

# An ITK Framework for Deterministic Global Optimization for Medical Image Registration

Florence Dru<sup>a,b</sup>, Mark P. Wachowiak<sup>\*a,c</sup>, Terry M. Peters<sup>a,c</sup>

<sup>a</sup>Robarts Research Institute, 100 Perth Drive, London, ON CANADA N6A5K8;

<sup>b</sup>Université Rennes 1, Rennes, FRANCE;

<sup>c</sup>Department of Medical Biophysics, University of Western Ontario, London, ON CANADA N6A5B8

## ABSTRACT

Similarity metric optimization is an essential step in intensity-based rigid and nonrigid medical image registration. For clinical applications, such as image guidance of minimally invasive procedures, registration accuracy and efficiency are prime considerations. In addition, clinical utility is enhanced when registration is integrated into image analysis and visualization frameworks, such as the popular Insight Toolkit (ITK). ITK is an open source software environment increasingly used to aid the development, testing, and integration of new imaging algorithms. In this paper, we present a new ITK-based implementation of the DIRECT (Dividing Rectangles) deterministic global optimization algorithm for medical image registration. Previously, it has been shown that DIRECT improves the capture range and accuracy for rigid registration. Our ITK class also contains enhancements over the original DIRECT algorithm by improving stopping criteria, adaptively adjusting a locality parameter, and by incorporating Powell's method for local refinement. 3D-3D registration experiments with ground-truth brain volumes and clinical cardiac volumes show that combining DIRECT with Powell's method improves registration accuracy over Powell's method used alone, is less sensitive to initial misorientation errors, and, with the new stopping criteria, facilitates adequate exploration of the search space without expending expensive iterations on non-improving function evaluations. Finally, in this framework, a new parallel implementation for computing mutual information is presented, resulting in near-linear speedup with two processors.

**Keywords:** Registration, Insight Toolkit (ITK), global optimization, dividing rectangles (DIRECT)

## 1. INTRODUCTION

Registering, or geometrically aligning images or volumes from the same or from different modalities is a vital step in image-guidance for minimally invasive surgery and therapy, pre-interventional planning, developing patient-specific models, and for assessing therapy response. For example, in the planning and execution of image-guided cardiac surgery, preoperative dynamic models are registered to the patient<sup>1</sup>, preoperative 3D vessel models are registered to intraoperative angiography<sup>2</sup>, endoscopic views are registered to the preoperative model<sup>3</sup>, and real-time ultrasound<sup>4</sup> or MRI<sup>5</sup> are registered to high quality preoperative MR volumes. Many recent registration techniques directly use image intensities to compute a similarity metric, such as mutual information<sup>5</sup>, between the source and target images or volumes. The similarity metric is then optimized with one of any number of optimization approaches. If derivatives are available or are easily estimated, then conjugate gradient descent or Levenberg-Marquardt trust-region methods provide fast convergence<sup>6</sup>. Widely used derivative-free methods used in medical image registration include the Nelder-Mead simplex algorithm and Powell's direction set technique, the latter being used most often for intensity based medical image registration<sup>7,8</sup>. Although these optimization approaches are efficient, robust, and accurate, they are inherently local in scope, and generally have limited capture ranges.

Global optimization attempts to avoid premature convergence to local extrema, which are common in many similarity metrics currently used in medical image registration, particularly those based on information theory. For example, Fig. 1 illustrates a mutual information (MI) objective function of a T1 high-noise source volume aligned with a PD low-

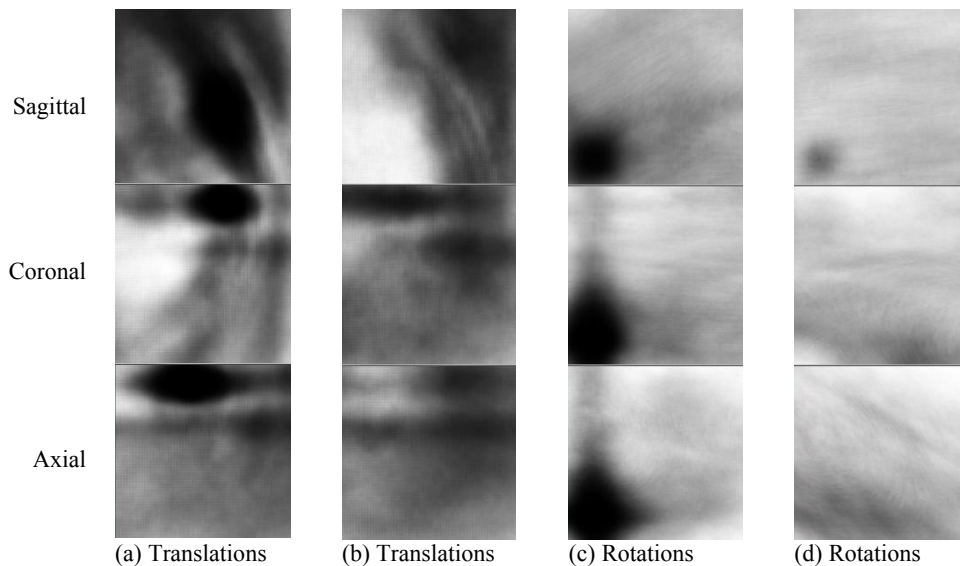
---

\* mwach@imaging.robarts.ca; Phone: 1 519 663 5777 x34280

noise target volume. The function is computed from the source-target overlap as the source is translated and rotated with respect to the target. As can be seen from Fig. 1, the “convergence basin” for the global minimum, represented by the large dark elliptical areas in Figures 1(a) and (c), are prominent when the source and target volumes are in proper alignment. However, the objective function also has many local minima and large, ill-defined areas of low function values, which are particularly evident when the source and target are misaligned, as seen in Figures 1(b) and (d).

Although it improves robustness, global optimization is not currently used widely in medical image registration because of its high computational cost in terms of increased similarity metric evaluations and slower convergence. The most common global optimization approaches are stochastic, and those used in registration include genetic algorithms and evolutionary strategies<sup>9-11</sup>, simulated annealing<sup>9</sup>, and particle swarm optimization<sup>12</sup>. However, the stochastic nature of many global optimization algorithms makes their use in clinical or time-critical applications problematic. Finally, many robust global optimization approaches are difficult to implement because of sensitivity to operating parameters, and because of complex data structures and bookkeeping requirements of some methods<sup>13</sup>.

