Deformable 2D-3D Medical Image Registration Using a Statistical Model: Accuracy Factor Assessment

Jianhua Yao\(^a\), Russel Taylor\(^b\)

\(^a\)Diagnostic Radiology Department, Clinical Center, the National Institute of Health, Bethesda, MD, USA
\(^b\)Computer Science Department, the Johns Hopkins University, Baltimore, MD, USA

Abstract – Deformable 2D-3D medical image registration is an essential technique in Computer Aided Surgery (CAS) to fuse 3D pre-operative data with 2D intra-operative data. Several factors may affect the accuracy of 2D-3D registration, including the number of 2D views, angle between views, view angle relative to anatomical objects, co-registration error between views, noise in the images, and image distortion. In this paper, we investigated the relationship between these factors and the accuracy of 2D-3D registration. We proposed a deformable 2D-3D registration method using a statistical model and conducted experiments on simulated x-ray images to assess the accuracy factors. Some discussions are provided on how to improve the accuracy of 2D-3D registration based on our assessment.

Key Words – deformable 2D-3D registration, accuracy assessment, statistical model

I. Introduction and Background

All submissions should follow the guidelines of this journal for submission. Medical imaging data in Computer Integrated Surgery (CIS) are mainly in three categories: 1) pre-operative data such as CT, MR and PET images, or 3D anatomical models; 2) intra-operative data such as X-rays, ultrasound, and endoscopes; and 3) post-operative data. Image registration is the process to transform different forms of data into one coordinate system, which is crucial for a wide variety of computer aided surgery applications. CT and MR images are frequently used in clinical diagnosis and surgical planning, but their use as interventional imaging modalities has been limited due to the space concern in the operating room and the real time requirement for most procedures. One of the common modalities for guiding surgical interventions is X-ray fluoroscopy. These images are acquired in real time, but only present 2D information. Important anatomical features, such as relative 3D location and orientation of anatomical landmarks, cannot be positively determined. One method to provide 3D information during the intervention is to register and fuse pre-operative 3D images/models with intra-operative fluoroscopic images. Basically, one must find the 2D-3D projective transformation that maps a 3D object onto one or more 2D projective images of the same object. This problem is called 2D-3D registration.

The easiest way to solve 2D-3D registration problem is to place artificial fiducial markers visible in both 2D and 3D images. However, this approach introduces many restrictions on the surgical procedures. 3D pre-operative images should often be acquired in temporal proximity to intra-operative 2D images, and the position of markers should be fixed. Furthermore, the placement of fiducial markers is often invasive. Most recent investigations have turned to non-invasive methods and used the image content to determine the 2D-3D transformation. Non-invasive 2D-3D registration techniques can be classified into two main categories: landmark-based and intensity-based [1]. Landmark-based methods register on salient anatomical features extracted from both images. Feature extraction is difficult to automate, and its error can lead to errors in the successive registration. Intensity-based methods compare the voxel values using a similarity measure based on image intensity distribution. Intensity-based methods are typically slower than landmark-based methods, but require little or no segmentation.

image segmentation and 2D-3D registration. Weese et al. [7] presented a pattern intensity method for 2D-3D registration. Lemieux et al. [8] proposed a method for frameless stereotactic brain surgery using a 2D-3D registration technique based on image intensity and used fiducial markers for validation. Zollei et al. [9] proposed a mutual information based 2D-3D registration algorithm which establishes the proper alignment via a stochastic gradient ascent strategy. Fleute et al. [10] proposed a deformable 2D-3D registration technique based on a statistical surface model. Dey et al. [11] presented a Monte Carlo-based hybrid algorithm to increase the robustness and success rate of 2D-3D registration. Novosad et al. [12] demonstrated that 3D reconstruction of vertebra can be achieved through 2D-3D registration using a single X-Ray image and prior models computed from preoperative CT. Tomazevic et al. [13] conducted experiments on 2D-3D registration of X-Rays with CT, 3DRX and MR. They evaluated the capture range, success rate and required number of X-Rays in each modality. Penney et al. [14] applied 2D-3D registration technique to accurately measure the position and orientation of an acetabular cup implant from postoperative X-rays. Several investigations [15] [16] [17] had been conducted for real-time X-ray/CT registration. LaRose [15] precomputed projection values and stored them in a “Transgraph”. Russakoff [16] used pre-computed attenuation fields to accelerate the generation of 2D projections, and Rohlfing [17] further employed progressive attenuation fields to avoid pre-computation. More recently, to set a common ground for researchers in this field, Crum et al. [18] proposed a generalized overlap measure for evaluation of 2D-3D registration, and Kraats et al. [19] created a publicly available database of X-ray and CT of vertebra and gold standard transformation for researchers to evaluate their 2D-3D registration techniques.

