Fast Deformable Image Registration with Non-Smooth Dual Optimization

Martin Rajchl

Dept. of Computing, Imperial College London, London, UK

`m.rajchl@imperial.ac.uk`

John S.H. Baxter, Wu Qiu, Ali R. Khan, Aaron Fenster, Terry M. Peters

Robarts Research Institute, Western University, London, ON

Daniel Rueckert

Dept. of Computing, Imperial College London, London, UK

Jing Yuan

Robarts Research Institute, Western University, London, ON

Abstract

Optimization techniques have been widely used in deformable registration, allowing for the incorporation of similarity metrics with regularization mechanisms. These regularization mechanisms are designed to mitigate the effects of trivial solutions to ill-posed registration problems and to otherwise ensure the resulting deformation fields are well-behaved. This paper introduces a novel deformable registration (DR) algorithm, RANCOR*, which uses iterative convexification to address DR problems under non-smooth total-variation regularization. Initial comparative results against four state-of-the-art registration algorithms and under smooth regularization, respectively, are presented using the Internet Brain Segmentation Repository (IBSR) database.*

1. Introduction

Registration is the systematic spatial deformation of one medical image as to align it with another, either in the same or a different modality. Although registration of images common to a single patient can rely largely on rigid transformations, registration between patient images, common in techniques such as atlas construction or atlas-based segmentation, have relied on highly non-linear deformations often in the absence of highly detectable and localizable landmarks. Deformable registration (DR) aims to address these problems, using a similarity metric to judge the quality of the alignment after deformation, and a regularization mechanism to ensure that the deformation field avoid trivial and otherwise undesirable components such as gaps or singularities. Deformable registration is indeed a challenging problem with many competing facets. Our algorithm is intended to provide an additional option that could facilitate atlas building and segmentation techniques.

Many components of our DR algorithm display a large amount of inherent parallelism between image voxels. Such algorithms have been of growing interest to the medical imaging community because of the ability to implement them on commercially available general purpose graphics processing units (GPGPUs) to dramatically improve their speed and computational efficiency for both registration[13] and segmentation problems [17, 16, 18].

1.1. Contributions

We propose a novel DR algorithm, *RANCOR* (RegistrAtioN via COnvex Relaxation), that allows for the combination of any pointwise error metric (*e.g.* as the sum of absolute intensity differences (SAD) for intra-modality registration, and mutual information (MI) [24, 19] or modality independent neighbourhood descriptors (MIND) [9] for inter-modality registration) while regularizing the deformation field by its total variation. Employing total variation regularization (TVR) in registration problems has been successfully investigated in former studies [10, 15, 27, 23]. The proposed method builds upon recent work [22], from which we generalise the formulation to incorporate any pointwise similarity metrics and replace and compare the employed quadratic regularizer (QR). The newly proposed regularizer employs an $L1$ penalty, thus allowing for non-smooth regularization of the deformation field, leading to accuracy improvements in registration of structures with high variability. To ensure high performance, the proposed algorithm is implemented using GPGPU.

1.2. Previous studies

Recent surveys provide a good overview of existing DR methods [6, 8, 21] and we would like to emphasize the study performed by Klein et al. [12], where 14 DR algorithms were compared across four open brain image databases. We

will compare our proposed method against the four highest ranked methods identified in [12]:

Advanced Normalization Tools (ANTs): The *Symmetric Normalization (SyN)* DR method in [3] uses a multi-resolution scheme to enforce a bi-directional diffeomorphism while maximizing a cross-correlation metric. It has been shown in several open challenges [7, 12, 14] to outperform well established methods. *SyN* regularizes the deformation field through Gaussian smoothing and enforcing transformation symmetry.

Image Registration Toolkit (IRTK): The well-known *Fast Free-Form deformations (F3D)* method in [20] defines a lattice of equally spaced control points over the target image and, by moving each point, locally modifies the deformation field. Normalized mutual information combined with a cubic b-spline bending energy is used as the objective function. Its multi-resolution implementation employs coarsely-to-finely spaced lattices and Gaussian smoothing.

Automatic Registration Toolbox (ART): [1] presents a homeomorphic DR method using normalized cross-correlation as similarity metric in a multi-resolution framework. The deformation field is regularized via median and low-pass Gaussian filtering.

