# Standardized Evaluation Methodology for 2-D–3-D Registration

Everine B. van de Kraats*, Graeme P. Penney, Dejan Tomaževič, Theo van Walsum, and Wiro J. Niessen, *Associate Member, IEEE*

***Abstract*—In the past few years, a number of two-dimensional (2-D) to three-dimensional (3-D) (2-D–3-D) registration algorithms have been introduced. However, these methods have been developed and evaluated for specific applications, and have not been directly compared. Understanding and evaluating their performance is therefore an open and important issue. To address this challenge we introduce a standardized evaluation methodology, which can be used for all types of 2-D–3-D registration methods and for different applications and anatomies. Our evaluation methodology uses the calibrated geometry of a 3-D rotational X-ray (3DRX) imaging system (Philips Medical Systems, Best, The Netherlands) in combination with image-based 3-D–3-D registration for attaining a highly accurate gold standard for 2-D X-ray to 3-D MR/CT/3DRX registration. Furthermore, we propose standardized starting positions and failure criteria to allow future researchers to directly compare their methods. As an illustration, the proposed methodology has been used to evaluate the performance of two 2-D–3-D registration techniques, *viz.* a gradient-based and an intensity-based method, for images of the spine. The data and gold standard transformations are available on the internet (http://www.isi.uu.nl/Research/Databases/).**

***Index Terms*—2-D–3-D registration, evaluation, gold standard, ground truth, validation.**

## I. Introduction

TWO-DIMENSIONAL (2-D) to three-dimensional (3-D) (2-D–3-D) registration has been proposed to help in a number of clinical areas, such as radiotherapy planning and treatment verification, spinal surgery, hip replacement, neurointerventions and aortic stenting. The 2-D–3-D registration can be a means to noninvasively register the patient to an image volume used for image-guided navigation by finding the best match between one or more intraoperative X-ray projections of the patient and the preoperative 3-D volume.

Several researchers have described and evaluated 2-D–3-D registration methods, which can roughly be subdivided in feature-based [1]–[8], intensity-based [9]–[18], hybrid [15], [19] and, most recently, gradient-based [20], [21] methods. Their results are not directly comparable, owing to the use of different datasets, different starting positions, different failure criteria and different error calculation methods, as was also commented on in [20]. Only a few efforts have been aimed at comparing 2-D–3-D registration approaches [12], [15], [22], mainly comparing similarity measures for the intensity-based approach. Moreover, many of the commonly used error measures do not adequately represent the accuracy of registration performance [23].

Evaluating the performance of 2-D–3-D registration approaches requires a reference dataset with a known gold standard registration. Previously, gold standards for 2-D–3-D registration were either obtained manually [9], [22], by controlled motion [7], [8], [24], by independent measuring devices [3], [7], [11], by back-projection of images features (mainly implanted fiducials) [4], [9], [12], [15], [17]–[19], [21], [24], [25] or by using simulated data [1], [2], [7], [8], [10], [14], [16], [21]. These methods all have their own drawbacks: manual methods are often subjective, and can require a large amount of user interaction. Methods which use controlled motion, independent measuring devices or fiducials usually require the use of rigidly attached frames or markers which often limits their use to phantom or cadaver studies. Simulated projection images often lack a number of characteristics of real X-rays (e.g., coarser resolution) and may introduce a bias as the projection images are usually generated from the 3-D volume, often using the same method as is employed by the registration algorithm. Therefore, most of the described methods are cumbersome, and so are not easily repeated for other anatomies.

To address the challenge of comparing 2-D–3-D registration algorithms, we introduce a standardized evaluation methodology, which can be used for different algorithms and different applications, and can easily be extended to a large database including multiple anatomies and patient data. The proposed methodology exploits the precalibrated relationship that exists between acquired 2-D X-ray projections and reconstructed 3-D data in a 3DRX imaging system to generate a gold standard dataset. The standard 3DRX system calibration is fiducial based and the calibration is checked every six months. The motion of the 3DRX system is highly reproducible, and there is a fixed relationship between each of the X-ray projections used to reconstruct the 3DRX image and the resultant 3DRX volume. This means that the distortion corrected projections, the 3DRX volume and the calibrated projection geometry can be used as a gold standard for 2-D–3-D registration. This has the big advantage that no markers need to be rigidly attached to the object being imaged. Therefore, it is much easier to

Manuscript received February 17, 2005; revised June 6, 2005. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was P. Suetens. *Asterisk indicates corresponding author.*

*E. B. van de Kraats is with the Image Sciences Institute, University Medical Center Utrecht, room Q0S.459, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands (e-mail: everine@isi.uu.nl)

G. P. Penney was with the Image Sciences Institute, University Medical Center Utrecht, The Netherlands. He is now with the Centre for Medical Image Computing, University College London, London WC1E 6BT, UK.

D. Tomaževič is with the Faculty of Electrical Engineering, University of Ljubljana, SI-1000 Ljubljana, Slovenia.

T. van Walsum and W. J. Niessen were with the Image Sciences Institute, University Medical Center Utrecht, The Netherlands. They are now with the Biomedical Imaging Group Rotterdam, Erasmus MC, University Medical Center, Departments of Radiology and Medical Informatics, 3000 DR Rotterdam, The Netherlands.

Digital Object Identifier 10.1109/TMI.2005.853240

acquire gold standard data from patients compared to using conventional fiducial-based methods. The 2-D–3-D gold standard registration can be extended to other 3-D modalities, such as computed tomography (CT) or magnetic resonance (MR), by 3-D–3-D image registration of the CT or MR volume to the 3DRX data using established and well-validated intensity-based registration [26]. Our methodology also comprises a standardized procedure for generating starting positions, and appropriate definitions for error measures and failure criteria.

To show the usefulness of such an evaluation methodology we apply our methodology to the registration of 3DRX, CT and MR volumes of vertebral bodies to mono- and biplane X-ray projections using two registration algorithms: gradient-based [20] and intensity-based [12]. These two approaches were selected because they have been developed for 2-D–3-D spine registration. From the information present in the papers it was not possible to compare them. Since the original implementations were available to us, these two methods served well in the demonstration of the proposed standardized evaluation methodology.

## II. Evaluation Methodology

This section describes in detail how the gold standard data was generated, how registration accuracy was assessed, how the capture range and failure criteria were defined, how starting positions were determined and how the data was presented to the registration algorithms.

### A. Generation of Gold Standard Data

In the proposed evaluation methodology, gold standard data is acquired with a clinical floor-mounted 3DRX C-arm system (Integris BV5000, Philips Medical Systems, Best, The Netherlands). During an 8-s run of 180° around the patient the 3DRX system acquires 100 projection images, which are used to reconstruct a high-resolution 3-D volume using a filtered back-projection reconstruction technique [27]. The C-arm system is calibrated for 3DRX imaging [28] at installation. Two calibration runs are required: the first run determines the image intensifier distortion and the position of the focal spot of each projection with respect to the detector. The second run establishes the projection parameters (rotations around the $x$-, $y$-, and $z$-axis of the calibration phantom and shifts in the $x$, $y$, and $z$ directions) for each projection. The most crucial aspect for our evaluation methodology is the accurate projection geometry calibration of the system, which has been evaluated in a number of studies. The accuracy of the derived geometrical parameters has been demonstrated in [29], where the 3-D image quality was evaluated in terms of spatial resolution and geometrical accuracy. It was shown that the system can obtain an almost isotropic 3-D resolution of up to 22 lp/cm (line pairs per cm). Furthermore, Rasche *et al.* [29] showed that the 3-D geometrical accuracy of a reconstructed volume was in the order of its spatial resolution. Using the same 3DRX system, Movassaghi *et al.* [30] investigated the geometrical accuracy of 3-D point reconstruction in a bullet phantom from pairs of 2-D projections using the calibrated projection geometry. Their results showed only a small error as a function of the projection pair used ($0.26 \pm 0.08$ mm). The small standard deviation of 0.08 mm shows that the projection geometry produces spatially consistent volumes. Also, using the same system, Baert *et al.* [31] performed guide wire reconstruction in 3-D from 2-D projections, achieving a high accuracy. Moreover, they showed that when introducing small errors into the projection geometry information, a large guidewire reconstruction error was obtained. This again indicates that the geometrical calibration process is highly accurate.

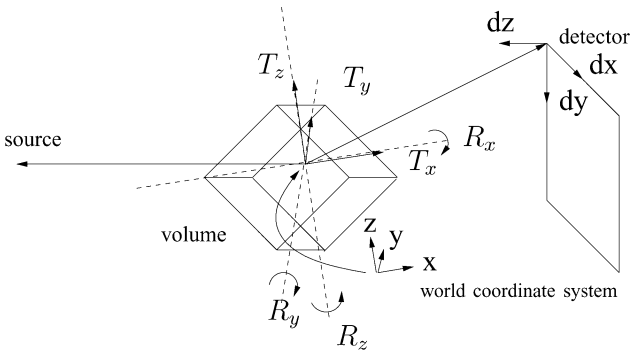


Fig. 1. The six degrees of freedom ($T_x$, $T_y$, $T_z$, $R_x$, $R_y$, and $R_z$) in our rigid body 2-D–3-D registration problem. The intrinsic parameters, which are the source and detector positions and orientations, are known. The world coordinate system is intrinsically linked to the center of the initially reconstructed 3DRX volume, and it is completely determined by the 3DRX system calibration.

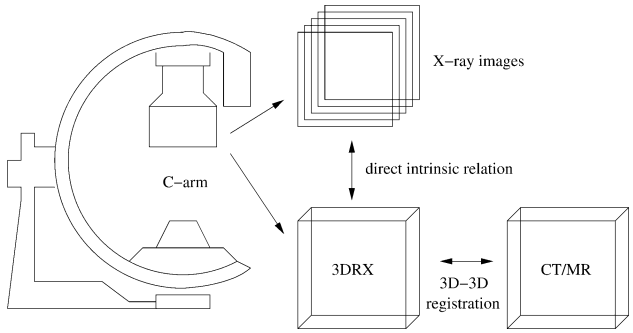


Fig. 2. Gold standard data obtained with 3DRX C-arm. The X-ray images are used to reconstruct the 3DRX volume and they are, thus, directly related to the 3DRX volume. The CT and MR data are registered to the 3DRX data using 3-D–3-D registration, thereby obtaining the gold standard for CT/MR data to X-ray images.

