Nonrigid Registration Combining Global and Local Statistics

Zhao Yi Stefano Soatto
University of California, Los Angeles
Los Angeles, CA 90095, USA
zyi@cs.ucla.edu soatto@cs.ucla.edu

Abstract

In this paper we exploit normalized mutual information for the nonrigid registration of multimodal images. Rather than assuming that image statistics are spatially stationary, as often done in traditional information-theoretic methods, we take into account the spatial variability through a weighted combination of global normalized mutual information and local matching statistics. Spatial relationships are incorporated into the registration criterion by adaptively adjusting the weight according to the strength of local cues. With a continuous representation of images and Parzen window estimators, we have developed closed-form expressions of the first-order variation with respect to any general, nonparametric, infinite-dimensional deformation of the image domain. To characterize the performance of the proposed approach, synthetic phantoms, simulated MRIs, and clinical data are used in a validation study. The results suggest that the augmented normalized mutual information provides substantial improvements in terms of registration accuracy and robustness.

1. Introduction

Exploratory medical studies are most informative when they probe tissues and organs with different modalities, each providing complementary anatomical or physiological information. Before any integrated analysis can be applied, imaging data generated by each modality must be related to a common reference, which poses a challenge to current image registration algorithms. Multimodal registration has been the subject of a considerable amount of research, since the pioneering work of Maes *et al*. [6] and Viola *et al*. [26], among others [18]. The idea is not to compare image intensities directly, but instead to compare their distributional properties. Because such distributions are estimated from pixel samples collected in image domains, the joint statistics between two deformed images are extremely sensitive to their region of overlap. This has been recognized in the past, and various methods have been proposed to address the problem, including normalized mutual information [25], entropy correlation coefficients [10], etc.

On the other hand, one could forgo intensity statistics altogether, and instead look at the spatial relationships between points in the image domain that are selected using some extremal mechanism [2, 4, 13], also known as a "feature detector." The simplest methods assume that images are affinely or projectively deformed versions of a common template. This is not general enough to capture the subtleties of inter-subject anatomical variations, and one has to consider more general, nonparametric, infinite-dimensional deformations.

While deformable registration based on intensity statistics is now commonplace, only parametric, finite-dimensional transformations have been considered in normalized mutual information based registration, either affine [11, 14, 25, 28] or splines [12, 14, 21, 28]. This may be partly to blame on the fact that analytically computing the gradient of normalized mutual information with respect to free-form deformations is non-trivial. *In this paper we present*[1] *the analytical expression of the functional gradient of normalized mutual information with respect to a general, nonparametric, infinite-dimensional deformation, and illustrate its application in nonrigid multimodal registration.* We demonstrate the performance of an ensuing registration algorithm on synthetic, simulated, and clinical images.

[1]Two of the work related to our research are proposed by D'Agostino [7] and Hermosillo *et al*. [8]. The authors derived the analytical expression of the mutual information gradient in continuous case. However, in their formulation the variation of the region of overlap is not counted and therefore the derivative is only a first-order approximation. Moreover, it only considers mutual information, but not normalized mutual information. Contrarily, we calculated the correct derivative of the gradient of normalized mutual information.

978-1-4244-3991-1/09/$25.00 ©2009 IEEE

More importantly, we address another shortcoming of normalized mutual information and other information-theoretical discrepancy measures for registration, by *introducing spatial constraints in the matching, via an adaptive weighting of local and global statistics*. Where the latter are sufficient to support registration, they are used to guide the local structure of the deformation field. Elsewhere, local statistics take over.

1.1. Relation to prior work

A number of extensions have been explored to include spatial information since the introduction of (normalized) mutual information [18]. The simplest is to consider local features, such as region labeling by Studholme *et al*. [24], local gradients by Pluim *et al*. [17], or shape information by Yi and Soatto [29]. An alternative strategy is to consider neighboring pixels instead of a single pixel in the formulation of information theoretical measures, such as Jumarie entropy by Rodriguez and Loew [19], second-order mutual information by Rueckert *et al*. [20], and region mutual information by Russakoff *et al*. [22] and Bardera *et al*. [1]. Improved robustness has been reported for the augmented measures over standard ones. Despite the general promising results, the aforementioned approaches are still based on the assumption of spatial stationarity; they assume that the intensity distribution in the given images are spatially invariant; therefore, the same global statistics, computed on the entire image domain, are used to guide the registration process. This is often an unrealistic assumption in medical images.

In this paper, we seek a different approach to incorporate spatial information by examining the spatial variability of local statistics in both images jointly. Computed within a neighborhood of corresponding pixels, local statistics are specific for each pixel pair. They encode the neighborhood information and are useful cues for registration. However, previous efforts to exploit local statistics [8, 9, 23, 27] have shown some drawbacks: Local statistics are not effective beyond a pre-selected scale, and thus the registration process is likely to be trapped in local minima. Also, local statistics are often estimated from a small sample size, and therefore more sensitive to noise and outliers. We propose to incorporate spatial information by a weighted combination of the global matching functional on the whole domain, and of the local statistics around individual pixels. Registration is then performed by densely blending the normalized mutual information of local neighborhoods with that of the global images, according to how strong local cues are in that neighborhood.

