A Robust Method for Registration of Three-Dimensional Knee Implant Models to Two-Dimensional Fluoroscopy Images

Mohamed R. Mahfouz*, William A. Hoff, Richard D. Komistek, and Douglas A. Dennis

Abstract—A method was developed for registering three-dimensional knee implant models to single plane X-ray fluoroscopy images. We use a direct image-to-image similarity measure, taking advantage of the speed of modern computer graphics workstations to quickly render simulated (predicted) images. As a result, the method does not require an accurate segmentation of the implant silhouette in the image (which can be prone to errors). A robust optimization algorithm (simulated annealing) is used that can escape local minima and find the global minimum (true solution). Although we focus on the analysis of total knee arthroplasty (TKA) in this paper, the method can be (and has been) applied to other implanted joints, including, but not limited to, hips, ankles, and temporomandibular joints. Convergence tests on an *in vivo* image show that the registration method can reliably find poses that are very close to the optimal (i.e., within $0.4°$ and 0.1 mm), even from starting poses with large initial errors. However, the precision of translation measurement in the Z (out-of-plane) direction is not as good. We also show that the method is robust with respect to image noise and occlusions. However, a small amount of user supervision and intervention is necessary to detect cases when the optimization algorithm falls into a local minimum. Intervention is required less than 5% of the time when the initial starting pose is reasonably close to the correct answer, but up to 50% of the time when the initial starting pose is far away. Finally, extensive evaluations were performed on cadaver images to determine accuracy of relative pose measurement. Comparing against data derived from an optical sensor as a "gold standard," the overall root-mean-square error of the registration method was approximately $1.5°$ and 0.65 mm (although Z translation error was higher). However, uncertainty in the optical sensor data may account for a large part of the observed error.

Index Terms—3-D to 2-D registration, X-ray fluoroscopy, TKA knee implants, simulated annealing.

I. INTRODUCTION

TOTAL KNEE ARTHROPLASTY (TKA) is a common operation in which the knee joint is replaced with artificial implants. The implant consists of two metallic components that
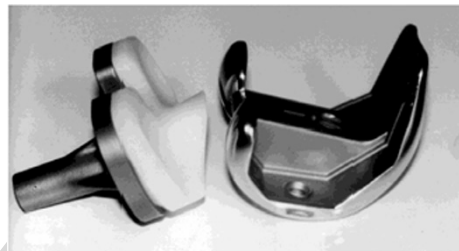
Fig. 1. Artificial knee implant, with tibial component (left) and femoral component (right). The white material is a polyethylene insert.

Fig. 2. Fluoroscopy image of *in vivo* TKA.

replace the bearing surfaces on the tibia and femur, separated by a high molecular weight polyethylene insert (Fig. 1). While many evaluations have shown excellent relief of pain and improved function, there are still problems associated with premature failure [1], [2]. It is believed that abnormal kinematics of implanted knees may lead to excessively high shear stresses on the polyethylene inserts, thus accelerating wear [3]. More knowledge of *in vivo* implant kinematics may allow implants to be designed that have less polyethylene wear.

Recently, X-ray fluoroscopy has been shown to be a useful tool for analyzing joint kinematics *in vivo* [4], [5]. The fluoroscopic process creates a perspective projection, where the metallic implants appear much darker than the soft tissues surrounding them (Fig. 2), allowing for direct observation and analysis of the implant components' silhouettes and their movements. Unlike methods that optically track skin-mounted markers, there is no error due to soft-tissue motion, since the components are observed directly.

There are many advantages of fluoroscopy as a measurement tool over previous methods. Joint kinematics can be measured *in vivo* under dynamic, weight-bearing activities. This is important
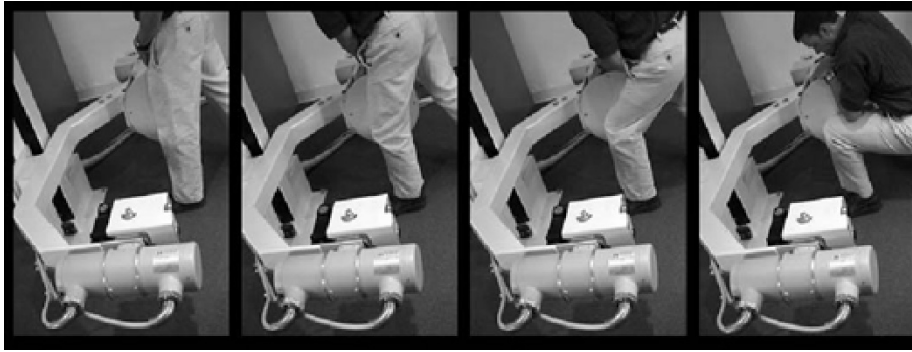
Fig. 3.   Single plane fluoroscopy allows the patient free motion in the plane between the X-ray source and the image intensifier.

in order to observe the effect of muscle loading and soft-tissue constraints. Many past evaluation techniques do not provide either *in vivo* or dynamic capabilities. These have included cadaveric simulations [6], optically tracked skin-mounted markers [7], externally worn goniometric devices [8], and static X-rays with tantalum bead markers [9]. Fluoroscopy is also noninvasive and relatively low risk to the patient. A typical measurement protocol of one minute gives the patient a radiation exposure on the order of 1.8 to 3.6 "rad equivalent man" (rem).

Since we wish to measure kinematics during activities such as gait, stair step, and chair rise, the patient's movement must be sufficiently unconstrained to allow them to perform these activities unimpaired. We use single-plane fluoroscopy because it allows the patient free motion in the plane between the X-ray source and the image intensifier (Fig. 3). Bi-planar fluoroscopy, using two orthogonal units, may lead to more accurate results, but would unacceptably constrain the motion of the patient.

Although fluoroscopy images are only two-dimensional (2-D) images, we can recover all six degrees of freedom (DOFs) of the pose; i.e., three translational (XYZ) and three rotational angles (roll, pitch, yaw). We can do this if we have an accurate geometric model of the object (implant component). Also it is important to have an accurate model of the imaging sensor, from which the image was formed. This is based on the fact that, given the model of the object and the model of the image formation process, the appearance of the object in the image can be predicted (Fig. 4). The predicted image is dependent on all six DOFs of the pose. For example, moving the object away from the sensor reduces the size of the object in the image (for a perspective projection imaging model). By searching the space of possible poses, one can find the pose for which the predicted image best matches the actual image of the object. When the poses of both the femoral and the tibial implant component models have been determined (with respect to the fluoroscope), the relative pose between the models can then be calculated. By repeating this process for each image (or for selected images) of a fluoroscopic sequence, we can reconstruct the kinematics of the joint during a complete motion cycle (gait, stair rise, deep knee bend, etc.).

Challenges that arise in this problem domain include noise, clutter, occlusions, and low object-to-background contrast. Clutter arises from extraneous objects in the image with appearance similar to the implant components. For example, metal calibration objects [such as the steel ball in Fig. 5(b)] and
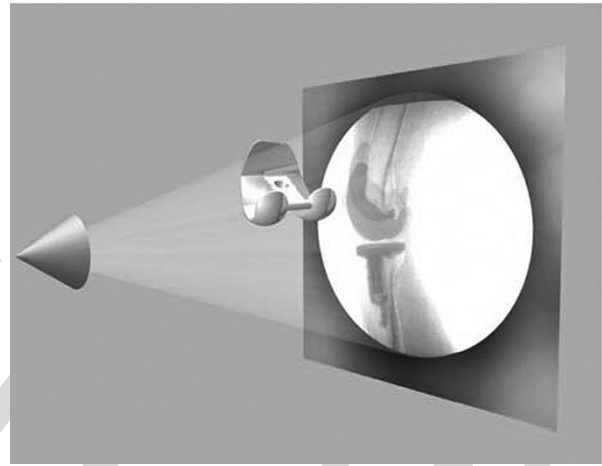

Fig. 4.   Using a perspective projection imaging model, the silhouette of the model can be predicted and compared with the observed silhouette in the image.

(a)   (b)
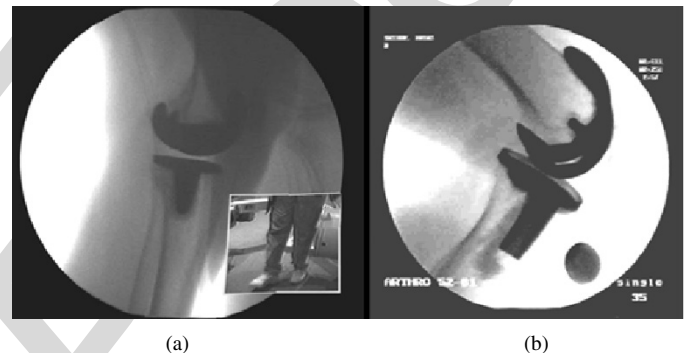
Fig. 5.   Occlusions and low object-to-background contrast may occur when the projection of the implant component overlaps (a) the other leg in the image or (b) another implant.

screws have high contrast, similar to the implant components. Occlusions and low object-to-background contrast may occur when the projection of the implant overlaps the other leg [Fig. 5(a)], or another implant [Fig. 5(b)]. Low object-to-background contrast may also occur when there is adjacent material with similar contrast, such as the bone cement next to the tibial component in Fig. 2. In these cases, it is difficult to automatically extract a complete contour of the object.

