An Iterative Framework for Improving the Accuracy of Intraoperative Intensity-Based 2D/3D Registration for Image-Guided Orthopedic Surgery

Yoshito Otake[1], Mehran Armand[2], Ofri Sadowsky[1],
Robert S. Armiger[2], Peter Kazanzides[1], and Russell H. Taylor[1]

[1] Department of Computer Science, the Johns Hopkins University,
3400 North Charles Street, Baltimore, MD, 21218, USA
[2] The Johns Hopkins University Applied Physics Laboratory,
11100 Johns Hopkins Road, Laurel, Maryland 20723, USA
{otake,osadow,pkaz,rht}@jhu.edu,
{Mehran.Armand,Robert.Armiger}@jhuapl.edu

Abstract. We propose an iterative refinement framework that improves the accuracy of intraoperative intensity-based 2D/3D registration. The method optimizes both the extrinsic camera parameters and the object pose. The algorithm estimates the transformation between the fiducials and the patient intraoperatively using a small number of X-ray images. The proposed algorithm was validated in an experiment using a cadaveric phantom, in which the true registration was acquired from CT data. The results of 50 registration trials with randomized initial conditions on a pair of X-ray C-arm images taken at 32° angular separation showed that the iterative refinement process improved the translational error by 0.32 mm and the rotational error by 0.61 degrees when compared to the 2D/3D registration without iteration. This tool has the potential to allow routine use of image guided therapy by computing registration parameters using only two X-ray images.

Keywords: Intensity-based 2D/3D registration, Image-guided therapy.

1 Introduction

In an image-guided surgery system using external tracking devices, the process of patient registration involves calculating a transformation that maps preoperative datasets to the patient's physical anatomy. Point-based registration is performed using corresponding points between the preoperative data and the surface of target patient anatomy. This broadly adopted registration method is used in most commercial navigation systems. The intraoperative workflow for point-based registration is well established and registration error properties such as fiducial localization error (FLE), fiducial registration error (FRE) and target registration error (TRE) are extensively analyzed (e.g., [1]). Point-based methods, however, have a major disadvantage: they require physical contact with multiple points over the target anatomy using a navigated instrument. Consequently, the surgeon often needs to create a larger exposure than required for navigation-free minimally invasive surgery. Also, due to

the time involved in performing the registration, this procedure is usually conducted once at the beginning of the surgery, and the tracker reference body is carefully handled until the end of the surgery to avoid losing the registration. Ideally, the registration process should be fast enough to repeat several times during the surgery just before each procedure which requires the navigation.

In order to overcome the shortcomings of point-based registration (invasiveness and processing time), image-based registration has been investigated [2]. In image-based registration, intraoperative X-ray images and a 2D/3D registration technique are used for deriving patient position. The method has been often used for radiation therapy, usually with a fixed X-ray device capable of accurate control of the detector position [3]. Image-based registration has also been employed for orthopaedic procedures (e.g., [4]), because it requires less time and no additional surgical exposure compared to the point-based method. The tradeoff, however, is increased radiation exposure.

