33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Ranking of Stopping Criteria for Log Domain Diffeomorphic Demons Application in Clinical Radiation Therapy

M. Peroni, P. Golland, G.C. Sharp, *Member, IEEE,* and G. Baroni

Abstract— **Deformable Image Registration is a complex optimization algorithm with the goal of modeling a non-rigid transformation between two images. A crucial issue in this field is guaranteeing the user a robust but computationally reasonable algorithm. We rank the performances of four stopping criteria and six stopping value computation strategies for a log domain deformable registration. The stopping criteria we test are: (a) velocity field update magnitude, (b) vector field Jacobian, (c) mean squared error, and (d) harmonic energy. Experiments demonstrate that comparing the metric value over the last three iterations with the metric minimum of between four and six previous iterations is a robust and appropriate strategy. The harmonic energy and vector field update magnitude metrics give the best results in terms of robustness and speed of convergence.**

I. Introduction

In recent years, image guidance has gained popularity in radiation therapy for planning and treatment monitoring. From modeling and compensation of setup uncertainties, the focus of research has gradually moved to deformations induced both from anatomic and physiological motion (e.g. bladder deformation, respiratory motion) and from therapy response (e.g. tumor growth or regression) for the implementation of adaptive radiotherapy. In addition, inter-subject probabilistic segmentation of anatomical volumes has become more important, with the end goal of supporting physician contouring. Each of these applications depends on Deformable Image Registration (DIR), which is an optimization problem with the goal to recover non-rigid deformations between a fixed or reference image I_f and a moving or target image I_m. Several algorithms have been developed which differ in physical or mathematical formulation, disparity metric and optimization strategy. Amongst DIR algorithms, one of the most popular algorithm is the *demons* algorithm, introduced by Thirion [1]. Demons is an automatic, image based algorithm, that optimizes displacement forces that pull voxels in I_m to match voxels of similar intensity in I_f , according to local characteristics of the images in a similar way Maxwell used for solving the Gibb's paradox. To provide invertibility and physical meaning to the computed deformation field, demons and other DIR algorithms include a smoothing step and/or inverse consistency procedure. One of the main limitation to the clinical use of DIR is the need for reasonable computational time and a robust parameter set. Several efforts have been made towards the first objective [2,3], thanks in part to the introduction of powerful and affordable GPUs [4]. Besides the most advanced technologies and the most complex algorithm, a good stopping condition that prevents extra optimizer iterations when convergence is reached would decrease the computational effort. The choice of what should be employed as escape condition is, however, not obvious given the complexity of the problem.

In this work we compare four different stopping rules for DIR, each focusing on a different aspect of the registration algorithm, as explained in section IIB. In [5], the similarity metric is used as the escape condition, without taking into account how the similarity metric is related to the vector field. In [6], a metric based on updates to the deformation field is used, which we now extend to the velocity field. Also in [7], the authors used the harmonic energy and the number of vector field voxels with negative Jacobian elements to assess algorithm performance. Here we propose to use these as an escape condition. Experimental evaluation of these stopping criteria were performed with three synthetic non-rigid deformations applied to clinical quality CT image volumes of an anthropomorphic phantom.

II. Materials and Methods

A. Log-Domain Diffeomorphic Demons

Since Thirion's paper [1], the so-called *demons* registration has gained popularity for intra-modality image registration and several variations have been implemented. An interesting and mathematically robust variation of the classical demons schema has been proposed by Vercauteren et al in 2009 [7]. The Log-Domain Diffeomorphic Demons combines the advantage of optimizing a diffeomorphic transformation with a computationally efficient framework, in which the optimization of the cost function happens in two different steps. The algorithm minimizes the cost function $E(\Phi, s; I_f, I_m) = \frac{1}{\sigma_i^2} simil\,(I_f, I_m \circ \Phi) + \frac{1}{\sigma_c^2} dist(\Phi, s) + \frac{1}{\sigma_T^2} reg(s)$, where $simil\,(\mathrm{I_f}, \mathrm{I_m} \circ \Phi) = \left\|\mathrm{I_f}, \mathrm{I_m} \circ \Phi\right\|^2$ is the image disparity measure, $dist\,(\Phi, \mathrm{s}) = \left\|\Phi - \mathrm{s}\right\|^2$ quantifies the similarity between the

Manuscript received on April 17th , 2011. The authors acknowledge the Rocca Foundation and NIH NIBIB NAMIC U54-EB005149 for the support. The first author would like to thank C. Gianoli for the stimulating discussions on this work.