This paper describes a new method for measuring the kinematics of TKA knees from single plane fluoroscopy images. Our method is robust with respect to image noise, occlusions,

and low object-to-background contrast. Unlike previous work in this area, we do not require an accurate segmentation of the implant silhouette in the image (which can be prone to errors). Instead, we use a direct image-to-image similarity measure, such as has been used in work on computed tomography (CT)-to-fluoroscopy registration [10], [11]. In this approach, a synthetic fluoroscopy image of the implant in a predicted pose is generated, and this image is correlated to the original input image. Although this method avoids explicit segmentation, it can result in numerous local minima that can lead to false registration solutions. We avoid this problem by using a robust optimization algorithm (simulated annealing) that can escape local minima and find the global minimum (true solution). Although we focus on knees in this paper, the method can be (and has been) applied to other joints, including hips, ankles, and temporomandibular joints (TMJ).

The rest of this paper is organized as follows. Section II provides a review of previous work in this area. Section III describes our new method in detail. Section IV gives the results of performance analyses on accuracy, reliability, and convergence rate. Section V provides a discussion and conclusion.

II. RELATED WORK

The system described in this paper is an example of registering a three-dimensional (3-D) model to a 2-D image. The problem of determining the position and orientation (pose) of a known 3-D object from a single 2-D image is a common problem in the field of computer vision. Typically, one starts with a known geometric model of the object, and a known model of the image formation process. The object is assumed to lie somewhere in the image, although it is not known which image features belong to the object of interest, and which features arise from other objects or structures in the scene (clutter). The problem is to identify the features that belong to the object of interest, and estimate the six-DOF pose of that object with respect to the sensor.

There are three approaches for determining the position and orientation (pose) of a known 3-D object from a single 2-D image. The first method, used frequently in robotic and machine vision, is based on identifying individual 2-D features in the image and matching them to 3-D features on the model [12]. These are usually point-type features (such as holes, protrusions, or distinctive markings) or line-type features (such as long straight edges). The correspondence between model features and image features can be determined using a tree search or with a hashing scheme (e.g., Hough transform). However, this approach is difficult to use in our problem domain because individual distinct features are difficult to extract. One reason is that only the silhouettes or extremal contours of the implant components are visible in the X-ray images, with no internal features or surface markings showing. Another reason is that the objects typically have smooth curved surfaces, and there are few (if any) easily recognizable features along the silhouette (such as a corner).

The second approach is to match the exterior surface of the object to the projected silhouette in the image. Methods have been developed for polyhedral models [13] and objects that can be represented with a small number of parameterized surface patches [14]. Other methods precompute a library of expected silhouettes of the object, or templates [5], [15]. Each template is created by graphically rendering the object at a known pose. The input image is then processed to extract a silhouette. The silhouette with the closest match in the library is taken to represent the pose of the object. Alternatively, a "hypothesize-and-test" approach can be used, where a pose is hypothesized, a test is performed to see how well the actual data matches the predicted data, and an optimization algorithm adjusts the pose as necessary [16]–[18]. The cycle is repeated until there is a close match between the predicted data and the actual data. This allows a continuous adjustment of the pose, instead of limiting the adjustments to the resolution of a precomputed library. However, these methods have the disadvantage that the object's contour must be accurately segmented from the image. This may be difficult in some images, due to noise, low contrast, and occlusions.

The third approach is to match the image values directly to a predicted image of the object. A predicted image is generated of the object in a hypothesized pose, and the pixel values are compared directly to the values in the actual input image, without trying to presegment the object from the image. For X-rays, the predicted images are known as digitally reconstructed radiographs. With this approach, a 3-D volumetric model, rather than a surface model, can be used. Researchers have matched 3-D volumetric models derived from CT, magnetic resonance imaging (MRI), or positron emission tomography data to static X-rays [10] or fluoroscopy images [19], [11]. A variety of image difference measures can be used [20], such as pattern intensity [19], gradient difference [11], and cross-correlation [10]. Since the measures are global in nature, they are robust to small amounts of clutter and occlusions. Although past approaches have focused on CT-derived volumetric models, there is no reason why direct image comparison methods could not be used for surface models (i.e., implant models). In fact, this is the approach that we use, as is discussed in Section III.

With all hypothesize-and-test methods, there is a need for an optimization algorithm to adjust the pose of the object until its predicted data matches the actual data. Optimization algorithms search for the best (e.g., minimum) value of a cost function. Many researchers (e.g., [10], [19]) use a local search algorithm such as gradient descent. This is fast but is prone to getting stuck in a local minimum. A hierarchical (i.e., coarse-to-fine) approach can be used, which improves the likelihood of finding the global minimum [21]. Nevertheless, the initial guess for the solution must be fairly close to the actual solution.

Robust optimization algorithms attempt to find the global minimum of a cost function, even in the presence of local minima. Typical algorithms in this category include simulated annealing and genetic algorithms [22], [23]. Although one cannot guarantee that they will find the global minimum, they greatly improve the likelihood of finding the global minimum. The disadvantage of these methods is that they require many function evaluations (i.e., iterations of the hypothesize-and-test loop).

One can improve the likelihood of reaching the true solution by starting from a good initial guess. *A priori* information can be

provided manually by an operator; or in some cases, automatically using domain knowledge. For example, when processing each image in a video sequence, the pose of the object in each image should be fairly close to its pose in the previous image (assuming small velocity).

III. DETAILED DESCRIPTION OF METHOD

Our overall approach is to use a robust optimization algorithm to minimize the error between a predicted and an actual X-ray image. We avoid explicit presegmentation of the object silhouette in the image, since this may be difficult to perform automatically.

The choice of using a metric based on 2-D measurements (rather than 3-D) was motivated by the fact that modern computer graphics workstations (such as the Silicon Graphics, Inc. Octane we used) can very quickly render 2-D images of 3-D objects at video frame rates, even for highly complex models. By doing all computations in 2-D image space, we avoid expensive 3-D computations of ray-to-surface distances.

Although we generate a predicted image of the object in a hypothesized pose, this does not need to be a high fidelity image, which would be expensive to compute. The reason is that most of the information is in the location of the projected contour; the exact values of the image pixels are not necessary to predict.

Our approach incorporates the following elements: 1) an initialization step; 2) a matching algorithm which evaluates the match between the observed image and the predicted image from the current hypothesized pose; 3) a robust optimization algorithm; and 4) a method of supervisory control. These elements are described in the following sections. Preliminary descriptions were given in a conference paper [24] and theses [25], [26].

A. Initialization

Prior to performing 2-D to 3-D registration to estimate pose, we must create geometric models of the objects. Detailed drawings or CAD models are available for most implants; another possibility is to digitize a physical prototype using a laser scanner. The result is a surface model composed of triangular facets, stored in "Open Inventor" format. Although piecewise planar, this model can accurately represent smooth surfaces if enough triangles are used.

The fluoroscope can be modeled by a perspective projection image formation model, which treats the fluoroscope sensor as consisting of an X-ray point source and a planar phosphor screen upon which the image is formed. Although image distortion and nonuniform scaling can occur, these can be compensated for by careful calibration (which only needs to be done once for each fluoroscope). The first step is to estimate any 2-D image geometric distortion. By taking a picture of a known rectangular grid of beads, we can estimate a 2-D spatial transform for each small square subimage that is bounded by four beads. Using standard techniques in geometric distortion removal (e.g., [27]), a local bilinear model is used for the mapping, as well as for the gray level interpolation method. Fig. 6 shows an image of the calibration pattern, before and after distortion removal.

Once the 2-D distortion has been removed, the effective source-to-image plane distance (focal length) can be computed
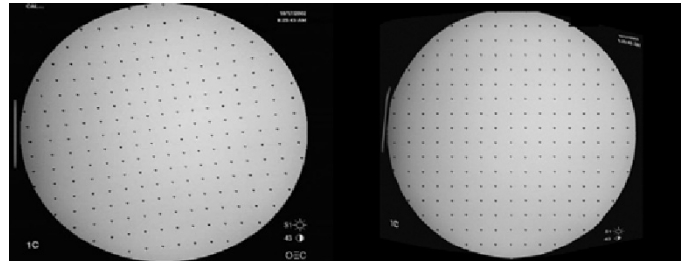

Fig. 6. Fluoroscopic image of calibration grid, before distortion removal (left) and after (right). Note the rotation and "pincushion" effects visible in the left image.

Fig. 7. Predicted rendered image of the femoral model (left) and its silhouette (center). The right image shows an expanded version of the silhouette where pixels are encoded with their closeness to the contour (within a small band).

by a two-plane calibration grid, with a known displacement between the planes (e.g., [5]).

In the experiments shown in this paper, a VF-2000 fluoroscope was used, from Radiographic and Data Solutions, Inc. (Minneapolis, MN). Images were captured using a progressive scan video camera and subsequently digitized to 8 bits and 640×480 pixels using a frame grabber attached to a PC. This fluoroscope had an image intensifier with a diameter of 12 inches, and a principal distance of 1200 mm. The digital images were preprocessed by a 7×7 median filter to reduce noise.