In Sect. 2, under the general framework of information-theoretical image registration, we justify the need to combine global and local statistics for multi-modality registration (§2.1), which leads to the formulation of weighted normalized mutual information (§2.2). A probabilistic interpretation of our approach is also provided (§2.3), followed by the mathematical derivation of its first order variation (§2.4). Then we test the proposed model on various data sets (see Sect. 3) including synthetic phantoms (§3.1) and simulated MRIs (§3.2) with known transformations, as well as clinical images (§3.3) with unknown deformations.

2. Methodology

Each (normalized) image $I : D \mapsto [0, 1]$ is associated with a probabilistic model by introducing a location random variable X, which is uniformly distributed in D, and its related intensity random variable $I(X)$. Given two images, a template $I_1 : D_1 \mapsto [0, 1]$ and a target $I_2 : D_2 \mapsto [0, 1]$, a deformation is represented by a mapping $w : D_2 \mapsto D_1$ which belongs to a Hilbert space. The joint intensity distribution of $I_1(w(X))$ and $I_2(X)$ in the overlap region $V = D_2 \cap w^{-1}(D_1)$ can be estimated from the pixel samples using a Parzen window

$$p(i_1, i_2; w) = \frac{1}{|V|} \int_V G_\sigma(i_1 - I_1(w(x)), i_2 - I_2(x))\, dx, \tag{1}$$

where G_σ denotes the isotropic Gaussian kernel with a standard deviation of σ. Commonly used information-theoretical discrepancy measures include the joint entropy $\mathcal{H}$, mutual information (MI) $\mathcal{M}$, normalized mutual information (NMI) $\mathcal{N}$, etc.

$$\begin{aligned} \mathcal{H}(I_1, I_2; w) &= -\iint p(i_1, i_2; w) \log p(i_1, i_2; w)\, di_1 di_2, \\ \mathcal{M}(I_1, I_2; w) &= \mathcal{H}(I_1; w) + \mathcal{H}(I_2; w) - \mathcal{H}(I_1, I_2; w), \\ \mathcal{N}(I_1, I_2; w) &= \frac{\mathcal{H}(I_1; w) + \mathcal{H}(I_2; w)}{\mathcal{H}(I_1, I_2; w)}. \end{aligned} \tag{2}$$

To register $I_1 \circ w$ with I_2, $\mathcal{H}$ is to be minimized, while $\mathcal{M}$ and $\mathcal{N}$ are to be maximized. It has been pointed out [18, 24] that NMI is less sensitive than MI to changes in the region of overlap. So in this paper we choose NMI as the candidate discrepancy measure. For simplicity, we drop I_1, I_2 in all formulae and use a subscript to denote marginal entropies. For example, $\mathcal{H}(w) \doteq \mathcal{H}(I_1, I_2; w)$, while $\mathcal{H}_1(w) \doteq \mathcal{H}(I_1; w)$.

2.1. Global vs. Local Statistics

Global statistics are computed in the region where the target and the deforming template images overlap. They

change during registration, and thus can be used to estimate current transformations. However, global statistics only characterize the general trend (image level) of intensity dependence, therefore are unaware of the spatial variability (patch level) of overlap statistics. For the very reason, we need also to look at local statistics.

Local statistics are calculated in the same way as global ones, except that the region of overlap is reduced to a local window centered at each pixel. Global statistics are coarse but robust, whereas local statistics encode the structure of the distribution on a finer granularization of the domain. We illustrate the power of local statistics through a registration example. Shown in Fig. 1 are aligned synthetic phantoms (similar to those used in [13]) where both images contain an elliptical region. We examine the changes of global NMI and mean local NMI with respect to horizontal translations. It can be seen clearly that global NMI is useless for registering the given images, while mean local NMI shows a clear peak corresponding to the ground truth registration.

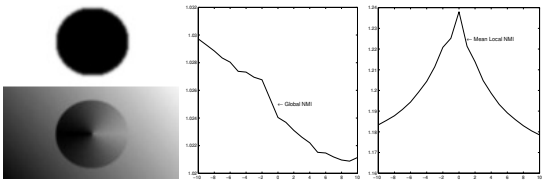

Figure 1. *Comparison between global and local Statistics: (left) aligned synthetic phantoms, (middle-right) changes of global and mean local NMI with horizontal translations. Notice that the global NMI keeps decreasing when the given images are brought into alignment, while mean local NMI has a unique peak corresponding to the correct registration.*

As we have already pointed out, local statistics have their own drawbacks: They are aggregated on regions of a certain scale, and therefore they are not affected by deformations at different scales, making the resulting registration algorithm more sensitive to local minima. Therefore, we propose to combine global and local statistics, instead of employing only one or the other.

