Intraoperative Image-based Multiview 2D/3D Registration for Image-Guided Orthopaedic Surgery: Incorporation of Fiducial-Based C-Arm Tracking and GPU-Acceleration

Yoshito Otake,
Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218 USA

Mehran Armand,
Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 20723 USA

Robert S. Armiger,
Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 20723 USA

Michael D. Kutzer,
Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 20723 USA

Ehsan Basafa,
Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218 USA

Peter Kazanzides, and
Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218 USA

Russell H. Taylor, IEEE [Fellow]
Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218 USA

Yoshito Otake: otake@jhu.edu; Mehran Armand: mehran.armand@jhuapl.edu; Robert S. Armiger: robert.armiger@jhuapl.edu; Michael D. Kutzer: michael.kutzer@jhuapl.edu; Ehsan Basafa: basafa@jhu.edu; Peter Kazanzides: pkaz@jhu.edu; Russell H. Taylor: rht@jhu.edu

Abstract

Intraoperative patient registration may significantly affect the outcome of image-guided surgery (IGS). Image-based registration approaches have several advantages over the currently dominant point-based direct contact methods and are used in some industry solutions in image-guided radiation therapy with fixed X-ray gantries. However, technical challenges including geometric calibration and computational cost have precluded their use with mobile C-arms for IGS. We propose a 2D/3D registration framework for intraoperative patient registration using a conventional mobile X-ray imager combining fiducial-based C-arm tracking and graphics processing unit (GPU)-acceleration. The two-stage framework 1) acquires X-ray images and estimates relative pose between the images using a custom-made in-image fiducial, and 2) estimates the patient pose using intensity-based 2D/3D registration. Experimental validations using a publicly available gold standard dataset, a plastic bone phantom and cadaveric specimens have been conducted. The mean target registration error (mTRE) was 0.34 ± 0.04 mm (success rate: 100%, registration time: 14.2 s) for the phantom with two images 90° apart, and 0.99 ± 0.41

Correspondence to: Yoshito Otake, otake@jhu.edu.

mm (81%, 16.3 s) for the cadaveric specimen with images 58.5° apart. The experimental results showed the feasibility of the proposed registration framework as a practical alternative for IGS routines.

Keywords

C-arm pose tracking; GPU-acceleration; image-guided surgery; intraoperative 2D/3D registration

I. Introduction

Registering the patient to a digitized navigation system has been extensively investigated in image-guided surgery (IGS). The registration process in an IGS system solves the transformation between the coordinates of the preoperatively acquired data and the patient lying on the operating table. Registration is one of the key elements affecting the accuracy of the IGS system since this maps digitized information to surgical reality [1].

Although point-based registration using physically contacted anatomical features is dominant in current IGS systems, an image-based 2D/3D registration approach, where intraoperative X-ray images are used to solve the registration, has several advantages. First, it does not require physical contact with the anatomical region of interest in the patient, in contrast to the conventional point-based method which requires invasive exposure and contact with large portions of target anatomy. Second, the preoperative images do not need to be segmented to isolate contact features; point-based registration requires the segmentation of the target anatomy to identify the location of anatomical landmarks or surfaces, and it usually involves manual interaction by an experienced operator, making it time consuming and subjective by nature. Third, the procedure time is reduced since the image-based approach requires the surgeon to acquire only a few X-ray images.

There are a few industry solutions (e.g., ExacTrac (BrainLAB, Heimstetten, Germany), CyberKnife Robotic Radiosurgery System (Accuray Incorporated, Sunnyvale, CA) and various implementations of 2D/3D registration [2]–[5] in the field of image-guided radiation therapy (IGRT) where the goal is to verify patient position for each treatment and register the treatment plan to the patient using a room-mounted fixed X-ray gantry. However, in IGS applications, there are two major technical obstacles that preclude practical use of 2D/3D registration in a clinical routine: 1) uncertainty of extrinsic/intrinsic calibration of the X-ray imager; and 2) the expensive computational load associated with 2D/3D image registration. The former is caused by instability of mobile-type imaging devices (e.g., due to gravity-induced mechanical flex [6]) commonly used in the operating room as compared to stable and motion-encoded imaging devices used for IGRT. The latter is due to the fact that the generation of digitally reconstructed radiographs (DRR) is computationally expensive. In this work, we hypothesized that our GPU-accelerated DRR generation algorithm can affordably decrease computation time and provide a practical solution for clinical IGS routines.

This paper proposes a clinically practical framework for intraoperative image-based 2D/3D registration using a mobile C-arm imaging device. The technical contributions of the present

work include: 1) a streamlined system integration of intraoperative 2D/3D registration incorporating fiducial-based C-arm tracking and GPU-acceleration, 2) extensive experimental validation using a publicly available gold standard dataset, plastic phantom and cadavers in a clinically realistic surgical scenario, and 3) comparative analysis between different numbers of images, DRR generation algorithms, similarity measures, and optimization strategies. Notably, our approach is amenable to a time-critical clinical scenario because it does not require manual segmentation of the fiducial from 2D X-ray images and target anatomy from CT volume data.

In this paper, we demonstrate the performance of the proposed framework specifically within the context of an IGS application in femoral bone augmentation surgery, where the 2D/3D registration is used for registering the intraoperative tracker coordinate system to the preoperative CT used for bone cement injection planning [7] (Fig. 1).

II. Related Work

The workflow of a typical DRR-based 2D/3D registration algorithm includes: generating an initial transformation estimate, creating DRR images from the 3D volume and the current estimate of the transformation, computing a similarity measure between DRRs and 2D X-ray images, and updating the current estimate to increase similarity until a convergence criterion is achieved. The Sections II-A and II-B focus on two specific areas of image-based 2D/3D registration that are relevant to the contribution of this paper: C-arm pose estimation and DRR generation.

A. C-Arm Pose Estimation

Many interventional applications, including cardiac and neuro-angiography catheter manipulations, TIPS stent surgery, and IGRT typically use fixed, room-mounted X-ray imaging systems [5], [8]–[10]. In this case, since the device is geometrically stable, the relative pose between the acquired X-ray images can be accurately calibrated prior to the therapy.

In IGS applications using a mobile C-arm imager, however, estimation of the relative pose between multiple X-ray images is significantly more difficult than room-mounted imaging systems because of reduced mechanical stability. To avoid this difficulty, 2D/3D registration using a single image has been explored for rigid [11]–[14] and nonrigid [15] cases in IGS. However, in these cases the registration accuracy is limited in a direction perpendicular to the image plane. Because the movement of the target anatomy in that direction results in a scale change only in 2D projection, the objective function is not sensitive to translation along that axis and optimization algorithms may converge on local maxima rather than global maxima. Projection images from other viewpoints, therefore, can increase the accuracy and robustness [9].

Previously proposed approaches to estimate the relative pose of multiple images in a mobile C-arm imager include 1) using an external tracker and 2) using an X-ray opaque fiducial marker within the field of view and imaging it together with the target anatomy. As the approach 1), optical and electromagnetic external trackers are typically used since most of

IGS applications require continuous tool tracking in any case [16], [17]. Accuracy, however, remains an issue as the C-arm movement is relatively large compared to the center of the calibrated tracking volume (typically 1.0–1.2 m radius) and/or limited line of sight may impose frequent equipment adjustments. As the approach 2), several authors (e.g., [18]–[20]) have considered specially designed fiducial structures for estimating C-arm pose from intraoperative X-ray images. At Johns Hopkins University, Taylor *et al.* [18] developed a corkscrew-like fiducial mounted around the shaft of a surgical instrument, and Jain [19], [20] subsequently developed a mathematically optimized fiducial marker (FTRAC) consisting of ellipses, lines and points. The 6DOF pose of the C-arm imager with respect to the fiducial is computed from the X-ray image of fiducial features (beads and wires) and its known geometry. The relative pose between images can be derived from the per-image pose estimation result.

In this paper we employ a fiducial based approach based on the FTRAC and extend the previously proposed estimation algorithm [19] so that it works robustly without the feature segmentation process.

B. DRR Generation

Generation of a DRR from a CT volume is a computationally expensive process, and the computation time associated with DRR generation has been the main limiting factor that precludes routine clinical use of intensity-based 2D/3D registration. Some have used a precomputed set of parameters in a certain data structure, such as an attenuation field [21], [22], Transgraph [23] or gradient vectors [24], [25] to enhance DRR generation. There is a body of recent work using GPU processors to improve the efficiency of 2D/3D registration. Ruijters *et al.* [26] tried to improve precision of the summation process using a high precision z-buffer. Cardoso [27] used polygon-mesh formatted data and generated DRRs by modification of polygon rendering on the GPU assuming uniform intensity distribution within the polygon model. Xu *et al.* [28] and Yang *et al.* [29] used a GPU to enhance iterative reconstruction and lung tumor tracking respectively. One popular method explored mainly in the computer graphics community is ray-tracing using trilinear interpolation. The volume data that resides in the texture memory on the GPU is sampled with trilinear interpolation at a constant step length along a ray connecting the virtual X-ray source and each virtual detector element, and then the sampled voxel intensities are accumulated [30], [31]. Several open source software packages (e.g., Plastimatch, [32], VTK [33]) have an implementation of DRR generation using a GPU with the simple volume rendering technique mentioned above. While this method is fast and easy to implement, the accuracy of DRR will depend on the sampling step length. Another algorithm that provides a relatively more accurate DRR is known as Siddon's ray-tracing method [34], [35], which computes the intersection length between the ray and each voxel element and accumulates the voxel intensities weighted by the length. We have implemented both trilinear interpolation and Siddon's method on a GPU. A contribution of this paper is the comparison of the effect of different DRR generation algorithms on the performance of 2D/3D registration. Siddon's ray-tracing algorithm as described in detail in Section III-D) is used as the ground truth. At the moment, none of the existing open source packages have implemented Siddon's algorithm.

III. Methods

A. Problem Definition

Given multiple projection images of an anatomical structure with a fiducial marker of known geometry, our objective is to compute the relative pose between 1) each image, and 2) the images and the anatomical structure.

