HHS Public Access

Author manuscript

Phys Med Biol. Author manuscript; available in PMC 2016 June 21.

Published in final edited form as:

Phys Med Biol. 2013 December 7; 58(23): 8535–8553. doi:10.1088/0031-9155/58/23/8535.

Robust 3D–2D image registration: application to spine interventions and vertebral labeling in the presence of anatomical deformation

Yoshito Otake[1,2], **Adam S Wang**[1], **J Webster Stayman**[1], **Ali Uneri**[2], **Gerhard Kleinszig**[3], **Sebastian Vogt**[3], **A Jay Khanna**[4], **Ziya L Gokaslan**[5], and **Jeffrey H Siewerdsen**[1,2]

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

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

[3]Siemens Healthcare, Erlangen, Germany

[4]Department of Orthopaedic Surgery, Johns Hopkins University, Baltimore MD, USA

[5]Department of Neurosurgery, Johns Hopkins University, Baltimore MD, USA

Abstract

We present a framework for robustly estimating registration between a 3D volume image and a 2D projection image and evaluate its precision and robustness in spine interventions for vertebral localization in the presence of anatomical deformation. The framework employs a normalized gradient information similarity metric and multi-start covariance matrix adaptation evolution strategy optimization with local-restarts, which provided improved robustness against deformation and content mismatch. The parallelized implementation allowed orders-of-magnitude acceleration in computation time and improved the robustness of registration via multi-start global optimization. Experiments involved a cadaver specimen and two CT datasets (supine and prone) and 36 C-arm fluoroscopy images acquired with the specimen in four positions (supine, prone, supine with lordosis, prone with kyphosis), three regions (thoracic, abdominal, and lumbar), and three levels of geometric magnification (1.7, 2.0, 2.4). Registration accuracy was evaluated in terms of projection distance error (PDE) between the estimated and true target points in the projection image, including 14 400 random trials (200 trials on the 72 registration scenarios) with initialization error up to ±200 mm and ±10°. The resulting median PDE was better than 0.1 mm in all cases, depending somewhat on the resolution of input CT and fluoroscopy images. The cadaver experiments illustrated the tradeoff between robustness and computation time, yielding a success rate of 99.993% in vertebral labeling (with `success' defined as PDE <5 mm) using 1,718 664 ± 96 582 function evaluations computed in 54.0 ± 3.5 s on a mid-range GPU (nVidia, GeForce GTX690). Parameters yielding a faster search (e.g., fewer multi-starts) reduced robustness under conditions of large deformation and poor initialization (99.535% success for the same data registered in 13.1 s), but given good initialization (e.g., ±5 mm, assuming a robust initial run) the same registration could be solved with 99 993% success in 6.3 s. The ability to register CT to fluoroscopy in a manner robust to patient deformation could be valuable in applications such as

jeff.siewerdsen@jhu.edu

radiation therapy, interventional radiology, and an assistant to target localization (e.g., vertebral labeling) in image-guided spine surgery.

1. Introduction

3D–2D image registration computes the transformation between a 3D volume and a 2D projection image. In image-guided intervention, it has been used to incorporate prior information (e.g., preoperative CT-based planning data) into intraoperative x-ray fluoroscopy. Applications of 3D–2D registration include improving target localization accuracy in radiation therapy (Gendrin *et al* 2012), real-time kinematic analysis of skeletal movements using fluoroscopy (Xin *et al* 2013, Zhu *et al* 2012, Jerbi *et al* 2012), guidance of various surgical interventions such as cardiac surgery (Rivest-Henault *et al* 2012), and transrectal ultrasound guided surgery (De Silva *et al* 2013). Recent application includes spine surgery (Otake *et al* 2012c) where automatic labeling of vertebral levels in intraoperative fluoroscopy was used to localize the target vertebrae and help prevent wrong-site surgery (Palumbo *et al* 2013). The lack of uniquely identifiable features and the periodic appearance of spine levels in projection images makes level localization especially challenging and even more so in minimally invasive surgery where the surgeon operates thorough a small incision. The reported incidence of wrong-level error is estimated to be 1 in every 3,110 spine surgeries (Mody *et al* 2008), with an approximate financial cost of $55 000 per case according to The Joint Commission. Initial implementation of the `LevelCheck' method (Otake *et al* 2012c) involved a single-image 3D–2D registration applied to spine interventions and validated its performance to automatically label spine levels in fluoroscopy images using simulation and rigid physical phantom studies.

3D–2D registration has been extensively investigated in various technical aspects including image modalities, methods for achieving dimensional correspondence, similarity metrics, and registration basis (e.g., extrinsic, feature-based, intensity-based and gradient-based methods) (Markelj *et al* 2012). Intensity-based approaches utilize all the intensity information contained in the 2D and 3D images, as opposed to the feature-based approach where, image features extracted from 2D and 3D images are used for the registration, and the gradient-based approach which utilizes the fact that rays connecting the x-ray source and edges in the projection image are tangent to surfaces of distinct anatomical structures, and have demonstrated higher accuracy and reliability (McLaughlin *et al* 2005, Mitrovic *et al* 2013). An intensity-based 3D–2D registration can be formulated as a numerical optimization, which solves for the transformation parameters (i.e., six degree-of-freedom (DOF) transformation in a rigid case) that maximize the similarity metric (objective function) between the 2D projection image and a digitally reconstructed radiograph (DRR) of the 3D volume generated based on the transformation.

Evaluation of the objective function is computationally expensive due to the projection operation in DRR generation, and improvements in computational efficiency have been an active area of research with proposed solutions including various software-based approaches (Birkfellner *et al* 2005, Russakoff *et al* 2005) that improved computation speed at the expense of slight degradation of image quality and hardware-based acceleration approaches

using GPU (Tornai *et al* 2012, Spoerk *et al* 2012). The latter approach with a relatively straightforward software implementation was employed in this study. The proliferation of GPU or alternative parallel computation environments (CPU/GPU clusters etc) and the fact that simplifying and scaling down the core kernel of the implementation generally improves parallelization efficiency offered a better load balancing which improves occupancy of processors.

Various types of similarity metrics have been proposed in the literature, and may be grouped according to the correspondence of global intensity such as mutual information (Maes *et al* 1997), normalized cross-correlation (Penney *et al* 1998), and metrics based on local intensity such as pattern intensity (Penney *et al* 1998), gradient difference (Penney *et al* 1998), gradient correlation (Hipwell *et al* 2003) and/or gradient information (GI) (Pluim *et al* 2000). Although the best performing similarity metric is application-dependent, comparative studies (Birkfellner *et al* 2009, Russakoff *et al* 2003) suggest that local metrics tend to demonstrate better performance. In this study, we normalized the conventional GI metric to a range of [0.0, 1.0] with 1.0 meaning the two images are identical. This normalized GI (NGI) provided an absolute measure of correspondence of edges between two images. The GI metric was previously shown to be robust against content mismatch between two images (Otake *et al* 2012c). However, regardless of the similarity metric used, it is inevitable that a particular transformation—different from the true registration—may happen to yield a local optimum, for example, in the case where the image contains symmetric or periodic structures (such as vertebrae) or when there is anatomical deformation. As a result, the objective function becomes noisy (i.e., a large number of false local optima), and the optimization algorithm must be able to localize the global optimum from a noisy objective function. In this paper, we use the term robustness to refer to the degree to which the registration algorithm can contend with a difficult and noisy optimization—including local optima—and converge upon the true transformation.

Both derivative-based optimization methods (e.g., gradient descent (McLaughlin *et al* 2005)) and derivative-free optimization (Powell 1964), downhill simplex (Nelder and Mead 1965), and the covariance matrix adaptation evolution strategy (CMA-ES) (Hansen and Kern 2004) have been used in solving 3D–2D registration. In comparison to the derivative-based approach and the classic derivative-free approaches, an evolutionary strategy such as CMA-ES tends to be more robust against local optima while requiring a larger number of function evaluations, such that robustness generally improves by tuning its optimization parameters to perform more function evaluations (e.g., increasing the population size).