B. Matching Algorithm

The matching algorithm compares two images—the predicted X-ray image and the actual input X-ray image. The predicted X-ray image is rendered using the implant CAD model, using an SGI graphics workstation and the Open Inventor graphics library. Fig. 7(left) shows an example of a rendered image. Using a self-illumination lighting model, the model is rendered as completely white against a black background. The boundary between the white and black regions is then extracted from this image [Fig. 7(center)]. Next, a growing operation is performed, which encodes each pixel within a small distance (3 pixels[1]) of the contour with a score that is inversely proportional to its distance to the contour. This allows points that are near to the contour to contribute to the matching score to an amount that is proportional to their nearness [Fig. 7(right)].

The second input image is the actual X-ray image taken from the fluoroscope. Before matching, this image is inverted so that implant component pixels are white (as in the predicted image). Then an edge enhancement operation (Sobel) is also performed (Fig. 8), to estimate the norm of the local image gradient.

The match between the input X-ray image and the predicted X-ray image is evaluated using weighted combination of two

---

[1]This distance was empirically chosen to improve the capture radius of the algorithm and reduce local minima.
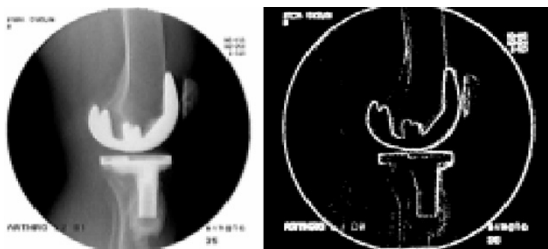
Fig. 8.   Input X-ray image, inverted (left) and gradient image (right).

Fig. 9.   (Left) Matching score for an implant as its pose is rotated about the vertical (y) axis, showing two large minima. (Right) A magnified subset of the curve showing many shallow local minima.

metrics. The first metric compares the pixel values of the two images, and the second metric evaluates the overlap of their contours (edges). Both scores are obtained by multiplying the two images together, summing the result, and normalizing by the sum of the predicted image. If $G(x,y)$ is the input X-ray image [Fig. 8(left)] and $H(x,y)$ is the predicted X-ray image [Fig. 7(left)], then the intensity matching score is

Intensity Matching Score

$$= \sum_{(x,y)} G(x,y)H(x,y) \bigg/ \sum_{(x,y)} H(x,y)$$

This score is similar to a cross correlation between the two images, except that the score is not normalized (since we are interested in finding the maximum of the score, normalization is not necessary). We can interpret our matching score as follows: The model image H is a binary image with nonzero values in the region of the silhouette, and zero values everywhere else. The intensity matching score is the average gray level intensity of the image G inside the projection of the silhouette of the model. This score should be high when the projection of the silhouette covers a bright region in the original image (such as an implant component).

The contour matching score is similarly calculated. If $J(x,y)$ is the input edge-enhanced image [Fig. 8(right)] and $K(x,y)$ is the predicted (expanded) edge image [Fig. 7(right)], then the contour matching score is:

Contour Matching Score

$$= \sum_{(x,y)} J(x,y)K(x,y) \bigg/ \sum_{(x,y)} K(x,y)$$

This score is similar to a cross correlation between the two edge images. The score is a maximum when the peaks in the predicted edge image coincide with the peaks in the input edge image. Our contour matching algorithm is a form of chamfer matching, commonly used in computer vision [28]. Image edge points farther than a certain threshold distance from the hypothesized contour are given zero weight, and thus do not contribute to the matching score. This means that outliers (erroneous data points) do not affect the resulting fit. This method is similar to other outlier rejection strategies such as least-median-squares regression [29] or M-estimators [30].

These two scores are then combined, with the contour matching value weighted more heavily than the area matching value. By weighting the contour score more heavily than the
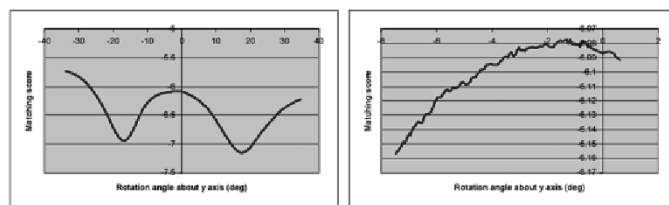
area value, the contour score dominates when the CAD models are close to the true solution. The weights of the intensity and contour scores were set to -1 and -2.67, respectively. These weights were determined experimentally to achieve good results on typical images[2] . The resulting total matching score, or similarity measure, produces a distinct minimum when the CAD model is exactly aligned with the image of the implant in the input X-ray image.

C. Optimization

The choice of optimization algorithm depends on the characteristics of the function space to be searched. Our function space is six-dimensional (corresponding to the number of DOFs in the model pose) and contains numerous local minima.

Fig. 9 is a one-dimensional exhaustive plot of the matching score for an *in vivo* image (Fig. 2) where the pose of the femur model was rotated about the vertical (Y) axis. At each hypothesized position, the matching score was recorded. Note the two large minima and many smaller local minima. The global minimum (the correct solution) is the deeper of the two large minima. The other large minimum is caused by the symmetry of the model (a femoral knee component), causing the silhouette to be very similar for two different orientations. The small minima are caused by the nature of the matching function—as the model is translated or rotated across the image, points constantly enter and leave the support set.

Fig. 10(a) shows the femoral implant in the correct overlay position, which corresponds to the global minimum. Fig. 10(b) shows the femoral implant component in the incorrect position, which corresponds to the other large minimum in Fig. 9. Fig. 10(c) shows that the two silhouettes are very similar, but different. The similarity of the silhouettes is due to the highly symmetrical shape of the implant.

To avoid these local minima, a robust optimization algorithm is needed that can find the global minimum. A local search algorithm such as Levenberg–Marquardt will simply find the nearest local minimum. Possible choices of global optimization techniques include simulated annealing (SA) [31] and genetic algorithms (GA) [23]. We chose SA due to its simplicity of implementation—GA may be more efficient in terms of the number of function evaluations required, but both are slow compared with local search methods.

Our SA algorithm is a modified version of the Nelder–Mead [30] (downhill simplex) optimization method. A simplex is a set of seven points where each point represents a possible pose,

[2]Negative weights are used so that the best fit corresponds to a *minimum* of the objective function.

along with the value of the function space at that point. The algorithm iteratively perturbs the worst (highest) point of the simplex, in the direction of a lower score. Gradually, after many iterations, the entire simplex converges upon a minimum of the function space. Generally, this method converges upon the closest local minimum; however, it may randomly escape small local minima based upon the size of the perturbation that it makes.

Simulated annealing employs a virtual temperature in order to escape larger local minima. Specifically, the algorithm adds or subtracts a positive, logarithmically distributed random variable, proportional to the temperature, to the function values in the simplex. Effectively, the temperature makes some solutions appear more favorable than they actually are. As the temperature iteratively decreases, this possibility of "uphill" movement decreases as well. The temperature is lowered according to a defined schedule. In our application, the virtual temperature is decreased every 40 iterations by multiplying the current temperature by 0.99. Periodically, the simplex is expanded to enable the algorithm to test more distant solutions. By slowly lowering the virtual temperature in this manner, the algorithm can escape local minima. On a Silicon Graphics Octane workstation, one iteration takes about 0.2 s. The algorithm automatically terminates when the difference between the best and worst points in the simplex is less than a small threshold.

Fig. 11 shows the error score as a function of the number of iterations, for the same image used earlier. The upper curve shows the instantaneous score at each iteration, as the optimization algorithm tests nearby model poses to see if the score will improve in various directions. Most of these hypothesized poses are rejected since they result in higher (worse) scores. Note the occasional increases in error score when the simplex is expanded. The lower curve is the "best ever" score found so far, and monotonically decreases. Also shown is the temperature, and snapshots of the model overlaid on the image.

D. Supervisory Control

A graphical user interface (GUI) was developed using the Silicon Graphics RapidApp user interface builder tool (Fig. 12) that allows the user to visualize and control the model fitting process. Visualization is accomplished by displaying the model as an overlay on the X-ray image, and by continuously updating the model's position with the current optimization solution. The GUI allows a user to initialize the model pose prior to optimization, which speeds convergence to the correct solution. Although the optimization algorithm can escape most (shallow) local minima, it occasionally gets stuck in a deeper minimum. With a large enough temperature the algorithm could even escape these large minima; however, we have found it to be more efficient to use an interactive approach. In this approach, if the system gets stuck in an incorrect pose, the user can temporarily seize control of the model (using the mouse) and move it toward the correct pose. Releasing the mouse then allows the optimization algorithm to resume at the new pose. This initializes the state vector closer to the solution, saving time by reducing the number of iterations needed to be performed.

