NIH Public Access
Author Manuscript
Phys Med Biol. Author manuscript; available in PMC 2013 September 07.

Published in final edited form as:
Phys Med Biol. 2012 September 7; 57(17): 5485–5508. doi:10.1088/0031-9155/57/17/5485.

Automatic Localization of Vertebral Levels in X-Ray Fluoroscopy Using 3D-2D Registration: A Tool to Reduce Wrong-Site Surgery

Y. Otake[1,2], **S. Schafer**[1], **J. W. Stayman**[1], **W. Zbijewski**[1], **G. Kleinszig**[3], **R. Graumann**[3], **A. J. Khanna**[4], and **J. H. Siewerdsen**[1,2]

[1]Department of Biomedical Engineering, Johns Hopkins University, Baltimore MD 21202

[2]Department of Computer Science, Johns Hopkins University, Baltimore MD 21218

[3]Siemens Healthcare XP Division, Erlangen Germany

[4]Department of Orthopaedic Surgery, Johns Hopkins Medical Institute, Baltimore MD 21202

Abstract

Surgical targeting of the incorrect vertebral level ("wrong-level" surgery) is among the more common wrong-site surgical errors, attributed primarily to a lack of uniquely identifiable radiographic landmarks in the mid-thoracic spine. Conventional localization method involves manual counting of vertebral bodies under fluoroscopy, is prone to human error, and carries additional time and dose. We propose an image registration and visualization system (referred to as LevelCheck), for decision support in spine surgery by automatically labeling vertebral levels in fluoroscopy using a GPU-accelerated, intensity-based 3D-2D (viz., CT-to-fluoroscopy) registration. A gradient information (GI) similarity metric and CMA-ES optimizer were chosen due to their robustness and inherent suitability for parallelization. Simulation studies involved 10 patient CT datasets from which 50,000 simulated fluoroscopic images were generated from C-arm poses selected to approximate C-arm operator and positioning variability. Physical experiments used an anthropomorphic chest phantom imaged under real fluoroscopy. The registration accuracy was evaluated as the mean projection distance (mPD) between the estimated and true center of vertebral levels. Trials were defined as successful if the estimated position was within the projection of the vertebral body (viz., mPD < 5mm). Simulation studies showed a success rate of 99.998% (1 failure in 50,000 trials) and computation time of 4.7 sec on a midrange GPU. Analysis of failure modes identified cases of false local optima in the search space arising from longitudinal periodicity in vertebral structures. Physical experiments demonstrated robustness of the algorithm against quantum noise and x-ray scatter. The ability to automatically localize target anatomy in fluoroscopy in near-real-time could be valuable in reducing the occurrence of wrong-site surgery while helping to reduce radiation exposure. The method is applicable beyond the specific case of vertebral labeling, since any structure defined in pre-operative (or intra-operative) CT or cone-beam CT can be automatically registered to the fluoroscopic scene.

Keywords

Wrong-site surgery; surgical planning; vertebral level localization; spine surgery; 3D-2D registration; GPU-acceleration; fluoroscopy; cone-beam CT

1. INTRODUCTION

Wrong-site surgery (Mulloy & Hughes, 2008) refers to a surgery performed on the wrong side or site of the body. Although the reported frequency of occurrence is subject to broad variation and possible under-reporting, an estimated 1 out of 113,000 surgeries meets with a wrong-site error (Kwaan, Studdert, Zinner, & Gawande, 2006), with 331 insurance claims

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

filled in the United States in the period 1985–1995 based on a review involving ~110,000 physicians (Canale, 2005). The spine is among the areas most prone to wrong-site surgery due to error in vertebral level localization (referred to as “wrong-level” surgery). The challenge in vertebral level localization arises in large part from the lack of uniquely identifiable features and periodic appearance of spine levels in projection imaging (i.e., x-ray fluoroscopy). In a survey of 3,505 physicians, Mody et al. reported that 50 % of spine surgeons experience wrong-level surgery at least once during their career, and wrong-level surgery is reported to occur approximately once in 3,110 spine surgeries (Mody et al., 2008).

The conventional site localization method in spine surgery is “level counting” – i.e., counting the number of vertebrae from readily identifiable anatomical landmarks in fluoroscopy (Hsiang, 2011). For example, localizing levels in the C-spine and upper T-spine may be accomplished by first localizing the occipitocervical junction and “counting” inferiorly to the target level; similarly for the lower T-spine or L-spine, level localization may be accomplished by counting “up” from the lumbosacral joint. Such a method involves an undesirable amount of time, ionizing radiation, and risk of human error. As an alternative, preoperative placement of radio-opaque fiducials [e.g., metallic screws (Upadhyaya, Wu, Chin, Balamurali, & Mummaneni, 2012) or PMMA markers (Hsu et al., 2008)] can be performed under real-time CT guidance as an additional preoperative procedure. Although this approach has shown a higher level of level localization accuracy, it carries additional time, cost dose, risk of infection, and inter-departmental hospital logistics.

In this paper, we propose an alternative approach for localization of the surgical site in intraoperative fluoroscopy using a fast CT-to-fluoroscopy registration. Specifically, the target vertebral levels (and other structures or trajectories of interest as desired) are defined in a 3D volume image in a preoperative planning step. The structures so defined in CT are subsequently registered and labeled automatically in the fluoroscopy image to provide near-real-time intraoperative visualization of the target level.

Registration of a 3D volume image to a 2D projection image, referred to as 3D-2D registration, has been extensively investigated (Markelj, Tomaževič, Likar, & Pernuš, 2012). The goal of 3D-2D registration is to estimate the transformation between the 2D imaging geometry and the 3D volume coordinate system from one or more 2D projection images of the 3D object. Such methods are typically classified into feature-based and intensity-based approaches. Feature-based approaches, e.g., (Gueziec, Kazanzides, Williamson, & Taylor, 1998), require the definition of image features (e.g., segmentation of anatomical structures) and feature correspondences which are non-trivial to automate and are susceptible to parameter tuning and segmentation error. Intensity-based approaches, on the other hand, utilize all the information contained in the 2D and 3D images and have demonstrated increased accuracy and reliability in comparison to the feature-based approaches (McLaughlin et al., 2005).

An intensity-based 3D-2D registration often consists of the following processes: 1) computing a simulated x-ray image, referred to as a digitally reconstructed radiograph (DRR), from the 3D volume at a given pose with respect to the 2D imaging geometry; 2) comparing the DRR with the (“real”) reference 2D image based on a similarity metric; and 3) performing a numerical optimization (usually iterative in steps 1 and 2) to find the optimum pose that maximizes similarity between the DRR and the real image. Despite advantages in accuracy and reliability compared to feature-based approaches, the high computational complexity of DRR computation and pose optimization have been a major hurdle somewhat inhibiting application of intensity-based approaches in routine clinical practice. To overcome such barriers, improvements in computational efficiency in DRR generation have formed a focus of research using both software and hardware solutions for

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

fast projection simulation. The software improvements mainly rely on pre-computation of a set of parameters such as an attenuation field (D. Russakoff et al., 2005; D. A. Russakoff, Rohlfing, Maurer, & Jr., 2003), transgraph (LaRose, & Kanade, 2000), or gradient vectors (Tomazevic, Likar, Slivnik, & Pernus, 2003). Despite its computational efficiency within the optimization process in the registration, pre-computation tends to be costly and involves various approximations in order to store the pre-computed information efficiently, potentially degrading the quality of the resulting DRR. Hardware-based enhancement has been largely focused on utilization of graphics processing units (GPUs) (Kubias et al., 2008; Spoerk, Bergmann, Wanschitz, Dong, & Birkfellner, 2007) and has benefited significantly from rapid in hardware, including the degree of integration and the clock speed of each core. Otake et al. reported GPU-accelerated fast 3D-2D registration approach using a traditional grid-interpolation ray-tracing method without pre-computation, thereby completing a single-image registration trial in about 10 seconds (Otake et al., 2012). The computation of a 512×512 pixel DRR using a 512×512×512 voxel CT volume required about 10 ms on a midrange GPU.

Various types of similarity metrics have been proposed in the literature, classified in two broad categories: 1.) metrics based on correspondence of global intensity in the images, such as mutual information (Maes, Collignon, Vandermeulen, Marchal, & Suetens, 1997), normalized cross-correlation (Penney et al., 1998), and sum of squared differences (Khamene, Bloch, Wein, Svatos, & Sauer, 2006); and 2.) metrics based on correspondence of local intensity, such as pattern intensity (Penney et al., 1998), gradient correlation (Penney et al., 1998), and gradient information (GI) (Pluim, Maintz, & Viergever, 2000). Comparative studies of these similarity metrics (Birkfellner et al., 2009; Penney et al., 1998; D. Russakoff et al., 2003) showed that metrics in latter category typically demonstrated better accuracy and robustness in 3D-2D registration, although the best similarity metric was dependent on the target application. For numerical optimization, local search algorithms have been widely employed, since the search space is well-posed (i.e., monotonic and convex in the vicinity of the true registration) when an appropriate similarity metric is chosen. These include gradient-based optimization approaches, such as gradient descent (McLaughlin et al., 2005) and non-gradient based approaches, such as Powell's method (Powell, 1964), the classic downhill simplex algorithm (Nelder & Mead, 1965), and the covariance matrix adaptation evolution strategy (CMA-ES) (Hansen & Kern, 2004; Hansen, Niederberger, Guzzella, & Koumoutsakos, 2009). In the work reported below, we employed the GI similarity metric and the CMA-ES optimizer by virtue of their inherent parallelizability (i.e., implementation on GPU) and robustness against image noise (detailed in 2.1.3.).

The scope of applications of 3D-2D registration in medical interventions is fairly broad, for example, in image-guided radiation therapy (IGRT). Apart from the classic visualization of radiotherapy "port films" in comparison to DRRs generated from planning CT, two modern IGRT systems, Accuray Cyberknife (Fu & Kuduvalli, 2008) and BrainLab Novalis (Jin et al., 2006), employ an intensity-based 3D-2D registration for target localization. Each system uses one or more x-ray tubes to acquire x-ray projections for registration to planning CT at the time of treatment. Since setup procedures for patient positioning are fairly well developed, the registration algorithm can employ an initial estimate fairly close to the true pose. For example, the Cybernife system uses DRRs pre-computed in a subset of the search space (translations and rotations) to reduce computation time at the expense of capture range. A classical steepest descent optimization is employed, since the similarity metric behaves well within the small capture range. The application addressed in this paper requires a larger capture range due to a highly uncertain initial pose estimate. We also consider 3D-2D registration based upon a single x-ray image, which is much more challenging than

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

scenarios involving two (nearly orthogonal) projections due to the potential for false local optima in the search space.

As detailed below, a system is proposed to transfer information defined in the preoperative CT coordinate system (viz., labels of the vertebral levels and other data defined in presurgical planning) onto intraoperative fluoroscopy using a fast intensity-based 3D-2D registration. The proposed method offers a potentially useful decision support system for quick localization of the target surgical site, reduction in radiation exposure, and reduction of the risk of wrong-site surgery. The specific application considered in this paper is the labeling of vertebral levels, and the application in this context is referred to as "LevelCheck." The spine levels of interest are labeled in preoperative CT and transferred automatically to intraoperative fluoroscopy acquired using a C-arm from any realistic pose about the patient. Mid-thoracic vertebrae were considered specifically in this paper, since they are the most challenging to accurately, unambiguously identify in x-ray projections due to the lack of unique anatomical landmarks and the periodic bony structure from level to level. The mid-thoracic vertebrae are also the most distant from the occipitocervical and/or lumbosacral landmarks used in "level counting" and are therefore the most prone to error. Section 2 describes an overview of the proposed method and the LevelCheck implementation of 3D-2D registration using the GI similarity metric and the 6 degree of freedom (6DOF) CMA-ES optimizer initialized by a 2DOF brute force search. Two types of experiments are detailed below as initial testing and validation of the approach – the first using a large body of image simulations (50,000 trials generated from 10 patient CT datasets) and the second using real projection images of an anthropomorphic phantom. Section 3 discusses the experimental results, the success rate of LevelCheck localization of spinal vertebrae, a sensitivity analysis of one of the main parameters in CMA-ES optimizer (population size), and an analysis of failure modes (i.e., local optima) and means to reduce or eliminate registration failure. A discussion of limitations, future work, and translation to clinical studies is provided in Section 4.

