Robust registration for computer-integrated orthopedic surgery: Laboratory validation and clinical experience

B. Ma, R.E. Ellis*

School of Computing, Queen's University, Kingston, Ontario, Canada K7L 3N6

Abstract

In order to provide navigational guidance during computer-integrated orthopedic surgery, the anatomy of the patient must first be registered to a medical image or model. A common registration approach is to digitize points from the surface of a bone and then find the rigid transformation that best matches the points to the model by constrained optimization. Many optimization criteria, including a least-squares objective function, perform poorly if the data include spurious data points (outliers). This paper describes a statistically robust, surface-based registration algorithm that we have developed for orthopedic surgery. To find an initial estimate, the user digitizes points from predefined regions of bone that are large enough to reliably locate even in the absence of anatomic landmarks. Outliers are automatically detected and managed by integrating a statistically robust M-estimator with the iterative-closest-point algorithm. Our in vitro validation method simulated the registration process by drawing registration data points from several sets of densely digitized surface points. The method has been used clinically in computer-integrated surgery for high tibial osteotomy, distal radius osteotomy, and excision of osteoid osteoma.
© 2003 Elsevier B.V. All rights reserved.

Keywords: Computer-integrated orthopedic surgery; Validation; Clinical experience

1. Introduction

Registration of a patient to a medical image or model is a fundamental requirement of computer-integrated surgery (CIS) systems that provide intraoperative navigational guidance. A registration method for general use in orthopedic surgery should be effective, fast, and simple to use. It should not depend on readily locatable anatomic landmarks because these are often difficult to find or nonexistent such as when performing a percutaneous procedure.

The iterative-closest-point (ICP) algorithm of Besl and McKay (1992) is a well-known method for registering a 3D set of points to a 3D model that minimizes the sum of squared residual errors between the set and the model, i.e. it finds a registration that is locally best in a least-squares sense. This algorithm requires an initial estimate of the registration; the computation speed and registration accuracy depend on how this initial estimate is chosen. Two acknowledged problems with ICP-like algorithms are (1) the requirement of a good initial estimate, and (2) that minimization of the sum of squared errors is optimal only under the assumption that measurement errors are independent and have identical Gaussian distributions. If noise in the measurements is non-Gaussian, a least-squares error measure can produce poor results. A common source of non-Gaussian noise is the presence of statistical outliers which can be present in the measurements if, for example, a measurement is accidentally collected far from the target anatomy or is taken from a point outside the medical image. Fig. 1 illustrates the effect of outliers on registration accuracy.

Two attractive features of ICP are that it is guaranteed to converge to a minimum in its objective function, and convergence is very fast when started from a good initial estimate. An efficient implementation by Rusinkiewicz and

*Corresponding author.
E-mail address: ellis@cs.queensu.ca (R.E. Ellis).

points with outliers   least squares ICP   robust ICP
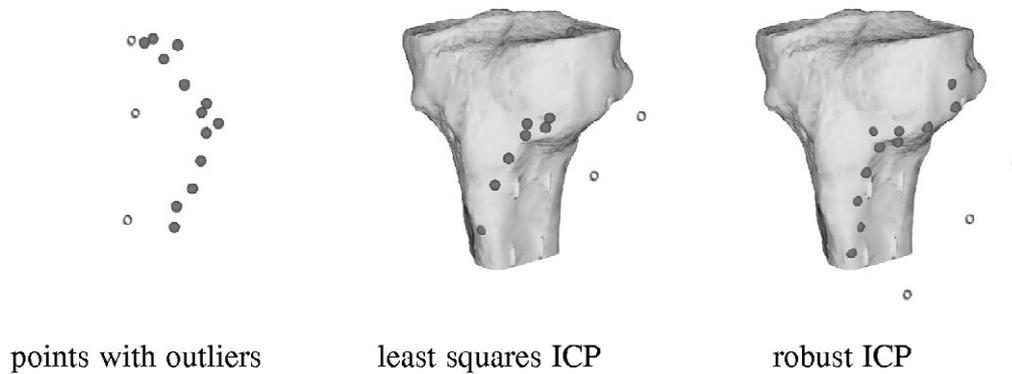
Fig. 1. The effect of outliers on registration accuracy. On the left are 16 measured contact points from a phantom tibia; three of the points are outliers. In the middle is a registration obtained using ICP initialized with a good estimate of a reasonable registration. On the right is a registration obtained using our robust implementation of ICP started from the same initial estimate.

Levoy (2001) that was not robust to outliers has been shown to register 2000 data points to models comprised of 100 000 points in tens of milliseconds. The computational efficiency of ICP makes it an attractive starting point for a robust registration algorithm.

We have previously described a robust registration algorithm in (Ma et al., 1999). In this article, we describe our experiences using this algorithm for orthopedic procedures. In Section 2 we present a brief summary of ICP-like methods and robust registration methods. In Section 3 we describe our registration algorithm. In Section 4 we describe our in vitro experiments used to validate registration accuracy. In Section 5 we describe the clinical application of our algorithm for high tibial osteotomy, distal radius osteotomy, and excision of osteoid osteoma. We conclude by discussing the results and limitations of our work.

2. Related work

Lavallée (1995) and Maintz and Viergever (1998) have extensively reviewed the literature on general registration algorithms. Here, we will review the literature of ICP-like methods and robust registration methods.

Several research groups have reported using ICP in CIS applications. Palombara et al. (1997), who used ICP for in vitro experiments related to total knee arthroplasty, reported that outliers were detrimental to the accuracy of the computed registrations. Betting et al. (1995) used a modified version of ICP that took into account surface-normal information and that used a $k-d$ tree to accelerate the nearest-neighbor search. Maurer et al. (1996) also used a $k-d$ tree to accelerate ICP, and discussed how to decompose a set of surface primitives into a weighted point-set representation. Applications of their algorithm have also been reported in (Herring et al., 1998) and (Maurer et al., 1998). Simon et al. (1995) used a $k-d$ tree and other techniques to decrease the computation time

required by ICP, and discussed how to choose an optimal set of points to use for registration. Cuchet et al. (1995) used a chamfer map to accelerate the nearest-neighbor search.

Rusinkiewicz and Levoy (2001) studied several variations of ICP for model-based tracking and 3D scanning. Greenspan and Godin (2001) have described ICP variants that use nearest-neighbour search methods which are significantly faster than $k-d$ tree methods for surfaces represented by point sets.

Rangarajan et al. (1997) extended the Procrustes method to matching point sets of different cardinality with unknown correspondences. Their Softassign method jointly solves for the rotation and point correspondances, treating non-homologies as outliers. There is empirical evidence in (Rangarajan et al., 1999) suggesting that the Softassign objective function is much smoother than the root mean squared error function minimized by ICP. Chui and Rangarajan (2000) re-interpreted their previous work as a maximum likelihood problem and used an approach similar to expectation maximization to solve the feature registration problem. Their expectation-like step updates the feature correspondences and the maximization step updates the registration transformation. Similar methods were independently developed by Granger et al. (2001). Dellaert (2001) argued that the feature correspondences were actually nuisance parameters and that one needs to compute the probability distribution of the transformation parameters over all possible correspondences. Since there are a combinatorial number of correspondences, the exact solution is computationally intractable and was approximated using Monte Carlo sampling. All of these methods can be made robust against outliers.

Zhang (1994) independently described an algorithm that is very similar to ICP. He attempted to identify outliers by examining the standard deviation of the residual matching errors and removing those points that had errors greater than some multiple of the standard deviation. We can identify three problems with this thresholding approach.

First, there is no mathematically sound way of choosing the threshold value as some multiple of the standard deviation. Second, a least-squares solution assumes a Gaussian distribution of the residual errors, yet an outlier may have a *smaller* residual error than a valid observation after a least-squares procedure. Finally, thresholding does not address the issue of the true underlying distribution of the measurement errors. Despite these limitations, other researchers Blais and Levine (1995) and Feldmar and Ayache (1994) have applied thresholding in an attempt to remove outliers in their registration methods.