Statistical Parametric Mapping DARTEL Toolbox (SPM_D): The *DARTEL* algorithm presented in [2] employs a static finite difference model of a velocity field. The flow field is considered as a member of the Lie algebra, which is exponentiated to produce a deformation inherently enforcing a diffeomorphism. It is implemented in a recursive, multi-resolution manner.

2. Methods

In this section, we propose a multi-scale dual optimization based method to estimate the non-linear deformation field $u(x) = [u_1(x), u_2(x), u_3(x)]^T$, between two given images $I_1(x)$ and $I_2(x)$, which explores the minimization of the variational optical-flow energy function. The approach is based on recent advances in DR using a Gauss-Newton (GN) optimizer with smooth regularization [9, 22]:

$$\min_u \quad P(I_1, I_2; u) + R(u) \qquad (1)$$

where the function term $P(I_1, I_2; u)$ stands for a dissimilarity measure of the two input images $I_1(x)$ and $I_2(x)$ under deformation by u, and $R(u)$ gives the regularization function to single out a smooth deformation field. In this paper, we use the sum of absolute intensity differences (SAD):

$$P(I_1, I_2; u) := \int_\Omega |I_1(x + u) - I_2(x)| \, dx, \qquad (2)$$

as a simple similarity metric for two input images from the same modality.

The proposed framework can also be directly adapted for more advanced image dissimilarity measures designed for registration between different modalities.

A regularization term, $R(u)$, is often incorporated to make the minimization problem (1) well-posed. Otherwise, minimizing the image dissimilarity function $P(I_1, I_2; u)$ can result in trivial or infinite solutions. We consider the total variation of the deformation field as the regularization term:

$$R(u) := \alpha \int_\Omega \left(|\nabla u_1| + |\nabla u_2| + |\nabla u_3| \right) dx \, . \qquad (3)$$

The expected non-convexity of $I_1(x)$ and $I_2(x)$, makes it challenging to directly minimize (1), even with convex regularization. To address this issue, we introduce an incremental convexification approach, which lends itself to a standard coarse-to-fine framework and allows for a more global perspective and avoiding local optima by capturing large deformations.

In Section 2.1, we develop the multi-scale optimization framework, developing a sequence of related minimization problems. Each of these problems are solved through a new non-smooth Gauss-Newton (GN) approach introduced in Section 2.2. which employs a novel sequential convexification and dual optimization procedure.

2.1. Coarse-to-Fine Optimization Framework

The first stage in our approach is the construction of the image pyramid. Let $I_1^1(x) \dots I_1^L(x)$ be the L-level pyramid representation of $I_1(x)$ from the coarsest resolution $I_1^1(x)$ to the finest resolution $I_1^L(x) = I_1(x)$, and $I_2^1(x) \dots I_2^L(x)$ the L-level coarse-to-fine pyramid representation of $I_2(x)$. The optimization process is started from the coarsest level, $\ell = 1$, which extracts the deformation field $u^1(x)$ between $I_1^1(x)$ and $I_2^1(x)$ such that:

$$\min_{u^1} \quad P(I_1^1(x), I_2^1(x); u^1) + R(u^1) \, . \qquad (4)$$

The vector field $u^1(x)$ denotes the optimal deformation field at the coarsest scale. It is upsampled to the next finer-resolved level, $\ell = 2$, to compute the optimal finer-level deformation field $u^2(x)$. The process is repeated, obtaining the deformation field $u^3(x) \dots u^L(x)$ at each level sequentially.

Second, at each resolution level ℓ, $\ell = 2 \dots L$, we compute an incremental deformation field $t^\ell(x)$ based on the two image functions $I_2^\ell(x)$ and $I_1^\ell(x + u^{\ell-1})$, where $I_1^\ell(x + u^{\ell-1})$ is warped by the deformation field $u^{\ell-1}(x)$ computed at the previous resolution level $\ell - 1$, *i.e.*

$$\min_{t^\ell} \quad P(I_1^\ell(x + u^{\ell-1}), I_2^\ell(x); t^\ell) + R(u^{\ell-1} + t^\ell) \, . \quad (5)$$

The optimization problem (4) can be viewed as a special case of (5), *i.e.* for $\ell = 1$, we define $u^0(x) = 0$ and

$u^1(x) = (u^0 + t^1)(x)$. Therefore, the proposed coarse-to-fine optimization framework sequentially explores the minimization of (5) at each image resolution level, from the coarsest $\ell = 1$ to the finest $\ell = L$.

