Correspondence Transfer for the Registration of Multimodal Images

Zhao Yi Stefano Soatto
Computer Science Department, University of California
Los Angeles, CA 90095
{zyi,soatto}@cs.ucla.edu

Abstract

Gene expression data provide information on the location where certain genes are active; in order for this to be useful, such a location must be registered to an anatomical atlas. Because gene expression maps are considerably different from each other – they display the expression of different genes – and from the anatomical atlas, this problem is currently addressed either manually by trained experts, or by neglecting all image information and only using the pre-segmented boundaries. In this manuscript we concentrate on data discrepancy measures that take into account image information when this is present in both the target and template images. We exploit such "bi-lateral" structures to drive the correspondence process in regions where the intensity information is inconsistent, analogously to a "motion inpainting" task. Although no ground truth can be established, and prior information clearly plays a key role, we show that our model achieves desirable results on subjective tests validated by expert subjects.

1. Introduction

Establishing correspondence between different images is key for us to infer properties of the underlying scene. The basic assumption is that there is *something* common between the images, modulo domain deformations (*e.g.* induced by viewpoint changes or by scene deformations) and range deformations (*e.g.* contrast transformations induced by changes in illumination, or by changes of imaging modality). Such commonality may be abstract, rather than physical, for instance when the images portray objects in the same category, say "hippocampus," even though each image portrays a different physical object. A crucial component of any approach to registration is the mechanism used to compare two (deformed) images: While range (intensity) similarity is a natural choice, for instance measured in the sense of $\mathbb{L}^2$ [30] or Total Variation [26], extreme contrast changes have been successfully tackled using Mutual Information [22].

The most recent developments in medical imaging, however, are challenging these premises altogether: Gene expression data are generated with different stains, highlighting different genes, with the express goal of making each resulting image as different as possible from the others, in order to maximize their information content. Nevertheless, the practitioner requires registering such images to anatomical atlases, in order to ascribe the activity of a gene to a particular anatomical structure (Fig. 1). The same goes for registering functional imaging (*e.g.* F-MRI) to anatomical atlases, a task that is by and large performed manually by trained physicians. While this is doable for a handful of subjects, systematic statistical studies of gene expression data in large populations call for some degree of automation.

But what does it mean to establish correspondence, when there is no common underlying structure, and when the data are designed to be as different ("independent") from each other as possible? Clearly expert prior knowledge of anatomy and biological functionality is indispensable, and several research groups are actively engaged in modeling, learning and enforcing shape priors in segmentation and registration [16, 25]. Nevertheless, any registration algorithm must also take into account the available data, and this problem has been largely overlooked in the literature, where mostly standard data terms are used [15, 24], or where only the boundary information is taken into consideration and the rest of the deformation field is determined by generic regularization [8, 20, 27]. Therefore, in this manuscript we focus our attention on devising suitable data terms for registering multi-modal images. Our goal is to design a scheme to take visible geometric structures (one could call them "landmark regions") into account when they are present

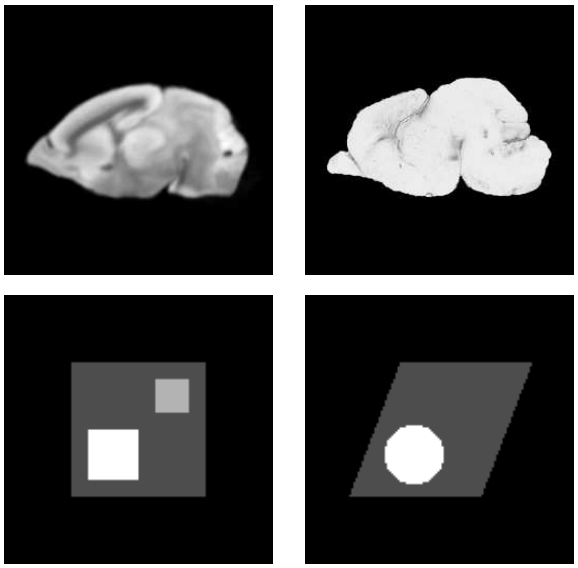

Figure 1. *Gene expression data (top-right) and a manually selected anatomical template (top-left). Detailed correspondence is currently performed largely by hand by trained physicians. Synthetic phantoms (bottom): Certain regions are visible in both the target and the template, although deformed, whereas other regions are visible only in one of the two. The goal is to exploit "bilateral" regions to drive the correspondence process for "unilateral" regions.*

