Evaluation of low-dose limits in 3D-2D rigid registration for surgical guidance

To cite this article: A Uneri et al 2014 Phys. Med. Biol. 59 5329

View the article online for updates and enhancements.

Related content
- 3D–2D registration for surgical guidance: effect of projection view angles on registration accuracy
  A Uneri, Y Otake, A S Wang et al.
- Known-component 3D–2D registration for quality assurance of spine surgery pedicle screw placement
  A Uneri, T De Silva, J W Stayman et al.
- Multi-stage 3D–2D registration for correction of anatomical deformation in image-guided spine surgery
  M D Ketcha, T De Silva, A Uneri et al.

Recent citations
- Intraoperative evaluation of device placement in spine surgery using known-component 3D–2D image registration
  A Uneri et al
- Sparse 3D Radon Space Rigid Registration of CT scans: Method and Validation Study
  Guy Medan et al
- 3D–2D registration in mobile radiographs: algorithm development and preliminary clinical evaluation
  Yoshito Otake et al
Evaluation of low-dose limits in 3D-2D rigid registration for surgical guidance

A Uneri1, A S Wang2, Y Otake3, G Kleinszig3, S Vogt3, A J Khanna4, G L Gallia5, Z L Gokaslan5 and J H Siewerdsen1,2

1 Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218, USA
2 Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205, USA
3 Siemens Healthcare XP Division, Erlangen, Germany
4 Department of Orthopaedic Surgery, Johns Hopkins Medical Institute, Baltimore, MD 21287, USA
5 Department of Neurological Surgery, Johns Hopkins Medical Institute, Baltimore, MD 21287, USA

E-mail: jeff.siewerdsen@jhu.edu

Received 10 April 2014, revised 4 July 2014
Accepted for publication 15 July 2014
Published 22 August 2014

Abstract

An algorithm for intensity-based 3D-2D registration of CT and C-arm fluoroscopy is evaluated for use in surgical guidance, specifically considering the low-dose limits of the fluoroscopic x-ray projections. The registration method is based on a framework using the covariance matrix adaptation evolution strategy (CMA-ES) to identify the 3D patient pose that maximizes the gradient information similarity metric. Registration performance was evaluated in an anthropomorphic head phantom emulating intracranial neurosurgery, using target registration error (TRE) to characterize accuracy and robustness in terms of 95% confidence upper bound in comparison to that of an infrared surgical tracking system. Three clinical scenarios were considered: (1) single-view image + guidance, wherein a single x-ray projection is used for visualization and 3D-2D guidance; (2) dual-view image + guidance, wherein one projection is acquired for visualization, combined with a second (lower-dose) projection acquired at a different C-arm angle for 3D-2D guidance; and (3) dual-view guidance, wherein both projections are acquired at low dose for the purpose of 3D-2D guidance alone (not visualization). In each case, registration accuracy was evaluated as a function of the entrance surface dose associated with the projection view(s). Results indicate that images acquired at a dose as low as 4 μGy (approximately one-tenth the dose of a typical fluoroscopic frame) were sufficient to provide TRE comparable or superior
to that of conventional surgical tracking, allowing 3D-2D guidance at a level of dose that is at most 10% greater than conventional fluoroscopy (scenario #2) and potentially reducing the dose to approximately 20% of the level in a conventional fluoroscopically guided procedure (scenario #3).

Keywords: 3D-2D registration, image-guided surgery, surgical navigation, neurosurgery, dose, x-ray fluoroscopy

(Some figures may appear in colour only in the online journal)

1. Introduction

Surgical guidance has been central to neurosurgical applications starting with the first stereotactic procedures performed in the 1940s (Spiegel et al 1947). Since then, developments in image guidance and computer-assisted technologies have extended the efficacy of neurosurgery and the scope of pathologies that can be successfully addressed by surgical intervention. Applications range from intracranial tissue biopsy (Woodworth et al 2006) and tumor resection (Schulder et al 1998) to ventricular shunt (Kobayashi et al 2012) and electrode placement (Bjartmarz and Rehncrona 2007), establishing surgical guidance as the modern standard of care in neurosurgery. The last two decades have seen a shift from the use of clamps and fixation devices to frameless alternatives that track a sparse set of markers (Smith et al 1994), thereby allowing navigation within the preoperatively acquired image of the patient. Although such systems provide a visual context for guidance (c.f. ‘blind’ approach using stereotactic frames), they are limited in their ability to adapt to changes in target anatomy, and they present workflow bottlenecks in both the preoperative processes (e.g. additional scanning of the patient with fiducial markers) and intraoperative processes (e.g. additional custom equipment and manual fiducial registration).