One of the difficulties precluding routine use of image-based registration during image-guided therapies is its relatively low accuracy and uncertain robustness. Most clinically available C-arm devices lack the position information for the acquired images, which is crucial for an accurate 2D/3D registration, due to poor physical stability of the large supported masses and no motion encoding. The pose estimation of the C-arm detector is a key technical challenge. Some researchers proposed attaching optical tracking fiducials on the detector itself [5]; however, this approach has several drawbacks in real clinical settings such as maintaining the line-of-sight of the optical tracker, remaining within the field of view of the tracker, and synchronizing the C-arm and tracker (due to the inherent physical instability of the C-arm). Our approach uses instead a previously developed fiducial [6] that is visible within the field of view and contains optical markers. Knowing the unique geometry of the fiducial, it is possible to resolve the 3D position of that fiducial relative to the detector from a 2D projection image. However, image distortion and background clutter adversely affect the registration accuracy in this approach. While some of these issues are mitigated by using lighter-weight and distortion-free flat-panel detectors, recovery of the C-arm pose and the patient pose still applies and can benefit from an iterative method for registration.
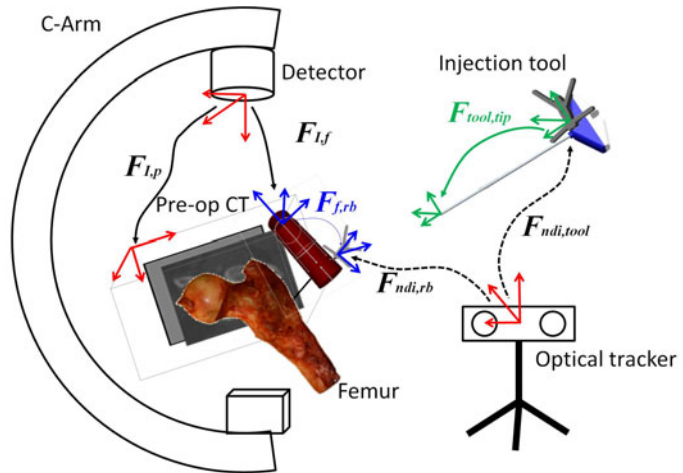
Fig. 1. An application scenario in orthopedic surgical navigation system

In this paper, we propose an iterative refinement framework for image-based registration that recovers the geometric transformations between the C-arm detector, images and CT coordinates while compensating for the error induced by image distortion and background clutter associated with intraoperative X-ray imaging. The motivation of this paper is to allow routine use of image-based
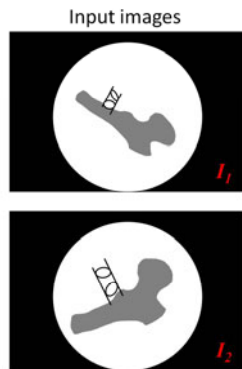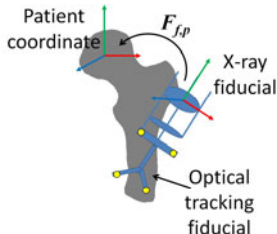
Fig. 2. Problem description of the X-ray image-based fiducial registration

2D/3D registration in image-guided therapy, and to avoid the time-consuming point-based registration methods that require larger surgical exposures. In order to improve the accuracy and the stability of the image-based registration process given image distortion and background clutter, we propose a framework in which the C-arm pose estimation and object pose estimation are iteratively updated.

2 Materials and Methods

2.1 An Application Scenario in Orthopedic Surgical Navigation System

A case-study scenario of our X-ray image-based registration algorithm is shown in Fig.1. The schematic shows an orthopedic surgical navigation system for delivering a planned amount of bone cement to the proximal femur, as a proposed method for reducing the risk of fracture in osteoporotic bones. Prior to the surgery, calibration of the intrinsic camera parameters and distortion correction are performed on a conventional C-arm using a calibrated phantom proposed by Sadowsky *et al.* [7]. A fiducial structure [6] is then attached to the patient using a bone screw. Next, between two to five X-ray images are taken from various directions. Using the preoperative CT data and the known geometry of the fiducial, the transformation between the tracking fiducial and the patient anatomy is computed. The transformation is then used to determine the pose of the surgical tools (e.g. needle, drill, etc.) relative to the patient as seen by an optical tracker.

