it in the experiments. On the one hand, the DynaCT served as a 3D scanner to derive volumetric images of the phantoms. On the other hand, we employed the DynaCT as a standard 2D fluoroscopic device during the experiments, which was then compared to the proposed system.

For the remaining part of the experimental series, a conventional Siemens CT scanner was used to create a 3D volume for the phantom right before the experiment. With a standard soft tissue setup, volumes with a slice thickness of 0.6 mm were created. In this setup, the trial was conducted in an intervention room with Siemens LITHOSKOP equipment[4] which contains a partly rotatable C-Arm device for intra-operative 2D X-ray imaging.

As for the comparison with ultrasound, a standard ultrasound unit (BK Medical) routinely used for PCNL was at hand for each of experiments. The ultrasound was used as

[4] For a detailed specification, please refer to the manufacturer's website http://www.siemens.com.

Fig. 4 The experimental setup using the Siemens syngo DynaCT device

in clinical routine, i.e., standalone and in combination with fluoroscopy.

Methods

The Methods part is split into the algorithmic part and the description of the experimental settings. The whole image streaming and processing pipeline is summarized in the paragraph *Augmented Reality Guidance*. The experiments consist firstly of the experimental series comparing the three respective guidance approaches for PCNL (tablet navigation, ultrasound, fluoroscopy) and secondly of an investigation of the visualization accuracy of the proposed navigation system. Both trials are compared in the last two paragraphs of the Methods section.

Augmented reality guidance

Image streaming and processing

To allow real-time processing of the image data, an important aspect of the presented mobile AR application is the rapid image streaming and processing pipeline. The software on the server side is implemented as part of the Medical Imaging and Interaction Toolkit (MITK [17]). MITK for embedded systems (MITK MES) which is an extension to MITK for mobile devices is installed on the tablet [18]. As depicted in Fig. 5, the navigation system relies on streaming images

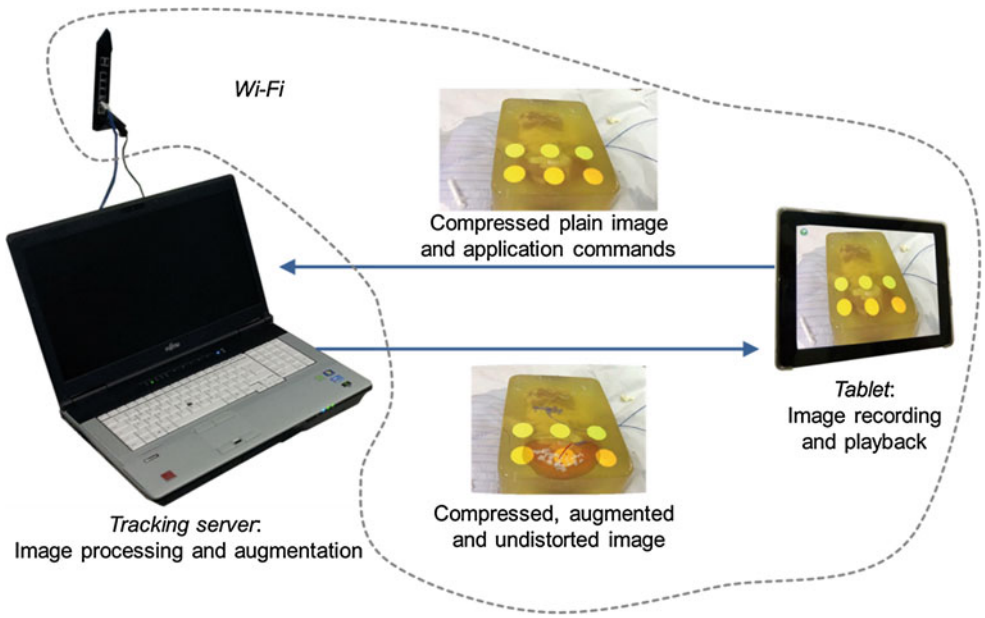

Fig. 5 Client–server communication: the tablet records images with its camera, compresses them and sends them to a server via Wi-Fi, possibly combined with further application commands which are created by user interaction with the tablet. The image processing pipeline is then executed on the server, resulting in an augmented image which is also compressed and sent back to the tablet where it is displayed

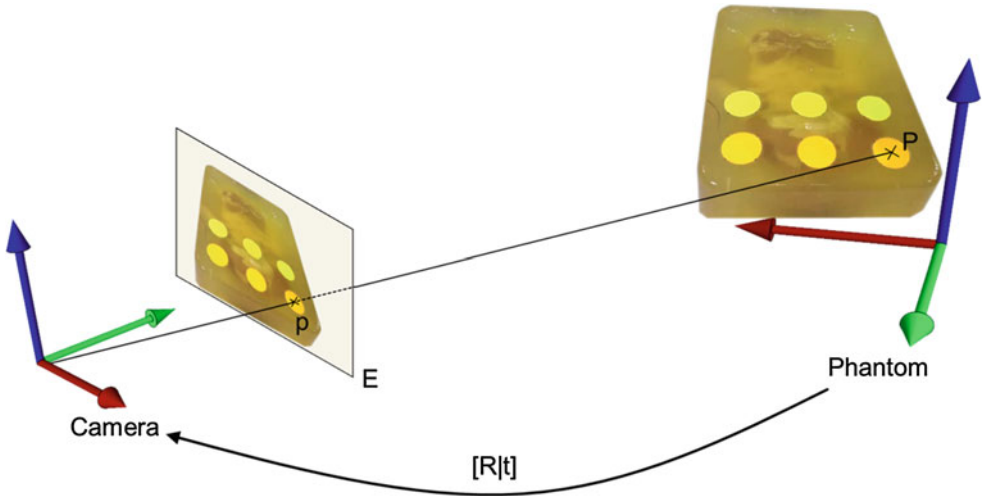

Fig. 6 Pose of a phantom with respect to a camera. The pose consists of a rotation matrix R and a translation vector t. The figure also shows a fiducial marker point P and the projected point p on an image plane E