2. METHODS

2.1 Overview of the Method

Figure 1 illustrates the proposed system. Preoperative processes include structure definition/labeling in preoperative CT as well as geometric calibration of the C-arm. Intraoperative registration processes include fluoroscopic imaging, 3D-2D registration, and overlay of registered information (e.g. labels) on the fluoroscopic image. The following subsections detail each step in the process.

2.1.1 Preoperative Processes: Label Definition and Geometric Calibration—

Preoperative processes include target identification in preoperative CT and geometric calibration of the x-ray imaging system. The definition of structures in preoperative CT can be quite general (e.g., labeling of skeletal structures, major vessels, and planned trajectories), but focus in the current work is specifically on definition of vertebrae as a means of augmenting intraoperative fluoroscopy in spine surgery. The target identification can be performed either by specifying the center of vertebrae as a series of 3D points or as a segmented CT volume of each vertebra (figure 1 lower left). In the current work, we defined each vertebral level simply as a "point" (labeled T1, T2, etc.) placed manually in preoperative CT at the approximate center of each vertebral body. A skilled operator using a 3D workstation was able to apply such labels in ~1 min. Verification/QA of the label definitions by the surgeon and/or an attending fellow required an additional ~1 min.

The geometric calibration of the C-arm determines the 3D location of the x-ray source and the detector with respect to the world coordinate system based on a calibration phantom with

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

known geometry (Navab et al., 1998). The proposed 3D-2D registration system uses only the x-ray source position with respect to the detector, since it uses a single projection image. In the case where an image intensifier is used, distortion correction is required as reported using a calibration phantom with metallic beads for 3D-2D registration (Otake et al., 2012) and for 3D reconstruction (Fahrig & Holdsworth, 2000).

The definition of coordinate systems used in this paper is as follows. The C-arm coordinate system was defined at the isocenter of the C-arm, with X and Y axes aligned with the horizontal and vertical axes of the detector (figure 1 upper right). The preoperative CT coordinate system was defined at the center of the CT volume, with X and Y axes aligned with the medio-lateral axis and posterior-anterior axis of the patient, respectively (figure 1 upper left). In order to describe the 3D-2D registration process, we defined the *"nominal anterior-posterior (AP) position"* as the relative position between the C-arm and CT coordinate systems in which: 1) the origins of the two coordinate system were aligned; 2) the Z-axis of the CT coordinates (craniocaudal axis) was aligned with the Y-axis of the C-arm coordinates; and 3) the X-axis of CT (mediolateral axis) was aligned with the X-axis of the C-arm. As in figure 1, this *nominal AP position* corresponds to the C-arm orientation in which the x-ray tube is under the table with the patient in a prone position (facing down, which is the usual case in spine surgery).

2.1.2. Intraoperative Process: Robust 2D Initialization—The workflow of the proposed registration method is shown in figure 2. A 2DOF brute-force search initializes a subsequent 6DOF optimization step. For the 2DOF initialization, the original CT volume was downsampled by a factor of 8 ($4 \times 4 \times 4$ mm^3 voxels for the clinical CT datasets; see Table 1 for details). The fluoroscopy image was downsampled by a factor of 16 [i.e., 768×768 native format (0.388×0.388 mm^2/pixel) downsampled to 48×48 pixels (6.2×6.2 mm^2/pixel)]. Such highly downsampled images were used in the initialization step, since registration with lower resolution images constrains large changes in similarity metric for small perturbations in pose, thereby reducing false local optima and improving robustness, although the accuracy of registration is coarse. This principle was studied and experimentally demonstrated in medical image registration using a hierarchical multi-resolution approach (Munbodh et al., 2009; Studholme, Hill, & Hawkes, 1997).

For the 2DOF search initialization, multiple DRRs of the preoperative CT were calculated with C-arm pose translated in 10 mm intervals in the X and Y directions from the *nominal AP position*. DRR calculation was performed by the grid-interpolation ray-tracing method (Cabral, Cam, & Foran, 1994) implemented in C/C++ using CUDA libraries for execution on GPU. The texture memory on the GPU was used to store the volume data to reduce memory access latency and the global memory was used to store the line integral. For the grid-interpolation algorithm, the step length parameter (i.e., the distance between each sample point along a ray) was adjusted to balance computation speed (faster for larger step lengths) and DRR artifacts (reduced for smaller step lengths by the denser sampling along a ray). Otake et al. (Otake et al., 2012) investigated the tradeoff between speed and accuracy and reported that although the registration accuracy depends on the target images, in all types of target images investigated, the accuracy was nearly constant up to a step length of ~3.0 voxels (despite visible degradation in image quality for step length larger than ~1.5 voxels). In this study, we chose a step length of 1.0 voxels.

Calculation of DRRs first converted the CT data from Hounsfield units (HU) to linear attenuation coefficient (mm^{-1}) based on the attenuation coefficient of water (μ_{water}) at the effective energy at which the CT data was acquired. Thus the absolute value of the line integral matched the fluoroscopy image after log normalization. Although the polyenegertic nature of x-rays imparts several important effects in the x-ray projection, the effects are

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

assumed to be small with respect to the 3D-2D registration task considered herein (analogous to beam-hardening artifacts in CT data that are typically reconstructed under an assumption of monoenergetic x-rays), and a monoenergetic x-ray model was employed for DRR calculation. The possible mismatch to real x-ray images was investigated in physical experiments using an anthropomorphic phantom and polyenergetic x-ray beam as described in section 2.3.

Image similarity between the fluoroscopy image and each DRR was computed, and the C-arm pose yielding the largest similarity was chosen as the initial estimate. We used gradient information (GI) as a similarity metric originally proposed by Pluim et al. (Pluim et al., 2000). GI was defined as

$$\mathrm{GI}(p_1, p_2) = \sum_{i,j \in \Omega} w(i,j) \min(|\nabla p_1(i,j)|, |\nabla p_2(i,j)|) \tag{1a}$$

$$\nabla p(i,j) \triangleq \left(\frac{d}{di} p(i,j), \frac{d}{dj} p(i,j) \right) \tag{1b}$$

where $p_{1(i,j)}$ and $p_{2(i,j)}$ are the pixel values in each image at the (i, j) pixel, min(A,B) indicates the smaller value of A and B, Ω represents the entire image domain, ∇p is a gradient vector, and the weighting function w was defined as:

$$w(i,j) = \frac{\alpha_{i,j}+1}{2}, \alpha_{i,j} = \frac{\nabla p_1(i,j) \cdot \nabla p_2(i,j)}{|\nabla p_1(i,j)| \cdot |\nabla p_2(i,j)|} \tag{2}$$

where $\alpha_{i,j}$ represents the cosine between gradient vectors at location (i,j). Higher weight is therefore applied to pixels for which the angle of two gradient vectors is close to zero (and lower weight is applied at locations for which α is close to 180°). The original definition of GI (Pluim et al., 2000) applied higher weight on gradient vectors of both 0° and 180° to make it insensitive to the sign of the gradient intensity in different imaging modalities. The modified form described above better suits the single-modality (i.e., x-ray CT to x-ray projection) registration considered here (as opposed to, for example, registration of T1 and T2 MR images).

GI is robust against a broad variety of mismatches, including deformations and image feature discrepancies, between intraoperative fluoroscopy and DRRs computed from preoperative CT, because it weights only those pixels exhibiting a strong gradient in both images by introducing the min() operator in (1a). Such image mismatches include, for example: the presence of surgical tools (e.g., screws and needles) in intraoperative fluoroscopy which were not present in preoperative CT; motion of the diaphragm – typically at the position of comfortable inspiration breathhold in preoperative CT, whereas it could present at any breathing phases in intraoperative fluoroscopy; and anatomical deformation of the spine due to different patient positioning (e.g., supine position in CT and prone position intraoperatively). Another advantage of using GI in 3D-2D registration is that the computation of the 2D gradient of an image is highly parallelizable, and therefore amenable to GPU implementation, compared for example, to mutual information, which requires the computation of a joint histogram [non-trivial to parallelize without a variety of approximations (Shams & Barnes, 2007)].

The fluoroscopy image is typically recorded as a linear intensity which is proportional to the number of x-ray photons (actually the total x-ray energy) absorbed at each detector element.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The intensity is logarithmically proportional to the line integral of the attenuation coefficient according to Beer's law. The logarithm of the pixel intensity I_d converts detector signal to a form proportional to the line integral. Since the GI similarity metric is based on the intensity gradient, an offset of the pixel value has no effect.

The initial 2DOF search is therefore described as follows:

$$\widehat{T}_{init} = \underset{T}{\operatorname{argmax}}\ GI\left(p_{fluoro}, p_{DRR}\left(T(t_x, t_y)\right)\right) \tag{3}$$

where p_{fluoro} is the fluoroscopy image, $p_{DRR}(T)$ is a DRR computed at a given C-arm pose, T. In this search, the transformation T is parameterized by just 2 parameters: translation in X (t_x) and Y (t_y) Although the search did not consider rotations or Z translation, experiments below demonstrate a reasonably robust initialization for AP fluoroscopy provided the C-arm within ~±20° of the true AP view. For lateral (LAT) fluoroscopy, the 2DOF search would operate over a similar range in Y(t_y) and Z(t_z).

2.1.3. Intraoperative Process: 3D-2D Registration—In the subsequent 6DOF optimization step, we employed downsample ratios of 8 for both the CT data (128×128×180 voxels, ~2×2×2 mm^3/voxel) and fluoroscopy images (96×96 pixels, 3.1×3.1 mm^2/pixel). GI was used as a similarity metric. The 6DOF search used a non-gradient optimization approach, CMA-ES (Hansen et al., 2009). The algorithm generates candidate sample points randomly around the current estimate according to a multivariate normal distribution with a certain mean and a covariance matrix. The number of sample points in one step (generation) is called the "population size" (denoted λ). The objective function is evaluated at each sample point, and according to the function values at all sample points in one generation, the algorithm modifies (adapts) the mean and the covariance matrix of the sample distribution for the successive generation in such a way that the distribution aligns with the gradient direction of the objective function. The choice of λ is generally a compromise between convergence speed and robustness. A smaller λ leads to faster convergence, and a larger λ helps avoid false local optima and gives a wider capture range. Since the function evaluations in each generation are performed independently, the algorithm is highly parallelizable, compared with, for example, downhill simplex (Nelder & Mead, 1965), where only N+1 sample points (function evaluation) can be parallelized in an N dimensional optimization problem.