Most 2D-3D registration works have focused on rigid 2D-3D registration of a single-subject image or model, i.e., the goal is to determine a rigid transformation between a coordinate system associated with a set of projection images and another coordinate system associated with a 3D volumetric scan or anatomical model of the same patient. Anatomical changes over time or deformations caused by patient motion can adversely affect the accuracy of such methods. A more fundamental limitation is that rigid registration would be invalid if a patient-specific scan or model is not available. The work reported in this paper focuses on the problem of deformable 2D-3D registration of a set of X-ray projection images with generic anatomical models incorporating statistical information of anatomical variation within a patient population. Deformable 2D-3D registration not only transforms the model in space, but also morphs the model or warps the image to obtain optimal matches.

Applications of deformable 2D-3D registration include creation of patient-specific 3D models from X-rays for the purpose of surgical planning, postoperative analysis, or intraoperative registration in the presence of (predictable) deformations. In a typical orthopedic surgery, a patient CT image is usually not available [20-22]. By registering a statistical model to a set of X-ray images, a patient specific model can be instantiated. The model can then be cross-sectioned to create a virtual patient CT. Another important application is to provide complement data for incomplete X-ray images. In real-life surgical scenario, X-ray images are often incomplete due to occlusion of surgical instrument and limited or partial field of view. Although multiple X-ray images may be acquired from different angles, it is difficult to fuse the information. If an anatomical model is adapted to the X-ray images, the registered model can then be projected to the 2D X-ray planes to fill in the missing information.

Accuracy is essential in 2D-3D registration in order to correctly transfer pre-operative knowledge into surgical procedures or fuse information from 2D and 3D images. Compared to 3D-3D registration between two 3D images, 2D-3D registration between a 3D image and a set of 2D images has fewer correspondence features and more parameters to optimize. There are several factors that may affect the accuracy of 2D-3D registration, including the number of 2D views, angles between 2D views, view angles relative to the anatomy, co-registration between 2D views, noise in the image, and image distortion. Very few investigations have been conducted to assess how these factors may affect the accuracy of 2D-3D registration. Assessment of these accuracy factors could provide valuable information to help researchers improve their surgical setup and protocol.

In our investigation, we first built a statistical pelvis model from a collection of training CT images, and generated simulated x-ray images (also known as Digitally Reconstructed Radiographs, DRRs) from a patient CT data set. We then performed a deformable 2D-3D registration between the pelvis model and the DRRs. We manipulated the parameters in the DRR generation to control the accuracy factors, and compared the registered model with a reference model to assess the accuracy factors. This paper is the
expansion of a previously published conference paper [23]. More experiments and analysis have been included in this paper.

II. Statistical pelvis model and its construction

We proposed a unique model representation to characterize both boundary surface and internal density distribution of bone structures. The model is represented as a tetrahedral mesh equipped with embedded Bernstein polynomial density functions on the barycentric coordinates of each tetrahedron. Multiple level-of-details (LOD) of the anatomical structure are characterized by a hierarchical representation. Prior information of both shape and density properties is incorporated in the model.

We developed an efficient “tetrahedral mesh reconstruction from contours” method to construct tetrahedral meshes for bone structures from a CT image. The method produces tetrahedral meshes with high flexibility and is capable to accommodate any anatomical shape. The meshes are built from contours of cortical bone boundaries. The contours are extracted from the image using an active contour technique [24]. They are then tiled into a tetrahedral mesh by solving a series of tiling, correspondingly and branching problems [25]. An analytical density function is assigned to every tetrahedron to minimize the residual errors in the density distribution. Currently, the density functions are defined as n-degree Bernstein polynomials [26] in the barycentric coordinate of a tetrahedron:

\[ D(\mu) = \sum_{i,j,k,l} C_{i,j,k,l} B^n_{i,j,k,l}(\mu), \]

where \( C_{i,j,k,l} \) is the polynomial coefficient, and \( \mu = (\mu_x, \mu_y, \mu_z, \mu_w) \).

\[ B^n_{i,j,k,l}(\mu) = \frac{n!}{i!j!k!l!} \mu_i^i \mu_j^j \mu_k^k \mu_l^l \]

is a barycentric Bernstein basis function, and \( (\mu_x, \mu_y, \mu_z, \mu_w) \) are the barycentric coordinates, with \( \mu_x + \mu_y + \mu_z + \mu_w = 1 \).

The advantages of such a representation are: 1) it is an explicit form; and 2) it is a continuous function in 3D space. Therefore, it is convenient to integrate, differentiate and interpolate. We also developed a tetrahedral mesh simplification algorithm based on edge collapsing operations [27] to build a multiple level-of-detail (LOD) model representation.

We designed a training scheme to compute a statistical model from a collection of training models. A model aligning procedure is first performed to map all training models into a common mesh topological structure. The procedure proceeds by first aligning the exterior surfaces, followed by morphing the internal structures. Then a Principal Component Analysis (PCA) is applied to compute the variability of both shape and density properties of the anatomical structure. The model can then be approximated by a set of statistical mode parameters \( \{d_i\} \):

\[ Y = \bar{Y} + Pd \]

where \( Y \) is a model instance, \( \bar{Y} \) is the average model representation, and \( P \) is the eigenvector matrix incorporating the prior information.