Since evaluation of the objective function can be costly in 3D–2D registration, computation time and robustness of the registration form a tradeoff. Several recent studies focusing on fast implementation of 3D–2D registration reported computation speed in terms of the number of function evaluations per second (eval s^{-1}). For example: Otake *et al* (2012c) reported ~500 eval s^{-1}, Fisher *et al* (2013) reported ~400 eval s^{-1} (DRR rendering speed), Dorgham *et al* (2012) reported ~167 eval s^{-1}, Varnavas *et al* (2013) reported ~21.5 eval s^{-1}, Steininger *et al* (2012) reported ~13.3 eval s^{-1}, and Spoerk *et al* (2012) ~9–22 eval s^{-1}. A fair comparison of computational performance is difficult, since image resolution, the approach to DRR generation, and similarity metric significantly vary in each method. Most

reported approaches required ~50–2000 function evaluations in one optimization, and ongoing improvements in both GPU speed and the form a particular 3D–2D registration implementation suggest the capability for near-real-time 3D–2D registration. For example, in the work reported below, we report an implementation offering >22 000 eval s^{-1} and registration in <6.3 s in cases involving a constrained search space. However, the robustness of the optimization is still a major factor that hinders routine clinical use of intraoperative 3D–2D registration.

The challenge in 3D–2D registration in a clinical scenario featuring anatomical deformation is illustrated in figure 1, causing the optimization to become ill-posed due to a reduction in the peak magnitude of the objective function, the emergence of local optima, and insensitivity to pertinent DOFs (e.g., out-of-plane translation and/or rotation about rotationally symmetric structures). A further complexity that can challenge intraoperative 3D–2D registration is mismatch in image content (e.g., surgical instrumentation in the intraoperative image but not in the preoperative CT). Previous work (Otake *et al* 2012c) demonstrated the LevelCheck algorithm to be reasonably robust in each of the requisite DOFs by virtue of invariance against coordinate-wise scaling in CMA-ES, and robustness to image content mismatch is in part achieved in that the GI similarity metric does not weigh inconsistent gradients (i.e., disregards gradients evident in only one of the images).

In this paper, we extend the previously reported method as follows. (1) To further improve robustness against initialization errors and anatomical deformation, a global multi-start strategy is implemented in the optimization which was facilitated by orders-of-magnitude acceleration in computation time compared to the original implementation. (2) We evaluate the method in cadaver experiments in the presence of gross anatomical deformation approximating that in spine surgery (viz, prone versus supine, kyphosis, lordosis, and motion of internal soft-tissue and air-filled structures).

2. Methods

2.1. 3D–2D registration method

2.1.1. Implementation strategy—The 3D–2D registration process is summarized in figure 2. As detailed below, the DRR and similarity metric were computed on GPU, while the CMA-ES optimization was implemented on CPU since computation of the covariance matrix update and generation of random samples are inherently sequential. To minimize CPU–GPU memory copies, no image was transferred during the optimization iteration. Only the transformation T (a single-precision, 6-element vector, 6 × 4 = 24 bytes) was transferred from CPU to GPU in each function evaluation, and the computed similarity metric (a single-precision scalar, 4 bytes) was copied back from GPU to CPU.

2.1.2. Multi-resolution pyramid—A multi-resolution pyramid (Munbodh *et al* 2009) was employed in the optimization to improve the convergence rate and help avoid local optima; however, we also found that downsampling the image can suppress certain image features and make other non-corresponding features more similar, thus creating false local optima in the objective function. We analyzed various 2D projection-domain downsampling schedules of 8 (96^2 pixels, 3.1^2 mm^2 $pixel^{-1}$), 4 (192^2 pixels, 1.3^2 mm^2 $pixel^{-1}$) and 2 (384^2

pixels, 0.78^2 mm^2 pixel^{-1}), referred to as DS8, DS4, DS2 respectively. The volume was downsampled accordingly, resulting in the volume-domain pyramid: 128^3 voxels, 2.7^3 mm^3 voxel^{-1} (DS8); 256^3 voxels, 1.3^3 mm^3 voxel^{-1} (DS4); and 512^3 voxels, 0.68^3 mm^3 voxel^{-1} (DS2), respectively.

2.1.3. DRR generation—A traditional grid-interpolation ray-tracing algorithm (Cabral *et al* 1994) was employed for DRR generation. Each ray was computed on a single CUDA thread and traced the volume at a certain interval (step length). The intensity at each sample point was computed by interpolation of eight neighboring grid points accumulated to obtain a line integral along the ray. The volume image and the projection matrices resided on texture memory to benefit acceleration by the cache mechanism on GPU. Multiple DRRs were generated simultaneously using a 3-dimensional CUDA block. Since the projection views in neighboring iterations were close to each other in 3D–2D registration, concurrent ray-tracing of pixels at the same location in neighboring projections improves the cache hit rate, and thus is advantageous especially when multiple DRRs were generated concurrently in parallel function evaluations. The step length was chosen as 2.0 voxels to achieve faster computation while not degrading accuracy, since previous work (Otake *et al* 2012a) showed that registration accuracy was nearly constant up to a step length of ~3.0 voxels.

2.1.4. Similarity metric—We implemented a normalized version of GI (Pluim *et al* 2000) as follows:

$$\mathrm{NGI}\left(I_0, I_1\right) = \frac{GI\left(I_0, I_1\right)}{GI\left(I_1, I_1\right)}. \quad (1)$$

Where the numerators is the usual GI between the fixed image and the moving image, and the denominator is the GI between the fixed image and itself. In each case, GI was defined as:

$$GI\left(p_1, p_2\right) = \sum_{i,j \in \Omega} m\left(i,j\right) w\left(i,j\right) \min\left(\left|\nabla p_1\left(i,j\right)\right|, \left|\nabla p_2\left(i,j\right)\right|\right), \quad (2a)$$

$$\nabla p\left(i,j\right) \triangleq \left(\frac{\partial}{\partial i} p\left(i,j\right), \frac{\partial}{\partial j} p\left(i,j\right)\right), \quad (2b)$$

$$w\left(i,j\right) = \frac{\alpha_{i,j}+1}{2}, \qquad \alpha_{i,j} = \frac{\nabla p_1\left(i,j\right) \cdot \nabla p_2\left(i,j\right)}{\left|\nabla p_1\left(i,j\right)\right| \left|\nabla p_2\left(i,j\right)\right|}. \quad (2c)$$

The range of NGI is therefore 0 – 1 and the case NGI = 1.0 means that the moving image gradients are identical to those in the fixed image.

All computations, except the image gradient (∇p) and the summation (Σ), were implemented as pixel-wise operations where each pixel was assigned to a single CUDA thread, providing a highly efficient parallelization (i.e., a large number of small, concurrently executable computations). The image gradient (∇p) along each axis was implemented as the kernel-based Gaussian derivative filter (Florack *et al* 1994) with $\sigma = 1.0$. The kernel convolution was implemented as a product of the Fourier transforms of the kernel and the target image. The Fourier transform of multiple target images was computed in parallel using the batch execution in the CUFFT library. Summation (Σ) of the pixel values over each image was implemented as a matrix-vector multiplication using cublasSgemv() in the CUBLAS library to achieve further speed-up (i.e., $M \times u^2$ matrix times $u^2 \times 1$ vector, where u^2 is the total number of pixels in one image and M is the number of images).

A spatial weighting mask, $m(i, j)$, could be optionally applied either to more heavily weigh NGI within a region of interest in which the target anatomy is clearly distinguishable or known *a priori* (e.g., the region containing the spine, typically occupying the central region of the image in AP fluoroscopy) and/or to diminish contributions to NGI from regions believed to present major mismatch in content or deformation (e.g., the skinline or a gas bubble). For spine images as in the application below, a weighting mask with a higher value along the center line of the spine and falling off at the edges of the image was optionally employed. The mask was taken as a Gaussian nominally centered on the projection with variable width (ranging 10–25% of the image width) tailing off laterally (x) and constant longitudinally (z). Turning off the weighting mask amounts simply to $m(i, j) = 1.0 \; \forall \; i, j$. The mask was implemented as a simple pixel-wise multiplication and therefore had almost no effect on computation speed.

The effect of the weighting mask on the objective function space was analyzed in a cadaver study in which deformation of a large, internal air-filled structure (gas pocket) shifted the global optimum from the true transformation (T_{true}) to a false local optimum (T_{false}).

2.1.5. Optimization—The 3D–2D rigid registration optimization is formulated as follows:

$$\hat{T} = \underset{T \in S_T}{argmax} \, \text{NGI}\left(I_1, I_0\left(T\left(t_x, t_y, t_z, \theta_x, \theta_y, \theta_z\right)\right)\right) \tag{3}$$

where T represents the six DOF transformation parameters, including translation along the x, y, z axes (t_x, t_y, t_z) and rotation around the x, y, z axes (θ_x, θ_y, θ_z), and S_T is the space of allowable transformations. We addressed this optimization problem using a derivative-free global optimization approach.