2.2. Weighted NMI

To integrate local cues into a global discrepancy criterion, we employ a simple weighting scheme. Let $\mathcal{N}_g$ denote the global NMI, $\mathcal{N}_l(x)$ denote the local NMI associated with pixel x, then the weighted NMI is defined as

$$\mathcal{N}_+ = \frac{1}{|V|}\int_V (1-\alpha)\mathcal{N}_g + \alpha\mathcal{N}_l(x)\,dx, \tag{3}$$

where the parameter $\alpha \in [0,1]$ trades off local and global NMI. The extreme cases $\alpha = 0, 1$, correspond to global and local NMI respectively.

The choice of α is a matter of modeling, and there is no right or wrong answer; it should depend on the spatial variability of overlap statistics. Instead of assuming α to be a constant, we spatially adjust it according to how strong local cues are present in both images. Since NMI by definition is bounded between 1 and 2, local NMI can be adapted to be the weight, *i.e.*, $\alpha(x) = \mathcal{N}_l(x) - 1$. Thus, (3) becomes

$$\mathcal{N}_+ = \frac{1}{|V|}\int_V (2-\mathcal{N}_l(x))\mathcal{N}_g + (\mathcal{N}_l(x)-1)\mathcal{N}_l(x)\,dx. \tag{4}$$

Note that not only $\mathcal{N}_g$, but also $\mathcal{N}_l(x)$ are functionals of the deformation w. If during registration the change in V can be neglected, the variation of $\mathcal{N}_+$ equals

$$\frac{\delta\mathcal{N}_+}{\delta w} = (2-\mathcal{N}_l(x))\frac{\delta\mathcal{N}_g}{\delta w} + (2\mathcal{N}_l(x) - 1 - \mathcal{N}_g)\frac{\delta\mathcal{N}_l(x)}{\delta w}. \tag{5}$$

The idea underlying this model is that, when local statistics provide strong cues for matching the neighborhood of a pixel, say x, *i.e.*, when $\mathcal{N}_l(x)$ is close to 2, we use them to drive the registration around x. On the other hand, when local statistics are weak, i.e., $\mathcal{N}_l(x)$ is close to 1, global statistics are employed as an alternative registration criterion. By adaptively adjusting the weighting spatially, we are able to accommodate large spatial variations of overlap statistics.

2.3. Probabilistic Interpretation

In formal terms, our goal can be stated as maximizing the posterior distribution of the diffeomorphic warping, that is

$$\begin{aligned}\hat{w} &\doteq \arg\sup_w \log p(w|I_1, I_2)\\ &= \arg\sup_w \log p(I_1, I_2|w)p(w).\end{aligned} \tag{6}$$

The second term, $\log p(w)$, can be easily recognized as the generic regularization $\mathcal{R}(w)$, in our model the fluid regularizer [3, 7]. So we concentrate on the log-likelihood term $\log p(I_1, I_2|w)$. This can be obtained via $p(I_1, I_2|w) = p(I_1|I_2, w)p(I_2)$. To this end, the models proposed by [15] could be employed, so what we need to compute is $p(I_1|I_2, w)$. Again in purely formal terms, we could represent the probability of matching given images using local statistics as $P(\mathcal{L})$; then what we wish to compute becomes

$$\begin{aligned}&p(I_1|I_2, w) =\\ &p(I_1|I_2, w, \mathcal{L})P(\mathcal{L}) + p(I_1|I_2, w, \not\mathcal{L})(1 - P(\mathcal{L})).\end{aligned} \tag{7}$$

Now, this is just formal notation. The difficulty comes in when we try to write explicitly the probabilities above, because the condition $\mathcal{L}$ (indicating that local statistics

should be used to guide registration) is specific to each pixel x. We have to specify the spatial statistics of the image, which would lead to an inference problem where all possible combinations of states are possible and time-consuming Monte Carlo methods become necessary rather than simple local descent algorithms.

So, instead of attempting to compute the above likelihood, we will approximate it by assuming that all pixels are independent, and computing an average (expectation) over pixels of the probability

$$p(I_1|I_2,w) = \prod_{x\in D_2} p(I_1(w(x))|I_2(x),w(x)) \propto \prod_{x\in D_2:\mathcal{L}} e^{\mathcal{N}_l(I_1(w(x)),I_2(x))} \prod_{x\in D_2:\not\mathcal{L}} e^{\mathcal{N}_g(I_1(w(x)),I_2(x))}, \tag{8}$$

which, modulo technicalities, should converge to (3).