B. Preoperative Calibration

Calibration of intrinsic parameters was conducted preoperatively. In this study, we used two types of C-arm imagers: an OEC9600 (GE OEC Medical Systems, Salt Lake City, UT) with a nominal 9-in X-ray image intensifier (XRII); and a prototype portable C-arm cone-beam CT (CBCT) system (Siemens Medical Solutions, Germany) equipped with a flat-panel detector [36]. Different intrinsic calibration methods were employed for each C-arm. In the C-arm with XRII, image distortion correction was conducted prior to the intrinsic parameter calibration. We followed Sadowsky's [37], [38] method using a custom-designed two-plane calibration phantom. The calibration phantom was fixed to the C-arm detector and an X-ray image was acquired. A uniformly distributed grid of 1 mm metallic ball bearings on one plane of the calibration phantom was segmented in the image and Bernstein polynomials were fit to the segmented BBs in order to estimate the distortion pattern. The focal length and principal point were estimated using square features on both planes of the calibration phantom. For the flat-panel detector C-arm, we employed a calibration method proposed in [29] using a phantom with BBs aligned in a helix. A distortion correction step was not necessary for the flat-panel C-arm because the distortion is known to be negligible compared to XRII [39].

C. Registration Workflow

The following describes the workflow for registration (Fig. 2): the X-ray images and DRR generated from CT image are compared using a similarity measure and the pose of the CT image that gives a maximum similarity measure is determined through an optimization step. The proposed registration framework involves: 1) estimation of relative pose between X-ray images using the custom-made fiducial; 2) preprocessing of the preoperative CT data and X-ray images; 3) determination of an initial guess on the relative pose between X-ray images and CT data for the optimization; 4) DRR generation; 5) computation of an objective function; and 6) optimization. We describe each step in Sections III-CI–III-C-VI.

1) C-Arm Pose Estimation—The custom-designed hybrid fiducial used for C-arm pose estimation in this study is shown in Fig. 3. The fiducial contains two subsystems, one that makes it visible to an optical tracking system, and another (FTRAC [20]) that makes it visible during fluoroscopy tracking. The body of the $18 \times 18 \times 72$ mm FTRAC is made of polycarbonate and contains nine stainless steel beads, four straight stainless steel wires and two stainless steel wires shaped elliptically. It was designed for estimating the 6DOF pose of the C-arm detector from its projection image [Fig. 3(c)] using the known feature geometry. We followed the intensity-based registration approach similar to Fallavollita *et al.* [40], [41] which does not require the segmentation of the features. While the size of the FTRAC may affect pose estimation accuracy, we used a smaller FTRAC than Fallavollita *et al.* in order to

satisfy the clinical need for reducing the overall size and weight of the FTRAC attached to the osteoporotic bone.

In the proposed approach, the location of fournoncoplanar beads on the FTRAC is first roughly identified in the X-ray image coordinate system by using morphological filtering and thresholding. The 6DOF pose of the FTRAC is estimated using the POSIT algorithm [42] where the pose of the object with respect to the camera is estimated in an iteration loop under the assumption of scaled orthographic projection using point correspondence between the known 3D model and points in the 2D image. Then a projection image of the CAD model of the fiducial at the estimated pose [Fig. 4(b)] is generated and compared to the original X-ray image [Fig. 4(a)] using a similarity metric. The estimated pose is repeatedly updated to maximize the similarity metric using an optimization based on the Nelder–Mead Downhill Simplex Algorithm [43] to compute

$$T_{\text{estimated}} = \arg \max_{T} \left(SM \left(I_{\text{Xray}}, I_{\text{prj}} \left(T \right) \right) \right) \quad (1)$$

where $I_{\text{Prj}}(T)$ is the virtually projected image of the CAD model, I_{Xray} is the acquired X-ray image and $SM(I_1, I_2)$ represents a similarity measure between images I_1 and I_2. In the experiment described in Section IV, we used mutual information [44] as the similarity measure.

2) Preprocessing of the CT Data and the X-ray Images—Preoperatively acquired diagnostic CT [represented in Hounsfield Units (HU)] was converted into a volume image represented by linear attenuation coefficients based on the following definition [45]

$$\mu = \left\{ \frac{(1000 + HU)}{1000} \right\} \mu_{\text{water}} \quad (2)$$

where μ is the linear attenuation coefficient of the voxel and μ_{water} is the coefficient of water at the X-ray energy which was used for the CT scanning. For this work, we assumed a monoenergetic X-ray to convert HU to μ in (2). This assumption is commonly made in most of the reported 2D/3D registration techniques between CT and X-ray images [31]. Although the polyenegertic nature of X-rays yields several important effects on the projection images (e.g., beam hardening), CT data contains information about attenuation only at a certain X-ray energy levels, and generally there is no way to know the spectrum of the X-ray attenuation for each voxel from a CT value.

X-ray projection images acquired intraoperatively were also preprocessed to get a line integral of the linear attenuation coefficient from the detector intensity. The line integral of the linear attenuation coefficient is represented as

$$g_d = \int_0^d \mu(s) \; ds = -\ln \left(\frac{I_d}{I_0} \right) \quad (3)$$

where the middle term represents the integral along a line connecting the X-ray source and the detector element, I_d is the intensity at the detector, and I_0 is the intensity of the incident beam [45]. We defined I_0 as the intensity of a pixel in the area with no object between the

X-ray source and detector. Here, we assumed a point source where all X-ray beams are emitted from a point [whereas, in an actual X-ray generator, the focal point is of finite size (~ 0.5 mm)].

In the 2D/3D registration process, the line integral image of g_d in (3) was computed from each X-ray image and the DRR was computed from the linear attenuation coefficient volume computed by (2).

3) Initial Guess for Optimization—After estimating the relative pose between the acquired images using the fiducial, a pose of the patient anatomy was roughly approximated as an initial guess for the optimization process. The approximate pose was estimated from the surgical protocol, which usually indicates the position of the patient on the operating table (e.g., supine position, prone position, etc.) and the pose of the imager. (Alternatively, it could be determined by selecting some anatomical landmarks on the patient's skin prior to the procedure and conducting a point-based registration using those points.)

4) DRR Generation—DRR generation is one of the key components in intensity-based 2D/3D registration. Since the process is repeated typically hundreds or thousands of times during the entire optimization process, speeding up the DRR generation was of great importance. In most of the previous 2D/3D registration methods, ray-tracing with trilinear interpolation was used to generate DRRs [5], [11], [31] [Fig. 5(a)]. Trilinear interpolation computes an intensity value at each sample point in the 3D volume with a certain interval (step length) along a ray. Zollei *et al.* [46] attempted to reduce the computation time by sparsely sampling the rays, but the reduced image quality resulted in a less accurate registration. Spoerk *et al.* [47] used an efficient antialiasing method called "wobbled splatting" to approximate a higher resolution DRR with a low computational cost. In the case where a GPU is used, trilinear interpolation method is especially fast since it is conducted on the GPU hardware (rather than in software) for most of commercially available GPUs as a texture fetching process. The ray-tracing process also can be sped up by using a large step length, while the accuracy of the line integral along each ray is degraded.

In radiation therapy, mathematically more accurate methods are used to compute X-ray dose by computing radiological path length for each ray. Siddon's method [34], [35] is one of the methods where the length of intersection between the ray and each voxel is calculated and the intensity value of the voxel is weighted by the intersection length before summation [Fig. 5(b)]. This approach is computationally expensive compared to the trilinear interpolation approach, but the computed line integral is more accurate. Fig. 6 shows a simple 2D example illustrating the difference in accuracy when computing the line integral using different approaches. The example assumed an X-ray source 100 mm apart from a 1D detector array (30 mm long, 1 mm per element) and an object (30 mm × 10 mm) with a uniform attenuation coefficient ($\mu = 0.04$ mm^{-1}). The line integral computed at each detector element by Siddon's algorithm, and by the trilinear interpolation algorithm with step lengths of 0.01 mm and 0.001 mm were compared with the analytically computed result and demonstrates that Siddon's algorithm provides less error than the trilinear interpolation algorithm even for small step lengths. Note that the slight difference between Siddon's algorithm and the analytical result comes from the limitation of numerical precision on the

GPU. The computation on the GPU was carried out in single precision floating point (32 bits) whereas the analytical result was computed in double precision (64 bits).

Since both approaches have trade-offs between computational cost and the quality of the resulting DRR, we have implemented both approaches in order to determine the optimum compromise for our 2D/3D registration application.

5) Objective Function—Mutual information (MI) [44], normalized MI (NMI) [48] and gradient information (GI) [49] between the DRR and the X-ray image was used as the objective function in the optimization. GI was defined as

$$\alpha_{i,j}=\cos^{-1}\frac{\nabla p_1\left(i,j\right)\cdot\nabla p_2\left(i,j\right)}{\left|\nabla p_1\left(i,j\right)\right|\left|\nabla p_2\left(i,j\right)\right|}\ \left(\nabla p_1\left(i,j\right)=\left(\frac{\partial}{\partial i}p_1\left(i,j\right),\frac{\partial}{\partial j}p_1\left(i,j\right)\right)\right) \quad (4)$$

$$GI\left(p_1,p_2\right)=\sum_{i,j}w\left(i,j\right)\min\left(\left|\nabla p_1(i,j)\right|,\left|\nabla p_2(i,j)\right|\right)\ \left(w\left(i,j\right)=\frac{\cos\left(2\alpha_{i,j}\right)+1}{2}\right) \quad (5)$$

where $p_1(i, j)$, $p_2(i, j)$ are the pixel values in the fixed (X-ray) and floating (DRR) images at the (i, j) pixel. The angle (α) of the gradient vector of each pixel in each image is computed in (4). Then the magnitudes of the gradient vector at each pixel are summed up with weight of w computed from the angle [see (5)]. In order to take into account strong gradients that appear in both images, the minimum of the gradient magnitude is selected in (5). The derivative images (∇p_1, ∇p_2) were computed by convolving the image with the first derivatives of a Gaussian kernel of scale $\sigma = 1$ (detail of the selection of the σ is discussed in [50]). As shown in Fig. 7, this similarity measure effectively eliminates the distracting effect of external objects presented in one of the images during registration (in this case the fiducial and the fixture device of the fiducial appeared only in the fixed image). This is done by selecting the smaller gradient as described above, whereas the information theoretic similarity measures (e.g., MI, NMI) yield extra local minima distracting the optimization (see Fig. 15 in Section VI). In computation of GI, the GPU mainly enhances the convolution process in the computation of the kernel-based Gaussian derivative, which can be efficiently parallelized.