This approach is a form of traded control [32], which is used in supervised robot control. The philosophy is that the system
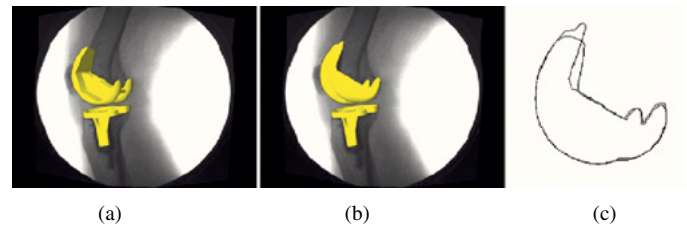

Fig. 10. Overlays corresponding to the two large minima in the previous figure. (a) Femoral implant in correct pose. (b) Incorrect pose. (c) Difference in silhouettes.
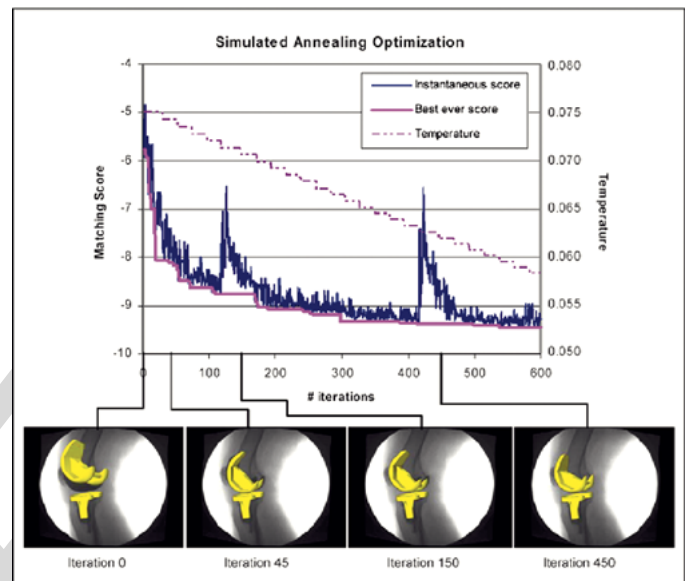

Fig. 11. Matching score as a function of the number of iterations, on the same fluoroscopy image.
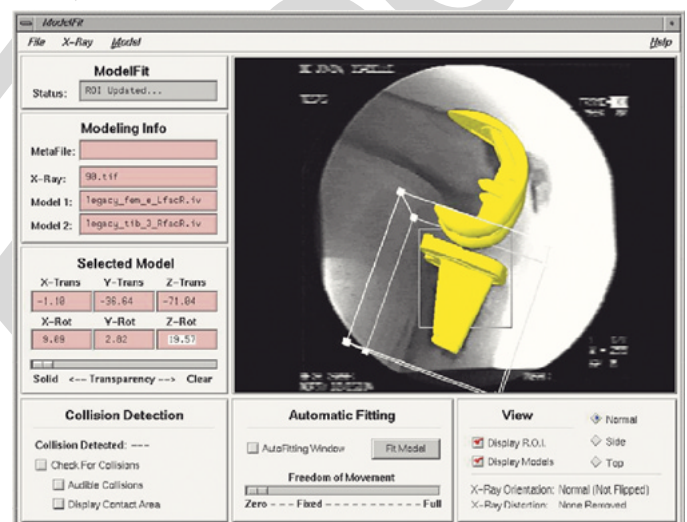
Fig. 12. The graphical user interface allows the user to easily visualize and control the progress of the model fitting process.

incorporates the best of both approaches. It combines the speed and precision of the computer with the common sense and domain knowledge of the human. The resulting system is both efficient and reliable.

Often the user can easily detect that the system has found a local minimum (incorrect solution), because the overlay of the

model does not fit the silhouette in the image. In other cases, the user can see that the resulting femur-tibia pose is physically impossible and/or biomechanically implausible. For example, Fig. 13 shows a tibia model that is in an incorrect pose. By observing the relative femur-tibia pose from the front, it is obvious that the models intersect (i.e., a physical impossibility). Looking at the models from the top, the femur has an excessive rotation with respect to the tibia (i.e., biomechanically implausible). In such cases, the user can move the model closer to the correct solution and restart the optimization algorithm.

IV. PERFORMANCE RESULTS

The registration method has been used to determine knee kinematics *in vivo* in numerous studies involving hundreds of patients and dozens of implant models [33]–[36]. Our results are consistent with measurements made by other methods, including roentgen stereophotogrammetry [37], [38], intercortical pins [39], [40], and quasi-static MRI [41], [42].

In this section, we give the results of tests performed to asses the accuracy and repeatability of the method on *in vivo* and cadaver images. Many previous analyses of 3-D to 2-D registration methods have used "clean" images of objects outside the body, or synthetic data [5], [16], [18]. However, to be sure of testing the method fairly, the images should be as close as possible to those that would be clinically available. Specifically, the implant components should be imaged within an environment of bone and soft tissue, whose presence can affect the accuracy of the results.

A. Convergence Analysis

Experiments were conducted to analyze the convergence properties of the algorithm. First, a representative fluoroscopy image was selected, containing a knee implant [Fig. 14 (left)]. Surface models were created for the femoral and tibial components. The best poses of the models were found (in terms of the matching score), and are shown in Fig. 14 (center).

Note that these poses are not a "gold standard," because we do not know the true poses of the models from another, independent measurement method. They are simply the poses which yield the best value of the similarity measure. Therefore, the experiments described in this section do not test registration accuracy, but rather the ability of the search strategy to find the optimum value of the similarity measure. Accuracy against a gold standard is addressed in the next section.

Tests were performed to measure how well the registration algorithm could find the best pose of the femoral component, when it was run from different starting conditions. Four classes of starting conditions were designated, representing different amounts of deviation of the initial pose to the best pose. Class 1 represents the smallest amount of initial deviation, and Class 4 the largest. The classes correspond to the quality or accuracy of the *a priori* knowledge of the pose. The actual initial deviation values for each parameter are shown in Table I. To set these values, we performed an informal test. A human operator was asked to manually position the femoral component model to best align the graphical overlay with the image. Class 3 represents the approximate deviation we measured after a few seconds of

Fig. 13. (Left) Improper registration of a tibial model. (Center) Viewing from front, component interference is visible. (Right) Viewing from top, excessive rotation is visible.

Fig. 14. (Left) Fluoroscopy image used for convergence analysis. (Center) Best model poses, in terms of the matching score. (Right) Initial starting pose, for a "Class 4" initial deviation.

manual positioning. Class 2 represents the deviation after about a minute of careful manual adjustment. Class 1 and Class 4 represent even more extreme starting conditions—either very accurate or very inaccurate initial placement, respectively. An example, Fig. 14 (right) shows the model in its initial starting pose, for one of the Class 4 starting poses.

1) Results—No Noise or Occlusion: Thirty runs were made for each class. The algorithm was allowed to run unsupervised to completion. For each run, the difference between the final pose and the best pose was recorded. The root mean square (rms) final differences are shown in Table II.

To test the effect of the random search component of the simulated annealing optimization algorithm, we repeated the runs with the initial temperature set to zero. Effectively, with temperature equal to zero, the algorithm becomes a simple gradient descent algorithm that just traverses downhill to the nearest local minimum. We found, as expected, that the algorithm quickly fell into a local minimum near its starting point. Thus, the algorithm did not improve the initial pose estimate significantly.

A closer look at the results from full temperature optimization shows that certain runs had large differences between the final pose and the best pose. Nearly all of these incorrect final poses were the "reflected" pose shown in Fig. 10(b). As described in Section III-C, this corresponds to a large local minimum in the objective function space, due to the symmetrical shape of the implant component. This incorrect pose is rotated approximately 35° from the correct pose, primarily about the vertical (Y) axis.

Since these final differences are very large, they are easy for a human to detect. For example, such an erroneous pose would imply a knee configuration that was biomechanically impossible, or it would result in the two components intersecting in space. Therefore, in practice when such a result occurs, the human operator pushes the model away from that pose and continues the optimization algorithm. This is usually enough to let the model settle into the correct pose. The percentages of times that large (outlier) final differences occurred in Class 1, Class

TABLE I
VALUES OF INITIAL STARTING POSE PARAMETERS

Parameter	Class 1	Class 2	Class 3	Class 4
Magnitude of the initial deviation in the XY (image) plane; the direction of the vector was randomly varied	2 mm	4 mm	8 mm	16 mm
Magnitude of the initial deviation in the Z (out of plane) direction	16 mm	32 mm	64 mm	128 mm
Magnitude of the initial rotational deviation; the direction of the rotation axis was randomly varied	2°	4°	8°	16°

TABLE II
RMS DIFFERENCES BETWEEN FINAL POSE AND BEST POSE

Initial deviation	Magnitude of XY (in-plane) difference	Magnitude of Z (out-of-plane) difference	Magnitude of angular difference
Class 1	0.091 mm	1.367 mm	0.401°
Class 2	1.503 mm	2.528 mm	4.040°
Class 3	1.965 mm	25.34 mm	15.11°
Class 4	2.692 mm	38.91 mm	23.35°

TABLE III
RMS DIFFERENCES BETWEEN FINAL POSE AND BEST POSE, IF OUTLIERS (OBVIOUS FAILURES) ARE REMOVED

Initial deviation	Magnitude of XY (in-plane) difference	Magnitude of Z (out-of-plane) difference	Magnitude of angular difference
Class 1	0.091 mm	1.367 mm	0.400°
Class 2	0.086 mm	1.236 mm	0.326°
Class 3	0.082 mm	1.146 mm	0.325°
Class 4	0.075 mm	0.936 mm	0.347°

2, Class 3, and Class 4 trials were 0%, 3%, 33%, and 50%, respectively.

