Automated Registration of 3-D Knee Implant Models to Fluoroscopic Images Using Lipschitzian Optimization

Paris D. L. Flood and Scott A. Banks

Abstract—**This paper describes an automated method for registering 3-D models of metallic knee implants to single-plane radiographic images. We develop a multistage approach that identifies the correct pose by matching altered dilations of an edge-detected image with the silhouette of an implant model. The location of the similarity function's minimum is found using a novel optimization routine that combines the Dividing Rectangles algorithm with properties of the registration metric. Depending on the implant type (tibial or femoral), this technique reliably converges under maximum displacements of approximately 25 to 55 millimeters for translation components and 25° to 55° for Euler angles. The method proves to be robust to noise from bones and soft tissue. After an initial guess for the first image in the sequence, subsequent frames can be automatically registered from the optimum pose in the previous image.**

Index Terms—**X-ray imaging and computed tomography, optimization, registration.**

I. Introduction

ACCURATE measurements of implant kinematics have been critical to the improvement of joint replacement designs [1] and simulation of joint mechanics [2]. Many techniques used to quantify joint kinematics lack the ability to provide either *in vivo*, dynamic or precise information [3]. Model-image registration has been successfully used in orthopaedic research for over twenty years [4] to overcome these limitations.

Common to these model-image registration methods, a computer-generated perspective projection of a three-dimensional implant model is compared to a digitized radiograph of the implant using a similarity metric. The pose (location and orientation) of the computer model is adjusted by an optimization routine until the metric is minimized and the model is correctly registered to the image (Figure 1). X-ray
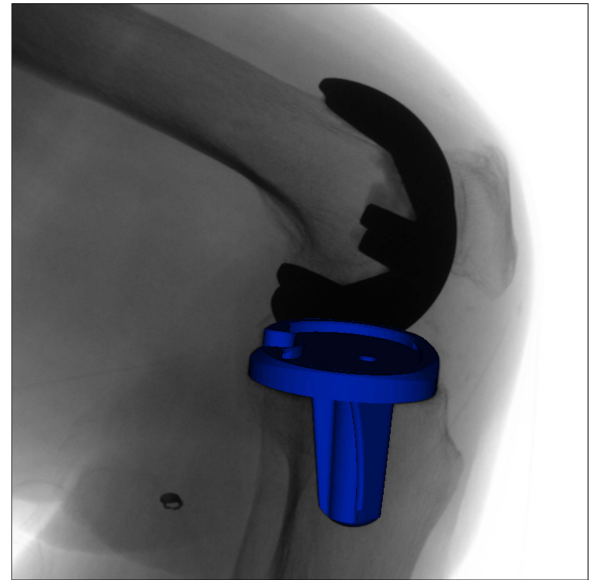
Fig. 1. Radiographic image of the moving knee with the correct projection of a tibial implant superimposed.

fluoroscopy is used to generate a sequence of radiographic images during a patient activity such as chair rise, stair stepping or arm abduction. Registration of the model to each radiographic image frame provides a description of the implant kinematics over the patient's motion cycle.

Quantification of the similarity between two images is a well-studied problem in computer vision [5], [6]. Strategies for determining the matching score are often categorized as feature-based or intensity-based comparisons [7]. Feature-based image registration is a technique commonly used in robotics that identifies point-type or line-type features in the image and locates them on the model. Intensity-based registration techniques compare grayscale patterns between pixels in the projected and original images.

Both feature-based and intensity-based techniques have formed highly successful similarity metrics for an array of registration routines [3], [4], [8]. These metrics often compare the silhouette of the implant with the projected edges of the model. These silhouettes are easily identified in radiographic images due to the high contrast between metallic implants and the surrounding biological tissue. Methods combining weighted edge and grayscale correlation have also achieved robust registration convergence on a variety of joints [3].
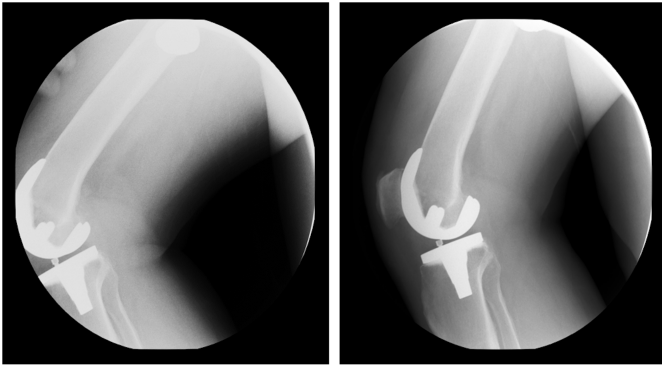
Fig. 2. Sequential fluoroscopy image frames from a lunge activity.

Once a matching score has been determined, model-image registration routines must find the pose that minimizes disparity. A large variety of optimization schemes and numerical solvers have been implemented including the Levenberg-Marquardt algorithm [9], simulated annealing [3], and the Durand-Kerner method [8]. Despite the variation in approach, the kinematic accuracy is relatively consistent across these optimization algorithms [10]. Due in part to image clutter, occlusions, and object design, the similarity metric is a noisy function with a high number of local minima. This fact severely limits the convergence range of the optimization scheme from a user-defined initial guess. Registration success rates are under-reported but can be around 90% when the initial guess is 4mm from the in-plane solution, and 4° from the rotation solution [3], [10]. However, this success rate can drop to as low as 50% when the initial guess is within 16mm and 16°, respectively [3].

In high quality fluoroscopy, even slow movements can cause a shift in pose from frame-to-frame that far exceeds the capture region of the optimizer (Figure 2). Consequently, a human operator is needed to initialize every image in the fluoroscopy with an accurate guess. This process is a time-consuming and error-prone task which severely limits the amount of data that can be processed for kinematics. In this paper we describe a novel method for expanding the convergence range and the success rate by incorporating a mutistage similarity metric with a Lipschitzian optimization scheme. Our algorithm was evaluated for femoral and tibial implants on a variety of images. By increasing the displacement of the starting pose from the true pose, we were able to determine the average error and registration success rate for a variety of search ranges. The testing indicated an optimizer tolerance that allows for automated registration by initializing each frame with the previous frame's solution. The estimated range of our algorithm (approximately 25 to 55 millimeters/degrees for each pose component, depending largely on implant type) compares favorably to current popular methods which focus on a much more restricted subset of the solution space [11].