Using the CMA-ES optimizer with the initialization described in the previous section, the 6DOF C-arm transformation $T(t_x,t_y,t_z,\theta_x,\theta_y,\theta_z)$ that maximizes similarity between fluoroscopy and DRR was estimated:

$$\widehat{T} = \underset{T}{\operatorname{argmax}}\ GI\left(p_{fluoro}, p_{DRR}\left(T(t_x, t_y, t_z, \theta_x, \theta_y, \theta_z)\right)\right) \tag{4}$$

The CMA-ES algorithm was implemented in Matlab as in Hansen (Hansen, 2006). The optimizer included a function call to an externally compiled shared library coded in C++ and CUDA for DRR generation and similarity metric computation. CMA-ES has been used for 3D-2D registration in other work (Gong & Abolmaesumi, 2008; Otake et al., 2012) and found potentially advantageous over classical optimizers. The studies below include investigation of the sensitivity of registration accuracy and speed to the population size parameter (λ) using a large number of simulation trials (detailed below). We also experimentally demonstrated the advantage in parallelization efficiency as a key factor in speed enhancement using massively parallelized computation on the GPU. We also

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

compared speed and robustness of the methods with the downhill simplex approach (Nelder & Mead, 1965).

The stopping criterion was chosen to avoid premature termination while minimizing unnecessary iterations at excessively fine increments in pose that exceed the geometric accuracy requirements in our application (specifically, <5 mm accuracy in the projected location of the vertebrae, as detailed below). The stopping criterion was therefore chosen as the tolerance projected in the pose parameter space (t_x,t_y,t_z,θ_x,θ_y,θ_z), for which the change in each coordinate in one generation becomes less than 1 mm for translations and 1 degree for rotations.

2.1.4. Intraoperative Process: Automatic Labeling and Visualization in Real-Time Fluoroscopy—Projection of the vertebral level ("label") location defined in the 3D CT coordinate system to the 2D coordinates *(u,v)* of real-time fluoroscopy was computed follows:

$$P^L_{fluoro}=(\ u \quad v \quad 1\ )^T \sim C \cdot {}^{C-arm}T_{CT} \cdot P^L_{CT}\ (L=1,2,\cdots,n) \tag{5}$$

where L denotes the Lth vertebral level "label", $P^L{}_{fluoro}$ is the homogeneous 2D coordinates of the 2D point in fluoroscopy coordinates (3 element vector), $P^L{}_{CT}$ is the homogeneous 3D coordinates of the point in the CT coordinates (4 element vector), n is the number of labels, ${}^{C\text{-}arm}T_{CT}$ is a 4×4 homogeneous matrix representing the transformation computed in (4), C denotes the intrinsic parameter matrix of the C-arm, and the ~ symbol denotes that the left and right sides are equal to within scalar multiplication – i.e., (a b c) ~ (A B C) implies a = A/C and b = B/C. In addition to projecting the location of vertebral labels defined in CT, a variety of planning information described as a series of 3D points (P_i) can be similarly transformed onto the fluoroscopy coordinate system in the same way – e.g., planned tool trajectories and the locations (or volumetric segmentations) of the surgical target and adjacent critical anatomy.

As illustrated in figure 1 for the case of a simple point localization of (the center of) vertebral levels, the "true" position of each level (as defined in the 3D CT and projected to 2D fluoroscopy) is marked as a cyan crosshair and label in all figures. The "estimated" position of each level (i.e., its position in 2D fluoroscopy as computed by the registration method detailed above) is marked as a yellow crosshair and label.

2.2. Simulation Studies

Simulation studies using an ensemble of patient CT images were conducted to test the feasibility, accuracy, robustness, and computation speed of the proposed method over a broad range of possible C-arm poses about the patient. Ten clinical CT datasets were randomly selected from The Cancer Imaging Archive (TCIA) provided by National Cancer Institute (Armato et al., 2011) as illustrated in figure 3. Details pertaining to each image along with the hardware specifications and optimization parameters used in the experiments are summarized in table 1. Each CT image covered at least 10 vertebral levels. Patient information such as age, sex, and body mass-index were not available in the de-identified datasets.

For the simulation study, fluoroscopic images were simulated by applying a high-fidelity Siddon's ray-tracing algorithm (Siddon, 1985) to the CT images, computing accurate line integrals based on the intersection length between CT voxels and the ray passing through each voxel. Note that Siddon's ray-tracing algorithm for simulating the fluoroscopy images is distinct from the coarse grid interpolation algorithm used for DRR calculation in the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

registration process. The former is more accurate than the latter and it involves considerably higher computational complexity, but it was considered a better choice for simulation of fluoroscopy images.

The nominal C-arm projection geometry involved a source-detector distance = 1200 mm and a source-isocenter distance = 600 mm. To allow for human operator variability in positioning the C-arm about the patient, the C-arm pose for each fluoroscopy image was perturbed from the *nominal AP* orientation – i.e., allowed to vary in 6 degrees of freedom provided the target vertebrae was still in the fluoroscopic field of view (FOV). The perturbations were intended to emulate realistic variability in clinical setup and C-arm positioning, which is a function of operator experience, positioning aids (e.g., field lights or lasers), and the complexity of the operating setup. Perturbations from the nominal pose (X, Y, Z = 0 mm; θ_x, θ_y, θ_z = 0°) were described by Gaussian probability distributions as shown in figure 4 with three standard deviations (3σ) in each distribution as follows: $3\sigma_X$ = 50 mm, $3\sigma_Y$ = 150 mm, $3\sigma_Z$ = 100 mm, and $3\sigma_{\theta x} = 3\sigma_{\theta Y} = 3\sigma_{\theta z}$ = 10°. These values were set according to estimates of setup variability by an experienced spine surgeon. The smallest variation ($3\sigma_X$ = 50 mm) is in the lateral direction, suggesting that the operator is reasonably capable of centering the spine medially in the fluoroscopic FOV (±~50 mm). Larger variations were allowed in the longitudinal (Y) and source-detector (Z) directions – the former, since it is difficult to place the C-arm on the target vertebral level (related to the very problem of level localization), and the latter, since a C-arm operator would usually not consider it imperative to place the spine at isocenter (as long as it is within the fluoroscopic FOV) and would not be particularly sensitive to variations in magnification. The variation in angles are associated primarily with the angle of C-arm setup at tableside. In each case, we assumed that the spinal vertebrae are somewhere within the fluoroscopic FOV; therefore extrema in σ that would project the spine outside the image were not permitted. The distributions in figure 4 appear somewhat non-Gaussian as a result of this requirement.

The simulation study allowed assessment of accuracy and robustness of the LevelCheck registration algorithm and investigation of algorithmic parameters without disturbance by factors such as x-ray scatter, quantum noise, polyenergetic spectrum effects, electronic noise, and other detector nonidealities. Such effects were assessed in a physical experiment described in the next section.

For each of the 10 patient images (figure 3), 5,000 fluoroscopic images were simulated (50,000 images in total) from C-arm pose determined by sampling of the distributions in figure 4. The proposed 3D-2D registration algorithm was applied to each case, and the accuracy of registration was evaluated using the mean projection distance (mPD) as described below (Section 2.4). It should be acknowledged that the 50,000 images tested do not represent independent trials; rather, the data represent 10 independent (randomly selected) patients covering a spectrum in body habitus and anatomical variations, each with 5,000 simulated fluoroscopic views. As noted below, the small number of “failures” (i.e., registration giving mPD > 5mm) among the 50,000 samples was too low to permit a statistically valid sub-group analysis of possible sources of bias (e.g., patients for whom the registration was more prone to failure). The pilot study reported below represents initial assessment of feasibility and performance and is subject to future investigation in a greater number of individual patients.

2.3. Experiment with Real X-ray Projections

A physical experiment using real fluoroscopic images was conducted to assess the effects of x-ray scatter, quantum and electronic noise, polyenergetic x-ray spectrum, etc. that appear in a real fluoroscopy and were not simulated in the study described above. The experimental setup is shown in figure 5. The subject was an anthropomorphic phantom containing a

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

natural human skeleton in soft-tissue-equivalent plastic (a custom Rando™ phantom (The Phantom Laboratory, Greenwich NY)). A prototype mobile C-arm developed for cone-beam CT/fluoroscopy guidance of minimally invasive surgery was used for image acquisition. The C-arm incorporated a flat-panel detector (FPD) and motorized rotation, with geometric calibration performed with a previously described calibration phantom (Navab et al., 1998). The FPD (PaxScan 3030 CB, Varian Imaging Products, Palo Alto, CA, USA) provided distortionless readout (30×30 cm^2 FOV, 768×768 pixel format, 0.388 mm^2/pixel) at 3.3 frames per second, and the imaging technique was 100 kVp, 1.3mA.

As in the simulation study, the target vertebral levels were identified in an existing 3D CT dataset (0.98×0.98×0.6 mm^3/voxel, 512×512×659 voxels), and the projection of each (according to the known C-arm geometry and the ground truth registration between the C-arm and the phantom) defined the "true" location of each level in fluoroscopy. The ground truth registration was computed by a 3D-2D registration between the 3D CT dataset and 4 projection images with known relative geometries equally distributed over 200 degrees. Thirty-five fluoroscopic images were acquired over a 70° arc about the *nominal AP* view (i.e., θ_y from −35° to +35° at 2° intervals). The LevelCheck registration algorithm was employed as detailed above to automatically label the position of each vertebral level in the real fluoroscopic images.

2.4. Evaluation Methods

The geometric error in vertebral localization was evaluated as the distance between the true and estimated vertebral position on the 2D image plane, referred to as the projection distance (PD) (van de Kraats, Penney, Tomazevic, van Walsum, & Niessen, 2005). Five mid-thoracic vertebrae were chosen as targets, and the average PD across the five targets was defined as the mean projection distance (mPD):

$$\mathrm{mPD} \triangleq \frac{1}{N}\sum_{i=1}^{N} \mathrm{PD}_i \text{ (N:number of target vertebrae)} \tag{6}$$

as illustrated in figure 6.

To determine the overall accuracy and reliability of the registration algorithm from multiple trials, a failure criterion (Gendrin et al., 2011; van de Kraats et al., 2005) was conservatively defined such that any trial resulting in mPD greater than 5 mm was classified as a failure. The 5 mm value was chosen such that PD less than this threshold implied that the estimated location was well within the projected boundaries of the true vertebral level in the fluoroscopy image. For example, in C-arm geometry of figure 1 (magnification = 2) and with the target vertebrae at isocenter, this threshold implied that the true location was within 2.5 mm of the true center of the vertebra in the 3D coordinate system. We consider this threshold to be a very conservative estimate such that mPD < 5 mm definitely implies a successful registration and an accurate labeling of the vertebrae – i.e., the label would lie clearly within the projected boundaries of the vertebra. The ratio of the number of successful trials (i.e., trials for which mPD < 5 mm) to the total number of trials defined the success rate (%).

In cases where the registration failed (mPD > 5 mm), an analysis of failure modes was conducted to determine why the algorithm may have converged at an inaccurate location (or if it converged at all). The objective function (GI) was plotted as a function of the pose parameters at each iteration of the algorithm, and the shape of the 6D search space was investigated for local maxima. Specifically, local objective shape was analyzed through a dimension reduction whereby the six dimensional parameter associated with the estimated pose [denoted $T_{est}(t_x,t_y,t_z,\theta_x,\theta_y,\theta_z)$] and that of the true pose parameter [denoted

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

$T_{true}(t_x,t_y,t_z,\theta_x,\theta_y,\theta_z)$] was linearly interpolated with respect to a single variable, α. The one-dimensional pose surrogate T(α) was defined as:

$$T(\alpha)=T_{est}+\alpha(T_{true}-T_{est}) \quad (7)$$

where α=0 indicates the estimated pose (T_{est}), and α=1 indicates the true pose (T_{true}). For each failure case (mPD > 5 mm), the GI was computed as a function of α to elucidate changes in GI within the search space – i.e., to evaluate the steepness of change and identify local and global maxima.