and commands between the tablet and the server in real time. The images are recorded by the tablet camera, down sampled to a resolution of 640 × 360 pixels, compressed and sent to a server computer via Wi-Fi. The down sampling of the images is necessary to reach a sufficient frame-per-second rate. Besides the images, application commands can be sent over the network, e.g., the user can toggle between different visualization modes by interactive buttons on the tablet screen. The user may toggle between volume rendering of the image with different transfer function settings or a visualization of surface models.

After decompression on the server side, the image is analyzed to find the center points of the colored markers. The algorithm to perform this task is based on previous work [14,19] and can be summarized as follows: the RGB image is converted into HSV color space and split into its single channels. In each channel, a user-defined upper and lower threshold is applied. The binary images are then cleaned from noise by means of morphological operations. The subsequently extracted contours are filtered based on user defined or computed criteria, e.g., size, morphology or predicted locations (cf. [19]). The center point is then determined for the most probable candidates by means of a minimum enclosing circle algorithm. These center points are considered the correspondences of the center of the crosses which are visible in the CT image (Fig. 3d). These point correspondences are then used to estimate a *camera pose*, i.e., a 2D–3D registration, of the video image and the CT contents.

Camera pose estimation

The task of determining the orientation and position (pose) of a camera with respect to an object is referred to as *camera pose estimation*. Updating this transformation in real time is the fundamental basis of an AR application in order to superimpose virtual content on real video images. In our scenario, the pose of the tablet camera with respect to the phantom is estimated using the 3D coordinates of the radiodense, fiducial markers extracted from the CT image volume and their projected, colored counterparts in the video image (cf. Fig. 6).

Several approaches exist for calculating or approximating the camera pose for real-time AR applications. Estimating the camera pose by point correspondences is also known as the *Perspective-n-Point problem (PnP)*.

All PnP algorithms expect the projective camera parameters, i.e., the camera intrinsics focal length, principal point and distortions parameters, to be known. These parameters are obtained by conducting a camera calibration. We rely on the well-known method by Zhang [20] using a chessboard as calibration target for the calibration of the tablet camera.

As stated above, camera pose estimation plays an integral role in visual navigation or surgical navigation, robot localization, photogrammetry and other areas. Therefore, a lot of solutions for solving the PnP problem exist [21–24], and new approaches arise quickly [25,26]. A basic differentiation of these algorithms is to divide them into analytical solutions, iterative procedures or hybrid approaches. However, the choice of the appropriate solver for the PnP problem depends on the conditions of the target application: number of point correspondences, degree of uncertainty concerning the point locations, runtime requirements, geometry of the target object, etc. Therefore, we examined three different camera pose estimation algorithms for their applicability to our mobile AR system in terms of visualization accuracy. For the final experiments, as described in section *Comparison*

of iPad-, ultrasound- and fluoroscopy assistance, a RANSAC (Random sample consensus, [27]) based, iterative Levenberg–Marquardt approach [28] was used which has proved to be the most robust and stable algorithm in terms of the TVE (target visualization error).

Estimation of the target visualization error

As in [14], we will refer to the offset of the projected virtual anatomy from the real structures in the video image as the TVE. We will refer to the TVE on the image plane as TVE_{2D}. This error can be easily expressed and calculated in terms of pixels in the image: The 3D object points are re-projected onto the current image using the calculated camera pose. The average distance of these points and the points that have been located by image analysis represents the TVE_{2D} in pixels per video frame. However, calculating an error in terms of millimeters in our experiments is difficult since no ground truth of the camera pose is available and image pixels cannot be converted into millimeters. Therefore, we use a similar evaluation strategy as in Baumhauer et al. [14]. For the evaluation, we use a random leave-one-out test strategy. Since six fiducial markers are visible at a time, we randomly pick five of them for each frame to estimate the camera pose. The one left out is used to derive a rough estimation of the TVE in space, i.e., the TVE_{3D}, as depicted in Fig. 7.

Please note that this evaluation strategy only gives an estimation of the real error. Since it incorporates the optical axis and the distortion parameters, it implies a perfect camera calibration which can also be error prone. Besides that, the calculated error does not reflect a depth error, but rather projects this error onto the described plane. A good overview of registration error analysis is given in [29].

Visualization

For every frame, a new camera pose is estimated. By applying the parameters to the virtual camera, the CT image volume and all virtual contents, i.e., segmentations and points, are rendered with the video image as texture in the background. Please note that the video image is first undistorted by utilizing the camera distortion parameters as have been determined by camera calibration. For the visualization itself, the application provides the user with the choice between a transparent volume rendering of the image (using different transfer functions) and the representation of the 3D surface models. However, for the experiments, only the 3D surface rendering was used, with different opacity and color settings for the different structures. The augmented reality image created in this manner is compressed, sent back and shown on the tablet.