2.2. Sequential Convexification and Dual Optimization

Now we consider the optimization problem (5) for each image resolution level ℓ. Given the highly non-linear function $P(I_1^\ell(x + u^{\ell-1}), I_2^\ell(x); t^\ell)$ in (5), we introduce a sequential linearization and convexification procedure for this challenging non-linear optimization problem (5). This results in a series of incremental warping steps in which each step approximates an update of the deformation field $t^\ell(x) = (t_1^\ell(x), t_2^\ell(x), t_3^\ell(x))^{\mathsf{T}}$, until the updated deformation is sufficiently small, *i.e.*, it iterates through the following sequence of convex minimization steps until convergence is attained:

- Initialize $(h^\ell)^0(x) = 0$ and let $k = 1$;

- At the kth iteration, define the deformation field as

$$\tilde{u}^{\ell-1}(x) := \left(u^{\ell-1} + \sum_{i=0}^{k-1}(h^\ell)^i\right)(x)$$

and compute the update deformation $(h^\ell)^k$ to $\tilde{u}^{\ell-1}(x)$ by minimizing the following convex energy function:

$$\min_{(h^\ell)^k} \int_\Omega \left|\tilde{P}_0^k + \nabla\tilde{P}^k \cdot (h^\ell)^k\right| dx + R(u^{\ell-1} + (h^\ell)^k),$$
(6)

where

$$\tilde{P}^k((h^\ell)^k) = P(I_1^\ell(x + \tilde{u}^{\ell-1}), I_2^\ell(x); (h^\ell)^k)$$

and $\tilde{P}_0^k(x) = P(I_1^\ell(x + \tilde{u}^{\ell-1}), I_2^\ell(x); 0)$.

- Let $k = k + 1$ and repeat the second step till the new update $(h^\ell)^k$ is small enough. Then, we have the total incremental deformation field $t^{\ell(x)}$ at the image resolution level ℓ as:

$$t^\ell(x) = \sum_{i=0}^k (h^\ell)^i(x).$$

These steps can be viewed as a non-smooth GN method for the non-linear optimization problem (5), in contrast to the classical GN method proposed in [4, 22]. Moreover, the L_1-norm and the convex regularization term $\mathcal{R}(\cdot)$, (6) results in a convex optimization problem. The non-smooth L_1-norm from (6) provides more robustness in practice than the conventional smooth L_2-norm used in the classical GN method [22].

Solving the convex minimization problem (6) is the most essential step in the proposed algorithmic framework. The introduced primal-dual variational analysis not only provides an equivalent dual formulation to the optimization problem (6) but also derives an efficient solution algorithm. First, we simplify the expression of the convex problem (6) as:

$$\min_h \int_\Omega |P_0 + \nabla P \cdot h|\, dx + R(\tilde{u} + h),$$
(7)

where $\tilde{u}(x)$ represents the deformation field.

Through variational analysis, we can derive an equivalent *dual model* to (7):

Proposition 2.1 *The* convex minimization problem (7) *can be represented by its* primal-dual model (15) *and* dual model*:*

$$\max_{|w(x)|\leq 1, q} E(w, q) := \int \left(wP_0 + \sum_{i=1}^3 \tilde{u}_i \operatorname{div} q_i\right) dx - R^*(q)$$
(8)

subject to

$$F_i(x) := (w \cdot \partial_i P + \operatorname{div} q_i)(x) = 0, \quad i = 1, 2, 3. \quad (9)$$

The dual regularization function $R^*(q)$ is given by (13).