Robust registration is a process of attempting to diminish or eliminate the effects of outliers in estimating the registration transformation. Most previous work in robust registration has been done by the computer-vision community. Fischler and Bolles (1981) reported one of the first statistically robust algorithms for scene analysis. Their random-sample consensus (RANSAC) algorithm randomly draws a minimum number of data points to solve the estimation problem; this minimal set is used to generate an initial solution. The consensus set is the set of all data points that agree with the initial solution to within some tolerance value. If the cardinality of the consensus set is sufficiently large, then the consensus set is used to compute a new estimate of the solution. If the cardinality of the consensus set is too small, then a new random subset is drawn to obtain a different initial solution. Meer et al. (1991) have stated that RANSAC and least-median-of-squares (LMS) estimation yield very similar results.

Haralick et al. (1989) presented a solution to the two-dimensional registration problem for matched-point sets using the Tukey *M*-estimator. They tested their method on synthetic data with added Gaussian noise and gross outliers. They concluded that their robust method had better performance and stability than a least-squares solution, but that it did not always discriminate accurately between outliers and valid data points. Robust estimation using *M*-estimators is a well known topic described by Hampel et al. (1986), Hoaglin et al. (1983) and Huber (1981).

Kumar and Hanson (1990) solved the registration problem, given the correspondence between three-dimensional lines represented in a world coordinate frame and two-dimensional image lines represented in a camera coordinate frame, using a Tukey *M*-estimator algorithm and an LMS algorithm. The LMS estimator was described by Rousseeuw and Leroy (1987). Because the LMS estimator does not have an analytic solution, algorithms used to compute it typically resort to brute-force minimization of the LMS objective function. If the running time for such an algorithm is too high, one can use a probabilistic argument to limit the size of the minimization problem at the risk of obtaining a poor solution. For their registration problem, the authors found that they had to search through a much larger number of potential solutions to find a good registration than the probabilistic argument predicted. They

found that obtaining a good solution with their Tukey-based algorithm depended critically on the quality of the initial estimate of the registration.

Masuda and Yokoya (1995) reported a statistically robust version of ICP. Their algorithm used the LMS estimator, and it was applied to the segmentation and matching of range images. They claimed that, for range images of size 256×256 and larger, their method could segment and register images with up to 50% outlier contamination.

Luck et al. (2000) recently reported a robust version of ICP that uses simulated annealing in an attempt to find the registration that produces the global minimum matching error. They applied their algorithm to the segmentation and matching of range images. They claimed that their method could register images with up to 50% outlier contamination.

ICP-like algorithms estimate the rigid transformation by searching for correspondences between the data and the model. Instead of searching for correspondences, we can attempt to directly estimate the best transformation. One such algorithm described by Rucklidge (1996) minimizes the Hausdorff distance by evaluating the Hausdorff distance on a discretized subdivision of the the transformation space of rotation and translation. The Hausdorff distance can be modified to be robust against outliers, but is known to have local minima. Interesting timing comparisons between the Hausdorff distance and the absolute difference distance were reported by Hagedoorn and Veltkamp (1999) for 2D pattern matching. For rigid 2D transformations, matching patterns with hundreds of features required between 2 and 4 minutes for the Hausdorff distance. The accuracy of these methods depends on the coarseness of the discretization of the transformation space.

Bächler et al. (2001) described their restricted surface matching algorithm which searches for the registration transformation parameters using an evolution strategy combined with local minimum optimization algorithm (Powell's method). They initialized their algorithm by using paired-point matching to a few coarse landmarks. They also included a penalty function in their objective function that penalized transformations that resulted in poor matching of the coarse landmarks.

3. Algorithm

In CIS, rigid-body surface-based registration is the process of finding a transformation from a set of measured points on the target anatomy to the model surface derived from the medical image. Let $P = \{\vec{p}_i\}$ be a set of n surface-data points measured from the target anatomy by the surgeon, let $X = \{\vec{x}_i\}$ be the set of all points on the surface model, and let $\mathbf{T}\vec{z} = \mathbf{R}\vec{z} + \vec{t}$ be a rigid transformation of a point $\vec{z}$. The registration goal is to find both the

rigid-body transformation $\mathbf{T}$ and some n-element subset Y of model surface locations X to which the target anatomy locations P project under $\mathbf{T}$. In the presence of errors the anatomical points P will not in general project exactly onto Y. A least-squares solution to the surface-based registration can be stated as the minimum, over $\mathbf{T}$ and $Y \subseteq X$, of

$$F_2(Y) = \sum_{i=1}^{n} \|\vec{y}_i - \mathbf{T}\vec{p}_i\|^2 = \sum_{i=1}^{n} \|\vec{r}_i\|^2, \tag{1}$$

where $\vec{y}_i \in Y$, and $\|\vec{r}_i\|$ is the magnitude of the residual matching error. In the general case this is a non-convex minimization problem with multiple local minima.

Many robust estimation techniques use *M-estimation*, in which the L_2 norm in (1) is replaced with a robust norm to yield an objective function of the form

$$F_M(Y) = \sum_{i=1}^{n} \rho(r_i; \sigma), \tag{2}$$

where $\rho(r_i; \sigma)$ is the robust norm applied to the residual r_i, and σ is a scale parameter that depends on the form of the expected error distribution. One robust estimator that has reportedly provided good performance on 3D range data is the Tukey biweight used by Mirza and Boyer (1993):

$$\rho(r; \sigma) = \begin{cases} \dfrac{\sigma^2}{2}\left(1 - \left(1 - \dfrac{r^2}{\sigma^2}\right)\right)^3 & \text{if } r \leq |\sigma|, \\ \sigma^2/2 & \text{otherwise.} \end{cases} \tag{3}$$

3.1. Robust registration estimation

We produced a robust version of ICP by modifying the process of updating the registration. This required a solution to the absolute orientation problem, for which Horn's method provided a common least-squares solution.

To obtain an M-estimate of absolute orientation, we used an iteratively reweighted least-squares modification (Hoaglin et al., 1983; Haralick et al., 1989) of Horn's method (Horn, 1987). The scale parameter σ in Eq. (3) was estimated (Rousseeuw and Leroy, 1987) by using the median of absolute deviations of the residuals: $\vec{r}_i(\tau) = \vec{y}_i - \mathbf{T}\vec{p}_i$:

$$\sigma = 1.4826 \, \text{median}_{i=1\ldots n}\left(\|\vec{r}_i\| - \text{median}_{i=1\ldots n}\|\vec{r}_j\|\right). \tag{4}$$

The main stages of our surface-based registration method were:

1. Spotlight regions were selected on the model of the anatomy. A spotlight is a generalization of an anatomical landmark, which we use for anatomical sites on which distinctive landmarks are not available intraoperatively. A spotlight is region of the model that the surgeon can locate readily on the patient.
2. Spotlight data were gathered intraoperatively. The surgeon contacted points on the exposed anatomical regions that corresponded to the spotlights shown on a monitor, as in Fig. 2.
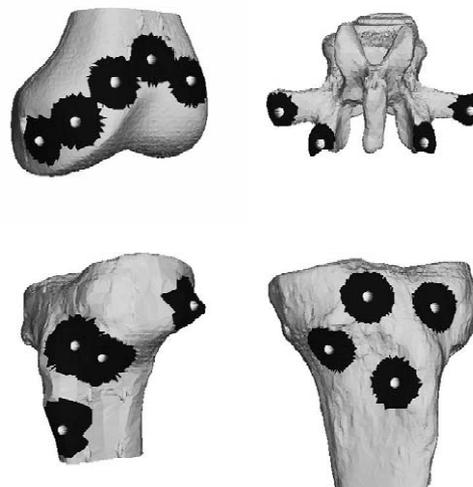
Fig. 2. Spotlight regions for registration shown on surface meshes of phantom bones derived from computed tomography. The spheres mark the centers of the spotlights. Clockwise from upper left: medial femur, posterior vertebra, medial tibia and proximal tibia.