2.2 Problem Description

The main problem in X-ray image-based registration is described in Fig. 2. Several X-ray source images (for simplicity, we consider 2 images, I_1 and I_2) are used to compute the transformation between patient and fiducial, $F_{f,p}$. (Note that we use the convention $F_{a,b}$ to describe a transformation from coordinate system a to coordinate system b throughout this paper). The registration process is divided into two parts: C-arm detector pose estimation and patient pose estimation. The following sections describe each of the two steps respectively.
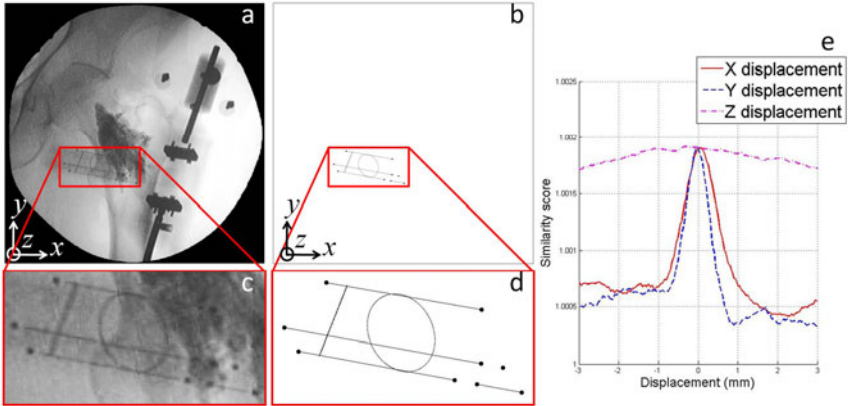
Fig. 3. C-arm detector pose estimation method. a) original fluoroscope image. b) projection image of the CAD model. c), d) magnified image of the fiducial. e) plot of the similarity score while displacing the object in x, y, and z direction.

2.3   C-arm Detector Pose Estimation

The six DOF transformation between the C-arm detector and the fiducial is computed from a 2D projection image using the custom hybrid fiducial that includes radio opaque features. The method for computing the transformation from segmented features (bead points, lines and ellipses) was described previously [6]. In clinical settings, however, due to severe background clutter, the segmentation of features is not always possible. The contributing factors to the background clutter and image noise may include patient anatomy, surgical tools and so on. Therefore we use an alternative approach to estimate the C-arm pose, eliminating the need for the segmentation process. In this approach, shown in Fig.3, first the pose is roughly estimated manually by point correspondence of the beads on the fiducial. Then a projected image of the CAD model of the fiducial is created based on the roughly estimated pose. A similarity score, mutual information [8], between the generated image and the intraoperative X-ray image is computed. A non-linear optimization method, the Nelder-Mead Downhill Simplex Algorithm [9], is applied to find the pose that yields the maximum similarity score per each image (I_1 and I_2). The pose is then used as an initial guess for our iterative framework.

2.4   Patient Pose Estimation Using Intensity-Based 2D/3D Registration Algorithm

The pose of the patient with respect to the fiducial is estimated using intensity-based 2D/3D registration, and is described in detail in our previously published work [10-11]. The 2D/3D registration algorithm makes use of a computationally fast DRR (Digitally Reconstructed Radiograph) created from CT data. To achieve this, the CT data is represented as a tetrahedral mesh model. Given any arbitrary source and detector location, a simulated radiograph is produced using the line integral of

intensities. Our objective is to vary the pose parameters to maximize the similarity score between the DRRs and X-ray projection images. In the DRR generation step, the projection of the X-ray tracking fiducial is merged in the DRR in order to create an image similar to the target image (Fig. 4 p_2). The similarity measure maximization process is described as:

$$\hat{F}_{f,p} = \arg\max_{F_{f,p}}\left[\sum_{i=1}^{N} S\left(I_i, DRR^+\left(F_{f,p}, F_{f,I_i}\right)\right)\right] \tag{1}$$

where $S(I_1, I_2)$ denotes a similarity score, which could be implemented as any type of similarity score. As described in 2.5, we implemented gradient information. $DRR^+(F_{f,p}, F_{f,Ii})$ denotes the DRR image created from the CT data and the X-ray fiducial transformed into the CT coordinate system using $F_{f,p}$. The DRR projection is in the coordinate system of the ith image, which is the same as the C-arm detector position, with respect to the fiducial coordinate system of $F_{f,Ii}$.