2.4. Variation of NMI

There is still one question remaining, i.e., how to calculate $\frac{\delta\mathcal{N}_g}{\delta w}, \frac{\delta\mathcal{N}_l}{\delta w}$ in (5). To the best of our knowledge, the use of NMI has been limited to parametric, finite-dimensional transformations [11, 12, 14, 21, 25], while the variation of NMI w.r.t. a general, nonparametric, infinite-dimensional deformation has never been reported before. To derive it, we follow similar steps to the calculation of the gradient of MI in [7, 8], and obtain the Gateaux derivative

$$\left.\frac{\partial\mathcal{N}(w+\epsilon u)}{\partial\epsilon}\right|_{\epsilon=0} = \frac{1}{\mathcal{H}(w)}\iint \left.\frac{\partial p(i_1,i_2;w+\epsilon u)}{\partial\epsilon}\right|_{\epsilon=0}\cdot L(i_1,i_2;w)\,di_1di_2 \tag{9}$$

where $L(w)=\log\frac{p(i_1,i_2;w)^{\mathcal{N}(w)}}{p(i_1;w)p(i_2;w)}$, and $p(i_1;w)$, $p(i_2;w)$ are the marginal intensity distributions. The expansion of the first term of the integrand in (9) will require a level set representation [16] for the region of overlap. Interested readers may refer to the appendix for details. Here we only present the final formula

$$\begin{aligned}&\left.\frac{\partial\mathcal{N}(w+\epsilon u)}{\partial\epsilon}\right|_{\epsilon=0}\\ &= \frac{1}{|V|\mathcal{H}(w)}\int_{D_2}\Big\{\left[\frac{\partial G_\sigma}{\partial i_1}*L\right](I_1(w(x)),I_2(x))\cdot\nabla I_1(w(x))\cdot\\ &\quad\cdot(1-H(\phi_1(w(x))))-[G_\sigma*L](I_1(w(x)),I_2(x))\cdot\\ &\quad\cdot\nabla\phi_1(w(x))\cdot\delta(\phi_1(w(x)))\Big\}\cdot u(x)\,dx.\end{aligned} \tag{10}$$

There are two terms in the above derivation. The first one reflects changes in the joint intensity statistics, while the other is caused by changes in the overlapping domain. The variation of $\mathcal{N}$ w.r.t. w thus becomes

$$\begin{aligned}&\frac{\delta\mathcal{N}(w)}{\delta w}\\ &= \frac{1}{|V|\mathcal{H}(w)}\Big\{\left[\frac{\partial G_\sigma}{\partial i_1}*L\right](I_1(w(x)),I_2(x))\nabla I_1(w(x))\cdot\\ &\quad\cdot(1-H(\phi_1(w(x))))-[G_\sigma*L](I_1(w(x)),I_2(x))\cdot\\ &\quad\cdot\nabla\phi_1(w(x))\cdot\delta(\phi_1(w(x)))\Big\}.\end{aligned} \tag{11}$$

3. Results

In this section, three sets of experiments are carried out. First, synthetic phantoms are shown where global NMI fails to yield successful registration, while weighted NMI does. This illustrates the advantage of our approach to accommodate large spatial variations in joint intensity statistics. Then we apply our approach to simulated MR images of human brains where ground truth transformations are available. A comparison is conducted to assess the accuracy and robustness of our approach relative to standard ones. Finally, we present a real world application on multimodality registration of MR and gene expression mouse brain images.

3.1. Synthetic Phantoms

The synthetic phantoms are similar to Fig. 1 except that the top and bottom of the ellipse are non-rigidly deformed to straight edges. As shown in Fig. 2, the weighted NMI correctly exploits the local cues near the edge, therefore can successfully capture the stage of alignment, while the global NMI is adversely affected by the strong intensity gradation in the target image and leads to a spurious registration. Though weighted NMI outperforms global NMI in this case, we should also point out that the rightmost part of the ellipse cannot be matched perfectly. This is expected due to the lack of local cues in this region.

3.2. Simulated MRIs

We validate the accuracy of our approach on simulated MR brain images, obtained from the BrainWeb MR simulator [5] with slice thickness $1mm$, noise level 3% , and intensity non-uniformity 20%. Triplets of pre-registered T1/T2/PD images are selected, where we artificially deform all T1 images according to randomly generated transformations w^*. To ensure regularity, w^* is smoothed by a Gaussian kernel with standard deviation $\sigma = 16$, followed by a normalization step to set the maximal displacement to 6 pixels. Fig. 3 shows representative cases prepared for validation.

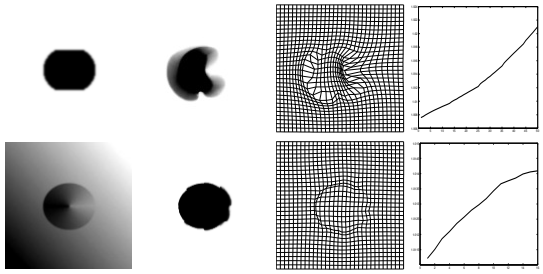

Figure 2. *Comparison between global (top) and weighted (bottom) NMI on synthetic phantoms. The left column shows the images of template and target, while the other columns present the results of deformed templates (middle-left), deformation fields (middle-right), and NMI values (right), respectively.*