II. METHOD

A. Similarity Metric

Contour-based similarity metrics have proven to be popular in image registration problems [7], [10]. In a radiographic image, the metallic components of an implant appear in stark contrast to the surrounding bone and tissues. This fact implies a steep intensity gradient around the silhouette of the implant which is conducive for robust edge detection - even in cases of high image noise. Because the gradient magnitude of the implant is relatively consistent across a fluoroscopic sequence, the threshold values of an edge detector (like the Canny algorithm) generally do not have to be specified more than once. The lack of individual image preparation allows for an increased level of automation in the routine.

The output of the edge detection algorithm is a binary image with nonzero values along the contours, and zero values everywhere else. The matching score is computed by comparing the pixel values of the original image and the synthesized projection image after each has been processed by the edge detector. If O is the edge-detected input image and P is the edge-detected candidate image, the matching score is:

$$\sum_x \sum_y |P_{xy} - O_{xy}| \tag{1}$$

The score is a simple ℓ^1 norm that rewards matching edges and punishes overlapping black space. There is no need for normalization because we are interested in finding the global minimum of the function. We use ℓ^1 in place of other p-norms because it can be computed by a graphics processing unit (GPU) extremely efficiently.

Automation requires the optimizer to traverse a large search domain within a limited budget of samples, thus, it is important to design a similarity metric that is "smooth" and free of excessive local extrema. Unfortunately, the contours of the projected model are sensitive to even minimal changes in the model pose. This situation results in a chaotic similarity function with an unacceptable amount of local minima. Therefore, we adopt a pyramidal approach to registering the model by transforming the similarity measure into an easily optimized function, then slowly rewriting the metric back to its original form. This task is achieved by dilating the edge detected images into thick lines, and subsequently running the optimization routine on thinner and thinner edges until the images are close to their original size.

A dilation algorithm takes a binary image and turns on values within a specified distance to an edge (Figure 3). Dilation greatly decreases the amount of noise in the similarity function by increasing the possibility for curve overlap (Figure 4). However, contours become less nuanced as their thickness is increased, often leading to exaggeration of local minima. To incorporate the lost information into the final location of the global minimum, we must use a very small amount of dilation in the final optimization run. In the next section we describe the integration of the pyramidal approach with the optimization scheme in detail.

B. Optimization Routine

The analytic form of our objective function is unknown for each pair of synthesized and radiographic images, therefore function values and gradient approximations can only be determined through sampling. Our search domain is bounded
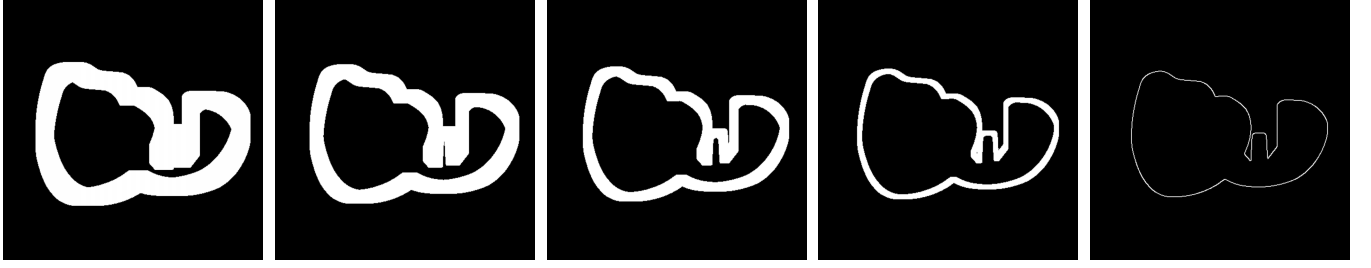
Fig. 3. Examples of morphologic dilation on a projected implant model, starting with high dilation (far left) and decreasing to no dilation (far right).
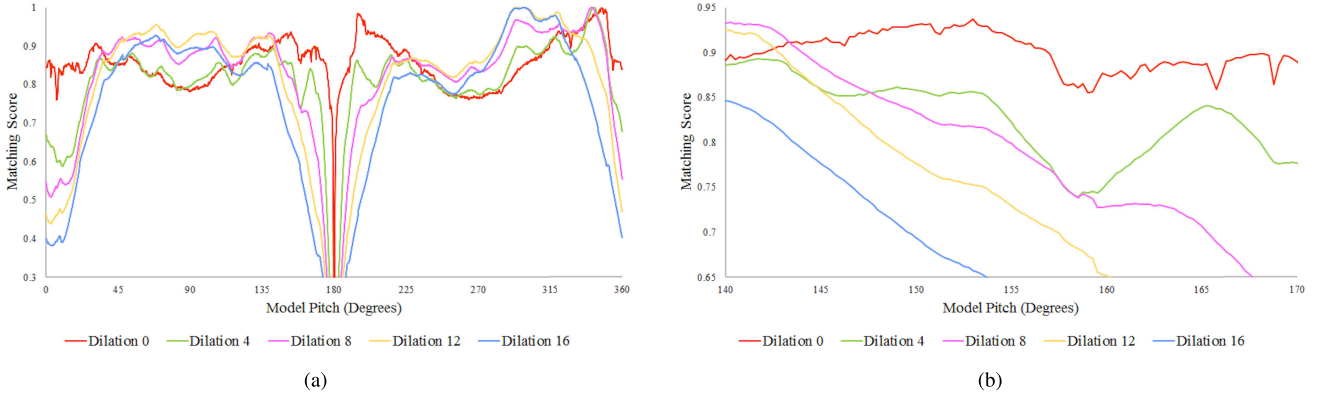

Fig. 4. (a) Matching score at varying dilation levels (same dilation for both input and comparison images) for a femoral implant rotated about the proximal-distal axis (optimum pose at 180°). Dilation reduces noise and eases optimization, but exaggerates local minima from symmetry. (b) Zoomed in view.

by tolerances on the position and orientation of the model after an initial pose is generated. Such *black box* optimization problems have been well-studied in the literature [12], [13] and a large number of high performance algorithms exist [14]. It is important to note that the similarity metric requires computationally expensive tasks (edge detection and 3D rendering) which significantly limit the total number of evaluations available to the optimizer. We, therefore, place great value on the convergence speed of the algorithm and the ability to make cost function calls in parallel. Lipschitzian optimization presents an attractive approach that satisfies our requirements of a black box optimizer and integrates intuitively with the pyramidal approach to image registration.