We have built a statistical density model of hemipelvis from eight training images. Table 1 lists the eigenvalues of the deformation modes. The first five modes comprise about 93% of the sum of all eigenvalues, which means they characterize about 93% of all variability. Figure 1a shows the exterior surface of a multiple-resolution hemipelvis model. Figure 1b shows the shape variation on mode 1, which controls the size in both horizontal and vertical directions. Applying +3 standard deviation on the first mode makes the model “taller” and “thinner”, while -3 standard deviation makes it “shorter” and “wider”. Figure 1c shows the density variation on mode 1 by projecting the model to a 2D plane using a ray casting technique. Applying +3 standard deviation on the first mode makes the model darker around the acetabular area, while -3 standard deviation makes the model brighter. A new model instance can be instantiated from the average model by adjusting the statistical mode parameters. Detailed description of our model and its reconstruction can be found in [28, 29].

III. Deformable 2D-3D registration scheme

Figure 2(a) illustrates the deformable 2D-3D registration. Given a set of 2D X-ray images and an anatomical model, transformation \( T \) and deformation \( D \) of the model need to be optimized so that the projections of the model align with the X-ray images. The registration can be treated as determining a set of transformation and projection parameters \( (t_x, t_y, t_z, r_x, r_y, r_z, s_x, s_y, s_z, \{d_i\}, c_x, c_y, p_x, p_y, f) \). The first nine parameters define an affine transformation, including translation \( (t_x, t_y, t_z) \), rotation angle \( (r_x, r_y, r_z) \), and scale \( (s_x, s_y, s_z) \). \( \{d_i\} \) is the deformation parameter set. The
last five parameters describe the perspective projection geometry. The coordinate \((c_x, c_y)\) is the projection center; \((p_x, p_y)\) is the pixel size; and \(f\) is the focal length. \((c_x, c_y, p_x, p_y, f)\) are intrinsic camera parameters which can be determined in a separate calibration stage before the registration [30].

The overall deformable 2D-3D registration scheme between a statistical model and a set of fluoroscopic X-ray images is outlined as follows:
1: The intrinsic parameters \((c_x, c_y, p_x, p_y, f)\) of the fluoroscope device are calibrated before the registration.
2: A set of fluoroscopic X-ray images \(\{I_f\}\) are acquired and the inherent geometric distortion is corrected.
3: The X-ray image planes are then co-registered to determine their relative poses with respect to a reference coordinate system (the first X-ray plane).
4: Digitally reconstructed radiographs (DRRs) of the statistical model \(M\) are generated on the co-registered X-ray planes, and a similarity measure between DRRs and fluoroscopic X-ray images is evaluated. The DRRs and X-ray images are normalized to 256 gray scales and 512*512 in dimension.
5: A non-linear optimization algorithm is employed to maximize the similarity measure between DRRs and X-ray images, and optimal transformation parameters for model \(M\) are obtained. We used Powell’s method [32] as our optimization algorithm. To further improve the efficiency and robustness of the registration, the process is run in a multiple-resolution framework. This involves first searching for the solution in coarser X-ray images and with lower LOD models, and then refining the solution in a series of finer resolution images and models. The algorithm is also implemented in a multiple-step-size manner, in which it starts with a large step size and gradually reduces the step size as it gets closer to the optimal solution.
6: Repeat step 4 and step 5 until the difference of similarity measures between two iterations is below a small threshold or maximum number of iterations are performed.

This scheme involves several key techniques. Among those, fluoroscopic image distortion correction and image co-registration were described in a previous paper [31]. Efficient DRR generations from the statistical model can be found in [29]. Several parameters are used to control this algorithm, which are listed in Table 2. We have tested different configurations of the control parameters. The values are selected empirically by consideration of both

<table>
<thead>
<tr>
<th>Control parameter</th>
<th>Default value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of resolution levels</td>
<td>3</td>
</tr>
<tr>
<td>Number of step sizes</td>
<td>5</td>
</tr>
<tr>
<td>Number of statistical modes</td>
<td>5</td>
</tr>
<tr>
<td>Maximum iterations</td>
<td>500</td>
</tr>
<tr>
<td>Minimum difference between iteration</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 2. Control parameters in the algorithm
accuracy and efficiency. Figure 2 illustrates the deformable 2D-3D registration between the statistical model and three X-ray images.

IV. Experiments and Results

This section presents experiments and results in assessing the accuracy factors in our proposed deformable 2D-3D registration. The intrinsic “camera” parameters of the fluoroscope were calibrated independently before the registration and achieved high accuracy [30], their impacts on registration were not evaluated in our investigation. The view configuration of 2D images, such as the number of views, angle between views, and view angle relative to the anatomical structure, were the main subjects. The artifacts in image acquisition and processing, including co-registration error of 2D views, image noise, and image distortion, were also investigated. These factors are referred as the “accuracy factors” of 2D-3D registration in this paper. Our objective was to investigate whether and to what extent these factors may affect the accuracy of 2D-3D registration. Since it was difficult to control the accuracy factor in real X-ray images, we used simulated X-rays (DRRs) in the experiments. By carefully manipulating the parameters in DRR generation, we can control the effects of accuracy factors and therefore obtain the assessment of their impact on the accuracy of 2D-3D registration. When one factor was evaluated, other factors were kept constant.