Intraoperative imaging allows frameless stereotactic neurosurgery with guidance in relation to an up-to-date view of patient anatomy, presenting an opportunity to account for changes due to patient motion or surgical manipulation and providing a dense set of anatomical features (c.f. a handful of fiducial markers) that can be used by surgical guidance algorithms. Intraoperative imaging modalities in common use are predominantly 2D, such as video endoscopy and x-ray fluoroscopy, with the latter utilized frequently for many tasks in brain surgery (Lee et al 2000). Recognizing the value of 3D localization, recent developments in neurosurgery include use of cone-beam CT (CBCT) (Lee et al 2012), 3D ultrasound (Unsgaard et al 2006), and intraoperative MRI (Lewin et al 2000). Depending on the clinical application, any of these modalities (or a combination thereof) may be best suited to the requirements and constraints of a particular clinical task. For example, intraoperative 3D CT or MRI may allow one to assess the completeness of target resection, but requirements and constraints in time, patient access, workflow, and (possibly) radiation dose may limit applicability to real-time (or near-real-time) guidance (Rafferty et al 2005). An alternative approach employs 3D-2D registration to position the previously acquired 3D image (e.g. preoperative or intraoperative CT) within an intraoperatively acquired 2D image (e.g. a single fluoroscopic frame), the first providing a 3D context with additional prior information, such as surgical planning data and other registered, preoperative, multi-modality images, and the latter providing an up-to-date 2D view of the patient (and surgical instrumentation) (Wein et al 2005, Birkfellner et al 2007, Frühwald et al 2009, Markelj et al 2012, Berkels et al 2014). Although many of these methods aim to provide accurate 2D overlays (e.g. overlay of planning data on the 2D fluoroscopic image as in (Uenohara and Kanade 1995) or (Otake et al 2012), previous work showed that accurate 3D
localization is also feasible using a single projection view or two views separated by a small angular separation (≈15° for 3D localization better than 2 mm), thus raising the possibility for accurate 3D surgical guidance (Uneri et al 2013).

An important consideration in using x-ray imaging modalities (projection fluoroscopy, CT, and/or CBCT and surgical staff). Minimization of dose in x-ray projection imaging is an active area of study (Strotzer et al 2002, Suleiman et al 2013), where the impact on image quality as assessed by a physical metrics (noise, spatial resolution, etc.), observer models, and/or human observers is used to optimize system design and image acquisition protocols. The impact of reduced dose on the performance of image guidance, however, is less understood, due to dependencies on the particular algorithm, application, and anatomy. Previous research includes the effect of dose reduction (i.e. increased noise in 3D images) on 3D-3D image registration of CBCT images (Sykes et al 2005) and 3D-2D registration of CBCT and endoscopic video (Mirota et al 2013).

In this work, we evaluate the effect of radiation dose reduction on the performance of 3D-2D registration of a preoperative CT (acquired at nominal dose for a high-quality preoperative scan protocol) and intraoperative 2D projection images acquired via C-arm fluoroscopy acquired at dose that is reduced in comparison to typical fluoroscopic exposures. The study draws from previous work that established the underlying 3D-2D registration algorithm (applied to identification of spine levels on 2D images in a single fluoroscopic frame by (Otake et al 2012)), later extended to 3D guidance using an optional secondary view to improve depth resolution (Uneri et al 2013). 3D-2D registration of CT and a single projection view (acquired at a nominal fluoroscopic dose level) demonstrated ≈3 mm accuracy in 3D target registration error (TRE), while a secondary view improved TRE to <2 mm. The study reported below specifically addresses the question of dose reduction in the 2D projection view and considers three potential implementations: (1) single-view image + guidance, wherein a single projection is used for visualization and 3D-2D guidance; (2) dual-view image + guidance, where one projection is acquired for visualization and combined with a second (lower-dose) projection acquired at a different C-arm angle for 3D-2D guidance; and (3) dual-view guidance, which uses two low-dose projections for the purpose of 3D-2D guidance alone (not visualization). Experiments emulating intracranial neurosurgery were performed in an anthropomorphic head phantom, with registration accuracy measured as a function of the dose in (single- or dual-view) x-ray projections, related to the dose of conventional fluoroscopic imaging, and compared to the accuracy obtained with an infrared surgical tracking system.

2. Methods

2.1. 3D-2D image registration algorithm

The fundamental component of 3D-2D guidance is the registration algorithm that solves for the rigid transformation of a 3D image by computing the transform that yields the highest similarity between a simulated 2D projection (digitally reconstructed radiograph, DRR) and the acquired 2D image. For the experiments described below, 2D images were acquired via C-arm flat-panel x-ray fluoroscopy (single-shot projection images), and a preoperative CT image was used as the 3D reference. The basic algorithm was described in detail by (Otake et al 2012) in application to labeling surgical targets on 2D images (viz. the ‘LevelCheck’ algorithm for labeling vertebral levels). The transform is that which maximizes the gradient information (GI) between the fluoroscopic view and (transformed) CT image:

$$T^{CT}_{Fluoro} = \arg \max_T \text{GI}(F, M_0(T)), \quad M_0(T) = B_0 T \cdot I_{CT},$$