1) Lipschitzian Optimization: A function $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is *Lipschitz continuous* if there exists a $K > 0$ such that

$$\|f(x) - f(y)\| \leq K \|x - y\| \tag{2}$$

for all $x, y \in U$.

K is known as the *Lipschitz constant* when it is the smallest number that satisfies the definition of Lipschitz continuity. If a Lipschitz continuous function has a known Lipschitz constant, Shubert showed that this information can be used to construct an iterative optimization routine to seek the global extremum [15].

Consider the one-dimensional function $f : [a, b] \to \mathbb{R}$. Assuming f is Lipschitz continuous with constant K, the following conditions apply for $x \in [a, b]$.

$$f(x) \geq f(a) - K(x - a) \tag{3}$$
$$f(x) \geq f(b) + K(x - b) \tag{4}$$

The intersection of the two inequalities at x_0 provides an initial estimate of the function minimum. The Shubert algorithm divides the interval at the intersection into $[a, x_0]$ and $[x_0, b]$ and performs the previous operation. The interval with the lowest function value at the point of intersection is selected for subdivision and the process is repeated until the sampling sequence converges within tolerance to the global minimum.

Shubert's method highlights the fundamental drawback to Lipschitzian optimization - knowledge of the Lipschitz constant. Without the analytic form of the function, the Lipschitz constant is difficult to determine. Additionally, the convergence speed of the algorithm is significantly reduced if an estimate is too high.

2) Overview of the DIRECT Algorithm: DIRECT was presented in [16] as a form of Lipschitzian optimization without the need for a known Lipschitz constant. The algorithm uses many possible values of K to determine if a region should be subdivided by framing the constant as a weighting parameter between global and local search. Consider a generic black-box optimization problem: let $l, u \in \mathbb{R}^N$, $\Omega = \{x \in \mathbb{R}^N : l_i \leq x_i \leq u_i\}$, and $h : \Omega \to \mathbb{R}$. Find

$$x^* = \arg \min_x h(x) \tag{5}$$

The first step in DIRECT is to map the domain Ω into the unit hyper-cube $\bar{\Omega}$ using the standard function $g : \Omega \to \bar{\Omega}$ defined by $g(x)_i = \frac{x_i - l_i}{u_i - l_i}$. We proceed to trisect $\bar{\Omega}$ along an arbitrary axis into three distinct hyper-rectangles. For each hyper-rectangle, the center is converted back to the original

domain and the function is evaluated at the point. From this step forward, the process of choosing a hyper-rectangle for further division is no longer trivial.

Let c_j be the center of a hyper-rectangle j and let d_j be a measure of size defined as the Euclidean distance from c_j to a vertex. In a purely global search scheme, j is selected for further division if

$$0 \leq d_j - d_i \quad \forall i \in \mathcal{J} \tag{6}$$

where $\mathcal{J}$ is the set of current hyper-rectangles.

After 3^{kN} evaluations we are left with a uniform grid of hyper-cubes, each with side length 3^{-k}. However, this method neglects function values when choosing space to search and consequently converges very slowly. In a purely local search scheme, the hyper-rectangle j is selected for further division if

$$0 \leq f(c_i) - f(c_j) \quad \forall i \in \mathcal{J} \tag{7}$$

where $f : \bar{\Omega} \to \mathbb{R}$ is defined by $f(x) = h \circ g^{-1}(x)$.

Choosing the hyper-rectangle with the lowest function value at the center can lead to quick solutions with a good initial guess, but the algorithm can easily fall into a local hole far from the global minimum. DIRECT strikes an efficient compromise by selecting several *potentially optimal* rectangles per iteration using all possible weightings of local and global search.

Let $\epsilon > 0$ and f^* be the current function minimum. A hyper-rectangle j is potentially optimal if there exists some $\widetilde{K}$ such that

$$f(c_j) - \widetilde{K} d_j \leq f(c_i) - \widetilde{K} d_i \quad \forall i, \text{ and} \tag{8}$$
$$f(c_j) - \widetilde{K} d_j \leq f^* - \epsilon |f^*| \tag{9}$$

The first line of the definition ensures that the set of potentially optimal hyper-rectangles contains the lowest possible function values for any value of the Lipschitzian constant. Figure 5 gives a geometric interpretation of this condition. We can determine the set of potentially optimal hyper-rectangles by selecting all the points over the lower convex hull. The second line of the definition uses a small value for ϵ (generally less than 10^{-7}) to ensure that our algorithm doesn't over-search near the current minimum. After all potentially optimal hyper-rectangles are identified, each one is trisected along the longest side and the algorithm is repeated until a limiting number of cost function calls has been reached. Figure 6 demonstrates the division process for a two-dimensional optimization problem.