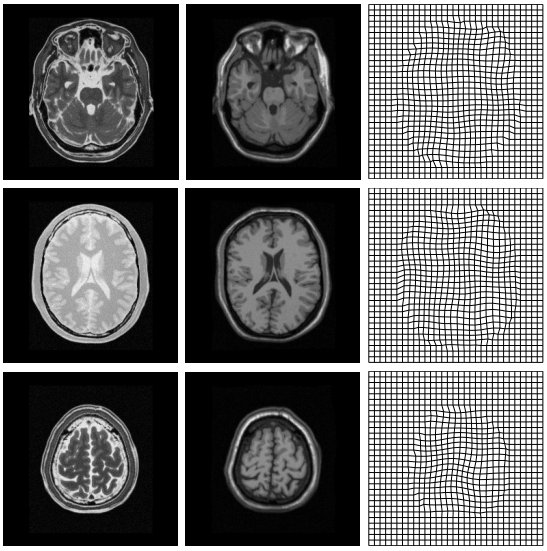

Figure 3. *Representative cases prepared for validation. Each row from left to right corresponds to preregistered T2/PD images (templates), artificially deformed T1 images (targets), and the ground truth transformations, respectively.*

Registration is then performed by choosing T2/PD images to be the templates, and the deformed T1 images to be the targets. The recovered deformations $\hat{w}$ are compared with the ground truth transformations w^* by the root mean squared (RMS) error within the brain region. We test the global, local, and weighted NMI on 11 triplets and obtain the following table (Tab. 1). It is clear that our approach quantitatively achieves more accurate (*i.e.*, mean RMS error smaller) and robust (*i.e.*, small RMS error in most cases) registration than the standard ones. Note that for T2-T1 registration, the improvement (up to 1.2 pixels) is naturally smaller than that of PD-T1 registration (up to 1.6 pixels) since there is less spatial variability of overlap statistics.

Case	T2-T1			PD-T1		
	GNMI	LNMI	WNMI	GNMI	LNMI	WNMI
01	1.107	0.888	0.657	1.686	1.016	0.819
02	0.946	0.705	0.590	2.273	0.762	0.673
03	1.080	0.833	0.868	2.063	1.043	1.013
04	1.108	1.368	0.875	1.629	1.214	0.924
05	0.920	0.843	0.649	0.897	0.873	0.799
06	0.886	0.722	0.648	0.914	1.113	0.730
07	0.945	0.818	0.745	0.900	1.159	0.777
08	0.632	0.539	0.470	0.907	0.575	0.576
09	0.844	0.816	0.735	1.098	0.674	0.665
10	1.275	1.203	1.006	2.473	1.127	1.051
11	1.898	0.799	0.725	2.393	0.791	0.931
Mean	1.058	0.867	0.724	1.567	0.941	0.814

Table 1. *RMS error in pixels between ground truth and recovered deformations within the brain region. 11 triplets are selected on which two types of multimodal registration are conducted. In almost all cases, weighted NMI yields a lower RMS error than both global and local NMI.*

Level	T2-T1			PD-T1		
	GNMI	LNMI	WNMI	GNMI	LNMI	WNMI
0%	0.970	0.900	0.700	1.217	0.966	0.730
20%	1.058	0.866	0.724	1.567	0.936	0.814
40%	1.234	0.881	0.788	2.356	0.988	0.972

Table 2. *Improvement of average RMS error at different non-uniformity levels. As the non-uniformity level increases, the improvement over global NMI also gets bigger. Local NMI generally behaves better than global NMI in the presence of non-uniformity, however, it is still worse than our approach especially when there is low non-uniformity.*

A registration example (case 04 PD-T1) is presented in Fig. 4 to illustrate the quality of our approach. We apply the recovered deformation, whose magnitude is visualized in fake colors, to the original T1 image. Then this deformed copy is subtracted from the given T1 target to obtain a final estimate of the difference after registration, also visualized in fake colors. As it can be seen, we achieve a difference image exhibiting less structure as well as a displacement map closer to the ground truth transformation by using weighted NMI, which suggests a qualitative improvement over both global and local NMIs.

To further demonstrate the performance of our approach, we test it on the same data with different non-uniformity levels $0\%, 40\%$, and obtain the average RMS error in Tab. 2. Since our approach is motivated to accommodate large spatial variations of overlap statistics, the improvement of the average RMS error over global