4.1 Experiments and Validation Metrics

We designed two experiments to validate our method and assess the accuracy factors. Experiment 1 used DRRs instantiated from a model instance; and experiment 2 used DRRs generated from a CT image.

In Experiment 1, a model instance \( M_s \) was instantiated by applying a given affine transformation and model deformation to the statistical model. \( M_s \) was then used to produce a set of DRRs. A deformable 2D/3D registration was carried out between the statistical model and those DRRs. A patient specific model \( M_s \) was computed as the outcome of the registration. Ideally, model \( M_s \) should be identical to model \( M_s \). Since all inputs in this experiment were simulated and controlled, this experiment can be used to test the basic methodology, although it did not verify the ability of the method to handle un-modeled deformation. It did, however, allow us to test the ability of the method to reconstruct 3D models from a limited number of 2D projections.

In Experiment 2, first a patient CT image other than those in the training set was selected. The CT image was cropped so that only the structures of interest were presented. A set of DRRs was generated from the CT image to simulate X-ray images. Then a deformable 2D-3D registration was conducted between the statistical model and DRRs to obtain a registered model \( M_g \). A manually segmented patient specific model \( M_s \) from the CT image was used as the reference model. The accuracy of the 2D-3D registration was validated by comparing model \( M_s \) with reference model \( M_g \).

In both Experiment 1 and 2, we assumed that the intrinsic camera parameters (focal length, image center, and pixel size) were perfectly calibrated. Experiment 1 was used to assess the basic algorithmic approach and compare similarity measures, and Experiment 2 was employed to assess the robustness of our method to image distortion, X-ray projection angles, and other accuracy factors.

We devised two metrics to evaluate the registration result. One metric is the volume overlap percentage. The model is scanned along X, Y, Z axes to produce a set of isotropic voxels within the model. The voxel size is 0.5mm x 0.5mm x 0.5mm. The volume overlap is computed as the percentage of overlapping voxels to the total number of voxels

\[
Overlap = \frac{\| V_s \cap V_g \|}{\| V_g \|} \times 100\% \quad (4)
\]

where \( V_s \) is the set of voxels in model \( M_s \), \( V_g \) is the set of voxels in model \( M_g \), and \( \| \) represents the size of a set.

The second metric is the distance between the exterior surfaces

\[
MeanD = \text{mean}_{\mu_i} (S(\mu_i, p(\mu_i))),\nMaxD = \max_{\mu_i} (S(\mu_i, p(\mu_i))) \quad (5)
\]

where \( \mu_i \) is one vertex on the exterior surface of \( V_s \), and \( p(\mu_i) \) is closest point of \( \mu_i \) on the exterior surface of \( V_g \), \( S(\mu_i, \mu_j) \) is the distance between two points. \( MeanD \) measures the mean distance, \( MaxD \) measures the maximum distance.

4.2 Evaluation on Transformation and Similarity Measures

We used the setup of Experiment 1 to evaluate the capability of our method in restoring transformations
between 2D and 3D images. We also evaluated similarity measures in 2D-3D registration, including normalized image subtraction, normalized cross correlation, entropy of difference image, mutual information, gradient correlation, pattern intensity, and gradient difference. The details of similarity measures can be found in the review paper written by Penney et al. [33].

The transformation to instantiate model instance included three translation parameters \((t_x, t_y, t_z)\), three rotation angles \((r_x, r_y, r_z)\), three scale factors \((s_x, s_y, s_z)\), and five deformation parameters \((d_1, d_2, d_3, d_4, d_5)\). These parameters built up a 14-dimensional parameter space. The value of deformation parameters was normalized to the range from −1 to +1, where −1 implied −3 standard deviations and +1 implied +3 standard deviations. We tested 15 configurations of transformations, which were listed in Table 3. Case 1 and 2 were translation only, case 3 and 4 were rotation only, case 5 and 6 were scale only, case 7 and 8 were deformation only, and case 9 to 15 were combinations of affine transformation and deformation. Three DRR views were used: first one was the A/P view, second one was the lateral view, and third one was in the middle of the first two views (135° apart in axial rotation) with 10° tilts toward the top of the pelvis. We tested all similarity measures for all 15 transformation configurations and reported registration accuracy and several statistics in various scenarios of transformations (Table 4). We then used experiment 2 to evaluate the similarity measures and reported the volume overlap and mean surface distance in Table 5. Results showed that “Mutual Information” and “Cross Correlation” produced the best registration accuracy in experiment 1, while “Mutual Information” achieved best result in experiment 2. Mutual Information had been employed in a few successful recent 2D-3D registration systems [13] [34] [11] [17] [16]. We used mutual information in all following experiments.