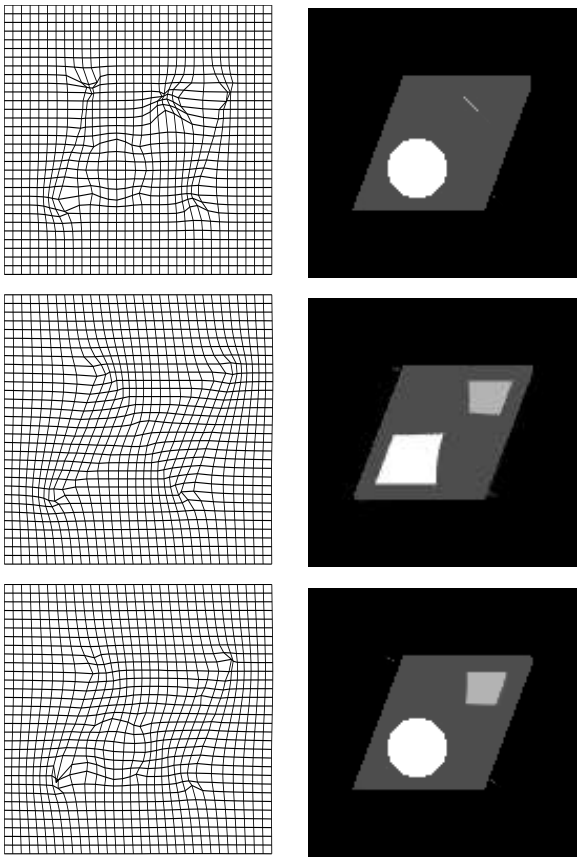

Figure 2. *Correspondence for the phantoms in Fig. 1 using only intensity information (top), using only geometric information (middle), and using the combined model we propose (bottom). In each case we show the deformation field (left) and the mapped template (right). In the case of intensity information alone, unilateral regions disappear (top). In the case of geometric information alone, bilateral regions are not deformed correctly (middle). In the combined model, bilateral regions are deformed according to the data, whereas unilateral regions are mapped according to geometric information (bottom).*

in both images, and use their registration as boundary condition to "guide" the structures that are present in one image (*e.g.* gene expression), but not the other (*e.g.* the atlas). The problem is best illustrated with a phantom, or "cartoon" (Fig. 1). A template image (left) exhibits some visible structures (*e.g.* grey region on the top left), and one is interested in determining where such structure is located on an atlas (right). Unfortunately, such a structure is *absent in the atlas!* Therefore, we need to "transfer" correspondence information from common (or "bilateral") structures in order to infer the motion and deformation of "unilateral" ones. One could think of this problem as "motion inpainting" [3], although one where domain knowledge plays a considerable role.

Any approach that relies on raw intensity information fails this task because, by design, one assumes that images are equivalent up to diffeomorphic domain transformations [2, 16]. In Fig. 2 we show the effects of a common intensity-based algorithm on registering the phantoms of Fig. 1: Bilateral structures are mapped correctly, but unilateral ones shrink to a point, inducing a singularity (sink) in the warping that is not physically plausible in this context (although it would be appropriate in a "growth" model [10]). This problem is mostly addressed in current literature by neglecting intensity information altogether, using instead the outer boundary of the slice. In this case, the deformation is smooth, but bilateral structures are not mapped correctly (Fig. 2). Our goal is to bridge this gap: Where bilateral structures are present, we want to use them to guide our warping. Unilateral structures, on the other hand, should be preserved and mapped onto the atlas. What we need is a spatially-varying criterion that uses intensity information only where available. We will formulate this problem as a probabilistic inference, where the likelihood of the data is weighted at each point by the probability of

there being a bilateral structure.

2. Formalization of the problem

Let $I_1, I_2 : D \subset \mathbb{R}^2 \to \mathbb{R}^+$ be two images $x \mapsto I_j(x),\ j = 1, 2$, and $w : D \to D$ a diffeomorphism of the domain of one onto the other. Within the domain of each image lies a region of interest $\Omega_j \subset D$, whereas the "background" $D - \Omega_j$ is assumed segmented (or equivalently I_j can be assumed to have zero value outside Ω_j.) We represent regions $\Omega \subset D$ using the signed distance function $\phi(x|\Omega) \doteq \pm \min_{y \in \partial\Omega}(|x - y|), x \in D$ with the positive sign for $x \in \Omega$ and negative otherwise. The function ϕ is at least Lifschitz continuous [19].