As a result of the calibration, both the projection geometry of the X-ray images with respect to the reconstructed volume and the distortion of the X-ray images are known [32]. Hence, the 3DRX volume and the 100 corresponding distortion corrected X-ray projections can be used as a gold standard dataset for 2-D–3-D registration. Fig. 1 shows all the degrees of freedom which are known for the calibrated 3DRX. For the 2-D–3-D registration, the intrinsic parameters are calculated using a calibration step and then the registration algorithms calculate the extrinsic parameters. The extrinsic parameters describe the position and orientation of the 3-D volume in space (translations in $x$, $y$, and $z$ directions ($T_x$, $T_y$, and $T_z$) and rotations around $x$-, $y$-, and $z$-axis ($R_x$, $R_y$, $R_z$)).

The gold standard can be extended to other 3-D modalities (such as CT or MR) by registering corresponding CT or MR data to the 3DRX data, thereby indirectly obtaining the transformation to the X-ray images (Fig. 2). An application based on the Multi-Modality Image Registration using Information Theory (MIRIT) software by the Laboratory for Medical Imaging Research, Leuven, Belgium [26] was used for 3-D–3-D registration. This algorithm optimizes mutual information. Voxel-based

3-D–3-D registration methods have been previously described in the literature and have been validated to be subvoxel accurate for certain applications [33]. In a previous study, we showed that, for spine images, mutual-information-based registration of MR data to 3DRX data, with an isotropic resolution of $2^3\ \mathrm{mm}^3$ and $0.62^3\ \mathrm{mm}^3$ respectively, yields a root-mean-squared error (RMSE) in the order of 0.8 mm at fiducial markers, and is at least as accurate as marker-based registration [34].

### *B. Method for Assessment of 2-D–3-D Registration Accuracy*

For our evaluation we introduce the following transformation matrices: $T_{\mathrm{gold}}$ is the $4 \times 4$ gold standard transformation matrix from CT/MR image space to the 3-D space of the 3DRX system obtained by registering the CT/MR volumes to the 3DRX volume; $T_{\mathrm{reg}}$ is the same transformation but obtained after having performed 2-D–3-D registration. $M_{\mathrm{gold}} = T_{\mathrm{proj}} T_{\mathrm{gold}}$ is the $3 \times 4$ gold standard projection matrix from 3-D space of the 3DRX system to 2-D space of the X-ray projection, where $T_{\mathrm{proj}}$ is obtained from the system calibration, and $M_{\mathrm{reg}} = T_{\mathrm{proj}} T_{\mathrm{reg}}$ is the same projection matrix but obtained after 2-D–3-D registration.

Various performance measures have been used in view of different registration tasks. LaRose *et al.* [23] categorized the measures as 2-D image space error, pose parameter error and 3-D registration error. Two-dimensional image space error is the error computed using 2-D image space features (e.g., fiducials or contours), where small errors in 2-D can be caused by large errors in 3-D. The pose parameter error, which consists of the six-dimensional vector $(T_x, T_y, T_z, R_x, R_y, R_z)$, can give misleading results as it depends on the position of the center of rotation, because the rotational and translational components of the transformation are not independent. Lastly, the 3-D registration error computes the pose position error in 3-D.

Since the presented standardized evaluation methodology is meant for 2-D–3-D registration in general, different performance measures suited for different tasks are used.

A widely used 3-D error measure is the target registration error (TRE) [35], where the "targets" in the TRE calculation can be predefined locations (either fiducials or landmarks), surface points [21], or arbitrary chosen points inside a region of interest. We propose to compute a mean TRE (mTRE) to determine the 3-D error of a registration. For $k$ points $\mathbf{p}_i$ in a fixed set $P$ of 3-D points on a regular grid (i.e., uniformly distributed), the distance between the point $\mathbf{p}_i$ transformed with $T_{\mathrm{gold}}$, the gold standard registration, and the same point transformed with $T_{\mathrm{reg}}$, the transformation determined by the registration algorithm, is computed. The average distance between the points defines the mTRE

$$\mathrm{mTRE}(P, T_{\mathrm{reg}}, T_{\mathrm{gold}}) = \frac{1}{k} \sum_{i=1}^{k} \| \, T_{\mathrm{reg}} \, \mathbf{p}_i - T_{\mathrm{gold}} \, \mathbf{p}_i \, \| \, . \quad (1)$$

The mTRE, and other errors described below, are depicted in Fig. 3.

The 3-D registration error can be best used when the application is patient-to-image registration for navigation, because it computes the error on specific points of interest within the registered volume where manipulations will take place.

However, when using a single X-ray image for assessing the 3-D position of an object, determining the 3-D spatial position in the projection direction is difficult. To assess the magnitude of the error in the projection direction, the mTRE in the projection direction, mTREproj, is computed. This is done by averaging the component of the displacement vector that is normal to the X-ray plane

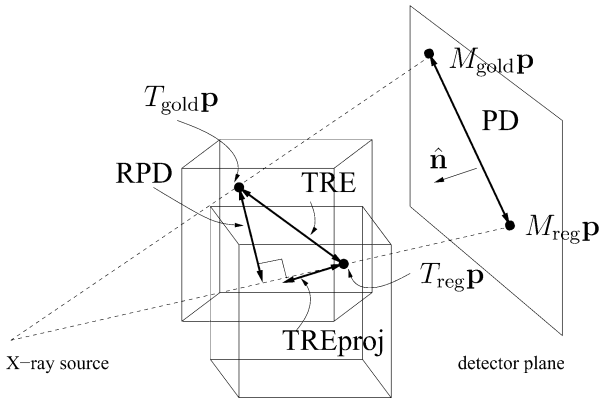


Fig. 3. Depiction of the error for each evaluation method displayed for a single point within the object, where TRE is the target registration error, TREproj the error in the projection direction, PD the projection distance, RPD the reprojection distance, $T_{\mathrm{gold}}$ the gold standard transformation, $T_{\mathrm{reg}}$ the registration transformation, $M_{\mathrm{gold}}$ the gold standard projection matrix, $M_{\mathrm{reg}}$ the registration projection matrix and $\hat{\mathbf{n}}$ the normal to the projection plane.

$$\mathrm{mTREproj}(P, T_{\mathrm{reg}}, T_{\mathrm{gold}}, \hat{\mathbf{n}}) = \frac{1}{k} \sum_{i=1}^{k} \| \, (T_{\mathrm{reg}} \, \mathbf{p}_i - T_{\mathrm{gold}} \, \mathbf{p}_i) \cdot \hat{\mathbf{n}} \, \| \quad (2)$$

where $\hat{\mathbf{n}}$ is the normal to the projection plane.

In applications where 2-D–3-D registration is used for projecting 3-D information onto a 2-D image plane, 2-D error measures are more appropriate than the mTRE. In this case, we use an error termed the mean projection distance (mPD) which is the error in the projection plane. More specifically, projection distance is the distance between the projection of a 3-D point at registration position and the projection of the same point at gold standard position

$$\mathrm{mPD}(P, M_{\mathrm{reg}}, M_{\mathrm{gold}}) = \frac{1}{k} \sum_{i=1}^{k} \| \, M_{\mathrm{reg}} \, \mathbf{p}_i - M_{\mathrm{gold}} \, \mathbf{p}_i \, \| \quad (3)$$

where $M_{\mathrm{reg}}$ and $M_{\mathrm{gold}}$ are the final perspective projection matrices for the registration and the gold standard position respectively, and where the distance is calculated after division by the homogeneous coordinate.

Rather than using the mean projection distance, some authors have used the reprojection distance [36], which computes the minimum distance between the line from the 2-D projected point at the registered position to the X-ray source and the gold standard 3-D position of that point. This line can also be described by the line through the 3-D point at the registered position and the X-ray source. The mean reprojection distance (mRPD) is calculated as follows:

$$\mathrm{mRPD}(P, T_{\mathrm{reg}}, T_{\mathrm{gold}}) = \frac{1}{k} \sum_{i=1}^{k} D(L_i(\mathrm{source}, T_{\mathrm{reg}}\mathbf{p}_i), T_{\mathrm{gold}}\mathbf{p}_i). \quad (4)$$

Here, $D(L_i, T_{\text{gold}}\mathbf{p}_i)$ is the minimum distance between the 3-D point $\mathbf{p}_i$ at gold standard position and a line $L_i$, where $L_i(\text{source}, T_{\text{reg}}\mathbf{p}_i)$ represents a 3-D line through the 3-D point at the registered position and the X-ray source. Unlike the mPD, the reprojection distance is independent of the distance to the projection plane.

The authors recommend that for evaluation of the performance of 2-D–3-D registration using a single X-ray image the mPD and mRPD are to be determined along with the more general mTRE.

### C. Capture Range and Failure Criteria

The capture range defines the range of starting positions from which an algorithm finds a sufficiently accurate transformation. As shown in the previous section, accuracy can be determined using various error measures. Starting ranges can also be determined using these error measures, describing the starting range as the initial accuracy from where registrations are still successful. Although the initial accuracy can be evaluated using different error measures, we always determine the capture range using the initial mTRE, which enables direct comparison of reported capture ranges. Moreover, the initial mTRE is most informative about the initial misalignment. Since the capture range naturally starts from zero, it is reported by stating the maximum start mTRE value of this range. Two factors are involved to calculate the capture range: the definition of a misregistration (using either mTRE, mPD or mRPD), and the fraction of allowed misregistrations. Both these factors depend on the application of the 2-D–3-D registration method.