Experiments

Estimation of the target visualization error using different camera pose algorithms

In previous work [14,30], a combination of the POSIT [22] and OrthIt algorithm [23] was implemented to derive a robust solution for the PnP problem. However, since the algorithm was designed for colored spheres attached to a prostate as fiducial markers (cf. [31]) and new approaches have been developed, we examined two further RANSAC-based approaches: an iterative Levenberg–Marquardt algorithm for finding the minimum reprojection error and the recent *Efficient Perspective-n-Point* (EPnP) approach [24]. Both algorithms are available as fast implementations in the OpenCV toolkit [32]. To give the reader an idea of the visualization accuracy of the proposed AR navigation system, we evaluated the three algorithms, i.e., combination of POSIT and OrthIt, RANSAC-based Levenberg–Marquardt approach and RANSAC-based EPNP, on recorded videos of our experiments. The video sequences include moving as well as static images of the phantom with six fiducial markers which are visible throughout the whole sequence. As a comparative measurement, the TVE_{3D} was assessed for each algorithm using the error estimation strategy as described in the section *Estimation of the target visualization error*.

Comparison of iPad-, ultrasound- and fluoroscopy assistance

The aim of the experiments was to assess the time for a successful puncture of the renal collecting system and the radiation dose exposed to the phantom when using the proposed navigation system, ultrasound and fluoroscopy, respectively. Altogether, we conducted 53 punctures in a series of five experiments on different days. Three experiments were carried out using standard CT imaging of the phantoms; two used the DynaCT setup. Accordingly, five identically constructed, but different phantoms with 10 different porcine kidneys were employed. The sequence in which the guidance techniques were used was changed each time.

The punctures were conducted by three urologists. An urological trainee and two experts (more than 5 years experience in PCNL) from different hospitals were asked to perform the punctures. Since the trainee performed more than half of all punctures, the results of the punctures conducted by the two experienced urologists are grouped together. The experiments were accompanied by one technician who observed the experiments in a room nearby and who took care of the navigation software. Furthermore, the technician was responsible for taking times and recording the radiation exposure as reported by the manufacturer's software of the fluoroscopic devices.

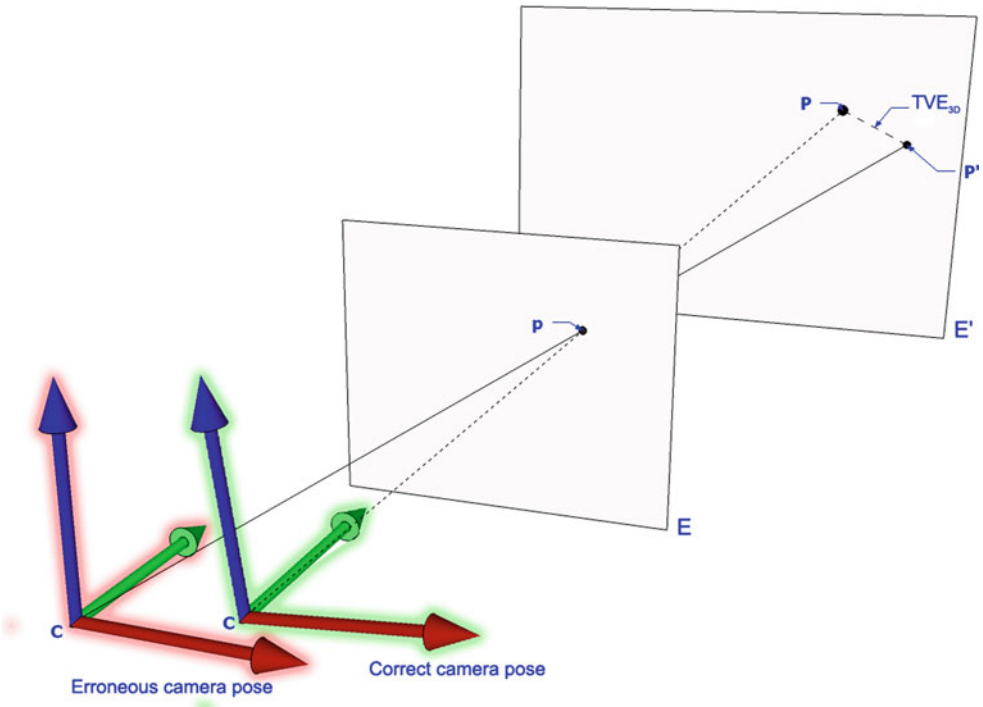

Fig. 7 Illustration of the error estimation: a distortion corrected point p (located by image analysis) in the image plane E and its corresponding 3D point P. The plane E', in which P lies, is parallel to the image plane, i.e., defined by the optical axis (green arrow). For the perfectly calculated camera pose (right coordinate system), the dashed ray which starts at the optical center C and goes through p intersects with the plane E' in the point P. For an estimated, erroneous camera pose (left coordinate system), the ray intersects at a Point P'. The distance between the two points P and P' then represents the TVE_{3D}

For the whole series of experiments, a single camera calibration for the iPad camera was conducted. At the beginning of each experiment, the equipment and the phantom were prepared, i.e., the iPad fixation was built up and a Wi-Fi connection between server and tablet was established. Before 3D imaging, which was either done with a CT device before the actual experiment or during the experiment using the DynaCT device, the fiducial markers were attached on the phantom's surface and contrast agent was injected into the porcine kidney collecting system to achieve good contrast (cf. Fig. 3c). As soon as the created image volume was imported to the server PC, the important structures and 3D locations of the fiducial markers were extracted by means of semi-automatic segmentation tools. The renal system filled with contrast agent and the fiducial markers were easily delineated by scalar thresholding procedures. Subsequently, the main experiments could be started.