In cases involving multiple images, one issue is how to combine similarity measures computed from multiple image pairs into one objective function. Details of this issue are discussed in Section V. In this study, we employed a numerical sum approach, where the sum of the similarity measures was used as the objective function value.

6) Optimization—Two optimization algorithms, Nelder–Mead Downhill Simplex [43] and CMA-ES [51], [52] were implemented and compared using the experimental data in terms of evaluation criteria described below. CMA-ES is a stochastic search algorithm known for robustness and efficiency in a rugged earch landscape. The CMA-ES algorithm implemented in MATLAB by Hansen [53] was used in this study. CMA-ES was used in registration problems previously by Winter *et al.* [54] for CT-US registration and by Gong *et al.* [55] for 2D/3D registration where CMA-ES was compared with UKF-based and Simplex-based optimizers. In this paper, we compare the use of CMA-ES and Simplex-

based optimization, as well as choice of objective function and DRR generation methods, using clinically realistic datasets and conduct a competitive validation on a publicly available dataset. The optimization used a coarse-to-fine multiresolution approach, where initially the fixed images were down-sampled to generate a coarse resolution image and gradually the resolution was increased. The optimization was reinitialized at each multi-resolution level. In this study, three levels of resolution were used (down-sampling factors were 4, 2, 1, respectively). Fig. 8 graphically explains the multiresolution strategy that we employed in this paper.

Each optimizer requires the user to manually tune several parameters that define the initial conditions (i.e., initial simplex size for Downhill Simplex and initial search distribution (σ) for CMA-ES) and convergence criteria.

Although there are optimization approaches using the gradient of the objective function [56], these approaches also do not guarantee a global solution. Additionally, an analytical gradient of the GI objective function (4) and (5) is difficult to compute and numerical approximation of the derivative is computationally costly. Therefore, in this paper, we evaluated two non-gradient based approaches and focused more on the choice of objective function.

D. GPU Implementation

Computations of DRRs and similarity measures were implemented on the GPU. The optimization algorithm was coded in Matlab but included function calls to an externally compiled dynamic link library coded in C++ using CUDA v4.0 (NVIDIA Inc.). In order to minimize data transfer between CPU and GPU during the optimization, input data (2D X-ray image and CT) were transferred to GPU at the initialization step and resided in the GPU memory throughout the registration. We used texture memory for the volume data to reduce memory access latency and global memory for storing the line integral. The neighboring rays are traced in the same thread block to achieve better performance by increasing cache hits. Since the line integral was stored in a separate memory space for each ray and tracing of voxel space involves only read-access, the possibility of race conditions was not a concern.

E. Evaluation Method

In order to assess robustness and sensitivity of the proposed algorithm, we followed the method similar to that of Gendrin *et al.* [57]. We performed 100 registration trials using randomly selected initial guesses. The initial guesses were offsets of the known ground truth registration (described later) composed of randomly selected 6DOF registration parameters resulting in mean target registration error (mTRE) [58] from 0 to 10 mm with a step of 1 mm. Ten initial guesses were computed for each 1 mm mTRE interval. The mTRE, success rate, number of function evaluations, and computation times were used to evaluate the methods. While the mTRE is widely used in the literature, the definition of the target points differs from study to study. Some use predefined locations (e.g., fiducials, anatomical landmarks) or surface points [5], [16], [59] of an object, whereas others suggested a fixed set of uniformly distributed points on a grid as the target points [58]. In this study, we used target points on a grid of 1 mm interval, within a volume of 50 mm^3, of which the origin

was located at the center of the femoral head. This is because the experiments in this study were conducted in the context of an image-guided surgery system for femoral bone augmentation, where the target anatomical structures were the femoral head and surrounding tissues.

The methods for acquiring ground truth registration are described for each experiment in Section IV. In reporting the error of 2D/3D registration using multiple images, an orthogonal, right-handed coordinate system that was fixed with respect to the first image was employed throughout this study. Registrations with a final mTRE smaller than 2.5 mm were classified as successful [5], [20].

IV. Experiments and Simulation

Four sets of experiments were conducted to evaluate the proposed algorithm. The experiments were designed in such a way that the complexity of the problem increased step by step. Table I summarizes the types of experiments, fixed images, floating images (one example set used in the experiment), types of C-arm detector and calibration phantom, and the method for acquiring the ground truth for the evaluation. Table II summarizes the specifications (resolution and pixel size) of the dataset used in the experiments. The Sections IV-A–IV-I detail each experiment. Specifications of the hardware used in the experiments are listed in Table III.

A. Evaluation Using Gold Standard Dataset

In order to compare our implementation with the literature, an experiment was done using a gold standard dataset published by Pawiro *et al.* [60] and Gendrin *et al.* [57]. Among several types of dataset including CT, CBCT, MR, kV/MV X-rays that are provided by the paper, we chose kV X-ray and CT datasets for our evaluation considering our IGS application scenario. The 2D/3D registration trial was repeated 250 times using the same protocol as Gendrin *et al.* We did not use the ROI mask that the paper suggested to use to make the algorithm more robust. Omitting this mask resulted in significant reduction of manual preprocessing on the intraoperative X-ray image. We used the trilinear interpolation method with step length of 1.5 mm to generate DRRs, GI as a similarity measure, and the CMA-ES algorithm without multiresolution approach for optimization. To do an equivalent comparison with the literature, the registration success threshold was set to 5 mm (as opposed to 2.5 mm for our other validation experiments).

B. Plastic Phantom Experiment

The purpose of this experiment was to evaluate the proposed registration framework with an ideal dataset excluding noise caused by soft tissue and pose estimation error. We mounted the hybrid fiducial on a plastic femur phantom (Fig. 9) to evaluate the feasibility of the fixation device and visibility of the fiducial in the X-ray images. A preoperative CT data (see Table II for the detail) was acquired with a diagnostic CT, Aquilion 64 (Toshiba Medical Systems, Japan). We then acquired two hundred X-ray projection images over a 180 degree orbit using the calibrated CBCT and reconstructed a 3D volume of 512^3 voxels ($150\ mm^3$) from these images. The system's built-in calibration using a helix phantom as

described in Section III D was used to obtain the C-arm pose for each image with respect to the CBCT volume. This was done in order to achieve an ideal evaluation condition that eliminates additional sources of error (i.e., pose estimation error, variation in intrinsic parameters for each image, unintended movement of the fiducial with respect to the bone, etc.). The registration between the preoperative CT volume and the CBCT was computed by an intensity based 3D/3D registration with mutual information as the similarity measure. We selected one to four evenly spaced images out of the 200 acquired images to evaluate the registration accuracy. Here, we used the trilinear interpolation method with step length of 1.5 mm for DRRs, GI as a similarity measure, and the CMA-ES algorithm for optimization. The optimization parameters that we used are listed in Table IV.

C. Cadaver Experiment Using Cone Beam CT

In this experiment we used a cadaver specimen to evaluate the proposed algorithm in the presence of soft tissue. CT data was acquired preoperatively and four hundred X-ray images were acquired by the CBCT using the specifications listed in Table II. The DRR images generated using a ground truth registration based on intensity-based 3D/3D registration are shown in Table I. For a similar reason to the one described for Experiment 2, we mounted the FTRAC to evaluate feasibility of the fixation device and visibility of the fiducial features on a cadaver specimen and used the built-in CBCT calibration to achieve an ideal evaluation environment. This situation assumes relatively advanced imagers such as an advanced mobile C-arm or Zeego (Siemens, Germany) which are motorized and equipped with encoder feedback and the complete geometric calibration parameters for each image are available. We selected one to four images out of 400 acquired images in order to evaluate the effect of number of images and range of view angle on the accuracy of the registration. Similar to Experiment 2, we used the trilinear interpolation method with step length of 1.5 mm, GI, and the CMA-ES algorithm for the registration.

D. Cadaver Experiment Using Conventional C-Arm

We evaluated the proposed algorithm using the conventional and non-motorized C-arm imager using a cadaver with soft tissue. Our motivation for this experiment was the continued prevalence of XRII-based conventional C-arms in most current routine surgeries. We acquired the ground truth registration from a preoperative CT taken with the X-ray fiducial attached (Fig. 10) by applying rigid-body point-based registration [61] to the nine BBs of the fiducial. In a realistic clinical application scenario, the hybrid fiducial would be mounted only during the operation in a less invasive manner. However, in this experiment, in order to establish a highly accurate ground truth registration, the hybrid fiducial was rigidly mounted to the middle of the femur bone using a bone screw prior to the experiment. We acquired two arbitrarily selected X-ray images from different directions while attempting to achieve as broad an angular separation as possible. In a realistic surgical setting of peripheral devices, angular separation is limited due to the limitation of the range of movement of the C-arm and obstruction by the operating table. In this experiment the angular separation between the two acquired images was 31.5°. We estimated the C-arm pose with respect to the fiducial using the method described in Section III-C1. In order to evaluate the system under a clinically realistic protocol, we calibrated intrinsic C-arm parameters only once prior to the experiment and used the same parameters for all of the

images used for pose estimation and registration. In this particular experiment, in order to cope with the low resolution and extremely noisy intraoperative X-ray images, we segmented the bone region in the CT data by thresholding (> 350 HU) prior to the 2D/3D registration. This experiment evaluated the proposed framework in the presence of several error sources as expected in a practical surgical scenario in a conventional operating theater, such as image distortion due to XRII, geometric calibration error coming from the fiducial-based pose estimation and pose-dependent changes of the intrinsic C-arm parameters. We used a similar approach to the previous two experiments to evaluate the accuracy of the 2D/3D registration.

E. Evaluation of Error Contribution in Each System Component

We also independently evaluated DRR generation algorithms, similarity measures, optimization strategies, and optimization algorithms to investigate their contributions to the registration error. 100 registration trials were conducted for each case following the protocol explained in Section III.