3. The initial contact points were first matched to the spotlight *centroids* on the model using a simple least-squares minimization method (Horn, 1987).
4. The initial contact points were then matched to the spotlight *surface regions* on the model, using a least-squares ICP method.
5. The surgeon then contacted another set of points on the relevant exposed anatomical region for refinement of the registration transformation. In practice, these locations should be chosen to cover the anatomy that will be involved in the image-guided surgery and should provide sufficient translational and rotational constraints on the registration.
6. The initial registration, along with the full set of contact points, was then scored. The initial registration was repeatedly perturbed, and the least-squares residual for each point was calculated. The perturbation with the largest number of residuals that were all less than a user-supplied threshold was taken as the best initial registration estimate; that is we seek to maximize the rank of the largest residual with magnitude smaller than some threshold value.
7. Finally, the perturbation registration estimate was refined further using a version of the ICP algorithm that incorporated the robust Tukey-biweight M-estimator.

Each iteration of the ICP algorithm actually involves two estimation steps: given a registration estimate, one needs to find the set of closest points on the surface to the transformed data points. From these closest points on the surface, one then needs to update the registration estimate. It is important that the search for the closest points be fast because it is one of the most computationally demanding steps of the algorithm.

3.2. Refinement of registration using perturbation

Even when started from a reasonable spotlight estimate, traditional ICP and simplistic robust variants suffer from 'trapping' by converging to a local non-global minimum of the registration parameters. The usual robustness remedy is to perturb the solutions at the first, and possibly subsequent registration estimates (Grimson et al., 1995; Henri et al., 1995). Our remedy for trapping was to use a perturbation technique to conduct a local search through the possible registrations, seeking the registration that gives the best least-squares fit for the largest number of points. This alternative was accomplished heuristically by means of a simple search procedure.

Our implementation sampled 64 points uniformly from a unit hemisphere to define 64 axes of rotation. The surgical data were rotated, about their mutual centroid, around each of these axes by $\pm 3°$ and the Euclidean residual errors were calculated. For each of the 128 rotations, if half of the transformed surgical data had residuals that were less than a provided threshold value (1 mm) then the rotation was noted. The perturbation that produced, for at least half the surgical data, the maximum rank of the largest residual below a threshold value was deemed to be the perturbation that gave the best initial fit to the refinement surgical data.

4. An in vitro study

Our in vitro experiments involved registering 3D digitized points to computer models of plastic bone phantoms. Fiducial markers were inserted into all of the phantoms in order to obtain a registration with known error bounds relative to ground truth. This section describes our experimental methods and results.

4.1. Materials and methods

The phantoms used in the experiments were urethane-foam anatomical models (Pacific Research Laboratories, Vashon, WA, USA). One left tibia, one left femur, and one lumbar vertebra phantom were used in the experiments.

Anchorlok® Leading Edge® soft tissue anchors (Wright Medical, Arlington, TN, USA), which were titanium-alloy anchor screws of 1.9 mm diameter, were used as fiducial markers. The marker locations were extracted from the CT images using a previously validated center-of-mass technique (Ellis et al., 1996).

Computed tomography (CT) scans were performed at Kingston General Hospital (Kingston, Ontario, Canada) using a HiSpeed CT scanner (GE Medical Systems, Milwaukee, WI, USA). The CT images were acquired using a protocol that was known to produce good patient images: axial mode, slice width 3 mm, and slice spacing 2 mm.

Three-dimensional triangular facet models of the phantoms were extracted from the CT scans using isosurface reconstruction. A decimation algorithm was used to reduce the number of very small facets that are an artifact of the isosurface algorithm (Schroeder et al., 1998). The model details are given in Table 1.

All surface measurements were obtained using a six-degree-of-freedom mechanical pointer (Faro Technologies, Lake Mary, FL, USA) instrumented with a sharp tip probe. The manufacturer's specification of accuracy for this device is ± 0.3 mm in position.

Experimental registrations were examined with respect to a fiducial registration derived from the implanted markers. When comparing an experimental registration transformation $\mathbf{T_j}$ to the marker-based registration $\mathbf{T}$, the rotation error in degrees was computed. This was done by first finding the residual rigid transformation, $\mathbf{D}$ that satisfied the equation

$$\mathbf{T_j}(\cdot) = \mathbf{D}(\cdot)\,\mathbf{T}(\cdot) \Rightarrow \mathbf{D}(\cdot) = \mathbf{T_j}(\cdot)\,\mathbf{T}(\cdot)^{-1}. \qquad (5)$$

$\mathbf{D}$ was then decomposed into a screw transformation from which the rotation error about the screw axis could be obtained (Ellis et al., 1997). In addition to calculating the rotational error, the experimental registration was applied to the set of N_P measured marker points, $P = \{\vec{p}_i\}$, and the root-mean-square (RMS) error was computed as

$$\text{RMS error} = \sqrt{\frac{1}{N_P}\sum_{i=1}^{N_P}\|\mathbf{T_j}\vec{p}_i - \vec{x}_i\|^2}, \qquad (6)$$

where $X = \{\vec{x}_i\}$ was the set of $N_X = N_P$ marker points in model coordinates.

In order to evaluate the accuracy of our algorithm, we densely sampled the surface of each phantom using the FARO mechanical arm. A series of non-overlapping square grids were drawn on each phantom where the spacing between the grid lines was approximately 1 mm. Each phantom was fixed in a frame that was mechanically rigid with respect to the base of the FARO arm. The surfaces of the phantoms were sampled by contacting the FARO probe tip to each intersection of grid lines, resulting in an approximately uniform sampling over an area consistent with the exposure of a particular surgical procedure. Details of the surface sampling are shown in Table 2.

Table 1
Number of vertices and triangular facets in the anatomic models generated using isosurface reconstruction

Phantom name	# Vertices	# Facets
Femur	15 395	30 152
Tibia-hto	34 537	68 564
Tibia-tkr	24 453	48 048
Vertebra	27 096	54 904

Table 2
Details of surface contact point sampling for the various phantoms used in the surface-based registration experiments

Phantom name	Approximate region area (mm×mm)	Number of sampled regions	Number of points per region
Femur	10×10	8	100
Tibia-hto	10×10	12	100
Tibia-tkr	10×10	11	100
Vertebra	6×6	8	36

The surface points for the femur were collected from the inferior mediolateral aspect in the surgical exposure of a total knee replacement procedure. On the tibia, two sets of points were used. One set was selected from the superior anterolateral aspect of the tibia in an area which would be accessible during a high-tibial osteotomy procedure. The other set was selected from the superior anterior and the superior anteromedial regions between the tibial plateau and the insertion of the patellar tendon in an area which would be accessible during a total knee replacement procedure. On the lumbar vertebra, the points were chosen from the posterior aspect on the transverse processes, superior articular processes, and the laminae at the base of the spinous process—these areas would be accessible during a pedicle-screw insertion procedure.

Spotlights of an appropriate size and location were selected from each model. Using the marker-based registration, the center of each spotlight was visually estimated on the phantom surface and then contacted with the mechanical pointer. The spotlight region was then sampled approximately uniformly with a datum spacing of approximately 1 mm. Details of the spotlight sampling are shown in Table 3. The spotlights used in the experiments are shown in Fig. 2.

For each phantom, one set of 120 outliers were collected. Each outlier was collected approximately 5 mm from the phantom surface by inserting a spacer between the phantom and the tip of the mechanical pointer.

Our laboratory experiments were conducted to determine the quality of the final surface registration obtained starting from a spotlight estimate. For each phantom, one datum from every digitized spotlight and square grid data set was randomly chosen to form a registration point set. Outliers were added by randomly choosing points from the

Table 3
Details of spotlight contact point sampling for the various phantoms used in the surface-based registration experiments

Phantom name	Approximate spotlight diameter	Number of spotlights	Number of points per spotlight
Femur	25 mm	5	131, 123, 127, 140, 123
Tibia-hto	25 mm	4	100, 100, 100, 100
Tibia-tkr	25 mm	4	111, 118, 117, 127
Vertebra	10 mm	4	99, 86, 116, 116

digitized outlier data set and appending them to the registration point set, varying the number of outliers from zero to five. The position of the fiducial markers were also measured to obtain a fiducial registration.