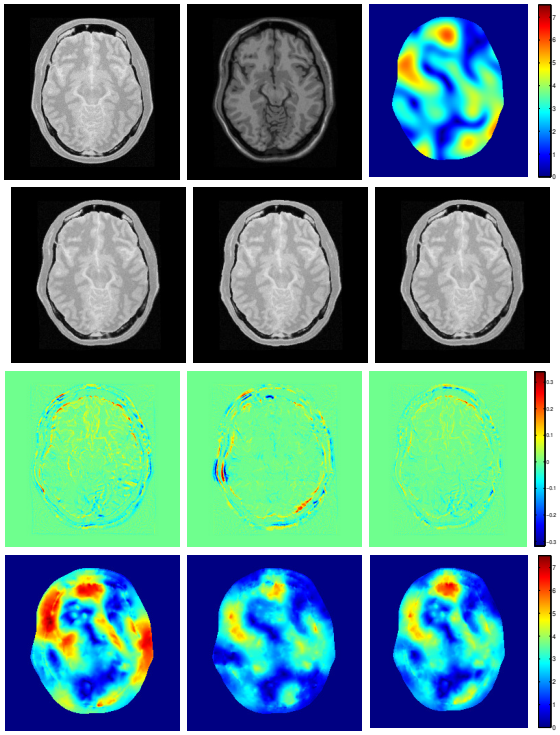

Figure 4. *Comparison between global (left), local (middle), and weighted (right) NMI on simulated MRIs. The top row corresponds to images of template, target, and ground truth displacement magnitude, while the other rows present the results of deformed templates (middle-top), difference images (middle-bottom), and displacement magnitudes (bottom), respectively. Both difference images and displacement magnitudes are visualized in fake colors, with green/blue-red representing small/large difference and blue/red representing small/large displacement.*

NMI increases with the level of non-uniformity. Local NMI generally behaves better than global NMI in the presence of non-uniformity, however, it is still worse than our approach especially when there is low non-uniformity. Moreover, local NMI is not robust and may fail in some cases (see case 04 T2-T1 and case 07 PD-T1 in Tab. 1 for example).

3.3. Clinical Data

We also apply our approach to clinical data where anatomical MR and gene expression mouse brain images need to be registered. The results are shown in Fig. 5, together with those obtained by a standard approach. Since the input images are obtained independently of each other, ground truth transformations are not known. Therefore, we propose to validate the accuracy of registration in three ways: visual observation, landmark tracking, and Jacobian calculation.

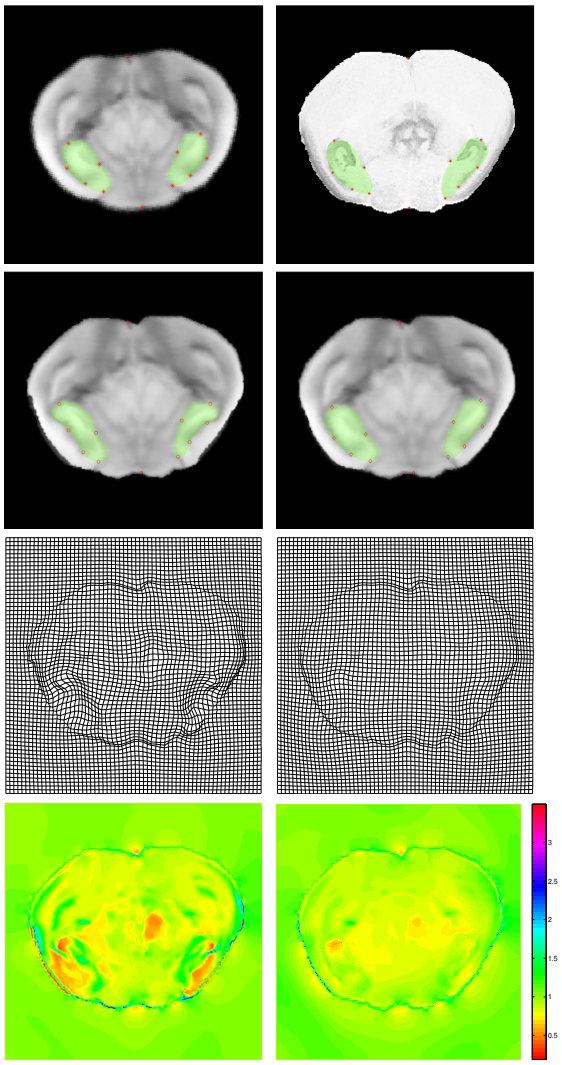

Figure 5. *Comparison between global (left) and weighted (right) NMI on clinical data. The top row corresponds to images of MR template and gene expression target, while the other rows present the results obtained by each registration, including deformed templates (middle-top), deformation fields (middle-bottom), and Jacobian maps (bottom). For better visualization, we delineate the underlying anatomies in each image by green masks. The manually placed point landmarks are marked by $\star$, while the forward-traced ones are marked by $\circ$ for global NMI and $\diamond$ for weighted NMI. The Jacobian maps are visualized in fake colors, with red representing values close to zero (singular), and yellow-green close to one (regular).*

By visually inspecting the deformed template images

(especially within the superimposed green masks), we can see that there are some undesired distortions by using the global NMI as a registration criterion. However, these artifacts are diminished if we switch to the weighted NMI criterion. To further examine the quality of both approaches, several corresponding pairs of point landmarks are manually placed on the template and target images according to photometrically or geometrically strong cues. Then we trace those on the target image back to the template image by applying the recovered deformations from registration. Distances between the obtained landmarks and their desired locations are calculated in pixels, with a RMS error of 6 pixels for our approach compared to nearly 7 pixels for the standard approach. More importantly, our approach achieves a deformation field with much better quality (smoother and more regular). This can been easily verified from the color-represented Jacobian maps (yellow-green: close to 1, red: close to 0). While the Jacobian map obtained by our approach (minimum 0.5, maximum 2.6) shows mostly yellow or green, the standard approach (minimum 0.2, maximum 3.5) yields a Jacobian map with many red regions.