(1)
where the iterative estimation of the 3D transform $T_{\text{Fluoro}}^\text{CT}$ in (1) used the covariance matrix adaptation evolution strategy (CMA-ES) as described in (Hansen et al 2009) to maximize the similarity between the log-normalized acquired projection (fixed, $F$) and the DRR (moving, $M$). A particular strength of CMA-ES motivating its use is robustness to local minima – for example, about the fairly shallow minimum in the depth direction when only a single projection view is used to resolve 3D pose, or when increased noise at low-dose conceals certain differentiating features or gradients. The DRRs are computed from the CT image ($I_{\text{CT}}$) using the 3D transform $T$ that relates the preoperative image to the C-arm coordinate system, such that multiple views may be computed using the pre-calibrated projective transforms $P_\theta$ defined by encoded angulations $\theta$ of the C-arm. The GI is defined in relation to the gradients of the fixed and moving images:

$$GI(F, M) = \sum_\theta \sum_{i,j} w_{ij} \min \left( \frac{|V F_{\theta,i,j}|}{|V M_{\theta,i,j}|}, \frac{|V M_{\theta,i,j}|}{|V F_{\theta,i,j}|} \right),$$

(2)

where the weights, $w_{ij}$, favor alignment of gradients as originally described by (Pluim et al 2000). The metric is computed over all image pixels $i$, $j$, and (optionally) over multiple images acquired at different $\theta$. Further weighting of the gradient intensities using a mask image ($M$) defined on the fixed image was described in (Otake et al 2013c) and shown to be of use in rejecting inconsistencies due to soft tissue deformation – e.g. to mask out unreliable gradients such as the skinline. The use of such gradient masks in intracranial neurosurgery could serve a similar purpose in mitigating the effects of brain shift and/or neck flexion on the registration. A brief assessment of masking functions in the context of images of the head and neck is provided in the Appendix. Finally, the nominal parameters of the algorithm were as defined in (Uneri et al 2013), such as the optimizer population ($\lambda = 50$) and step size ($\sigma = 5\, \text{mm}, 5^\circ$).

2.2. Physical experiments in a neurosurgery model

As illustrated in figure 1, an anthropomorphic head phantom (The Phantom Laboratory, Greenwich, NY) was constructed and filled with a gelatin mixture simulating brain ($50 \pm 10 \, \text{HU}$), and inserts shaped from wax to simulate ventricles ($0 \pm 10 \, \text{HU}$). Additionally, the phantom was implanted with 32 spherical Teflon targets (3.2 mm diameter), homogeneously distributed within the intracranial space for evaluation of TRE. A collection of larger ($5$–$12 \, \text{mm}$ diameter) low-contrast spheres (e.g. polyethylene) were also included in the intracranial space to simulate hemorrhage or soft-tissue targets but were not directly used in the experiments reported below.

A preoperative CT scan was acquired with a high-quality head imaging protocol (120kVp, 500mAs, reconstructed with $0.5 \times 0.5 \times 0.6 \, \text{mm}^3$ voxel size using the H30s reconstruction kernel, SOMATOM Definition Flash, Siemens Healthcare). The effects of CT voxel size on 3D-2D registration accuracy were evaluated in previous research (Uneri et al 2014). Single-shot fluoroscopic projections were acquired using a mobile C-arm incorporating a flat-panel detector (PaxScan 3030+, Varian, Palo Alto CA) operated in dual-gain mode (Schmidgunst et al 2007) and a motorized orbit (Siewerdsen et al 2005) for acquisition of projections at different view angles ($\theta$). The $\theta$ values were read from the gantry angle encoders built into the C-arm, though they can also be estimated using an external tracking system. These acquisitions were repeated for all dose levels identified in the following section. Each projection
Figure 1. Experimental setup. (a) Photograph of the C-arm, tracking system, phantom, ionization chamber, and interface for 3D-2D registration and navigation. X-ray projections are displayed on the navigation interface, along with the registered CT for comparison of registration accuracy. The ion chamber was used for dose measurements. (b) Photograph of the head phantom during construction, showing wax ventricles, target spheres, and low-contrast spheres placed within brain-simulating gelatin. (c) CT volume rendering of the head phantom.

was 768 × 768 pixels with 0.388 mm pixel pitch. The magnification was fixed at 2 (patient at the C-arm isocenter) as it was previously shown to be optimal in spine anatomy (Uneri et al 2014), and because the dependence of TRE on magnification was not particularly steep, we do not expect strong deviations from the observed trends at alternative selections of magnification. C-arm projection geometry was calibrated using a spiral BB phantom (Navab 1996), resulting in projective transforms $P_\beta$, with high reproducibility previously demonstrated for the C-arm used in this work (Daly et al 2008).