3. RESULTS

3.1. Simulation Studies

3.1.1. 2DOF Initialization—Results of the 2DOF initialization search in four examples are shown in figure 7. In each case, DRRs were computed at 209 poses (11 in the X-direction covering a range of 100 mm laterally, and 19 in the Y-direction covering a range of 180 mm in the SI direction), and GI was computed between each DRR and the target fluoroscopy image. For example #1 (*nominal AP pose*), a sharp peak is observed at the center of the similarity metric plot (figure 7b). For displacement in the X direction (example #2), the peak shifts accordingly as seen in figure 7c. Combined translations and rotations (examples #3 and #4; figures 7d and 7e) reduce the magnitude of the peak GI value and introduce weak local maxima in the search space, but still yield global optima at the X, Y position corresponding to the true registration. The 2DOF search therefore provided a fairly robust initialization, and results below indicate that initialization brought the registration to within ±25 mm in all cases considered.

Generation of the (209) DRRs and computation of GI at the downsampled resolution associated with the 2DOF initialization took 0.6 ms and 1.2 ms, respectively, therefore requiring 380 ms on average for the entire initialization step. Note that each DRR and GI computation could be performed independently from each other which significantly improved computational efficiency in a parallel computing environment.

3.1.2. Sensitivity of the Optimization Process—The sensitivity analysis of registration performance as a function of CMA-ES population size λ (section 2.1.3.) is summarized in figure 8. Fifty thousand trials with random simulated fluoroscopic images were analyzed using 5 different population sizes: λ = 30, 60, 90, 120, and 150. The results quantify the extent to which increasing population size improved the success rate at the expense of computation time. Larger population size (e.g., a factor of two increase from λ = 60 to 120) resulted in an increased computation time (3.27 s and 4.67 s, respectively), an increased number of function evaluations (1,675 and 2,425, respectively), and a reduction in failures (11/50,000 and 1/50,000, respectively). The improvement in success rate is asymptotic, while the increase in computation time is nearly linear. The downhill simplex showed a significantly lower success rate than CMA-ES even when the population size was chosen such that the computation time was nearly equal (λ = 30). Due to the difference in parallelization efficiency (described in 2.1.3.), CMA-ES could perform a larger number of function evaluations in the same amount of time compared to downhill simplex (about a factor of two in these experiments). Although the case λ = 150 showed 100% success in this experiment, it is worth noting that a stochastic optimization approach such as CMA-ES entails a non-deterministic solution and is always subject to a finite level of failure probability. Further experiments using an even larger number of trials would be necessary to reveal the frequency of such small probability stochastic failure.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

As a result of the sensitivity analysis, we chose λ = 120 as a reasonable population size for initial studies of vertebral level localization, offering a level of accuracy and computation time which is generally consistent with desired performance in the target application (i.e., localization well within the projected boundaries of the vertebra).

Computation time increased linearly. Failures seen at small population size (λ=30) tended to concentrate in a few particular subjects (#2, #4, #8, #9, #10), suggesting degeneracies (local maxima) in the 6DOF search space.

3.1.3. Simulation Study Results—Figure 9 shows the convergence typical of a given trial. Although the calculation is performed in parallel in the actual implementation as explained in 2.1.3., the plots show the iterations serially. Variations in the observed Z-translation estimate at each iteration were larger than the other directions because image features are relatively insensitive to the C-arm movement in the Z direction, affecting magnification rather than in-plane shifts and resulting in small changes in the similarity metric. The mean projection distance converged toward zero as the GI similarity metric was maximized at ~2,000 function evaluations (figure 9c,d). The plots suggest that although the 6 pose parameters converged to a range small enough to achieve a successful registration (i.e., mPD < 5 mm), the GI and mPD metrics were still decreasing with each iteration (not yet converged), implying that further iterations could be pursued if higher levels of accuracy are needed in a given application (i.e., other than coarse labeling of vertebral labels). As explained below in section 3.3, we observed no failures in the entire experiment due to premature convergence.

Table 2 and figure 10 detail registration results for each dataset. For the parameter selections detailed above (e.g., λ = 120), there was one failure in 50,000 trials, giving a success rate of 99.998 %. The average computation time was 4.67 seconds, suggesting suitability to labeling of “spot” (i.e., single-shot) fluoroscopy, but not real-time, multi-frame cine. Figure 10 summarizes the translational and rotational error of 50,000 registration trials: (a) before the registration, (b) after the initial 2DOF search, and (c) after the 6DOF optimization search. The box plots show the median, quartile range, and maximum and minimum of the entire trial. The standard deviation in X and Y translation was reduced by the 2DOF initialization search from (16.3, 33.8) mm to (3.6, 3.9) mm, while the other 4 pose parameters remained the same. After the 6DOF optimization search, the absolute errors in each direction were (0.02 ± 0.02, 0.02 ± 0.02, 0.54 ± 0.28) mm for translation, and (0.03 ± 0.03, 0.03 ± 0.03, 0.01 ± 0.01) degrees for rotation. The larger error in Z translation (the axis perpendicular to the image plane) was attributed to the fact that changes in magnification (i.e., motion along Z) changes the projection much less than shifts in X and Y, and the similarity metric is therefore less sensitive to Z translation. Slightly larger errors in rotations were observed about X and Y compared to Z, attributed to cylindrical symmetry of the spine – i.e., C-arm rotation around the long axis (Y direction) and lateral-medial axis (X direction) changes the projection less than the rotation around the anterior-posterior axis (Z axis). Figure 10d shows mPD before registration (75.9 ± 39.3 mm), after 2DOF initialization search (15.7 ± 8.3 mm), and after the optimization search (0.2 ± 0.2 mm). The single failure case (detailed below) exhibited an mPD of 44.4 mm.

3.2. Experiments with Real X-ray Images

Figure 11 summarizes results for the experiment using real fluoroscopic images. All 35 images (obtained over the angular range −35° to +35° about the *nominal AP view*) were successfully labeled with less than 1 mm mPD (0.4 ± 0.2 mm). This suggests that non-idealities such as a polyenergetic spectrum, x-ray scatter, quantum noise, and electronic noise in real fluoroscopic images did not have a significant impact on the LevelCheck

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

method. The algorithm was robust against C-arm misalignment up to ±35 degrees, which is considered to be a much larger range than expected with a trained fluoroscopy technician.

3.3. Analysis of Possible Failure Modes

As detailed in figure 8, smaller values of the CMA-ES population parameter (λ) resulted in higher probability of registration failure. Interestingly, all 47 failure cases found in the study (i.e., the total across all settings of the population parameter, λ=30–150) were associated with the same failure mode – a false local optimum in the 6DOF search space arising from the periodic structure of vertebral levels. The single typical failure case observed in the λ=120 case was representative of the failure mode and is illustrated in figure 12. The failed registration exhibits a translational shift in the superior-inferior direction by one vertebral height – potentially periodic (but at reduced strength in the local optimum) at multiples of the vertebral height. As detailed in 2.4., the shape of the search space around the false local optimum for the failure case was analyzed as shown in figure 13. The shallow local optimum is a result of the longitudinally periodic features in the image around the true pose. However, in all failures observed, the objective function value at the true pose showed a sharper global optimum. Increasing the CMA-ES population size from 30 to 120 provided a larger number of sample points in each generation, and the covariance matrix in successive generations was less sensitive to small perturbations in the objective function, helping to avoid the local optimum and reducing the number of failures from 32 down to 1 in 50,000 trials.

4. DISCUSSION

This paper proposed a method to automatically localize and label the vertebral levels in intraoperative fluoroscopy using vertebral positions defined in the preoperative CT. We specifically addressed the problem of a fast and highly robust intensity-based 3D-2D (CT-to-fluoroscopy) registration. The registration estimated by the initial implementation of the LevelCheck technique was successful in more than 99.99% of trials and required less than 5 seconds to compute. The analysis of failure modes revealed that all the failures observed (most notably with smaller CMA-ES population size) arose from the periodically symmetric structure of the spine and could be avoided with an increased population size at the expense of computation time.

The study demonstrates the potential of the proposed 3D-2D registration approach for automatic labeling of vertebral levels in fluoroscopy; of course, the study has a variety of limitations and areas deserving future investigation. In the simulation study, we used randomly generated fluoroscopic images assuming a Gaussian distribution on the possible variability of the C-arm setup by a human operator. The statistical properties of C-arm setup variability by a human operator in real clinical practice are unknown. However, AP fluoroscopy is a well defined, common setup that is frequently used in spine surgery, and the studies above assumed that the operator is able to align the C-arm at tableside with a reasonable degree of accuracy, especially in terms of rotation and horizontal translation. The distributions in C-arm setup variability shown in figure 4 are likely conservative (i.e., the ranges in the assumed pose are likely broader than those typical in real practice with a trained fluoro operator). A narrower range in C-arm setup variability would likely increase the success rate of the registration algorithm above that reported here.

Another frequently used fluoroscopy setup is the lateral view, where the x-ray projection is acquired in right-left direction to confirm alignment of structures in anterior-posterior direction. The lateral view is not the typical view used in spine localization (“level counting”), and the current work focused on registration in the AP view. Other preliminary experiments (Otake, Schafer et al., 2012) demonstrated the potential for application of the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

LevelCheck registration approach to the lateral view and demonstrated insensitivity to error in geometric calibration. This is a fairly intuitive result, since the similarity-based optimization is performed in the 2D projection domain, and a slight mismatch of the source position between the real and the virtual (calibrated) geometry would not change the relative position of the labels with respect to the anatomy in the 2D projection, although it does change the absolute position of the 3D volume coordinate system with respect to the C-arm coordinate system.

The physical experiments in this paper used a C-arm incorporated with a flat-panel detector, but the same approach could be applied to images acquired by the image intensifier (II) with an image distortion correction. As described in 2.1.1., the image distortion appearing in an II image can be corrected by well-studied correction methods, for example, using a phantom with beads in grid and approximating the distortion using the Bernstein polynomial (Sadowsky, 2008).

One limitation of the initial implementation is the requirement for preoperative CT data in which vertebral levels (and other structures, if desired) are defined. If only preoperative MR images are available, an additional image processing step would necessary to simulate an x-ray attenuation volume from the MR images. A possible solution in this regard involves a statistical atlas created from CT datasets – for example the “MR-CT synthesis” method, which creates a CT-like volume from an MR image by applying a piecewise-rigid registration on the CT-based statistical atlas. The resulting patient specific CT-like volume could then be used in the LevelCheck registration. Alternatively in the absence of preoperative CT, the definition of vertebral levels could be performed using an intraoperative CT scanner (which are becoming more common in CT-guided spine surgery(Patil, Lindley, Burger, Yoshihara, & Patel, 2012; Tonn et al., 2011)) and/or intraoperative CBCT (as with the prototype C-arm used in this work).