### D. Starting Positions

Several approaches are possible for determining the initial distance to the gold standard of a starting position. Tomaževič *et al.* [20] use a "normalized distance," that relates rotations to linear distances. Rohlfing *et al.* [14] use the effective motion in the projection plane per parameter for normalization. We choose to use the mTRE at the starting position as a measure for the initial distance to the gold standard, similar to Russakoff *et al.* [15]. The mTRE is determined over a fixed set of 3-D points uniformly distributed in an average-sized region (which can differ per anatomy). By using the same quantity for initial distance and final result, the effect of the algorithm is directly measurable. Of course, the effect on the end mPD or end mRPD is not directly measurable when plotting them against the start mTRE.

The starting positions are offsets from the gold standard position. Given our definition of capture range, and the use of the initial mTRE as the distance measure for the starting positions, the starting parameter values ($T_x$, $T_y$, $T_z$, $R_x$, $R_y$, $R_z$) must be generated such that several resulting transformations are within several ranges of mTRE. This is achieved by the following.

1) Choosing intervals for the starting position distance, e.g., 0–1, 1–2 mm, and determining the center of rotation, which generally is the center of the fixed set of 3-D points.
2) Then, for each interval, for each of the six transformation parameters, the range is determined that will yield an mTRE less than or equal to the interval upper bound, e.g., for a starting mTRE between 1 and 2 mm, the range for each translation is 0–2 mm. Since the region used to determine the starting mTREs is not necessarily cubic, rotations around different axes may not have the same effect on the mTRE. This is taken into account in the following way: the angle of rotation that results in an mTRE of 1 mm for each individual rotation is calculated. This angle is then linearly scaled to determine the maximum allowed rotation for the other intervals. In this calculation, the small angle approximation is used, i.e., the effective displacement is linearly related to the angle.
3) Next, for each interval, transformations are generated, where the transformation parameters are chosen randomly (uniformly distributed) from their predetermined range. Subsequently, the mTRE of the composite transformation is determined, and if that mTRE is within the interval it is kept, otherwise the transformation is discarded.

The last step is repeated until each interval contains a predetermined number of starting positions.

### E. Data Preparation

Most registration algorithms only use a part of the available image data, to focus the registration on the object of interest and to reduce the amount of data, resulting in lower computational costs. Since some algorithms reduce data by using regions of interest (ROIs) in the 2-D projection images and others by using volumes of interest (VOIs) in the 3-D volume, it is not possible to have exactly the same data input for both types of algorithms. For the presented evaluation methodology the VOIs are manually determined in the 3DRX volume, by determining a rectangular region of interest around the anatomy. The VOIs should at least contain the anatomy that is used and relevant for the registration. When performing inter-modality comparison the VOIs are transformed to the corresponding CT or MR dataset using the gold standard image-based 3-D–3-D registration. Rectangular ROIs are also manually determined in the X-ray image so that they contain approximately the same anatomy as is contained in the VOIs. However, the ROIs have no direct relation to the VOIs in the 3-D volumes. Corresponding ROIs could have been determined in the X-ray images by projecting the VOIs in the 3-D volume to the 2-D image using the gold standard calibration. However, the effect on the registration when using slightly different ROIs is expected to be minimal. Moreover, for algorithms that have both a data reduction in 2-D and in 3-D directly related VOIs and ROIs should not be used as this would bias the registration.

## III. Experiment 1: Comparison of Two Methods

### A. Methods

To illustrate the potential of the proposed evaluation framework, we applied it to compare two previously published 2-D–3-D registration methods, *viz.* a gradient-based and an intensity-based method. For both algorithms the original implementations were available, though the intensity-based algorithm had to be adapted so it could take multiple X-ray images as input.

The gradient-based 2-D–3-D registration method [20] registers 3-D volumes to 2-D images by searching for the best match

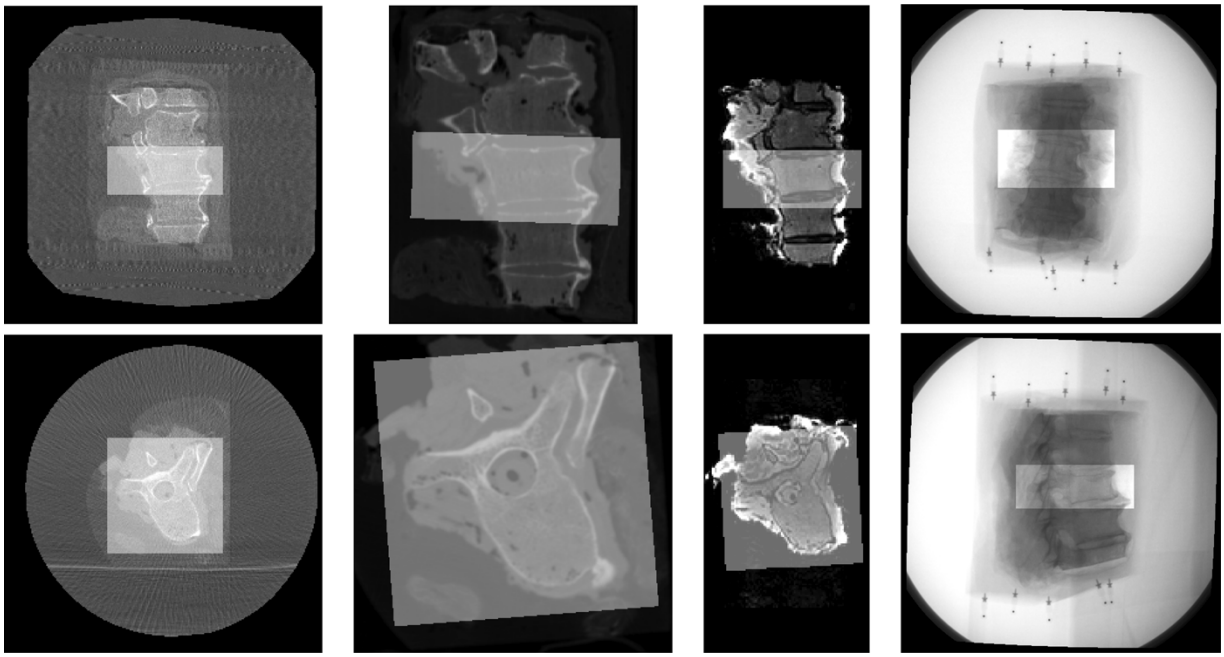

Fig. 4. Top row from left to right: coronal planes with VOI for 3DRX, CT, and MR data, and an anterior-posterior X-ray image with ROI. Bottom row from left to right: transverse planes with VOI for 3DRX, CT, and MR data, and a lateral X-ray image with ROI.

TABLE I
PARAMETER VALUES FOR THE REGISTRATION METHODS. THE "/" IS USED WHEN THE PARAMETER VALUES ARE DIFFERENT FOR THE TWO SPINAL SEGMENTS. "GB" REPRESENTS THE GRADIENT-BASED METHOD AND "IB" THE INTENSITY-BASED METHOD

| | Parameter | MR | 3DRX | CT |
|---|---|---|---|---|
| GB | ($\sigma$=Blur volume) | 0.3 | 0.5 | 0.3 |
| | ($\sigma$=Blur X-ray) | 0.5/1.0 | 0.5/1.0 | 0.5/1.0 |
| | SampleStep | 1.0 mm | 1.0 mm | 1.0 mm |
| | Threshold | 10 | 25 | 18 |
| IB | MinStep | | 0.125 | 0.125 |
| | Threshold | | 90/80 | 40 (≈400 HU) |

TABLE II
SUMMARY OF IMAGE DIMENSIONS AND VOXEL SIZES FOR THE IMAGES ACQUIRED OF TWO SPINE SEGMENTS USING FOUR DIFFERENT MODALITIES

| Modality | Segment | Resolution | Size |
|---|---|---|---|
| MR | 1 | $1.00\times0.75\times0.75\,\text{mm}^3$ | $100\times256\times256$ |
| | 2 | $1.00\times0.88\times0.88\,\text{mm}^3$ | $120\times256\times256$ |
| CT | 1 | $0.31\times0.49\times0.31\,\text{mm}^3$ | $320\times260\times320$ |
| | 2 | $0.31\times0.49\times0.31\,\text{mm}^3$ | $280\times300\times300$ |
| 3DRX | 1 | $0.87\times0.87\times0.87\,\text{mm}^3$ | $256^3$ |
| | 2 | $0.52\times0.52\times0.52\,\text{mm}^3$ | $256^3$ |
| X-ray | 1 | $0.63\times0.63\,\text{mm}^2$ | $512^2$ |
| | 2 | $0.53\times0.53\,\text{mm}^2$ | $512^2$ |

between surface normals in the 3-D volume and back-projected gradients from the 2-D X-ray images. The surface normals and their positions are extracted from the 3-D volume in a preprocessing step. In this preprocessing step, the volume is blurred with a Gaussian filter, isotropically resampled, and the locations with a gradient magnitude larger than a predefined threshold are extracted. The gradient-based registration method needs the following parameters: volume blur, X-ray image blur, sample step and a gradient threshold. Although the original authors have only used their method for registration with multiple X-ray images, we also applied it in registrations using only one X-ray image. See Table I for the parameter values used in this evaluation. The parameters were chosen such that correct points were extracted in the 3-D volume, which required some tuning especially for the MR volume.

The intensity-based registration method [12] compares digitally reconstructed radiographs (DRRs or simulated projections) to X-ray images using a similarity measure called gradient difference. The algorithm has two parameters: bone threshold and minimal step-size. The bone threshold is manually determined so that only voxels containing bone contribute to the DRRs. The minimal step-size is the smallest step that the algorithm takes in the optimization procedure, which is a factor by which the position parameter values can change in an iteration. The rotations, which are expressed in radians, need an extra factor of 0.02 to make the steps comparable to the translation steps, which are in millimeters. The optimization uses a multi-resolution approach, where initially coarse resolution $256 \times 256$ pixel DRR images are generated, and then finer $512 \times 512$ pixel images. See Table I for a summary of the parameter values used, which were considered typical by the authors of the algorithm.