2.5 Similarity Measure

Gradient information (defined in [12]) was used as a similarity measure in our algorithm (Fig. 4). First, the gradient vector is computed for each sample point in each image using a Gaussian gradient filter with kernel of scale σ. Section 4 discusses the determination of the proper σ value. The angle α and weighting function w between the gradient vectors is defined as:

$$\alpha_{i,j} = \arccos\frac{\nabla I_1(i,j)\cdot\nabla I_2(i,j)}{\left|\nabla I_1(i,j)\right|\left|\nabla I_2(i,j)\right|}, w(i,j) = \frac{\cos\left(2\alpha_{i,j}\right)+1}{2} \tag{2}$$

where $\nabla I(i,j)$ denotes the gradient vector of the pixel located at (i,j) in image I. In order to take into account strong gradients that appear in both images, the weighting function is multiplied by
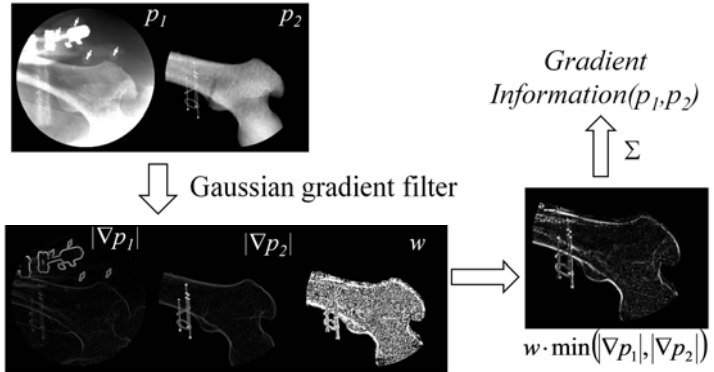
Fig. 4. Computation of the gradient-based similarity measure that is used as a cost function for the optimization

the minimum of the gradient magnitudes. Summation of the resulting product is defined as a gradient information $GI(I_1,I_2)$.

2.6 Iterative Refinement Framework

Both C-arm and patient pose estimation algorithms suffer from the error induced by C-arm intrinsic parameter estimation error, incomplete distortion correction and background clutter in the images. These errors are most severe in the axis

Fig. 5. Workflow of the proposed iterative refinement approach

perpendicular to the image plane as the similarity score does not change significantly when the object is moved in this direction.

This is shown in Fig. 3(e) where the maximum of the z-displacement similarity score is not as distinct as that of the x- and y- displacement. Because of this limitation, the optimization process is less stable and may converge incorrectly to a local maximum. If the C-arm pose estimation includes error, the patient pose estimation algorithm becomes less accurate and vice versa. In order to overcome this issue, we developed an iterative refinement framework as shown in Fig. 5. In the iterative framework, both the C-arm pose and the patient pose are refined with respect to the fiducial. Note that the intrinsic parameters were assumed constant throughout the procedure. Given the estimated transformation between the fiducial and patient, $(F_{f,p})$ in equation (1), the C-arm pose for each image is refined using both patient anatomy and fiducial information as the following.

$$\hat{F}_{f,I_i} = \arg \max_{F_{f,I_i}} \left[S\left(I_i, DRR^+\left(F_{f,p}, F_{f,I_i}\right)\right)\right] \tag{3}$$

where the parameters are the same as in (1). $\hat{F}_{f,p}$ is computed by (1) and inserted into (3) as a dependent variable in order to compute a refined registration F_{f,I_i} denoted as $\hat{F}_{f,I_i}$. The two optimization processes (1) and (3) are iteratively repeated (Fig. 6). The intuition behind this iterative approach is that incorrectly estimated camera parameters can be adjusted by using the patient itself as a fiducial.