3) Integration With Pyramidal Optimization and Dynamic Edge Dilation Through DIRECT-JTA: Although DIRECT has found success as a black box optimizer for low-dimension engineering problems (including medical image registration) [17], [18], the algorithm suffers from slow asymptotic convergence [19]. The balance parameter ϵ introduces a global bias to DIRECT that is necessary to prevent the optimizer from overvaluing hyper-rectangles close to the current sampled minimum. However, even small values of ϵ diminish the algorithm's ability to efficiently exploit promising areas of the domain. Consequently, DIRECT is quick to identify
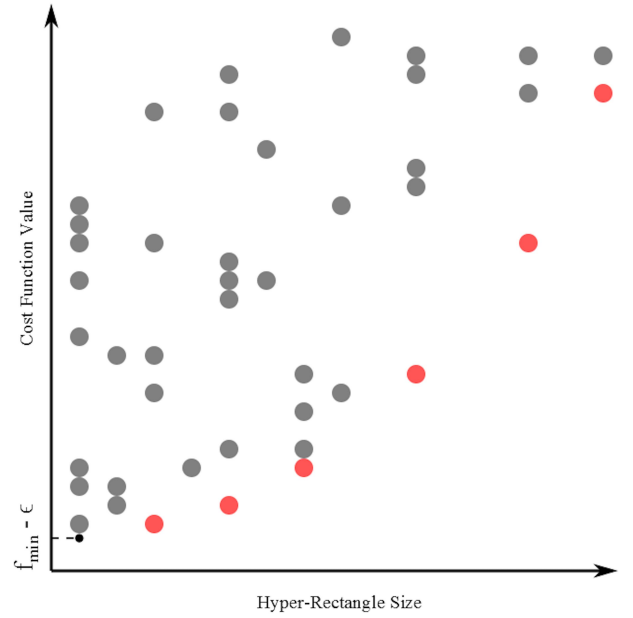

Fig. 5. Hyper-rectangles plotted by size and cost function evaluation at center. Red points mark potentially optimal rectangles along the lower convex hull that represent an efficient weighting of global and local search.
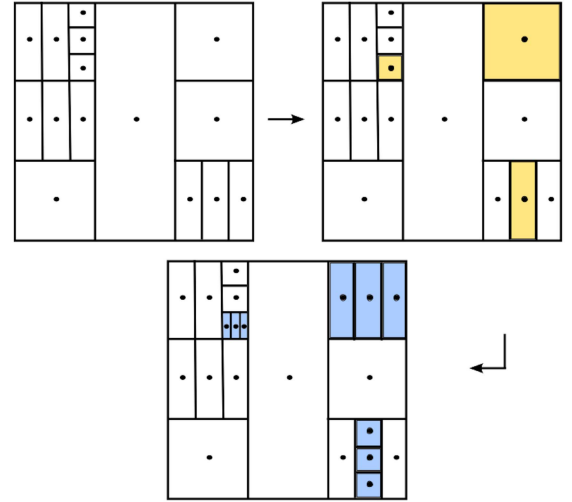
Fig. 6. DIRECT iteration over a two-dimensional optimization problem. Potentially optimal yellow rectangles are selected and subsequently trisected into blue rectangles. Trisection occurs along the longest side or, in the case of a tie, a predefined order.

basins containing the optimum, but slow to pick out the global extremum. Adjustments devised to improve the speed of convergence by altering the role of the balance parameter have demonstrated moderate success [20]. We propose a modified algorithm, called DIRECT-JTA (Joint Track Automated), that overcomes the convergence issue by integrating an explore and exploit scheme with our pyramidal approach to image registration.

DIRECT-JTA is a multi-start implementation of DIRECT that explores the domain for regions of immediate improvement before exploiting them to find the minimum. The first step in the algorithm is to run DIRECT on the initial search

Fig. 7. Set of hyper-rectangles after an initial DIRECT-JTA level. Color gradient shows distance from current function minimum.
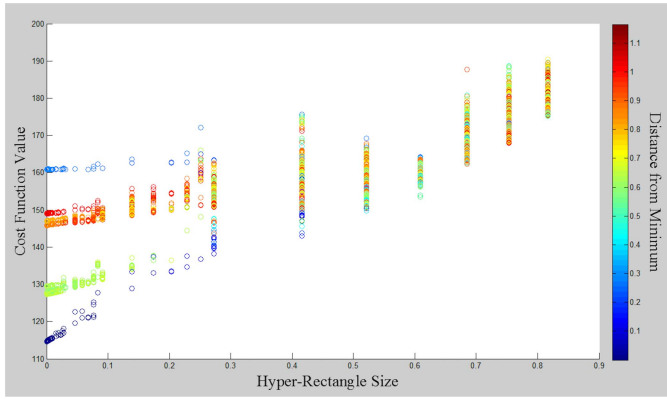

Fig. 8. Correct registration of a femoral implant (left) and an incorrect registration of the same implant due to rotational symmetry (right).

range using a high level of dilation in the edge-detected input images (this is known as the "trunk" stage). As demonstrated in Figure 4, high levels of dilation generate smoother similarity metrics than contours with little to no dilation. In this stage the generated silhouette is dilated to a value of one as this was found to reduce the sensitivity of the maximum dilation constant. DIRECT's swift initial convergence speed combined with an easily traversed similarity metric allows for the fast identification of promising subdomains. To mitigate the distortion introduced by the dilation, the algorithm is restarted (after a budget of cost function calls is exhausted) at a lower level of dilation with a smaller hyper-rectangle. In this stage both comparison and input images have the same dilation level.

At the end of the initial level in DIRECT-JTA we are left with a set of hyper-rectangles similar to those presented in Figure 7. The selection of the initial hyper-rectangle at the restart should reflect the smoothness of the similarity metric. While it is intuitive that an excessive amount of local minima will impede optimizer performance, it remains important to characterize the navigability of the function. The Lipschitz constant presents an attractive approach to measure local turbulence and the global convergence speed. Functions with rapid changes in small neighborhoods and steep descents to the optimum will have relatively high Lipschitz constants. The similarity metric with no dilation in Figure 4 has an estimated Lipschitz constant nearly twice as large as the one corresponding to the dilation factor of 16.

In general, the Lipschitz constant of the similarity function is inversely proportional to the dilation of the edges. We use this fact to determine which hyper-rectangles to create at the restart. At the end of a cycle, a hyper-rectangle of predetermined size is fixed around the location of the current function minimum. This position corresponds to the best potentially optimal hyper-rectangle for low Lipschitz constants (which we would expect to see at higher dilation levels).