Given the conjugate representation of the absolute function:

$$|v| = \max_w w \cdot v, \quad \text{s.t. } |w| \leq 1,$$
(10)

we can rewrite the first L_1-norm term of (7) as follows:

$$\int_\Omega |P_0 + \nabla P \cdot h|\, dx = \max_{|w(x)|\leq 1} \int_\Omega w(P_0 + \nabla P \cdot h) dx.$$
(11)

Additionally, given $R(\tilde{u} + h)$ in terms of (3), we have

$$\alpha \sum_{i=1}^{3} \int_{\Omega} |\nabla(u_i + h_i)| dx$$

$$= \max_{|q|^2 = \alpha^2} \sum_{i=1}^{3} - \int_{\Omega} q \cdot \nabla(u_i + h_i) dx$$

$$= \max_{|q| \leq \alpha} \sum_{i=1}^{3} - \int_{\Omega} q \cdot \nabla(u_i + h_i) dx$$

$$= \max_{|q| \leq \alpha} \sum_{i=1}^{3} \int_{\delta\Omega} (u_i + h_i) q_i \cdot ds - \int_{\Omega} q \cdot \nabla(u_i + h_i) dx$$

$$= \max_{|q| \leq \alpha} \sum_{i=1}^{3} \int_{\Omega} \operatorname{div}(u_i + h_i) q_i) dx - \int_{\Omega} q \cdot \nabla(u_i + h_i) dx$$

$$= \max_{|q| \leq \alpha} \sum_{i=1}^{3} \int_{\Omega} (u_i + h_i) \operatorname{div} q_i dx$$

$$= \max_{q} \sum_{i=1}^{3} \int_{\Omega} (u_i + h_i) \operatorname{div} q_i dx - R^*(q)$$

$$(12)$$

where each dual variable $q_i(x)$, $i = 1, 2, 3$, has characteristic function of the constraint $|q_i(x)| \leq \alpha$, $i = 1, 2, 3$:

$$R^*(q) = \chi_{|q_{1,2,3}(x)| \leq \alpha}(q). \qquad (13)$$

Considering (11) and (12), one can see that the convex minimization problem (6) is equivalent to the minimax problem:

$$\min_{h} \max_{|w(x)| \leq 1, q} \int w(P_0 + \nabla P \cdot h) dx$$
$$+ \sum_{i=1}^{3} \int \operatorname{div} q_i (\tilde{u}_i + h_i) dx - R^*(q) \qquad (14)$$

that is

$$\min_{h} \max_{|w(x)| \leq 1, q} \int (w P_0 + \sum_{i=1}^{3} \tilde{u}_i \operatorname{div} q_i) dx$$
$$+ \sum_{i=1}^{3} \int h_i (w \cdot \partial_i P + \operatorname{div} q_i) dx - R^*(q) \qquad (15)$$

which is called the *primal-dual formulation* in this paper.

After variation by the free variable $h_i(x)$, $i = 1, 2, 3$, the minimization of the *primal-dual formulation* (15) over $h_i(x)$, $i = 1, 2, 3$, results in the linear equalities' constraints

$$(w \cdot \partial_i P + \operatorname{div} q_i)(x) = 0, \quad i = 1, 2, 3, \qquad (16)$$

and the maximization problem

$$\max_{|w(x)| \leq 1, q} E(w, q) := \int (w P_0 + \sum_{i=1}^{3} \tilde{u}_i \operatorname{div} q_i) dx - R^*(q)$$

thereby proving Prop. 2.1.

Each component of the deformation field $[h_1(x), h_2(x), h_3(x)]^{\mathsf{T}}$ works as the optimal multiplier functions to their respective constraints, (9). Therefore, the energy function of the primal-dual model (15) is exactly the Lagrangian function to the *dual model* (8):

$$L(h, w, q) = E(w, q) + \sum_{i=1}^{3} \langle h_i, F_i \rangle,$$

where $E(w, q)$ and the linear functions $F_i(x)$, $i = 1, 2, 3$, are defined in (8) and (9) respectively. We can now derive an efficient duality-based Lagrangian augmented algorithm based on modern convex optimization theory (see [5, 25, 26] for details), using the augmented Lagrangian function:

$$L_c(h, w, q) = L(h, w, q) - \frac{c}{2} \sum_{i=1}^{3} \|F_i\|^2, \qquad (17)$$

where $c > 0$ is a positive constant and the additional quadratic penalty function is applied to ensure the functions (9) vanish. Our proposed *duality-based optimization algorithm* is:

- Set the initial values of w^0, q^0 and h^0, and let $k = 0$.