To select the gradients in the volume of interest, i.e., at the bone interface, an interactive windowing step is required for the gradient-based method when using 3DRX and MR data. In order to make the evaluation comparable, the same data was used for the intensity-based method. Both methods were directly applied to the CT data, whose intensity range was converted to comply with the input criteria of the methods.

### *B. Image Data*

Two defrosted segments of a spinal column, both fixed in foam, were imaged with three modalities: 3DRX, CT, and MR. The 3DRX images were reconstructed using 100 projection images acquired with a clinical 3DRX system (Integris BV5000, Philips Medical Systems). As mentioned before, the geometric relation and distortion correction for each of these images are known from the standard 3DRX system calibration. Spinal segment 1 was obtained with an image intensifier (II) size of 38 cm and spinal

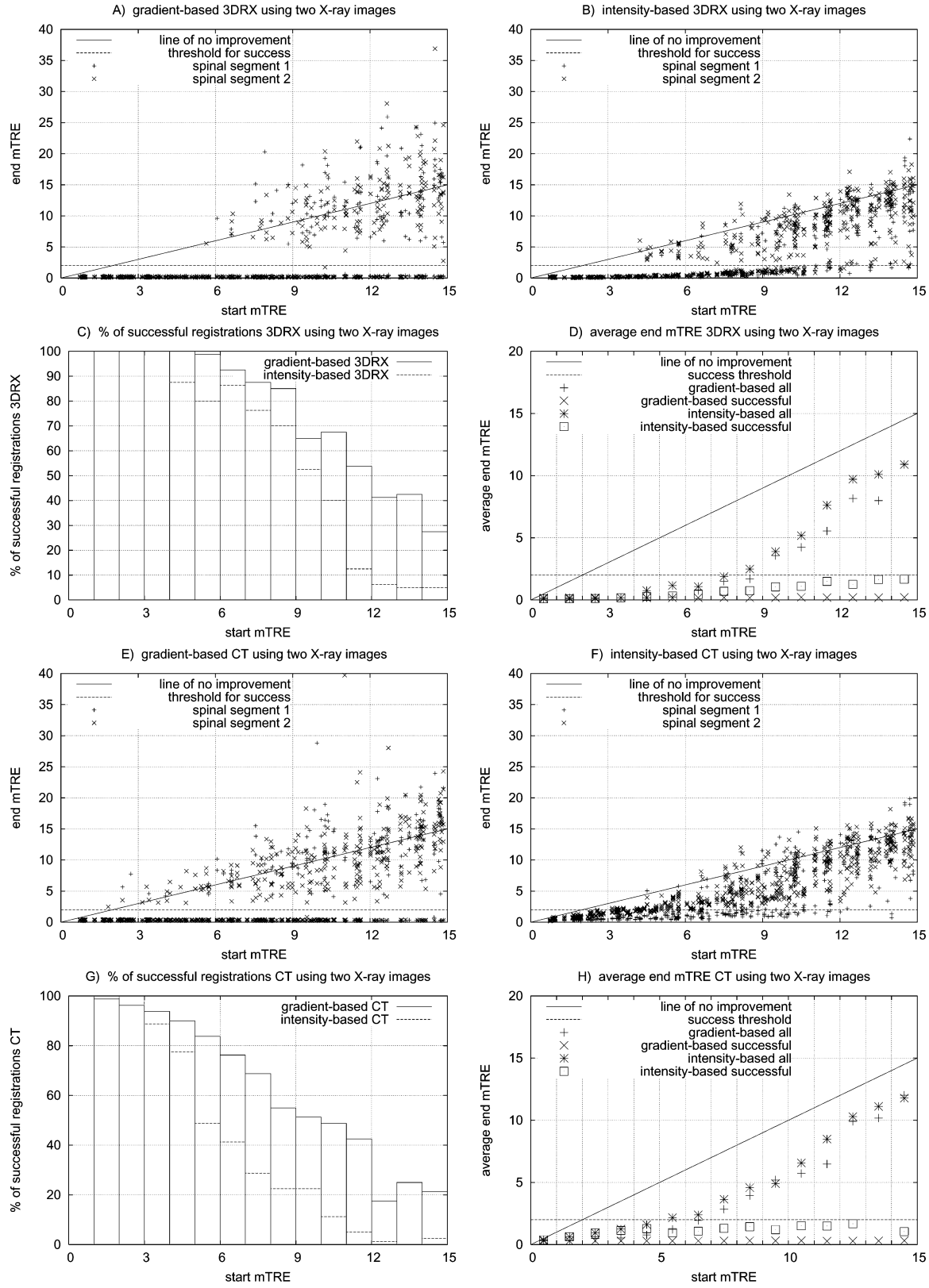


Fig. 5. Results for 3DRX and CT to X-ray registration using two X-ray images. For brevity, all 450 registrations (150 × 3 vertebral bodies) and all 750 registrations (150 × 5 vertebral bodies) for both spinal column segments were labeled as spinal segment 1 and spinal segment 2, respectively.

segment 2 with an II size of 31 cm. The CT volumes were acquired with a clinical multi-slice CT scanner (MX8000, IDT 16, Philips Medical Systems). The MR volumes were acquired with a clinical 1.5-Tesla MR scanner (Gyroscan NT, Philips Medical Systems) using a sagittal 3-D turbo spin echo acquisition, with a turbo factor of 29, TR/TE of 1500 ms/90 ms.

Spinal segment 1 consists of 3 thoracolumbal vertebral bodies and segment 2 consists of 5 thoracic vertebral bodies. Some soft

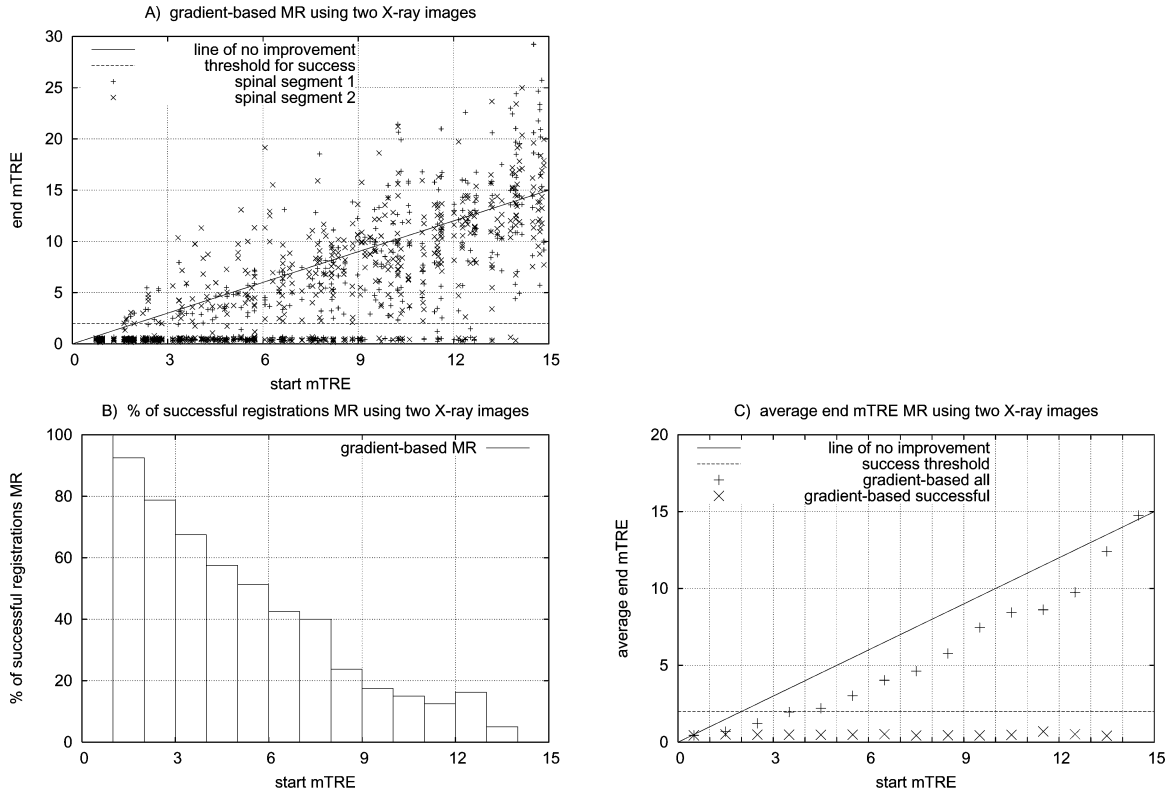


Fig. 6. Results for MR to X-ray registration using the gradient-based method using two X-ray images. For brevity, all 450 registrations (150 × 3 vertebral bodies) and all 750 registrations (150 × 5 vertebral bodies) for both spinal column segments were labeled as spinal segment 1 and spinal segment 2, respectively.

tissue around the spinal segments was still present. See Table II for details on image sizes and resolutions.

The CT and MR volumes were registered to the corresponding 3DRX volumes to establish the gold standard for both CT to X-ray registration and MR to X-ray registration using the previously described maximization of mutual information [26].

The rectangular VOIs used by the gradient-based method were manually determined in the 3DRX volumes for each vertebra. The size of these VOIs was approximately $100 \times 50 \times 100$ voxels in segment 1 and $120 \times 50 \times 150$ voxels in segment 2. The VOIs were transformed to the CT and MR data using the gold standard image-based 3-D–3-D registration. In the X-ray images, the rectangular ROIs that were used by the intensity-based method were manually determined around the vertebral bodies (approximately $170 \times 70$ pixels). See Fig. 4 for examples of the data with corresponding VOIs and ROIs. For each of the eight vertebral bodies, centers of rotation were determined in the world coordinate system (directly linked to the 3DRX data).