We further conducted the same experiment with many more cases using “mutual information” similarity measure. We chose 4 translation sets, which were \((0, 0, 0)\), \((20, 20, 20)\), \((-10, -5, 15)\), and \((15, 0, -5)\); 4 rotation angle sets, which were \((0, 0, 0)\), \((10, 10, 10)\), \((-5, -5, 5)\), and \((5, 0, -5)\); 3 scale sets, which were \((1, 1, 1)\), \((1.05, 1.05, 1.05)\), and \((0.95, 0.95, 0.95)\); and 3 deformation sets, which were \((0, 0, 0, 0, 0)\), \((0.4, 0.4, 0.3, 0.3, 0.2)\), and \((-0.4, -0.4, -0.3, -0.3, -0.2)\). The complete combination of translation, rotation, scales and deformation sets generated a total of \(4^4 * 3^3 * 3 = 144\) transformations. Furthermore, we randomly generated 50 transformations. We evaluated the capability of our method to restore the transformation by computing the difference between the registered transformation and the transformation used to generate the simulation X-ray images (which was the target transformation). The difference between two transformations was reported as a percentage ratio, written as:

\[
\Delta P = \frac{p^{(c)} - p^{(i)}}{|p^{(c)} - p^{(i)}|} \times 100\%
\]

(6)

Here, \(p^{(i)}\) is the initial transformation, \(p^{(c)}\) is the registered transformation, and \(p^{(t)}\) is the target transformation. Since the parameters had different units, the translation, rotation, scale and deformation parameter difference \((\Delta T, \Delta R, \Delta S, \text{ and } \Delta D)\) were computed separately, written as:

\[
\Delta T = \frac{t^{(t)} - t^{(c)}}{|t^{(t)} - t^{(c)}|}
\]

\[
\Delta R = \frac{r^{(t)} - r^{(c)}}{|r^{(t)} - r^{(c)}|}
\]

\[
\Delta S = \frac{s^{(t)} - s^{(c)}}{|s^{(t)} - s^{(c)}|}
\]

\[
\Delta D = \frac{d^{(t)} - d^{(c)}}{|d^{(t)} - d^{(c)}|}
\]

(7)

Here, \(t^{(i)} = (t_x^{(i)}, t_y^{(i)}, t_z^{(i)})\) is the initial translation, \(t^{(c)} = (t_x^{(c)}, t_y^{(c)}, t_z^{(c)})\) is the translation at the end of registration, and \(t^{(t)} = (t_x^{(t)}, t_y^{(t)}, t_z^{(t)})\) is the target translation. Similarly, we can define other parameter differences in Eq. 7. The initial transformation is \((0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)\). Table 6 reports the registration accuracy (volume overlap and mean surface distance), as well as the transformation restoration percentage. It also reports the statistics (average, standard deviation, and maximum (worst case)) from all 194 experiment cases. From the result, translation parameters were restored most accurately (with 3.94% difference), while deformation parameters were restored least accurately (with 21.55% difference).

4.3 Number of 2D Views

We used Experiment 2 setup with “Mutual Information” as the similarity measure to assess the accuracy factors (Sections 4.3-4.8).

We first investigated the relationship between the number of X-ray image views and the accuracy of 2D-3D registration. If the number of views is large enough
(e.g., greater than 50), the problem should be considered as a reconstruction problem rather than a registration problem. Table 7 shows the registration accuracy and the running time of using from one 2D view to six views. The results demonstrated that as the number of views increases, the registration accuracy improved, and the running time also increased proportionally. The improvement was significant from using one view to using two views, but was relatively small when more than three views were used. The results also suggested that 2D-3D deformable registration was impracticable with a single view. One possible reason may be that the scale parameter and translation along the axis perpendicular to the viewing plane are ambiguous from one projection, so is the rotation around the axis parallel to the viewing plane (the rotation in/out of the plane).

4.4 Angles between Views

Figure 3(a) illustrates the experimental setup to assess the impact of the angle between X-ray views on 2D-3D registration. The CT data set of a pelvis was placed in the center of two view planes. One view plane was fixed and the other view plane was rotated around the pelvis from 0° to 180° with 10° increments. At each angle, two DRR images were generated on the two view planes. Figure 3(b) plots the registration accuracy (Overlap and MeanD) versus the angle. As one would expect, the result showed that two X-ray views can produce best registration results when they were approximately orthogonal to each other. It produced the least accurate result at angle 0° since it was equivalent to using just one view.

4.5 View Angles Relative to the Anatomical Structure

We also investigated the impact of the view angles of X-ray image planes relative to the anatomical structure. As in Figure 4(a), a pelvis CT data set was put in the center of two view planes orthogonal to each other. The two view planes were rotated together around the pelvis from 0° to 180° with 10° increments. 0° was the angle where one view was the AP view and the other view is the lateral view. Figure 4(b) plots the registration accuracy (Overlap and MeanD) versus the angle. The overlap percentages over different angles fluctuated around 86%. The result showed that the view angle relative to the anatomical structure did not apparently affect the registration accuracy if two views orthogonal to each other were used.