M. Peroni* and G. Baroni are with the department of Bioengineering at Politecnico di Milano, 20133 Milano, Italy (corresponding author e-mail: marta.peroni@ mail.polimi.it). P. Golland is with the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA. G. C. Sharp is with the department of Radiation Oncology and Harvard Medical School, Boston 02114, MA, USA.

deformations Φ and s, $reg(s)$ is the degree of smoothness of the deformation and σ_i^2, σ_c^2, σ_T^2 balance the contribution of the three terms of the cost function. This energy is optimized in a symmetric log domain by an efficient second-order minimization in which the warp Φ is represented with a smooth and stationary velocity field v. The relationship between the two is defined as $\Phi(x) = \exp(v)(x)$ which is a diffeomorphic deformation, with inverse $\Phi^{-1}(x) = \exp(-v)(x)$. At each iteration, an update velocity field u is computed minimizing $E_{diff}(\Phi, s; I_f, I_m) = simil(I_f, I_m \circ \Phi \circ \exp(u)) + \|u\|^2$.

Therefore, the algorithm proceeds as follows:

1. Compute $u_{forward}$ and $u_{backward}$ that minimize E_{diff}.
2. Smooth with a Gaussian kernel K so that $u = \frac{1}{2} K * (u_{forward} - u_{backward})$
3. Update $v = v + u + \frac{1}{2}[u, v]$. Here, the Lie is computed as $[u, v] = |J^v(x)| \cdot u(x) - |J^v(x)| \cdot v(x)$, where $|J^v(x)|$ is the determinant of the Jacobian of the velocity field.

B. Stopping Criteria

The results of demons registration depends greatly on the chosen stopping criteria. The most basic stopping condition is to terminate the algorithm after a predetermined number of iterations. Although very simple, this criteria is only weakly related to the actual convergence. To relate the stopping criteria with registration convergence, we consider conditions based on image intensities and the deformation update field. Criteria based on image intensities are generally aimed at observing the change in similarity between I_f and I_m at current iteration t with respect to the previous iteration $t-1$. Using Mean Square Error (MSE) of image intensity as the image similarity measure, we define a threshold ε such that

$$\frac{MSE_{current} - MSE_{previous}}{MSE_{previous}} < \varepsilon_{MSE} .$$

Here, ε is an user-controlled parameter, which quantifies the error reduction between iterations. If the registration is converging (i.e. the warped image at t is more similar to I_f than the one at $t-1$), the ratio would be negative, while it goes to zero if the algorithm has plateaued or converged. The ratio will be positive if the chosen step direction is suboptimal, but the registration can be allowed to continue if the degree of non-convergence stays below the user acceptance threshold ε.

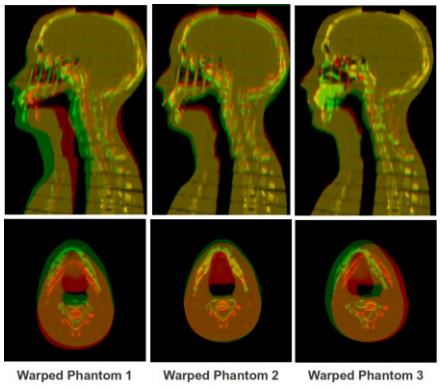

Fig. 1 - The phantom used in this work. In red: the original RANDO ® phantom. In green: the warped phantoms.

TABLE I
CURRENT AND PREVIOUS STOPPING CONDITION VALUE (SCV) FOR MSE, HE AND QUANTITY OF UPDATE

	Current value	Previous value
A	$SCV(t)$	$SCV(t-1)$
B	$SCV(t)$	$SCV(t-3)$
C	$SCV(t)$	$SCV(t-5)$
D	$SCV(t)$	$min(SCV(t-a)), with\ a = [1;6]$
E	$min(SCV(t-a)), with\ a = 0,1,2$	$min(SCV(t-a)), with\ a = [1;3]$
F	$min(SCV(t-a)), with\ a = 0,1,2$	$min(SCV(t-a)), with\ a = [3;6]$

The current and previous SCV tested in this work in terms of convergence speed. For conditions from D to F and stopping criteria HE, the minimum shall be substituted with a maximum, given HE definition.