Global optimization approaches have been employed in 3D–2D registration in prior work, including the exhaustive search approach as in Birkfellner *et al* (2003), where the objective function is evaluated at a few (3–5) grid points in each of six directions with three resolution levels selecting the point with the highest function value as the solution. Alternatively, Mahfouz *et al* (2003), Dennis *et al* (2005), and Turgeon *et al* (2005) used a modified

downhill simplex optimization in which the optimization search is restarted after convergence or a randomly perturbed starting point is used. In Dey and Napel (2006), a Monte Carlo-based optimization used an initial estimate pose provided manually or based on prior knowledge. All of these approaches were shown to improve the global search performance over a conventional local search.

Otake *et al* (2012c) proposed an automatic initialization method based on an exhaustive search of 2D planar grid points at an (~20 mm) interval in-plane to the projection image, assuming rotation and out-of-plane translation to be zero (i.e., at the iso-center, with geometric magnification = 2.0). The objective function was evaluated at the ~200 grid points, and the point scoring the highest objective value was chosen as the initialization. While this simple initialization worked well in a rigid phantom, it was subject to false local optima in cases of anatomical deformation with a large number of local optima. The method proposed here extends the concept to multiple CMA-ES searches starting from a collection of 3D grid points.

The global optimization strategy consists of two stages: (1) a global multi-start; and (2) a local restart. The multi-start strategy is traditionally used in global optimization problems to broaden the capture range. Rinnooy Kan and Timmer (1987) theoretically proved that a stochastic optimization method converges to the global optimum with probability 1.0 when the number of function evaluations is increased infinitely. Analogously, the local restart strategy was employed in Auger and Hansen (2005b, 2005a), where application to CMA-ES demonstrated improved performance with the IEEE CEC 2005 benchmark functions (Suganthan *et al* 2005). In this study, the basic CMA-ES algorithm was implemented in MATLAB as in Hansen (2006), including three tuning parameters—population size (λ), upper/lower bounds of *x*, *y*, *z* translation (mm) and *x*, *y*, *z* rotation (deg) (S_T in equation (3)), and the stopping criterion (chosen as the tolerance in the pose parameter space (mm or deg)). We stopped the iteration when the standard deviation of the sample distribution in one generation became smaller than the chosen threshold in each coordinate. The initial search distribution was set as 20% of the range of bounds (upper – lower bound) as suggested by Hansen (2006). The two-staged global optimization is detailed as follows.

> *(1) Global multi-start*. The basic global multi-start strategy is a *divide-and-conquer* approach in which the entire 3D search space is partitioned into multiple subspaces (*divide*), and a separate CMA-ES search is run in each subspace (*conquer*). A variety of space-partitioning algorithms can be employed in the division step. In this study, we employed a standard kD-tree partitioning strategy (Bentley 1975) in 3D, where the largest box (space) was recursively partitioned into two boxes at the middle of its longest edge until the number of boxes reaches a desired number, *N* (equal to the number of multi-starts). A larger value of *N* confines each local search to a smaller space, thus making each local search more robust, whereas it increases overall computation time in a manner proportional to *N*. A range *N* = 1 to 250 was chosen in the current implementation to limit computation time to a few seconds, compatible with intraoperative use. The center of each box (leaf of the kD-tree) was used as an initialization for each local search. Rotational elements of the initialization were determined separately from translation as uniformly distributed

random 3D vectors with specified upper and lower bounds, because the rotational search range was small in our application (±10° was generally sufficient as detailed in below) and a 6-dimensional kD-tree with a small number of leaf nodes (e.g., N = 250) does not provide efficient partitioning of the space (e.g., 3^6 = 729 nodes required to partition a 6-dimensional space into three partitions in each dimension). The upper and lower bounds of translational elements in each local search were set to twice the extent of the box. This created an overlap of 26 adjacent search spaces, each of which was initialized with a random rotation, providing increased robustness since the random rotation increased the chances that one of the 26 rotations is close to the true registration (i.e., good initialization). The N independent CMA-ES searches were performed in parallel by storing multiple covariance matrices and evolution paths throughout the optimization. This increased the number of parallelizable function evaluations significantly (λN parallel evaluations) compared to the conventional sequential CMA-ES (λ parallel evaluations), improving performance.

(2) Local restart. To improve the robustness of each local search (i.e., the *conquer* step), another CMA-ES search was performed using the converged point as the starting point (i.e., a restart) with the same optimization parameters as the previous step. Although a larger number of restarts would further improve robustness, we performed only one restart to limit computation time. Finally, the resulting local optimum scoring the best objective value was selected, and another search with a higher resolution level was performed.

2.2. Experiments

2.2.1. Cadaver setup and deformation scenarios—As illustrated in figure 3, a fresh (unfrozen, unfixed) cadaver was imaged with a prototype mobile C-arm (Siemens Healthcare, Erlangen, Germany) incorporating a flat-panel detector (FPD) and motorized rotation. Geometric calibration was performed with a previously described phantom (Navab *et al* 1998). The FPD (PaxScan 3030 CB, Varian Imaging Products, Palo Alto, CA, USA) provided distortion-free readout (30^2 cm^2 FOV, 768^2 pixel format, 0.388^2 mm^2 pixel^{-1}) at 3.3 frames per second, and the imaging technique was 100 kVp, 2.05 mAs per projection. A workstation with an Intel Xeon 2 processor (2.4 Gz) with Windows 7 64-bit and GTX690 was used for the experiment.

Two preoperative CT images of the cadaver were acquired—one in supine position (referred to as `supine CT') and the other in prone position (`prone CT'). Details of the volumes are listed in table 1. A series of C-arm projection images were acquired, including three regions of interest (thoracic, abdominal, and lumbar), three geometric magnifications (m = 1.7, 2.0, 2.4), and two positions (supine, denoted by the symbol `↑', and prone, denoted `↓'). In addition, two types of anatomical deformation were imparted using a form block (figure 3(c))—lordosis in supine position (denoted `↑⋂') and kyphosis in prone position (denoted `↓⋂'). In total, 36 projection images were thus acquired, and 3D–2D registration was performed on each image using the two CT datasets, resulting in 72 CT-projection pairs, comprising the eight deformation scenarios summarized in table 1. For example: scenario #1 involved supine CT (↑) to supine projection (↑), representing a case of minimal deformation;

conversely, scenario #8 involved prone CT (↓) to supine projection with kyphosis (↑⋂), representing a case with large anatomical deformation. The scenarios were numbered in increasing rank order approximating the magnitude of deformation.

We defined target points in each vertebra (19 vertebrae from C6 to L5) in each CT dataset. To consistently define the points relative to each vertebral structure in two CTs acquired in different positions, we segmented each vertebra separately (figures 3(d) and (e)) and the geometrical centroid of each segmented vertebra was used as the target point. Note that in a real clinical workflow, the target point can simply be a series of points identified in CT (e.g., a single point for each vertebra).

Movement of each vertebral joint between prone CT and supine CT was computed by rigid ICP registrations (Besl and McKay 1992) between each segmented vertebra. Larger deformations (±3°–5° at each joint) were observed in the lumbar region, while deformations in thoracic and abdominal regions were smaller (<2° at each joint) likely due to constraint by the rib cage.

2.2.2. Definition of truth—Ground truth registration was defined by way of consistency analysis (Jenkinson *et al* 2002), in which multiple, independently performed registrations converged to the same result, and that result was judged sensible as assessed by a clinical expert. We note that consistency is a surrogate for truth and a necessary (but not sufficient) condition for truth definition appropriate to real data (cf, image simulation).

We performed two independent 3D–2D registrations between each projection image and supine CT and prone CT. For each projection, we considered the registration to be true when the two registrations showed consistent results in terms of: (1) visual verification by an expert operator; and (2) projection distance error (PDE) (in mm) as detailed in section 2.2.3. Visual verification was performed by checking the consistency of edges in the NGI images between the fixed image and DRR at the solution. Each registration for truth definition was manually initialized close to the true alignment and optimized with parameters tuned for robustness (i.e., a large number of multi-starts, requiring ~300–600 s) with a small search range (±30 mm, ±5°) and using a weighting mask (section 2.1.4). Under such conditions, all cases demonstrated PDE < 3 mm and were visually verified as correct alignment. Figure 4 illustrates an example registration in such definition of truth.