An important consideration in application of the registration method to a realistic clinical scenario is the deformations and other perturbations (e.g., introduction of surgical tools) to the patient occurring between the preoperative CT and intraoperative fluoroscopy. Such changes include surgical tools introduced in the patient during surgery or changes of the diaphragm level due to respiratory motion. As described in 2.1.2., the GI similarity metric is robust to foreign structures that appear in only one of the images – i.e., the GI metric exhibits high value at pixels that have large gradient in *both* images, so if a gradient is absent in one image or the other, the similarity metric is low. This property is advantageous to suppress adverse effects associated with the local shape changes of the target structure. Figure 14 illustrates an example surgical scenario in which a pedicle screw was placed in the T5 vertebral level of the phantom during fluoroscopy (but was not present in the preoperative CT). The LevelCheck algorithm was applied to two simulated x-ray images (approximately AP and LAT), and as evident in the images of figure 14, the screws were successfully ignored in the registration process, and the vertebral labels and planned transpedicle trajectories defined in the preoperative CT coordinate system were successfully overlaid onto the fluoroscopy image. Using such overlay images as a means of decision support, surgeons can obtain visual localization of planning structures within the fluoroscopic scene along with other quantitative metrics (e.g., screw positions and angles with respect to the target anatomy). Similarly, mismatch in the level of the diaphragm between the preoperative CT (breath-hold) and intraoperative fluoroscopy (free-breathing) may not pose a major limitation by virtue of the robust similarity metric. We also note that changes associated with gross patient deformation (e.g., differences in spine curvature) due to posture difference or anatomical changes across large areas (e.g., disc resection) may also adversely affect the registration. In such cases, the LevelCheck algorithm is likely more useful at the beginning of the procedure when such changes are less prominent and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

introduce less mismatch between fluoroscopy and preoperative CT. Methods incorporating deformable models that encode shape deformation as the principal component modes of a statistical atlas (Sadowsky, Chintalapani, & Taylor, 2007)(Hurvitz & Joskowicz, 2008), control points in a free-form deformation (Qi, Gu, & Xu, 2008), and/or a piece-wise rigid deformation is subject of future work.

It is worth noting that the LevelCheck algorithm is intended as an assistant (and not an absolute guide) in decision support for interventional procedures. Although the initial implementation demonstrated a very high success rate (e.g., 1 failure in 50,000 trials, which is less the error rate associated with manual level counting), there is always the potential for failure with a stochastic optimization scheme such as CMA-ES, and there is considerable work to be done to better validate registration accuracy under real clinical conditions. Even then, the labels registered by the algorithm should be interpreted as guide (and not an absolute indication of “truth”) as with any aspect of a surgical navigation system. By analogy, the labels applied by the LevelCheck algorithm are useful in a similar manner as the virtual position of a tracked tool in a surgical navigation system: specifically, each is subject to error, and it is ultimately up to the surgeon’s experience, expertise, and physical senses to determine the true situation and trap potential errors. Since the registration algorithm employs a quantitative similarity metric within the optimization, it is possible to include an auxiliary “confidence” measure along with the registration yielded by the algorithm. For example, in addition to labeling the vertebral levels as illustrated above, a metric of confidence related to the strength of the GI optimum achieved in the registration process could be displayed. In cases where such a metric is low (and thereby possibly associated with a false local optimum), the system could withhold label overlay, report a possible error, and/or request a repeat. Such considerations fall to future work in clinical translation.

The current work focuses on the basic methodology and initial implementation, with next steps toward clinical studies involving cadaver study and application to retrospective clinical data.

A straightforward extension of the proposed method includes applications to other types of target anatomy and surgical procedures in which data are defined in 3D preoperative data and could be visualized in intraoperative fluoroscopy. The proposed framework allows any type of preoperative information defined in the 3D image coordinate system to be registered and projected onto the 2D fluoroscopy. Potential applications include: postoperative verification of pedicle screw placement; analysis of joint replacement relative to the planned implant orientation; analysis of scoliosis correction relative to planned levels of straightening; overlay of planning data in image-guided radiation therapy (e.g., tumor, normal tissue volumes, and treatment beam apertures) on intra-treatment projections; and visualization of planned stent/catheter locations and target anatomy in interventional radiology.

Acknowledgments

This research was supported by academic-industry partnership with Siemens Healthcare and National Institutes of Health Grant No. R01-CA-127444. Many thanks to Dr. Elliot McVeigh (Department of Biomedical Engineering, Johns Hopkins University) and Dr. Ziya L. Gokaslan (Department of Neurosurgery, Johns Hopkins University) for valuable discussions relating to the development and application of the proposed 3D-2D registration.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

References

Armato, SG.; McLennan, G.; Bidaut, L.; McNitt-Gray, MF.; Meyer, CR.; Reeves, AP., et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. AAPM; 2011.

Birkfellner W, Stock M, Figl M, Gendrin C, Hummel J, Dong S, et al. Stochastic rank correlation: A robust merit function for 2D/3D registration of image data obtained at different energies. Medical Physics. 2009; 36(8):3420–3428. [PubMed: 19746775]

Cabral, B.; Cam, N.; Foran, J. Accelerated volume rendering and tomographic reconstruction using texture mapping hardware. VVS '94: Proceedings of the 1994 Symposium on Volume Visualization; 1994. p. 91-98.

Canale ST. Wrong-site surgery: A preventable complication. Clinical Orthopaedics and Related Research. 2005; (433):26–29. [PubMed: 15805933]

Fahrig R, Holdsworth DW. Three-dimensional computed tomographic reconstruction using a C-arm mounted XRII: Image-based correction of gantry motion nonidealities. Medical Physics. 2000; 27(1):30–38. [PubMed: 10659735]

Fu D, Kuduvalli G. A fast, accurate, and automatic 2D-3D image registration for image-guided cranial radiosurgery. Medical Physics. 2008; 35(5):2180–2194. [PubMed: 18561693]

Gendrin, C.; Markelj, P.; Pawiro, SA.; Spoerk, J.; Bloch, C.; Weber, C., et al. Validation for 2D/3D registration II: The comparison of intensity- and gradient-based merit functions using a new gold standard data set. AAPM; 2011.

Gong RH, Abolmaesumi P. 2D/3D registration with the CMA-ES method. Medical Imaging 2008: Visualization, Image-Guided Procedures, and Modeling. 2008; 6918(1):69181M.

Gueziec A, Kazanzides P, Williamson B, Taylor RH. Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot. Medical Imaging, IEEE Transactions on. 1998; 17(5):715–728.

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. Springer; 2006. p. 75-102.

Hansen N, Kern S. Evaluating the CMA evolution strategy on multimodal test functions. Parallel Problem Solving from Nature {PPSN VIII}. 2004; 3242:282–291.

Hansen N, Niederberger ASP, Guzzella L, Koumoutsakos P. A method for handling uncertainty in evolutionary optimization with an application to feedback control of combustion. Evolutionary Computation, IEEE Transactions on. 2009; 13(1):180–197.

Hsiang J. Wrong-level surgery: A unique problem in spine surgery. Surgical Neurology International. 2011; 2:47. [PubMed: 21660270]

Hsu W, Sciubba DM, Sasson AD, Khavkin Y, Wolinsky JP, Gailloud P, et al. Intraoperative localization of thoracic spine level with preoperative percutaneous placement of intravertebral polymethylmethacrylate. Journal of Spinal Disorders & Techniques. 2008; 21(1):72–75. [PubMed: 18418141]

Hurvitz A, Joskowicz L. Registration of a CT-like atlas to fluoroscopic X-ray images using intensity correspondences. International Journal of Computer Assisted Radiology and Surgery. 2008; 3(6): 493–504.

Jin J, Ryu S, Faber K, Mikkelsen T, Chen Q, Li S, et al. 2D/3D image fusion for accurate target localization and evaluation of a mask based stereotactic system in fractionated stereotactic radiotherapy of cranial lesions. Medical Physics. 2006; 33(12):4557–4566. [PubMed: 17278807]

Khamene A, Bloch P, Wein W, Svatos M, Sauer F. Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy. Medical Image Analysis. 2006; 10(1): 96–112. [PubMed: 16150629]

Kubias A, Deinzer F, Feldmann T, Paulus D, Schreiber B, Brunner T. 2D/3D image registration on the GPU. Pattern Recognition and Image Analysis. 2008; 18(3):381–389.

Kwaan MR, Studdert DM, Zinner MJ, Gawande AA. Incidence, patterns, and prevention of wrong-site surgery. Archives of Surgery. 2006; 141(4):353–358. [PubMed: 16618892]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

LaRose D, Bayouth J, Kanade T. Transgraph: Interactive intensity-based 2D/3D registration of x-ray and CT data. Medical Imaging 2000: Image Processing. 2000; 3979(1):385–396.

Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Transactions on Medical Imaging. 1997; 16(2):187–198. [PubMed: 9101328]

Markelj P, Tomaževič D, Likar B, Pernuš F. A review of 3D/2D registration methods for image-guided interventions. Medical Image Analysis. 2012; 16(3):642–661. [PubMed: 20452269]

McLaughlin RA, Hipwell J, Hawkes DJ, Noble JA, Byrne JV, Cox TC. A comparison of a similarity-based and a feature-based 2-D-3-D registration method for neurointerventional use. Medical Imaging, IEEE Transactions on. 2005; 24(8):1058–1066.

Mody MG, Nourbakhsh A, Stahl DL, Gibbs M, Alfawareh M, Garges KJ. The prevalence of wrong level surgery among spine surgeons. Spine. 2008; 33(2):194–198. [PubMed: 18197106]

Mulloy, DF.; Hughes, RG. Wrong-site surgery: A preventable medical error. In: Hughes, RG., editor. Patient safety and quality: An evidence-based handbook for nurses. Rockville (MD): 2008.

Munbodh R, Tagare HD, Chen Z, Jaffray DA, Moseley DJ, Knisely JPS, et al. 2D-3D registration for prostate radiation therapy based on a statistical model of transmission images. Medical Physics. 2009; 36(10):4555–4568. [PubMed: 19928087]

Navab, N.; Bani-Hashemi, A.; Nadar, M.; Wiesent, K.; Durlak, P.; Brunner, T., et al. 3D reconstruction from projection matrices in a C-arm based 3D-angiography system. In: Wells, W.; Colchester, A.; Delp, S., editors. Medical image computing and computer-assisted interventation — MICCAI'98. Springer; Berlin/Heidelberg: 1998. p. 119-129.

Nelder JA, Mead R. A simplex method for function minimization. The Computer Journal. 1965; 7(4): 308–313.

Otake Y, Armand M, Armiger RS, Kutzer MD, Basafa E, Kazanzides P, et al. Intraoperative image-based multiview 2D/3D registration for image-guided orthopaedic surgery: Incorporation of fiducial-based C-arm tracking and GPU-acceleration. Medical Imaging, IEEE Transactions on. 2012; 31(4):948–962.

Otake Y, Schafer S, Stayman JW, Zbijewski W, Kleinszig G, Graumann R, et al. Automatic localization of target vertebrae in spine surgery using fast CT-to-fluoroscopy (3D-2D) image registration. Proc SPIE Medical Imaging 2012. 2012; 8316(1):83160N.

Patil S, Lindley EM, Burger EL, Yoshihara H, Patel VV. Pedicle screw placement with O-arm and stealth navigation. Orthopedics. 2012; 35(1):e61–5. [PubMed: 22229616]

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. Medical Imaging, IEEE Transactions on. 1998; 17(4):586–595.

Pluim JP, Maintz JB, Viergever MA. Image registration by maximization of combined mutual information and gradient information. IEEE Transactions on Medical Imaging. 2000; 19(8):809–814. [PubMed: 11055805]

Powell MJD. An efficient method for finding the minimum of a function of several variables without calculating derivatives. The Computer Journal. 1964; 7(2):155–162.

Qi, W.; Gu, L.; Xu, J. Non-rigid 2D-3D registration based on support vector regression estimated similarity metric. In: Dohi, T.; Sakuma, I.; Liao, H., editors. Medical imaging and augmented reality. Springer; Berlin/Heidelberg: 2008. p. 339-348.