2.7 Experiment

We have evaluated our method on images acquired from a cadaveric specimen. The X-ray tracking fiducial was fixed to the femur bone using a bone screw (Fig. 7). CT data was acquired with a spatial resolution of 0.835× 0.835 × 0.3mm. We defined a local coordinate system for the femur bone at the center of the femoral head by fitting

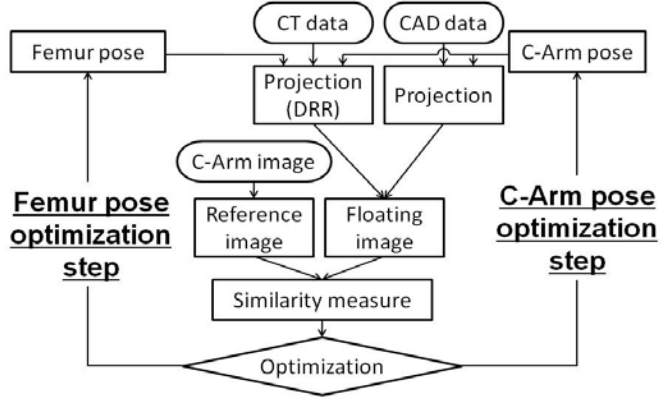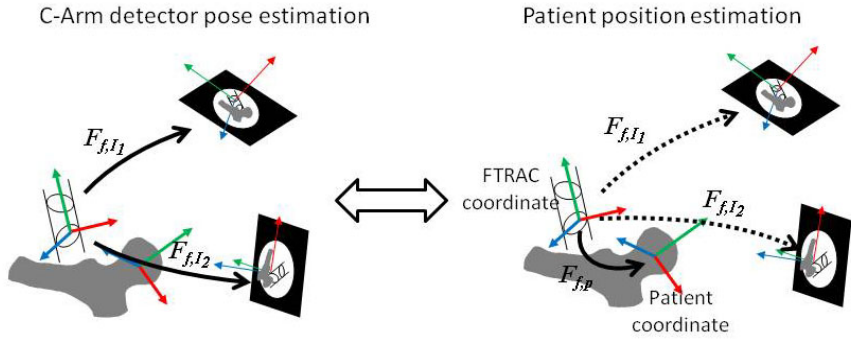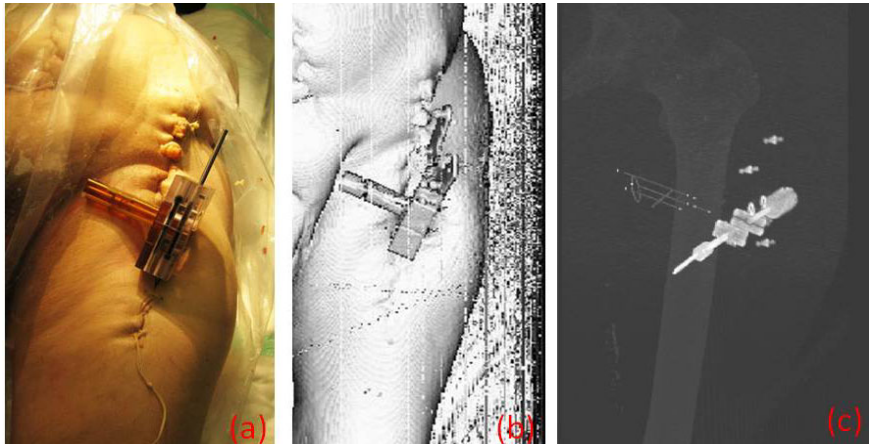**Fig. 6.** Proposed iterative refinement process. C-Arm detector pose estimation and patient position estimation are iterated in order to improve the accuracy of registration between fiducial coordinates and CT coordinates ($F_{f,p}$).