4.6 Image Noise
In the formation of an X-ray image, noise is inevitably presented. To assess the effect of image noise on registration accuracy, 2D Gaussian white noise was intentionally added to the intensity values of DRR images. The statistical model was then registered with the DRRs contaminated with noise. Different noise magnitudes (standard deviation of the Gaussian operator) were tested, and the relationship between the noise magnitude (in the units of grayscale) and the registration accuracy was plotted in Figure 5. Three views were used in this experiment. Figure 5 showed that 2D-3D registration was insensitive to the noise when the noise magnitude was below 10 grayscale. However, the accuracy decreased dramatically for noise larger than 20 grayscale magnitude.

4.7 Image distortion

It is well known that fluoroscopic X-ray images have inherent spatial distortion due to the pin cushion effect of image intensifier and the interference of earth magnetic field. Several techniques have been developed to correct the spatial distortion in X-ray images, but small residual errors still remain after the correction [31]. We investigated the effect of spatial distortion in X-ray images by imposing spatial distortion to DRR images. The way to add spatial distortion was as follows: first a distortion map was obtained using a checkerboard based distortion correction algorithm [31] on a pre-acquired X-Ray image, and the distortion map was normalized to the range of \(-1…1\); then a new distortion map was computed by multiplying with a distortion magnitude (in the unit of pixels); finally the new distortion map was applied on the DRRs to generate distorted DRRs. Figure 6 plots the relationship between the magnitude of image distortion (in the unit of pixels) and the registration accuracy. The results showed that 2D-3D registration was sensitive to image distortion. Distortion of 10-pixelsize magnitude would deteriorate the registration dramatically.

4.8 Co-registration error of 2D projections

The X-ray images on different view planes need to be co-registered before employed in 2D-3D registration. The co-registration is usually conducted using magnetic or optical tracking devices, or calibration objects [20]. We conducted experiments to assess the sensitivity of 2D-3D registration to the residual error in 2D image co-registration. Three views were used in our experiment. Perturbations were added...
to the view matrices in DRR generation to simulate the errors in image co-registration. The perturbation ranged from 0° to 20°. DRRs were then generated using the contaminated view matrices and registered with the statistical model. Figure 7 plots the relationship between the perturbation angles and the registration accuracy. The results showed that 2D-3D registration was sensitive to the error in 2D image co-registration. A perturbation of 5° worsened the registration accuracy by about 12%.

V. Discussion

We have proposed a deformable 2D-3D registration technique using a statistical density model and validated our method through a series of experiments. We were able to achieve about 88% volume overlap and 0.93mm mean surface distance using simulated images from pelvis CT images. We also assessed several accuracy factors that might affect the accuracy of 2D-3D registration. We have identified several accuracy factors and assessed their relationship with 2D-3D image registration. The results of this investigation can help researchers better understand the 2D-3D registration process and improve the surgical setup and imaging protocol to achieve more accurate registration.

The more number of x-ray views are used, the more accurate the registration can be. From our results, three or four views are generally sufficient. As expected, there is a big gain in going from one view to two views, with diminishing returns after that. Several recent investigations [12] [11] [34] [35] reported that one-view could generate reasonably good 2D/3D registration. They were using either landmark fiducials [12] or tracking devices [35]. Some of them had relatively small capture range [34] or big registration error [11]. More importantly, all of them can only handle rigid transformation. [13] conducted similar experiments to investigate the effect of number of X-Rays on the registration. Their results also demonstrated a big improvement from one image to two images and from two images to three images, and the improvement diminished after that. Also as expected, the experiments show that if only two views are available, the best results are achieved when they are orthogonal. Since this is not always feasible, it is reassuring that the results are not greatly degraded so long as the viewing angle is between 75 and 105 degrees (see Figure 3). The reason may be that the orthogonal views can most reliably compensate each other in recovering the content from 3D objects. The view angle relative to the anatomical structures has relatively little impact on the registration. Since the angles to acquire X-ray images in the operating room should accommodate with the patient’s position and other surgical instruments, sometimes there are not many feasible view configurations. The effect of the presence of other anatomical structures or surgical instruments within the X-ray field of view, and their partial occlusion of the concerned anatomy, is yet to be investigated.

The insensitivity to small image noise in the 2D-3D registration may be explained with the “mutual information” similarity measure. “Mutual information” utilizes the statistics of intensity distribution and smoothes out some noise. However excessive noise (more than 10% of the pixel intensity) will degrade the registration accuracy. Our study also shows that 2D-3D registration is relatively sensitive to the distortion in X-ray images. Fortunately, current distortion correction techniques can achieve about one-pixel accuracy [31], and new fluoroscope devices can generate X-ray images without spatial distortion. The accuracy in 2D X-ray co-registration is also a critical factor in 2D-3D registration. A 5-degree error in co-registration will introduce significant errors in the following 2D-3D registration. The co-registration error can be minimized if a biplanar device is used. It is mentioned in literatures that angiography suites [36] and tracker-based systems [20] can typically produce pose estimation errors on order of about 1 degree or better. In real-life X-ray images, image noise, spatial distortion and co-registration error are likely to appear at the same time. The combinatorial effect of these factors is expected to worsen the registration accuracy. However, the effect is more likely to be added up than be combined. Image noise is in the domain of image intensity. Spatial distortion is in the domain of space. And the co-registration error is a systematic bias of the projection matrix. The results in this paper demonstrate an insignificant degradation of performance over a reasonable range of perturbation in co-registration and image distortion. Therefore, it is highly possible to achieve acceptable 2D-3D registration if all the accuracy factors are controlled under a reasonable limit.