2.2.3. Analysis, precision and robustness—The registration accuracy was quantified in terms of PDE (Van De Kraats *et al* 2005), defined as follows to confine the target points to those appearing in the image:

$$\mathrm{PDE} \triangleq \frac{1}{K}\sum_{i=1}^{K}\mathrm{PDE}_i \qquad (4)$$

where PDE_i is PDE at the ith target point, and K is the number of target points appearing in the image. We defined the success criterion as PDE < 5.0 mm as in Otake *et al* (2012c), and the success rate as the percentage of successful registrations within a given trial.

Precision was analyzed in terms of the reproducibility in the result, which for stochastic optimization algorithms, such as CMA-ES, is not guaranteed. 100 registration trials were performed on each projection-CT pair (100 trials × 72 pairs = 7200 registrations) with initialization including a perturbation to the true registration (±5 mm, ±2° uniformly random distribution). Parameters for the CMA-ES optimizer were: population size 20, stopping criterion 0.001 (mm or deg), upper/lower bounds of the search ±10 mm, ±4°. A (deterministic) downhill simplex optimization was also performed as a comparison, with the initial simplex size of ±10 mm in translation, ±4° in rotation and a stopping criterion of 0.001 (mm or deg). For fair comparison, the number of function evaluations was fixed in the CMA-ES search (~3,000) and starting multiple simplex searches from eight randomly selected starting points.

In addition, the dispersion (Lunacek and Whitley 2006) of the objective function was computed to quantify reproducibility in computing the local optimum. We evaluated the objective function at 2×10^5 uniformly distributed random sample points within ±0.1 mm, ±0.1° around the true registration and ranked the points according to the function value. Dispersion was defined as the mean pair-wise distance between M best points (with M varied from 6 – 25 000) as follows:

$$\text{dispersion} \triangleq \frac{1}{M\,(M-1)} \sum_{i=1}^{M} \sum_{j=1}^{M} d\,(x_i, x_j) \qquad (5)$$

where $d(a, b)$ is the L2-norm between vectors a and b. If the dispersion decreases as the sample is restricted to better regions (i.e., decreasing M), the function is considered to have *low dispersion* and suggests smooth convergence at the solution. On the other hand, if the dispersion is constant or increases (*high dispersion*), it suggests that local optima around the solution are broadly dispersed, implying a higher chance that the optimization algorithm will sporadically reach a false local optimum and reduce precision.

The robustness of the method against initialization error (e.g., operator variability) and local optima arising from anatomical deformation was evaluated with 200 random registration trials on each projection-CT pair (200 × 72 pairs = 14 400 registrations) with a variable number of multi-starts (N = 1 to 250). The range of random perturbation to the initialization was chosen to emulate realistic operator variability in C-arm positioning for AP fluoroscopy similar to Otake *et al* (2012c). We considered two scenarios for the use of 3D–2D registration in clinical workflow: (1) when registration is performed for the first time in a case, a large search space around the initialization is allowed, referred to as the `large search' scenario; and (2) in subsequent registration, the search space is reduced, since the previous result can be used as an initialization, referred to as the `small search' scenario. Penney *et al* reported that patient rotation between CT and fluoroscopy position was 3°–4° on average (max 8°) based on 23 patient studies (Penney *et al* 2011). Hence, we allowed a rotational range with σ of 5° (where σ is the standard deviation in the normal distribution) for the large search and ±3° for the small search. The translational uncertainty allowed a σ = ±15 mm in right–left direction, ±50 mm out-of-plane, and ±100 mm in the superior–inferior direction

for the large search scenario, whereas each was set to ±15 mm for the small search. The distribution was clamped at ±2σ to avoid extrema. The two-level multi-resolution pyramid (DS8 and DS4) was used with the population size of 200, stopping criteria of 3.0 (mm or deg) for DS8, 0.1 (mm or deg) for DS4.

3. Results

3.1. Effect of the NGI weighting mask

Figure 5 shows the effect of the weighting mask on the optimization space with large deformation (scenario #8 in table 1) including a large air pocket moving in the abdomen. The edges of the deformed air-filled structure present a strong gradient that incorrectly aligns the DRR when the weighting mask is not applied (figure 5(b)), but as shown in figure 5(c), a weighting mask centered on the fluoroscopy image (i.e., assuming the gradients of interest—namely, the spine—to be in the center of the projection) reduces the contribution of the air pocket to NGI and yields correct registration of the vertebrae. Figure 5(d) shows the objective function landscape along one line in the 6-dimensional space connecting the two poses, where the intermediate pose was computed by linear interpolation, resulting in a one-dimensional pose surrogate $T(\alpha)$ defined as $T(\alpha) = T_{\text{false}} + \alpha(T_{\text{true}} - T_{\text{false}})$. The weighting mask is seen to improve the optimization landscape by elevating the global optimum. Dependency of the registration results on the lateral displacement of the weighting mask is shown in figure 5(e). TRE significantly increased when the mask included regions with large deformation (air-filled structure in the right side of the image), whereas stayed less than 5 mm for displacement to the left when the weighted structure was rigid.

Alternative functions (e.g., normalized pointwise mutual information (Rogelj *et al* 2003)) could be employed in the weighting mask, but a simple pointwise scalar was easily implemented in the parallelized computation, whereas a histogram-based method could be difficult to efficiently parallelize.

3.2. Analysis of precision

The objective function landscape around the true registration is illustrated in figure 6 for an example trial: fluoroscopic image acquired with the cadaver prone with kyphosis; thoracic region; magnification = 2.4; registered to prone CT. The registration was performed with hierarchical downsampling DS8, DS4, and DS2 with two in-plane translations (X and Z). Because the optimization landscape at DS8 (first inset in figure 6) exhibits numerous sharp local optima (spikes), the optimization does not follow the exact solution in repeated trials. The spikes are caused by discretization in 3D voxels in the volume and 2D pixels in the projection. For such large voxels, a small change in transformation parameter can yield large change in line integral and thus in NGI. The discontinuities are present even in the DS2 level at a very fine scale (final inset in figure 6) and present a potential source of imprecision for both the CMA-ES (which is intrinsically stochastic) and downhill simplex (which is otherwise deterministic).

The PDE and coordinate-wise error are shown as a function of the downsampling ratio in figure 7, and the dispersion about the true global optimum is plotted as a function of the

number of sample points (M in equation (5)). At DS8, median PDE was 0.101 mm for downhill simplex, compared to 0.050 mm for CMA-ES. The median out-of-plane translational error (Y-direction) was reduced with CMA-ES from 0.093 to 0.021 mm from DS8 to DS2. Moreover, the dispersion decreased and precision improved at successively finer levels of the multiresolution pyramid. Overall, CMA-ES achieved better precision than downhill simplex (despite its intrinsic stochasticity), suggesting a higher level of robustness against local optima and leading to more consistent convergence on the optimum. Even with a good initialization, downhill simplex showed lower precision than CMA-ES.

3.3. Analysis of robustness

Successful registration (i.e., PDE < 5 mm) was achieved in >99.99% of all trials for all scenarios of anatomical site and degree of deformation, but there is an important tradeoff between robustness and computation time as illustrated in figures 8 and 9. For cases of minor deformation (scenarios #1 and #2), a success rate >99.99% was achieved in 0.5 s for the small search (N = 1 multi-start), and in 6.8 s for the large search (N = 20). For cases of strong deformation (scenarios #1 to #8), the small search scenario required 6.3 s and the large search 54.0 s to achieve >99.99% success. For N = 250, the overall success rate, including all deformation scenarios, was 99.993% (one failure in 14 400 trials). The median PDE at initialization was 129.65 mm, and that at the solution was 0.025 mm, suggesting a highly consistent registration result. To consider a different clinical application scenario such as pedicle screw placement, the PDE was also computed using target points defined on each pedicle (Tomazevic *et al* 2004), showing initial median PDE of 129.47 mm and the solution of 0.026 mm. The average edge length of the local search space (i.e., leafs of the kD-tree) was 106.7 mm for N = 5, and reduced to 26.9 mm for N = 250. The results suggest that the possibility to tune the single parameter N (the number of multi-starts) to control the robustness–time tradeoff. For example, in clinical use, the operator could adjust robustness according to the surgical scenario to achieve >99.99% confidence in the result, using a longer search (large N) in the first registration performed in the case (poor initialization) or situations where a strong deformation or content mismatch is suspected, and using a shorter search in subsequent registrations (good initialization and/or minor deformation from prior runs).