One thousand registration point sets were registered to each phantom model using our robust algorithm. We also used the ordinary least-squares ICP algorithm starting from our spotlight-based estimate to obtain registrations. The rotation and RMS errors compared to the fiducial registration were computed for each of the one thousand trials.

The total number of registration points used for the femur, tibia-hto, tibia-tkr and vertebra phantoms were 13, 16, 15 and 12, respectively (the sum of the third columns from Tables 2 and 3). The number of registration points was comparable to that used by Bächler et al. (2001) and are consistent with the number of points that can readily be collected through a minimally invasive surgical exposure. With five additive outliers, the fraction of outlier contamination was between 24 and 29%.

4.2. Results and discussion

Histograms of the rotation error results are shown in Figs. 3–6. The RMS errors calculated using Eq. (6) have similar distributions, and their statistics are tabulated in Table 4.

One interesting result of this experiment was that, given only a relatively small number of digitized points distributed over the area of interest, ICP initialized from a spotlight-based estimate converged to a registration close to the fiducial registration. There were median rotation differences of 2–3° and median RMS errors of 2–3 mm between the surface-based and fiducial-based registrations. We were surprised that by using only a dozen or so measurements we could obtain registrations with this level of error, especially given the fact that the measured points were chosen at random.

The histograms (Figs. 3–5) show that the performance of the least-squares-based ICP algorithm deteriorated as the number of additive outliers was increased. The performance of the Tukey-based ICP algorithm remained consistent as the number of additive outliers was increased. Also, the performance of the Tukey-based algorithm is similar to that of the least-squares algorithm for the zero outlier case.

The histograms also show that the spread or variance of the distributions is large for both the least-squares and Tukey-based algorithms. Several factors were responsible for this: the large size of the spotlights limited the accuracy of the initial estimate, the small number of points underconstrained the surface-based registration problem, and ICP is only guaranteed to converge to a local minima solution.

The results for the vertebral phantom had an unusually large number of very poor registrations even though the
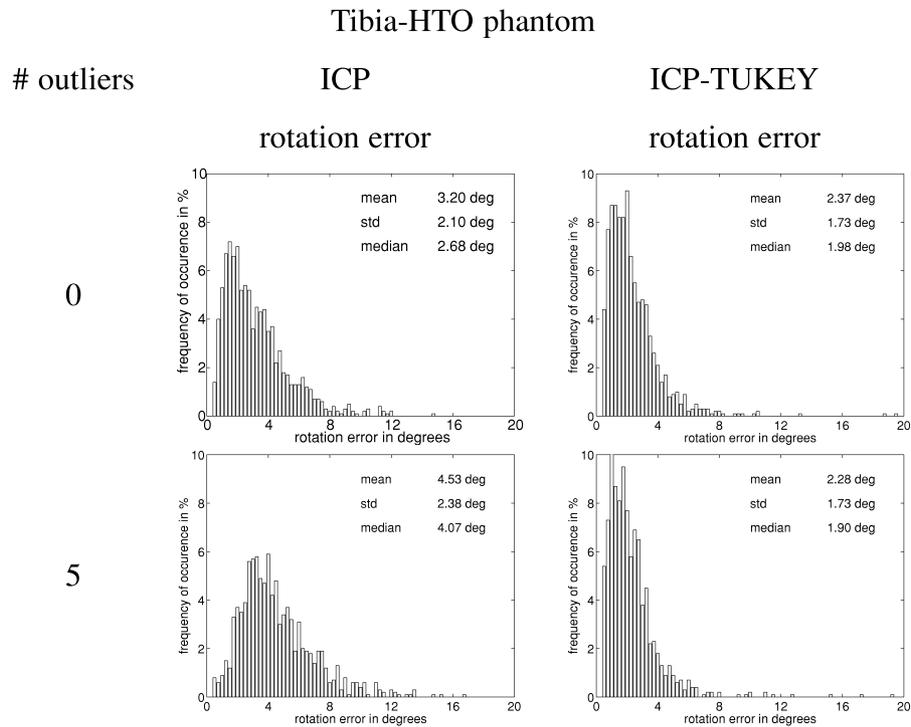
Tibia-HTO phantom

Fig. 3. Rotation error distributions of the ICP and ICP–Tukey algorithms for the tibia-HTO phantom.

median results were good. This was because our surface model included the entire exterior surface of the phantom rather than just the plausible surgically accessible surfaces. For the vertebral phantom, this is problematic because the *anterior* surfaces of the transverse processes and the spinal canal are valid surfaces for the registration algorithm. This situation occurred when the initial spotlight estimate of the registration was poor. Fig. 7 shows an example where this
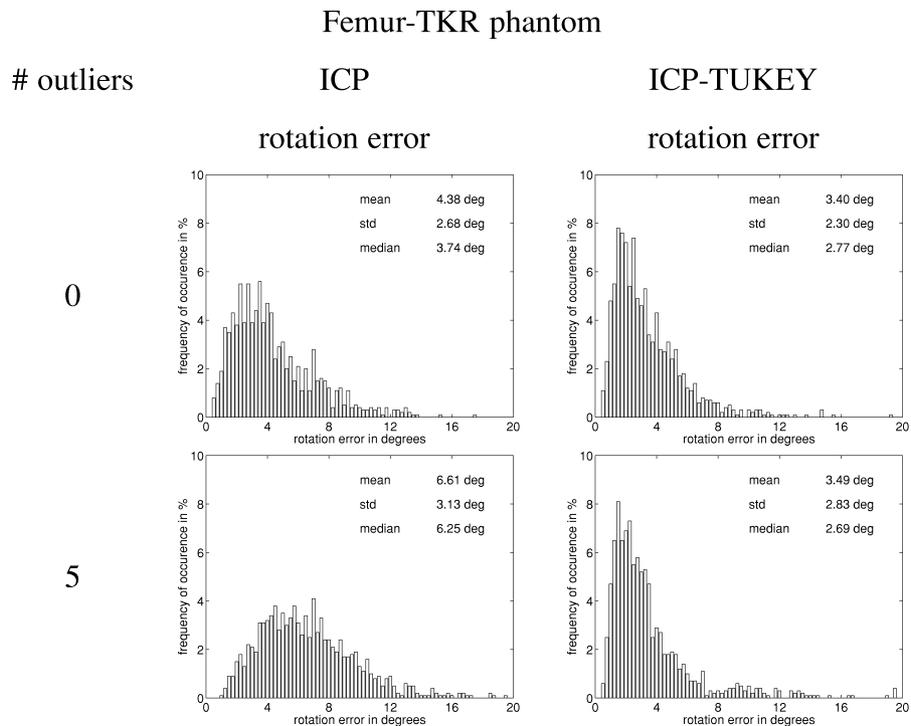
Femur-TKR phantom

Fig. 4. Rotation error distributions of the ICP and ICP–Tukey algorithms for the femur-TKR phantom.
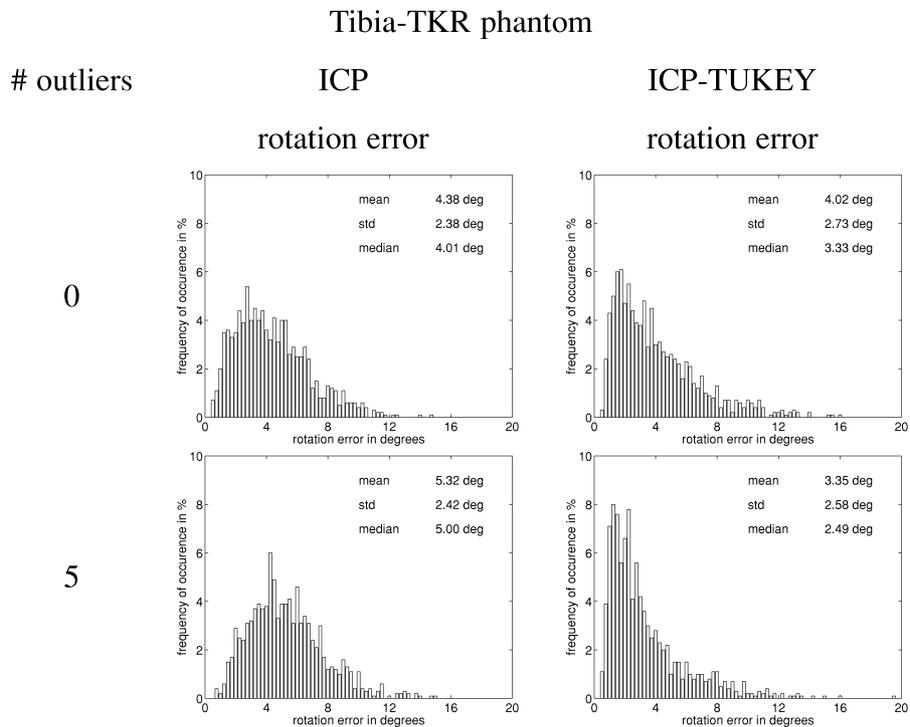
Tibia-TKR phantom

Fig. 5. Rotation error distributions of the ICP and ICP–Tukey algorithms for the tibia-TKR phantom.

problem is further compounded by the ability of the robust estimator to disregard points.