In this paper, we present a deterministic approach using the Dividing Rectangles (DIRECT) algorithm implemented within the open source ITK (Insight Toolkit) framework. ITK is becoming increasingly important in the medical imaging community, as evidenced by the number of reported applications<sup>14</sup>, including educational purposes<sup>15</sup>. Furthermore, there is a great need for automated medical image analysis techniques, and the public exchange of software tools can potentially accelerate improvement of health care<sup>14</sup>. Although the public release of ITK contains many registration approaches and optimization methods, it currently supports only one global approach (the one-plus-one evolutionary technique). This paper extends our previous work with DIRECT by: (1) Introducing this deterministic global approach into ITK; (2) Implementing new convergence criteria and adaptive adjustment of a locality parameter, and (3) Parallelizing the computation of MI within the ITK framework. In previous work by the authors, a parallelized adaptation of DIRECT was proposed, wherein the similarity metrics for different regions of the search space were computed simultaneously. Parallelization of MI computation was also developed for Powell’s method<sup>8,13</sup>.



**Figure 1.** MI cost function for alignment of T1 MRI to PD MRI BrainWeb volumes. (a) MI with respect to translations. The origin at the 3 orthogonal is the orientation at correct alignment. (b) MI with respect to translations. Origin at a point 30 mm from correct alignment. (c) MI with respect to rotations. Origin at correct alignment. (d) MI with respect to rotations. Origin placed at 30° about all coordinate axes.

## 2. ITK AND THE REGISTRATION PROCESS

In 1999, the US National Library of Medicine of the National Institute of Health awarded six three-year grants to develop an open-source registration and segmentation toolkit that eventually came to be known as the Insight Toolkit (ITK), and formed the basis of the Insight Toolkit Software Consortium. In 2002, the first official public release of ITK was made available. As it is fairly recent, ITK is in a state of perpetual evolution. ITK relies on object-oriented software engineering principles, and is readily extendible by its community of users.

ITK is implemented in C++. It is cross-platform, using the CMake build environment to manage the compilation process. CMake generates native makefiles and workspaces that can be used in the compiler environment of the user's choice. In addition, an automated wrapping process generates interfaces between C++ and interpreted programming languages such as Tcl, Java, and Python. This enables developers to create software using a variety of programming languages. ITK's C++ implementation style, which makes use of templated code, is referred to as generic programming. C++ templating facilitates highly efficient code, and many software problems are discovered at compile-time, rather than at run-time.

In ITK, registration is performed within a framework of components that can be easily plugged and interchanged, allowing users to select the tools most appropriate for their specific application. Optimization algorithms are encapsulated as `itk::Optimizer` objects. These objects are generic and can be used for a variety of applications. In its current release, ITK supports the following optimization methods: Levenberg-Marquardt, Nelder-Mead simplex, gradient descent, including a specialized version for quaternion transforms, Fletcher-Reeves conjugate gradient descent, the LBFGS algorithm, regular step gradient descent, and Powell's method<sup>16</sup>. For registration, the basic input to an optimizer is a cost function object, whose corresponding root class is the `itk::ImageToImageMetric`. Any subclass that inherits from the `itk::ImageToImageMetric` class thus provides this functionality. As the value of similarity metrics is being optimized in the present application, the DIRECT optimizer class, `itk::DirectOptimizer`, has been derived from the `itk::SingleValuedNonLinearOptimizer` base class.

## 3. THE DIRECT ALGORITHM

### 3.1 Description of DIRECT

DIRECT algorithm is a Lipschitzian-based pattern search method<sup>17</sup>. Considering the one-dimensional problem of finding the global minimum of a function  $f(x)$  defined on the closed interval  $[l, u]$ , the standard Lipschitzian approach assumes that there exists a finite bound on the rate of change  $f$ . In other words, there exists a positive constant  $K$ , the Lipschitz constant, such that:

$$|f(x) - f(x')| \leq K|x - x'|, \quad x, x' \in [l, u] \quad (1)$$

DIRECT is based on an iterative division of  $n$ -dimensional space, represented as a hypercube or hyperrectangle. The essential feature of DIRECT is that  $K$  need not be explicitly specified, but is computed on the basis of the size of the hyperrectangles and the function value at the center of the hyperrectangles. The basic algorithm follows:

1. Normalize the search space to be the unit hyperrectangle. Set iteration counter  $k$  to 0. Define this hyperrectangle as "potentially optimal" (see below).
2. For each potentially optimal hyperrectangle:
  - a. Let  $\mathbf{c}$  be the center of the hyperrectangle, and evaluate  $f(\mathbf{c})$ . Set  $f_{\min} = f(\mathbf{c})$ .
  - b. For each dimension of the hyperrectangle having the longest side length, sample two points on either side of  $\mathbf{c}$ .
  - c. Evaluate  $f$  at these points.
3. Divide the hyperrectangles so that new hyperrectangles centered at all the sampled points are generated.
4. Identify the set of potentially optimal hyperrectangles that will be further divided in the next iteration.
5. Determine the new current best objective function value  $f_{\min}$ .

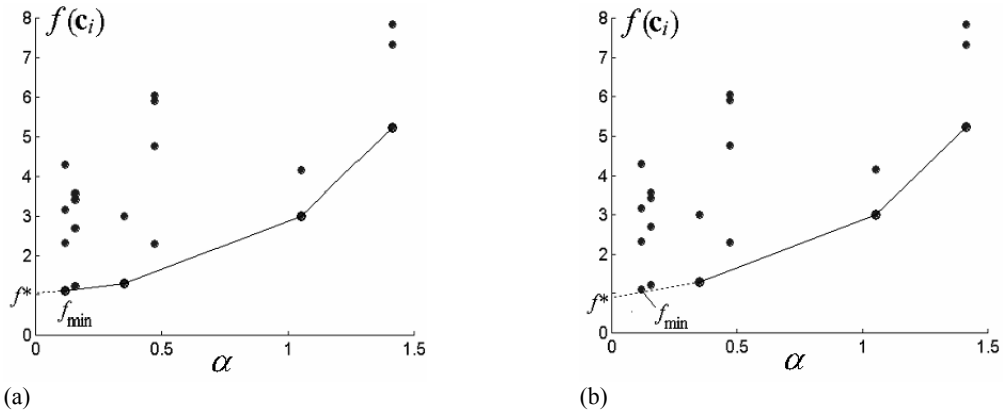
6. Check the convergence criteria. If they are not satisfied, increment  $k$  and go to Step 2.