F. Comparison Between Different DRR Generation Algorithms

As we described earlier, there is a trade-off between accuracy and computational cost in DRR generation using Siddon's algorithm and the trilinear interpolation algorithm. In order to evaluate the performance of each algorithm, a series of benchmark tests was conducted. A CT volume of a plastic bone phantom was used for the benchmark. In order to test the performance under higher computational load, we used a 512^3 volume (0.3 mm^3/voxel) instead of a cropped volume of 256^3 (0.3 mm^3/voxel) as used in Experiment 2. In this test, DRRs computed with Siddon's algorithm were used as a ground truth and image quality of the DRRs computed by the trilinear interpolation algorithm with various step lengths was evaluated by computing their absolute difference from the ground truth and required computation time. Since both algorithms traverse every voxel along a ray, computation time of DRR depends on the length that each ray travels through the volume, which is determined by the orientation of the volume with respect to the detector. To find the trend, a series of DRR images were generated for different view directions spanning all around the volume (360 view angles on a 1° interval), and the mean computation time for 1800 images (five rotations) for various image sizes was computed. The effect of the image quality difference on the final registration error was then evaluated using data acquired in the three experiments described above. The evaluation protocol followed the one described in Section III-D. Two images were used in the registration, GI was used as a similarity measure, and the CMA-ES with multilevel strategy was used for optimization.

G. Comparison Between Similarity Measures

We conducted the same evaluation protocol for three different types of similarity measures, MI, NMI, and GI using the data acquired in the three experiments. We used two images, trilinear interpolation with 1.5 mm step length for DRR generation, and CMA-ES with multilevel strategy for optimization.

H. Effect of Coarse-to-Fine Multi Level Optimization Strategy

In order to evaluate the effect of the coarse-to-fine multilevel strategy in the optimization process, we compared the registration result with and without the strategy. We used two images, trilinear interpolation with 1.5 mm step length, GI, and CMA-ES for optimization.

I. Comparison Between Optimization Algorithms

We also compared two optimization algorithms, Nelder-Mead Downhill Simplex [43] and CMA-ES [51], [52], using the same protocol and data with trilinear interpolation with 1.5 mm step length, GI, and multilevel strategy.

V. Results

A. Evaluation Using Gold Standard Dataset

The results of the experiment using the gold standard dataset are summarized in Table V. The proposed method using GI as a similarity measure showed lower mean and standard deviation of mTRE than the reported result which used MI and BGR, and significantly lower registration time than reported gradient-based method (the backprojection gradient-based registration, BGB, and the reconstruction gradient-based registration, RGB). Both GI and MI in the proposed method showed lower standard deviation of mTRE than the reported results which suggests an advantage of the optimizer that we used (CMA-ES) over the one used in the literature (Powell's method). The soft tissue deformation of the jaw between CT scanning and X-ray image acquisition (see Table I) caused the larger mTRE in MI in the proposed method, whereas GI was more robust to the deformation.

B. Plastic Phantom Experiment

The results of the experiment using a plastic bone phantom are summarized in Table VI. Fig. 11 shows a plot representing the initial and final mTRE of each registration trial. The average mTRE using a single image showed the largest error in terms of mean and standard deviation. Thirty one percent (31%) of the trials failed when using one image. The success rate, however, was 100% when more than one image was used. The computation time increased in approximately linear proportion to the number of images.

C. Cadaver Experiment Using Cone Beam CT

Table VI summarizes the results of the cadaveric experiment using CBCT. All registration trials failed when only one image was used. As seen in Table I, the femur bone in the image from Experiment 3 was obscured by the soft tissue. Since the images from the view close to lateral direction showed almost no contrast between the femur bone and the surrounding soft tissue, we confined our analysis to the images over approximately 90° and selected equally spaced sets of 2, 3, and 4 images (in Table VI   shows the angle between consecutive images). Using 2, 3, and 4 images resulted in an average error in the range of 0.8–1.0 mm and a success rate of greater than 80%. Similar to the previous experiment, the computation time increased in proportion to the number of images.

D. Cadaver Experiment Using Conventional C-Arm

Results of the cadaver experiment using a conventional C-arm are summarized in Table VI. Similar to the previous cadaver experiment, single-image 2D/3D registration failed in all trials. The error with two images was 2 times larger than the previous flat-panel C-arm experiment suggesting that the error was caused by inaccuracies due to using the conventional C-arm device.

E. Evaluation of Error Contribution in Each System Component

We evaluated the effect of using different DRR generation algorithms, similarity measures, multilevel approaches, and optimization algorithms on the accuracy of the 2D/3D registration. Fig. 12 summarizes these results.

F. Comparison Between Different DRR Generation Algorithms

As shown in Fig. 12(a), the accuracy of the registration depends on the DRR generation algorithm (S in Fig. 12(a) indicates Siddon's method, and the numbers 0.15–7.0 represent the trilinear interpolation algorithm using the indicated step length). Experiment 3 showed a clear trend of improvement of the accuracy with smaller step length. Fig. 13 shows the differences in image quality between Siddon's algorithm and the trilinear interpolation algorithm with 3 different step lengths, 0.15 mm, 1.5 mm, and 3.0 mm. From the subtraction of images (lower row), a slight difference can be seen in the 1.5 mm step length while a significant difference can be observed in the 3.0 mm step length image. Fig. 14 shows quantitative results for the same benchmark tests. The absolute difference value (thick line) clearly shows a rapid increase of quality below the 1.5 mm step length, which is 5 times larger than the edge length of a voxel, while computation time shows a plateau above 1.5 mm.

G. Comparison Between Similarity Measures

In Fig. 15, we plotted the similarity measure values as a function of the deviation from the ground truth for each pose parameter (translation and rotation). As shown here, in a simple plastic bone experiment, all three similarity measures work desirably (i.e., smooth landscape with distinct local maxima), whereas the rugged landscape can be seen in cadaver images, likely due to the added image noise associated with the presence of soft tissue. Registration using three different similarity measures [Fig. 12(b)] supports the above-mentioned observation showing the trend where the registration algorithm was more robust in the plastic bone experiment than in the cadaver experiments (Experiments 3 and 4). The box plots show an apparent advantage of GI in cadaver experiments compared to MI and NMI based on the much lower mTRE.

H. Effect of Coarse-to-Fine Multi Level Optimization Strategy

The upper graph in Fig. 12(c) shows comparison between the results using the data from Experiment 3 (cadaver with flat-panel C-arm) with and without the coarse-to-fine scheme. The results of Experiments 2 and 3 showed statistically significant improvement ($P < 0.001$, $P = 0.040$ respectively) by introducing the multi level scheme. However, Experiment 4

showed statistically significant decrease of the accuracy ($P < 0.001$) suggesting that on noisier data sets, multilevel optimization may enlarge the error in the low-resolution stages.

I. Comparison Between Optimization Algorithms

Fig. 12(d) shows the result of comparison of the optimization algorithms. In Experiment 2, all trials were successful in both algorithms, whereas in Experiment 3, CMA-ES achieved a higher success rate (81%) than Downhill Simplex (76%), a smaller variance (6.9 mm versus 18.9 mm), and showed statistically significant difference ($P = 0.020$) in the average error. In Experiment 4, both algorithms showed low success rates (45% and 26%) and CMA-ES showed lower average error ($P = 0.020$). Computation time of the CMA-ES algorithm was nearly double that of the Downhill Simplex.

VI. Discussion

In this paper, we have presented a novel framework for intraoperative image-based 2D/3D registration incorporating fiducial-based C-arm tracking and GPU-acceleration. We have evaluated the accuracy and feasibility of the proposed framework using laboratory experiments varying from an experiment on a gold standard dataset to clinically realistic experiments using cadavers.

The results of Experiment 1 (Table V) showed superior performance of our implementation to the literature in terms of registration time and accuracy, specifically in the GI similarity metric.

Despite clear differences in the DRR quality between the two DRR generation algorithms shown in Fig. 14, Fig. 12(a) suggests that the registration result varies by the target images. As shown in Table I, the fixed images are covered with a black background in Experiment 2 where a plastic bone phantom was used, and the factor that dominates registration results was the edges of the bone, i.e., intensity difference between air and the plastic rather than the detailed texture within the bone. The step length, therefore, did not affect the registration result. In Experiment 3, since soft tissues surrounding the hip covered large parts of the X-ray image, the quality of the DRR defined by the step length affected the registration results [Fig. 12(a)] upper graph). Siddon's ray tracing algorithm showed the best result with approximately 5 times increase in computation time when compared to the trilinear interpolation algorithm with step length of 1.5–7.0 mm. In Experiment 4, since we extracted the bone region to enhance the edge of the bone, the step length did not have a large effect on the result, similar to the experiment using plastic bone. The average computation time for one registration trial [red marker in the lower graph of Fig. 12(a)] decreased as the step length increased in all experiments. Reducing the ray tracing step length to less than 1.5 mm had little effect on the registration quality. We attribute this effect to the fact that the the contrast resolution in the X-ray image (fixed image) was larger than the subtle quality difference in the floating image, which resulted in no improvement in the shape of the objective function. Although DRR generation using double precision arithmetic has the potential to improve quality of the floating image, this level of subtle improvement is unlikely to significantly affect the registration result.

As other researchers have noted (e.g., [12], [59], [62]) the choice of similarity measure has a great impact on the registration result. However, to our knowledge, a decisively versatile similarity measure that works for any type of 2D/3D registration problem has not yet been introduced. Furthermore, several comparative studies have reported inconsistent results. As suggested in previous studies, the performance of the registration is heavily dependent on the application (e.g., anatomical target region) and the quality of both 2D images and 3D volume. In this study, we evaluated GI, MI and NMI for 3 different types of experimental data. We consider that one reason why GI performed superior to MI and NMI was that the smoothing effect by Gaussian gradient suppressed adverse effects by the quantum noise in the X-ray images (fixed images). Also as mentioned in Section III-C-V, MI and NMI suffered from the external object (fiducial fixture device), whereas GI eliminated the distracting effect by taking into account only the pixels that have a large gradient magnitude in both fixed and floating images. Therefore, for our application scenario, the GI approach outperformed the MI and NMI approaches in all experiments.

One additional consideration arises when multiple images are used in one 2D/3D registration trial. Although we can compute the similarity measure for each image pair, the question is how we should merge those similarity measures to establish one objective function for the optimization.