### C. Experiments and Evaluation

The starting positions were generated as described in Section II, with 10 starting positions per 1 mm interval up to 15 mm mTRE, resulting in 150 starting positions. Ten starting positions per bin were chosen because we considered 80 ($10 \times 8$ vertebral bodies) registrations per 1-mm bin sufficient to assess the capture range. In this starting position generation process, the mTRE was determined over a fixed set of 3-D points in an average vertebra-sized region of $95 \times 45 \times 95$ voxels (size $0.87^3$ mm$^3$) (as described in Section II). For the accuracy evaluation this fixed grid was also used. The same starting positions (offsets from the gold standard) were used for each vertebra.

From the 100 acquired X-ray images per spinal segment two images were selected, one anterior-posterior (AP) and one lateral (LAT), which were used for the registration experiments for both 3DRX, CT, and MR to X-ray registration. For 3DRX and CT, we also performed experiments using only one X-ray image, which was AP for spinal segment 1 and LAT for spinal segment 2.

In this example evaluation, registrations were classified as successful when the end mTRE over the fixed set of 3-D points (same as was used for the generation of the starting positions) was smaller than 2 mm. This threshold was also used in [20]. It is however application specific as the required accuracy for spine procedures varies from 0.0 mm to 3.8 mm [37] in different levels of the spine. The capture range was defined as the 95% success range. When using a single X-ray image the accuracy evaluation was extended to include end mPD and end mRPD, where success was also defined as an error smaller than 2 mm to illustrate the effect of different accuracy measures on the observed performance of an algorithm.

## IV. Experiment 2: Evaluation of the Gold Standard

The results of the evaluation depend on the accuracy of the gold standard. Earlier studies [30], [31] as discussed in Section II-A indicated that the calibration is very accurate. Previous investigation shows that errors in multi-modality 3-D–3-D registration can be in the order of the voxel size of the images [33]. To investigate how errors in the calibration and in the 3-D–3-D registration propagate into 2-D–3-D registration errors, a number of experiments was performed.

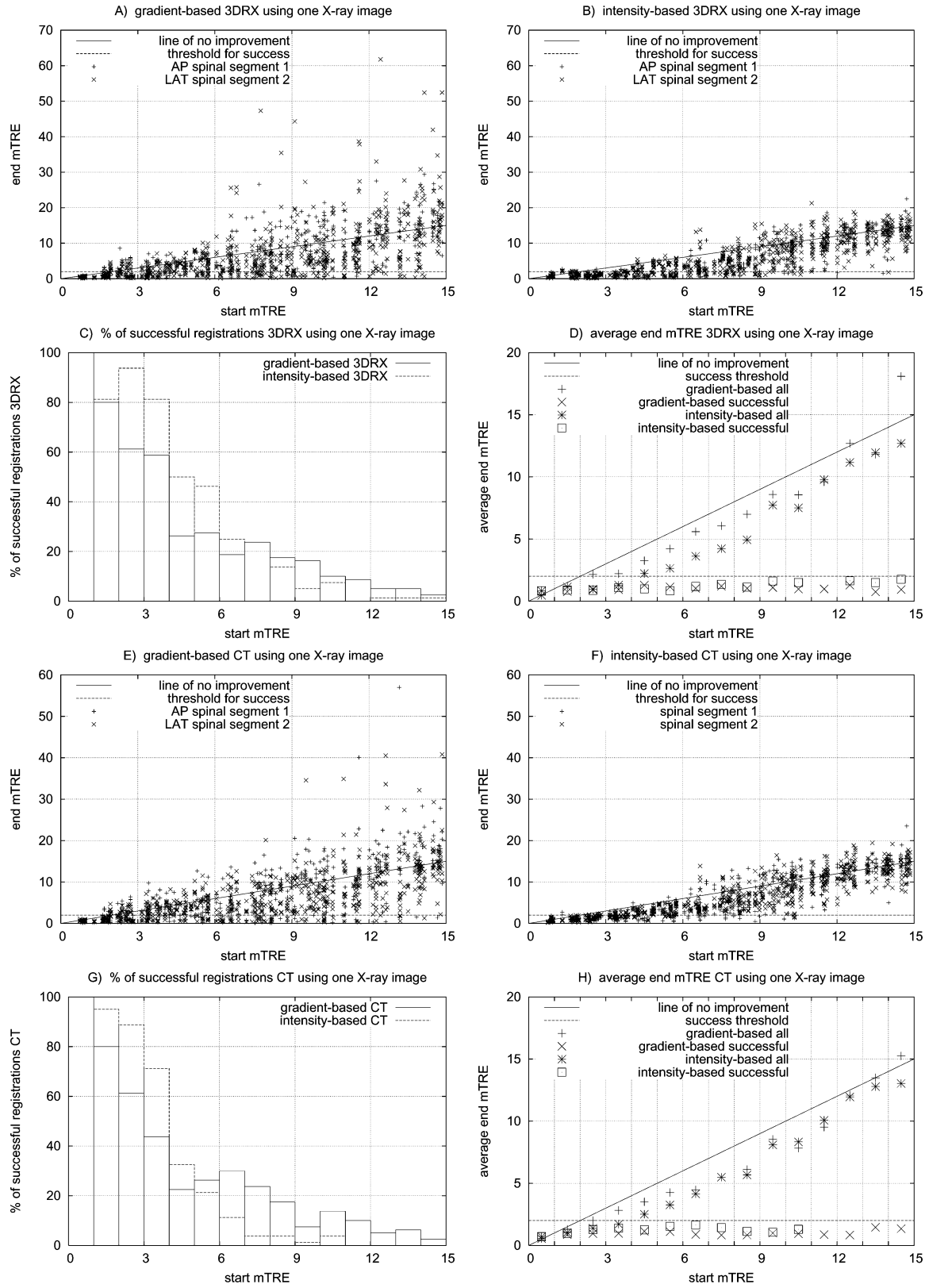


Fig. 7. Results for 3DRX and CT to X-ray registration using one X-ray image. For brevity, all 450 registrations (150 × 3 vertebral bodies) and all 750 registrations (150 × 5 vertebral bodies) for both spinal column segments were labeled as spinal segment 1 and spinal segment 2, respectively.

Errors in the calibration were simulated by adding 0.5° and 1.0° to the rotation angle (roll direction of C-arm), and 0.5° and 1.0° to the angulation angle (tilt direction of C-arm) of the calibrated projection geometry determined in the standard system calibration. These errors are larger than the expected gold standard error, as has been shown in previous studies [31].

Errors in the 3-D–3-D registration were simulated by adding random transformations to the gold standard that introduce an

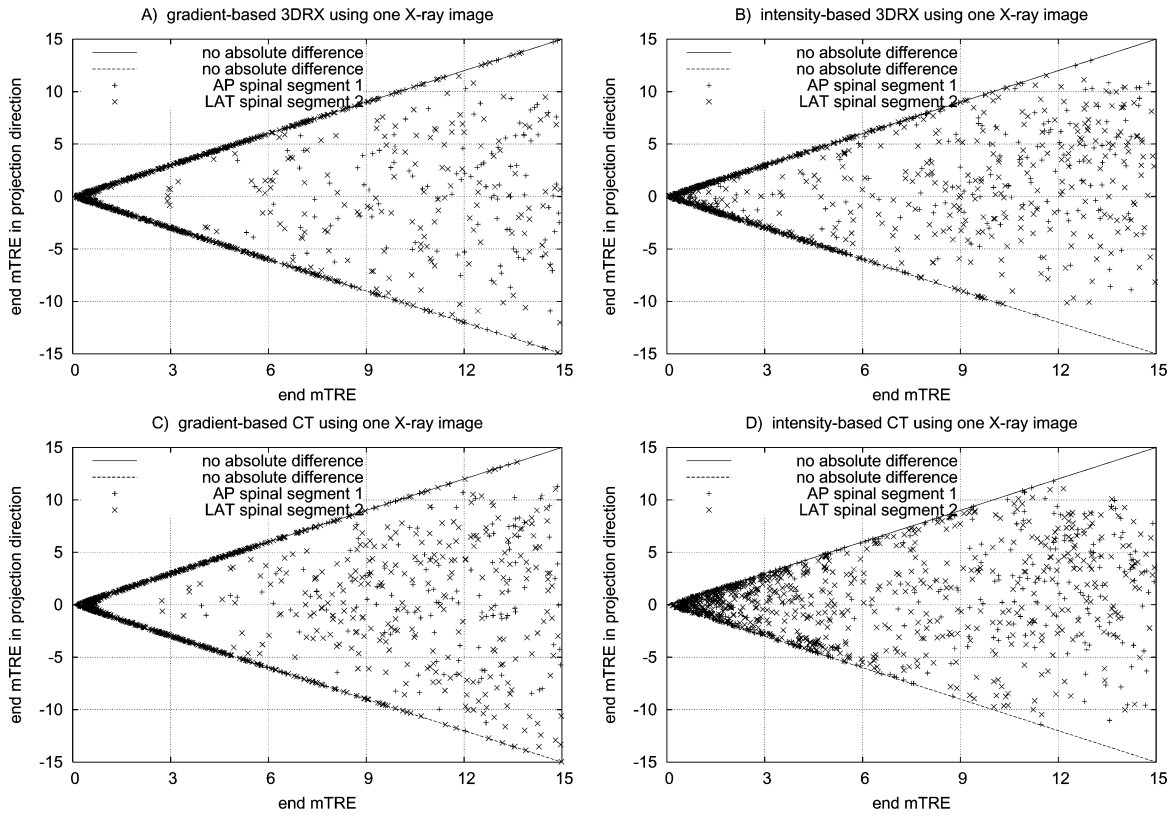


Fig. 8. Mean TREs in the projection direction. For brevity, all 450 registrations (150 × 3 vertebral bodies) and all 750 registrations (150 × 5 vertebral bodies) for both spinal column segments were labeled as spinal segment 1 and spinal segment 2, respectively.

error of up to 1 mm mTRE (which is approximately the voxel size). These error transformations were generated in a similar fashion as the 2-D–3-D starting positions. Ten 3-D–3-D error transformations were produced per bin of size 0.2 mm.