TABLE II
CHOSEN GAUSSIAN PARAMETERS

Test number		Application Voxel	Radius [#voxels]	Relative Weight	Standard Deviation [#voxels]
1	x	[30 30 30]	15	5	12
	y	[50 50 50]	30	5	18
	z	[80 80 80]	10	100	36
2	x	[60 60 60]	25	1	12
	y	[50 50 60]	30	1	18
	z	[60 60 55]	45	100	36
3	x	[60 60 60]	25	10	12
	y	[50 50 60]	30	10	18
	z	[60 60 55]	45	1000	36

Parameters of the Gaussian distribution used to generate three artificial non-rigid 128*128*128 vector fields. Besides dislocating the centers and changing the standard deviation, we further weighted the distributions by a scaling factor to enhance either one of the components (Relative Weight).

We next consider convergence based on the Harmonic Energy (HE) of the deformation field. HE is defined as the average over all voxels of the squared Frobenius norm of the Jacobian of the vector field. As with MSE, we compute:

$$\frac{HE_{current} - HE_{previous}}{HE_{previous}} > \varepsilon_{HE} .$$

HE is expected to increase with convergence, and therefore the ratio should be always positive unless the optimization is diverging.

The third quantity we consider is derived from the Jacobian of the velocity field. Because the Jacobian matrix of an unstable deformation vector field is negative, we compute the ratio between the number of voxels with a negative Jacobian to the total number of patient voxels and compare it with a user defined threshold, thus

$$\frac{NrJacobianElem < 0}{NrVoxelOfPatient} < \varepsilon_{JAC} .$$

Finally, as in [3], we analyze the Quantity of Update (QU) between iterations. QU is defined as:

$$l_T = \frac{\sum |dr_t|}{\sum |r_{t-1}|}$$

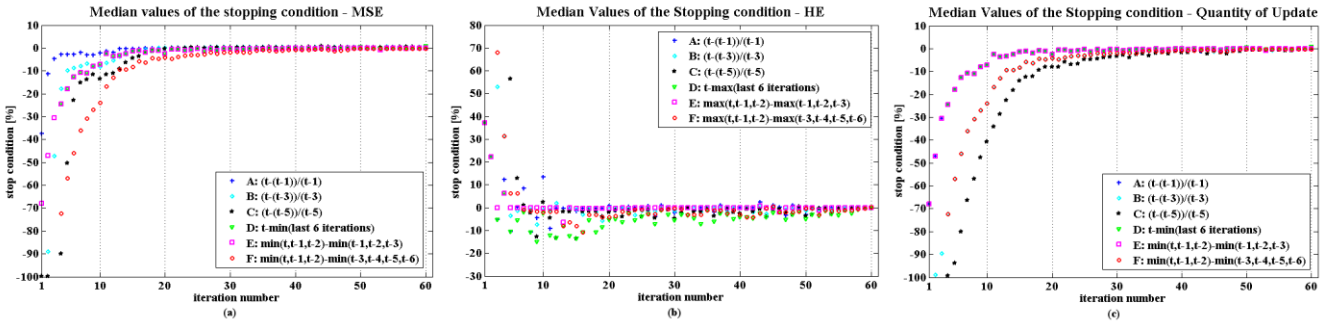

Fig. 2 - Median value of all the considered stopping criteria (Panels a – c) for each of the computation strategies in Table I. The most critical condition is HE, where several of the calculation strategies happears unstable. Computational strategy D seems to be a very good compromise, but it is unstable with HE.

where $\boldsymbol{r}_{t-1}$ is the velocity field at a previous iteration and $d\boldsymbol{r}_t$ is the update field at current iteration. $\boldsymbol{l}_t$ decreases with convergence and therefore we can compute

$$\left|\frac{l_{current}-l_{previous}}{l_{previous}}\right|<\varepsilon_{INCR}.$$

At convergence, $l_{current}\to 0$ and as a consequence, the ratio decreases below the allowed percentage. Each of the metrics, besides the Jacobian, requires the comparison with a previous Stopping Condition Value (SCV). Table I summarizes the methods we analyzed in terms of convergence speed.

C. Testing dataset

We tested our approaches on an image of a RANDO® phantom acquired on a clinical CT scanner, using supine setup and clinical acquisition protocols. The volume acquired is 512*512*123 voxels and [0.94, 0.94, 3] mm element spacing. Ground truth deformations were generated from three artificial non-rigid vector fields, simulating the distribution of left-right, anterior-posterior and inferior-superior components with a smooth, localized Gaussian deformation. The chosen Gaussians are centered at different voxels in a matrix of size 128x128x128 and have diverse standard deviation and amplitude, as reported in Table II. After re-sampling the Gaussian deformations on the image grid, they are superimposed to define a continuous and smooth vector field. This vector field is then recovered with the demons algorithm. The three warped phantoms are shown in Fig. 1.