- Fix q^k and h^k, optimize w^{k+1} by

$$w^{k+1} := \arg \max_{|w(x)| \leq 1} L_c(h^k, w, q^k) \qquad (18)$$

generating the convex minimization problem:

$$\min_{|w(x)| \leq 1} \int w P_0 dx + \frac{c}{2} \sum_{i=1}^{3} \int (w \partial_i P - T_i^k)^2 dx; \qquad (19)$$

where $T_i^k(x)$ ($i = 1, 2, 3$) is computed from the fixed variables q^k and h^k. w^{k+1} is computed by thresholding:

$$w^{k+1} = \operatorname{Threshold}_{|w(x)| \leq 1}(w^{k+1/2}(x)), \qquad (20)$$

where

$$w^{k+1/2} = \frac{c \sum_{i=1}^{3}(\partial_i P \cdot T_i^k) - P_0}{c \sum_{i=1}^{3}(\partial_i P)^2}.$$

- Fixing w^{k+1} and h^k, optimize q^{k+1} by

$$q^{k+1} := \arg \min_{q} L_c(h^k, w^{k+1}, q); \qquad (21)$$

which amounts to three convex minimization problems:

$$\min_{q_i} \int q_i \cdot \nabla \tilde{u}_i dx + \frac{c}{2} \int (\operatorname{div} q_i - U_i^k)^2 dx + R^*(q);$$
$$i = 1, 2, 3; \qquad (22)$$

where U_i^k is computed from the fixed variables w^{k+1} and h^k. Hence, q_i^{k+1}, $i = 1, 2, 3$, can be approximated by a gradient-projection step corresponding to (13).

- Once w^{k+1} and q^{k+1} are obtained, update h^{k+1} by

$$h_i^{k+1} = h^k - c\left(w^{k+1} \cdot \partial_i P + \operatorname{div} q_i^{k+1}\right);$$
$$i = 1, 2, 3; \qquad (23)$$

- Increment k and iterate until converged, *i.e.*

$$c \int \left| w^{k+1} \cdot \partial_i P + \operatorname{div} q_i^{k+1} \right| dx \leq \delta, \qquad (24)$$

where δ is a chosen small positive parameter (5×10^{-4}).

3. Experiments

3.1. Image Database

The image data consisted of an open multi-center T1w MRI dataset with corresponding manual segmentations, the Internet Brain Segmentation Repository (IBSR) database, totalling 18 labelled image volumes at 1.5T available on *Mindboggle*[1] in a pre-processed form with labelling protocols and transforms into MNI space.

The experiments were performed in a pair-wise manner. For each image in the database, seventeen registrations were performed using the chosen image as the reference image and one of the remaining as the floating image. Thus, our experiment consisted of 306 registration problems in total.

3.2. Compared Algorithms

We compare the four best performing DR methods according to [12] (*i.e.* IRTK, ART, SPM_D and ANTS) and the Gauss-Newton (GN) optimization based method in [22] using quadratic regularization (QR). For all evaluations, we retrieved the numerical registration results reported in [12] from *Mindboggle*1.

3.3. Initialization & Pre-processing

Prior to registration, all images were skull stripped by constructing brain masks from manual labels using morphological operations [12] and then affinely registered using the *FMRIB Software Library's (FSL) FLIRT* package [11] into the space of the *MNI152_T1_1mm_brain*. These affine transformations were made available on *Mindboggle*1 and used to initialize the DR algorithms. This guarantees that the same initialization is used for the algorithms in [12] and allows for quantitative comparisons. As a pre-processing step, both affinely registered images were robustly normalized to zero mean and standard deviation units to ensure a constant regularization weight α could be used.

---

[1] http://www.mindboggle.info

3.4. Implementation & Parameter Tuning

The proposed DR method was implemented in MATLAB (Natick, MA) using the Compute Unified Device Architecture (CUDA) (NVIDIA, Santa Clara, CA) for GPGPU computing. Each level in the coarse-to-fine framework consists of multiple warps invoking the proposed GPGPU accelerated regularization algorithm. Parameter tuning of the regularization weight α was done on two randomly picked dataset pairs similar to the tuning of the comparative methods in [12], where a small number of images was provided to the authors of each compared method. All other parameters, such as the number of levels (L), the number of warps (N_t) and the maximum number of iterations (k_{max}) were determined heuristically on a single image volume not used in this study. Table 1 contains all set parameter values.