4. Conclusion

We have presented a derivation of the functional gradient of normalized mutual information with respect to any infinite-dimensional deformation. This has been used to define a variational algorithm to perform multimodal image registration. Some of the fundamental shortcomings of global and local mutual information as a discrepancy measure for image registration have been pointed out, and they have been addressed by introducing a model that weights local and global statistics in normalized mutual information, allowing the local structures of the image, when present, to drive the registration, and the global statistics to take over elsewhere. We have illustrated the performance of our approach both quantitatively and qualitatively on synthetic phantoms and simulated MRIs with ground truth transformations, as well as on real clinical data of multimodal (anatomical MR and gene expression) mouse brain images with unknown deformations.

Acknowledgments

This research was supported by NIH U54 RR021813.

Appendix

In the following we derive the Gateaux derivative of the joint intensity distribution $p(i_1, i_2; w)$ with respect to a general, nonparametric, infinite-dimensional deformation w. Recall the Parzen window estimator in (1) and let $f(w)$ denote the numerator, then

$$f(w) = \int_{D_2} G_\sigma(i_1 - I_1(w(x)), i_2 - I_2(x)) \cdot \cdot (1 - H(\phi_1(w(x)))) \, dx. \tag{A-1}$$

Here H is the Heaviside function

$$H(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise,} \end{cases} \tag{A-2}$$

and ϕ_1 is the signed distance function of boundary ∂D_1

$$\phi_1(x) = \begin{cases} -\min_{y \in \partial D_1} ||x - y||, & \text{if } x \in D_1 \\ +\min_{y \in \partial D_1} ||x - y||, & \text{otherwise.} \end{cases} \tag{A-3}$$

Similarly, let $z(w)$ denote the denominator, we have

$$z(w) = |V| = \int_{D_2} (1 - H(\phi_1(w(x)))) \, dx. \tag{A-4}$$

Now let w be perturbed by an infinitesimal amount in an arbitrary direction u, we have the Gateaux derivatives of $f(w)$

$$\begin{aligned} &\left.\frac{\partial f(w + \epsilon u)}{\partial \epsilon}\right|_{\epsilon=0} \\ &= \int_{D_2} \Big\{ -\left[\frac{\partial G_\sigma}{\partial i_1}\right] (i_1 - I_1(w(x)), i_2 - I_2(x)) \cdot \nabla I_1(w(x)) \cdot \\ &\quad \cdot (1 - H(\phi_1(w(x)))) - G_\sigma(i_1 - I_1(w(x)), i_2 - I_2(x)) \cdot \\ &\quad \cdot \nabla \phi_1(w(x)) \cdot \delta(\phi_1(w(x))) \Big\} \cdot u(x) \, dx, \end{aligned} \tag{A-5}$$

and $z(w)$

$$\left.\frac{\partial z(w + \epsilon u)}{\partial \epsilon}\right|_{\epsilon=0} = \int_{D_2} -\nabla \phi_1(w(x)) \cdot \delta(\phi_1(w(x))) \cdot u(x) \, dx, \tag{A-6}$$

where $\delta(x) = H'(x)$ is Dirac's delta. Consequently, the Gateaux derivative of $p(i_1, i_2; w)$ has the following form

$$\begin{aligned} &\left.\frac{\partial p(i_1, i_2; w + \epsilon u)}{\partial \epsilon}\right|_{\epsilon=0} \\ &= \frac{1}{|V|} \int_{D_2} \Big\{ -\left[\frac{\partial G_\sigma}{\partial i_1}\right] (i_1 - I_1(w(x)), i_2 - I_2(x)) \cdot \nabla I_1(w(x)) \cdot \\ &\quad \cdot (1 - H(\phi_1(w(x)))) - \big(G_\sigma(i_1 - I_1(w(x)), i_2 - I_2(x)) - \\ &\quad - p(i_1, i_2; w)\big) \cdot \nabla \phi_1(w(x)) \cdot \delta(\phi_1(w(x))) \Big\} \cdot u(x) \, dx. \end{aligned} \tag{A-7}$$

If the variation of V can be neglected during registration (as in [7]), the last two terms vanish and (A-7) thus reads

$$\begin{aligned} &\left.\frac{\partial p(i_1, i_2; w + \epsilon u)}{\partial \epsilon}\right|_{\epsilon=0} \\ &= \frac{1}{|V|} \int_V \Big\{ -\left[\frac{\partial G_\sigma}{\partial i_1}\right] (i_1 - I_1(w(x)), i_2 - I_2(x)) \cdot \\ &\quad \cdot \nabla I_1(w(x)) \Big\} \cdot u(x) \, dx. \end{aligned} \tag{A-8}$$

References

[1] A. Bardera, M. Feixas, I. Boada, and M. Sbert. High-dimensional normalized mutual information for image registration using random lines. In *Workshop on Biomedical Image Registration*, pages 264–271, 2006.