For the experiments detailed below, projections were acquired in a continuous orbit spanning 178° to yield 200 images at roughly equal increments of 0.9°. One (‘single-view’) or two (‘dual-view’) projections were selected from the 200 images for 3D-2D registration at any particular view angle. For the results reported below, all projections (i.e. 200 × number of dose levels) were used in evaluating the single-view scenario. CBCT volumes were reconstructed from the same projections using the Feldkamp algorithm (Feldkamp et al 1984), at 512³ voxels and 0.3 mm isotropic voxel size, solely for the purpose of reliable definition of true target locations.

Registration performance was compared to that obtained with a conventional infrared surgical tracking system (Polaris Spectra, NDI, Waterloo ON), operating at 60 Hz frame rate, and with an expected localization error of 2 mm contingent on marker visibility within the tracker field of view. Twelve fiducial divots were placed on the surface of the head phantom, distributed evenly across anterior and posterior regions, and their locations were repeatedly captured
(10 times) with a pointer tool, using a reference marker to account for specimen motion, and median filtered to reject potential outliers. The tracker was registered to CBCT images using Horn’s method of rigid registration (Horn 1987). To assess the accuracy of the tracking system, the pointer tool was inserted (30 times) into different regions within the intracranial space, the tooltip position was identified in CBCT as ground truth, and the TRE was computed as the difference between the true tooltip position and that measured by the tracker.

2.3. Dose measurements

The dose per frame in conventional fluoroscopy is typically 6–28 μGy (low-dose) or 56–111 μGy (higher-dose) depending on the selection of kVp, mA, and frame rate (Mahesh 2001). For studies below, the nominal dose of conventional fluoroscopy was taken to be ~30 μGy/frame, in fairly close agreement to the measured value of 34 μGy/frame for the C-arm imaging system in figure 1 and nominal protocol of 80 kVp and 1.6 mA. Entrance skin dose (ESD) was measured with a 60 cm³ ‘pancake’ ion chamber (Radcal 10 × 6–60, Radcal Corporation, Monrovia CA). All measurements were performed with the head phantom at C-arm isocenter, giving source-to-skin distance of ~52 cm (with source-to-isocenter distance of 60 cm and source-to-image distance of 120 cm). To vary the radiation dose, x-ray tube current was adjusted from 0.1 to 3.3 mA at doubling intervals, resulting in measured dose of 3–62 μGy/frame. To further reduce the dose beyond the minimum system tube current, an 8 mm aluminum (Al) filter was added to the source, giving measured ESD of 1–27 μGy/frame across the same range of mA. Al was chosen to filter non-penetrating low-energy photons (thus reducing ESD), while minimally affecting the shape of the x-ray spectrum, which in turn could affect registration in uncontrolled aspects. The anthropomorphic head phantom was placed prone on the operating table with the ion chamber placed on the lateral or anterior (beam entrance-side) surface such that the incident beam was normal to the chamber. The dose per fluoroscopic frame was measured in three C-arm positions (left-lateral, right-lateral, and AP) from the accumulated dose of 40 fluoroscopic frames. Due to the shielding effect of the carbon fiber table in the AP position (~5% reduction in dose), the average lateral dose, which was unobstructed by the table is reported in all results below to absolve the uncontrolled effects of table attenuation.

The image noise was measured in each projection as the standard deviation in pixel values (normalized by the mean pixel value) in a region of unattenuated beam. Relative noise was computed via normalization by the noise measured from the highest-dose scan. Example AP projections are shown in figure 2 at varying levels of ESD.

2.4. 3D-2D guidance scenarios

As detailed below, three scenarios of 3D-2D guidance are considered: (1) single-view image + guidance; (2) dual-view image + guidance; and (3) dual-view guidance. In these three scenarios, the moniker ‘image’ implies acquisition of a projection at a conventional dose level that is suitable for qualitative visualization by the surgeon, and the moniker ‘guidance’ refers to a projection that is acquired (at potentially reduced dose) for the purpose of 3D-2D registration (but not necessarily visualization). In the first scenario, 3D guidance is provided from a single projection view (i.e. a projection as acquired normally in single-shot fluoroscopic visualization). This scenario involves the conventional fluoroscopy-guided clinical workflow and uses only a single projection for 3D localization, shown previously to provide 3D localization by virtue of magnification effects for extended anatomy, albeit not as accurately or precisely as using at multiple views, particularly in the depth direction (Uneri et al 2013). An important
point of motivation in this scenario is compatibility with the natural workflow of a conventional fluoroscopic procedure, while adding the capability for quantitative 3D localization and guidance. In this scenario, 3D-2D guidance is intended to operate at dose levels comparable to or less than conventional fluoroscopically guided procedures. Assuming a standard procedure to require $N$ shots resulting in $ND$ total dose, an estimate for the dose reduction in scenario #1 would be, $nD + (N - n)(D - d)$, where $n$ is the number of biplane views acquired that would be obviated by the capability for 3D guidance, and $d < D$ is the minimum dose required for registration. For example, in a hypothetical estimate whereby the dose per view is unchanged, $d = D$, and 3D guidance reduces the number of biplane views by 2, $n = N/2$, then the total dose is reduced to a level 0.5$ND$ (i.e. dose reduced by half).