The size of the new hyper-rectangle is largely determined by the type of implant we are optimizing. In general, the initial stage will find basins around either the correct pose or a highly similar local minimum stemming from the symmetry of the implant (Figure 8). These "symmetric traps" are specific to the type of implant (femur or tibia) and almost always differ from
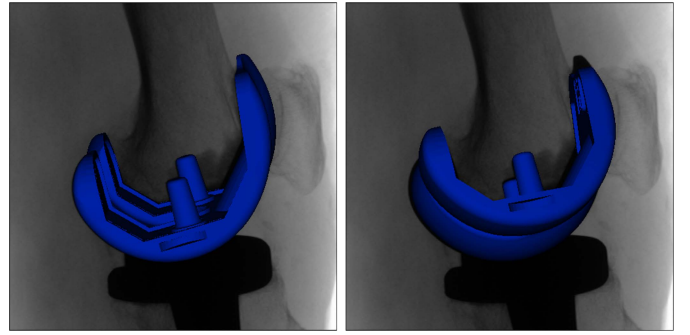
the optimum by a proper subset of the degrees of freedom (generally a rotational symmetry). Although setting the size of the hyper-rectangles to overcome symmetric traps is not difficult for those familiar with the implant, we test a generic size for users unfamiliar with the concept.

The new hyper-rectangle is evaluated in the "branch" stage using DIRECT at a relatively high dilation level (albeit lowered from the initial search) in order to take advantage of the lower associated Lipschitz constant. Several branch restarts are performed, all at the same level of dilation. After all of these searches have used their budget of evaluations, the location of the current optimum is refitted with a hyper-rectangle designed to thoroughly search the out-of-plane translation. This is achieved by holding the search range for the in-plane translation and Euler angles to a few millimeters and degrees, respectively. The z-translation search range is set much higher, generally on par with its value in the initial search range. DIRECT is once again performed for a fixed amount of cost function calls and the optimum result is returned as the final pose. This extra search in the out-of-plane translation is necessary because it is the least accurate degree of freedom for single-plane fluoroscopy.

The eventual guarantee of global convergence for the DIRECT algorithm does not necessarily hold for DIRECT-JTA due to the restart. At the end of each cycle, the new hyper-rectangle will not form a cover for the initial normalized domain $\bar{\Omega}$. To adopt the convergence properties of DIRECT for a new dilation level, we could make the restart hyper-rectangle large enough to cover this domain. However, this requirement would negate the explore and exploit structure of DIRECT-JTA by expanding the refined domains of the lower dilation levels. Therefore, by removing unpromising regions of $\bar{\Omega}$, DIRECT-JTA sacrifices a notion of global convergence for improved asymptotic performance.

III. PERFORMANCE ANALYSIS

In this section we give a performance analysis of DIRECT-JTA that addresses the convergence range from an initial pose, the values of the constants described in the algorithm, and the effects of image noise and model shape on accuracy. Images were collected from clinical trials studying motion cycles of the tibiofemoral joint. The baseline kinematics for the algorithm were established by a human operator in a

supervisory GUI environment at a dilation level of one. The initial poses supplied by the operator were sufficiently accurate to ensure the correct convergence of the routine after a large number of iterations (often several million) were evaluated. Note that because we do not know the true poses of the models, comparisons to these baseline kinematics test the ability of the search algorithm to find the optimum value of the metric.

A. Convergence Testing

To determine the furthest distance from the true pose at which DIRECT-JTA can consistently converge using ideal constants, we began by collecting 40 fluoroscopic images from two studies on stair stepping motion cycles at a capture rate of 30 frames per second with a resolution of approximately 0.75 megapixels and a density of .255 mm per pixel edge. Patients had been fitted with either GMK Sphere implants from Medacta International SA, Switzerland (Group A) or 3DKnee implants from DJO Surgical, Austin, Texas (Group B). Note each patient received two implants (one femoral and one tibial) leading to a total of four cases: Group A femoral/tibial and Group B femoral/tibial. After removing images where one or more of the implants were not fully in frame, ten images were randomly selected from each study. 100 random poses (ten per image) were subsequently generated for each of the four implant models at a specified maximum displacement range. We evaluated the algorithm for the femoral implants at ±25, ±35, ±45, and ±55 mm/degrees for each of the six pose components using independent Uniform distributions across the given ranges. Tibial implants were studied using ranges of ±25, ±30, ±35, and ±40 mm/degrees. This meant that each model group had 400 unique starting poses for a total of 1,600 trial runs. The performance of the algorithm was measured by three metrics: strict success rate, relaxed success rate, and the average convergence error from the baseline data of strict successes. The strict success rate is the percentage of random poses that DIRECT-JTA moved to within 1 mm for each of the in-plane translations and 1 degree for each of the rotation angles. Accuracy within ±1 mm/degree provides clinically usable information about joint kinematics [10], therefore we use this threshold as our margin of tolerance. However, this does not necessarily mean that strict success rate is a measure of clinical tolerance (for further discussion please see the concluding section of the article). The relaxed rate expands these boundaries to ±3 mm/degrees in order to gauge the ability of the algorithm to generate a close attempt. Note that out-of-plane translation is not considered in our definition of success. The motivation for this exclusion stems from the fact that single-plane, or monocular, radiographic images are extremely unreliable for determining out-of-plane position [10]; this is expected as the translation occurs along the line of sight. An orthogonal biplane comparison is likely the best resolution for this inaccuracy. It should be noted that differences in the registration success rates between groups are certainly not an indicator of implant quality. Each group also used different imaging devices which contributed to a slightly better contrast between the surrounding tissue/bone and the implants in Group B than in Group A.

Successful convergence of a black-box optimization problem is largely dependent on the complexity of the cost function and the size of the search range. Although in radiographic model-image registration each X-ray/implant pair has a unique similarity mapping, function complexity varies primarily with image noise and model type. This situation implies that the difficulty of finding the optimum pose for an implant is relatively consistent across an image sequence because the quality of images and intricacy of the silhouettes are generally the same (notable exceptions include an out-of-frame implant and interference from metallic objects).

B. Algorithm Constants

Like most optimization routines, the correct specification of the algorithm parameters was important to the success of DIRECT-JTA. Through trial and error the high-performance values for the dilation level (in pixels) and the Canny edge detector values were determined for each of the four cases. Significant effort was made to ensure the values for these constants were as general as possible, however a sensitivity analysis was also undertaken in a later section. The two other constants, iteration count and search range size, were set in accordance with available computing power and experiment pose displacement, respectively.