[2] P. Cashier, E. Bardinet, D. Dormont, X. Pennec, and N. Ayache. Iconic feature based nonrigid registration: the pasha algorithm. *Computer Vision and Image Understanding*, 89:272–298, 2003.

[3] G. Christensen, R. Rabbitt, and M. Miller. Deformable templates using large deformation kinematics. *IEEE Trans. Image Processing*, 5:1435–1447, 1996.

[4] H. Chui, L. Win, R. Schultz, J. Duncan, and A. Rangarajan. A unified non-rigid feature registration method for brain mapping. *Medical Image Analysis*, 7:113–130, 2003.

[5] C. Cocosco, V. Kollokian, R.-S. Kwan, and A. Evans. Brainweb: online interface to a 3d mri simulated brain database. In *Int. Conf. Functional Mapping of the Human Brain*, page S427, 1997.

[6] A. Collignon et al. Automated multimodality medical image registration using information theory. In *Proc. Int. Conf. Information Processing in Medical Imaging: Computational Imaging and Vision*, pages 263–274, 1995.

[7] E. D'Agostino, F. Maes, D. Vandermeulen, and P. Suetens. A viscous fluid model for multimodal non-rigid image registration using mutual information. *Medical Image Analysis*, 7:565–575, 2003.

[8] G. Hermosillo and O. Faugeras. Dense image matching with global and local statistical criteria: a variational approach. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 73–78, 2001.

[9] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, and P. Suetens. Nonrigid image registration using conditional mutual information. In *Int. Conf. Information Processing and Medical Imaging*, pages 725–737, 2007.

[10] F. Maes, A. Collignton, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Trans. Medical Imaging*, 16:187–198, 1997.

[11] F. Maes, D. Vandermeulen, and P. Suetens. Comparative evaluations of multiresolution optimization strategies for multimodality image registration by maximization of mutual information. *Medical Image Analysis*, 3.

[12] D. Mattes, D. Haynor, H. Vesselle, T. Lewellen, and W. Eubank. Pet-ct image registration in the chest using free-form deformations. *Medical Image Analysis*, 22:120–128, 2003.

[13] M. Mellor and M. Brady. Phase mutual information as a similarity measure for registration. *Medical Image Analysis*, 9:330–343, 2005.

[14] C. Meyer and etal. Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. *Medical Image Analysis*, 1:195–206, 1997.

[15] D. Mumford and B. Gidas. Stochastic models for generic images. *Communications in Pure and Applied Mathematics*, 54(1):85–111, 2001.

[16] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer, 2002.

[17] J. Pluim, J. Maintz, and M. Viegever. Image registration by maximization of combined mutual information and gradient information. *IEEE Trans. Medical Imaging*, 19:809–814, 2000.

[18] J. Pluim, J. Maintz, and M. Viegever. Mutual information based registration of medical images: a survey. *IEEE Trans. Medical Imaging*, 22:986–1004, 2003.

[19] C. Rodriguez-Carranza and M. Loew. A weighted and deterministic entropy measure for image registration using mutual information. In *SPIE Conf. Image Processing*, pages 155–166, 1998.

[20] D. Rueckert, M. Clarkson, D. Hill, and D. Hawkes. Non-rigid registration using higher-order mutual information. In *Proc. SPIE Image Processing: Medical Imaging*, pages 438–447, 2000.

[21] D. Rueckert et al. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE Trans. Medical Imaging*, 18:712–721, 1999.

[22] D. Russakoff, C. Tomasi, T. Rohlfing, and C. Maurer. Image similarity using mutual information of regions. pages 596–607, 2004.

[23] C. Studholme, C. Drapaca, B. Iordanova, and V. Cardenas. Deformation-based mapping of volume change from serial brain mri in the presence of local tissue contrast change. *IEEE Trans. Medical Imaging*, 25:626–639.

[24] C. Studholme, D. Hill, and D. Hawkes. Automated 3-d registration of mr and ct images of the head. *Medical Image Analysis*, 1:163–175.

[25] C. Studholme, D. Hill, and D. Hawkes. An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, 32:71–86.

[26] P. Viola and W. Wells. Alignment by maximization of mutual information. *Int. Journal of Computer Vision*, 24:137–154, 1997.

[27] J. Weese, P. Rosch, T. Netsch, T. Blaffert, and M. Quist. Gray-value based registration of ct and mr images by maximization of local correlation. In *Medical Image Computing and Computer-Assisted Intervention*, volume 1679, pages 656–664, 1999.

[28] R. Xu, Y.-W. Chen, S.-Y. Tang, S. Morikawa, and Y. Kurumi. Parzen-window based normalized mutual information for medical image registration. *IEICE Trans. Information and Systems*, 91:132–144, 2008.

[29] Z. Yi and S. Soatto. Correspondence transfer for the registration of multimodality images. In *IEEE Int. Conf. Computer Vision*, 2007.