For the second and third (dual-view) scenarios, a secondary view is acquired at a given angular separation ($\Delta \theta$) from the first to improve depth resolution. A biplane system is not required for this – and is, in fact, unnecessary, considering the key finding of previous work that showed small angular separation ($\Delta \theta \sim 15^\circ$) is sufficient in providing the same level of 3D registration accuracy as a biplane views (Uneri et al 2013). For the second scenario (dual-view image + guidance), the first image is acquired at nominal dose (and is presumably sufficient for visualization by the surgeon), and the auxiliary view is acquired at a lower dose, hypothesizing that the 3D-2D registration algorithm is sufficiently robust against quantum noise to maintain registration accuracy. The motivation in this scenario is to improve the accuracy and robustness of 3D-2D registration (particularly in cases of strong anatomical deformation and/or small field of view) using a low-dose auxiliary view acquired at a small angle from the primary view. In this scenario, 3D-2D guidance incurs an increment in total dose for the procedure associated with that of the auxiliary view, although total dose may yet be lowered if $nD - (N - n)d$ is positive. For example, if the auxiliary view is acquired at a dose $d = D/5$, and 3D guidance reduces the number of views required by 2, then the total dose would be reduced to a level 0.6$ND$.

For the third scenario (dual-view guidance), a workflow is envisioned in which both views are acquired solely as a basis for 3D-2D registration (and not necessarily for visual interpretation by the surgeon), thus providing accurate guidance (c.f. single-view) at the lowest possible dose. Following the hypothesis that the registration algorithm is robust to elevated levels of...
quantum noise incurred at low dose, this scenario involves both views acquired at a dose that is reduced in comparison to the nominal dose of a single fluoroscopy frame. Estimated dose reduction would then be, \( nD + (N - n)(D - 2d) \), where it is expected that \( d \ll D \), thus operating at a total dose less than that of conventional fluoroscopy. For example, if dose per view is \( d = D/5 \), but the number of views required is half that of a conventional fluoroscopic procedure, then the total dose is reduced to a level 0.2 ND (i.e. one-fifth the dose of a conventional fluoroscopic procedure).

2.5. Initial registration and evaluation of capture range

The 3D-2D guidance scenarios delineated above involve ‘step-and-shoot’ acquisition of x-ray images at the surgeon’s discretion. This allows initialization of the current registration from previous results. For the experiments below, an initial TRE of \( \approx 5–10 \) mm was used, consistent with the TRE demonstrated in (Uneri et al 2013) and emulating the maximum expected motion between repeat acquisitions. To prevent bias associated with a fixed starting point, initial registration was computed from the iterative closest point match of targets implanted in the head, randomly perturbed by \( 3n \pm 1n \) mm in translation and \( 0.75n \pm 0.25n^\circ \) in rotation, where the ‘perturbation factor’ \( n = 1 \). To assess the sensitivity of the algorithm to initialization and quantify its capture range, an analysis was also performed where the perturbation factor was varied from \( n = 1 \) to \( n = 9 \).

Aside from the expected motion between acquisitions, the initial pose estimate within \( \approx 5–10 \) mm may be achieved by a global optimization that is solved once at the beginning of a case or at major milestones during the case in which a longer registration runtime is permissible for good initialization. For example, (Otake et al 2013a) showed that larger optimization population size and search space, coupled with a multi-start strategy is able to solve complex registration with TRE better than 5–10 mm at the cost of increased computation time. In addition, using more than 1–2 projections would further ensure accurate 3D initialization. Other means of initialization exist, such as the method by (van der Bom et al 2010) which uses the projection-slice theorem in combination with phase correlation to estimate initial transform parameters. Alternatively, an optional initial 3D scan (e.g. CBCT) could be acquired and registered to the CT, which would serve as initialization for 3D-2D registration.