D. Experiments

We register the original image volume (I_f) to each of the three warped phantoms (I_m) trying to reproduce the same deformation vector fields we simulated. We run 300 iterations at one-eighth of the image resolution (i.e. a coarse registration stage). For computing the convergence values, we exclude background (air) voxels from the computation, because, on one side, we are only interested in patient volumes and, on the other side, nothing can be said from deformation calculated in homogeneous areas. At each iteration, the algorithm computes the deformation vector field and we compute the stopping condition values described in Table 1. To rank the stopping condition performance, we compute the vector length of the residual deformation at each iteration, by means of component-wise subtraction of the ground-truth displacements from the demons vector field. Given the image resolution, we set the threshold on the vector length residual error to 1.875 mm (i.e. 2 voxels at full resolution, or half a voxel at one-eigth resolution), and assess the number of iteration required to obtain the desired level of accuracy ($t_{required}$). We combine this information with the SCV at threshold iteration and we call it $SCV_{critical_i}$ with i=[1;N] ($N=3$ in this case). Ideally, if a given escape condition is optimal for the problem, the $SCV_{critical_i}$ would match $t_{required}$ for all experiments. We then rank the performance of the stopping criteria in terms of number of extra iterations needed for all experiments to reach $SCV_{critical_i}$. In addition, we rank the SCV calculation strategies in terms of convergence speed and stability.

III. Result and Discussion

A. Convergence properties

Fig. 2 illustrates the median over the three phantoms of MSE, HE and QU values at each iteration with each of the computational strategies in Table I. We notice that all of the strategies are potentially good for the assigned registration problem, but some of them are not as computational efficient as the others. For example, strategies F and C are suboptimal, as they require more iteration than e.g. E. Strategy A might be undesirable, because it monitors just the immediately previous iteration, and the user has to find the best tradeoff between risk of stopping early and extra iterations, to achieve a similar registration quality. Strategy D is a good candidate for MSE (Fig.1, Panel a) and QU (Fig.1, Panel c) but it is unstable for HE (Fig. 1, Panels c and d). Strategies B and E seem to capture the best tradeoff between stability and speed of convergence. We suggest that E is the most stable one especially for HE, though might require a few extra iterations when employed in MSE and QU.

B. Deformation recovery capability

Figure 3 shows the plot of the median, 25th and 75th percentile for the residual of the displacement field . It is evident that more iterations do not necessarily imply better convergence, which reinforces the need for a stopping condition. The chosen threshold should stop the algorithm near the minimum residual.

We report the results on the metric performances in Tables III and IV. Table III illustrates the $SCV_{critical}$ at $t_{required}$ (1.875 mm). First of all, notice that the Jacobian has a constant value. This is probably due to the large denominator (i.e. the number of voxels belonging to the patient) with respect to numerator and a phantom study might not be appropriate to study this property. In addition, this stopping condition might be more appropriated for non diffeomorphic [7] and/or very large deformations. Future work will include testing threshold on the Jacobian elements greater than 0. This is motivated from the observation that, if an element goes below zero, the deformation is actually already unacceptable. We then compute the number of extra iterations required to reach the minimum of the $SCV_{critical}$ at $t_{required}$ (Table IV). For example, considering MSE, we look at the $SCV_{critical}$ for first phantom in the other two simulation. The best performance algorithms are HE and QU, while MSE tends to need more iterations at a given threshold value. The two metrics HE and QU describe two different aspects of the DIR problem. The value of the first one can indeed be related to the smoothness of the field, while QU quantifies how much forces are pulling the transformation. Therefore we propose in our future work to investigate a combination of HE and QU. Further effort is also needed to study the effects of image artifacts and inhomogenieties in the image Hounsfield Units.