Approaches to remedy this problem can be classified into three categories. The first category is a numerical sum approach [5], [8], [10], [38], [46], [63], where the sum of the similarity measures for each image pair is considered as the objective function to maximize. The next category uses a composite approach [12], where multiple fixed images and multiple floating images are considered as one large image and one similarity measure between the large images is computed. The last category is an alternating approach [66], where the similarity measure of only one image pair is used in one iteration of optimization to compute the next estimate. Then the target image pair is changed in the next iteration and repeatedly alternated in successive iterations. In this paper, we used the numerical sum approach. The comparison between those approaches is in the scope of our future research.

The experimental results suggested a significant advantage with the multi level strategy in Experiments 2 and 3. We believe that the reason why Experiment 4 did not show an improvement was due to the inaccuracy caused by the error-prone nature of XRII C-arm.

The statistically significant advantage of CMA-ES in Experiment 3 suggests that robustness of 2D/3D registration in the presence of soft tissue increased by using the CMA-ES algorithm. Whereas in Experiment 4, similar to the previous section, the error due to using XRII dominated the effect of using different optimization algorithms. The difference in computation time was due to the number of function evaluations required by each algorithm, i.e., genetic algorithms generally require more function evaluation. Although our focus in this paper was on the choice of objective function as mentioned in Section III, one of the possible future extensions of the work includes comparative study with other types of optimization methods using the first-order derivative of the objective functions. A comparative analysis could assess the trade-off between increasing the optimization efficiency using gradient information, versus the additional computation time of computing

the gradient at each optimization step. The first order derivative of MI and NMI was derived in [67]. For GI, for which it is hard to derive an analytical form of the first order derivative, finite difference based approaches can be employed.

The larger average error (2.04 mm) despite the small standard deviation (0.85 mm) in Experiment 4 implies that the majority of registration trials converged to a similar position which was different from the ground truth. This suggests that part of the error might be the result of error in the ground truth registration. One potential source of error is the movement of the fiducial with respect to the femur bone between preoperative CT scanning and the intraoperative X-ray capturing during the transportation process of the specimen from diagnostic CT to the operating table.

Tracking of the patient movement is an unavoidable requirement in any type of IGS system unless a decent (sometimes invasive) patient fixation device (e.g., head clamp, screw-type leg fixation device) is used. Depending on the types of patient fixation, we can consider a looser fixation using a small thin bone screw or noninvasive fixation, such as attachment to the patient's skin or to the patient fixation device.

While it was feasible to fixate the hybrid fiducial to the bone using a small bone pin in our experiments, as an alternative approach, we may consider that the fast computation time achieved by the proposed framework offers a possibility of repeating registration as the need arises (i.e., right before each important step in the surgery, provided that the patient's target organ does not move during the procedure). This may eliminate the need for attaching the fiducial tightly to the bone.

Nevertheless, we consider the size and weight of our current fiducial and the fixation device (which required a relatively large incision and tight fixation to the bone) to be a limitation of our current implementation that must undergo design optimization before application in a clinical environment. Future work will include exploring the optimal design of the fiducial structure, with special attention to size and weight. As is the case with most IGS systems, patient tracking with a less invasive fiducial fixation is an important consideration in clinical application.

Although our experiment was targeted to an orthopaedic surgery application, we believe the methods reported here can be applied to broader applications of image-based registration including craniomaxillofacial surgery, neurosurgery, and dentistry.

VII. Conclusion

This paper demonstrated the 2D/3D registration framework using fiducial-based pose estimation and GPU-acceleration. By combining the C-arm tracking and intensity-based 2D/3D registration, we achieved an intraoperative image-based multiview registration using a conventional non-motorized C-arm. As a result of extensive experimental validations, it has been shown that the intraoperative registration using 2 X-ray images could be completed within about 15 s with reasonable accuracy and robustness without manual preoperative processing. Based on the analysis described, the registration was performed most effectively when using GI as the similarity metric, Siddon's ray tracing algorithm, and the CMA-ES

algorithm with coarse-to-fine multilevel approach for optimization. The trilinear interpolation ray tracing algorithm improves the computation time while slightly sacrificing the registration accuracy.

In summary, we believe that the integrated framework proposed here addresses the key barriers that have precluded clinical use of image based registration, and the potential to enable routine use of minimally-invasive intensity-based 2D/3D registration.

Acknowledgment

References

1. Widmann G, Stoffner R, Bale R. Errors and error management in image-guided craniomaxillofacial surgery. Oral Surg., Oral Med., Oral Pathol., Oral Radiol., Endodontol. 2009; 107:701–715.

2. Kessler ML. Image registration and data fusion in radiation therapy. Br. J. Radiol. 2006 Sep. 79:S99–S108. [PubMed: 16980689]

3. Balter JM, Kessler ML. Imaging and alignment for image-guided radiation therapy. J. Clin. Oncol. 2007 Mar.25:931–937. [PubMed: 17350941]

4. Dawson LA, Jaffray DA. Advances in image-guided radiation therapy. J. Clin. Oncol. 2007 Mar. 25:938–946. [PubMed: 17350942]

5. Russakoff DB, Rohlfing T, Adler JR Jr, Maurer CR Jr. Intensity-based 2D–3D spine image registration incorporating a single fiducial marker. Acad. Radiol. 2005 Jan.12:37–50. [PubMed: 15691724]

6. Daly MJ, Siewerdsen JH, Cho YB, Jaffray DA, Irish JC. Geometric calibration of a mobile C-arm for intraoperative cone-beam CT. Med. Phys. 2008 May.35:2124–2136. [PubMed: 18561688]

7. Otake Y, Armand M, Sadowsky O, Armiger RS, Kutzer MD, Mears SC, Kazanzides P, Taylor RH. An image-guided femoroplasty system: Development and initial cadaver studies. Proc. SPIE. 2010 Mar.7625:76250P–76250P.

8. Jomier J, Bullitt E, Van Horn M, Pathak C, Aylward SR. 3D/2D model-to-image registration applied to TIPS surgery. Med. Image Comput. Comput. Assist. Interv. Int. Conf. Med. Image Comput. Comput. Assist. Interv. 2006; 9:662–669.

9. Khamene A, Bloch P, Wein W, Svatos M, Sauer F. Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy. Med. Image Anal. 2006 Feb.10:96–112. [PubMed: 16150629]

10. Turgeon GA, Lehmann G, Guiraudon G, Drangova M, Holdsworth D, Peters T. 2D–3D registration of coronary angiograms for cardiac procedure planning and guidance. Med. Phys. 2005 Dec.32:3737–3749. [PubMed: 16475773]

11. Weese J, Penney GP, Desmedt P, Buzug TM, Hill DL, Hawkes DJ. Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery. IEEE Trans. Inf. Technol. Biomed. 1997 Dec.1:284–293. [PubMed: 11020832]

12. Penney GP, Weese J, Little JA, Desmedt P, Hill DLG, hawkes DJ. A comparison of similarity measures for use in 2-D-3-D medical image registration. IEEE Trans. Med. Imag. 1998 Aug. 17(no. 4):586–595.

13. Chen X, Gilkeson RC, Fei B. Automatic 3D-to-2D registration for CT and dual-energy digital radiography for calcification detection. Med. Phys. 2007 Dec.34:4934–4943. [PubMed: 18196818]

14. Zheng G, Steppacher S, Zhang X, Tannast M. Precise estimation of postoperative cup alignment from single standard X-ray radiograph with gonadal shielding. Med. Image Comput. Comput. Assist. Interv. Int. Conf. Med. Image Comput. Comput. Assist. Interv. 2007; 10:951–959.

15. Zheng G. Statistical shape model-based reconstruction of a scaled, patient-specific surface model of the pelvis from a single standard AP X-ray radiograph. Med. Phys. 2010 Apr.37:1424–1439. [PubMed: 20443464]

16. Livyatan H, Yaniv Z, Joskowicz L. Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. IEEE Trans. Med. Imag. 2003 Nov.22(no. 11):1395–1406.

17. Sadowsky O, Chintalapani G, Taylor RH. Deformable 2D–3D registration of the pelvis with a limited field of view, using shape statistics. Med. Image Comput. Comput. Assist. Interv. Int. Conf. Med. Image Comput. Comput. Assist. Interv. 2007; 10:519–526.

18. Taylor RH, Joskowicz L, Williamson B, Gueziec A, Kalvin A, Kazanzides P, Van Vorhis R, Yao J, Kumar R, Bzostek A, Sahay A, Borner M, Lahmer A. Computer-integrated revision total hip replacement surgery: Concept and preliminary results. Med. Image Anal. 1999 Sep.3:301–319. [PubMed: 10710298]

19. Jain A, Fichtinger G. C-arm tracking and reconstruction without an external tracker. Med. Image Comput. Comput. Assist. Interv. Int. Conf. Med. Image Comput. Comput. Assist. Interv. 2006; 9:494–502.

20. Jain AK, Mustafa T, Zhou Y, Burdette C, Chirikjian GS, Fichtinger G. FTRAC—A robust fluoroscope tracking fiducial. Med. Phys. 2005 Oct.32:3185–3198. [PubMed: 16279072]

21. Russakoff, DA.; Rohlfing, T.; Maurer, CR, Jr. Fast intensity-based 2D–3D image registration of clinical data using light fields; Proc. 9th IEEE Int. Conf. Comput. Vis. (ICCV); 2003. p. 416-423.

22. Russakoff DB, Rohlfing T, Mori K, Rueckert D, Ho A, Adler JR Jr, Maurer CR Jr. Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D–3D image registration. IEEE Trans. Med. Imag. 2005 Nov.24(no. 11):1441–1454.

23. LaRose D, Bayouth J, Kanade T. Transgraph: Interactive intensity-based 2D/3D registration of X-ray and CT data. Proc. SPIE. 2000 Jun.3979:385–396.

24. Tomazevic D, Likar B, Slivnik T, Pernus F. 3-D/2-D registration of CT and MR to X-ray images. IEEE Trans. Med. Imag. 2003 Nov.22(no. 11):1407–1416.

25. Tomazevic D, Likar B, Pernus F. 3D/2D image registration: The impact of X-ray views and their number. Med. Image Comput. Comput. Assist. Interv. Int. Conf. Med. Image Comput. Comput. Assist. Interv. 2007; 10:450–457.

26. Ruijters, D.; Haar-Romeny, BMT.; Suetens, P. GPU-accelerated digitally reconstructed radiographs; Proc. 6th IASTED Int. Conf. Biomed. Eng. Innsbruck, Austria; 2008. p. 431-435.