2.6. Characterization of registration accuracy

3D registration accuracy was assessed in terms of TRE as defined by (Fitzpatrick et al 1998). The 32 spherical targets implanted within the head phantom were identified by manual segmentation in CT and CBCT, followed by a 3D radial Hough transform (Mosaliganti et al 2009) to quantitatively define the center of each sphere as \( s_{CT} \) and \( s_{CBCT} \), respectively. To prevent potential bias from the (small) gradients associated with the target spheres in the registration process, each sphere was digitally masked in the CT image using an inpainting algorithm adapted from (Garcia 2010). The 3D TRE was then given by:

\[
TRE_{3D2D} = \left\| T_{Fluoro}^{CT} \cdot T_{Fluoro}^{CT} \cdot s_{CT} - s_{CBCT} \right\|_2,
\]

where \( T_{Fluoro}^{CT} \) is the solution to 3D-2D registration, and \( T_{Fluoro}^{Fluoro} \) is identity, since the CBCT image was reconstructed from the same views as the fluoroscopy projections and is therefore in the same coordinate frame. For the results presented below, the upper bound of the 95% confidence interval in TRE (TRE_{95%}) was reported as an estimate for algorithm robustness and
gives a more conservative metric of overall performance than, for example, the median TRE. Similarly, the 3D TRE for the tracking system was given by:

$$T_{\text{Tracker}} = \left\| T_{\text{Tracker}} (\Theta) - T_{\text{CBCT}} \right\|_2,$$

where a total of 30 recorded tooltip locations $t_{\text{Tracker}}$ were compared to locations defined in CBCT, $T_{\text{CBCT}}$, using the tracker-to-CBCT registration, $T_{\text{Tracker}}$.

3. Results

Previous studies in (Uneri et al 2013) identified the angular separation ($\Delta \theta$) between projection views as a critical dependence on 3D registration accuracy; hence, the results below first considered the effects of dose reduction in relation to $\Delta \theta$. Taking nominal dose (80 kVp, 1.6 mA, 34 $\mu$Gy/frame) as the baseline performance, the three scenarios were evaluated with the aim of identifying the lowest dose level that maintained comparable registration performance.

A summary of results is presented in figure 3, showing consistent trends as in previous experiments performed in the context of pelvic, abdominal, and thoracic anatomy (here, in the context of intracranial procedures). For the results below, the initial TRE resulting from the random perturbation described above was mean TRE 5.3 mm and TRE95% 7.5 mm. The final TRE overall was observed to be lower than in previous work, potentially due to reduced deformation. The central results are consistent: TRE comparable or superior to that of a surgical tracker can be acquired with angular separation of $\Delta \theta \sim 15^\circ$ in each of the dual-view scenarios and, as investigated in greater detail below, at dose less than that of a single fluoroscopic frame.

3.1. Scenario 1: single-view image + guidance

As evident in the diagonal of figure 3(b)–(d), a single fluoroscopic projection is sufficient to provide 3D registration accuracy at a level of TRE $\sim 2.5$ mm. As shown in figure 4, two trends are revealed in analysis of the TRE as a function of reduced dose in the single-view fluoroscopy.
scenario. Over nearly an order of magnitude range about the nominal dose level (~34 μGy), the dependence of TRE on dose is very weak, showing a gradual logarithmic degradation in TRE at reduced dose despite a strong increase in quantum noise (quadratic). Some adjacent dose measurements (e.g. at 8 and 9 μGy) depict an opposite trend in which TRE reduces slightly with decreased dose; however these are attributable to spectral effects associated with cases in which the dose was reduced by filtering the beam with added Al filtration as described in section 2.3. At the lowest dose levels considered (<4 μGy), although the median TRE is still fairly low, a strong reduction in robustness is observed as an increased number of failures (outliers). At this seemingly minimum tolerable dose level, the projection image is degraded not only by an increased level of quantum noise but also a stronger appearance of electronic noise and imperfectly corrected gain-offset corrections, evident as subtle horizontal and vertical line defects in the image that present gradients that can confound the GI similarity metric. The minimum dose value (~4 μGy) is likely system-specific and may be further reduced with higher performance flat-panel detector readout electronics and gain-offset correction methods. Post-processing methods (e.g. an edge-preserving noise reduction filter applied to the low-dose projections) could also improve accuracy and robustness in the low-dose regime, as discussed below. Benefiting from the complex, rigid anatomy and strong gradients of the cranium, the single-view scenario provides TRE<sub>95%</sub> < 2 mm (better than the infrared tracking system) at nearly an order of magnitude lower dose than a single fluoroscopic frame.

3.2. Scenario 2: dual-view image + guidance

Figure 5 shows the registration accuracy for the 3D-2D guidance scenario (#2) in which one projection is acquired at nominal dose (34 μGy, suitable for visualization at a nominal level
of image quality) and an auxiliary projection is acquired at a lower dose. This dual-view scenario demonstrates improved accuracy compared to the single-view scenario of figure 4. Evident in figure 5 are several trends that demonstrate a complex dependence of registration accuracy and robustness on the angle and dose of the auxiliary view ($\Delta \theta$ and ESD, respectively). First, at extremely low dose (ESD $\leq$ 2 $\mu$Gy), the auxiliary view was found to degrade robustness (for $\Delta \theta \leq 10^\circ$) or give no added benefit (for $\Delta \theta \sim 10–20^\circ$) compared to the single-view scenario (#1). For sufficiently wide angular separation ($\Delta \theta \geq 30^\circ$), however, the auxiliary view acquired at such low dose improved registration accuracy. At higher dose (ESD $>$ 2 $\mu$Gy), the auxiliary view improved accuracy and robustness (compared to the single-view scenario) for all degrees of angular separation ($\Delta \theta$). For example, an auxiliary view acquired at $\Delta \theta 10^\circ$ and ESD 4 $\mu$Gy provided mean TRE 1.1 mm, TRE$_{95\%}$ 1.7 mm, and was robust across all measurements (no outliers).