The most challenging scenarios were those in the region of the thorax (presenting a larger number of false local optima in periodic structures of the rib cage) under large deformation (diaphragm, mediastinum and air pocket) at higher magnification (smaller FOV and reduced overall NGI).

Still such registration was solved to better than 5 mm PDE in >99.99% cases given a sufficiently long search (N = 250) and a weighting mask about the center of the projection. Although registration with N = 250 showed a high rate of success, 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 probability of failure. However, the analysis of precision in comparison to downhill simplex in the previous section suggests that errors due to stochasticity are less than those associated with the image data themselves (e.g.,

downsampling). Further experiments with an even larger number of trials are necessary to reveal the frequency of such stochastic failures.

Figure 10 shows the NGI at the true registration along with the median PDE and 95% confidence interval in PDE (all for the case $N = 250$). NGI was highest in scenarios of smaller deformation (deformation scenario #1–#4) and lower PDE, and it was reduced in cases of larger deformation (image index #7–#8 in deformation scenario #6–#8). The single failure case (PDE: 46.2 mm, NGI: 0.537) was associated with the lowest overall NGI, suggesting a possibility to be investigated more fully in future work that NGI could be used as a confidence metric or, at least, a measure of the difficulty of registration.

4. Discussion

This paper advanced a robust intensity-based 3D–2D image registration using NGI as a similarity metric and a global optimization algorithm consisting of multi-start search and restart strategy. The multi-start strategy partitioned the search space into multiple subspaces and performed multiple independent searches confined to the small space that improved each local search. The strategy broadened the capture range to allow a fairly rough estimate of patient positioning with respect to the C-arm (e.g., tolerance was ~ ±200 mm in the superior–inferior direction). A cadaver experiment demonstrated the robustness of the method against realistic changes in patient position (i.e., supine, prone) and anatomical deformations (i.e., kyphosis, lordosis) between preoperative CT and intraoperative x-ray images. Analysis of the objective function landscape revealed that the challenging (spiky) landscape induced by spatial discretization in 2D projection (pixelization) and 3D volume (voxelization) caused degradation in registration precision, and the optimization algorithm proved robust against such false local minima. The multi-start method was facilitated by an orders-of-magnitude improvement in the number of function evaluations performed via GPU acceleration. The implementation maintained high occupancy in thousands of processor cores, suggesting that performance will further improve in proportion to future increase in the number of available cores (e.g., an increase of processor cores on GPU or in computation nodes in CPU/GPU cluster).

The main challenge in 3D–2D registration is twofold: the similarity metric and the optimization approach. If the similarity metric does not demonstrate a global optimum at the true registration (e.g., due to anatomical deformation), then the optimization algorithm will not converge on the true registration. The experiments detailed in this work showed that NGI yielded the correct global optimum in 93% of the cases examined (even with gross deformation) owing to local rigidity of the spine, and the remaining 7% of cases were absolved by application of an NGI weighting mask placing greater weight on the center of the projection image (i.e., the area occupied by the spine) to down-weight structures such as the skinline or soft tissues not associated with the task of vertebral localization.

Correlation between NGI at the solution and the degree of deformation suggested the potential use of NGI as a metric indicating the difficulty of the registration. In clinical translation, an interface can be envisioned that displays the registered target points and various planning structures along with this metric to intuitively convey the difficulty of

registration to the operator (e.g., points displayed in green when the NGI is high, and yellow if the NGI is low). In cases where the NGI is low, the operator would be free to, for example, increase the number of multi-starts (N) to assure a confident registration.

Other research reported in the literature employs 3D–2D registration incorporating a non-rigid transformation. For example, the transformation can be parameterized using prior information such as the principal components and eigen-values of motion derived from a training dataset (Wang *et al* 2013) or properties of length preservation of vessels (Groher *et al* 2009). Similarly, finite element modeling could drive such a non-rigid transform (Hopp *et al* 2012). Alternatively, deformation vectors could be derived according to the centerline of vessels (Rivest-Henault *et al* 2012) or control points in a free-form deformation (Qi *et al* 2008). While the registration framework proposed in the work reported here can also be applied to such parameterizations by modifying the representation of the transformation and increasing the number of optimization variables, we focused in the current study on rigid transformation including only six parameters to yield a more stable optimization and reduce computational cost. As demonstrated in the experiments, the rigid transformation approach was still able to find the locally rigid structures (vertebrae) in both images and utilize consistent features to robustly solve the registration even in the presence of strong anatomical deformation and within a clinically realistic computation time.

The current analysis focused on registration in the AP projection view as common in minimally invasive spine surgery (either single plane or biplane). Future work will extend experimentation and analysis to the lateral projection view. Other preliminary simulation studies (Otake *et al* 2012b) demonstrated that the LevelCheck registration was accurate in the lateral projection view, but further investigation is required to confirm robustness against deformation. While the method is intrinsically robust to image content mismatch (e.g., surgical tools in the fluoroscopy image but not in the preoperative CT), future work will include experiments and analysis in such scenarios in anticipation of a typical clinical workflow where an x-ray opaque tool is included in the fluoroscopy image together with the patient to correlate the image to externally visible landmarks.

Parameters employed in the optimization could also be extended to include intrinsic parameters of the x-ray projection geometry, namely source-to-detector distance and image center. This extension would be essentially useful in images acquired with an unconstrained geometry, such as a mobile x-ray device with a wireless x-ray detector (Otake *et al* 2013). The optimization framework is being extended to accommodate the additional DoFs (i.e., 6 DOF transformation + 3 DOF intrinsic parameters) for registration in mobile radiographs. Finally, the system is being translated to use with C-arms featuring x-ray imaging intensifiers, potentially involving a distortion correction as described in Yan *et al* (2011) and extended to clinical pilot studies.

Acknowledgments

This research was supported by academic–industry partnership with Siemens Healthcare (XP Division Erlangen Germany). The authors thank Dr Mehran Armand (Johns Hopkins Applied Physics Lab) for assistance with computational resources, Mr Ronn Wade (University of Maryland Anatomy Board) for assistance with cadaver specimens, and Dr Elliot McVeigh (Department of Biomedical Engineering, Johns Hopkins University) as well as

Dr Jonathan Lewin and Ms Lori Pipitone (Department of Radiology, Johns Hopkins University) for support and research infrastructure.

References

Auger A, Hansen N. Performance evaluation of an advanced local search evolutionary algorithm. IEEE Congress on Evolutionary Computation. 2005a; 2:1777–84.

Auger A, Hansen N. A restart CMA evolution strategy with increasing population size. IEEE Congress on Evolutionary Computation. 2005b; 2:1769–76.

Bentley JL. Multidimensional binary search trees used for associative searching. Commun. ACM. 1975; 18:509–17.

Besl PJ, Mckay ND. A method for registration of 3-D shapes. IEEE Trans. Pattern Analysis and Machine Intelligence. 1992; 14:239–56.

Birkfellner W, Seemann R, Figl M, Hummel J, Ede C, Homolka P, Yang X, Niederer P, Bergmann H. Fast DRR generation for 2D/3D registration. Med. Image Comput. Comput. Assist. Interv. 2005; 8:960–7. [PubMed: 16686053]

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; 36:3420–8. [PubMed: 19746775]

Birkfellner W, Wirth J, Burgstaller W, Baumann B, Staedele H, Hammer B, Gellrich NC, Jacob AL, Regazzoni P, Messmer P. A faster method for 3D/2D medical image registration—a simulation study. Phys. Med. Biol. 2003; 48:2665–79. [PubMed: 12974581]

Cabral, B.; Cam, N.; Foran, J. Accelerated volume rendering and tomographic reconstruction using texture mapping hardware. Proceedings of the 1994 Symposium on Volume visualization; Tysons Corner, VA: ACM; 1994. p. 91-8.