27. Cardoso, A. M.S. thesis. Porto, Portugal: Univ. Porto; 2010. Generation of planar radiographs from 3D anatomical models using the GPU.

28. Xu F. Fast implementation of iterative reconstruction with exact ray-driven projector on GPUs. Tsinghua Sci. Technol. 2010; 15:30–35.

29. Yang, Y.; Zhong, Z.; Rong, G.; Guo, X.; Wang, J.; Solberg, T.; Mao, W. Real-time GPU-aided lung tumor tracking; Proc. 4th Pacific-Rim Symp. Image Video Technol. (PSIVT); 2010. p. 495-500.

30. Cabral, B.; Cam, N.; Foran, J. Accelerated volume rendering and tomographic reconstruction using texture mapping hardware; Proc. 1994 Symp. Volume Visualizat; 1994. p. 91-98.

31. Kubias A, Deinzer F, Feldmann T, Paulus D, Schreiber B, Brunner T. 2D/3D image registration on the GPU. Pattern Recognit. Image Anal. 2008; 18:381–389.

32. Plastimatch. [Online]. Available: http://plastimatch.org.

33. VTK User's Guide Version 11. Clifton Park, NY: Kitware; 2010.

34. Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. Med. Phys. 1985 Mar-Apr;12:252–255. [PubMed: 4000088]

35. de Greef M, Crezee J, van Eijk JC, Pool R, Bel A. Accelerated ray tracing for radiotherapy dose calculations on a GPU. Med. Phys. 2009 Sep.36:4095–4102. [PubMed: 19810482]

36. Siewerdsen JH, Moseley DJ, Burch S, Bisland SK, Bogaards A, Wilson BC, Jaffray DA. Volume CT with a flat-panel detector on a mobile, isocentric C-arm: Pre-clinical investigation in guidance of minimally invasive surgery. Med. Phys. 2005 Jan.32:241–254. [PubMed: 15719975]

37. Sadowsky, O. Ph.D. dissertation. Baltimore: Johns Hopkins Univ.; 2008. Image registration and hybrid volume reconstruction of bone anatomy using a statistical shape atlas.

38. Sadowsky O, Lee J, Sutter EG, Wall SJ, Prince JL, Taylor RH. Hybrid cone-beam tomographic reconstruction: Incorporation of prior anatomical models to compensate for missing data. IEEE Trans. Med. Imag. 2011 Jan.30(no. 1):69–83.

39. Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. Int. J. Radiat. Oncol. Biol. Phys. 2002; 53:1337–1349. [PubMed: 12128137]

40. Fallavollita P, Burdette EC, Song DY, Abolmaesumi P, Fichtinger G. Technical note: Unsupervised C-arm pose tracking with radiographic fiducial. Med. Phys. 2011; 38:2241–2245. [PubMed: 21626958]

41. Fallavollita, P.; Burdette, C.; Song, D.; Abolmaesumi, P.; Fichtinger, G. C-Arm Tracking by Intensity-Based Registration of a Fiducial in Prostate Brachytherapy. Vol. 6135. Berlin, Germany: Springer; 2010. p. 45-55.

42. Dementhon DF, Davis LS. Model-based object pose in 25 lines of code. Int. J. Comput. Vis. 1995; 1415:123–141.

43. Nelder JA, Mead R. A simplex method for function minimization. Comput. J. 1965 Jan.7:308–313.

44. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imag. 1997 Apr.16(no. 2):187–198.

45. Prince, JL.; Links, JM. Medical Imaging Signals and Systems. Upper Saddle River, NJ: Pearson Prentice Hall; 2006.

46. Zollei L, Grimson E, Norbash A, Wells W. 2D–3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators. Comput. Vis. Pattern Recognit. 2001; 2:II-696–II-703.

47. Spoerk J, Bergmann H, Wanschitz F, Dong S, Birkfellner W. Fast DRR splat rendering using common consumer graphics hardware. Med. Phys. 2007 Nov.34:4302–4308. [PubMed: 18072495]

48. Studholme C, Hill DLG, Hawkes DJ. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognit. 1999; 32:71–86.

49. Pluim JP, Maintz JB, Viergever MA. Image registration by maximization of combined mutual information and gradient information. IEEE Trans. Med. Imag. 2000 Aug.19(no. 8):809–814.

50. Otake Y, Armand M, Sadowsky O, Armiger R, Kazanzides P, Taylor R. An iterative framework for improving the accuracy of intraoperative intensity-based 2D/3D registration for image-guided orthopedic surgery. Lecture Notes in Computer Science. 2010; 6135:23–33.

51. Hansen, N.; Kern, S. Parallel Problem Solving From Nature. New York: Springer; 2004. Evaluating the CMA evolution strategy on multimodal test functions; p. 282-291.{PPSN VIII}.

52. Hansen N, Niederberger ASP, Guzzella L, Koumoutsakos P. A method for handling uncertainty in evolutionary optimization with an application to feedback control of combustion. IEEE Trans. Evolutionary Computat. 2009 Feb.13(no. 1):180–197.

53. Hansen, N. The CMA evolution strategy: A comparing review. In: Lozano, JA.; Larranaga, P.; Inza, I.; Bengoetxea, E., editors. Towards a New Evolutionary Computation. Advances on Estimation of Distribution Algorithms. New York: Springer; 2006. p. 75-102.

54. Winter S, Brendel B, Pechlivanis I, Schmieder K, Igel C. Registration of CT and intraoperative 3-D ultrasound images of the spine using evolutionary and gradient-based methods. IEEE Trans. Evolutionary Computat. 2008 Jun.12(no. 3):284–296.

55. Gong RH, Abolmaesumi P. 2D/3D registration with the CMA-ES method. Proc. SPIE. 2008 Mar. 6918:69181M–69181M.

56. Fletcher, R. Practical Methods of Optimization. New York: Wiley; 1981.

57. Gendrin C, Markelj P, Pawiro SA, Spoerk J, Bloch C, Weber C, Figl M, Bergmann H, Birkfellner W, Likar B, Pernuš F. Validation for 2D/3D registration II: The comparison of intensity- and gradient-based merit functions using a new gold standard data set. Med. Phys. 2011; 38:1491–1512. [PubMed: 21520861]

58. van de Kraats EB, Penney GP, Tomazevic D, van Walsum T, Niessen WJ. Standardized evaluation methodology for 2-D-3-D registration. IEEE Trans. Med. Imag. 2005 Sep.24(no. 9):1177–1189.

59. Birkfellner W, Stock M, Figl M, Gendrin C, Hummel J, Dong S, Kettenbach J, Georg D, Bergmann H. Stochastic rank correlation: A robust merit function for 2D/3D registration of image data obtained at different energies. Med. Phys. 2009 Aug.36:3420–3428. [PubMed: 19746775]

60. Pawiro SA, Markelj P, Pernuš F, Gendrin C, Figl M, Weber C, Kainberger F, Nöbauer-Huhmann I, Bergmeister H, Stock M, Georg D, Bergmann H, Birkfellner W. Validation for 2D/3D registration I: A new gold standard data set. Med. Phys. 2011; 38:1481–1510. [PubMed: 21520860]

61. Horn BKP. Closed-form solution of absolute orientation using unit quaternions. J. Opt. Soc. Am. A. 1987; 4:629–642.

62. Wu J, Kim M, Peters J, Chung H, Samant SS. Evaluation of similarity measures for use in the intensity-based rigid 2D–3D registration for patient positioning in radiotherapy. Med. Phys. 2009 Dec.36:5391–5403. [PubMed: 20095251]

63. Lemieux L, Jagoe R, Fish DR, Kitchen ND, Thomas DG. A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. Med. Phys. 1994 Nov. 21:1749–1760. [PubMed: 7891637]

64. Bodensteiner C, Darolti C, Schumacher H, Matthaus L, Schweikard A. Motion and positional error correction for cone beam 3D-reconstruction with mobile C-arms. Med. Image Comput. Comput. Assist. Interv. Int. Conf. Med. Image Comput. Comput. Assist. Interv. 2007; 10:177–185.

65. Chintalapani G, Ellingsen LM, Sadowsky O, Prince JL, Taylor RH. Statistical atlases of bone anatomy: Construction, iterative improvement and validation. Med. Image Comput. Comput. Assist. Interv. Int. Conf. Med. Image Comput. Comput. Assist. Interv. 2007; 10:499–506.

66. Leventon ME, Wells WM III, Grimson WEL. Multiple view 2D–3D mutual information registration. Proc. Image Understand. Workshop. 1997:625–630.