3.3. Scenario 3: dual-view guidance

The final scenario (#3) involved dual-view registration in which both projections are acquired at dose less than that of a nominal fluoroscopic frame. The images involve elevated levels of quantum noise and may therefore be unsuitable for qualitative visualization; even so, we hypothesized the registration algorithm to be robust against such noise to the extent that this dual-view scenario represents a ‘guidance-only’ mode in which the images are acquired solely for purpose of registration and guidance (and not necessarily for visualization).

As shown in figure 6, the dual-view guidance scenario exhibits TRE that is lower to that of single-view registration (figure 4) but somewhat higher to dual-view image + guidance (figure 5). With the dose per projection reduced to less than $\sim 17 \mu$Gy, the total dose for dual-view guidance is less than a single fluoroscopic frame ($\sim 34 \mu$Gy), while providing 3D registration accuracy with TRE $< 2$ mm in each case. For example, with $\Delta \theta = 20^\circ$ and a dose per view of 2 $\mu$Gy (4 $\mu$Gy total dose), the system yielded registration accuracy of TRE$_{95\%}$ $< 1.5$ mm.
3.4. Evaluation of capture range

The sensitivity of the algorithm to a given initial registration was assessed using the nominal parameters as identified above for scenario #2. Specifically, dual views were used, a lateral primary acquired at $34.07\,\mu\text{Gy}$, and an auxiliary view separated by $\Delta \theta \sim 20^\circ$ and acquired at $3.97\,\mu\text{Gy}$ ($\sim 80\%$ reduced dose). The optimization parameters were kept fixed at reported values, and the initialization perturbation of $3n \pm 1n \,\text{mm}$, $0.75n \pm 0.25n^\circ$ was varied from the reported value of $n = 1$, up to $n = 9$.

Figure 7 shows that for the given parameters, the algorithm is robust against initial registration errors of up to $\sim 30 \,\text{mm} \,\text{TRE}$, after which it begins to present outliers. Initial TRE also shows a direct correlation with the number of iterations required for convergence, since the starting point is further from the solution. As noted in section 2.1, both robustness and convergence speed may be improved for these extreme cases through case-specific modifications to the optimizer parameters, and at the expense of computational resources.

4. Discussion

Intraoperative 2D x-ray projections acquired in a single- or dual-view manner were found to provide accurate 3D registration (TRE < 2 mm) at dose levels down to nearly an order of magnitude below that of a single fluoroscopy frame. Potential scenarios for 3D-2D guidance were presented that leverage conventional fluoroscopy (single-view scenario #1, with dose equal to that of conventional fluoroscopic imaging) or extend to include an extra view for more accurate 3D localization (dual-view, with total dose incremented by +10\% (scenario #2) in comparison to conventional fluoroscopy or reduced to as little as 20\% of the dose of conventional fluoroscopy (scenario #3)). The dual-view scenarios could be achieved by an automated rocking motion of the C-arm or through an auxiliary low-power source placed at a slight angle from the primary source. The results include the
configuration $\Delta \theta = 90^\circ$ corresponding to biplane systems, which can in principle support both dual-view guidance scenarios. For $\text{ESD} > 4 \mu\text{Gy}$ (compared to $\sim 34 \mu\text{Gy}$ typical of a single fluoroscopic frame), all scenarios demonstrated mean TRE within 3 mm (better than the infrared surgical tracking system) and a high degree of robustness with minimal or no outliers.

The algorithm was implemented on GPUs using CUDA architecture (nVidia Corporation, Santa Clara CA), since both the DRR and similarity metric calculation are amenable to parallelization. Runtimes of $\sim 1–5$ s were achieved and may be sufficient for a near-real-time, ‘step-and-shoot’ interface in which the surgeon identifies a point in the anatomy, performs a (single- or dual-view) image acquisition, and then can visualize 3D planning data registered to the 2D radiographic scene. This discrete interaction with the navigation system may fit well with actual surgical workflow in that identification of anatomical landmarks can be accomplished without the real-time aspect of conventional trackers. Investigation of step-and-shoot navigation workflow is the subject of future work.