Russakoff DB, Rohlfing T, Mori K, Rueckert D, Ho A, Adler JR Jr, et al. Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration. Medical Imaging, IEEE Transactions on. 2005; 24(11):1441–1454.

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. Computer Vision (ICCV); 2003. p. 416-423.

Russakoff, D.; Rohlfing, T.; Ho, A.; Kim, D.; Shahidi, R.; Adler, J., et al. Evaluation of intensity-based 2D-3D spine image registration using clinical gold-standard data. In: Gee, J.; Maintz, J.; Vannier, M., editors. Biomedical image registration. Springer; Berlin/Heidelberg: 2003. p. 151-160.

Sadowsky, O. PhD Dissertation. the Johns Hopkins University; 2008. Image registration and hybrid volume reconstruction of bone anatomy using a statistical shape atlas.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Sadowsky O, Chintalapani G, Taylor RH. Deformable 2D-3D registration of the pelvis with a limited field of view, using shape statistics. Medical Image Computing and Computer-Assisted Intervention: MICCAI…International Conference on Medical Image Computing and Computer-Assisted Intervention. 2007; 10(Pt 2):519–526. [PubMed: 18044608]

Shams, R.; Barnes, N. Speeding up mutual information computation using NVIDIA CUDA hardware. Digital Image Computing Techniques and Applications, 9th Biennial Conference of the Australian Pattern Recognition Society on; 2007. p. 555-560.

Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. Medical Physics. 1985; 12(2):252–255. [PubMed: 4000088]

Spoerk J, Bergmann H, Wanschitz F, Dong S, Birkfellner W. Fast DRR splat rendering using common consumer graphics hardware. Medical Physics. 2007; 34(11):4302–4308. [PubMed: 18072495]

Studholme C, Hill DLG, Hawkes DJ. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. Medical Physics. 1997; 24(1):25–35. [PubMed: 9029539]

Tomazevic D, Likar B, Slivnik T, Pernus F. 3-D/2-D registration of CT and MR to X-ray images. Medical Imaging, IEEE Transactions on. 2003; 22(11):1407–1416.

Tonn JC, Schichor C, Schnell O, Zausinger S, Uhl E, Morhard D, et al. Intraoperative computed tomography. Acta Neurochirurgica Supplement. 2011; 109:163–167. [PubMed: 20960337]

Upadhyaya CD, Wu J, Chin CT, Balamurali G, Mummaneni PV. Avoidance of wrong-level thoracic spine surgery: Intraoperative localization with preoperative percutaneous fiducial screw placement. Journal of Neurosurgery: Spine. 2012; 16(3):280–284. [PubMed: 22054638]

van de Kraats EB, Penney GP, Tomazevic D, van Walsum T, Niessen WJ. Standardized evaluation methodology for 2-D-3-D registration. IEEE Transactions on Medical Imaging. 2005; 24(9):1177–1189. [PubMed: 16156355]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

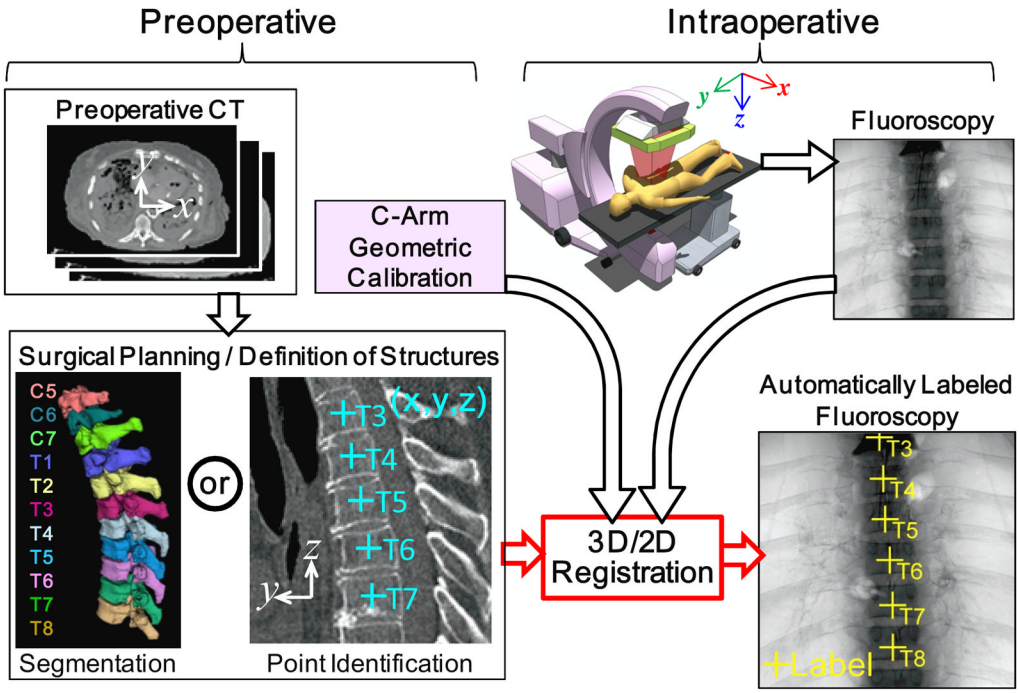

Figure 1.
Overview of the proposed system.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

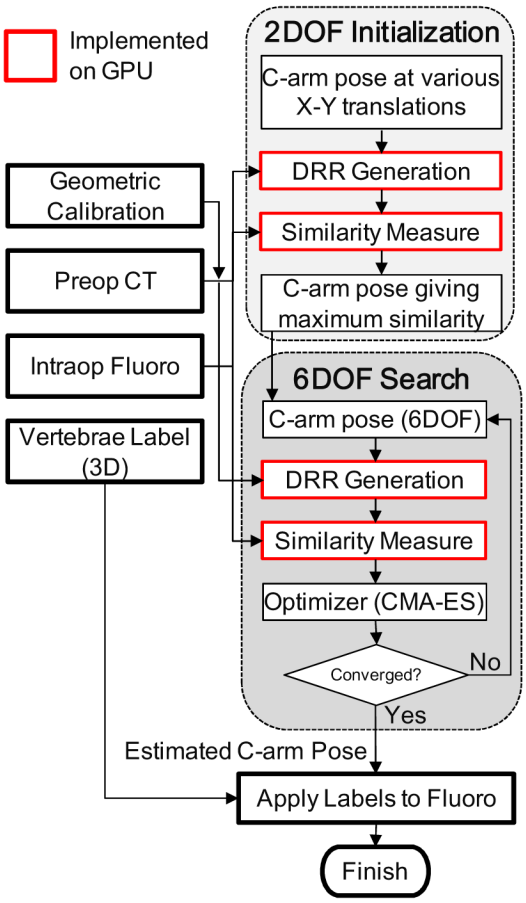

Figure 2.
Workflow diagram of the proposed method showing input data (preop CT data with labels, geometric calibration, and intraoperative fluoroscopy) and two main components of the 3D-2D "LevelCheck" registration approach – an initial 2DOF search and an iterative 6DOF optimization.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

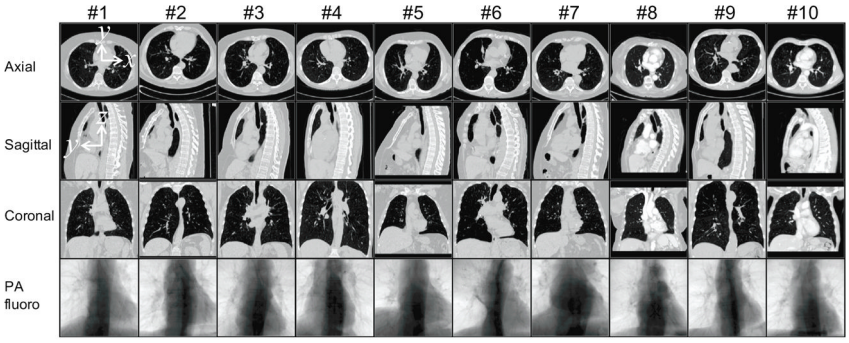

Figure 3.

Clinical CT datasets used in the simulation study. Each image was randomly selected from the NCI TCIA database, providing a fairly broad range in body habitus and normal anatomical variations.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

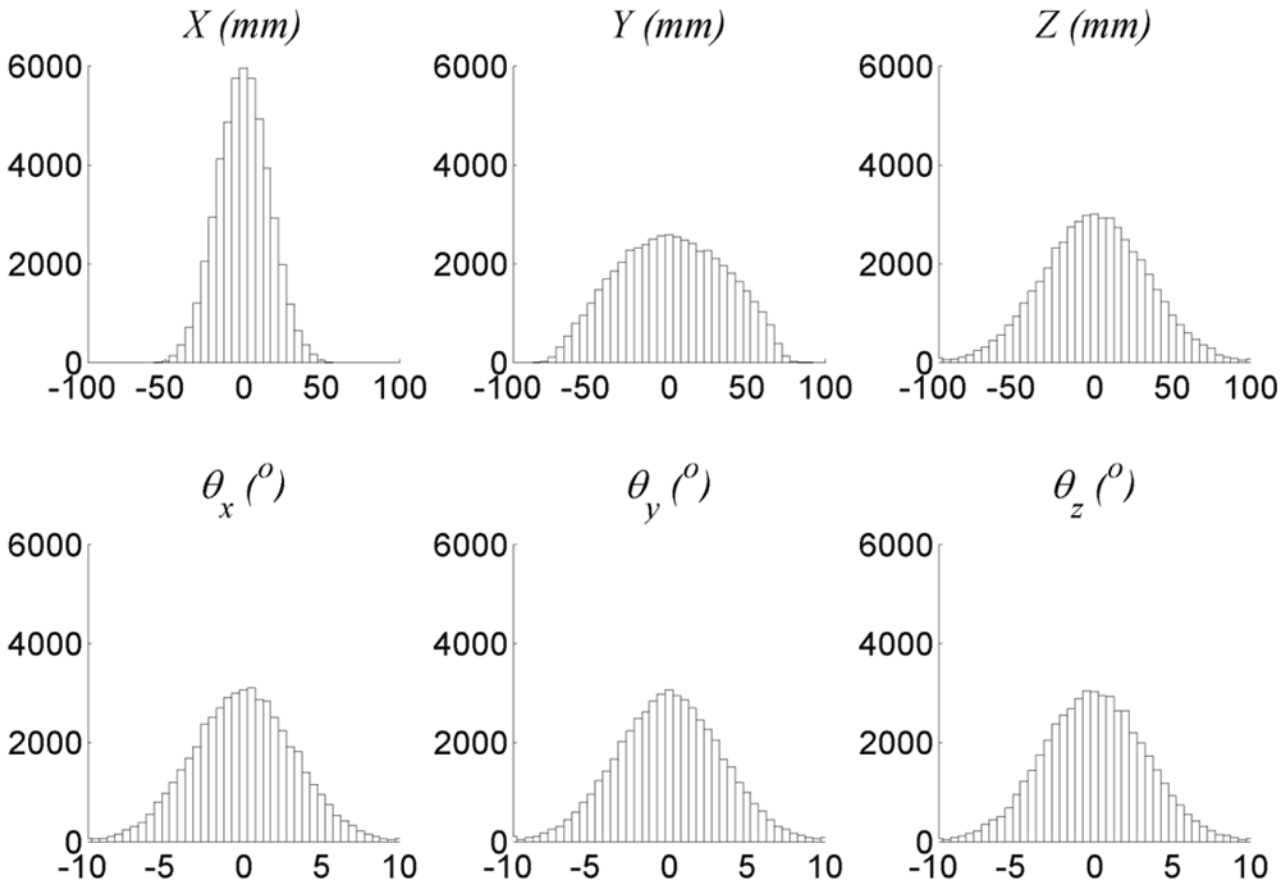

Figure 4.