We used volume overlap percentage as one measure to evaluate the registration accuracy. Our overlap percentage is closely related to the generalized overlap measures proposed in [18]. It is also closely related to the extensively used Dice and Tanimoto coefficients. In fact, it is identical to Dice coefficient if the volumes of ground truth model and registered model are the
same. Target registration error (TRE) is another commonly used error measure in many 2D-3D registration systems [13] [34] [19] [16]. However, TRE is only valid for rigid transformation, where the targets do not move during the registration. In deformable registration, since the model changes shape, no fixed targets can be used to evaluate the error. In our experiment, we also employed mean surface distance as one error measurement.

It must be noted that DRRs and real X-ray images could be very different. Penney et al. [1] argued that the differences are due to different modalities and changes in the imaged objects. Moreover, several error sources may be combined in real X-ray images. We expect that the registration result would be worse when real X-ray images are used. The similarity measure may differ when real X-ray images are used or it is applied to different applications. The projected value is a function of the projection angle and CT value, which is a nonlinear function with respect to the CT value. Thus, DRRs from the statistical model and simulated X-ray images from the CT have non-linear intensity relation when they are projected from different view angles. This was one of the reasons that we chose “mutual information” as the similarity measure in our experiment since it does not rely on linear intensity relation. “Mutual information” had been proven to be a robust measure in many recent 2D-3D registration applications [13] [34] [11] [17] [16]. “Cross correlation” also achieved high accuracy in our experiment 1. This may thank to the identity relationship map when we used DRR as 2D images. Once real X-ray images are applied, “cross correlation” may not achieve as promising results. The discovered trend of accuracy factors, such as number of x-ray views, image noise etc, should still stand when real X-ray images are applied. More rigorous validation using clinical data are desirable to verify our findings.

The technique presented in this paper is based on a statistical model created from a population of patients. The advantage is that the model can be built beforehand, and no patient specific CT image is required in the registration and no segmentation on preoperative CT is necessary. Furthermore, a patient specific model can be obtained as the result of the registration. One problem with this technique is that un-modeled variability cannot be restored. Therefore, the registered patient specific model may not perfectly reflect the anatomic structures. This has been demonstrated in the results of experiment 1 and experiment 2 (94.5% vs. 87.9% volume overlap). In experiment 1, all the variability presented in the simulated DRRs is incorporated in the model, while in experiment 2 the testing CT image is not in the training set and thus includes variability that cannot be extracted from the statistical model. One way to increase the variability of the model is to incorporate more training images. However, no matter how large the training set is, the model still cannot characterize all the variability especially for complex shapes like human anatomy. Another way to solve this problem is to add a local refinement step after the statistical deformation to address the discrepancy due to the lack of variability. The difficulty lies in the determination of local correspondence in the 2D projection images. Increasing the variability of the model will improve registration accuracy in general. However, it should not change the trend in the relationship between the accuracy factors and 2D-3D registration. Our model has three level-of-details (LOD), which is a scheme to improve the efficiency of the algorithm and avoid local minimal. We believe that three LOD is sufficient for most 2D-3D applications. In our investigation, we focused on the factors related to 2D images. We haven’t explored the factors related to 3D images/models, such as image resolution and statistical variability of the model. Since 3D images/models are acquired or built before the intervention, their quality can be better controlled and guaranteed. Nevertheless, investigating how the variability and resolution of statistical model might affect the accuracy factors in 2D-3D registration is worthwhile in future researches.

Clinical applications of deformable 2D-3D registration include surgical planning, surgical guidance and postoperative evaluation. Gueziec et al. [37] suggested that 1.5mm accuracy was tolerable in orthopedic surgeries such as total hip replacement. Most reported 2D-3D techniques [8-10, 15] [19] [17] [13] achieved this requirement. However, most current techniques were tested under tightly controlled environment or required intensive human interaction. In real-life surgical scenario, the problem could be much more complex. In order to achieve acceptable registration accuracy, first of all, the accuracy factors assessed in this paper need to be considered. Other than these factors, there are more issues need to be addressed in the operating room. For example, the number of available views and the angle of views are often limited by the operating room circumstance, surgical protocol and patient position. X-ray images may suffer from occlusions from surgical instruments and partial field of view. Patients may move during the image acquisition. Strict surgical protocol needs to be
Acknowledgements

This work was partially funded by NSF Engineering Research Center grant EEC9731478. We thank Branko Jaramaz and Tony DiGioia, of Carnegie Mellon University and the Institute of Computer Assisted Orthopedic Surgery at Western Pennsylvania for providing the pelvis data. We specially thank Dinggang Shen for his suggestions on deformable registration and for the use of his programs in one step of our 3D model building process. We thank Ofri Sadowsky for reviewing the paper and providing suggestions to improve the paper. We thank Jerry Prince and Christos Davatzikos for many useful discussions. We also thank the anonymous reviewers for suggestions that greatly improved the paper.