De Silva T, Fenster A, Cool DW, Gardi L, Romagnoli C, Samarabandu J, Ward AD. 2D-3D rigid registration to compensate for prostate motion during 3D TRUS-guided biopsy. Med. Phys. 2013; 40:022904. [PubMed: 23387775]

Dennis DA, Mahfouz MR, Komistek RD, Hoff W. *In vivo* determination of normal and anterior cruciate ligament-deficient knee kinematics. J. Biomech. 2005; 38:241–53. [PubMed: 15598450]

Dey J, Napel S. Targeted 2D/3D registration using ray normalization and a hybrid optimizer. Med. Phys. 2006; 33:4730–8. [PubMed: 17278825]

Dorgham OM, Laycock SD, Fisher MH. GPU accelerated generation of digitally reconstructed radiographs for 2-D/3-D image registration. IEEE Trans. Biomed. Eng. 2012; 59:2594–603. [PubMed: 22801484]

Fisher M, Dorgham O, Laycock S. Fast reconstructed radiographs from octree-compressed volumetric data. Int. J. Comput. Assist. Radiol. Surg. 2013; 8:313–22. [PubMed: 22821505]

Florack LMJ, Haar Romeny BM, Koenderink JJ, Viergever MA. General intensity transformations and differential invariants. J. Math. Imaging Vis. 1994; 4:171–87.

Gendrin C, et al. Monitoring tumor motion by real time 2D/3D registration during radiotherapy. Radiother. Oncol. 2012; 102:274–80. [PubMed: 21885144]

Groher M, Zikic D, Navab N. Deformable 2D-3D registration of vascular structures in a one view scenario. IEEE Trans. Med. Imaging. 2009; 28:847–60. [PubMed: 19131296]

Hansen, N. The CMA evolution strategy: a comparing review. In: Lozano, J.; Larrañaga, P.; Inza, I.; Bengoetxea, E., editors. Towards a New Evolutionary Computation. Springer; Berlin: 2006. p. 75-102.

Hansen, N.; Kern, S. Evaluating the CMA evolution strategy on multimodal test functions. In: Yao, X.; Burke, E.; Lozano, J.; Smith, J.; Merelo-Guervós, J.; Bullinaria, J.; Rowe, J.; Tiňo, P.; Kabán, A.; Schwefel, H-P., editors. Parallel Problem Solving from Nature - PPSN VIII. Springer; Berlin: 2004. p. 282-91.

Hipwell JH, Penney GP, Mclaughlin RA, Rhode K, Summers P, Cox TC, Byrne JV, Noble JA, Hawkes DJ. Intensity-based 2-D-3-D registration of cerebral angiograms. IEEE Trans. Med. Imaging. 2003; 22:1417–26. [PubMed: 14606675]

Hopp T, Dietzel M, Baltzer PA, Kreisel P, Kaiser WA, Gemmeke H, Ruiter NV. Automatic multimodal 2D/3D breast image registration using biomechanical FEM models and intensity-based optimization. Med. Image Anal. 2012; 17:209–18. [PubMed: 23265802]

Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage. 2002; 17:825–41. [PubMed: 12377157]

Jerbi T, Burdin V, Leboucher J, Stindel E, Roux C. 2D-3D frequency registration using a low-dose radiographic system for knee motion estimation. IEEE Trans. Biomed. Eng. 2012; 60:813–20. [PubMed: 22361657]

Lunacek, M.; Whitley, D. The dispersion metric and the CMA evolution strategy. Proc. 8th Annu. Conf. on Genetic and Evolutionary Computation; Seattle, DC: ACM; 2006. p. 477-84.

Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. IEEE Trans. Med. Imaging. 1997; 16:187–98. [PubMed: 9101328]

Mahfouz MR, Hoff WA, Komistek RD, Dennis DA. A robust method for registration of three-dimensional knee implant models to two-dimensional fluoroscopy images. IEEE Trans. Med. Imaging. 2003; 22:1561–74. [PubMed: 14649746]

Markelj P, Tomazevic D, Likar B, Pernus F. A review of 3D/2D registration methods for image-guided interventions. Med. Image Anal. 2012; 16:642–61. [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. IEEE Trans. Med. Imaging. 2005; 24:1058–66. [PubMed: 16092337]

Mitrovic U, Spiclin Z, Likar B, Pernus F. 3D-2D registration of cerebral angiograms: a method and evaluation on clinical images. IEEE Trans. Med. Imaging. 2013; 32:1550–63. [PubMed: 23649179]

Mody MG, Nourbakhsh A, Stahl DL, Gibbs M, Alfawareh M, Garges KJ. The prevalence of wrong level surgery among spine surgeons. Spine (Phila Pa 1976). 2008; 33:194–8. [PubMed: 18197106]

Munbodh R, Tagare HD, Chen Z, Jaffray DA, Moseley DJ, Knisely JP, Duncan JS. 2D-3D registration for prostate radiation therapy based on a statistical model of transmission images. Med. Phys. 2009; 36:4555–68. [PubMed: 19928087]

Navab, N.; Bani-Hashemi, A.; Nadar, MS.; Wiesent, K.; Durlak, P.; Brunner, T.; Barth, K.; Graumann, R. 3D reconstruction from projection matrices in a C-arm based 3D-angiography system. In: Wells, W.; Colchester, A.; Delp, S., editors. MICCAI'98: Medical Image Computing and Computer-Assisted Interventation. Springer; Berlin: 1998. p. 119-29.

Nelder JA, Mead R. A simplex method for function minimization. Comput. J. 1965; 7:308–13.

Otake Y, Armand M, Armiger RS, Kutzer MD, Basafa E, Kazanzides P, Taylor RH. Intraoperative image-based multiview 2D/3D registration for image-guided orthopaedic surgery: incorporation of fiducial-based C-arm tracking and GPU-acceleration. IEEE Trans. Med. Imaging. 2012a; 31:948–62. [PubMed: 22113773]

Otake Y, Schafer S, Stayman JW, Zbijewski W, Kleinszig G, Graumann R, Khanna AJ, Siewerdsen JH. Automatic localization of target vertebrae in spine surgery using fast CT-to-fluoroscopy (3D-2D) image registration. Medical Imaging 2012: Image-Guided Procedures, Robotic Interventions, and Modeling. 2012b:8316.

Otake Y, Schafer S, Stayman JW, Zbijewski W, Kleinszig G, Graumann R, Khanna AJ, Siewerdsen JH. Automatic localization of vertebral levels in x-ray fluoroscopy using 3D-2D registration: a tool to reduce wrong-site surgery. Phys. Med. Biol. 2012c; 57:5485–508. [PubMed: 22864366]

Otake Y, Wang AS, Stayman JW, Kleinszig G, Vogt S, Khanna AJ, Wolinsky JP, Gokaslan ZL, Siewerdsen JH. Verification of surgical product and detection of retained foreign bodies using 3D-2D registration in intraoperative mobile radiographs. Proc. Computer Assisted Radiology and Surgery (June). 2013:S185–6.

Palumbo MA, Bianco AJ, Esmende S, Daniels AH. Wrong-site spine surgery. J. Am. Acad. Orthop. Surg. 2013; 21:312–20. [PubMed: 23637150]

Penney, G.; Varnavas, A.; Dastur, N.; Carrell, T. An image-guided surgery system to aid endovascular treatment of complex aortic aneurysms: description and initial clinical experience. In: Taylor, R.;

Yang, G-Z., editors. Information Processing in Computer-Assisted Interventions. Springer; Berlin: 2011. p. 13-24.

Penney GP, Weese J, Little JA, Desmedt P, Hill DL, Hawkes DJ. A comparison of similarity measures for use in 2-D-3-D medical image registration. IEEE Trans. Med. Imaging. 1998; 17:586–95. [PubMed: 9845314]

Pluim JP, Maintz JB, Viergever MA. Image registration by maximization of combined mutual information and gradient information. IEEE Trans. Med. Imaging. 2000; 19:809–14. [PubMed: 11055805]

Powell MJD. An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput. J. 1964; 7:155–62.

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: 2008. p. 339-48.

Rinnooy Kan AHG, Timmer GT. Stochastic global optimization methods part I: clustering methods. Math. Program. 1987; 39:27–56.

Rivest-Henault D, Sundar H, Cheriet M. Nonrigid 2D/3D registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions. IEEE Trans. Med. Imaging. 2012; 31:1557–72. [PubMed: 22531755]

Rogelj P, Stanislav K, James CG. Point similarity measures for non-rigid registration of multi-modal data. Comput. Vis. Image Und. 2003; 92:112–40.