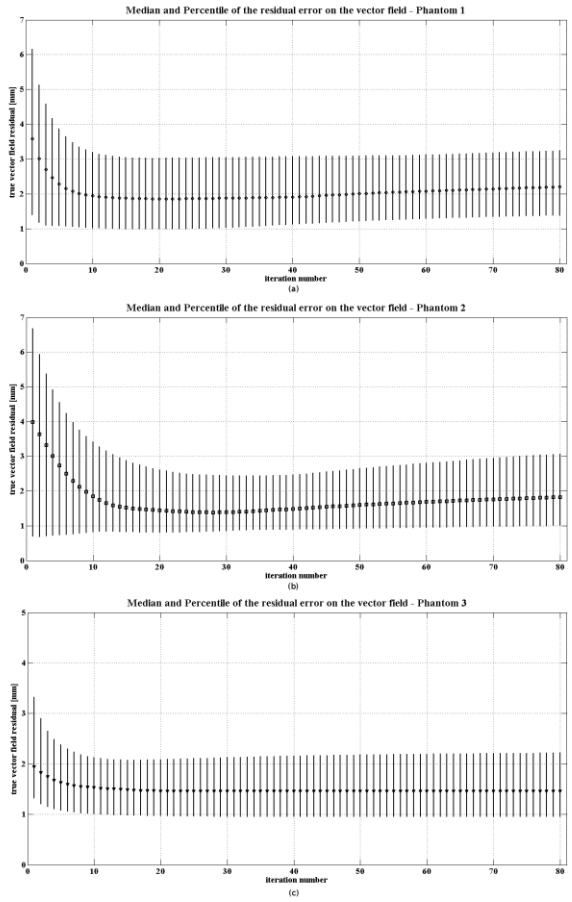

Fig. 3 - Median, 25th and 75th percentile for the residual of the displacement field difference for the three phantoms was used to compute the required number of iterations needed to achieve the desired accuracy on the vector field.

TABLE III
METRIC PERFORMANCES

Test number	$t_{required}$	MSE	HE	QU	Jacobian
1	16	-0.003	0.00	-0.03	0
2	10	-0.03	0.00	-0.17	0
3	2	-0.83	0.17	---	0

$SCV_{critical}$ at the iteration corresponding to the chosen threshold (1.875

TABLE IV
EXTRA ITERATIONS

Test number	MSE	HE	QU	Jacobian
1	0	0	0	0
2	22	0	5	0
3	15	3	6	0

Number of extra iteration the algorithm needed to reach the $SCV_{critical}$.

IV. CONCLUSION

Studying the performances of a stopping criteria is fundamental to the final implementation of DIR in clinical context. We studied four common stopping criteria and analyzed the data both for convergence and for deformation recovery capability. For each metric, we first answer to the question which $SCV_{previous}$ should the $SCV_{current}$ compared with and how. Results were presented which supported the hypothesis that calculation strategies based on the previous iteration are very likely to be outperformed by more complex strategies like E (see table III).

We also analyzed how well these methods can recover a known vector field, relative to minimum number of iteration that satisfiy our accuracy requirements (Table IV). Based on these results, we suggest that the hypothesis combining HE and QU would lead to computationally reasonable and accurate results.

REFERENCES

[1] Thirion JP. *Image matching as a diffusion process: an analogy with maxwell's demons.* Medical Image Analysis, 2(3):243–260, 1998.

[2] Holden M., *A review of geometric transformations for nonrigid body registration.,* IEEE Trans. Med. Imaging 27:111–28, 2008

[3] Yang D.,Li H.,Low D.,Deasy J.O. and El Naqa I., *A fast inverse consistent deformable image registration method based on symmetric optical flow computation.,* Phys. Med. Biol., 53:6143-65, 2008.

[4] Gu X., Choi D., Men C., Pan H., Majumdar A. and Jiang S., *Implementation and evaluation of various demons deformable image registration algorithms on a GPU.* Phys Med Biol,7;55(1):207-19, 2010.

[5] Schmidt-Richtberg A., Ehrhardt J., Werner R., Handels H., *Diffeomorphic Diffusion Registration of Lung CT Images*. In Medical Image analysis for the clinic: A grand Challenge, Workshop Proceedings of the 13th International Conference MICCAI, Beijing, 2010

[6] Nithiananthan, Brock , Daly, Chan , Irish, Siewerdsen, *Demons deformable registration for CBCT-guided procedures in the head and neck: convergence and accuracy.* Med Phys., 36(10):4755-64, 2009.

[7] Vercauteren T, Pennec X, Perchant A, and Ayache N. *Diffeomorphic Demons: Efficient Non-parametric Image Registration.* NeuroImage, Vol. 45, No. 1, Supp. 1, March 2009