Fig. 7. Cadaver phantom used for the validation experiment. (a) specimen, (b)(c) Volume rendering and Maximum Intensity Projection images of CT data.

a sphere to a manually segmented femur surface model. Two X-ray images were acquired from two different directions (Fig. 8) with an angle of 32 degrees between the normal vectors of the two images. This relatively small angular separation between images is sub-optimal for registration accuracy, but is typical of the practical constraints often encountered intraoperatively. The ground truth registration was obtained by applying rigid-body point-based registration [13] to nine beads on the FTRAC. Fifty registration trials were performed using fifty different initial guesses in order to simulate the operator dependent error during the manual process. The initial guesses were obtained by concatenating randomly selected transformations with uniform and independently distributed translations within ±10 mm and within ±10° of the ground truth registration. The average error between the transformation estimated by the proposed algorithm and the ground truth registration was computed relative to the coordinate system of the image plane associated with the first image. The

rotational part
of the error
was first
computed as
the arithmetic
mean of the
rotations
represented as
unit quarter-
nions, then
converted
into Euler
angle repre-
sentation to
give an intuit-

Fig. 8. C-Arm images captured for validation experiment

tive sense on the direction of rotation. While this estimate of the mean rotation using a
unit quaternion ignores the fact that rotations belong to a nonlinear manifold, it was
shown previously that a linear approximation of a set of unit quaternions that spans a
small area on the four-dimensional sphere is adequate [14].

3   Results

The average errors and their standard deviations along each axis for 50 registration
trials are shown in Table 1. We compared the results using a conventional approach to
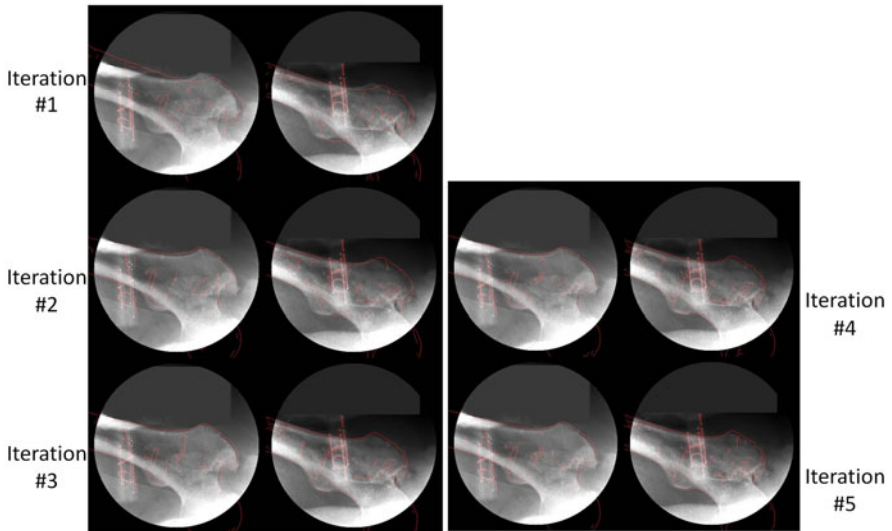the proposed iterative method after five iterations. The directions of each axis are

Fig. 9. Exemplar result of the iterative refinement process. Red line indicates the edge of the
DRR which was generated based on the femur position estimated at the previous iteration step.

Table 1. Results of the 50 registration trials using randomly selected initial guess within ±10 mm translation and within ±10° rotation

		X	Y	Z	X-Y plane
Translation (mm)	Without iteration	1.18±0.74	1.73±1.04	3.50±2.07	2.09±0.87
	After 5 iterations	0.97±0.62	1.80±1.02	3.06±1.74	2.05±0.89
Rotation (deg)	Without iteration	1.72±2.13	2.19±3.38	1.04±2.06	
	After 5 iterations	1.36±1.82	1.77±2.65	1.00±1.99	

shown in Fig. 8(lower-left). The iterative approach improved errors for all parameters except Ty, and decreased the Euclidean translation error from 4.1 to 3.7mm. The translational error along the Z-axis, which is perpendicular to the image plane, improved 0.43mm after 5 iterations. As noted, the greatest errors are usually observed along the Z-axis. Table 1 shows that the standard deviations of the error were reduced in all axes, implying increased stability in the presence of an error in the initial guess. Fig. 9 shows an example of the iteration process. The outline of the DRR images, shown in red, was created based on the C-arm pose. Each outline was created using its prior iteration step. In the figure, the outlines are overlaid onto the original C-arm images. The results clearly show that the DRR approached to the correct registration, as the iterations continued.