The operator in charge (either the trainee or one of the two experienced urologists) was then asked to perform 3–4 punctures with each guidance technique (tablet navigation, fluoroscopy and ultrasound) consecutively. The overall number of conducted punctures per experiment varied depending on the time that was available in the intervention room. The time and the radiation exposure from the fluoroscopic device were recorded from the moment of needle insertion until successful puncturing of the collecting system or until the operator canceled the trial. An attempt was regarded as canceled if the needle had to be pulled out completely. To be compliant with the clinical setup in which a puncture cannot be canceled, the recorded data were nevertheless added to the analysis. Puncture time of canceled trials was treated as a censored observation. An important aspect of the experimental design was the fact that the operators could use fluoroscopy at any time to verify the progress of the needle insertion. Thus, the mobile navigation and ultrasound guidance could always be combined with fluoroscopic information if the operator wanted to. Verification that the collecting system was reached was achieved by pulling out fluid, i.e., the contrast agent, through a syringe. As soon as it was possible to extract the fluid, the puncture was deemed successfully finished. The whole experiment was finished after an overall number of nine to twelve punctures—depending on the available time, as stated earlier. Figure 8 shows the tablet navigation system in the experimental setup. Online Resource 1 shows a sample video sequence of the AR view seen on the tablet display.

Radiation exposure was measured in microgray per square meter (μGym^2), and treated as a continuous, normally distributed variable. Mean radiation differences were compared for the three guidance conditions using Student's *t* test, with fluoroscopy guidance as the reference. Time until successful

Fig. 8 Urological trainee evaluating the needle insertion after puncture trial with the mobile navigation system

puncture was measured in seconds, with puncture times of canceled trials being treated as censored observations. Differences between the three guidance conditions were analyzed using Cox proportional hazards model, with a discrete covariate indicating the condition. Separate analyses were carried out for experts and trainees.

Due to the exploratory nature of the current feasibility trial, results of statistical analyses should not be interpreted in a strict decisive sense, and P values are reported with correction for multiplicity.

Results

The following is a summary of the results of the conducted experiments. As described above, the results are split into the comparison of iPad-, ultrasound- and fluoroscopy assistance and the estimation of the target visualization error using different Camera Pose algorithms.

The overall time to prepare the navigation system, i.e., setup and image processing, was approximately 12 min. The segmentation process of delineating the kidneys and the renal collecting system in our phantoms took about 8 min.

The inexperienced urologist made 16 punctures using the iPad assistance (81 % successful), 10 using ultrasound (70 % successful) and 11 with standard fluoroscopy support (64 % successful).

Mean radiation exposure for the trainee was 7.8 μGym^2 (SD = 6.8), with fluoroscopy guidance. With iPad guidance, we measured substantially less radiation (mean = 2.6 μGym^2, SD = 3.6, t = −2.70, P = .011), whereas

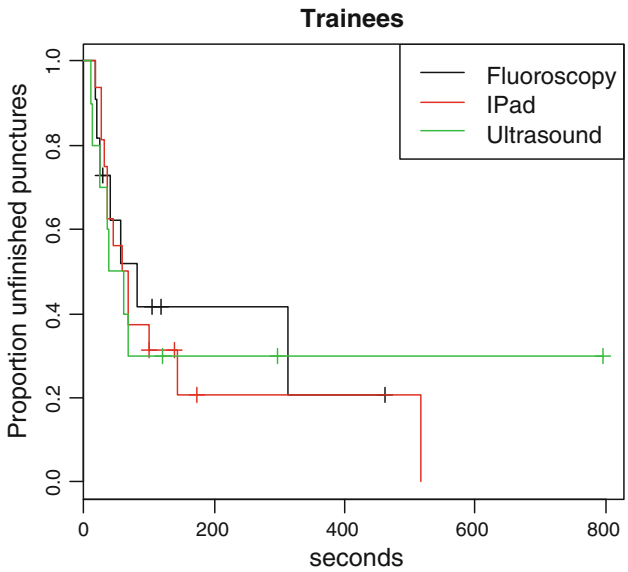

Fig. 9 The proportion of unfinished punctures in the Cox model for the trainee

radiation was slightly more reduced when using ultrasound guidance (mean 2.4 μGym^2, SD = 4.8, t = −2.49, P = .018). Median puncture times amounted to 81, 65 and 50 s for fluoroscopy-, iPad-, and ultrasound-guided punctures. Compared to fluoroscopy, iPad guidance accelerated the procedure by 20 % (hazard ratio = 1.20, z = −0.39, P = 0.70, not significant). Ultrasound accelerated the procedure by 14 % (hazard ratio = 1.20, z = −0.39, P = 0.80). Kaplan–Meier estimates of puncture time are depicted in Fig. 9.

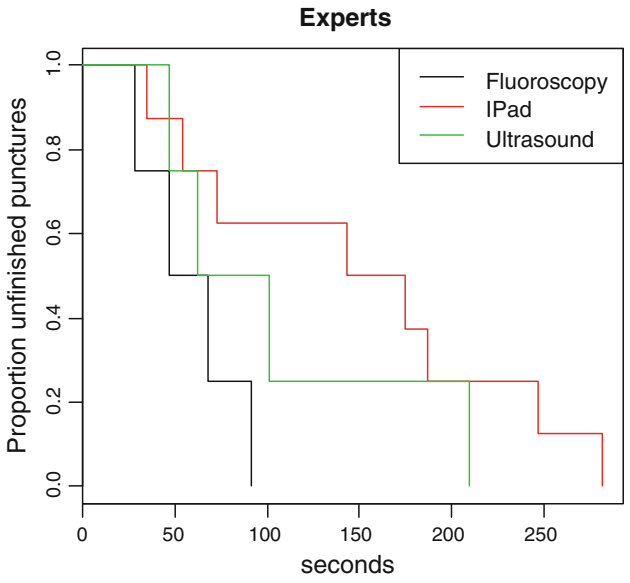

Fig. 10 The proportion of unfinished punctures in the Cox model for the expert group