Another noteworthy point is that the surface model must be comprised of only the exterior surface of the object. A

naive application of an isosurface algorithm, such as the threshold-based marching-cubes algorithm (Lorensen and Cline, 1987), typically results in a model with internal surfaces. For example, the internal interfaces between
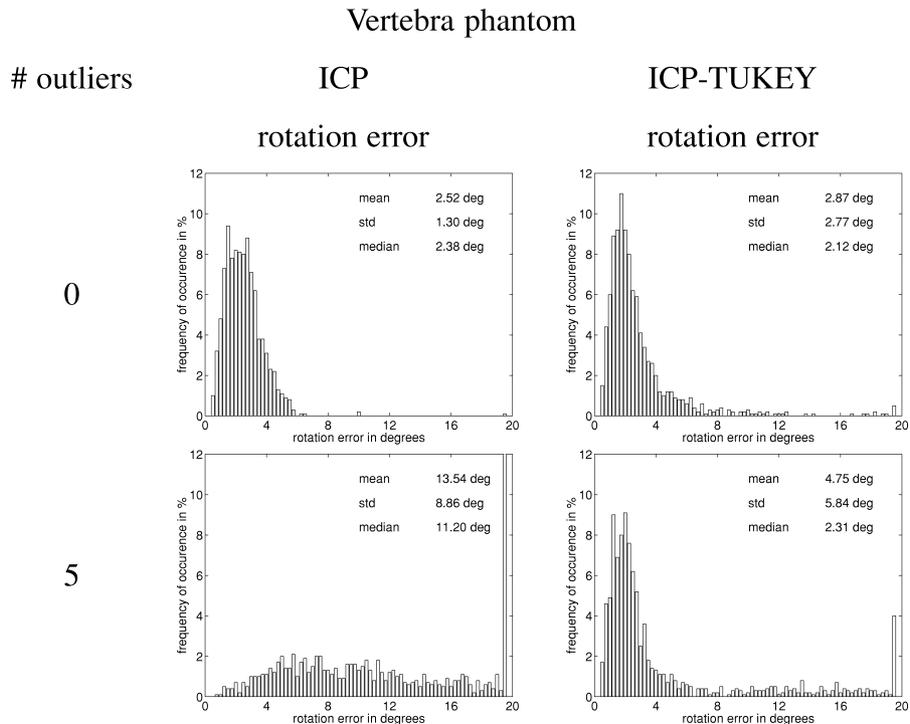
Vertebra phantom

Fig. 6. Rotation error distributions of the ICP and ICP–Tukey algorithms for the vertebra phantom.

Table 4
Rotation error statistics in degrees for plastic phantom registration experiments

Phantom name	# Outliers	ICP			ICP–Tukey		
		Mean	S.D.	Median	Mean	S.D.	Median
Tibia-hto	0	3.20	2.10	2.68	2.37	1.73	1.98
	1	3.76	2.29	3.21	2.40	1.62	2.03
	3	4.46	2.55	3.98	2.41	1.89	2.07
	5	4.53	2.38	4.07	2.28	1.73	1.90
Femur-tkr	0	4.38	2.68	3.74	3.40	2.30	2.77
	1	5.55	2.68	5.14	3.29	2.26	2.68
	3	6.68	3.26	6.20	3.10	2.09	2.58
	5	6.61	3.13	6.25	3.49	2.83	2.69
Tibia-tkr	0	4.38	2.38	4.01	4.02	2.73	3.33
	1	4.56	2.34	4.13	3.38	2.47	2.55
	3	5.27	2.54	4.87	3.26	2.42	2.56
	5	5.32	2.42	5.00	3.35	2.58	2.49
Vertebra	0	2.52	1.30	2.38	2.87	2.77	2.12
	1	6.71	5.92	5.03	3.02	3.13	2.24
	3	12.03	8.41	9.34	3.84	5.06	2.23
	5	13.54	8.86	11.20	4.75	5.84	2.31

correct (fiducial) registration   incorrect robust registration
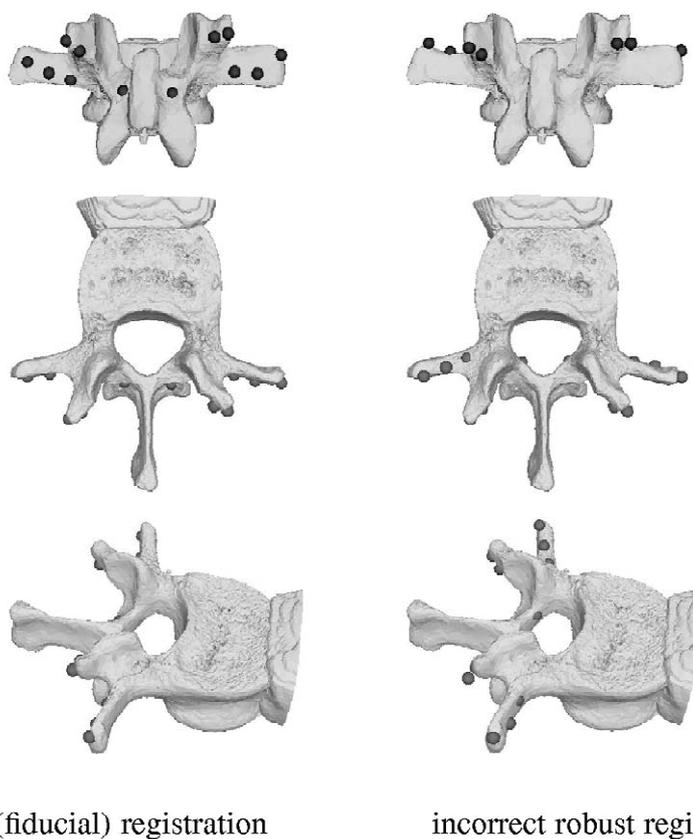
Fig. 7. Two different possible registrations from the same digitized point set for the lumbar vertebra phantom. In the left column are three views of a fiducial registration applied to a set of 12 contact points measured from the posterior aspect of a phantom lumbar vertebra. In the right hand column are three similar views of an incorrect robust registration. The robust algorithm has converged to a registration where points match surfaces from the spinal canal and the anterior aspect of the transverse processes. It has also incorrectly rejected a point as an outlier.

cortical bone and the marrow, and the internal structures of cancellous bone, are typically present in an isosurface model. We removed these non-exterior surfaces prior to estimating the registration.

Our experiments did not quantify the accuracy of the entire process of registration. In general, registration errors are caused by a combination of factors, including inaccuracies in imaging, model generation, digitization of surface points, and the registration algorithm. Our study only included errors in digitization caused by intentional outliers, by the physical accuracy of the measurement equipment, and by the registration algorithm.

Finally, our results bring into question the suitability of a local search method (such as ICP) for surgical registration. On average we obtained good registration results, but the error distributions showed that both ICP and our robust variant could converge to poor registrations.