4   Discussion and Conclusions

This paper has presented an iterative refinement framework for intensity-based 2D/3D registration without external tracking of the C-arm pose. Because of the cost, technical difficulties, and insufficient accuracy associated with tracking the pose of the imaging device, intensity-based 2D/3D registration is not commonly used in commercial surgical navigation systems. By using the image information for tracking the device, and by improving the accuracy and robustness of the registration technique, the proposed method increases the potential to use intensity-based 2D/3D registration routinely in image-guided therapy.

The method combines two different optimization problems: C-arm pose estimation and patient pose estimation. The experimental results show the iterative approach improves the registration accuracy along the Z-axis (normal to the image plane), which is important to accurately resolve tool position in three-dimensional space. Note that maximizing the similarity measure along the Z-axis was found to be most problematic because the global maximum is not clearly defined (see Fig. 3(e)). The actual amount of error also depends on experimental conditions including tissue quality and on the relatively narrow difference in the viewing angle of the acquired images. The images acquired in this experiment were not optimized to show the best performance of the proposed algorithm. Additional experiments to expand this initial result are planned. The key contribution of this paper is the demonstration of a novel iterative framework for reducing geometric uncertainties resulting from imperfect C-arm calibration and patient registration errors. Our framework utilizes the patient CT information not only for estimating the patient pose but also for improving the C-arm pose estimation by using the patient as a type of reference *fiducial*.
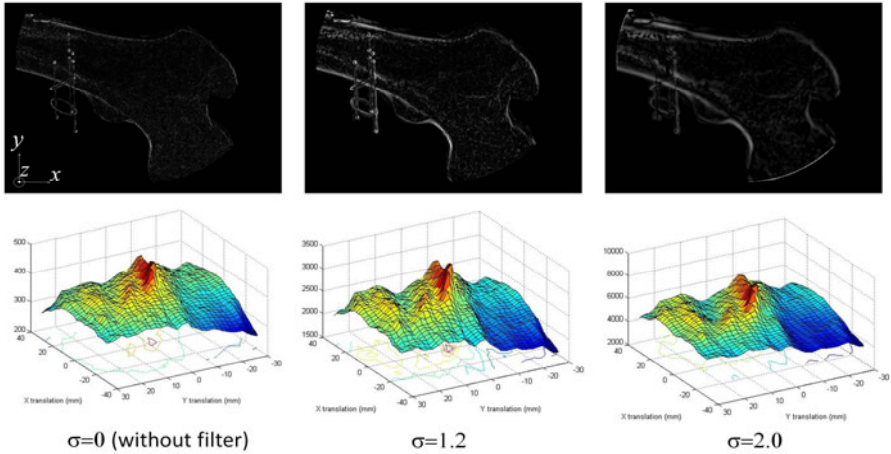
Fig. 10. Comparison of the similarity measure between different sigma values for Gaussian gradient filtering. (upper row: gradient image *w* computed by (2) was used as a cost function. lower row: Plot of the cost function, summation of the intensities of the upper row images.)

The use of a statistics-based similarity measure without considering spatial information results in an ill-defined metric that can contain local maxima in the presence of significant background clutter [12]. Therefore, instead of the mutual information which is widely used in intensity-based 3D/3D and 2D/3D registration, we used the gradient information as the similarity measure.

The scale of the kernel (σ) in the Gaussian filter has a great influence on the similarity measure computation described in 2.4. In order to have information about the preferable σ value in our application, we tested three different σ values. In this analysis, we first generated the DRR using the femur pose derived from the ground truth registration. We then translated the femur position from -30 to +30 mm along the X and Y directions. The gradient information between the DRR and the original C-arm image was computed for each pose and plotted (Fig. 10). For convergence to the global maximum, the cost function should be smooth and have a sharp peak at the center. Based on the results, we used σ=2.0 for our experiments.