Russakoff, D.; Rohlfing, T.; Ho, A.; Kim, D.; Shahidi, R.; Adler, J., Jr; Maurer, C, Jr. Evaluation of Intensity-based 2D-3D spine image registration using clinical gold-standard data. In: Gee, J.; Maintz, JBA.; Vannier, M., editors. Biomedical Image Registration. Springer; Berlin: 2003. p. 151-60.

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. Imaging. 2005; 24:1441–54. [PubMed: 16279081]

Spoerk J, et al. High-performance GPU-based rendering for real-time, rigid 2D/3D-image registration and motion prediction in radiation oncology. Z. Med. Phys. 2012; 22:13–20. [PubMed: 21782399]

Steininger P, Neuner M, Weichenberger H, Sharp GC, Winey B, Kametriser G, Sedlmayer F, Deutschmann H. Auto-masked 2D/3D image registration and its validation with clinical cone-beam computed tomography. Phys. Med. Biol. 2012; 57:4277–92. [PubMed: 22705709]

Suganthan, PN.; Hansen, N.; Liang, JJ.; Deb, K.; Chen, Y.; Auger, A.; Tiwari, S. KanGAL Report. Indian Institute of Technology; Kanpur: 2005. Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization. 2005005

Tomazevic D, Likar B, Pernus F. `Gold standard' data for evaluation and comparison of 3D/2D registration methods. Comput. Aided Surg. 2004; 9:137–44. [PubMed: 16192053]

Tornai GJ, Cserey G, Pappas I. Fast DRR generation for 2D to 3D registration on GPUs. Med. Phys. 2012; 39:4795–9. [PubMed: 22894404]

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; 32:3737–49. [PubMed: 16475773]

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. Imaging. 2005; 24:1177–89. [PubMed: 16156355]

Varnavas A, Carrell T, Penney G. Increasing the automation of a 2D-3D registration system. IEEE Trans. Med. Imaging. 2013; 32:387–99. [PubMed: 23362246]

Wang H, Stout DB, Chatziioannou AF. A method of 2D/3D registration of a statistical mouse atlas with a planar x-ray projection and an optical photo. Med. Image Anal. 2013; 17:401–16. [PubMed: 23542374]

Xin C, Graham J, Hutchinson C, Muir L. Automatic inference and measurement of 3D carpal bone kinematics from single view fluoroscopic sequences. IEEE Trans. Med. Imaging. 2013; 32:317–28. [PubMed: 23193232]

Yan S, Nie S, Zheng B. Improving accuracy of XRII image distortion correction using a new hybrid image processing method: performance assessment. Med. Phys. 2011; 38:5921–32. [PubMed: 22047357]

Zhu Z, Massimini DF, Wang G, Warner JJ, Li G. The accuracy and repeatability of an automatic 2D-3D fluoroscopic image-model registration technique for determining shoulder joint kinematics. Med. Eng. Phys. 2012; 34:1303–9. [PubMed: 22285714]

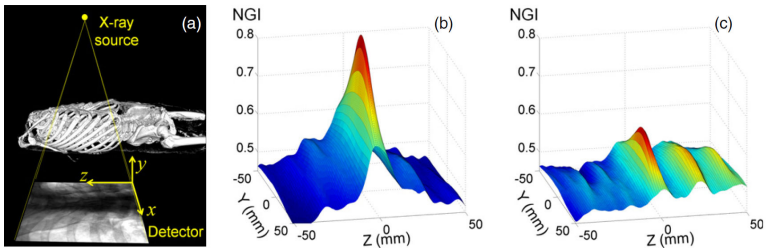

Figure 1.

Illustration of the 3D–2D registration objective function landscape. (a) Illustration of the projection geometry. (b) and (c) Objective function value plotted as a function of two variables (*Y*, out-of-plane, and *Z*, in-plane, translations; the complete registration involves six variables of translation and rotation) illustrating the challenge of anatomical deformation. In (b) without deformation, the objective function exhibits a clear global optimum. In (c) with deformation (spine lordosis), the optimization becomes poorly posed due to reduction in the peak value of the objective function and the emergence of multiple local optima.

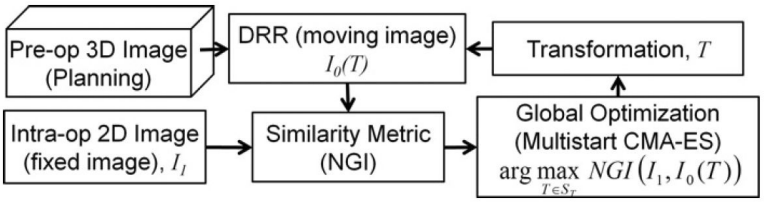

Figure 2.
Flowchart for the 3D–2D registration framework.

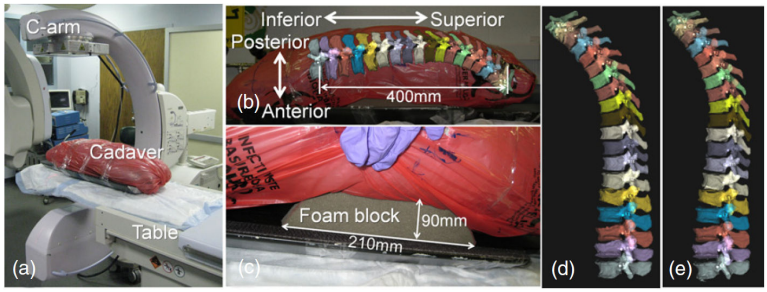

Figure 3.

Experimental setup. (a) Photograph of the C-arm and cadaver. (b) Cadaver torso shown in the prone position, with kyphosis or lordosis simulated using (c) a foam block placed below the belly or back, respectively. (d) Segmentation of vertebrae in prone CT and (e) supine CT, illustrating the degree of overall deformation between the two CT datasets.

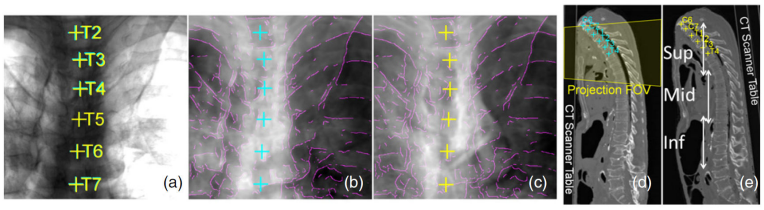

Figure 4.

An example registration for truth definition. (a) Real fluoroscopy (thoracic region) in prone position. (b) DRR of prone CT (vertebral labels in cyan). (c) DRR of supine CT (vertebral labels in yellow). Magenta overlays in (b) and (c) are the image gradients from (a), and the cyan and yellow vertebral labels are in turn overlaid in (a)—all suggesting accurate alignment at the `true' registration. (d) Sagittal slice of the prone CT. (e) Sagittal slice of the supine CT registered to (a). Note the rotation of the volume to align within the projection FOV. The three anatomical regions tested in the experiment are marked as thoracic (Sup), abdominal (Mid), and lumbar (Inf).

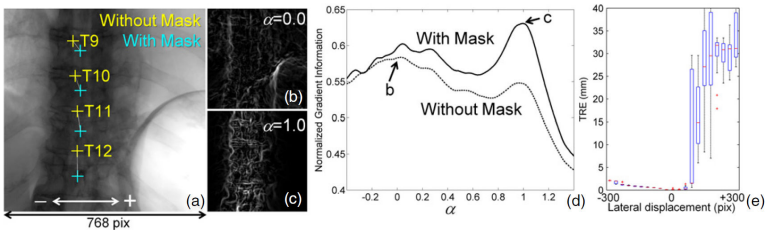

Figure 5.

Effect of a weighting mask. (a) Fluoroscopy image overlaid with the labels computed by the registration with a weighting mask (cyan) and without the weighting mask (yellow). NGI images are shown (b) without and (c) with a weighting mask. (d) Objective function landscape along a 1D line connecting the two poses. (e) Dependency of registration result on the lateral displacement of the weighting mask in the case of a very strong gradient introduced by a large bubble and/or the skinline. Each box plot represents result of 20 random trials. The box marks first and third quartiles, and the whiskers mark the minimum and maximum values.