If we remove these obvious failures from the results, the algorithm achieves results that are very close to the correct pose. Table III shows the rms difference between the final pose and best pose, as a function of the initial deviations. The magnitude of the final differences is almost independent of the initial deviation.

Thus, given a small amount of user supervision and intervention, the optimization algorithm can reliably find poses that are very close to the optimal (as measured by the matching score), even from starting poses with large initial deviations.

Note that the final difference is higher for translation in the out-of-plane direction (the Z axis) because the registration method is much less sensitive to displacements in that direction. Any single-plane registration method will be less accurate along the line of sight direction. If registration were performed with multiple viewing directions, using say, bi-plane or stereo fluoroscopy, this would reduce the Z axis error. However, in some applications the larger Z axis error is acceptable. For example, physical constraints in the knee joint prevent large translational motion between the femur and tibia in the medial-lateral direction. If we view the knee joint in the sagittal (XY) plane, then the relative translational motion in the Z direction should be small and is not of interest in clinical applications.

2) Results—Noise and Occlusion: To show the robustness of the method, we conducted additional tests using images with added noise and occlusion. We added synthetic Gaussian distributed noise $(\mu = 0, \sigma = 20)$ to the image used in the previous section (Fig. 2). This is much higher that the observed measured pixel intensity noise on our fluoroscope and image capture equipment, which is approximately three gray levels. The noisy image is shown in Fig. 15.

To test the effect of occlusion and low foreground-to-background contrast, we next generated an image with simulated occlusions, by combining an image of another knee with the test image used in the previous section (Fig. 16). Such occlusion effects occur frequently in actual practice due to the patient swinging one leg in front of the other during gait [see, for example, Fig. 5(a)]. The net effect is to drastically reduce the contrast between the object of interest (the implant component) and the background.

We next re-ran the same convergence tests on the noise image and the occlusion image. For each image, thirty runs were done for each of the four classes of starting poses. The results show that the convergence was not greatly affected by noise or occlusion. Fig. 17 shows the percentage of runs that converged to the best pose (or a close vicinity of the best pose) instead of the false local minimum. In fact, convergence was slightly better for the noise and occlusion images, indicating possibly that a multi-resolution approach could improve robustness.

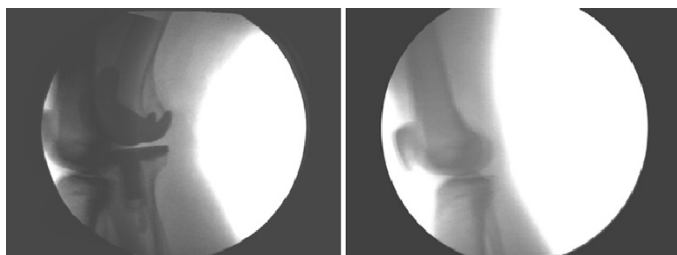Fig. 15. The test fluoroscopy image with a large amount of added noise.

Fig. 16. A simulated occluded image (left) was created by combining the test image in Fig. 2 with an actual image of another leg (right).
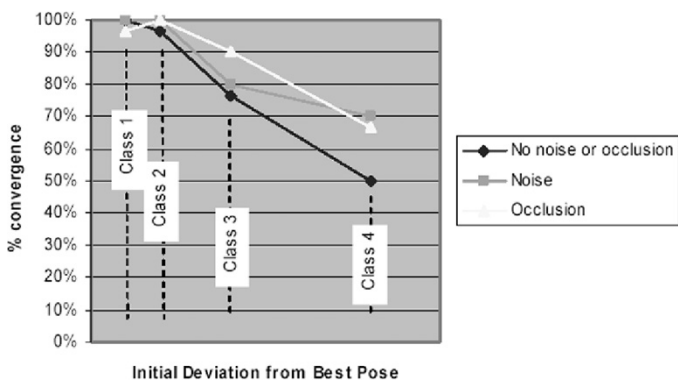

Fig. 17. The percentage of runs that converged to the correct pose instead of the false local minimum.

For those runs that converged to the correct pose, the final pose differences were somewhat higher in the noise and occlusion images. Table IV shows the rms magnitude of the difference between the final pose and best pose, in XY translation, Z translation, and orientation. These were almost independent of the starting pose, so only the results for the "Class 4" initial poses are shown.

The Z (out-of-plane) difference was significantly higher for the occlusion image than for either the noise image or the no-noise image. A close look at the results showed that the registration method consistently estimated the Z position of the implant as being too far away. In other words, there was a consistent bias in the results. The Z position is primarily determined by the apparent size of the silhouette in the image. Since in the occlusion image only a portion of the silhouette is visible, this may degrade the accuracy of the Z position computation.

B. Accuracy Tests

The previous section showed that the registration method can reliably find a pose that is very close to the optimal (as measured by the matching score), even from starting poses with large initial errors. However, a separate issue is the absolute accuracy of the resulting pose, in terms of 3-D translation and rotation. We performed tests on a cadaver knee because it provided the same imaging effects of bone and soft tissue that would be found in *in vivo* images. The resulting images were completely equivalent to *in vivo* images except that the objects were not moving while the images were taken (however, if the exposure time is small, motion blur is not a factor even for moving objects). We also took images over a wide range of flexion and viewing angles, to simulate as closely as possible the range of imaging conditions that are actually used in clinical X-ray fluoroscopy images.

In order to perform a test of absolute accuracy, we need a source of known "ground truth" pose data. Ideally, the ground truth data should be very accurate—at least as accurate as the method being tested, and hopefully much better. Some researchers have located fiducial markers (beads) in X-ray images to obtain ground truth [19], [11]. However, since our registration method itself is also based on X-rays, this is not really an independent method. Other researchers used mechanical positioning systems to place the objects at a known relative pose [5], [16]. However, this method is difficult to use on implants located within the body. Also, there are difficult issues in calibrating the mechanical positioning system itself, such as determining axes of rotation and the position of the components on the actuators.

Instead, our approach was to physically measure specific points on the implant to determine ground truth. In cadavers, it is possible to expose a portion of the surface of the implant so that it can be touched by a hand-held probe. The probe can be automatically tracked using an optical sensor, and the 3-D coordinates of the probe tip can be recorded by a computer. Given a sufficient number of registration points measured in this way, the pose of the implant can be estimated. This pose can be compared with the pose that is estimated from fluoroscopy images, as measured by our registration method.

One problem with obtaining ground truth is that it is difficult to know the precise position and orientation of the fluoroscope coordinate system with respect to the coordinate system of the optical sensor. Even small errors in measuring this alignment would result in large errors in the ground truth data, which would dwarf the errors from our registration method. Therefore, we instead chose to concentrate on measuring the *relative* pose between two implants (i.e., femoral and tibial knee components), and compare the relative pose obtained using the fluoroscopy registration method with the relative pose obtained from the optical sensor. This avoided the need to calibrate the fluoroscopy sensor with respect to the optical sensor. Also, the relative pose between two implant components is often the quantity of interest in applications such as measuring the kinematics of the knee. However, one should note that our test would unable to detect a systematic registration error that would cancel out when computing relative pose.

1) Experimental Data Collection Procedure: Coordinate reference frames were defined for each object in the data collection procedure (Fig. 18). There are two implant components, denoted {Femur} and {Tibia}. The pose of each implant is measured with respect to an optical sensor {Opto}, and also

TABLE IV
RMS VALUES FOR DIFFERENCES BETWEEN THE FINAL POSE AND BEST POSE, USING "CLASS 4" STARTING POSES

Image	Magnitude of XY (in-plane) difference	Magnitude of Z (out-of-plane) difference	Magnitude of angular difference
No noise or occlusion	0.075 mm	0.936 mm	0.347°
Noise	0.210 mm	3.811 mm	0.624°
Occlusion	0.436 mm	40.80 mm	1.142°

with respect to the fluoroscope {Fluoro}. From these measured poses, the relative pose of the femur to tibia is derived.

Shallow registration holes (5/64″) were drilled into the femoral and tibial components of a posterior cruciate retaining knee implant (6 in the tibial component, 10 in the femoral component). The holes were distributed throughout the components in a noncollinear arrangement (Fig. 19). These holes were used to establish a reference frame for each implant component using the optical sensor.

The next step was to create a surface model of the implants. The implants were laser scanned using a Digibot III machine (DIGI-BOTICS Austin, TX), which recorded a large number of 3-D points on the object's surface. A polygonal surface model was then created, resulting in approximately 12 000 polygons for the femoral component, and 28 000 for the tibial component (Fig. 20). Although CAD models of the implants were available, the reason we created these "as is" models, was to determine the exact locations of the registration holes in the implant model coordinate reference frames.

The optical sensor used was an Optotrak 3020 system (Northern Digital Inc., Waterloo, Canada), which tracks infrared light emitting diodes (LEDs) by three fixed linear array charged coupled device (CCD) cameras [Fig. 21(left)]. The distance between the outermost two cameras is approximately 90 cm. The Optotrak sensor was used to track a handheld probe, consisting of 6 LED markers rigidly mounted on one side of a planar aluminum rectangular grid [Fig. 21(right)]. Attached to the probe tip is a 2 mm diameter ball at a known position with respect to the grid. Thus, by locating the 3-D coordinates of the LEDs, the 3-D coordinates of the probe tip can be computed. By placing the probe tip in a registration hole, the location of the registration hole with respect to the Optotrak can be determined.