In Step 1, the first function value is computed at the center  $\mathbf{c}$  of the  $n$ -D hyperrectangle, normalized to unit side length. As there is only one hyperrectangle at this point, the hypercube is considered to be potentially optimal. That is, the global optimum is potentially contained within this hyperrectangle. Points are then evaluated around  $\mathbf{c}$  at locations  $\mathbf{c} \pm 1/3 \mathbf{e}_i$ , for  $i = 1, \dots, n$ , and where  $\mathbf{e}_i$  is the  $i$ -th unit vector (Step 2b). The original hyperrectangle is then divided, first along the dimension that produced the smallest function value, then along the dimension with the next smallest value, and continuing in this manner until all sides have been divided (Step 3). Thus, a new set  $S$  of hyperrectangles are generated with centers at  $\mathbf{c}_i = \mathbf{c}_0 \pm 1/3 \mathbf{e}_i$ , and the minimum is denoted as  $f_{\min}$ . In subsequent iterations, it must be determined which of the newly formed hyperrectangles are potentially optimal (Step 4). A hyperrectangle  $R \in S$  with center  $\mathbf{c}_R$  and size  $\alpha(R)$  is considered as potentially optimal if there is a positive constant  $K'$  such that:

$$f(\mathbf{c}_R) - K' \alpha(R) \leq f(\mathbf{c}_T) - K' \alpha(T) \quad \text{for all } T \in S, \text{ and} \quad (2)$$

$$f(\mathbf{c}_R) - K' \alpha(R) \leq f_{\min} - \varepsilon |f_{\min}|, \quad \text{where } \varepsilon \geq 0. \quad (3)$$

Here,  $K'$  is the estimate of the Lipschitz constant, and  $\varepsilon$  is a parameter – the only one required by DIRECT – that controls the locality of the search. Whether  $R$  satisfies Eqs. (2) and (3), and therefore, whether it is potentially optimal and selected for division, can be determined from the convex hull of the numerical values at the center of all hyperrectangles generated, plotted against the corresponding hyperrectangle size, as shown in Fig. 2. This figure also illustrates the effect of  $\varepsilon$ . In Fig. 2(a),  $\varepsilon$  is set to zero, and thus  $f^* = f_{\min} - \varepsilon |f_{\min}| = f_{\min}$ . The hyperrectangle containing  $f_{\min}$ , in this case also the one with the smallest size, is selected for division, resulting in a more local search. In Fig. 2(b),  $\varepsilon = 0.01$ , and no hyperrectangle with the smallest size is selected for division, although one contains  $f_{\min}$ . This has the effect of dividing larger hyperrectangles, resulting in a more global search. The size  $\alpha(R)$  of hyperrectangle  $R$  is usually computed as the diameter,  $\alpha(R) = \alpha(l, p) = 3^{l(n-8p/9)^{1/2}}$ , where  $R$  has long sides of length  $3^{-l}$  and  $p$  short sides of length  $3^{-(l+1)}$ . Details of DIRECT and its convergence properties are described in the literature<sup>17-19</sup>.



**Figure 2.** Convex hull formed by values of  $f(\mathbf{c}_i)$  plotted against hyperrectangle sizes.  $f^* = f_{\min} - \varepsilon |f_{\min}|$ . (a),  $\varepsilon = 0$ . (b),  $\varepsilon = 0.01$ .

### 3.2 Adaptive Adjustment of the Locality Parameter

As stated above, in its basic implementation DIRECT requires only one parameter,  $\varepsilon$ , which quantifies the locality of the search. As can be seen from Fig. 2(a), if  $\varepsilon \approx 0$ , then there is a local bias, as small  $R$  having centers with good function values are subdivided. This intensive search around the  $\mathbf{c}^*$  (the point with the best minimum found thus far:  $f(\mathbf{c}^*) = f_{\min}$ ) may result in incremental improvement of  $f_{\min}$ , but at the expense of potentially better solutions that are contained in larger and more distant hyperrectangles. To address this issue,  $\varepsilon$  can be adaptively adjusted during the course of the search<sup>18</sup>. If stagnation is detected (i.e.  $f_{\min}$  does not decrease appreciably in a specified number of iterations), then  $\varepsilon$  can be set to a different value. In the present implementation,  $\varepsilon$  is set to 0.01, biasing the search to global scope. If there is no improvement in five iterations,  $\varepsilon$  is reduced to 0.001. If there is still no improvement after

five additional iterations, then  $\varepsilon$  is set to 0 to force exploitation of the search space around  $\mathbf{c}^*$ . However, if there is no further improvement when  $\varepsilon = 0$ , then the algorithm is likely trapped in a local minima (or is possibly in the vicinity of the global minima), and  $\varepsilon$  is reset to 0.01 to facilitate more global exploration.

### 3.3 Convergence Criteria

In most implementations of DIRECT, the procedure is stopped when a maximum number of function evaluations have been expended. However, for problems in high dimensions, a large number of evaluations is required to adequately explore the search space. In clinical applications, where registration must be performed very quickly, a fixed number of evaluations is not feasible because of the high computation cost, and because of the difficulty in empirically assigning a maximum number of evaluations. It is therefore desirable that stopping criteria reduce the number of function evaluations (specifically, non-improving iterations) without sacrificing accuracy. In the current implementation, termination criteria are as follows:

1. The current  $f_{\min}$  is within a small tolerance of the previous  $f_{\min}$ .
2. The hyperrectangle centered at  $\mathbf{c}^*$  is sufficiently small.
3. The current  $\mathbf{c}^*$  and previous  $\mathbf{c}^*$  are sufficiently close, as measured by some metric (i.e. Euclidean distance), indicating that it is unlikely that any further exploration of the search space will result in a better function value.