The experienced urologists conducted 8 punctures with the iPad navigation system (100 % successful), 4 punctures with ultrasound (100 % successful) and 4 trials using fluoroscopy (100 %). The mean radiation exposure was 6.9 μGym^2 (SD = 4.2) with fluoroscopy guidance. With iPad guidance, we observed a slight change in radiation (mean=3.9 μGym^2, SD = 3.7, t = −1.44, P = .174), whereas radiation was lowest in guidance with ultrasound (mean 0.5 μGym^2, SD = 0.7, t = −2.65, P = .020). Median puncture times amounted to 58, 159 and 82 s for fluoroscopy-, iPad-, and ultrasound-guided punctures. Thus, compared to fluoroscopy, iPad guidance led to increased puncture times (hazard ratio = 0.22, z = −2.05, P = 0.04). The calculated Cox model is depicted in Fig. 10.

The following diagram depicts the results of the evaluation of the target visualization error TVE_{3D} for the three different Camera Pose Estimation algorithms POSIT and Orthit, RANSAC-based Levenberg–Marquardt approach and RANSAC-based EPNP.

Discussion

In this study, we present a detailed description and systematic evaluation of a novel augmented reality system that provides an intuitive virtual insight into the patient using a tablet computer such as an iPad. The system proved to be capable in assisting percutaneous needle access to the renal collecting system during PNCL. A statistical analysis on the outcomes of the phantom study revealed that puncturing could be accomplished slightly faster and with substantially less radiation exposure to the patient than with conventional methods using ultrasound or fluoroscopy, if puncturing was performed by an inexperienced urologist. On the basis of this analysis, we could not find major improvements by the system when used by experienced urologists. Besides that, we evaluated three different Pose Estimation algorithms that were used for the proposed AR system in terms of visualization accuracy. The RANSAC-based Levenberg–Marquardt approach performed best with a mean accuracy of 2.5 mm.

The navigation method was designed to be as simple and intuitive as possible to not alter the clinical workflow of the intervention drastically. Therefore, the system setup is divided in five easy steps:

1. Attachment of fiducial markers
2. Hardware setup
3. Localization of fiducial markers
4. Segmentation of relevant structures in the image volume
5. Color adjustments, if needed, and start of AR visualization

The overall preparation time (about 12 min) is acceptable. Of course, in clinical reality, the preparation time depends on the target structures being delineated. Preliminary experiments showed that ballistic gelatin is not easily distinguishable from the kidney tissue. We applied a high concentration of contrast agent into the gelatin mixture to facilitate the delineation of the kidney. We want to emphasize that the system is also feasible without contrast agent. The human kidney and the collecting system are identifiable in unenhanced CTs. Although our software framework offers easy-to-use, semi-automatic segmentation methods for these tasks, an automatic segmentation for certain entities, e.g., the kidney, spleen, colon, is desirable.

Apart from that, the integration of the system was possible at two different clinical sites without problems highlighting the system's mobility, which contrasts with device-dependent or stationary navigation approaches such as the laser guidance provided by the DynaCT device [13], mechanical add-ons, such as the needle stabilizing holder for a PCNL needle in [11] or the robotic device for prostate brachytherapy presented in Song et al. [8]. The generic design of the presented navigation system enables the clinician to couple it with any kind of 3D imaging device. This should be especially considered when discussing radiation exposure. Of course, one aim of the system is to reduce the amount of radiation the patient is exposed to. In our experiments, the iPad assistance lowered the radiation exposure by a factor of 3 for the inexperienced physician and by a factor of 1.8 for the experts in comparison with fluoroscopy usage. However, the experts were able to conduct successful punctures with ultrasound and very sparse fluoroscopy usage. Since we can employ diagnostic CT images for our mobile navigation system, we

are not altering the clinical standard workflow for treatment of kidney stones [33] and we do not expose the patient to additional radiation while lowering the need for fluoroscopic control imaging during the intervention.

Concerning the puncturing time and the overall number of successful punctures, we want to state that the urological trainee performed better with the iPad assistance than the experts did. The most reasonable interpretation here is the fact that the experts are much more trained in conducting PCNLs with standard equipment. As reported in the literature [34], experience with the procedure itself as well as with the equipment used is a very important factor in PCNL treatment. Thus, training is obviously also mandatory for the proposed mobile AR navigation system in its current state. Puncturing time is the commonly used measure in the literature to judge the benefit of a puncturing aid in PCNL quantitatively. However, a decisive advantage of our navigation approach, which is to provide spatial information of surrounding anatomical structures in relation to the collecting system, can only be judged qualitatively. The representation of a 3D anatomy on top of the phantom during puncturing was regarded as very helpful during the experiments. Of course, this feature is especially helpful in the real clinical setup [16].

A main drawback is the lack of guidance information, i.e., the augmented reality view does not provide real depth information yet. This led to difficulties in reaching the target in some cases and may also explain the high variance of the intervention time and the radiation exposure for some cases. To address this issue, video-based tracking of the PCNL needle is currently being tested. With a tracked needle, the augmented reality view can be enhanced with depth information. Of course, this additional information has to be presented in an appropriate way. Different visualization and interaction schemes for computer-assisted needle insertions have already been examined [35]. This will possibly resolve the need for training and make the handling of the system more intuitive. In general, novel technical solutions often lack widespread acceptance on the part of the physicians who tend to be reluctant to change their habits. The systems should require a minimum of user interaction, setup time and training.