We measured the uncertainty of the probe tip location by recording the inter-point distance between every pair of registration points on each component 99 times. Since the distance between any two-registration points should be fixed, any deviation must be due to measurement noise. The total measured variance of the distance between any two points was $\sigma^2 = 0.1674$ mm^2. This variance is the sum of the variance in locating point P_i and the variance in locating point P_{i+1}. Therefore, the variance in locating a single point is $\sigma^2/2 = 0.0837$ mm^2, and the standard deviation of a single point is $\sigma_{single} = 0.2893$ mm.

The poses of the femoral and tibial components with respect to the Optotrak sensor were calculated using the "absolute orientation" algorithm [43]. Given the measured 3-D data points from the Optotrak $\{^{Opto}P_1, {}^{Opto}P_2, \ldots, {}^{Opto}P_N\}$,
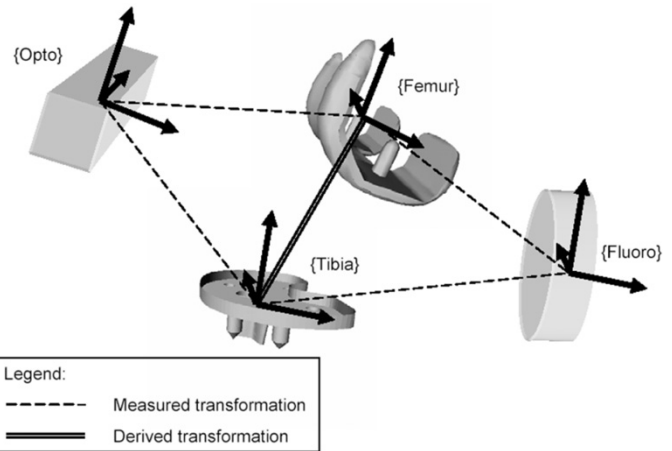

Fig. 18. Principal coordinate frames involved in the accuracy analysis tests.
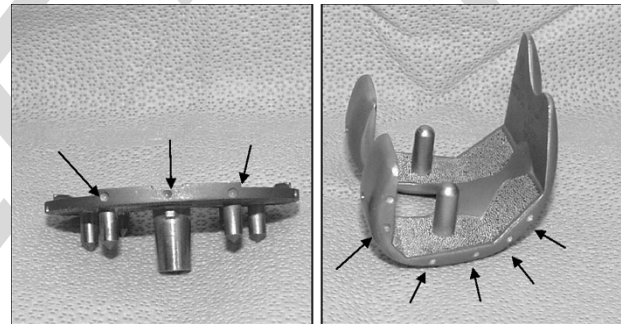

Fig. 19. PCR **<Author: Please define acronym PCR?>** knee implant components, showing registration holes.
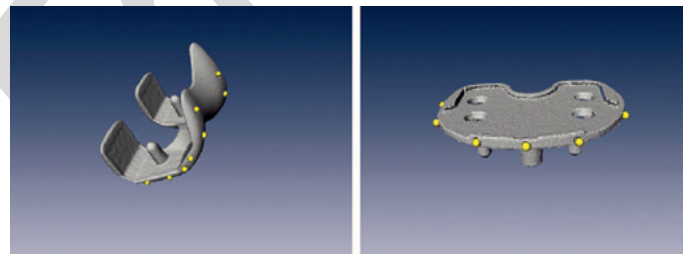

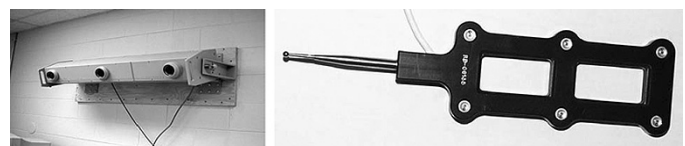Fig. 20. The surface models created by laser scanning, showing the position of the registration holes.

Fig. 21. (Left) Northern Digital Optotrak 3020 position tracking sensor. (Right) Optotrak probe, consisting of six LED markers that are tracked by the sensor.

and the corresponding 3-D points on the model $\{^{\text{Model}}P_1, {}^{\text{Model}}P_2, \ldots, {}^{\text{Model}}P_N\}$, the algorithm finds the pose of the model that minimizes the least squared error between the measured 3-D point locations and the predicted 3-D point locations. The result is the pose of the model with respect to the Optotrak, expressed as a 6-vector $^{\text{Opto}}_{\text{Model}}\mathbf{x} = (x, y, z, \theta_x, \theta_y, \theta_z)^T$, where (x, y, z) is the translation of the model and $(\theta_x, \theta_y, \theta_z)$ represents the rotation angles about the x, y, and z axes. The pose can also be expressed as a 4×4 homogeneous transformation matrix $^{\text{Opto}}_{\text{Model}}\mathbf{H}$ [44].

The Optotrak probe was used to measure the (x, y, and z) coordinates of the registration holes in the femoral and tibial components while simultaneously video recording the X-ray fluoroscopy (Fig. 22). Using this data, we estimated the pose of the femur ($^{\text{Opto}}_{\text{Femur}}\mathbf{H}$) and the tibia ($^{\text{Opto}}_{\text{Tibia}}\mathbf{H}$) separately with respect to the Optotrak sensor. Finally, the relative pose of the femur with respect to tibia in the Optotrak sensor coordinate system was calculated, using the equation $^{\text{Tibia}}_{\text{Femur}}\mathbf{H}^{(\text{opto})} = {}^{\text{Femur}}_{\text{Opto}}\mathbf{H}^{\text{Opto}}_{\text{Tibia}}\mathbf{H}$. This represents the "ground truth" relative pose, by which we can compare the results from our registration method.

The cadaver knee was manipulated into 11 different angles of flexion, ranging from approximately 0° to approximately 120°. Three views were taken at each flexion angle, called "frontal," "sagittal," and "oblique," according to the predominant viewing angle with respect to the fluoroscope (Fig. 23).

Thus, there were a total of 33 distinct poses. Each pose was measured three times using the Optotrak sensor, and also captured using video fluoroscopy. The 3-D to 2-D registration system estimated the pose of the CAD models from the fluoroscopy images. A human operator placed the models in an initial position, and the system subsequently operated fully autonomously without human intervention. The result was the pose of each implant component with respect to the fluoroscope, $^{\text{Fluoro}}_{\text{Femur}}\mathbf{H}^{(\text{fluoro})}$ and $^{\text{Fluoro}}_{\text{Tibia}}\mathbf{H}^{(\text{fluoro})}$. The relative pose of the femur with respect to the tibia was then calculated using the equation $^{\text{Tibia}}_{\text{Femur}}\mathbf{H}^{(\text{fluoro})} = {}^{\text{Femur}}_{\text{Fluoro}}\mathbf{H}^{\text{Fluoro}}_{\text{Tibia}}\mathbf{H}$.

2) Accuracy Analysis: We had two sets of relative poses of the femur with respect to the tibia, one computed by the 3-D-to-2-D registration system ($^{\text{Tibia}}_{\text{Femur}}\mathbf{H}^{(\text{fluoro})}$) and the other the "ground truth" pose as derived from Optotrak data ($^{\text{Tibia}}_{\text{Femur}}\mathbf{H}^{(\text{opto})}$). Each pose represented a pair of implants, where the absolute location of the pair in space was unknown, but the relative pose between them was known. We transformed each pair so that the tibia of each was co-located and aligned with the origin, and looked at the resulting locations of the femurs (Fig. 24). Ideally, if the two relative poses were identical, the two femoral components should be co-incident. Thus, the relative pose between the two femoral components was the error between the two sets of measurements. The relative pose between the two femoral components was computed using $^{\text{Femur-fluoro}}_{\text{Femur-opto}}\mathbf{H}^{(\text{diff})} = {}^{\text{Femur}}_{\text{Tibia}}\mathbf{H}^{(\text{fluoro})\text{Tibia}}_{\text{Femur}}\mathbf{H}^{(\text{opto})}$.

This represented the difference (or error) between the two sets of results—one from our registration system and the other from the ground truth. The relative pose consisted of a translation error and a rotation error. Before analyzing these errors, we first rotated them into the coordinate system of the fluoroscope,
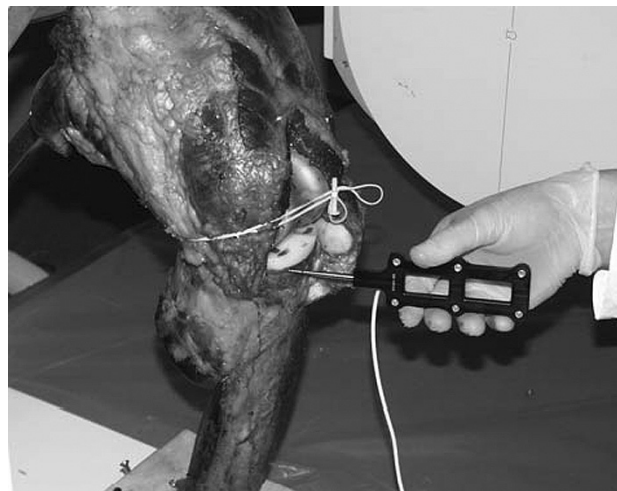

Fig. 22. The Optotrak probe was used to determine the location of the registration points.