5. In vivo clinical experience

We have integrated spotlight registration into a custom written image-guided surgical navigation system that has been used to guide surgeons in performing orthopedic extremity procedures. The patients were drawn from existing waiting lists, and freely consented to participate in studies approved by the Research Ethics Board of Kingston General Hospital and Queen's University. Over the past four years we have performed more than 100 cases of a wide variety of procedures.

Surgical technique in each case was similar. A preoperative CT scan was processed, using custom software, to obtain surface models for visualization and registration. Where appropriate, a preoperative plan was made. Intraoperatively, a 3D optoelectronic tracking system was used (OptoTrak, Northern Digital, Waterloo, Canada). A dynamically tracked reference frame was attached to the target bone using 4.0 mm Schantz pins and external-fixation devices (AO/Synthes, Bern, Switzerland). Registration data were collected with a preoperatively calibrated, optically tracked probe. Registration was validated by the operating surgeon(s), who contacted distinctive anatomical features when possible and ensured that the computed point in the CT scans and the surface models corresponded to the anatomy.

This section describes our results for three different clinical applications: closing-wedge high tibial osteotomy, distal radius osteotomy, and excision of deep bone tumors. For the osteotomies we report the postoperative alignment as measured from plain radiographs and, for the wrist, from postoperative functional evaluation. For the tumor excision we report postoperative clinical evaluation and pathology confirmation of navigation accuracy.

5.1. High tibial osteotomy

Closing-wedge high tibial osteotomy is a procedure used to treat a relatively young or active patient having osteoarthritis that is confined to the medial compartment of the knee. The goal of this procedure is to correct the mechanical axis of the leg so that load is carried mainly by the unaffected lateral compartment of the knee. This is accomplished by removing a wedge of bone from the lateral side of the proximal tibia and closing the wedge like a hinge. In the modified Coventry technique for this procedure, one or more Kirschner wires are drilled into the proximal tibia to define one or both planes of the wedge. The correct placement of the wires is often confirmed fluoroscopically, and then the wires are used as guides for sawing the bone.

Our approach to computer-integrated high tibial osteotomy was to plan the correction preoperatively and then to use an optically tracked surgical drill to implant the Kirschner wires into the planned positions. Details of our planning software (Tso et al., 1998) and in vitro results (Ellis et al., 1999) have been previously reported. Our in vitro studies indicated a reduction in error magnitude of over 50% ($p < 0.05$).

In a consecutive series of 15 patients, we used our robust registration algorithm to implant the guide wires. A radiologist measured the actual angle between the tibial plateau and tibial shaft from preoperative and postoperative frontal-plane X-ray films. Comparing the achieved angle to the planned angle, we found that the maximum error was $\pm 2°$. In one case the superior plane was located 5 mm more proximal than planned, and in one case the computer technique was abandoned because the registration was deemed inaccurate. The suspected problem in the inaccuracy was inadvertent motion of the dynamic reference body attached to the tibial shaft by Schantz pins.

5.2. Distal radius osteotomy

Fractures of the distal radius are common, constituting approximately 15% of all fractures seen in the emergency room. A malunited fracture often leads to pain, reduced range of motion, reduced strength of the wrist, and arthritic changes. Distal radius osteotomy may be performed to correct such a malunion. Traditional technique requires a freehand bone cut near the site of the original fracture, followed by visual alignment using X-ray fluoroscopy. The alignment process is greatly complicated by the soft tissue contracture that often accompanies malunions and by the 2D nature of the X-ray images. Once the desired alignment of the distal bone fragment is achieved, a trapezoidal bone graft or substitute is fashioned to fill the bone gap. A fixation plate is contoured to fit the shape of the radius and secured with bone screws.

Our approach to computer-integrated distal radius osteotomy required CT scans of both the malunited and unaffected wrists of the patient. Isosurface models of the two wrists were generated from the scans. The healthy wrist model was reflected as a mirror image, to serve as a template for the correction of the malunion. The template

and model of the affected wrist were aligned and then the model was cut with a virtual osteotomy to produce the proximal and distal fragments of the radius. The distal radial fragment was aligned with the template of the healthy distal radius. When the surgeon was satisfied with the new alignment of the deformed distal fragment, a model of the fixation plate was placed on the models so that the plate would hold the bone fragments in place. The locations of the pilot holes for the plate were saved relative to the uncut deformed radius.

Intraoperatively, the surgeon drilled the pilot holes with an optically tracked drill and used the optically tracked probe to determine the plane of the osteotomy. The distal fragment was shaved to permit the plate to fit as per the plan. The plate was then attached to the distal fragment and the plate/fragment assembly was progressively distracted. When the actual holes in the plate matched the pilot holes in the bone, the alignment was complete and the defect was filled with viable autologous bone. Details of our planning software and in vitro results have been previously reported in (Croitoru et al., 2001). Our in vitro studies indicated a reduction in error magnitude of over 50% ($p < 0.01$).

In a consecutive series of six patients, we used the computer-integrated technique to perform distal radius osteotomy and conducted detailed postoperative clinical evaluation as described in (Athwal et al., 2002). At an average follow-up of 25 months the patients underwent evaluation that included functional tests. In addition, true posteroanterior and true lateral radiographs of the wrist were assessed in a blinded fashion by one independent observer.

All patients were pleased with the surgical outcome and would have the procedure again in a similar situation. All reported decreased pain and improved functionality and cosmesis. Functionally, the average postoperative range of motion measured 87% of the motion of the contralateral wrist. The average grip strength was 30 kg compared to 38 kg in the contra-lateral hand. The radiographic indices used to assess correction improved dramatically: radial inclination improved from 12° preoperatively to 21° postoperatively (normally 22), ulnar variance improved from $+7.5$ mm preoperatively to $+1.9$ mm postoperatively (contralateral mean $+1.5$ mm), and volar tilt improved from $-30°$ for dorsal malunions and $+20°$ for volar malunions preoperatively to 9° (normally 11°) postoperatively. Fig. 8 shows representative results.

5.3. Osteoid osteoma excision

An osteoid osteoma is a small, benign, painful osteoblastic lesion of cortical bone. Osteoid osteoma is most frequently observed in young individuals. If the lesion causes pain that does not respond adequately to nonsteroidal anti-inflammatory medications, or if the patient is
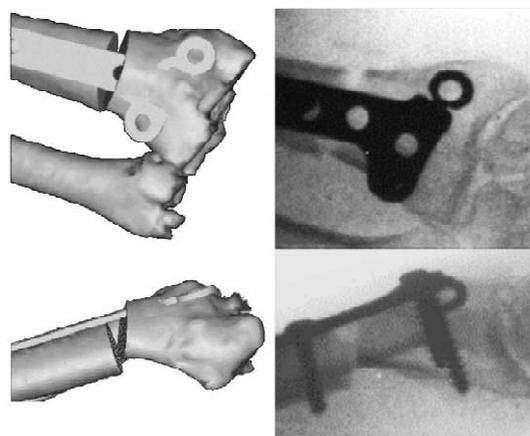
Fig. 8. A preoperative plan, and postoperative fluoroscopic images, for the second patient in the pilot clinical study.

unable to tolerate such medication, then the lesion must be entirely excised.

A variety of operative techniques have been recommended, with en bloc resection traditionally being the technique of choice (Canale, 1998). En bloc resection of a substantial mass of bone requires that the defect be filled with graft and fixated with wires, staples, screws or plates. With cortical lesions, there is an appreciable risk of subsequent fracture near the operative site.

We have operated on three consecutive patients (21, 19 and 14 years old) diagnosed with osteoid osteoma (Ellis et al., 2001). Our computer-integrated technique for this procedure involved semitransparent rendering of the cortical bone, for improved visualization of the lesion, and percutaneous registration for minimally invasive excision. The semitransparent rendering showed the nidus of the lesion in bright red (as it appears in the actual bone), and the supralesional cortical bone in light grey. Fig. 9 shows the cortical bone and a semitransparent rendering of the lesion area of the second case in the pilot clinical study.

We registered these cases percutaneously, by intraoperatively calibrating a tracked probe that had a 16 gauge hypodermic needle as its tip. In the second case, spotlight regions were chosen in the malleoli and on the anterior aspect of the bulge over the lesion site. Additional data were collected percutaneously and in the surgical exposure. Fig. 10 shows the surface model and the registered data.