Both femur groups were set at dilation levels of 10 for the trunk search stage, while the tibia groups used a dilation level of 6. In all four cases, the dilation was decreased by a value of four in the branch search stage, and the final z-search stage was conducted at a dilation level of one. The Canny edge detection algorithm was used to remove noise from the biological materials, whilst retaining the edges of the implants. Using a fixed value of three for the aperture size of the Sobel operator (higher values led to unacceptable levels of interference) the threshold values for both implants in Group B and the tibial implants in Group A were set to 40 and 120. The femoral implants in Group A needed threshold values of 30 and 80 in order to display similar levels of clarity to the other implants.

Iteration count for the algorithm was pegged to take approximately one minute of operating time. The iteration budget was 50,000 for the trunk search, 15,000 for each of the three branch searches, and 50,000 for the z-translation search. Using a 4GHz Intel Core i7 4790K in conjunction with a Nvidia GeForce GTX 970, a GPU based version of the registration routine achieved speeds of approximately 3,000 iterations per second (50 seconds per frame).

The trunk search range sizes were set to the search range being tested, while the branch restart sizes were 15 mm for the in-plane translation, 25 mm for the out-of-plane translation, and 25° for each of the Euler angles. This allowed the branch search to easily overcome local minima that the highly dilated trunk search had identified due to symmetry. The in-plane translation components were set slightly smaller than the rest as the algorithm typically found good values for the x and y translation rather quickly. During the z-translation search phase the out-of-plane range was set to 20 mm while all other components were set to 5 mm/degrees.

TABLE I
AVERAGE CONVERGENCE ERROR AND SUCCESS RATE OF FEMUR IMPLANTS IN GROUP A

Search Range	Translation			Rotation			Success Rate	
	X	Y	Z	X	Y	Z	Strict	Relaxed
± 25	0.0342 mm	0.0382 mm	1.0689 mm	0.0490°	0.0413°	0.1524°	100%	100%
± 35	0.0387 mm	0.0420 mm	1.0280 mm	0.0553°	0.0416°	0.1453°	100%	100%
± 45	0.0354 mm	0.0365 mm	1.0598 mm	0.0388°	0.0430°	0.1455°	100%	100%
± 55	0.0334 mm	0.0442 mm	1.1223 mm	0.0649°	0.0397°	0.1502°	94%	94%

TABLE II
AVERAGE CONVERGENCE ERROR AND SUCCESS RATE OF FEMUR IMPLANTS IN GROUP B

Search Range	Translation			Rotation			Success Rate	
	X	Y	Z	X	Y	Z	Strict	Relaxed
± 25	0.0503 mm	0.0247 mm	1.3300 mm	0.0488°	0.0450°	0.1427°	100%	100%
± 35	0.0502 mm	0.0257 mm	1.2782 mm	0.0554°	0.0427°	0.1457°	100%	100%
± 45	0.0451 mm	0.0191 mm	1.2475 mm	0.0569°	0.0437°	0.1200°	100%	100%
± 55	0.0502 mm	0.0241 mm	1.3647 mm	0.0537°	0.0391°	0.1301°	95%	97%

TABLE III
AVERAGE CONVERGENCE ERROR AND SUCCESS RATE OF TIBIA IMPLANTS IN GROUP A

Search Range	Translation			Rotation			Success Rate	
	X	Y	Z	X	Y	Z	Strict	Relaxed
± 25	0.0277 mm	0.0631 mm	1.0665 mm	0.0741°	0.0964°	0.0382°	97%	98%
± 30	0.0274 mm	0.0650 mm	1.1408 mm	0.0558°	0.0740°	0.0273°	93%	96%
± 35	0.0322 mm	0.0683 mm	1.1561 mm	0.0634°	0.0977°	0.0472°	95%	96%
± 40	0.0245 mm	0.0666 mm	1.0163 mm	0.0670°	0.0934°	0.0365°	94%	97%

TABLE IV
AVERAGE CONVERGENCE ERROR AND SUCCESS RATE OF TIBIA IMPLANTS IN GROUP B

Search Range	Translation			Rotation			Success Rate	
	X	Y	Z	X	Y	Z	Strict	Relaxed
± 25	0.0383 mm	0.0496 mm	1.0616 mm	0.0331°	0.2642°	0.0614°	88%	94%
± 30	0.0396 mm	0.0542 mm	1.2448 mm	0.0370°	0.2198°	0.0647°	77%	89%
± 35	0.0386 mm	0.0502 mm	1.2610 mm	0.0432°	0.2739°	0.0645°	67%	85%
± 40	0.0379 mm	0.0485 mm	1.2310 mm	0.0422°	0.2417°	0.0664°	50%	65%

C. Convergence Range Results

Tables I - IV display the the success rates and the average convergence error of successful registrations for each of the four groups. The femoral implants performed extremely well through the ± 55 test set, with near perfect success rates and an average out-of-plane translation around one millimeter. The Group A tibial implants were similarly excellent at all ranges tested, however, the Group B set saw a sharp drop off for the strict success rate after the ± 25 search range. This was largely due to the highly symmetric nature of the tibial implant's position in many of the frames, which created a deep local minimum just over a degree along the Y Euler angle from the optimum. The relaxed success rate proved to be far more robust, although it also tapered off after the ± 35 search range. It should be noted that this data set also contained significant obstruction from the femoral implant leading to a reduction in the detectable tibial edges (although this problem did not seem to affect the femoral implant's convergence range).