The current study was limited to a rigid phantom, which is considered a fair starting point for neurosurgery applications considering the rigidity of the skull. However, further investigation should consider the effect of brain shift or deformation arising from neck flexion, combined with the possible implementation of a gradient weighting mask as described in the Appendix. In order to evaluate the baseline, most conservative performance characteristics of the 3D-2D registration technique, the current work did not employ additional post-processing of the projections, which may be anticipated to further improve the low-dose performance. For example, an edge-preserving noise reduction approach such as total variation denoising or anisotropic filtering could mitigate the (quadratic) increase in quantum noise and further shift the minimum dose to still lower values. Another potential area of improvement involves retuning the optimization parameters for low-dose image data, whereas the current study used nominal parameter settings defined in earlier work (Uneri et al 2013) at nominal dose. The sensitivity of the algorithm to surgical tools introduced in the patient has been previously studied (Otake et al 2013a), but their impact in relation to the low-dose limits deserves future investigation. Finally, the algorithm in its current form employs a geometrically calibrated C-arm for which the
registration transform amounts to 6 degrees-of-freedom (6 DoF), although extension to other platforms (e.g. mobile radiography) has been demonstrated in a 9 DoF implementation (Otake et al 2013b).

Future work is planned in which the methodology is extended to automatically localize surgical tools with known shape information (e.g. a needle or screw) through forward projection of known 3D models as in (Stayman et al 2012), thus achieving known tool tracking similar to that provided by a conventional surgical tracker. Ongoing work includes the evaluation of alternative similarity metrics and optimizers. For example, correlation-based metrics such as in (Gottesfeld Brown and Boult 1996) may be more robust against intensity differences and increased quantum noise (i.e. reduced detectability of gradients), and an optimizer such as multilevel coordinate search (Huyer and Neumaier 1999) that does not include a stochastic component may provide sufficient speed and robustness given good initialization (e.g. initialization by the registration computed in the previous shot).

Acknowledgements

The authors extend their thanks to Dr Aaron Kyle Jones (The University of Texas MD Anderson Cancer Center) for helpful discussion on measurement of fluoroscopy dose. Mr Nicholas Uebele (Johns Hopkins University) developed the wax ventricle models, which were shaped according to a population model derived from brain images drawn from The Cancer Imaging Archives (http://cancerimagingarchive.net). Financial support for this work was provided by an academic-industry partnership with Siemens XP (Erlangen, Germany).

Appendix A. Gradient masks

Previous work (Otake et al 2013c) showed that a gradient mask (\(M\)) can improve 3D-2D registration performance, particularly to mitigate the effect of deformable structures with strong gradients, such as the abdominal skinline. Within the context of intracranial neurosurgery considered in the experiments reported in this paper, we investigated a variety of gradient weighting masks with respect to the head and neck, as shown in figure A: (a) No Mask; (b) a Peripheral Mask down-weighting the cranial boundary; (c) a Central Mask down-weighting the cranial interior; (d) a Superior Mask down-weighting the cranial interior; and (e) an Inferior Mask down-weighting the temporal bones, mandible and neck. Such masks could potentially improve or degrade performance, depending on the strength of the gradient (e.g. cranium-to-air boundary), the extent to which the gradients are expected to deform (e.g. the mandible and neck), and the amount of noise in the unweighted region of the image (e.g. higher noise in higher attenuation medial regions). As shown in the TRE measurements, masking the peripheral or superior aspect of the cranium resulted in reduced performance, since the cranium presented a strong, non-degenerate gradient. On the other hand, masking the central or inferior aspects associated with more subtle gradients and higher image noise in regions about the temporal bone or sinuses provided a slight improvement in performance. The effect was fairly small and was not globally optimal (i.e. did not hold for all cases of \(\Delta \theta\)), so the use of a mask was not employed in the measurements reported in Section 3. Future work will consider generation of such masks in a more data-specific manner, such as those that are weak (susceptible to noise), inconsistent (susceptible to deformation), and/or degenerate (do not change with \(\theta\)).
Figure A. The effect of a gradient weighting mask. (a)–(e) Various masks designed to down-weight certain aspects of the image that may be more susceptible to noise or deformation. (f)–(i) TRE measured for each form of gradient masks (each corresponding to the dual-view (scenario #3) with 4 μGy total entrance surface dose) and four values of angular separation (Δθ).

References

Bjartmarz H and Rehncrona S 2007 Comparison of accuracy and precision between frame-based and frameless stereotactic navigation for deep brain stimulation electrode implantation Stereotact. Funct. Neurosurg. 85 235–42
Garcia D 2010 Robust smoothing of gridded data in one and higher dimensions with missing values Comput. Stat. Data Anal. 54 1167–78


Huyer W and Neumaier A 1999 Global optimization by multilevel coordinate search J. Glob. Optim. 14 331–55


Mahesh M 2001 Fluoroscopy: patient radiation exposure issues Radiographics 21 1033–45


Wein W, Roeppe B and Navab N 2005 2D/3D registration based on volume gradients Proc. of SPIE ed J M Fitzpatrick and J M Reinhartd (International Society for Optics and Photonics) pp144–50