Table 1. Registration algorithm parameters

Method	α	L	N_t	k_{max}
GN QR [22]	0.05	3	4	220
GN TVR	0.30	3	4	220
All parameters were kept constant across all experiments.				

3.5. Evaluation Metric

To compare our registration method against other DR algorithms, we used the *target overlap* (*TO*) as a regional metric:

$$TO = \frac{\sum_L |F_L \cup R_L|}{\sum_L |R_L|} \qquad (25)$$

where F is the floating image, R the reference image, and L a labelled region, as indicated in [12]. This parallels our motivation of using DR to port segmentation labels to incoming datasets, and takes advantage of the manual segmentations providing in the IBSR database.

Results were considered significant if the probability of making a type I error was less than 1% ($p < 0.01$). For this purpose, we employed a series of two-tailed, pairwise Student's T-test.

4. Results

4.1. Accuracy

Figure 2 shows box plots of the TO accuracy for each of the registration methods. The results were averaged across all regions, according to [12]. The corresponding numerical results are provided in Table 3. Example registrations for GN TV and GN QR are depicted in Fig. 1.

4.2. Run times

The experiments were conducted on a Ubuntu 12.04 (64-bit) desktop machine with 144 GB memory and an NVIDIA Tesla C2060 (6 GB memory) graphics card. The
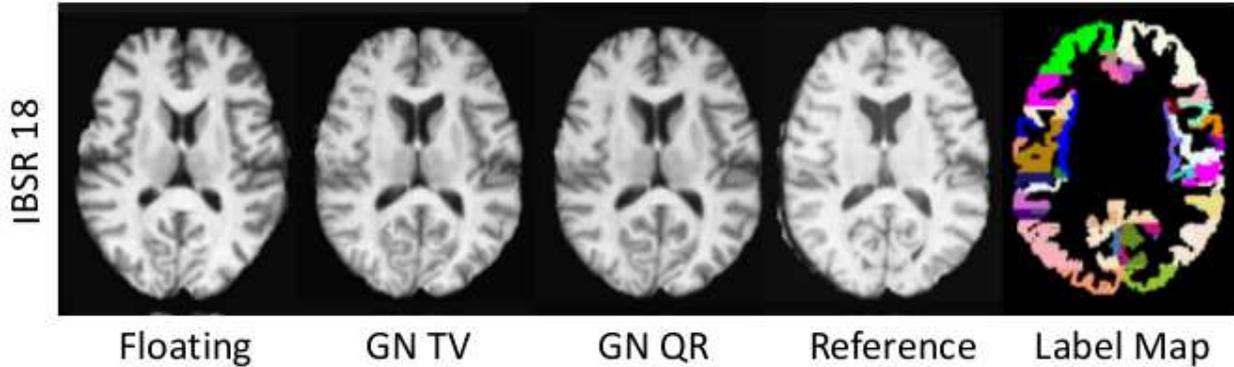
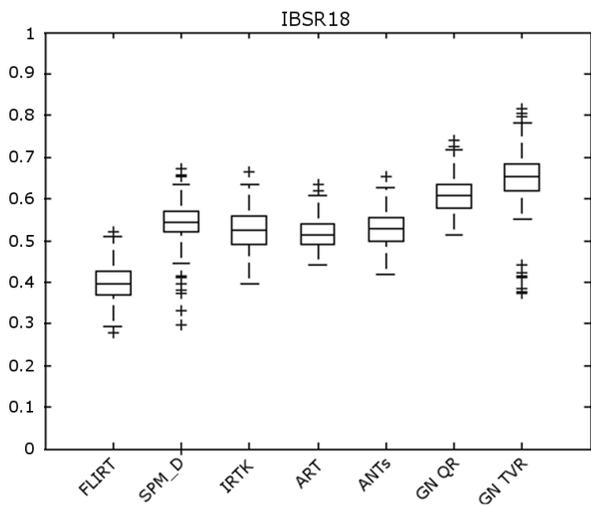Figure 1. Example registration results using the proposed non-smooth GN TV and the smooth GN QR in [22].