Within each domain Ω_j, assume that there are regions $B_j \subset \Omega_j, j = 1, 2$ (not necessarily simply connected) that have a distinct photometric signature so that they can be detected by a low-level image processing algorithm. We will make this precise, and indeed we will relax this assumption later; for now, assume that the regions $B_j,\ j = 1, 2$ are known. We call these *bilateral* regions, in the sense that they are detected in both images (Fig. 1). On the other hand, there are regions $U \subset \Omega_j$ that are detected in one image but not the other, which we call *unilateral*. For the sake of illustration we will assume that $U \subset \Omega_1$. This scenario is displayed in its most elementary form in Fig. 1.

Our model is based on the premise that, locally around bilateral regions, the warping w is determined by intensity information, whereas away from bilateral regions, where intensity is constant or inconsistent between the two images (*e.g.* around unilateral regions), the diffeomorphism is determined by the geometry of the regions Ω_1, Ω_2, as well as by generic regularizers. These assumptions can be translated into a simple generative model

$$\begin{cases} I_1(w(x)) = I_2(x), & x \in \Omega_2 \cap \mathcal{B}_\sigma(B_2) \\ B_1 = w(B_2) \\ \phi(w(x)|\Omega_1) = \phi(x|\Omega_2), & x \in D \end{cases} \tag{1}$$

where $\mathcal{B}_\sigma(C)$ is a region including C by a margin $\sigma > 0$ (*e.g.* the union of C with a covering of balls of radius σ around ∂C). We will first review criteria to infer the diffeomorphism w based on geometric information only in Sect. 2.2 (i.e. neglecting the value of $I_j(x),\ x \in \Omega_j$); then based on photometric information only (i.e. on the intensity value of the images) in Sect. 2.3, and finally discuss our model in Sect. 2.4, for which we will provide a probabilistic interpretation in Sect. 2.5. Before doing so, however, we discuss the issue of validation.

2.1. On validation

Naturally, because there is no data to support correspondence of unilateral structures, the result will be a direct consequence of our assumptions (or "model", "prior" or "regularizers", depending on the parlance of the scientific domain of preference). In this sense, the problem is both scientifically ill-defined (i.e. non-falsifiable), and mathematically ill-posed (there are infinitely many solutions that are non-continuously dependent on the initial conditions). As a consequence, "ground truth" cannot be established – similar to Inpainting [3] – and the problem is tautologically defined by its solution. Ultimately, the quality of our result can only be judged subjectively on experiments performed by highly trained anatomists that can establish such a correspondence based *not* on the available data alone (as we do), but based on high-level knowledge that is not available to the untrained eye, reflecting the clinical value of a proposed scheme.

2.2. Geometry-driven component cost

The simplest model to perform registration based only on the shape of the regions Ω_j is to minimize a discrepancy term between $\phi_1(x) \doteq \phi(x|\Omega_1)$ and $\phi_2(x) \doteq \phi(x|\Omega_2)$, for instance the $\mathbb{L}^2$ norm. Because the diffeomorphism is infinite-dimensional, the problem is ill-posed, hence we need to impose some regularization, for instance the $\mathbb{L}^2$ norm of its gradient:

$$\hat{w} \doteq \arg\min_w \ \Phi_{geom} + \beta\Phi_{reg} \doteq$$
$$\doteq \int_D \frac{1}{2}|\phi_1(w(x)) - \phi_2(x)|^2 + \frac{\beta}{2}|\nabla w(x)|^2 \, dx \tag{2}$$

where $\beta > 0$ is a tuning parameter and $|v|^2 \doteq v^T v$ denotes the squared two-norm of a vector. Of course more elaborate models and techniques can be employed, and the reader is referred to the literature, for instance [1]. However, this simple one suffices for us to introduce our model to combine geometric with intensity information. We review the sole-intensity model next.