Condition (1) ensures that the function values are appreciably decreasing (improving). Condition (2) ensures that the search space around  $\mathbf{c}^*$  has been adequately explored, because if a hyperrectangle  $R$  is the result of  $k$  divisions, then the neighboring hyperrectangles have resulted from either  $k$  or  $k - 1$  divisions. Therefore, small  $R$ s are necessarily clustered together, and the sizes of the  $R$ s are indicators of the degree of clustering. If conditions (1) and (2) are satisfied, then it is still possible that there is more than one cluster, indicating that the search space must be further explored. Therefore, condition (3) has been imposed to ensure that if multiple clusters are observed, they are sufficiently close to be considered to belong to the same convergence basin.

### 3.4 Parallelization

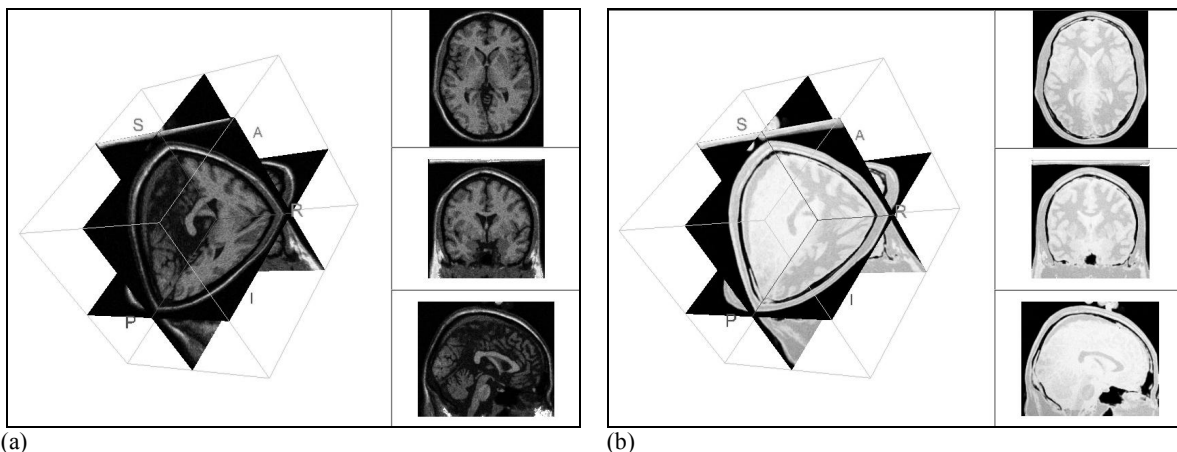
As the computations of each  $f(\mathbf{c})$  in Step 2c of the DIRECT algorithm are independent of each other, DIRECT is intrinsically parallel. Therefore, applying a geometric transformation, determining the overlap region, interpolation, estimation of joint and marginal densities, and computing the metric can be performed for many regions of the search space simultaneously<sup>13</sup>. However, in its present release, ITK is not conducive to this high level parallelism because of dependencies within the object-oriented class hierarchy. ITK does provide utilities for low-level parallelization through a *pthread*s mechanism. Therefore, the computation of Mattes mutual information was parallelized with a shared-memory multithreading approach. From run-time profiling, it was observed that joint histogram estimation consumes a substantial amount of computation time. Consequently, most of the efficiency gains from multithreading are obtained by parallelizing this step.

## 4. METHODOLOGY

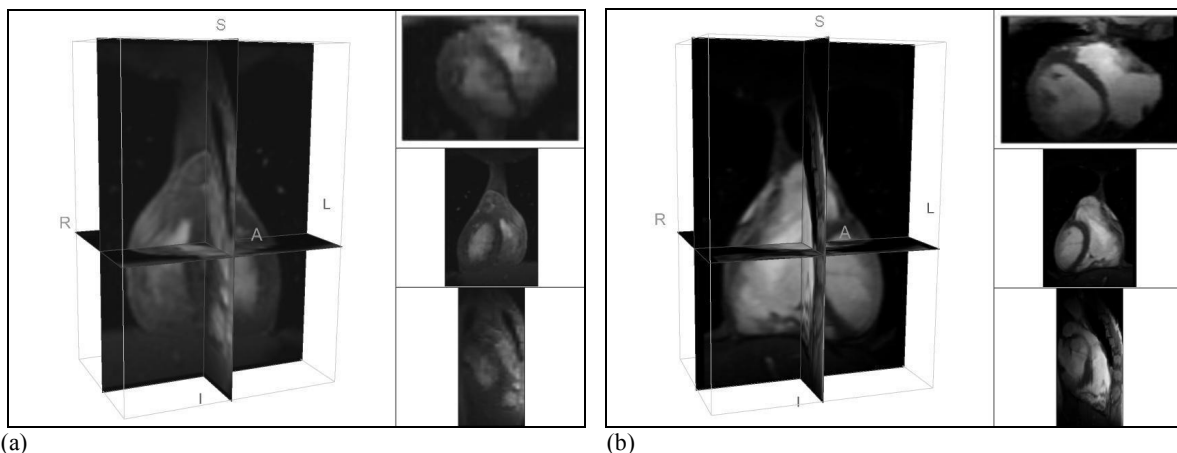
Rigid body (6D) registration experiments were performed on two sets of data. The first set consisted of simulated BrainWeb volumes (Montreal Neurological Institute) with  $181 \times 217 \times 181$  1-mm<sup>3</sup> voxels, shown in Fig. 3. BrainWeb volumes are realistic simulations generated from real MRI data, and are used extensively in the neuroscience community for developing and validating segmentation and registration algorithms<sup>20</sup>. The source image (Fig. 3(a)) was a simulation of a normal T1 MRI volume with 9% noise. The target (Fig. 3(b)) was a simulated PD MRI brain volume with MS lesions, with 3% noise. The noise percentage is the ratio of the standard deviations of the noise and signal for the reference tissue. The noise is Rayleigh distributed in background regions, and Rician in signal regions. The second data set consisted of real clinical MRI cardiac volumes with  $160 \times 107 \times 75$  1.5-mm<sup>3</sup> voxels, as shown in Fig. 4. The source was a low SNR MRI volume acquired at systole, originally with 1.5 mm slice thickness (Fig. 4(a)), and the target was a high SNR MRI volume acquired at diastole, with the same dimensions as the source (Fig. 4(b)). Ground truth alignments were known for all data sets, although with a small degree of uncertainty for the clinical MRI images.