Figure 2. Mean Target Overlap Results

Table 2. Mean target overlap (TO) accuracy averaged over all labelled regions. All results were statistically significant (p <0.01).

IBSR18	
Method	mean $\pm$ std.
FLIRT	39.7 ± 13.0
SPM_D	54.0 ± 14.7
IRTK	52.1 ± 15.0
ART	51.5 ± 14.1
Syn	52.8 ± 14.9
GN QR	60.8 ± 11.8
GN TVR	64.9 ± 12.5

maximum run times for the MATLAB code including pre-processing, optimization, and GPGPU enhanced regularization are given in Table 3.

5. Discussion

We propose a novel GPGPU-accelerated DR method, based on GN optimization and non-smooth total variation

Table 3. Maximum GPGPU regularization run times at each level l (*reg.*), maximum required and total registration time in seconds.

	reg. [s]			GN opt. [s]	Total [s]
l	1	2	3		
GN QR	0.26	1.82	13.59	9.08	71.40
GN TVR	0.37	2.70	20.66	10.62	104.70

regularization. The method is implemented within a coarse-to-fine optimization framework and compared on an open and publicly available database, IBSR. We employed the same initialization, tuning conditions, and evaluation scripts to quantitatively compare the proposed methods against four well-known DR methods in [12]. Additionally, we extended our evaluation to include the algorithm under quadratic regularization as used in Sun *et al.* [22] and obtain favourable registration results. We numerically report the TO accuracy on the IBSR database allowing for direct comparison.

The proposed method significantly outperforms the comparative methods in terms of TO ($p < 0.01$). We note, that both the proposed methods employed the simplest and most non-robust similarity metric, SAD, while *SPM_D*, *IRTK*, *ART* and *SyN* use advanced metrics (*c.f.* [12]). The choice of similarity metric was intentionally chosen for these experiments to demonstrate the potential of the proposed method without an advanced similarity metrics or optimizer (*e.g.* a Levenberg-Marquardt optimizer as used in *SPM_D* [2]).

The current *RANCOR* framework can be seen as a basic method to be extended over time, under the same open science credo, that allowed us to readily and quantitatively compare well-known open methods using public databases. As the current framework cannot currently guarantee diffeomorphic deformations, the next step is to enforce such constraints on the resulting deformation fields. Furthermore, to enable inter-modality DR, we aim to implement and test commonly used advanced similarity metrics, such as nor-

malized mutual-information, normalized cross-correlation, or more recently developed methods, such as the $L2$ norm of the MIND descriptor [9]. Since command-line tools, such as the compared open DR methods are needed for data analysis, *RANCOR* will be definitely included into such a package and, as a matter of course, be made available to the community.

6. Conclusions

We proposed a novel GPGPU-accelerated registration algorithm that optimizes any pointwise similarity metric and non-smooth total variation regularization within a Gauss-Newton optimization framework. This algorithm was then evaluated against the four highest ranking non-linear registration algorithms according to [12] and against a similar method employing smooth regularization [22] on an open image database. We report favourable accuracy results compared to all tested methods, suggesting further investigation of this promising DR approach.

References

[1] B. A. Ardekani, S. Guckemus, A. Bachman, M. J. Hoptman, M. Wojtaszek, and J. Nierenberg. Quantitative comparison of algorithms for inter-subject registration of 3d volumetric brain mri scans. *Journal of neuroscience methods*, 142(1):67–76, 2005.

[2] J. Ashburner. A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1):95–113, 2007.

[3] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis*, 12(1):26–41, 2008.

[4] M. Baust, D. Zikic, N. Navab, and G. Munich. Diffusion-based regularisation strategies for variational level set segmentation. In *BMVC*, pages 1–11, 2010.

[5] D. P. Bertsekas. Nonlinear programming. 1999.

[6] W. Crum, T. Hartkens, and D. Hill. Non-rigid image registration: theory and practice. 2004.

[7] K. Farahani, B. Menze, M. Reyes, E. Gerstner, J. Kirby, and J. Kalpathy-Cramer. Multimodal brain tumor segmentation (brats 2013), Sept. 2013.