67. Maes F, Vandermeulen D, Suetens P. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. Med. Image Anal. 1999; 3:373–386. [PubMed: 10709702]
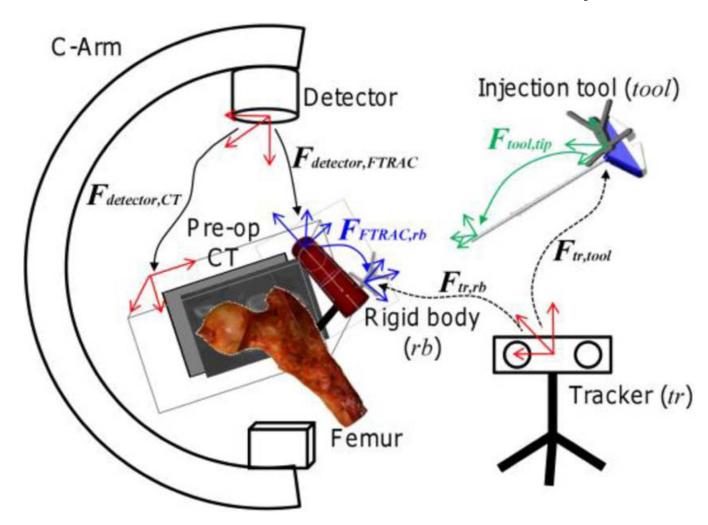
Fig. 1.
Exemplar application scenario in image-guided femoral bone augmentation surgery. Solid arrows mark preoperatively calibrated transformations, dashed arrows mark transformations that are updated intraoperatively. An arrow from coordinate system a to coordinate system b denotes the transformation $F_{a,b}$.
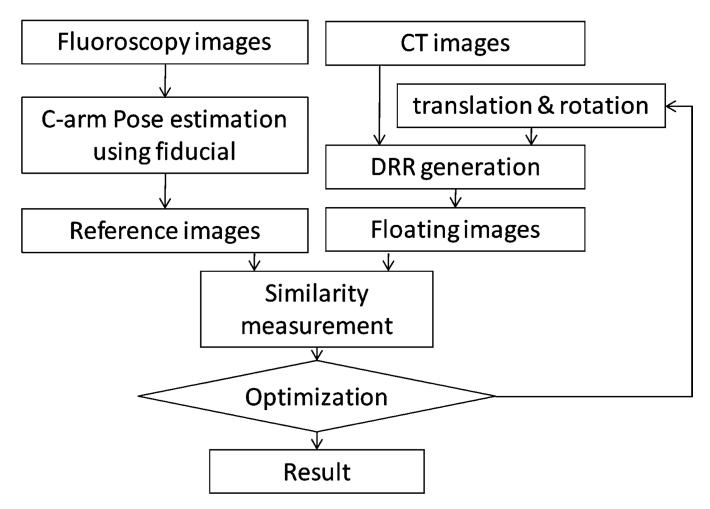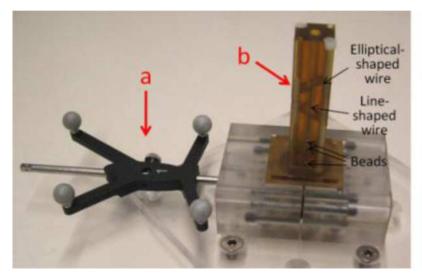
Fig. 2.
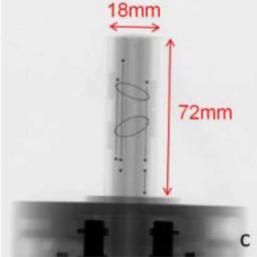Workflow of the intensity-based 2D/3D registration.

Fig. 3.
Hybrid fiducial. a: Optical tracking fiducial. b: Fluoroscopy tracking fiducial. c: An example X-ray image of the fluoroscopy tracking fiducial.
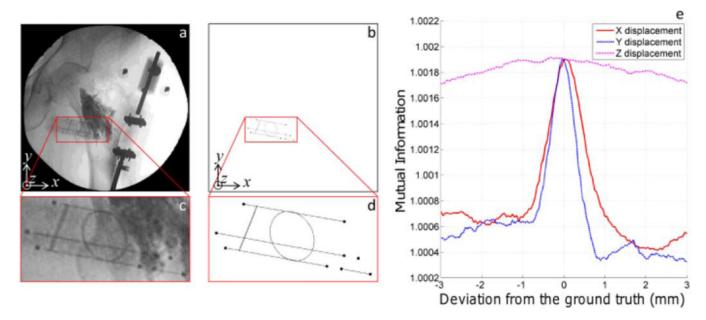
Fig. 4.
C-Arm fluoroscopy pose estimation algorithm. a: Original X-ray image. b: Projection image of the CAD model. c–d: Magnified image. e: Plot of mutual information (cost function) between the two images as a function of x, y, and z translational deviation from the ground truth.
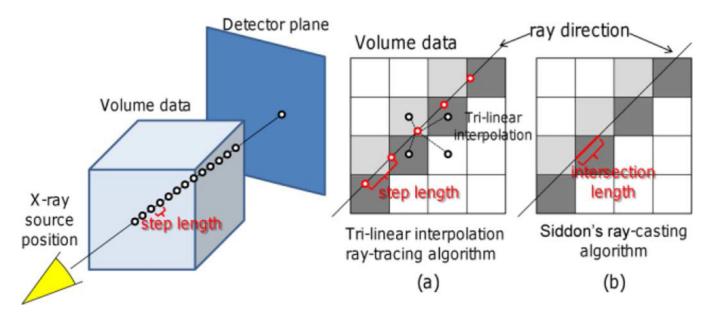
Fig. 5.
Two different types of DRR generation algorithms.
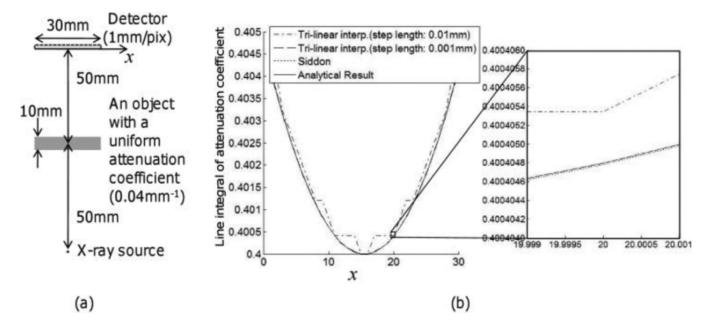
Fig. 6.
Simple 2D example illustrating differences of accuracy in computation of the line integral. a: Setup of detector and X-ray source in the example. b: Results of computation (horizontal axis: position of the detector element, vertical axis: computed line integral). See text for details.
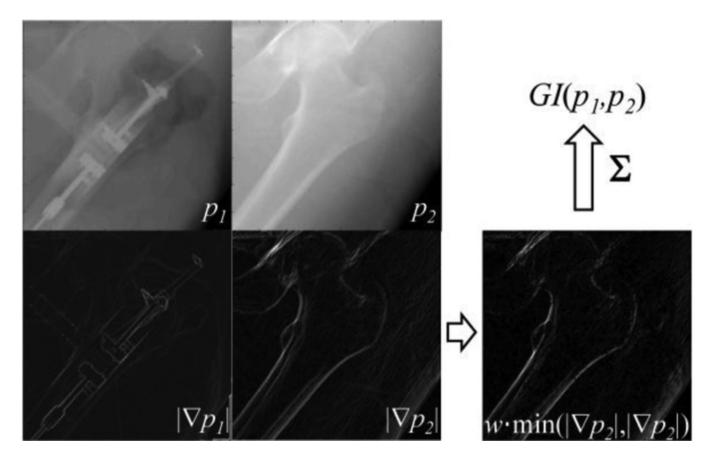
Fig. 7.
Graphical representation of computation of the gradient information (GI) similarity measure.
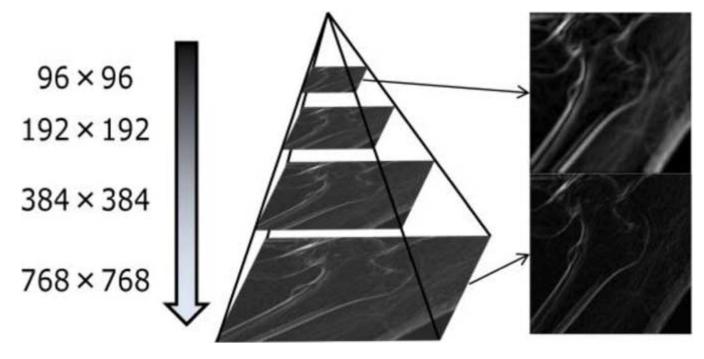p_1: fixed image (X-ray), p_2: floating image (DRR), w: weight (see text).

Fig. 8.
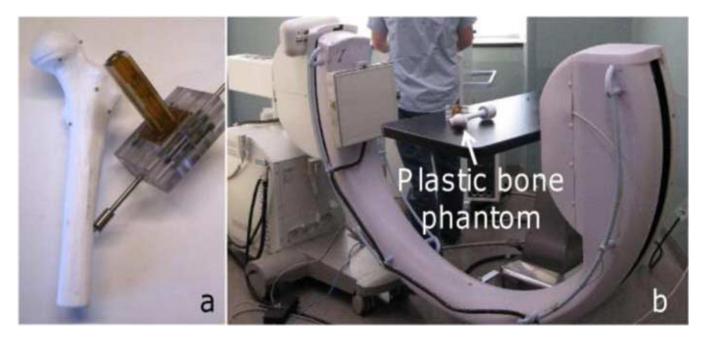Coarse-to-fine multiresolution optimization strategy.

Fig. 9.
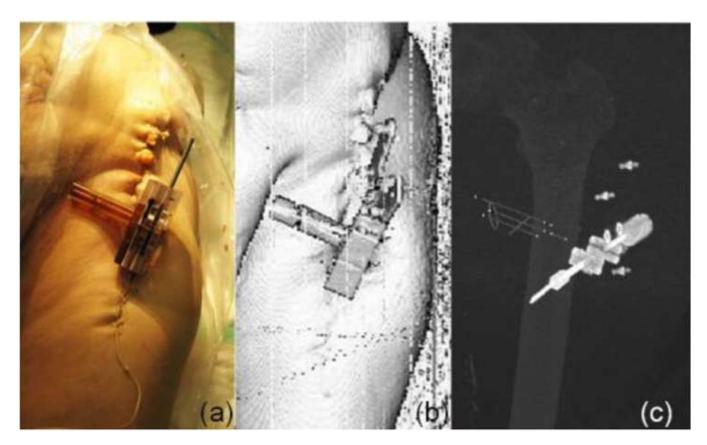Phantom experiment. a) Sawbone phantom with hybrid fiducial. b) Experimental setup.