Source volumes were initially mistranslated at  $d_0 = 5, 10, 20,$  and  $40$  mm, and misrotated at  $\pm 10, 20, 30,$  and  $40^\circ$  about the three coordinate axes.

Mattes mutual information (MI), as implemented in ITK, was used as the similarity metric. 160 experiments were performed for each initial distance. Performance was judged on translation/rotation errors and the number of function evaluations. The optimization methods were DIRECT, given a maximum budget of 50 iterations, and DIRECT with the new termination criteria described above. Because global techniques perform well in identifying the convergence basin of the global optimum, for both DIRECT methods, Powell's method was used to locally refine the result. Results were compared with Powell's method, used alone. Powell's method was chosen for comparison because of its ubiquitous use in medical image registration<sup>7,13</sup>. The DIRECT/Powell multiresolution approach was performed by computing MI from 100,000 samples from the source/target overlap during the global optimization phase with DIRECT, and the resulting transformation was used as the initial orientation for local refinement with Powell's method, using 200,000 samples from the source/target overlap. For Powell's method, 200,000 samples for single-resolution registration were used. MI was computed with 64 histogram bins. All experiments were performed with 1, 2, 4, 8, and 16 CPUs on an SGI Altix shared memory system, with each CPU running at 1.3 GHz.



**Figure 3.** BrainWeb MRI volumes. (a) Source: T1 MRI, Normal, 9% noise. (b) Target: PD MRI, MS lesion, 3% noise.



**Figure 4.** Cardiac MRI volumes. (a) Source: Systole, low SNR. (b) Target: Diastole, high SNR.

## 5. RESULTS

Results for the translation error, rotation error, and number of function evaluations for the BrainWeb registration experiments with a budget of 50 iterations are shown in Table 1. These values for the proposed termination criteria are shown in Table 2. The results for Powell's method, used alone, are given in Table 3. In these tables, results are reported over all initial misrotations. As can be seen in Tables 1-2, DIRECT, used in conjunction with Powell's method for local refinement, consistently had translation errors less than 1 mm and rotation errors of less than  $0.3^\circ$  for all initial mistranslation distances  $d_0$  and for all initial misrotations. The proposed termination method resulted in only marginally higher errors than with the fixed budget of iterations, but errors were still under 1 mm for all  $d_0$ . However, the number of function evaluations decreased substantially with the new method. The translation and rotation errors were much larger for Powell's method (Table 3), with the severity of the errors generally increasing with  $d_0$ . As a strictly local technique, Powell's method performed best for small  $d_0$  ( $d_0 = 5$  and 10 mm). Although the translation errors for these distances were higher than 9mm, errors were less with small initial misrotations ( $\pm 10, 20^\circ$ ).

An analysis of the results also showed that Powell's method, used alone, often exhibited erratic convergence behavior. This can be seen in Tables 3 and 6, where the standard deviations for the translation and rotation errors are higher than the means. For initial orientations outside the convergence basin of the global minimum, some very large errors were observed (for example, translation errors over 6mm), while many errors were much lower (less than 2 mm). Consequently, the distribution of translation and rotation errors were not Gaussian, but were highly skewed. Therefore, although Powell's method often converges to the correct registration, even for large initial misorientations, accuracy is inconsistent. The DIRECT algorithm, with lower standard deviations for the errors, is more consistent.

Results for the cardiac experiments are shown in Tables 4-6. Because of the inherent uncertainty in measuring ground truth with real clinical data, the reported translation and rotation errors were somewhat higher than for the BrainWeb experiments. The mean translation errors were in the range of 3 mm and the rotation errors were less than  $5^\circ$ . The errors with Powell's method were much higher, as seen in Table 6.

Computation times and the effect of parallelization were also analyzed, and are respectively displayed in Tables 7 and 8 for BrainWeb and cardiac registration. The speedup is computed from the total registration time, consisting of a low-resolution global search with DIRECT followed by local refinement with Powell's method. Speedup is defined as (Computation time for 1 CPU) / (Computation time for  $j$  CPUs), where  $j = 1, 2, 4, 8, \text{ and } 16$ . Ideally, speedup is linear. That is, using  $j$  CPUs should result in a speedup of  $j$ . In practice, however, most useful algorithms rarely achieve linear speedup because of thread synchronization and overhead costs, and because most algorithms contain inherently serial sections. For the BrainWeb and cardiac registrations, near linear speedup was achieved with two processors (speedup = 1.7 in both cases). Speedup decreased for a higher number of CPUs. However, computation time was still reduced.

**Table 1.** BrainWeb experiments. DIRECT/Powell, 50 iterations with DIRECT.

$d_0$	Translation error (mm)	Rotation error ( $^\circ$ )	Function evaluations
5	$0.72 \pm 0.19$	$0.24 \pm 0.10$	$874 \pm 176$
10	$0.76 \pm 0.17$	$0.28 \pm 0.09$	$795 \pm 164$
20	$0.72 \pm 0.19$	$0.28 \pm 0.10$	$790 \pm 121$
30	$0.72 \pm 0.19$	$0.26 \pm 0.10$	$810 \pm 137$
40	$0.71 \pm 0.19$	$0.26 \pm 0.10$	$820 \pm 139$

**Table 2.** BrainWeb experiments. DIRECT/Powell, new stopping criteria.

$d_0$	Translation error (mm)	Rotation error (°)	Function evaluations (DIRECT)	Function evaluations (Powell)
5	$0.81 \pm 0.24$	$0.15 \pm 0.14$	$286 \pm 39$	$204 \pm 90$
10	$0.78 \pm 0.23$	$0.16 \pm 0.13$	$337 \pm 46$	$248 \pm 105$
20	$0.79 \pm 0.26$	$0.14 \pm 0.15$	$373 \pm 54$	$262 \pm 119$
30	$0.79 \pm 0.22$	$0.14 \pm 0.14$	$394 \pm 71$	$278 \pm 100$
40	$0.79 \pm 0.24$	$0.13 \pm 0.14$	$398 \pm 61$	$264 \pm 103$