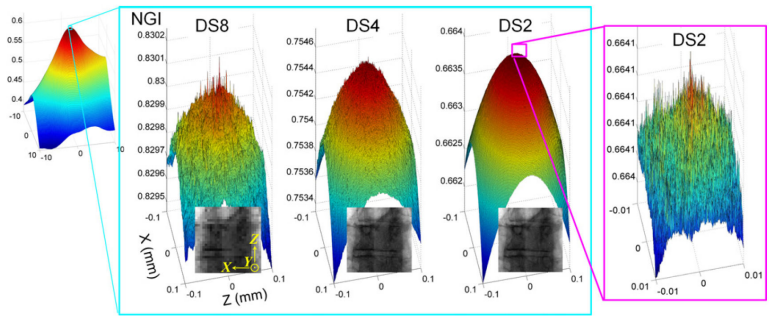

Figure 6.

Dependence of the objective function landscape on downsampling (DS) ratio: DS8 (3.1^2 mm^2 $pixel^{-1}$, 2.7^3 mm^3 $voxel^{-1}$); DS4 (1.6^2 mm^2 $pixel^{-1}$, 1.3^3 mm^3 $voxel^{-1}$); and DS2 (0.78^2 mm^2 $pixel^{-1}$, 0.68^3 mm^3 $voxel^{-1}$). Two in-plane translations (X and Z translations) in [−0.1, +0.1] mm around the true registration are plotted. The insets illustrate a zoomed-in region around one vertebra in each level of the multiresolution pyramid.

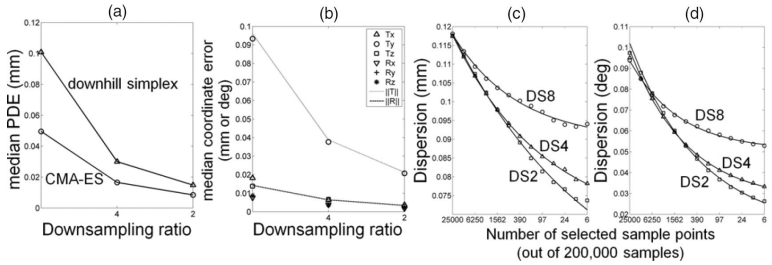

Figure 7.

Precision of 3D–2D registration. (a) Median PDE for CMA-ES and downhill simplex. (b) Median coordinate-wise error in each translation and rotation parameter in CMA-ES (T_y: out-of-plane translation, ||T|| and ||R|| indicate L2-norm of translation and rotation). (c), (d) Average dispersion computed as a function of the number of selected sample points, demonstrating a smoother objective function landscape in DS2 (decreasing dispersion) compared to DS8.

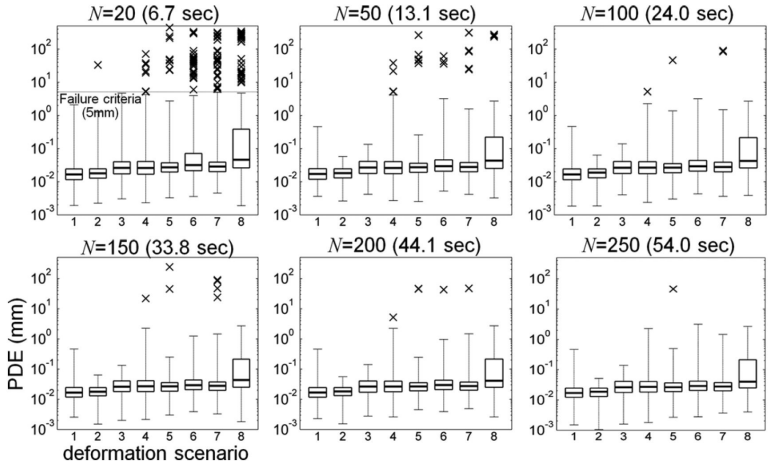

Figure 8.

Robustness of 3D–2D registration. The PDE for all registration trials and deformation scenarios is shown for various settings of N (number of multi-starts). The abscissa in each plot is the scenario # defined in table 1, with progressively stronger deformation for higher scenario #. Each box plot represents 1,800 random trials (9 images × 200 trials). The median PDE is marked by the bold line, the box marks first and third quartiles, and the whiskers mark the minimum and maximum values. Outliers (registrations defined as a failure due to PDE > 5 mm) are marked by the × symbols.

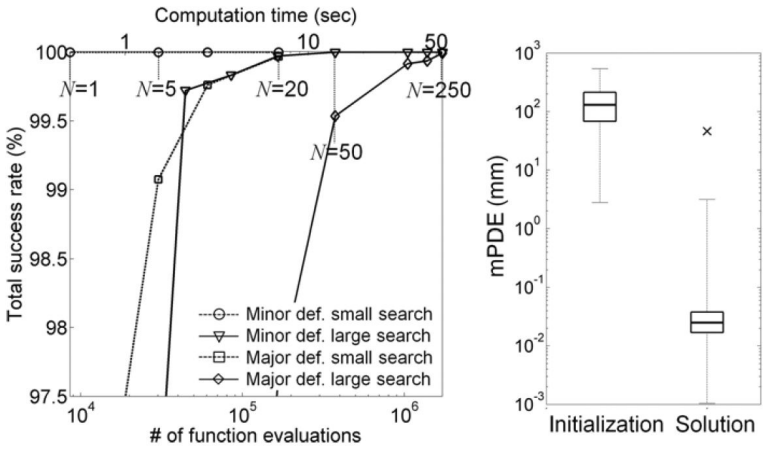

Figure 9.

(Left) The total success rate plotted as a function of computation time and number of function evaluations. (Right) PDE before and after registration with $N = 250$ (54.0 s) of 14 400 trials. Box plots mark the median (bold line), first and third quartiles (box), and range (whiskers).

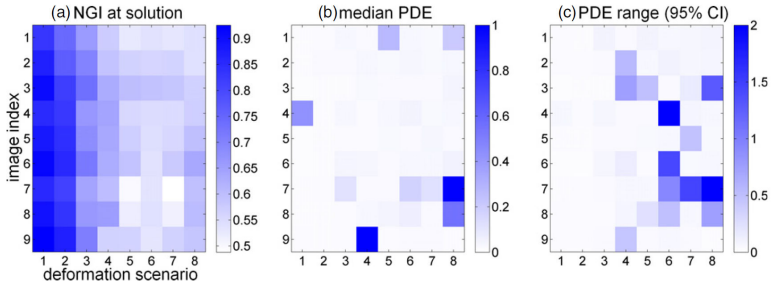

Figure 10.

For each deformation scenario (#1– #8) and image index (#1– #9) defined in table 1, plots of (a) NGI, (b) median PDE and (c) the 95% confidence interval of PDE resulting from 3D–2D registration ($N = 250$). The plots illustrates not only the registration accuracy in each anatomical region and deformation scenario (median PDE typically < 1 mm) but also the correlation between NGI and degree of deformation.

Table 1

Summary of datasets in the cadaver experiments. Notation is as follows: (↑ supine), (↓ prone), (↑∩ supine with lordosis), and (↓∩ prone with kyphosis).

Image characteristics		
Supine CT	512 × 512 × 1437 voxels	0.67 × 0.67 × 0.60 mm
Prone CT	512 × 512 × 1410 voxels	0.67 × 0.67 × 0.60 mm
Projections	768 × 768 pixels	0.388 × 0.388 mm

Target regions and magnifications

Image #	**Geometric magnification**	**Region**
1	2.4	Thoracic (C6-T7)
2	2.0	Thoracic (C6-T7)
3	1.7	Thoracic (C6-T7)
4	2.4	Abdominal (T8-L1)
5	2.0	Abdominal (T8-L1)
6	1.7	Abdominal (T8-L1)
7	2.4	Lumbar (L2-L5)
8	2.0	Lumbar (L2-L5)
9	1.7	Lumbar (L2-L5)

Deformation scenarios

Scenario #	**Position in CT**	**Position in fluoroscopy**	**Description**	
			Deformation	**Position between CT and fluoroscopy**
1	↑ (supine)	↑ (supine)	No	Same
2	↓ (prone)	↓ (prone)		Same
3	↑ (supine)	↑∩ (supine+lordosis)	Yes	Same
4	↓ (prone)	↓∩ (prone+kyphosis)		Same
5	↑ (supine)	↓ (prone)	No	Different
6	↓ (prone)	↑ (supine)		Different
7	↑ (supine)	↓∩ (prone+kyphosis)	Yes	Different
8	↓ (prone)	↑∩ (supine+lordosis)		Different