For each of these specific simulated errors (both calibration and 3-D–3-D registration errors) 100 registrations for one vertebral body in spinal segment 2 were performed using the gradient-based method. For the calibration error experiments 3DRX was registered to two X-ray projections, and for the 3-D–3-D registration error experiments CT was registered to two X-ray projections. The average registration errors (mean mTRE) for successful registrations per bin were determined and compared to the mean mTRE when using the real gold standard.

## V. Results and Discussion

### A. Experiment 1: Comparison of Two Methods

We performed 150 registrations for each of the eight vertebral bodies, for each modality and for a different number of X-ray images (either single plane or using two X-ray images simultaneously). Therefore the following experiments were carried out: 3DRX to one X-ray, 3DRX to two X-rays, CT to one X-ray, CT to two X-rays, and MR to two X-rays. For two X-ray images, the average registration time was approximately 25 s for the gradient-based method, running on a Windows 2000 Dell Workstation PWS340 Intel Pentium 4, 1.7 GHz, 1.6 GB RAM, and approximately 9 min for the intensity-based method, running on a multi-user Linux Dell PowerEdge 1600, dual Xeon 2.8 GHz with 4GB memory. The algorithms were not optimized for speed.

The results for registrations using two 2-D X-ray images are presented in Fig. 5 (for CT and 3DRX) and in Fig. 6 (for MR). Fig. 7 shows the results of registering CT and 3DRX to a single 2-D X-ray image. Each of these figures presents the registration results for both algorithms, except for Fig. 6, because only the gradient-based algorithm was used on MR. The intensity-based method was not performed on the MR dataset, as no DRRs can be generated from MR images. The end mTRE values for each registration, the percentage of successful registrations, the average end mTRE per bin for all registrations and for successful registrations are displayed. These evaluations are extended for the registrations based on a single X-ray image by also calculating end mTREproj (Fig. 8) and end mPD and end mRPD (Fig. 9).

From the results for registrations performed using two X-ray images several conclusions can be drawn. The average error for successful registrations is stable for the gradient-based method, while the intensity-based method has increasing difficulty in finding the optimal position when the initial offset from the gold standard position increases [Fig. 5(d),(h)]. Failure determination is most probably easier for the gradient-based method [Fig. 5(a),(e)] than for the intensity-based method [Fig. 5(b),(f)], because the gradient-based registration seems to be either correct or incorrect. In a clinical situation, a failure can be visually assessed, by displaying the DRR belonging to the registration next to (or on top of) the X-ray image. For both methods the average error is larger for CT to X-ray registration than for 3DRX to X-ray registration.

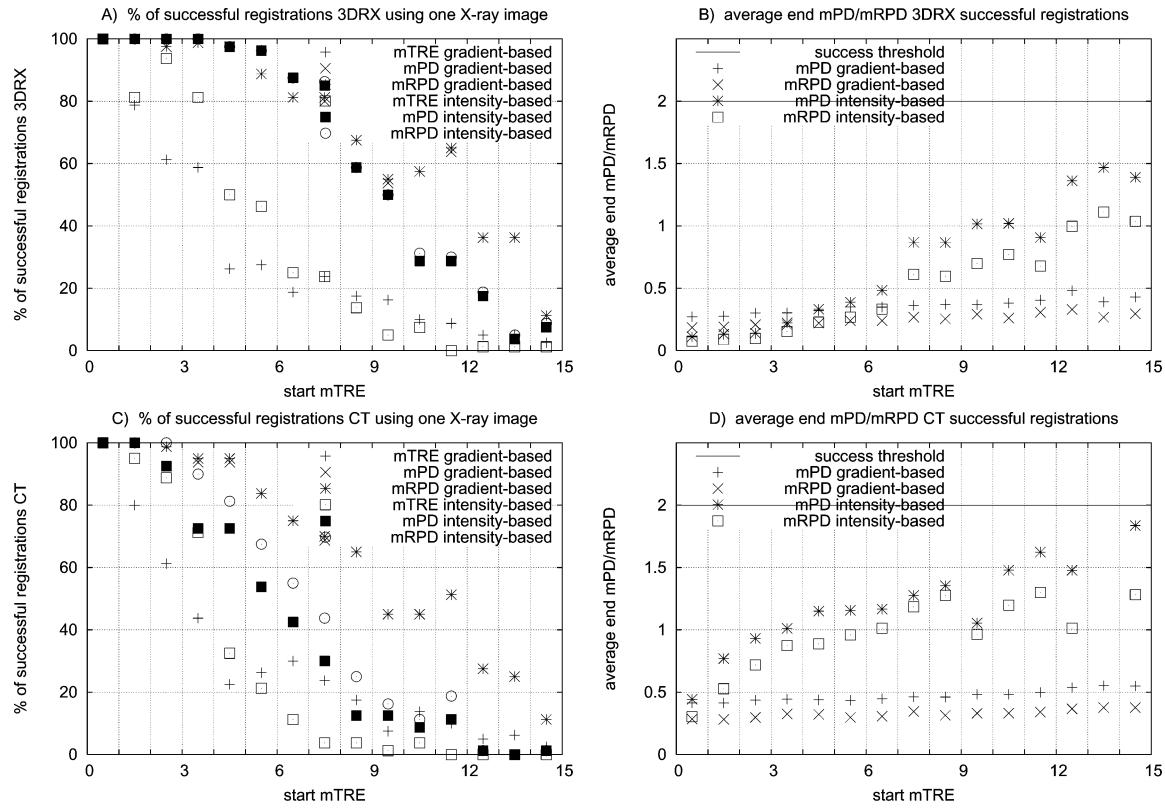


Fig. 9. Mean projection distance (mPD), mean reprojection distance (mRPD), and mean target registration error (mTRE) for registrations for 3DRX and CT based on a single X-ray image.

The capture ranges for 3DRX and CT, respectively, end at 6 mm and 3 mm for the gradient-based method and at 4 mm and 3 mm for the intensity-based method [Fig. 5(c),(g)]. Within the capture range the average end mTREs of successful registrations for 3DRX and CT, respectively, are 0.18 and 0.30 mm for the gradient-based method and 0.13 and 0.65 mm for the intensity-based method.

Better results for 3DRX to X-ray registration could be explained by greater similarity between the 2-D and 3-D modalities: both X-ray images and 3-D volume are produced using X-rays at the same energy (around 60 keV), whereas CT uses 80 keV, and the 3DRX volume is reconstructed using the X-ray projections. Another reason could be the difference between the gold standard for 3DRX and the gold standard for CT and MR, where the 3-D–3-D registration could have introduced additional errors.

Results for MR to two X-ray image registration are displayed in Fig. 6 for the gradient-based method. Since the gradient-based method is intended for more than one X-ray image, and since MR to X-ray registration is quite challenging, we decided to only perform MR to X-ray registration using two X-ray images. The capture range ends at 1 mm and within this range the average end mTRE of success is 0.45 mm. It must be noted that the gradient-based registration was not optimized for this MR protocol.

Results for the registrations based on one X-ray image are displayed in Fig. 7. The capture ranges for 3DRX and CT, respectively, end at 1 mm and 1 mm for the gradient-based method and 1 mm and 2 mm for the intensity-based method [Fig. 7(c), (g)]. Within the capture range the average end mTREs of successful registrations for 3DRX and CT, respectively, are 0.46 mm and 0.53 mm for the gradient-based method and 0.84 mm and 0.84 mm for the intensity-based method [Fig. 7(d), (h)]. The intensity-based method seems to perform better (higher capture range) on one X-ray image than the gradient-based method. The results for both methods are significantly worse than when using two X-ray images. The main reason is difficulty in estimating the correct position of the 3-D volume in the projection direction.

In order to assess whether the largest part of the error was in the direction of the normal to the projection plane, the end mTREproj was calculated as explained in Section II-B. In Fig. 8, the end mTRE in the projection direction is plotted as a function of the total end mTRE for registrations based on one X-ray image. From the graphs in Fig. 8 we can conclude that especially for small end mTREs (thus accurate registrations) the major part of the total end mTRE can be explained by the end mTRE in the projection direction only.

In order to quantify the in-plane error and the reprojection distance, the end mPD and end mRPD were also calculated for both registration methods for registrations using one X-ray image (as explained in Section II-B). The results are displayed in Fig. 9, where results for spinal segments 1 and 2 are combined and the end mTRE is shown as a reference. Only the averages per bin of the successful registrations were displayed to show the difference between the error measures more clearly. For the mPD, the 3DRX capture ranges for the gradient-based method and the intensity-based method, respectively, end at: 5 mm (0.29-mm average mPD for success) and 6 mm (0.22-mm average mPD for success), the CT capture ranges, respectively,

TABLE III
SUMMARY OF THE RESULTS OF ALL REGISTRATION EXPERIMENTS. "GB" REPRESENTS THE GRADIENT-BASED METHOD AND 'IB' THE INTENSITY-BASED METHOD, "ONE" AND "TWO" REPRESENT ONE X-RAY IMAGE AND TWO X-RAY IMAGES, RESPECTIVELY, AND "E" AND "C.R." ARE ABBREVIATIONS FOR ERROR AND CAPTURE RANGE

| | | | 3DRX | | CT | |
|---|---|---|---|---|---|---|
| | | | E (mm) | C.R. (mm) | E (mm) | C.R. (mm) |
| GB | two | mTRE | 0.18 | 6 | 0.30 | 3 |
| | | mTRE | 0.46 | 1 | 0.53 | 1 |
| GB | one | mPD | 0.29 | 5 | 0.42 | 3 |
| | | mRPD | 0.20 | 5 | 0.30 | 5 |
| IB | two | mTRE | 0.13 | 4 | 0.65 | 3 |
| | | mTRE | 0.84 | 1 | 0.84 | 2 |
| IB | one | mPD | 0.22 | 6 | 0.61 | 2 |
| | | mRPD | 0.15 | 6 | 0.52 | 3 |

end at: 3 mm (0.42-mm average mPD for success) and 2 mm (0.61-mm average mPD for success). For the mRPD, the 3DRX capture ranges for the gradient-based method and the intensity-based method, respectively, end at: 5 mm (0.20-mm average mRPD for success) and 6 mm (0.15-mm average mRPD for success), the CT capture ranges, respectively, end at: 5 mm (0.30-mm average mRPD for success) and 3 mm (0.52-mm average mRPD for success). The results show that the capture ranges are much larger when the error in the projection direction is not relevant. The most appropriate error measure to use depends on the application.