**Table 3.** BrainWeb experiments. Powell's method, used alone.

$d_0$	Translation error (mm)	Rotation error (°)	Function evaluations
5	$9.58 \pm 48.93$	$8.09 \pm 42.24$	$456 \pm 99$
10	$13.16 \pm 54.53$	$9.80 \pm 43.78$	$517 \pm 138$
20	$17.05 \pm 61.94$	$7.78 \pm 35.29$	$552 \pm 135$
30	$22.29 \pm 64.21$	$15.96 \pm 123.00$	$699 \pm 186$
40	$14.18 \pm 50.31$	$7.96 \pm 30.31$	$775 \pm 202$

**Table 4.** Cardiac experiments. DIRECT/Powell, 100 iterations with DIRECT.

$d_0$	Translation error (mm)	Rotation Error (°)	Function evaluations
5	$2.56 \pm 1.96$	$4.31 \pm 0.33$	$2057 \pm 201$
10	$2.40 \pm 1.79$	$4.41 \pm 0.42$	$2039 \pm 245$
20	$3.38 \pm 3.23$	$4.46 \pm 0.61$	$1980 \pm 300$
30	$3.25 \pm 2.45$	$4.35 \pm 0.50$	$2062 \pm 255$
40	$2.09 \pm 1.20$	$4.60 \pm 0.41$	$2043 \pm 311$

**Table 5.** Cardiac experiments. DIRECT/Powell, new stopping criteria.

$d_0$	Translation error (mm)	Rotation error (°)	Function evaluations
5	$3.51 \pm 2.78$	$4.38 \pm 0.47$	$1563 \pm 548$
10	$3.49 \pm 3.01$	$4.54 \pm 0.44$	$1419 \pm 666$
20	$5.27 \pm 3.92$	$4.64 \pm 0.81$	$1165 \pm 713$
30	$4.90 \pm 3.96$	$4.60 \pm 0.79$	$1129 \pm 680$
40	$3.30 \pm 2.89$	$4.45 \pm 0.52$	$1005 \pm 672$

**Table 6.** Cardiac experiments. Powell's method used alone.

$d_0$	Translation error (mm)	Rotation error (°)
5	$24.23 \pm 46.46$	$13.94 \pm 46.47$
10	$22.83 \pm 47.44$	$12.91 \pm 25.48$
20	$30.79 \pm 61.03$	$17.41 \pm 34.09$
30	$27.24 \pm 59.28$	$18.90 \pm 44.11$
40	$47.07 \pm 75.40$	$26.54 \pm 38.89$

**Table 7.** Execution time scaling for BrainWeb experiments.

Number of CPUs	DIRECT (sec)	Powell (sec)	Total DIRECT/Powell (sec)	Speedup
1	72.05 ± 21.80	107.90 ± 54.09	179.96 ± 69.64	1.0
2	43.23 ± 7.37	61.66 ± 25.54	104.89 ± 29.59	1.7
4	25.52 ± 4.35	32.06 ± 13.00	57.59 ± 15.48	3.1
8	20.07 ± 3.87	21.88 ± 9.25	41.94 ± 11.74	4.3
16	15.64 ± 3.17	14.63 ± 6.38	30.28 ± 8.44	5.9

**Table 8.** Execution time scaling for Cardiac experiments.

Number of CPUs	DIRECT (sec)	Powell (sec)	Total DIRECT/Powell (sec)	Speedup
1	149.13 ± 90.83	54.15 ± 27.66	203.27 ± 83.31	1.0
2	87.67 ± 52.63	29.10 ± 14.77	116.77 ± 48.39	1.7
4	51.80 ± 30.43	14.79 ± 7.35	66.59 ± 28.16	3.0
8	36.68 ± 21.09	8.64 ± 4.27	45.32 ± 19.73	4.5
16	30.15 ± 16.91	5.31 ± 2.53	35.46 ± 16.03	5.7

## 6. DISCUSSION

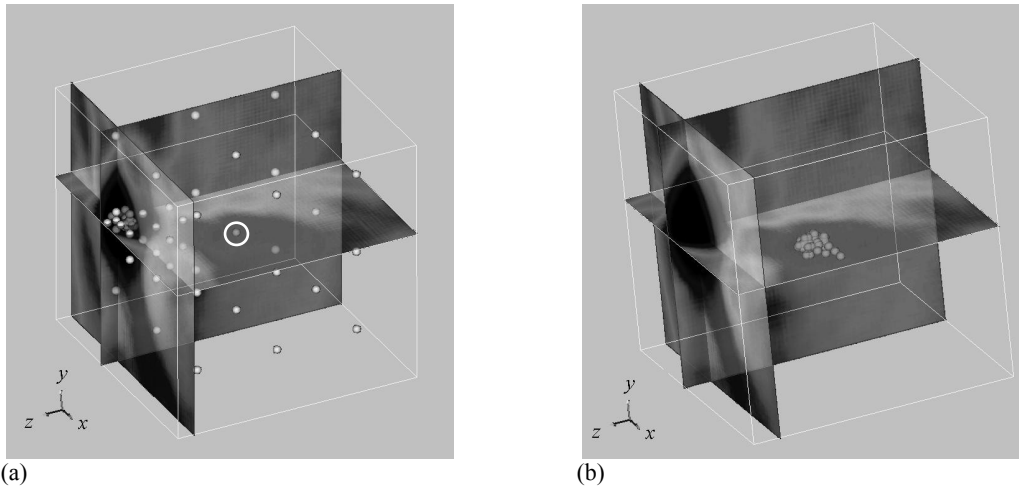
Global optimization has been shown to improve registration performance in a number of studies<sup>9-13</sup>, primarily because of the ability to escape the local minima that characterize many useful similarity metrics. The larger capture range of DIRECT, compared with Powell's method, is illustrated in Tables 1-3 for the BrainWeb registrations and Tables 4-6 for the cardiac experiments. Whereas the performance of Powell's method is dependent upon the closeness of the initial orientation to optimal alignment (errors increase with  $d_0$ ), with DIRECT, the magnitudes of the errors are consistent for all  $d_0$  and for all initial misrotations. The situation is illustrated in Fig. 5, which shows a 3D representation of the MI registration function (with respect to translations) for the BrainWeb MRI volumes (also see Fig. 1). The large dark elliptical area represents the global minimum, indicating correct alignment. The source volume has been initially placed 30mm from ground truth, as indicated by the central sphere. The source was also rotated 30° about all three axes. Because of its global scope, DIRECT samples the entire search space, eventually converging to the region containing the minimum. By contrast, Powell's method quickly converges to a local minimum. The effect of such local entrapment from Powell's method is seen in Fig. 6, where the source and target prior to and after registration are overlaid.