Fig. 23. Examples of fluoroscopy images from the frontal, sagittal, and oblique viewing angles (left to right).
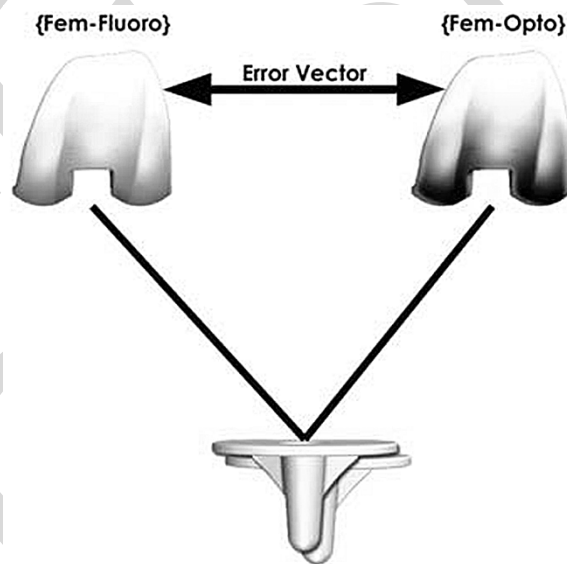

Fig. 24. To compare the two sets of femur-to-tibia poses, the tibias are exactly aligned and the relative pose between the femurs is computed.

because we expected that there would be a dependency of the translation error with the fluoroscope Z axis (note that this rotation does not change the magnitude of the errors, only their directions).

As previously mentioned, the 3-D to 2-D registration system operated fully autonomously, after initial model placement by the operator. The optimization process was allowed to converge, and the pose results were recorded without any adjustment or intervention. However, in actual operational practice, the results are always reviewed by a human operator to detect any large

TABLE V
RMSEs BETWEEN POSES DERIVED FROM OUR REGISTRATION METHOD AND THE OPTOTRAK-DERIVED RESULTS.

View	Magnitude of XY (in-plane) error	Magnitude of Z (out-of-plane) error	Magnitude of angular error
Frontal	0.687 mm	4.701 mm	1.474°
Sagittal	0.678 mm	1.713 mm	1.318°
Oblique	0.610 mm	2.277 mm	1.678°
Combined	0.658 mm	3.209 mm	1.502°

TABLE VI
MEANS AND STANDARD DEVIATIONS OF OBSERVED ERRORS. THE THIRD COLUMN GIVES THE ESTIMATED STANDARD DEVIATION OF
ERRORS OF POSES DERIVED FROM THE OPTOTRAK MEASUREMENT SYSTEM.

	Mean error	Standard deviation of error	Estimated standard deviation of Optotrak error
X translation	-0.023 mm	0.473 mm	0.365 mm
Y translation	-0.086 mm	0.449 mm	0.198 mm
Z translation	1.054 mm	3.031 mm	0.243 mm
X rotation	-0.068°	0.942°	0.479°
Y rotation	0.001°	0.771°	0.316°
Z rotation	0.253°	0.841°	0.554°

errors, caused by the system finding a local minimum. These errors are usually obvious because they manifest themselves as visible discrepancies between the graphical overlay and the silhouette in the underlying image. When such a large error occurs, the operator can simply "nudge" the model away from its current pose and toward the correct solution. The optimization process then resumes and can subsequently find the true solution (global minimum) on its own. We reviewed the results obtained by the registration system and found one (a sagittal view) that had a large registration error. Since in practice an operator would detect this, we removed this pose from the results.

The rms values of the final registration errors are shown in Table V. Results are given for each of the three views separately, as well as for the combined set of views. Note that error is significantly higher for translation along the Z axis (the direction perpendicular to the image plane) because the registration method is much less sensitive to displacements in that direction.

We can also look at the errors about the individual axes. Table VI shows the means and standard deviations of the errors about each of the axes, for the combined set of views. Clinically, these axes can be interpreted as follows. The three translations are the anterior/posterior (corresponding to translation along the X axis of the tibia), inferior/superior (corresponding to translation along the y axis of the tibia), and medial/lateral (corresponding to translation along the z axis of the tibia). The three rotations are abduction/adduction (rotation about the x axis of the tibia), internal/external (rotation about the y axis of the tibia), and flexion/extension (rotation about the z axis of the tibia).

The differences between the two methods are due not only to the errors in our process, but also to any errors in the Optotrak-derived measurements. Thus, these numbers represent an upper bound on the error on our ability to measure relative pose.

However, we know that the Optotrak measurements have significant uncertainty. As described in Section IV-B1, the observed noise in measuring points with the Optotrak probe was 0.2893 mm. Given this number, the uncertainty of the Optotrak poses can be analytically derived, using the method of propagation of errors [45], [46]. These values are shown in the third column of Table VI.

Note that errors in the Optotrak data may account for a large part of the observed error in the test of our registration method. If the errors from our registration system and the errors from the Optotrak system are uncorrelated, then we could subtract the contribution due to the Optotrak from the measured errors, to give a more realistic estimate of the error in our process. However, we did not investigate whether these errors are truly uncorrelated.

V. DISCUSSION

This paper has described a new method for measuring the kinematics of TKA knees from single plane fluoroscopy images. This method is robust with respect to image noise, occlusions, and low object-to-background contrast. We use a direct image-to-image similarity measure, taking advantage of the speed of modern computer graphics workstations to quickly render simulated (predicted) images. As a result, we do not require an accurate segmentation of the implant silhouette in the image (which can be prone to errors). The disadvantage of this method is that it can result in numerous local minima that make it difficult to find the correct solution. We avoid this problem by using a robust optimization algorithm (simulated annealing) that can escape local minima and find the global minimum (true solution). Although we focus on knees in this paper, the method
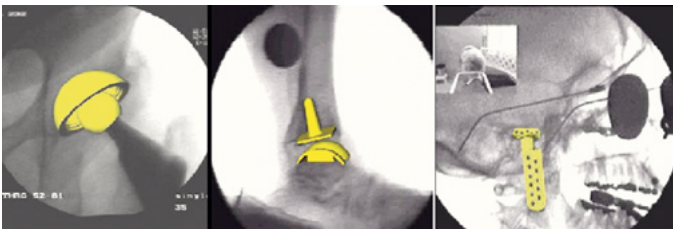
Fig. 25. Applications of the registration system to hips, ankles, and TMJ implants (left to right).

can be (and has been) applied to other implanted joints, including hips, ankles, and TMJ (Fig. 25).

Another advantage of the direct image-to-image similarity measure is that it allows for future extension of this work to non-implanted joints and registration of other organs besides bone. For nonmetallic objects, internal details will be visible in the X-ray images, and not just a black silhouette. This will allow prediction of the intensities inside the object silhouette, and possibly lead to more accurate prediction and registration. Preliminary work has already been done to extend the method to registering the bones of the normal knee [47].

Convergence tests showed that the registration method can reliably find poses that are very close to the optimal (as measured by the matching score), even from starting poses with large initial errors. The rms deviations from the optimal pose were approximately 0.4 (total angular error) and 0.1 mm of translation (in the XY plane). RMS deviations were somewhat larger in the Z translation direction (up to 1.4 mm), because the registration method is much less sensitive to changes in that direction. We also showed that the method is robust with respect to image noise and occlusions. However, a small amount of user supervision and intervention is necessary to detect cases when the optimization algorithm falls into a local minimum. Intervention is required less than 5% of the time when the initial starting pose is reasonably close to the correct answer, but up to 50% of the time when the initial starting pose is far away.

Accuracy tests were performed on cadaver images that were very similar to *in vivo* clinical X-ray fluoroscopy images, to allow a fair assessment of the algorithm. A completely independent method using an optical sensor was used for determining the ground truth, unlike other work that uses ground truth derived from X-ray data. The results showed that our registration method is highly accurate for measuring relative pose. With the exception of the Z translation, the overall rms error difference in translation was approximately 0.65 mm, and the rms error in rotation was approximately $1.5°$.

Although ideally the ground truth data should be much more accurate than the system being tested, in our case the ground truth data (derived from an Optotrak sensor) was comparable in accuracy to the system being tested. Therefore, the measured errors contain contributions from both the errors due to our system and the errors due to the Optotrak. We analytically estimated the contribution of the error in the relative pose measurement due to the Optotrak and showed that it is a significant fraction of the measured error. For example, the standard deviations of the measured translation errors in the XY plane were approximately 0.45 mm, and the estimated standard deviation due to

Optotrak was approximately 0.35 mm. Similarly, the standard deviations of the measured rotation errors were approximately $0.9°$, and the estimated standard deviation due to Optotrak was approximately $0.5°$. Thus, the Optotrak may be a large contribution to the total observed error. In our analytical model, we assumed independent, normally distributed errors in measured point locations. This does not take into account systematic or correlated errors.

We note that the accuracy of the registration results depends on having accurate geometric models of the implant components, as well as good calibration of the fluoroscope and imaging process. Finally, in our tests, we used stationary knees, unlike clinical fluoroscope sequences in which the knee may be moving. We have performed experiments with moving knees and have seen no difference in the results, but we have not tested extremely fast motions.