[8] A. Gholipour, N. Kehtarnavaz, R. Briggs, M. Devous, and K. Gopinath. Brain functional localization: a survey of image registration techniques. *Medical Imaging, IEEE Transactions on*, 26(4):427–451, 2007.

[9] M. P. Heinrich, M. Jenkinson, M. Bhushan, T. Matin, F. V. Gleeson, S. M. Brady, and J. A. Schnabel. Mind: Modality independent neighbourhood descriptor for multi-modal deformable registration. *Medical Image Analysis*, 16(7):1423–1435, 2012.

[10] L. Hömke, C. Frohn-Schauf, S. Henn, and K. Witsch. Total variation based image registration. In *Image Processing Based on Partial Differential Equations*, pages 343–361. Springer, 2007.

[11] M. Jenkinson, P. Bannister, M. Brady, and S. Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841, 2002.

[12] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain mri registration. *Neuroimage*, 46(3):786–802, 2009.

[13] M. Modat, G. R. Ridgway, Z. A. Taylor, M. Lehmann, J. Barnes, D. J. Hawkes, N. C. Fox, and S. Ourselin. Fast free-form deformation using graphics processing units. *Computer methods and programs in biomedicine*, 98(3):278–284, 2010.

[14] K. Murphy, B. Van Ginneken, J. M. Reinhardt, S. Kabus, K. Ding, X. Deng, K. Cao, K. Du, G. E. Christensen, V. Garcia, et al. Evaluation of registration methods on thoracic ct: the empire10 challenge. *Medical Imaging, IEEE Transactions on*, 30(11):1901–1920, 2011.

[15] T. Pock, M. Urschler, C. Zach, R. Beichel, and H. Bischof. A duality based algorithm for tv-l 1-optical-flow image registration. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 511–518. Springer, 2007.

[16] W. Qiu, J. Yuan, E. Ukwatta, Y. Sun, M. Rajchl, and A. Fenster. Fast globally optimal segmentation of 3d prostate mri with axial symmetry prior. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 198–205. Springer Berlin Heidelberg, 2013.

[17] M. Rajchl, J. Yuan, E. Ukwatta, and T. Peters. Fast interactive multi-region cardiac segmentation with linearly ordered labels. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages Page–s. IEEE Conference Publications, 2012.

[18] M. Rajchl, J. Yuan, J. White, E. Ukwatta, J. Stirrat, C. Nambakhsh, F. Li, and T. Peters. Interactive hierarchical max-flow segmentation of scar tissue from late-enhancement cardiac mr images. *IEEE Transactions on Medical Imaging*, 33(1):159–172, 2014.

[19] P. Rogelj, S. Kovačič, and J. C. Gee. Point similarity measures for non-rigid registration of multi-modal data. *Computer vision and image understanding*, 92(1):112–140, 2003.

[20] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *Medical Imaging, IEEE Transactions on*, 18(8):712–721, 1999.

[21] A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration: A survey. *Medical Imaging, IEEE Transactions on*, 32(7):1153–1190, 2013.

[22] Y. Sun, J. Yuan, W. Qiu, M. Rajchl, C. Romagnoli, and A. Fenster. Three-dimensional non-rigid mr-trus registration using dual optimization. *IEEE Transactions on Medical Imaging*, 34(5):1085–1095, 2014.

[23] V. Vishnevskiy, T. Gass, G. Székely, and O. Goksel. Total variation regularization of displacements in parametric image registration. In *Abdominal Imaging. Computational and Clinical Applications*, pages 211–220. Springer, 2014.

[24] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical image analysis*, 1(1):35–51, 1996.

[25] J. Yuan, E. Bae, and X.-C. Tai. A study on continuous max-flow and min-cut approaches. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2217–2224. IEEE, 2010.

[26] J. Yuan, E. Bae, X.-C. Tai, and Y. Boykov. A continuous max-flow approach to potts model. In *Computer Vision–ECCV 2010*, pages 379–392. Springer, 2010.

[27] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime tv-l 1 optical flow. In *Pattern Recognition*, pages 214–223. Springer, 2007.