A potential difficulty with DIRECT (and global methods in general) is the high number of function evaluation required. In general, global optimization methods are most successful at determining a small region containing the global minimum. However, once this region is identified, convergence is very slow. To address this issue, the approach used in a previous study by the authors<sup>13</sup> was adopted: An global exploration of the search space is performed at a low resolution with DIRECT so that the convergence basin for the global minimum (or the "best vicinity") is located, and then an efficient local optimization technique is used to find the minimum in this vicinity. Termination criteria for DIRECT, based on the clustering of the minima described above, facilitate an adequate search with reasonable computation times.

The results presented in this paper were obtained with the adaptive locality parameter  $\epsilon$ <sup>18</sup>. It was found that this feature marginally reduced registration time by localizing the search when promising regions have been located. In the early stages of the search, adaptive adjustment of  $\epsilon$  helped to avoid stagnation in local minima regions. However, in the later stages, the vicinities of the minima were more intensively explored, and fewer iterations were expended in non-improving regions. Consequently, termination criteria were achieved faster, allowing Powell's method to quickly

converge to the minimum. Although adaptively changing  $\varepsilon$  throughout the search did not result in a significant reduction in computation time, this feature is still recommended for marginal improvement.

Parallelization of similarity metric computation also improved efficiency, but only with 2 or 4 processors. With DIRECT, multiple processors can also be used to apply transformations, determine the overlap region, interpolate voxel intensities, and to compute the joint histogram and similarity metric for many different regions of the search space simultaneously<sup>13</sup>. This approach increases the computation-to-communication ratio, and could potentially lead to a further reduction in computation time. Powell's method, which is not intrinsically parallel, benefits from the low-level similarity metric parallelization described in this paper.



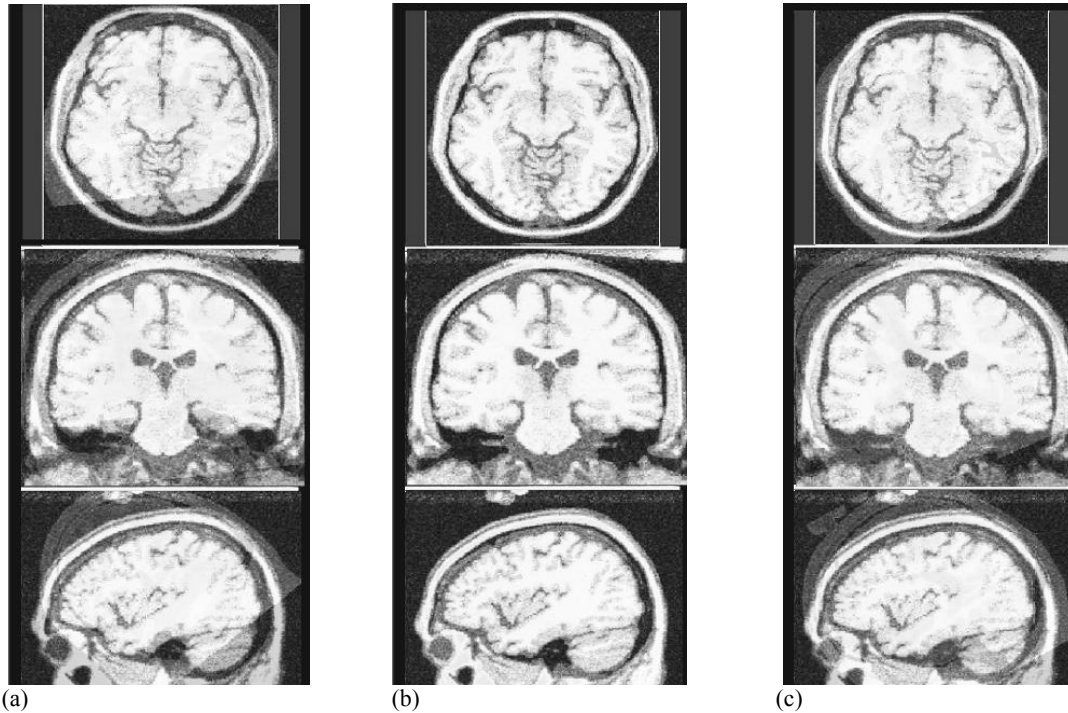
**Figure 5.** Sections of the translation search space explored by the optimization algorithms. The initial orientation, 30mm from ground truth, is circled in white. (a) DIRECT. (b) Powell's method.

## 7. CONCLUSION

The new open source/open science paradigm is increasing in popularity amongst medical imaging researchers<sup>14</sup>. It is expected that the number of ITK applications will continue to grow steadily, with a corresponding increase in algorithm sophistication. Of particular note is the ability to integrate ITK with visualization systems, such as VTK, and to build interactive computation/visualization frameworks from these open source toolkits<sup>21</sup>. In fact, many researchers feel that interactive segmentation and registration will facilitate acceptance of new algorithms for clinical use.

As a contribution to this effort, an efficient global optimization approach for medical image registration, DIRECT, has been implemented in the Insight Toolkit, a popular and widely used open source framework. The goal of this work is to provide ITK users and developers with an alternative optimization approach, and to encourage further improvements in the DIRECT algorithm, as well as the introduction of other deterministic (and stochastic) optimization techniques. This work also introduced parallel computation for mutual information within the ITK framework, which yielded at near-linear speedup for 2 processors, and additional efficiency improvements with additional processors. This study also illustrated termination criteria based on the clustering of minima, and incorporated an adaptive approach to locality parameter adjustment. These enhancements allow DIRECT to terminate once the convergence basin of the minima is reached, with local refinement with Powell's method, a robust, derivative-free local technique.