Histogram of randomly generated 6DOF C-arm pose parameters. A transformation computed from these parameters was used as an offset to the *nominal AP position* to simulate operator variability in C-arm positioning for fluoroscopy. Larger distributions were assumed for translations in Y and Z, since uncertainty along these axes is generally larger than the others.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

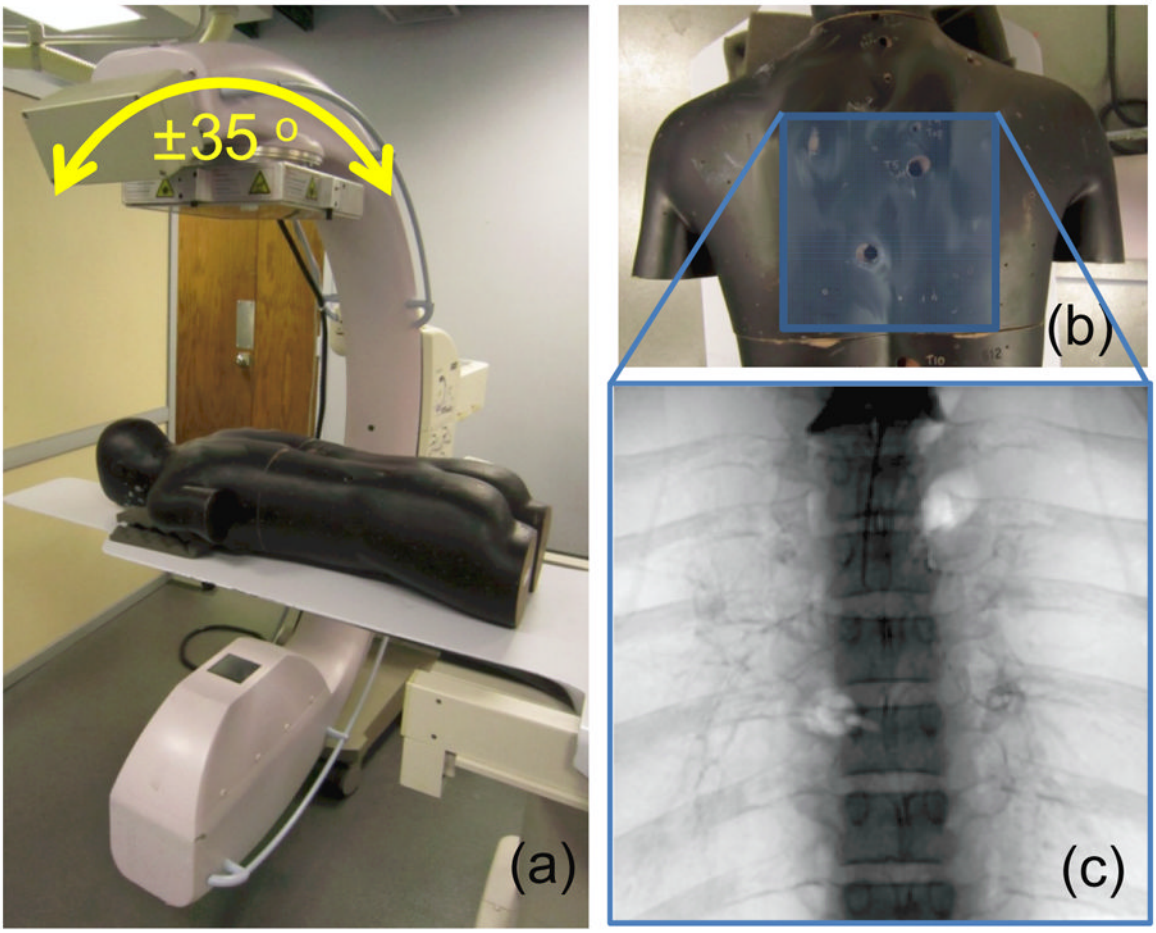

Figure 5.
Setup for experiments using real fluoroscopy images of an anthropomorphic phantom. (a) Mobile C-arm, (b) photograph of the phantom, and (c) example fluoroscopic image from the *nominal AP* view. Fluoroscopic images throughout are displayed in a "black-bone" grayscale typical of fluoroscopic displays.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

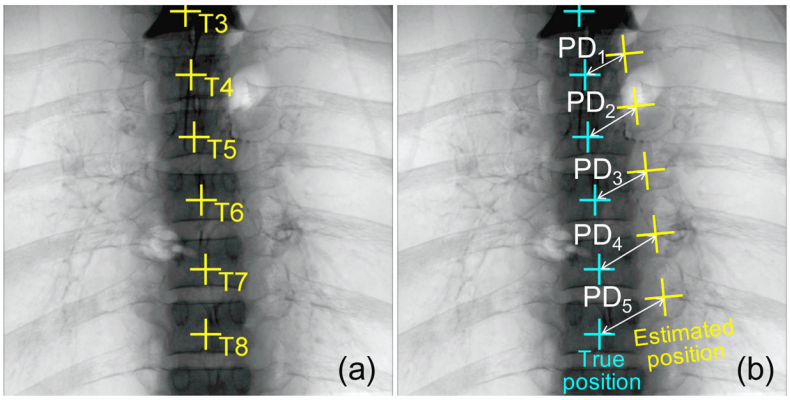

Figure 6.

Evaluation of geometric error in vertebral localization. (a) Estimated vertebral levels (yellow) as labeled by LevelCheck on a real fluoroscopic image. (b) Error in vertebral localization was characterized in terms of the mean projection distance (mPD) between estimated (yellow) and true (cyan) locations in the 2D image plane.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

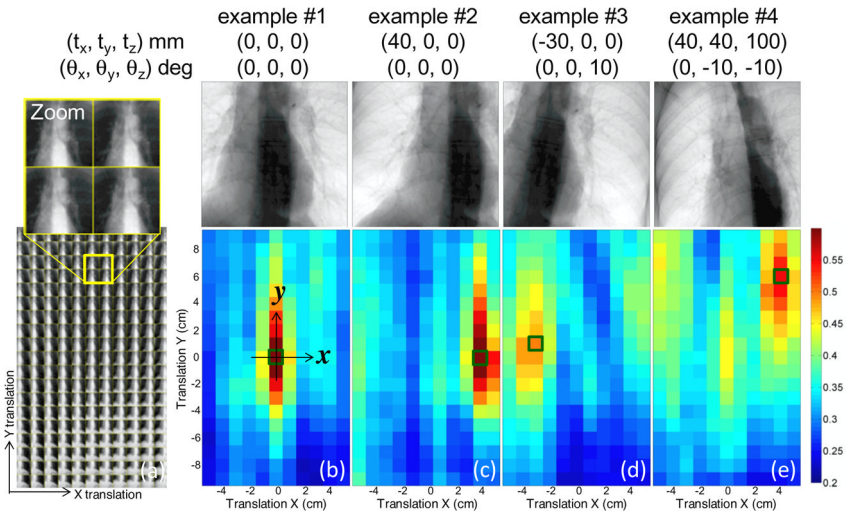

Figure 7.

Initial 2DOF search. (a) DRRs generated from preoperative CT data with varying X and Y translation (10 mm interval) from the *nominal AP position*. The top row shows target fluoroscopy images for four examples at various (true) C-arm poses, and the plot below shows the GI computed as a function of X and Y. (b) Example at the *nominal AP* pose. (c) Example with 40 mm translation in X. (d) Example with X translation and θ_z rotation. (e) Example with a combination of translations and rotations. The more complex poses reduce the magnitude of the GI peak but retain a global maximum at the true pose.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

(a)

Performance	Downhill Simplex	CMA-ES (λ=30)	(λ=60)	(λ=90)	(λ=120)	(λ=150)
# of Failures (in 50,000 trials)	272	32	11	3	1	0
Success Rate (%)	99.456	99.936	99.978	99.994	99.998	100.0
Ave. # of Function Evaluations	472	1105	1675	2079	2425	2783
Ave. Computation Time (sec)	1.79	2.22	3.27	4.04	4.67	5.30

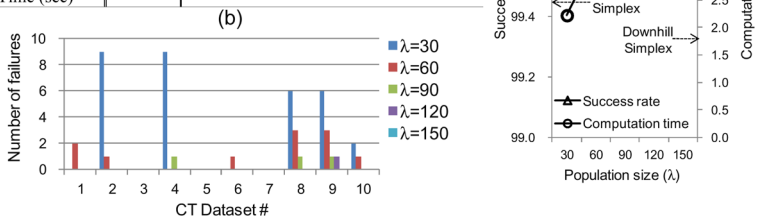

Figure 8.

(a) Success rate and computation time for various settings of CMA-ES population size (λ). (b) Distribution of the number of failures in each of 5,000 trials in each of the 10 CT datasets. (c) Success rate increased asymptotically with λ and approached unity at population size of 120.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

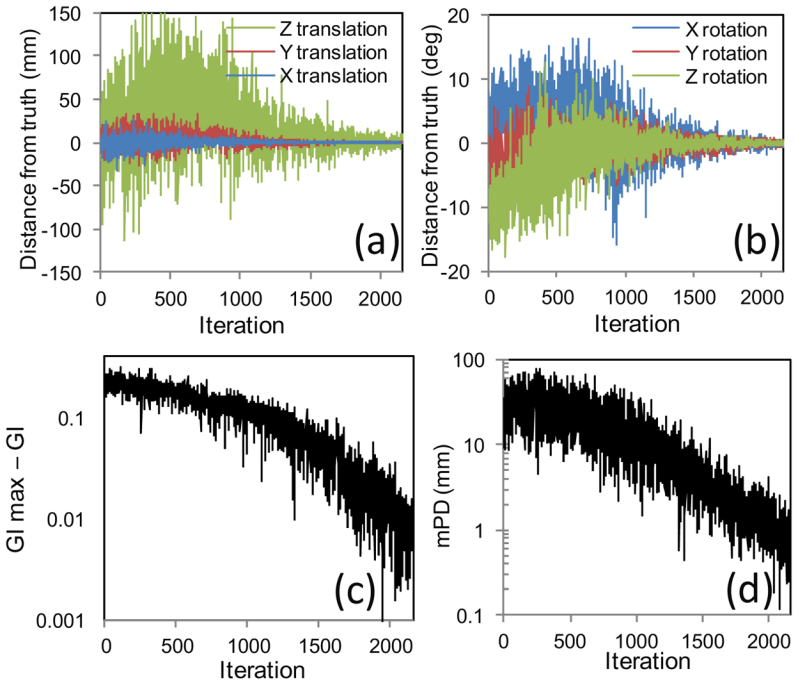

Figure 9.