Table III summarizes the main results for all 3DRX and CT experiments. The main results for MR to two X-ray projections experiments for the gradient-based method were a capture range of 1 mm with an average error for successful registrations of 0.45 mm. The evaluation method reveals that for both methods the capture ranges are limited. This could be a problem for fully automated 2-D–3-D registration. Manual initialization may be required to be within this capture range at the start of registration. By superimposing bony edge points in CT or MR on X-ray images, an operator can find an initial registration. The achievable accuracy of manual initial registration should be investigated further. The registration method may also be used in situations where despite measures to limit movement, some motion has occurred. In a previous study [38], it was shown that in many interventions the intraoperative movement is limited, which suggests that in these situations no further initialization is required. Nevertheless, the limited capture range remains a potential problem.

It has been shown that the capture range depends strongly on the evaluation measure used, and this in turn depends on the application. When the objective is to find the 3-D position of an object, we recommend the use of more than one projection and evaluation using the mTRE. When the objective is to display 3-D information in 2-D, one projection is sufficient, and the performance can be adequately evaluated using the mPD or the mRPD. For such applications a larger capture range will be obtained.

The evaluation also reveals that MR to X-ray registration exhibited larger errors and a smaller capture range than 3DRX and CT-based registration. This seems plausible as it is much more difficult to relate the intensities of MR data and X-ray images. Many more graphs and conclusions can be drawn from the generated data; this paper, however, only shows the potential of the presented evaluation methodology.

TABLE IV
AVERAGE REGISTRATION ERRORS (MEAN MTRE IN MILLIMETERS) FOR SUCCESSFUL REGISTRATIONS WHEN EXTRA ERRORS ARE INTRODUCED TO THE CALIBRATED PROJECTION GEOMETRY. REGISTRATIONS WERE PERFORMED FOR 3DRX TO TWO X-RAY PROJECTIONS. THE ORIGINAL GOLD STANDARD ERROR WAS 0.11 MM

| | Extra calibration errors | | | | |
|---|---|---|---|---|---|
| | 0 | ang 0.5° | ang 1.0° | rot 0.5° | rot 1.0° |
| mean mTRE | 0.11 | 0.51 | 0.94 | 0.36 | 0.66 |

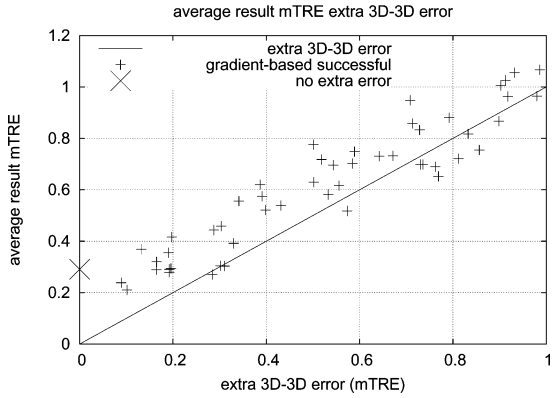


Fig. 10. Graph displaying the average result mTRE for successful registrations when an extra error is introduced in the 3-D–3-D registration. The registrations were performed for CT to two X-ray projections using the gradient-based method. The original gold standard error was 0.29 mm.

### *B. Experiment 2: Evaluation of the Gold Standard*

The obtained results depend on the gold standard which in turn depends on the calibrated geometry of the system and the 3-D–3-D registration (for MR and CT). As the reported mean errors are small (end mTRE) for successful registrations over a large number of experiments, there is high confidence in the accuracy of the gold standard. However, to investigate the influence of errors in the geometry calibration and 3-D–3-D registration on the evaluation, additional experiments were performed by determining the mTRE for a modified gold standard, as described in Section IV. The results of these experiments are displayed in Table IV and Fig. 10.

The error for successful registrations on vertebral body 1 of segment 2 was on average 0.11 mm when using our gold standard (with no additional errors) for registration of 3DRX to two X-ray projections. Extra geometry calibration errors increase the resulting mTREs for successful registrations significantly (e.g., a change of 0.5° angulation already increases the mTRE from 0.11 to 0.51 mm). Thus, it can be concluded that the calibration is very accurate, and that accurate calibration is essential.

The error for successful registrations on vertebral body 1 of segment 2 was on average 0.29 mm when using our gold standard (with no additional errors) for registration of CT to two X-ray images. As expected, the addition of an error in 3-D–3-D registration decreases the accuracy of the registration. It can be observed that even for small extra errors in 3-D–3-D registration, the calculated 2-D–3-D registration error significantly increases. For example, for an extra 3-D–3-D mTRE error of 0.50 mm the 2-D–3-D error increases from 0.29 to 0.78 mm. This indicates that the initial 3-D–3-D registration of CT to 3DRX was fairly accurate (less than a voxel). The extra error

does not have a significant effect on the capture range, since the maximum extra 3-D–3-D error is 1 mm, and so it can only decrease the capture range by a maximum of 1 mm.

In supplement to these experiments, we would like to make an additional remark. The error in the gold standard consists of projection calibration errors, distortion correction errors, and 3-D–3-D registration errors. Projection calibration and distortion calculation are performed separately. While 3-D–3-D registration is performed on whole segments, 2-D–3-D registration is performed on each vertebra separately. Moreover, the two evaluated 2-D–3-D registration methods use different optimization techniques, similarity criteria, and either ROIs or VOIs. Therefore, the errors in 2-D–3-D registration and the errors in the gold standard should be independent. In the case of independent errors, the total error is the root of the sum of squared individual errors. Together with the fact that the errors reported in Section V-A are small, these assumptions give us confidence that the gold standard is accurate.

## VI. Conclusion

A standardized evaluation methodology for objective inter-method and inter-modality comparison of 2-D–3-D registration methods has been introduced. Objective and systematic comparison is facilitated by providing a gold standard and using the same datasets, starting positions, regions of interest, centers of rotation, error measures, capture range and failure criteria. Also, by measuring the error in the starting position and the registration error in the same way, the effect of an algorithm can be clearly seen.

The calibration procedure used in generating the gold standard is very accurate and the 3-D–3-D registration necessary to obtain the gold standard for CT/MR only appears to introduce subvoxel errors. Our experiments showed that the total error (which consists of a calibration error and a 3-D–3-D registration error) is similar to the error in 3-D–3-D registration. Thus, accurate 3-D–3-D registration, such as used in this gold standard, is required.

We demonstrated the potential of the described evaluation methodology by comparing two 2-D–3-D registration methods. Not surprisingly, the results show that the performance of each method is significantly better when using two X-ray images, than when using only one X-ray image for finding the 3-D position of an object (quantified by the mTRE). Even for registrations based on two X-ray images the capture range of the algorithms is limited, but for successful registrations the attained accuracy is quite high (0.18 and 0.30 mm mTRE for 3DRX and CT, respectively, with the gradient-based method and 0.13 mm mTRE and 0.65 mm mTRE for 3DRX and CT, respectively, with the intensity-based method). For registrations based on one X-ray image, the capture range of the intensity-based method is larger than for the gradient-based method, but the gradient-based method has a larger capture range when using two X-ray images.

A possible extension of the evaluation method is to investigate how the capture range is influenced by different choices of the parameters for each method, and thereby it can help in parameter tuning. The effect of a different success threshold on the capture range can also be assessed, and another extension is to evaluate the performance of registration methods when using more than two X-ray images, since more X-ray images are available.

The 3DRX, CT, MR, and X-ray data, together with gold standard transformations, starting positions, ROIs, VOIs, centers of rotation, and evaluation criteria as published in this paper are available on the internet (http://www.isi.uu.nl/Research/Databases/). Since the described methodology uses a calibrated 3DRX system, extension to other anatomies which have high contrast on X-ray (e.g., bone) is straightforward and this data as well as patient data can be included in the database. In the field of rigid 3-D–3-D registration, the availability of the Vanderbilt dataset has already shown the importance of a common database, validation statistics, and error measure for comparison of multiple algorithms [33]. Similarly, our aim is to create a standardized dataset for 2-D–3-D registration which can be used for future evaluations by other researchers.

## Acknowledgment

The would like to thank Dr. R.L.A.W. Bleys of the Anatomy department at the University Medical Center Utrecht for making the spinal segments available. They would also like to thank W. van Wolferen, S. Plomp, and J. J. Verlaan for preparing the spinal segments for our research and W. Bartels, T. van Osch, and J. H. Seppenwoolde for helping with the MR acquisitions.

## References

[1] R. Bansal, L. Staib, Z. Chen, A. Rangarajan, J. Knisely, R. Nath, and J. S. Duncan, "A minimax entropy registration framework for patient setup verification in radiotherapy," *Comput. Aided Surg.*, vol. 4, pp. 287–304, 1999.

[2] Á Czopf, C. Brack, M. Roth, and A. Schweikard, "2D-3D registration of curved objects," *Periodica Polytechnica Ser. Elect. Eng.*, vol. 43, no. 1, pp. 19–41, 1999.

[3] J. Feldmar, N. Ayache, and F. Betting, "3D-2D projective registration of free-form curves and surfaces," *Comput. Vis. Image Understanding*, vol. 65, no. 3, pp. 403–424, 1997.

[4] A. Guéziec, P. Kazanzides, B. Williamson, and R. H. Taylor, "Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot," *IEEE Trans. Med. Imag.*, vol. 17, pp. 715–728, Oct. 1998.

[5] A. Hamadeh, S. Lavallee, and P. Cinquin, "Automated 3-dimensional computed tomographic and fluoroscopic image registration," *Comput. Aided Surg.*, vol. 3, no. 1, pp. 11–19, 1998.