Depending on the setup, puncturing times reported in the literature vary a lot: in a direct comparison of solely ultrasound- and fluoroscopy-guided PCNL, puncturing times of 14.5 $\pm$ 2.6 min and 9.4 $\pm$ 2.3 min are reported [36], whereas [37] examined ultrasound as an adjunct to fluoroscopy with puncturing times of 3.2 min for fluoroscopy and 1.8 min for the combination of both modalities. However, in contrast to those clinical studies, the puncturing times in our experiments for both experience levels are mostly lower for fluoroscopy and ultrasound. Puncturing a rigid phantom assuming a rigid patient anatomy is of course a much easier task than the real clinical procedure. A phantom design simulating motion, as used for example in Maier-Hein et al. [38], or animal experiments would provide a more realistic setup. However, a main benefit of the constructed phantom is that its haptic feedback is similar to that of human soft tissue when it is perforated with a standard PCNL needle.

To compare the proposed system with other navigation approaches for PCNL, we want to revisit the references mentioned in the introduction [11–13]. Lazarus et al. present a novel mechanical assistance for PCNL with reported puncturing times of 225 s without and 118 s with the respective device in an in vitro study. These numbers are comparable to our results, and the device seems to be clinically applicable in an easy way. Despite these aspects, we see a benefit of our system in using 3D imaging for PCNL since we are able to visualize surrounding structures and we are able to apply pre- or intra-operative planning. Huber at al. present a navigation approach based on electromagnetic tracking. With a puncturing time of 14 s in an ex vivo porcine model after a learning phase of 30 punctures, they can report the best results. Furthermore, they do not need fluoroscopic assistance at all, except for verification. But, despite advances in electromagnetic tracking [39], it is not always applicable in real environments without further effort due to its vulnerability to ferromagnetic materials [40]. In contrast to Huber et al. we see a main advantage of our system in its seamless integration into the clinical workflow. The most recent work is that of Ritter et al. [13], introducing the laser guidance capabilities of the Siemens DynaCT device. For 10 punctures in a biological model, they report a median time of 4.6 (2–10.2) min using this type of assistance, which is rather long compared to the times we measured in our experiments. As stated previously, we see an advantage in the fact that the proposed mobile navigation system is not bound to a certain device, but can rather deliver navigation assistance independent of location and manufacturer. As a final statement concerning the different systems, we want to emphasize that the comparability of these approaches, including ours, is limited due to the different experimental protocols. For a real comparison, a greater number of puncture trials and operators in a similar experimental setup for each guidance approach would be necessary.

For registering imaging data and the patient, i.e., to create an augmented reality view of the patient, fiducial markers are employed, as e.g., proposed in Nicolau et al. [5]. From a technical point of view, newer approaches exist. Markerless methods use natural feature tracking [41] or cameras equipped with a range sensor, such as Time-of-Flight (*ToF*) cameras [15,42] or the Microsoft Kinect camera [43,44]. Although these approaches are a reasonable orientation in research, marker-based approaches already provide a precise registration in real time. Also, the attachment of the presented fiducial markers is unproblematic, non-invasive and therefore justifiable. Concerning the visualization accuracy, an error of about 2.5 mm at a distance of about 40 cm from the target

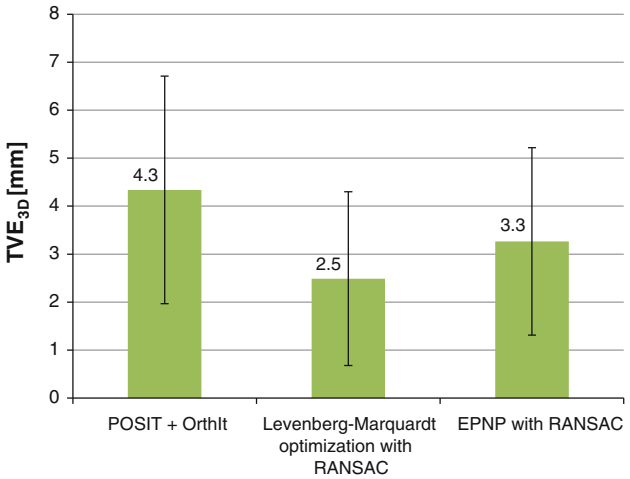

Fig. 11 The mean, approximated TVE_{3D} in millimeters for the three different types of Pose Estimation algorithms. The standard deviation is depicted as error bars in the diagram

seems satisfying for the visualization of the renal collecting system on the iPad. Based on the results shown in Fig. 11, we conclude that the RANSAC-based approaches are more suitable than the combined POSIT and OrthIt algorithm [30] for our type of application. Yet, the evaluation strategy is not ideal since it only gives a rough estimation for the target visualization error.

Future work will focus on the technical aspects of the system. Newer generations of tablet PCs may provide enough resources to drop the additional server PC. The most urgent improvement, however, is the development of a reliable tracking method for the needle to incorporate adequate depth information. Automated path planning, as suggested in Seitel et al. [45], i.e., the calculation of an optimal path to the collecting system based on the individual patient's anatomy, would be a reasonable extension to the system. The use of radio-dense markers yields another advantage for the future development of the system: Because the fiducial markers are also visible in the 2D fluoroscopic images, these images can also be fused with the video images. By that, a live deformation control and data verification component would be added to the system. The step subsequent to the technical improvements is an in vivo evaluation of the system in animal experiments to incorporate the influence of a non-rigid patient anatomy.