Future work includes clinical validation of this approach on more datasets, integration with other deterministic and stochastic optimization algorithms, distributed memory parallelization of other components of DIRECT, and the further development of an interactive visualization platform, employing ITK components for segmentation, registration, and analysis tasks for image guidance of minimally invasive therapeutic procedures.



**Figure 6.** Effect of convergence to local minima. (a) Initial orientation. (b) DIRECT, showing excellent registration. (c) Powell's method, clearly displaying misregistration.

## ACKNOWLEDGEMENTS

The authors thank Dr. Usaf Aladl for valuable assistance with the ITK platform, and Kevin Wang for assisting with parallelization issues. Thanks are also given to Chris Wedlake and John Moore for technical support. This work was supported by SHARCNet, NSERC R3146-A02, CIHR 14735, the Canadian Foundation for Innovation, the Ontario Innovation Trust, and the Ontario Research and Development Challenge Fund through the Ontario Consortium for Image-guided Therapy and Surgery.

## REFERENCES

1. M. Wierzbicki, M. Drangova, G. Guiraudon, and T. M. Peters, "Validation of dynamic heart models obtained using non-linear registration for virtual reality training, planning, and guidance of minimally invasive cardiac surgeries," *Med. Image Anal.* **8**, pp. 387-401, 2004.
2. G.-A. Turgeon, G. Lehmann, G. Guiraudon, M. Drangova, D. Holdsworth, and T. M. Peters, "2D-3D registration of coronary angiograms for cardiac procedure planning and guidance," *Med. Phys.* **32**, pp. 3737-3749, 2005.
3. S. Szpala, M. Wierzbicki, G. Guiraudon, and T. M. Peters, "Real-time fusion of endoscopic views with dynamic 3-D cardiac images: a phantom study," *IEEE Trans. Med. Imag.* **24**, pp. 1207-1215, 2005.
4. X. Huang, N. A. Hill, J. Ren, G. Guiraudon, D. Boughner, and T. M. Peters, "Dynamic 3D ultrasound and MR image registration of the beating heart," in *Proc. MICCAI 2005 (LNCS)* **3750**, pp. 171-178, 2005.
5. R. Smolikova-Wachowiak, M. Wachowiak, A. Fenster, and M. Drangova, "Registration of two-dimensional cardiac images to preprocedural three-dimensional images for interventional applications," *J. Magn. Resonance* **22**, pp. 219-228, 2005.

6. T. G. Kolda, R. M. Lewis, and V. Torczon, "Optimization by direct search: new perspectives on some classical and modern methods," *SIAM Review* **45**, pp. 385-482, 2003.
7. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual information-based registration of medical images: a survey," *IEEE Trans. Med. Imag.* **22**, pp. 986-1004, 2003.
8. M. P. Wachowiak and T. M. Peters, "Combining global and local parallel optimization for medical image registration," in *Proc. SPIE Medical Imaging (Image Processing)* **5747**, pp. 1189-1200, 2005.
9. G. K. Matsopoulos, N. A. Mouravliansky, K. K. Delibasis, and K. S. Nikita, "Automatic retinal image registration scheme using global optimization techniques," *IEEE Trans. Inf. Technol. B.* **3**, pp. 47-60, 1999.
10. N. P. Castellanos, P. L. D. Angel, and V. Medina, "Nonrigid medical image registration technique as a composition of local warpings," *Pattern Recogn.* **37**, pp. 2141-2154, 2004.
11. R. He and P. A. Narayana, "Global optimization of mutual information: application to three-dimensional retrospective registration of magnetic resonance images," *Comput. Med. Imag. Graph.* **26**, pp. 277-292, 2002.
12. M. P. Wachowiak, R. Smolikova, Y. Zheng, J. M. Zurada, and A. S. Elmaghraby, "An approach to multimodal biomedical image registration utilizing particle swarm optimization," *IEEE Trans. Evolut. Comput.* **8**, pp. 289-301, 2004.
13. M. P. Wachowiak and T. M. Peters, "High performance medical image registration using new optimization techniques," *IEEE Trans. Inf. Technol. B.* (To appear), 2006 (Available on-line through the IEEE Xplore Digital Library).
14. T. S. Yoo and D. N. Metaxas, "Open science – combining open data and open source software: Medical image analysis with the Insight Toolkit," *Med. Image Anal.* **9**, pp. 503-506, 2005.
15. D. Shelton, G. Stetten, S. Aylward, L. Ibáñez, A. Cois, and C. Stewart, "Teaching medical image analysis with the Insight Toolkit," *Med. Image Anal.* **9**, pp. 605-611, 2005.
16. L. Ibáñez, W. Schroeder, L. Ng, and J. Cates, *The ITK Software Guide (ITK 1.4)*, Kitware, Inc., 2003.
17. D. R. Jones, C. D. Perttunen, and B. E. Stuckman, "Lipschitzian optimization without the Lipschitz constant," *J. Optimiz. Theory App.* **79**, pp. 157-181, 1993.
18. D. E. Finkel and C. T. Kelley, "An Adaptive Restart Implementation of DIRECT," North Carolina State University, Center for Research in Scientific Computation, CRSC-TR04-30, 2004.
19. D. E. Finkel and C. T. Kelley, "Convergence Analysis of the DIRECT Algorithm," North Carolina State University, Center for Research in Scientific Computation, CRSC-TR04-28, 2004.
20. R. K. S. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image processing and classification methods," *IEEE Trans. Med. Imag.* **18**, pp. 1085-1097, 1999.
21. I. Wolf, M. Vetter, I. Wegner, Ingmar, T. Böttger, M. Nolden, M. Schöbinger, M. Hastenteufel, T. Kunert, and H.-P. Meinzer, "The Medical Imaging Interaction Toolkit," *Med. Image Anal.* **9**, pp. 594-604, 2005.