The computer then tracked the drill and superimposed an image of the drill on the visualization model and on axial, sagittal and coronal reformats of the CT scan. Fig. 11 shows a typical intraoperative display; as per radiological convention, the CT axial slices are displayed as though the patient was prone.

In each case the osteoma nidus was entirely excised and pathological examination confirmed that the lesion was an osteoid osteoma. The surgical wounds were closed with
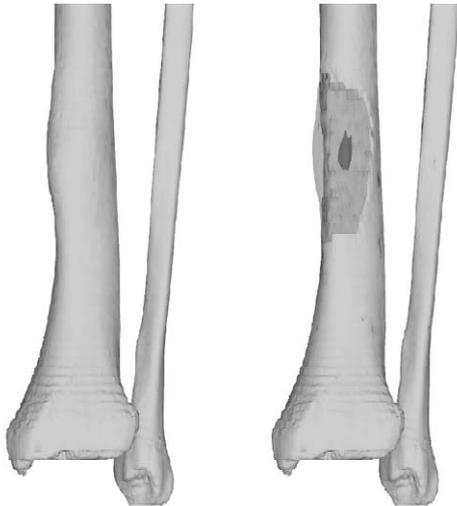
Fig. 9. Surface rendering and semitransparent rendering of the site of an osteoid osteoma in the posteromedial right tibia of a 19-year-old female. Note the absence of distinctive landmarks, which presented a considerable challenge for surface-based registration.

SteriStrips, and each patient was discharged within 24 h. Each patient had a complete and uneventful recovery, and on postoperative evaluation 2 weeks following surgery were pain-free.

6. Conclusions

Registration between the patient and the model is the mathematical cornerstone for computer-integrated surgery guided by preoperative medical images. For orthopedics, the principal sources of registration data are derived from physical contact with bone surfaces and models derived from CT images. Errors introduced by the process of
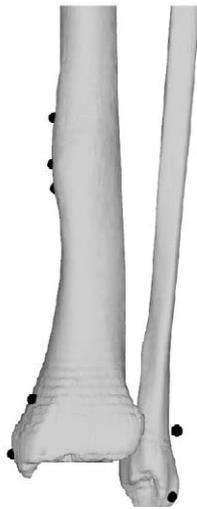
Fig. 10. Registration of percutaneously collected points to the distal tibia. Note that one point was rejected and was considerably distant from the bone surface (radius of spheres used to render points is 3 mm).

collecting these data do not necessarily follow a Gaussian distribution, so some technique for managing statistical outliers is needed.

Our work used a specific sequence of robust estimators to estimate the registration. The first step was to collect data from generally recognizable landmark regions or spotlights, and to compute from these initial points an initial registration. Further data points were then used to improve the estimate of the registration while minimizing the influence of points that were far from the model. A laboratory study, which included fiducial markers as a ground-truth reference, confirmed that the estimator sequence could accurately find a registration with submillimeter root-mean-square error in the presence of spurious data.

Many researchers have described CIS applications that use ICP with k–d trees for registration purposes (Section 2). We used a perturbation method to obtain a robust initial estimate of the registration, but the Hausdorff distance (Rucklidge, 1996) or absolute difference distance (Hagedoorn and Veltkamp, 1999) could also be used to find an initial robust estimate on a relatively coarse spatial subdivision. Although our algorithm has proven to be sufficiently fast for intraoperative use, alternative acceleration methods (Rusinkiewicz and Levoy, 2001; Greenspan and Godin, 2001) may also be used. To estimate the registration, methods such as RANSAC (Fischler and Bolles, 1981) and LMS (Kumar and Hanson, 1990; Masuda and Yokoya, 1995) algorithms are known to be robust to a greater fraction of outliers than is our M-estimator but the combinatorial time complexity of these algorithms is too slow for many intraoperative applications. Finally, the statistical methods described by Rangarajan et al. (1997), Rangarajan et al. (1999), Chui and Rangarajan (2000), Granger et al. (2001) and Dellaert (2001) claim to solve the problem of convergence to a non-global minimum of the registration objective function and are robust to statistical outliers. We cannot comment on their suitability for surgical guidance.

The robust method appeared to be clinically useful in achieving registration, particularly in cases where the exposure was very limited or where the anatomical target had a relatively featureless surface. In a series of 21 clinical cases the registration process failed only once, and in that case the surgical procedure was successfully completed by conventional technique. Each of the more than 100 of the computer-integrated cases performed to date have had very good or excellent technical outcomes, which suggests that errors due to registration were negligible. This is consistent with our laboratory findings.

Of particular note were the successful registrations in the cases of excision of osteoid osteoma from the tibia. The data were collected percutaneously, by tracking a thin deformable hypodermic needle that pierced substantial thickness of soft tissue before contacting the bone surface. Our robust estimation procedure was able to automatically

Fig. 11. Image guidance for drilling to intraoperatively localize an osteoma nidus.

detect and discard spurious data, leading to successful guidance and removal of these painful lesions. The clinical significance was changing a time-consuming surgical procedure, usually followed by a lengthy hospital stay and recovery, into a minimally invasive procedure with prompt discharge of the patient and complete, uneventful recovery.

One observation we made while developing our registration algorithm is that registration is a *process*, not a mathematical calculation. It is important to consider the sources of data, how they are collected, what sources of error may occur, and how to compensate for error. It is also critical that the user, here an orthopedic surgeon, be involved in the ergonomics and the human–computer interface throughout the development and testing of the registration process. Ultimately the surgeon is in control of the case, so the registration process must be both convenient and must clearly communicate to the surgeon how best to proceed. Our experience was that surgeons could readily identify spotlight regions on images and on patients, which contributed to the success in achieving useful registrations.

We have identified shortcomings in our work and opportunities for improvements. Most important is that we used a local estimation process and the calculated registration was not guaranteed to be the global minimal, correct registration. Our sequence of estimators cannot be demonstrated to be optimal, and thus represented only one possible means of converging on a locally minimal registration. There also may be other ways of achieving an initial registration other than by generalizing the concept of the anatomical landmark.

As long as preoperative medical images are used to plan and guide surgical procedures, registration of the patient and the images will be needed. Robust registration methods may be applicable to many other intraoperative sensing modalities, such as 3D ultrasound data or fluoroscopic images, and to applications other than orthopedics. What this work demonstrates is that robust registration provides a clinically convenient and effective way to treat a variety of problems of bones and joints with a computer-integrated surgical technique.

References

Athwal, G.S., Ellis, R.E., Small, C.F., Pichora, D.R., 2002. Outcomes of computer-assisted distal radius osteotomy. J. Hand. Surg. (accepted for publication).

Bächler, R., Bunke, H., Nolte, L.-P., 2001. Restricted surface matching—numerical optimization and technical evaluation. Comput. Aided Surg. 6 (3), 143–152.

Besl, P.J., McKay, N.D., 1992. A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. 14 (2), 239–256.

Betting, F., Feldmar, J., Ayache, N., Devernay, F., 1995. A new framework for fusing stereo images with volumetric medical images. In: Ayache, N. (Ed.), Proceedings of the International Conference on Computer Vision, Virtual Reality and Robotics in Medicine. Springer, Berlin, pp. 30–39.

Blais, G., Levine, M.D., 1995. Registering multiview range data to create 3D computer objects. IEEE Trans. Pattern Anal. 17, 820–824.

Canale, S.T., 1998. Campbell's Operative Orthopaedics, Vol. 1. Mosby-Year Book, St. Louis.

Chui, H., Rangarajan, A., 2000. A feature registration framework using mixture models. In: Proceedings of the IEEE Workshop on Mathematical Methods and Biomedical Image Analysis, pp. 190–197.

Croitoru, H., Ellis, R.E., Small, C.F., Prihar, R., Pichora, D.R., 2001. Fixation-based surgery: A new technique for distal radius osteotomy. Comput. Aided Surg. 6 (3), 160–169.