Suggestions for future research include: 1) using a more efficient optimization algorithm than simulated annealing; 2) determining whether the use of biplanar fluoroscopy yields results that are significantly more accurate than single plane fluoroscopy; and 3) measuring the sensitivity of the registration method to errors in calibrating the fluoroscope (the sensitivity of a similar calibration method was analyzed in [48]).

Finally, the relative weights in our matching score function (giving the relative importance of the region score versus the contour score) have fixed values, regardless of the object-to-background contrast. However, in images with low contrast, the magnitude of the gradient will change, and will affect the contribution of the contour score. It is possible that better performance could be achieved by dynamically adjusting the weights depending on the image contrast.

ACKNOWLEDGMENT

REFERENCES

[1] M. M. Landy and P. S. Walker, "Wear of ultra-high molecular weight polyethylene components of 90 retrieved knee prostheses," *J. Arthoplasty*, vol. 3, pp. s73–s85, 1988.

[2] D. A. Dennis, "Patellofemoral complications in total knee arthroplasty: A literature review," *Amer. J. Knee Surg.*, vol. 5, pp. 156–166, 1992.

[3] T. M. Wright, C. M. Rimnac, and S. D. Stulberg, "Wear of polyethylene in total joint replacement. Observations from retrieved PCA knee implants," *Clin. Orthop.*, vol. 276, pp. 126–134, 1992.

[4] J. B. Stiehl *et al.*, "Fluoroscopic analysis of kinematics after posterior-cruciate retaining knee arthroplasty," *Br. J. Bone Joint Surg.*, vol. 77, no. 6, pp. 884–889, 1995.

[5] S. A. Banks and W. A. Hodge, "Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy," *IEEE Trans. Biomed. Eng.*, vol. 43, pp. 638–649, June 1996.

[6] L. Blankevoort, R. Huiskes, and A. de Lange, "The envelope of passive knee joint motion," *J. Biomech.*, vol. 21, no. 9, pp. 705–720, 1988.

[7] M. C. Murphy *et al.*, "*In vivo* measurement of the three-dimensional skeletal motion at the normal knee," *Trans. Othop. Res. Soc.*, p. 142, 1995.

[8] E. Y. S. Chao, "Justification of triaxial goniometery for the measurement of joint rotation," *J. Biomech.*, vol. 13, pp. 989–1006, 1980.

[9] T. P. Andriacchi, T. S. Stanwyck, and J. O. Galante, "Knee biomechanics and total knee replacement," *J. Arthroplastyd*, vol. 1, pp. 211–219, 1986.

[10] L. Lemieux *et al.*, "A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs," *Med. Phys.*, vol. 21, no. 11, pp. 1749–1760, 1994.

[11] G. Penney *et al.*, "A comparison of similarity measures for use in 2-D-3-D medical image registration," *IEEE Trans. Med. Imag.*, vol. 17, pp. 586–595, Aug. 1998.

[12] W. E. L. Grimson, *Object Recognition by Computer*. Cambridge, MA: MIT Press, 1990.

[13] D. G. Lowe, "Fitting parameterized three-dimensional models to images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 441–450, May 1991.

[14] D. J. Kriegman and J. Ponce, "On recognizing and positioning curved 3-D objects from image contours," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 12, pp. 1127–1137, 1990.

[15] W. A. Hoff *et al.*, "Three-dimensional determination of femoral-tibial contact positions under "in-vivo" conditions using fluoroscopy," *Clinical Biomechanics*, vol. 13, no. 7, pp. 455–472, 1998.

[16] S. Lavallee and R. Szeliski, "Recovering the position and orientation of free-form objects from image contours using 3-D distance maps," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 378–390, Apr. 1995.

[17] A. Gueziec *et al.*, "Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot," *IEEE Trans. Med. Imag.*, vol. 17, pp. 715–728, Oct. 1998.

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

[19] J. Weese *et al.*, "Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery," *IEEE Trans. Inform. Technol. Biomed.*, vol. 1, pp. 284–293, Dec. 1997.

[20] P. Viola and W. Wells, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, 1997.

[21] P. Hellier *et al.*, "Hierarchical estimation of a dense deformation field for 3-D robust registration," *IEEE Trans. Med. Imag.*, vol. 20, pp. 388–402, May 2001.

[22] G. Matsopoulos *et al.*, "Automatic retinal image registration scheme using global optimization techniques," *IEEE Trans. Inform. Technol. Biomed.*, vol. 3, pp. 47–60, Mar. 1999.

[23] J. Rouet, J. Jacq, and C. Roux, "Genetic algorithms for a robust 3-D MR-CT registration," *IEEE Trans. Inform. Technol. Biomed.*, vol. 4, pp. 126–136, June 2000.

[24] M. E. Sarojak *et al.*, "An interactive system for kinematic analysis of artificial joint implants," presented at the *36th Rocky Mountain Bioengineering Symp.*, Copper Mountain, Colorado, 1999.

[25] M. Sarojak, "Interactive pose estimation of total joint arthroplasty via X-ray fluoroscopy," masters thesis, Eng. Div., Colorado Sch. Mines, Golden, CO, 1998.

[26] M. Mahfouz, "A new registration method for implanted and non implanted knee joints using single plane X-ray fluoroscopy," masters thesis, Eng. Div., Colorado Sch. Mines, Golden, 2002.

[27] W. K. Pratt, *Digital Image Processing*, 2nd ed. New York: Wiley, 1991.

[28] "Registration of contours using chamfer matching," M. Van Herk, in *Handbook of Medical Imaging, Processing, and Analysis*, I. Bankman, Ed. San Diego, CA: Academic, 2000, pp. 515–527.

[29] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. New York, NY: Wiley, 1987.

[30] W. H. Press *et al.*, *Numerical Recipes in C*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[31] S. Kirkpatrick, J. C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[32] T. B. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*. Cambridge, MA: MIT Press, 1992.

[33] J. Stiehl *et al.*, "*In vivo* determination of condylar lift-off and screw-home in a mobile-bearing total knee arthroplasty," *J. Arthroplasty*, vol. 14, no. 3, pp. 293–299, 1999.

[34] J. Stiehl *et al.*, "The cruciate ligaments in total knee arthroplasty: A kinematic analysis of 2 total knee arthroplasties," *J. Arthroplasty*, vol. 15, no. 5, pp. 545–550, 2000.

[35] D. Dennis *et al.*, "Femoral condylar lift-off *in vivo* in total knee arthroplasty," *Br. J. Bone Joint Surg.*, vol. 83, no. 1, pp. 33–39, 2001.

[36] J. Stiehl *et al.*, "Frontal plane kinematics after mobile-bearing total knee arthroplasty," *Clin Orthop.*, vol. 392, pp. 56–61, 2001.

[37] K. Nilsson, J. Karrholm, and L. Ekelund, "Knee motion in total knee arthroplasty," *Clin. Orthop. Related Res.*, vol. 256, pp. 141–161, 1990.

[38] J. Karrholm, S. Brandsson, and M. Freeman, "Tibiofemoral movement 4: Changes of axial tibial rotation caused by forced rotation at the weight-bearing knee studied by RSA," *Br. J. Bone Joint Surg.*, vol. 82, no. 8, pp. 1201–1203, 2000.

[39] M. A. Murphy, *Geometry and the Kinematics of the Normal Human Knee*. Cambridge, MA: Massachusetts Inst. Technol., 1990.

[40] M. A. Lafortune *et al.*, "Three dimensional kinematics of the human knee walking," *J. Biomech.*, vol. 25, no. 4, pp. 347–357, 1992.

[41] H. Iwaki, V. Pinskerova, and M. Freeman, "Tibiofemoral movement 1: The shapes and relative movements of the femur and tibia in the unloaded cadaver knee," *Br. J. Bone Joint Surg.*, vol. 82, no. 8, pp. 1189–1195, 2000.

[42] P. Hill *et al.*, "Tibiofemoral movement 2: The loaded and unloaded living knee studied by MRI," *Br. J. Bone Joint Surg.*, vol. 82, no. 8, pp. 1196–1198, 2000.

[43] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 679–698, June 1986.

[44] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Optical Soc. Amer.*, vol. 4, no. 4, pp. 629–642, 1987.

[45] J. Craig, *Introduction to Robotics: Mechanics and Control*, 2nd ed. Reading, MA: Addison-Wesley, 1990.

[46] R. Haralick and L. Shapiro, *Computer and Robot Vision*. Reading, MA: Addison-Wesley, 1993.

[47] W. A. Hoff and T. Vincent, "Analysis of head pose accuracy in augmented reality," *IEEE Trans. Visual. Comput. Graphics*, vol. 6, pp. 319–334, Oct.-Dec. 2000.

[48] D. Dennis *et al.*, "*In vivo* three-dimensional determination of kinematics for subjects with a normal knee or a unicompartmental or total knee replacement," *J. Bone Joint Surg. Amer.*, vol. 83-A, pp. 104–115, 2001. Suppl 2 Pt 2.

[49] S. A. Banks, *Model Based 3-D Kinematic Estimation from 2-D Perspective Silhouettes: Application with Total Knee Prostheses*. Cambridge, MA: Massachusetts Inst. Technol., 1992.