Convergence plots of a typical registration trial. (a) Translational coordinates of the pose estimate versus iteration number. (b) Rotational coordinates of the pose estimate versus iteration number. (c) Similarity metric plotted versus iteration number, where (GI_{max}–GI) denotes the difference between GI and its maximum value, GI_{max}. (d) Mean projection distance computed versus iteration number. The plots show that mPD reduced to < ~1 mm in ~1000–2000 iterations (which was well within requirements of the spine level labeling application) and could be reduced to still finer levels of accuracy with more iterations.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

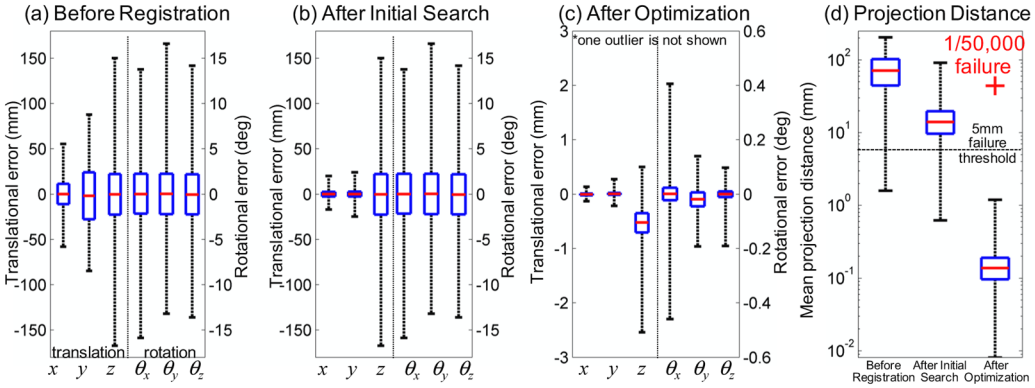

Figure 10.

Summary of registration accuracy in simulation studies. Translation errors (left axis) and rotation errors (right axis) from the *nominal AP position* are shown for cases: (a) before registration, (b) after initial 2DOF search, and (c) after 6DOF iterative optimization search. (d) The mean projection distance before registration, after initial search, and after optimization. The LevelCheck registration was accurate to within ~1 mm overall, with a single failure case for which mPD exceeded 5 mm.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

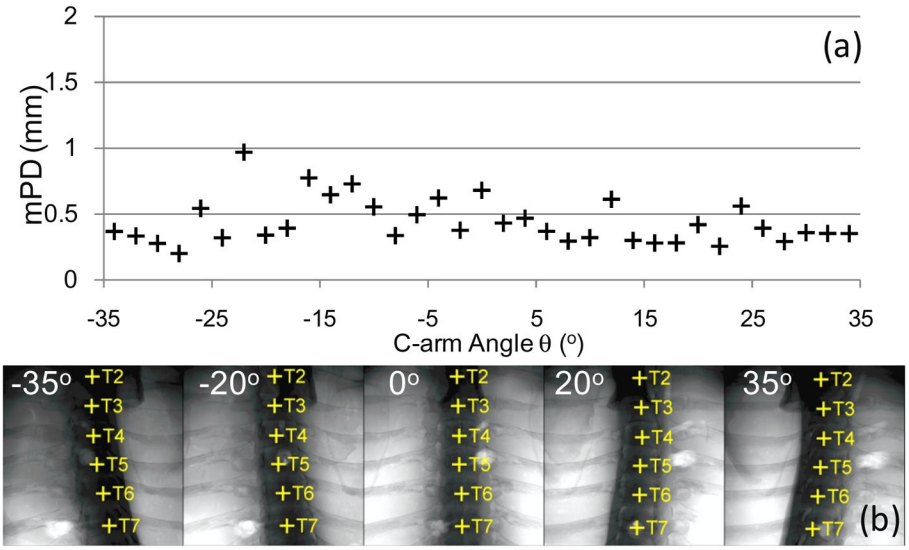

Figure 11.

LevelCheck registration performance in real fluoroscopy. (a) Mean projection distance for images acquired at various C-arm angles, θ, about the *nominal AP view* (b) Fluoroscopy images acquired at θ = −35°, −20°, 0°, 20°, 35° overlaid with the estimated label for each spine level (all of which were within 1 mm of the true level location).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

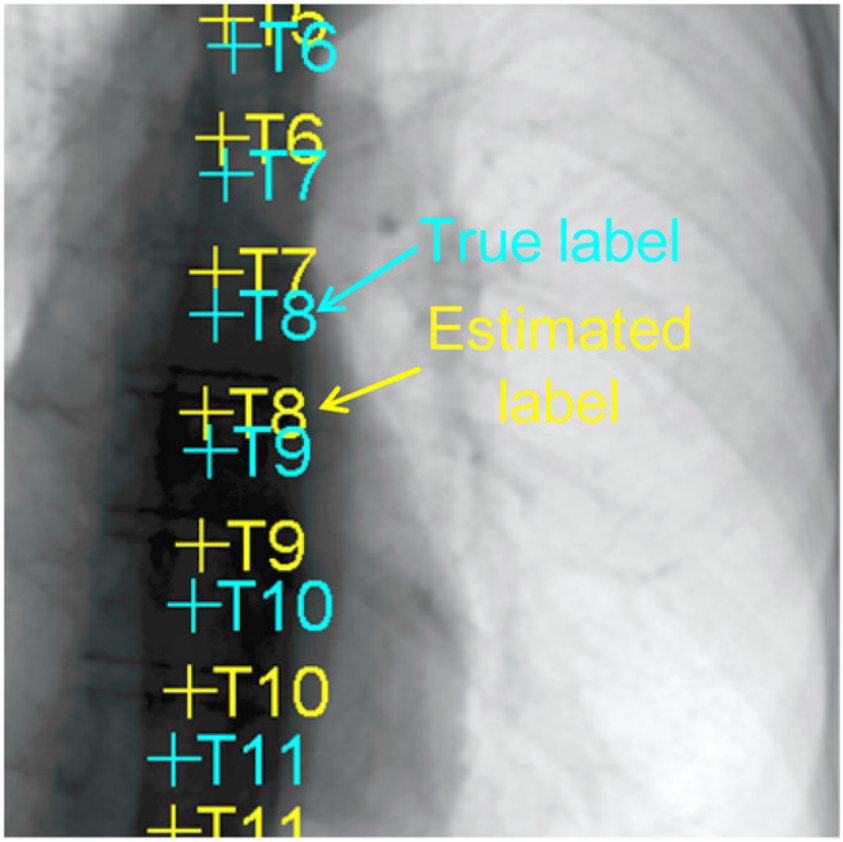

Figure 12.
An example failure case. The image corresponds to the single failure (1/50,000) observed for the CMA-ES optimizer parameter setting λ=120. The overlay of estimated level labels (yellow) are seen to be displaced from the true levels (cyan) by one vertebra.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

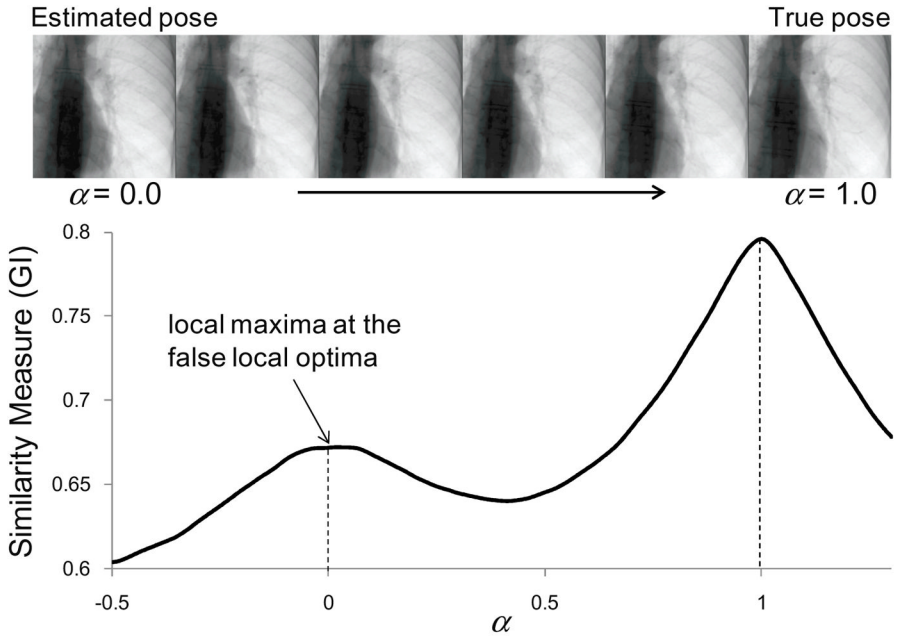

Figure 13.

One slice of the search space in the failed registration shown in figure 12. The plot shows the GI similarity metric between the fluoroscopy and DRRs computed at a pose $T(\alpha)=T_{est}+\alpha(T_{true}-T_{est})$, where T_{est} and T_{true} are the estimated and true pose, respectively (see 2.4. for detail). The shallow local optimum around $\alpha=0$ caused the optimization to converge at the wrong pose for smaller values of CMA-ES population size.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

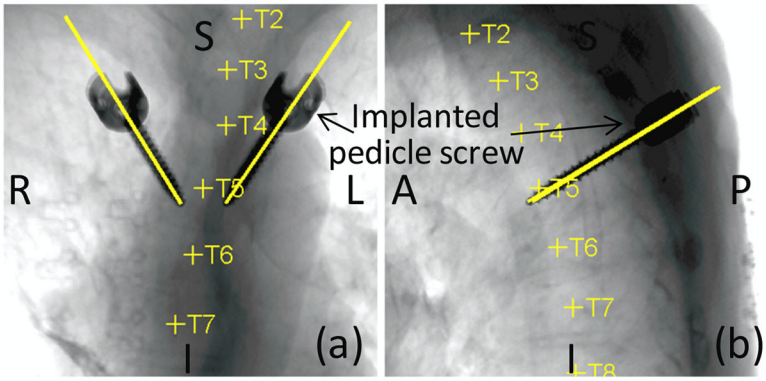

Figure 14.

An example clinical application scenario in pedicle screw placement. Vertebral labels and the planned trajectories defined in the preoperative CT were automatically overlaid onto the simulated fluoroscopy images using the LevelCheck algorithm, (a) AP view and (b) LAT view with LevelCheck labels overlaid in yellow. The algorithm was robust against the presence of screws in fluoroscopy which were present in preoperative CT.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Summary of experimental parameters.

CT Dataset #		CT Volume (voxels)	Voxel size (mm)
1		512×512×733	0.64×0.64×0.50
2		512×512×730	0.73×0.73×0.50
3		512×512×672	0.61×0.61×0.50
4		512×512×682	0.54×0.54×0.50
5		512×512×636	0.69×0.69×0.50
6		512×512×588	0.51×0.51×0.50
7		512×512×559	0.69×0.69×0.70
8		512×512×590	0.63×0.63×0.45
9		512×512×633	0.54×0.54×0.50
10		512×512×580	0.61×0.61×0.45
Hardware specification			
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	
Optimization parameters			
Downsampled resolution	2DOF	2D projection: 48×48 pixels	
	Initialization	3D volume: 64x64x90 voxels	
	6DOF Search	2D projection: 96×96 pixels	
		3D volume: 128×128×180 voxels	
CMA-ES population size (λ)		120	
Initial search distribution (σ) (x, y, z, θ_x, θ_y, θ_z)		(10 mm, 10 mm, 40 mm, 3 °, 3 °, 3 °)	

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 2

Summary of simulation study results from 50,000 trials (5,000 trials × 10 CT datasets).

CT Dataset #	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	Total
# of Failures	0	0	0	0	0	0	0	0	1	0	1
Success Rate (%)	100	100	100	100	100	100	100	100	99.98	100	99.998
Ave. # of Function Evaluations	2414	2430	2397	2404	2463	2463	2371	2451	2434	2427	2425
Ave. Computation Time (sec)	4.70	4.73	4.66	4.63	4.77	4.64	4.57	4.69	4.74	4.53	4.67