In conclusion, the proposed tablet computer-based augmented reality system has proven helpful in assisting percutaneous interventions such as PCNL and shows benefits compared to other state-of-the-art assistance systems. With the mentioned improvements, the system may even facilitate renal access during PCNL for experts without the need of additional training. In particular, the simple integration into the clinical workflow underlines the potential impact of this approach to such interventions. Moreover, due to its mobility, low costs and generic design, the presented system has a wide field of applications.

Acknowledgments The presented work was conducted within the setting of the "Research group 1126: Intelligent Surgery-Development of new computer-based methods for the future workplace in surgery" funded by the German Research Foundation (DFG). Furthermore, we want to thank the staff in the urological departments of our partner hospitals for the support during our experiments. The presented software was developed as part of the Medical Imaging Interaction Toolkit (MITK, http://www.mitk.org).

Conflict of interest None.

References

1. Romero V, Akpinar H, Assimos DG (2010) Kidney stones: a global picture of prevalence, incidence, and associated risk factors. Rev Urol 12(2–3):86–96
2. Skolarikos A, Alivizatos G, de la Rosette JJ (2005) Percutaneous nephrolithotomy and its legacy. Eur Urol 47(1):22–28
3. Michel MS, Trojan L, Rassweiler JJ (2007) Complications in percutaneous nephrolithotomy. Eur Urol 51(4):899–906
4. Andonian S, Scoffone C, Louie MK, Gross AJ, Grabe M, Daels F, Shah HN, De La Rosette J (2012) Does imaging modality used for percutaneous renal access make a difference? a matched case analysis. J Endourol 27(1):24–28
5. Nicolau S, Soler L, Mutter D, Marescaux J (2011) Augmented reality in laparoscopic surgical oncology. Surg Oncol 20(3): 189–201
6. Maier-Hein L, Tekbas A, Seitel A, Pianka F, Müller SA, Satzl S, Schawo S, Radeleff B, Tetzlaff R, Franz AM, Müller-Stich BP, Wolf I, Kauczor H-U, Schmied BM, Meinzer H-P (2008) In vivo accuracy assessment of a needle-based navigation system for CT-guided radiofrequency ablation of the liver. Med Phys 35(12):5385–5396
7. Fichtinger G, Deguet A, Fischer G, Iordachita I, Balogh E, Masamune K, Taylor RH, Fayad LM, de Oliveira M, Zinreich SJ (2005) Image overlay for CT-guided needle insertions. Comput Aided Surg 10(4):241–255
8. Song DY, Burdette EC, Fiene J, Armour E, Kronreif G, Deguet A, Zhang Z, Iordachita I, Fichtinger G, Kazanzides P (2011) Robotic needle guide for prostate brachytherapy: clinical testing of feasibility and performance. Brachytherapy 10(1):57–63
9. Wein W, Brunke S, Khamene A, Callstrom MR, Navab N (2008) Automatic CT-ultrasound registration for diagnostic imaging and image-guided intervention. Med Image Anal 12(5):577–585
10. Wood BJ, Kruecker J, Abi-Jaoudeh N, Locklin JK, Levy E, Xu S, Solbiati L, Kapoor A, Amalou H, Venkatesan AM (2010) Navigation systems for ablation. J Vasc Interv Radiol 21(8 Suppl):257–263
11. Lazarus J, Williams J (2011) The locator: novel percutaneous nephrolithotomy apparatus to aid collecting system puncture-a preliminary report. J Endourol 25(5):747–750
12. Huber J, Wegner I, Meinzer HP, Hallscheidt P, Hadaschik B, Pahernik S, Hohenfellner M (2011) Navigated renal access using electromagnetic tracking: an initial experience. Surg Endosc 25(4):1307–1312
13. Ritter M, Rassweiler MC, Hacker A, Michel MS (2012) Laser-guided percutaneous kidney access with the Uro Dyna-CT: first experience of three-dimensional puncture planning with an ex vivo model. World J Urol. 30:1–5
14. Baumhauer M, Simpfendörfer T, Stich BM, Teber D, Gutt C, Rassweiler J, Meinzer H-P, Wolf I (2008) Soft tissue navigation for