Cuchet, E., Knoplioch, J., Dormont, D., Marsault, C., 1995. Registration in neurosurgery and neuroradiotherapy applications. In: Proceedings of the International Symposium on Medical Robotics and Computer Assisted Surgery. Wiley-Liss, New York, pp. 31–38.

Dellaert, F., 2001. Monte Carlo EM for data-association and its application in computer vision. PhD thesis, Carnegie Mellon University.

Ellis, R.E., Toksvig-Larsen, S., Marcacci, M., Caramella, D., Fadda, M., 1996. Use of a biocompatible fiducial marker in evaluating the accuracy of CT image registration. Invest. Radiol. 31 (10), 658–667.

Ellis, R.E., Fleet, D.J., Bryant, J.T., Rudan, J., Fenton, P., 1997. A method for evaluating CT-based surgical registration. In: Troccaz, J., Grimson, E., Mösges, R. (Eds.), Proceedings of the International Conference on Computer Vision, Virtual Reality and Robotics in Medicine. Springer, Berlin, pp. 141–150.

Ellis, R.E., Tso, C.Y., Rudan, J.F., Harrison, M.M., 1999. A surgical planning and guidance system for high tibial osteotomy. Comput. Aided Surg. 4 (5), 264–274.

Ellis, R.E., Kerr, D., Rudan, J.F., Davidson, L., 2001. Minimally invasive excision of deep bone tumors. In: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, Berlin, pp. 1154–1156.

Feldmar, J., Ayache, N., 1994. Rigid and affine registration of smooth surfaces using differential properties. In: Eklundh, J.-O. (Ed.). Proceedings of the European Conference on Computer Vision, Vol. 2. Springer, Berlin, pp. 397–406.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24 (3), 381–395.

Granger, S., Pennec, X., Roche, A., 2001. Rigid point-surface registration using an EM variant of ICP for computer guided oral implantology. In: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 752–761.

Greenspan, M.A., Godin, G., 2001. A nearest neighbor method for efficient ICP. In: Proceedings of the International Conference on 3D Digital Imaging and Modeling, pp. 161–168.

Grimson, W.E.L., Ettinger, G.J., White, S.J., Gleason, P.L., Lozano-Pérez, T., Ill, W.M.W., Kikinis, R., 1995. Evaluating and validating an automated registration system for enhanced reality visualization in surgery. In: Ayache, N. (Ed.), Proceedings of the International Conference on Computer Vision, Virtual Reality and Robotics in Medicine. Springer, Berlin, pp. 3–12.

Hagedoorn, M., Veltkamp, R.C., 1999. Reliable and efficient pattern matching using an affine invariant metric. Int. J. Comput. Vis. 31 (2/3), 203–225.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. Robust Statistics. The Approach Based on Influence Functions. Wiley, New York.

Haralick, R.M., Joo, H., Lee, C., Zhuang, X., Vaidya, V.G., Kim, M.B., 1989. Pose estimation from corresponding point data. In: Freeman, H. (Ed.), Machine Vision for Inspection and Measurement. Academic Press, New York, pp. 1–84.

Henri, C.J., Colchester, A.C.F., Zhao, J., Hawkes, D.J., Hill, D.L.G., Evans, R.L., 1995. Registration of 3-D surface data for intra-operative guidance and visualization in frameless stereotactic neurosurgery. In: Ayache, N. (Ed.), Proceedings of the International Conference on Computer Vision, Virtual Reality, and Robotics in Medicine. Springer, Berlin, pp. 47–56.

Herring, J.L., Dawant, B.M., Maurer Jr, C., Muratore, D.M., Galloway, R.L., Fitzpatrick, J.M., 1998. Surface-based registration of CT images to physical space for image-guided surgery of the spine: A sensitivity study. IEEE Trans. Med. Imaging 17 (5), 743–752.

Hoaglin, D.C., Mosteller, F., Tukey, J.W. (Eds.), 1983. Understanding Robust and Exploratory Data Analysis. Wiley, New York.

Horn, B.K.P., 1987. Closed-form solution of absolute orientation using unit quaternions. J. Opt. Soc. Am. A 4 (4), 629–642.

Huber, P.J., 1981. Robust Statistics. Wiley, New York.

Kumar, R., Hanson, A.R., 1990. Analysis of different robust methods for pose refinement. In: Proceedings of the International Workshop on Robust Computer Vision, pp. 167–182.

Lavallée, S., 1995. Registration for computer integrated surgery: Methodology, state of the art. In: Taylor, R.H., Lavallée, S., Burdea, G.C., Mösges, R. (Eds.), Computer Integrated Surgery. MIT Press, Cambridge, MA, pp. 77–98.

Lorensen, W.E., Cline, H.E., 1987. Marching cubes: A high resolution 3D surface construction algorithm. Comput. Graph. 21 (4), 163–169.

Luck, J., Little, C., Hoff, W., 2000. Registration of range data using a hybrid simulated annealing and iterative closest point algorithm. In: Proceedings of the IEEE International Conference on Robotics and Automation, pp. 3739–3744.

Ma, B., Ellis, R.E., Fleet, D.J., 1999. Spotlights: A robust method for surface-based registration in orthopedic surgery. In: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention. Springer, Berlin, pp. 936–944.

Maintz, J.B.A., Viergever, M.A., 1998. A survey of medical image registration. Med. Image Anal. 2 (1), 1–36.

Masuda, T., Yokoya, N., 1995. A robust method for registration and segmentation of multiple range images. Comput. Vis. Image Underst. 61 (3), 295–307.

Maurer, C.R., Fitzpatrick, J.M., Maciunas, R.J., Allen, G.S., 1996. Registration of 3-D images using weighted geometrical features. IEEE Trans. Med. Imaging 15 (6), 836–849.

Maurer, Jr. C., Maciunas, R.J., Fitzpatrick, J.M., 1998. Registration of head CT images to physical space using a weighted combination of points and surfaces. IEEE Trans. Med. Imaging. 17 (5), 753–761.

Meer, P., Mintz, D., Rosenfeld, A., 1991. Robust regression methods for computer vision: A review. Int. J. Comput. Vis. 6 (1), 59–70.

Mirza, M.J., Boyer, K.L., 1993. Performance evaluation of a class of m-estimators for surface parameter estimation in noisy range data. IEEE Trans. Robotic Autom. 9 (1), 75–85.

Palombara, P.F.L., Fadda, M., Martelli, S., Nofrini, L., Marcacci, M., 1997. A minimally invasive 3-D data registration protocol for computer and robot assisted total knee arthroplasty. In: Troccaz, J., Grimson, E., Mösges, R. (Eds.), Proceedings of the International Conference on Computer Vision, Virtual Reality and Robotics in Medicine. Springer, Berlin, pp. 663–672.

Rangarajan, A., Chui, H., Bookstein, F.L., 1997. The softassign procrustes matching algorithm. In: Duncan, J., Gindi, G. (Eds.), Information Processing and Medical Imaging, pp. 29–42.

Rangarajan, A., Chui, H., Duncan, J.S., 1999. Rigid point feature registration using mutual information. Medical Image Analysis 3 (4), 425–440.

Rousseeuw, P.J., Leroy, A.M., 1987. Robust Regression and Outlier Detection. Wiley, New York.

Rucklidge, W., 1996. Efficient Visual Recognition Using the Hausdorff Distance. Springer, Berlin.

Rusinkiewicz, S., Levoy, M., 2001. Efficient variants of the ICP algorithm. In: Proceedings of the International Conference on 3D Digital Imaging and Modeling, pp. 145–152.

Schroeder, W., Martin, K., Lorensen, B., 1998. The Visualization Toolkit, 2nd Edition. Prentice Hall, Upper Saddle River, NJ.

Simon, D., Hebert, M., Kanade, T., 1995. Techniques for fast and accurate intra-surgical registration. J. Image Guided Surg. 1, 1.

Tso, C.Y., Ellis, R.E., Rudan, J., Harrison, M.M., 1998. A surgical planning and guidance system for high tibial osteotomies. In: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 39–50.

Zhang, Z., 1994. Iterative point matching for registration of free-form curves and surfaces. Int. J. Comput. Vis. 13 (2), 119–152.