1) Convergence Range and Automation: DIRECT-JTA demonstrated a femoral capture range of 55 millimeters for each translation component and 55 degrees for each rotation angle. While there was some variation between the data sets,

TABLE V
SENSITIVITY OF SUCCESS RATES TO CHANGES IN DILATION

Group Type	Search Range	Success Rate	Dilation Constant				
			$Value - 2$	$Value - 1$	$Value$	$Value + 1$	$Value + 2$
Group A Femur	$\pm\ 55$	*Strict*	96%	95%	94%	92%	84%
		Relaxed	96%	95%	94%	95%	90%
Group B Femur	$\pm\ 55$	*Strict*	92%	97%	95%	96%	96%
		Relaxed	92%	97%	97%	97%	98%
Group A Tibia	$\pm\ 25$	*Strict*	93%	92%	97%	95%	95%
		Relaxed	96%	95%	98%	96%	97%
Group B Tibia	$\pm\ 25$	*Strict*	66%	72%	88%	87%	89%
		Relaxed	85%	94%	94%	97%	100%

DIRECT-JTA optimized tibial implants at a maximum distance of approximately 25 to 40 millimeters for each translation component and 25 to 40 degrees for each rotation angle. Common pulsed x-ray fluoroscopy systems that are used for this type of imaging capture at speeds between 3.75 and 30 frames per second [21]. At these frame rates, the ranges for both implants should be sufficient to automate registration by using the previous frame's optimum pose as the current initial guess. In our own data the maximum average degree of freedom shift between frames for any of the implants was less than half the smallest convergence range of $\pm\ 25$ millimeters/degrees. Note that since we used Uniform variables, our experiments begin with an average degree of freedom error equal to half the convergence range.

D. Sensitivity of Dilation Constant and Noise Filtration

It is natural that potential users of DIRECT-JTA might be concerned about setting the array of different algorithm constants and the sensitivity of the method's performance to these constants. In Section III - B we gave high-performance femoral and tibial trunk dilation levels of 10 and 6, respectively. These values worked well for a variety of search ranges, however a potential user might have difficulty selecting a robust dilation value off the bat. This leads to the question, what level should the initial dilation be set at? To quantify the sensitivity of this constant, the original experiment was performed at $\pm\ 55$ for the femoral implants and $\pm\ 25$ for the tibial implants, but with trunk dilation $\pm\ 1$ and $\pm\ 2$ from the aforementioned values. The results, listed in Table V, indicate that the femur implants are relatively robust at trunk dilations of eight to twelve, while the Group A tibial implants performed well at dilation levels of four to eight. The Group B tibial implants favored a higher dilation in the range of six to eight. It is important to keep in mind that the ideal dilation values are likely to also depend on the size of the implant in pixels. A very small implant image (say 64 pixels across) would dilate to meaningless noise at the levels of dilation described in this paper. However,

in our experience most fluoroscopic images have similar implant pixel sizes to our data set.

Another factor to consider when preparing the DIRECT-JTA algorithm is the threshold values for the Canny edge detector. Ideally these values should be set to reduce noise from soft tissue and bone, whilst detecting as much of the implant boundary as possible. Finding the correct thresholds for high quality images is generally an easy task as long as there is no interference from other metallic objects. However, in low quality images a user might need to overshoot the sensitivity of the Canny algorithm in order to pick up the implant (leading to large amounts of background noise). Conversely, the Canny algorithm might need to be programmed to ignore the excessive amounts of noise in the low quality image - a setting which could fail to detect many of the implant's contours. In order to measure the effects of incorrectly specified Canny thresholds and/or low quality images, the original experiment was performed at $\pm\ 55$ for the femoral implants and $\pm\ 25$ for the tibial implants, but with altered Canny threshold values to simulate both highly noisy and insufficient edge detection. Threshold values were chosen to given consistent levels of noise/insufficient contours between the four implant classes tested. However, because of variations in image quality and implant contrast, different threshold values were used between several of the data sets. The highly noisy low/high threshold values were 20/65 for Group A femoral implants, and 30/80 for the remaining three implants. The insufficient edge detection low/high threshold values were 40/100 for Group A femoral implants, while the other implant groups used values of 100/140. The results are displayed in Table VI and show an adequate level of robustness to both forms of inaccurate implant edge detection.

IV. FINAL DISCUSSION

In this paper we have implemented a novel registration scheme that uses a pyramidal similarity metric and Lipschitzian optimization routine to find the three-dimensional

TABLE VI
SUCCESS RATES OF IMPLANTS WITH VARYING EDGE DETECTION

Group Type	Search Range	Highly Noisy Success Rate		Reduced Edges Success Rate	
		Strict	Relaxed	Strict	Relaxed
Group A Femur	± **55**	75%	82%	89%	93%
Group B Femur	± **55**	87%	89%	97%	98%
Group A Tibia	± **25**	79%	80%	77%	90%
Group B Tibia	± **25**	80%	95%	79%	97%

pose of a metallic implant from its radiographic projection. In an automated registration scheme that takes an initial guess from the previous image, DIRECT-JTA expands convergence to a margin sufficient for both femoral and tibial implants to successfully register even when there are large displacements between frames. The algorithm displays the ability to recover from significant interference from bones or tissues and can complete a registration on a consumer grade computer within a reasonable time frame. The increased convergence range represents a considerable improvement to existing methods which typically use tolerances about five times smaller per pose component [11]. Note the capture range is for the minimum of the cost function. Establishing the absolute accuracy from the true, physical pose would require alternate forms of testing. For example, Mahfouz et al. established the deviation of their method from the true pose was minimal using an optical sensor on a cadaver knee [3]. Translucent renderings of the optimal poses indicate that the cost function is a very good measure of success as the implants appear to be properly registered within a few millimeters/degrees. If a user desires confirmation that their registration is within clinical levels, the Mahfouz method could be run at the end of DIRECT-JTA as a form of cleanup. However, it should be noted that this does not necessarily mean DIRECT-JTA is less accurate; rather, it simply hasn't been validated against cadaver data.

Despite the promising capture range, there are cases when DIRECT-JTA might not achieve complete automation. If a large misregistration occurs, the pose can irreversibly drift in subsequent frames. This makes the algorithm as robust as the worst image - a severe issue for videos of movements that take an implant out of frame. The success of the algorithm is also dependent on the symmetry of the implant. Femoral implants are relatively simple to optimize because of their many protrusions, however tibial implants are easily trapped in local minima due to rotational symmetry. While the tibial range appears to be sufficient for automation purposes, DIRECT-JTA is greatly limited in its ability to register highly symmetric implants from other joints such as the glenosphere implant in a reverse shoulder replacement. DIRECT-JTA is also not suitable for bone registration as the similarity metric is edge-based and recovering the contours of a bone from a fluoroscopic image without significant noise is nontrivial.