Fig. 10.
Cadaver phantom used for Experiment 4. a: Specimen. b: Volume rendering. c: Maximum intensity projection images of CT data.
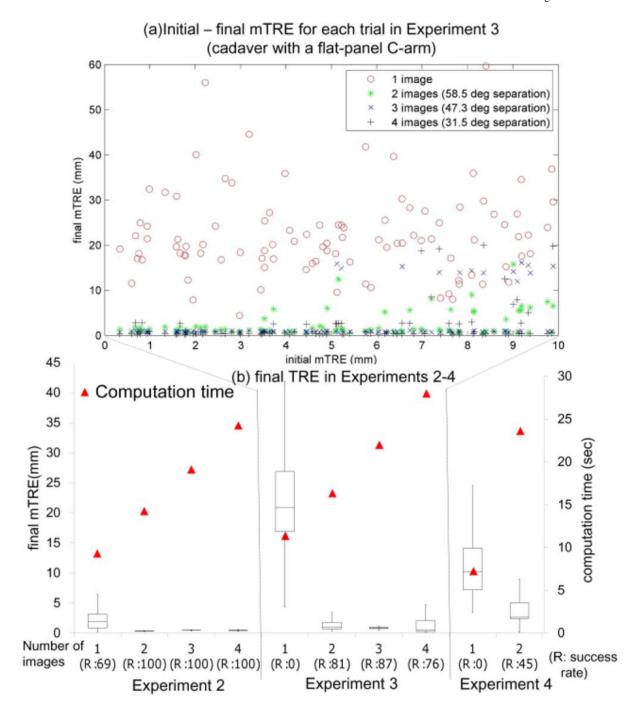
Fig. 11.
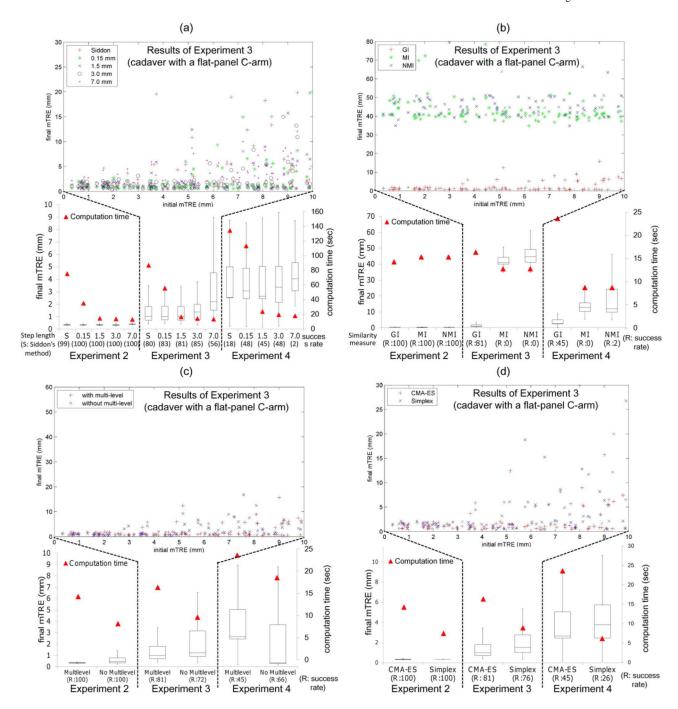Results for Experiments 2–4 with varying number of images.

Fig. 12.
Results of 2D/3D registration under various experimental conditions (see text for the detail). (a) Comparison between various step lengths. (b) Comparison between various similarity measures. (c) Effect of coarse-to-fine multilevel approach. (d) Comparison between optimization algorithms.
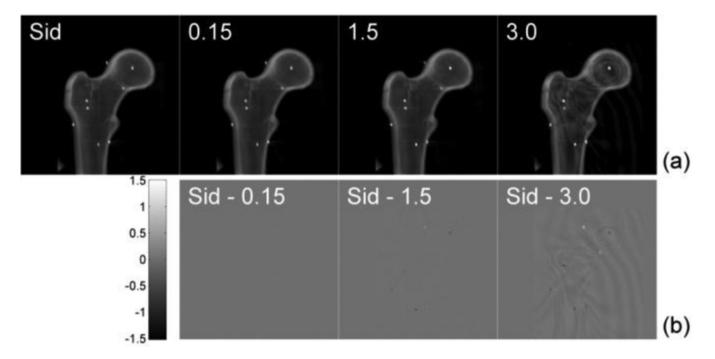
Fig. 13.
Comparison of the image quality between different ray-tracing methods. (a) comparison of the DRRs (the number shows step length, Sid: Siddon's algorithm). (b) difference images between the trilinear interpolation method and the ground truth (Siddon).
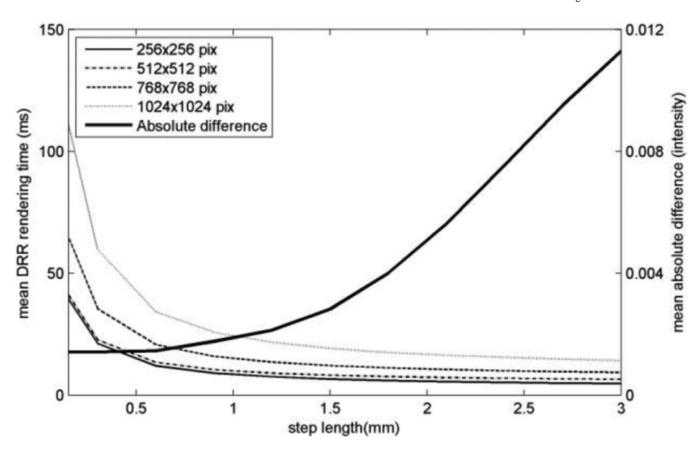
Fig. 14.
Computation time for DRR generation of 512^3 volume using trilinear interpolation method
and mean absolute difference of line integral from the ground truth (Siddon's method).

Fig. 15.
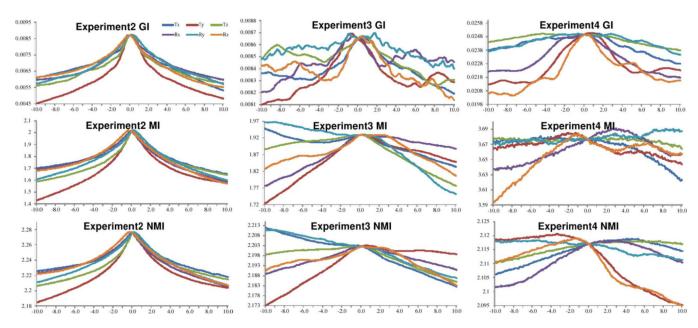Similarity measure values (vertical axis, sum of 2 image pairs) for individual pose parameters as a function of their deviation from the ground truth value (zero for all parameters.) Translational deviations (Tx, Ty, Tz) are in millimeters, rotational deviations (Rx, Ry, Rz) are in degrees. X, Y, Z axes correspond to the coordinate system associated with the first image in each experiment (see Table I).

TABLE I

Summary of the Dataset Used in the Experiments

Experiment type	Fixed images (X-ray images after log correction (eq. 3))	Floating images (DRRs of preoperative CT) at the ground truth position	Type of detector, calibration phantom	Ground truth
1) Gold standard dataset (a cadaver pig head) [60]			Flat panel (Electa synergy), Built-in calibration	paired point registration using fiducial markers
2) plastic phantom			Flat panel, Helical BB phantom	3D/3D registration (CBCT-PreopCT)
3) cadaver1			Flat panel, Helical BB phantom	3D/3D registration (CBCT-PreopCT)
4) cadaver2			Image intensifier, Two-plane phantom*	Paired point registration using fiducial screws

*Two-plane phantom contains BBs aligned in grid on the closer plane which enabled correction of image distortion

TABLE II

Specifications of the Dataset Used in Experiments

	X-ray image		Preoperative CT	(diagnostic CT)
	Resolution (pixels)	Pixel size (mm)	Resolution (voxels)	Voxel size (mm)
1)Gold standard	410×410	1.0×1.0	326×326×330	1.0×1.0×1.0
2)Plastic phantom	768×768	0.388×0.388	256×256×256	0.56×0.56×0.6
3)cadaver 1	768×768	0.388×0.388	256×256×256	0.78×0.78×2.0
4)cadaver 2	480×480	0.450×0.450	300×300×700	0.84×0.84×0.6

TABLE III

Specification of the Hardware Used for the Experiments

Operating system	Windows 7 64 bit
Processor type	Intel® Xeon® (2 processors)
CPU clock frequency	2.00 GHz
Graphics card type	NVIDIA® GeForce® GTX470
No. CUDA processor cores	448
Memory bandwidth	133.9 GB/s
Graphics memory	1280 MB

TABLE IV

Optimization Parameters Used in the Experiments 2–4

	Multi resolution level (down-sampling ratio)	3 (×1/4)	2 (×1/2)	1 (×1)
Down-hill simplex	Initial simplex size	4.0	4.0	3.0
	Convergence criteria [**]	0.1	0.01	0.01
CMA–ES	Initial search distribution (σ)	1.0	1.0	1.0
	Convergence criteria [**]	0.1	0.01	0.01

[*] All parameters are represented in mm for translations and in degrees for rotations.

[**] The iteration terminates when the change of parameters become smaller than this value.

TABLE V

Summary of Experiment 1 Using Gold Standard Dataset [60] (Failure Threshold of 5 mm was Used in This Experiment Only)

	Similarity measure	mTRE (mm) mean±std	Success rate (%)	Registration time (s)
Proposed method	GI	1.75 ± 0.32	100	4.4
	MI	4.11 ± 0.43	98.4	4.5
	MI	2.20 ± 0.92	98.4	5.7
Reported in [57] (CT vs kV x-ray without ROI)	BGB	2.30 ± 3.22	93.2	96
	RGB	1.48 ± 0.99	99.2	199

TABLE VI

Summary of the Results for Experiments 2, 3, and 4 With Varying Number of Images (100 Registration Trials With Initial Guess of 0–10 mm mTRE. Step Length: 1.5 mm, CMA-ES Optimizer, GI for Similarity Measure)

Number of fixed images	Experiment 2				Experiment 3				Experiment 4			
	mTRE (mm) for success	T(s)	R(%)	(deg)	mTRE (mm) for success	T(s)	R(%)	(deg)	mTRE (mm) for success	T(s)	R (%)	(deg)
1	1.33 ± 0.69	9.3 ± 0.1	69	n/a	*	11.4 ± 0.6	0	n/a	*	7.2 ± 0.2	0	n/a
2	0.34 ± 0.04	14.2 ± 0.2	100	90	0.99 ± 0.41	16.3 ± 0.1	81	58.5	2.04 ± 0.85	23.6 ± 2.8	45	31.5
3	0.49 ± 0.04	19.1 ± 0.2	100	60	0.82 ± 0.12	22.0 ± 0.2	87	47.3	n/a	n/a	n/a	n/a
4	047 ± 0.05	24.3 ± 0.2	100	45	0.43 ± 0.21	28.0 ± 0.3	76	31.5	n/a	n/a	n/a	n/a

T - registration time, R - success rate, - angle between consecutive images used in the experiment

*
mTRE for trials less than 10% success rate are not shown here