References


Table 3. Testing transformation configurations

<table>
<thead>
<tr>
<th>Case</th>
<th>t_x</th>
<th>t_y</th>
<th>t_z</th>
<th>r_x</th>
<th>r_y</th>
<th>r_z</th>
<th>s_x</th>
<th>s_y</th>
<th>s_z</th>
<th>d_x</th>
<th>d_y</th>
<th>d_z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-10</td>
<td>-5</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-5</td>
<td>-5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
<td>1.05</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-0.4</td>
</tr>
</tbody>
</table>

Table 4. Validation results of transformations and similarity measures (Experiment 1)

<table>
<thead>
<tr>
<th>Image subract</th>
<th>Cross correlation</th>
<th>Entropy</th>
<th>Mutual information</th>
<th>Pattern correlation</th>
<th>Pattern intensity</th>
<th>Gradient difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average (case 1-6)</td>
<td>95.9, 0.39</td>
<td>96.4, 0.29</td>
<td>96.5, 0.28</td>
<td>96.5, 0.26</td>
<td>94.0, 0.45</td>
<td>93.5, 0.49</td>
</tr>
<tr>
<td>Average (case 7-15)</td>
<td>89.5, 0.98</td>
<td>93.1, 0.46</td>
<td>91.7, 0.57</td>
<td>93.2, 0.45</td>
<td>90.7, 0.65</td>
<td>89.0, 0.80</td>
</tr>
<tr>
<td>Average (case 1-15)</td>
<td>92.1, 0.74</td>
<td>94.4, 0.39</td>
<td>93.6, 0.46</td>
<td>94.5, 0.38</td>
<td>92.0, 0.57</td>
<td>90.8, 0.68</td>
</tr>
<tr>
<td>Std Dev (case 1-15)</td>
<td>6.2, 0.63</td>
<td>3.1, 0.17</td>
<td>3.3, 0.21</td>
<td>2.9, 0.17</td>
<td>4.2, 0.24</td>
<td>4.1, 0.29</td>
</tr>
</tbody>
</table>

Note: the numbers in each table cell are in the format of "Overlap (%), MeanD (mm)"

Table 5. Similarity measure comparison using Experiment 2

<table>
<thead>
<tr>
<th>Image subract</th>
<th>Cross correlation</th>
<th>Entropy</th>
<th>Mutual information</th>
<th>Gradient correlation</th>
<th>Pattern intensity</th>
<th>Gradient difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overlap (%)</td>
<td>84.6</td>
<td>86.8</td>
<td>85.1</td>
<td>87.9</td>
<td>85.5</td>
<td>86.1</td>
</tr>
<tr>
<td>MeanD (mm)</td>
<td>1.32</td>
<td>0.98</td>
<td>1.23</td>
<td>0.93</td>
<td>1.16</td>
<td>1.07</td>
</tr>
</tbody>
</table>
Table 6. Statistics of validation results in restoring transformation (Experiment 1)

<table>
<thead>
<tr>
<th></th>
<th>Overlap (%)</th>
<th>MeanD (mm)</th>
<th>AT (%)</th>
<th>AR (%)</th>
<th>AS (%)</th>
<th>AD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>94.2</td>
<td>0.36</td>
<td>3.94</td>
<td>9.51</td>
<td>6.50</td>
<td>21.6</td>
</tr>
<tr>
<td>Std Dev</td>
<td>2.7</td>
<td>0.19</td>
<td>1.20</td>
<td>1.83</td>
<td>1.44</td>
<td>4.85</td>
</tr>
<tr>
<td>Worst Case</td>
<td>90.7</td>
<td>0.94</td>
<td>9.56</td>
<td>12.3</td>
<td>10.6</td>
<td>38.3</td>
</tr>
</tbody>
</table>

Note: $\Delta T$, $\Delta R$, $\Delta S$, and $\Delta D$ are in percentage

Table 7. Number of 2D views vs. registration results

<table>
<thead>
<tr>
<th>Number of views</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume overlap (%)</td>
<td>64.2</td>
<td>86.3</td>
<td>87.9</td>
<td>88.4</td>
<td>89.3</td>
<td>90.1</td>
</tr>
<tr>
<td>MeanD (mm)</td>
<td>3.54</td>
<td>1.12</td>
<td>0.93</td>
<td>0.82</td>
<td>0.73</td>
<td>0.69</td>
</tr>
<tr>
<td>Running time (s)</td>
<td>56</td>
<td>147</td>
<td>463</td>
<td>605</td>
<td>724</td>
<td>819</td>
</tr>
</tbody>
</table>