laparoscopic partial nephrectomy. Int J Comput Assist Radiol Surg 3:307–314
15. Maier-Hein L, Franz AM, Fangerau M, Schmidt M, Seitel A, Mersmann S, Kilgus T, Groch A, Yung K, dos Santos TR, Meinzer H-P (2011) Towards mobile augmented reality for on-patient visualization of medical images. In: Handels H, Ehrhardt J, Deserno TM, Meinzer H-P, Tolxdorff T (eds) Bildverarbeitung für die Medizin. Springer, Berlin, pp 389–393
16. Rassweiler JJ, Müller M, Fangerau M, Klein J, Goezen AS, Pereira P, Meinzer HP, Teber D (2012) iPad-assisted percutaneous access to the kidney using marker-based navigation: initial clinical experience. Eur Urol 61(3):628–631
17. Wolf I, Vetter M, Wegner I, Böttger T, Nolden M, Schöbinger M, Hastenteufel M, Kunert T, Meinzer H-P (2005) The medical imaging interaction toolkit. Med Image Anal 9(6):594–604
18. Fangerau M (2011) Medical imaging interaction toolkit for mobile devices. Master's thesis, Hochschule Mannheim, University of Applied Sciences
19. Müller M, Groch A, Baumhauer M, Maier-Hein L, Teber D, Rassweiler J, Meinzer H-P, Wegner I (2012) Robust and efficient fiducial tracking for augmented reality in HD-laparoscopic video streams. In: DRH III, Wong KH, (eds) SPIE medical imaging 2012: visualization, image-guided procedures, and modeling, vol 8316, No 1. SPIE, p 83161M
20. Zhang Z (2000) A flexible new technique for camera calibration. IEEE T Pattern Anal 22:1330–1334
21. Lowe DG (1991) Fitting parameterized three-dimensional models to images. IEEE Trans Pattern Anal Mach Intell 13:441–450
22. DeMenthon DF, Davis LS (1995) Model-based object pose in 25 lines of code. Int J Comput Vis 15:123–141
23. Lu C-P, Hager GD, Mjolsness E (2000) Fast and globally convergent pose estimation from video images. IEEE Trans Pattern Anal Mach Intell 22:610–622
24. Lepetit V, Moreno-Noguer F, Fua P (2008) EP*n*P: an accurate o(*n*) solution to the P*n*P problem. Int J Comput Vis 81(2):155–166
25. Sarkis M, Diepold K (2012) Camera-pose estimation via projective Newton optimization on the manifold. IEEE Trans Image Process 21(4):1729–1741
26. Li S, Xu C, Xie M (2012) A robust o(n) solution to the perspective-n-point problem. IEEE Trans Pattern Anal Mach Intell. 34(7):1444–1450
27. Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM 24(6):381–395
28. Nocedal J, Wright SJ (2006) Numerical optimization, 2nd edn. ser. Springer series in operations research and financial engineering. Springer, New York
29. Holloway RL (1997) Registration error analysis for augmented reality. Presence Teleoper Virtual Environ 6(4):413–432
30. Baumhauer M (2008) Real-time compensation of organ motion for augmented reality in laparoscopic surgery. Ph.D. dissertation, Ruprecht-Karls University Heidelberg
31. Simpfendörfer T, Baumhauer M, Müller M, Gutt CN, Meinzer HP, Rassweiler JJ, Guven S, Teber D (2011) Augmented reality visualization during laparoscopic radical prostatectomy. J Endourol 25:1841–1845
32. Laganière R (2011) OpenCV 2 computer vision application programming cookbook. Packt Publishing, Birmingham
33. Miller NL, Lingeman JE (2007) Management of kidney stones. BMJ 334(7591):468–472
34. Ritter M, Siegel F, Krombach P, Martinschek A, Weiss C, Hacker A, Pelzer AE (2012) Influence of surgeon's experience on fluoroscopy time during endourological interventions. World J Urol 31(1): 183–187
35. Seitel A, Maier-Hein L, Schawo S, Radeleff B, Müller SA, Pianka F, Schmied BM, Wolf I, Meinzer H-P (2007) In-vitro evaluation of different visualization approaches for computer assisted targeting in soft tissue. In: Lemke H, Inamura K, Doi K, Vannier M, Farman A (eds) International journal of Computer Assisted Radiology and Surgery. Berlin (Germany), pp 188–190, June 2007
36. Karami H, Rezaei A, Mohammadhosseini M, Javanmard B, Mazloomfard M, Lotfi B (2010) Ultrasonography-guided percutaneous nephrolithotomy in the flank position versus fluoroscopy-guided percutaneous nephrolithotomy in the prone position: a comparative study. J Endourol 24(8):1357–1361
37. Agarwal M, Agrawal MS, Jaiswal A, Kumar D, Yadav H, Lavania P (2011) Safety and efficacy of ultrasonography as an adjunct to fluoroscopy for renal access in percutaneous nephrolithotomy (PCNL). BJU Int 108(8):1346–1349
38. Maier-Hein L, Walsh CJ, Seitel A, Hanumara NC, Shepard J-A, Franz AM, Pianka F, Müller SA, Schmied B, Slocum AH, Gupta R, Meinzer H-P (2009) Human vs. robot operator error in a needle-based navigation system for percutaneous liver interventions. In: SPIE medical imaging 2009: visualization, image-guided procedures, and modelling, vol 7261, p 72610Y (12 p)
39. Franz A, März K, Hummel J, Birkfellner W, Bendl R, Delorme S, Schlemmer H-P, Meinzer H-P, Maier-Hein L (2012) Electromagnetic tracking for us-guided interventions: standardized assessment of a new compact field generator. Int J Comp Assist Radiol Surg 7:1–6
40. Yaniva Z, Wilson E, Lindisch D, Cleary K (2009) Electromagnetic tracking in the clinical environment. Med Phys 36(3):876–892
41. Lee T, Hollerer T (2009) Multithreaded hybrid feature tracking for markerless augmented reality. IEEE Trans Vis Comput Graph 15(3):355–368
42. Grimm R, Bauer S, Sukkau J, Hornegger J, Greiner G (2012) Markerless estimation of patient orientation, posture and pose using range and pressure imaging. Int J Comput Assist Radiol Surg 1:1–9
43. Newcombe RA, Davison AJ, Izadi S, Kohli P, Hilliges O, Shotton J, Molyneaux D, Hodges S, Kim D, Fitzgibbon A (2011) Kinectfusion: real-time dense surface mapping and tracking. In: Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on, pp. 127–136, Oct 2011
44. Mersmann S, Gergel I, Seitel A, Gaa J, Wegner I, Meinzer H-P, Maier-Hein L (2011) Microsoft kinect controller as intra-operative imaging modality. Int J Comput Assist Radiol Surg 6(Suppl 1): 251–252
45. Seitel A, Engel M, Sommer CM, Radeleff BA, Essert-Villard C, Baegert C, Fangerau M, Fritzsche KH, Yung K, Meinzer H-P, Maier-Hein L (2011) Computer-assisted trajectory planning for percutaneous needle insertions. Med Phys 38(6):3246–3259