The experiment described here used a C-arm device with a conventional X-ray image-intensifier, which is ubiquitous in the current operating theater. Flat-panel detector systems are emerging in operating rooms and have the advantage of less image distortion and high image contrast. Our iterative registration approach is still applicable and expected to be beneficial with these devices by resolving the ambiguity due to the errors related to the DRR generation process, C-arm intrinsic parameter calibration, and optimization of a cost function with inherent local minima. Although the cadaver experiment showed sufficient accuracy and robustness of the proposed method, we note that the movement of the fiducial with respect to the target organ after capturing the C-arm image may introduce additional errors in the actual surgery when compared to the cadaver experiments. Thus the framework would further the potential for performing minimally invasive registration using a small number of X-ray images (e.g. two to five images) during the routine image-guided orthopedic surgery. We believe that the methods reported here can apply to broader

applications of image-guided therapy including plastic surgery, neurosurgery, and dentistry and so on.

We have demonstrated a unique approach for intensity-based registration that improves the accuracy compared to conventional 2D/3D registration through iterative optimization. We plan to perform additional cadaveric studies in order to investigate the effects of the number of images and different angular separation. These additional studies will produce a more comprehensive analysis of the robustness of the reported method for intensity-based 2D/3D registration.

Acknowledgements

References

1. Moghari, M.H., Abolmaesumi, P.: Distribution of fiducial registration error in rigid-body point-based registration. IEEE Trans. Med. Imaging 28(11), 1791–1801 (2009)
2. Gueziec, A., Kazanzides, P., Williamson, B., Taylor, R.H.: Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot. IEEE Trans. Med. Imaging 17(5), 715–728 (1998)
3. Chen, X., Gilkeson, R.C., Fei, B.: Automatic 3D-to-2D registration for CT and dual-energy digital radiography for calcification detection. Med. Phys. 34(12), 4934–4943 (2007)
4. Scarvell, J.M., Pickering, M.R., Smith, P.N.: New registration algorithm for determining 3D knee kinematics using CT and single-plane fluoroscopy with improved out-of-plane translation accuracy. J. Orthop. Res. (October 1, 2009)
5. Livyatan, H., Yaniv, Z., Joskowicz, L.: Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. IEEE Trans. Med. Imaging 22(11), 1395–1406 (2003)
6. Jain, A., Fichtinger, G.: C-arm tracking and reconstruction without an external tracker. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006, Part I. LNCS, vol. 4190, pp. 494–502. Springer, Heidelberg (2006)
7. Sadowsky, O.: Image registration and hybrid volume reconstruction of bone anatomy using a statistical shape atlas [dissertation]. The Johns Hopkins University (2008)
8. Wells III, W.M., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal volume registration by maximization of mutual information. Med. Image Anal. 1(1), 35–51 (1996)
9. Nelder, J.A., Mead, R.: A simplex method for function minimization. The Computer Journal 7(4), 308–313 (1965)
10. Sadowsky, O., Cohen, J.D., Taylor, R.H.: Projected tetrahedra revisited: A barycentric formulation applied to digital radiograph reconstruction using higher-order attenuation functions. IEEE Trans. Vis. Comput. Graph. 12(4), 461–473 (2006)
11. Sadowsky, O., Chintalapani, G., Taylor, R.H.: Deformable 2D-3D registration of the pelvis with a limited field of view, using shape statistics. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part II. LNCS, vol. 4792, pp. 519–526. Springer, Heidelberg (2007)
12. Pluim, J.P., Maintz, J.B., Viergever, M.A.: Image registration by maximization of combined mutual information and gradient information. IEEE Trans. Med. Imaging 19(8), 809–814 (2000)
13. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. J. Opt. Soc. Am. A. 4(4), 629–642 (1987)
14. Gramkow, C.: On averaging rotations. International Journal of Computer Vision 42(1), 7–16 (2001)