The original DIRECT algorithm has parallel implementations [22] and many could easily be extended to DIRECT-JTA. Distributing the sampling process over a cluster computer would greatly reduce the registration time per frame. The number of hyper-rectangles deemed potentially optimal is at the very least two, and in practice can be much higher. A scheme that simultaneously renders the silhouette for each of these hyper-rectangles would scale almost proportionally up to this number.

There are several directions that future work on DIRECT-JTA could explore. The effectiveness of the algorithm on bi-plane fluoroscopy has not been properly determined. Although the constraint of an additional image plane can reduce error in the out-of-plane translation [10], it is possible that the resulting similarity function could be hard to traverse from a poor starting pose. There is also a need for a table of bounding boxes for pose shifts specific to common patient motion cycles. For instance, a seated leg extension gives similar in-plane locations but large changes in orientation. Tight specifications would reduce errors and allow for increased search in problematic areas like out-of-plane translation. Beyond DIRECT-JTA, further work on the effectiveness of pyramidal optimization for model-image registration could be very promising.

REFERENCES

[1] S. A. Banks, G. D. Markovich, and W. A. Hodge, "*In vivo* kinematics of cruciate-retaining and -substituting knee arthroplasties," *J. Arthroplasty*, vol. 12, no. 3, pp. 297–304, Apr. 1997.

[2] D. Zhao, S. A. Banks, and D. D. D'Lima, C. W. Colwell, Jr., and B. J. Fregly, "*In vivo* medial and lateral tibial loads during dynamic and high flexion activities," *J. Orthopaedic Res.*, vol. 25, no. 5, pp. 593–602, May 2007.

[3] M. R. Mahfouz, W. A. Hoff, R. D. Komistek, and D. A. Dennis, "A robust method for registration of three-dimensional knee implant models to two-dimensional fluoroscopy images," *IEEE Trans. Med. Imag.*, vol. 22, no. 12, pp. 1561–1574, Dec. 2003.

[4] 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, no. 6, pp. 638–649, Jun. 1996.

[5] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Comput. Surveys*, vol. 40, no. 2, pp. 5:1–5:60, Apr. 2008.

[6] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.

[7] A. A. Goshtasby, *2-D and 3-D Image Registration: For Medical, Remote Sensing, and Industrial Applications*. Hoboken, NJ, USA: Wiley, Apr. 2005.

[8] T. Yamazaki *et al.*, "Improvement of depth position in 2-D/3-D registration of knee implants using single-plane fluoroscopy," *IEEE Trans. Med. Imag.*, vol. 23, no. 5, pp. 602–612, May 2004.

[9] S. Zuffi, A. Leardini, F. Catani, S. Fantozzi, and A. Cappello, "A model-based method for the reconstruction of total knee replacement kinematics," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 981–991, Oct. 1999.

[10] S. Mu, "Skeletal kinematic measurement using model-image registration and mechanical constraints," Ph.D. dissertation, Dept. Mech. Aerosp. Eng., Univ. Florida, Gainesville, FL, USA, 2010.

[11] A. H. Prins, B. L. Kaptein, B. C. Stoel, D. J. P. Lahaye, and E. R. Valstar, "Performance of local optimization in single-plane fluoroscopic analysis for total knee arthroplasty," *J. Biomech.*, vol. 48, no. 14, pp. 3837–3845, Nov. 2015.

[12] S. Droste, T. Jansen, and I. Wegener, "Upper and lower bounds for randomized search heuristics in black-box optimization," *Theory Comput. Syst.*, vol. 39, no. 4, pp. 525–544, Jul. 2006.

[13] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *J. Global Optim.*, vol. 13, no. 4, pp. 455–492, 1998.

[14] N. Hansen, A. Auger, R. Ros, S. Finck, and P. Pošík, "Comparing results of 31 algorithms from the black-box optimization benchmarking BBOB-2009," in *Proc. 12th Annu. Conf. Companion Genet. Evol. Comput.*, New York, NY, USA, 2010, pp. 1689–1696.

[15] B. O. Shubert, "A sequential method seeking the global maximum of a function," *SIAM J. Numer. Anal.*, vol. 9, no. 3, pp. 379–388, 1972.

[16] D. R. Jones, C. D. Pertunnen, and B. E. Stuckman, "Lipschitzian optimization without the Lipschitz constant," *J. Optim. Theory Appl.*, vol. 79, no. 1, pp. 157–181, 1993.

[17] T. L. Nguyen and K.-S. Low, "A global maximum power point tracking scheme employing DIRECT search algorithm for photovoltaic systems," *IEEE Trans. Ind. Electron.*, vol. 57, no. 10, pp. 3456–3467, Oct. 2010.

[18] M. P. Wachowiak and T. M. Peters, "High-performance medical image registration using new optimization techniques," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 2, pp. 344–353, Apr. 2006.

[19] D. E. Finkel and C. T. Kelley, "Additive scaling and the DIRECT algorithm," *J. Global Optim.*, vol. 36, no. 4, pp. 597–608, Dec. 2006.

[20] D. E. Finkel and C. T. Kelly, "An adaptive restart implementation of DIRECT," Dept. Math., Center Res. Sci. Comput., North Carolina State Univ., Raleigh, NC, USA, Tech. Rep. CRSC-TR04-30, 2004.

[21] J. T. Bushberg, *The Essential Physics of Medical Imaging*. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2002.

[22] L. T. Watson and C. A. Baker, "A fully distributed parallel global search algorithm," *Eng. Comput.*, vol. 18, nos. 1–2, pp. 155–169, Feb. 2001.

[23] D. E. Finkel, "DIRECT optimization algorithm user guide," Center Res. Sci. Comput., North Carolina State Univ., Raleigh, NC, USA, Tech. Rep. CRSC-TR03-11, 2003.

[24] J. M. Gablonsky, "DIRECT version 2.0 user guide," Dept. Math., North Carolina State Univ., Raleigh, NC, USA, Tech. Rep. CRSC-TR01-08, 2001.

[25] C. A. Floudas and P. Pardalos, "Direct global optimization algorithm," in *Encyclopedia of Optimization*, 2nd ed. New York, NY, USA: Springer, 2009, pp. 725–735.