[6] Y. Kita, D. L. Wilson, and J. A. Noble, "Real-time registration of 3D cerebral vessels to X-ray angiograms," in *Lecture Notes in Computer Science*, W. M. Wells, A. Colchester, and S. Delp, Eds. New York: Springer, 1998, vol. 1496, Medical Image Computing and Computer-Assisted Intervention (MICCAI 98), pp. 1125–1133.

[7] S. Lavallée and R. Szeliski, "Recovering the position and orientation of free-form objects from image contours using 3D distance maps," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 4, pp. 378–390, Apr. 1995.

[8] M. J. Murphy, "An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotaxic radiosurgery," *Med. Phys.*, vol. 24, no. 6, pp. 857–866, 1997.

[9] J. H. Hipwell, G. P. Penney, R. A. McLaughlin, K. Rhode, P. Summers, T. C. Cox, J. V. Byrne, J. A. Noble, and D. J. Hawkes, "Intensity-based 2-D–3-D registration of cerebral angiograms," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1417–1426, Nov. 2003.

[10] S. Jonić, P. E. Thévenaz, and M. Unser, "Multiresolution spline-based 3D/2D registration of CT volume and C-arm images for computer-assisted surgery," *Proc. SPIE (Medical Imaging 2001: Image Processing)*, vol. 4322, pp. 1101–1109, 2001.

[11] L. Lemieux, R. Jagoe, D. R. Fish, N. D. Kitchen, and D. G. T. Thomas, "A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs," *Med. Phys.*, vol. 21, no. 11, pp. 1729–1760, 1994.

[12] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2-D-3-D medical image registration," *IEEE Trans. Med. Imag.*, vol. 17, no. 4, pp. 586–595, Apr. 1998.

[13] G. P. Penney, P. G. Batchelor, D. L. G. Hill, D. J. Hawkes, and J. Weese, "Validation of a two- to three-dimensional registration algorithm for aligning preoperative CT images and intraoperative fluoroscopy images," *Med. Phys.*, vol. 28, no. 6, pp. 1024–1032, 2001.

[14] T. Rohlfing, D. B. Russakoff, M. J. Murphy, and C. R. Maurer Jr., "An intensity-based registration algorithm for probabilistic images and its application for 2-D to 3-D image registration," *Proc. SPIE (Medical Imaging 2002: Image Processing)*, vol. 4683, pp. 581–591, 2002.

[15] D. B. Russakoff, T. Rohlfing, A. Ho, D. H. Kim, R. Shahidi, J. R. Adler Jr., and C. R. Maurer Jr. *et al.*, "Evaluation of intensity-based 2D-3D spine image registration using clinical gold-standard data," in *Lecture Notes in Computer Science*, J. C. Gee *et al.*, Eds. Berlin, Germany: Springer-Verlag, 2003, vol. 2717, WBIR 2003, pp. 151–160.

[16] D. Sarrut and S. Clippe, "Patient positioning in radiotherapy by registration of 2D portal to 3D CT images by a content-based research with similarity measures," in *Computer Assisted Radiology and Surgery*, H. Lemke, M. Vannier, K. Inamura, A. G. Farman, and K. Doi, Eds. Amsterdam, The Netherlands: Elsevier Science B.V., 2000, pp. 707–712.

[17] J. Weese, G. P. Penney, P. Desmedt, T. M. Buzug, D. L. G. Hill, and D. J. Hawkes, "Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery," *IEEE Trans. Inf. Technol. Biomed.*, vol. 1, no. 4, pp. 284–293, Dec 1997.

[18] J. Weese, R. Göcke, G. P. Penney, P. Desmedt, T. M. Buzug, and H. Schumann, "Fast voxel-based 2D/3D registration algorithm using a volume rendering method based on the shear-warp factorization," *Proc. SPIE (Medical Imaging 1999: Image Processing*, vol. 3661, pp. 802–810, 1999.

[19] M. Vermandel, N. Betrouni, G. Palos, J. Y. Gauvrit, C. Vasseur, and J. Rousseau, "Registration, matching, and data fusion in 2D/3D medical imaging: Application to DSA and MRA," in *Lecture Notes in Computer Science*, R. Ellis and T. Peters, Eds: Springer-Verlag, 2003, vol. 2878, Medical Image Computing and Computer Assisted Intervention – MICCAI 2003, pp. 778–785.

[20] D. Tomaževič, B. Likar, T. Slivnik, and F. Pernuš, "3-D/2-D registration of CT and MR to X-ray images," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1407–1416, Nov. 2003.

[21] H. Livyatan, Z. Yaniv, and L. Joskowicz, "Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1395–1406, Nov. 2003.

[22] R. A. McLaughlin, J. Hipwell, D. J. Hawkes, J. A. Noble, J. V. Byrne, and T. Cox, "A comparison of 2D-3D intensity-based registration and feature-based registration for neurointerventions," in *Lecture Notes in Computer Science*, T. Dohi and R. Kikinis, Eds: Springer, 2002, pt. 2, vol. 2489, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2002, pp. 517–524.

[23] D. A. LaRose, J. Bayouth, and T. Kanade, "Evaluating 2D/3D registration accuracy," in *Computer Assisted Radiology and Surgery*, H. Lemke, M. Vannier, K. Inamura, A. Farman, and K. Doi, Eds. Amsterdam, The Netherlands: Elsevier Science B.V., 2000, pp. 147–152.

[24] D. Tomaževič, B. Likar, and F. Pernuš, "'Gold standard' 2D/3D registration of X-ray to CT and MR images," in *Lecture Notes in Computer Science*, T. Dohi and R. Kikinis, Eds. New York: Springer, 2002, pt. 2, vol. 2489, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2002, pp. 461–468.

[25] L. Brunie, S. Lavallée, J. Troccaz, P. Cinquin, and M. Bolla, "Pre- and intra-irradiation multimodal image registration: Principles and first experiments," *Radiotheraphy Oncol.*, vol. 29, no. 3, pp. 244–252, 1993.

[26] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[27] M. Grass, R. Koppe, E. Klotz, R. Proksa, M. Kuhn, H. Aerts, J. Op de Beek, and R. Kemkers, "Three-dimensional reconstruction of high contrast objects using C-arm image intensifier projection data," *Computerized Med. Imag. Graphics*, vol. 23, no. 6, pp. 311–321, 1999.

[28] R. Koppe, E. Klotz, J. Op de Beek, and H. Aerts, "3D vessel reconstruction based on rotational angiography," in *Computer Assisted Radiology*, H. Lemke, K. Inamura, C. Jaffe, and V. Vannier, Eds. New York: Springer, 1995, pp. 101–107.

[29] V. Rasche, B. Schreiber, C. Graeff, T. Istel, H. Schomberg, M. Grass, R. Koppe, E. Klotz, and G. Rose, "Performance of image intensifier-equipped X-ray systems for three-dimensional imaging," in *Computer Assisted Radiology and Surgery*. ser. Int. Congr., H. Lemke, M. Vannier, K. Inamura, A. Farman, K. Doi, and J. Reiber, Eds. Amsterdam, The Netherlands: Elsevier Science B.V., 2003, vol. 1256, pp. 187–192.

[30] B. Movassaghi, V. Rasche, M. Grass, M. A. Viergever, and W. J. Niessen, "A quantitative analysis of 3-D coronary modeling from two or more projection images," *IEEE Trans. Med. Imag.*, vol. 23, no. 12, pp. 1517–1531, Dec. 2004.

[31] S. A. M. Baert, G. P. Penney, T. van Walsum, and W. J. Niessen, "Precalibration versus 2D-3D registration for 3D guide wire display in endovascular interventions," in *Lecture Notes in Computer Science*, C. Barillot, D. Haynor, and P. Hellier, Eds. New York: Springer, 2004, pt. 2, vol. 3217, Medical Image Computing and Computer-Assisted Intervention-MICCAI 2004, pp. 577–584.

[32] P. Haaker, E. Klotz, R. Koppe, and R. Linde, "Real-time distortion correction of digital X-ray II/TV-systems: An application example for digital flashing tomosynthesis (DFTS)," *Int. J. Cardiac Imag.*, vol. 6, no. 1, pp. 39–45, 1990.

[33] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawkes, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, X. Pennec, M. E. Noz, G. Q. Maguire Jr., M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods, "Comparison and evaluation of retrospective intermodality brain image registration techniques," *J. Comput. Assist. Tomogr.*, vol. 21, no. 4, pp. 554–566, 1997.

[34] E. B. van de Kraats, T. van Walsum, J. J. Verlaan, F. C. Oner, M. A. Viergever, and W. J. Niessen, "Noninvasive magnetic resonance to three-dimensional rotational X-ray registration of vertebral bodies for image-guided spine surgery," *Spine*, vol. 29, no. 3, pp. 293–297, 2004.

[35] J. M. Fitzpatrick and J. B. West, "The distribution of target registration error in rigid-body point-based registration," *IEEE Trans. Med. Imag.*, vol. 20, no. 9, pp. 917–927, Sep. 2001.

[36] Y. Masutani, T. Dohi, F. Yamane, H. Iseki, and K. Takakura, "Interactive virtualized display system for intravascular neurosurgery," in *Lecture Notes in Computer Science*, J. Troccaz, W. Grimson, and R. Mösges, Eds: Springer, 1997, vol. 1205, CVRMed-MRCAS '97, pp. 427–435.

[37] Y. R. Rampersaud, D. A. Simon, and K. T. Foley, "Accuracy requirements for image-guided spinal pedicle screw placement," *Spine*, vol. 26, no. 4, pp. 352–359, 2001.

[38] J. V. Byrne, C. Colomina, J. Hipwell, T. Cox, J. A. Noble, G. P. Penney, and D. J. Hawkes, "An assessment of a technique for 2D-3D registration of cerebral intra-arterial angiography," *Br. J. Radiol.*, vol. 77, no. 914, pp. 123–128, 2004.