2.3. Intensity-driven component cost

Correspondence based on intensity information is a field almost as broad as Computer Vision itself, so obviously no fair review of the literature can be provided in this venue. We will choose one of the simplest models that can serve our purpose, namely the $\mathbb{L}^2$ matching criterion that corresponds to restricting the classical Horn

and Shunk flow [11] to the regions of interest:

$$\hat{w} = \arg\min_{w} \Phi_{int} + \beta\Phi_{reg} \doteq$$
$$\doteq \int_{\Omega_2 \cap \mathcal{B}_\sigma(B_2)} \frac{1}{2}|I_1(w(x)) - I_2(x)|^2 + \frac{\beta}{2}|\nabla w(x)|^2 \, dx. \tag{3}$$

This model is rather restrictive, in that it does not allow intensity variations among the two images and assumes that they are simply diffeomorphically equivalent [10, 16]. It can be relaxed by allowing simple (global) changes in contrast and scaling, either via preprocessing, or by augmenting the model with additional parameters that can be inferred along with the warping $\hat{w}$. More generally, the cost function can be modified by allowing the intensities to be different, so long as the mutual information between the two images is maximized [15, 24, 29]. Of course other variations using other $\mathbb{L}^p$ norms [12, 28], total variation [26], Kullback-Liebler divergence [6], Bhattacharya distances between region histograms [7], or a myriad of different regularizers can also be employed. Also, the functional above can be made symmetric with respect to which image (in our case I_1) is mapped to which, by allowing two diffeomorphisms to warp both images to a common template [17].

2.4. Combined functional and bilateral region detection

As we have anticipated, our approach consists of using intensity information (or other intensity statistics, for instance gradient histograms) where salient regions are detected in both images. Following the model above, this simply translates into a functional of the form

$$\Phi(w) = \int_D \frac{1}{2}|I_1(w(x)) - I_2(x)|^2 \chi_{\Omega_2 \cap \mathcal{B}_\sigma(B_2)}(x) + \\ + \frac{\alpha}{2}|\phi_1(w(x)) - \phi_2(x)|^2 \chi_{\Omega_2^c \cup \mathcal{B}_\sigma^c(B_2)}(x) + \\ + \frac{\beta}{2}|\nabla w(x)|^2 \, dx \tag{4}$$

where $\chi_S(x)$ is the characteristic function of a set S, $\alpha > 0$ is a tuning multiplier, and the superscript c denotes the complement in D.

Now, the use of the characteristic functions above assumes that the bilateral regions B_j have been detected, and this is usually accomplished by a low-level vision algorithm. Like any other decision problem, this will involve selecting a threshold on some statistic of the image in the neighborhood of B_j, involving the probability that x belongs to it. For instance, one can compute likelihood ratios based on the gradient of the image, or better yet look for extrema of operators in scale-space [14]. Rather than assuming that this decision has been made for us, we will simply weight the geometric and intensity terms at a point x by the probability that such a point belongs to a "structure," using the same criterion that a low-level structure criterion would use. For the sake of illustration, we will use the normalized gradient of Gaussian scale-space of the image, following [14], that is equivalent to assuming $P(x \in B_2) = |\nabla_{norm} I_2(x)| \in [0, 1]$. Naturally, for B_2 to be a bilateral region, it will have to have a correspondence in image I_1, so it is not sufficient to evaluate the gradient at I_2, we must also evaluate it at I_1, warped via w, so that the criterion becomes $|\nabla_{norm} I_1(w(x))| \cdot |\nabla_{norm} I_2(x)|$. We will write this in terms of probabilities in the next subsection, and discuss how to extend it to more general discrepancy functions such as mutual information in Sect. 3; for now we just notice that the cost functional above becomes

$$\Phi(w) = \int_D \frac{1}{2}|I_1(w(x)) - I_2(x)|^2 |\nabla_{norm} I_1(w(x))| \cdot \\ \cdot |\nabla_{norm} I_2(x)| + \frac{\alpha}{2}|\phi_1(w(x)) - \phi_2(x)|^2 \cdot \\ \cdot (1 - |\nabla_{norm} I_1(w(x))| \cdot |\nabla_{norm} I_2(x)|) + \frac{\beta}{2}|\nabla w(x)|^2 \, dx. \tag{5}$$

This functional only considers intensity where the normalized scale-space gradient is large *both in the target image and in the warped template*. This only happens on and around bilateral regions, to an extent that depends on the scale of such regions (see [14] for details on automatic scale selection). Where such conditions are not satisfied, the geometric term and the generic regularizer drive the energy. An added benefit is that, because we have assumed that the images have been masked so that the background is zero, we can simply perform the integral on D without restricting portions of it to Ω_2.

To minimize $\Phi(w)$, variational calculus yields the first variation (for simplicity we only consider $U \subset \Omega_1$)

$$\frac{\delta\Phi}{\delta w} = (I_1(w(x)) - I_2(x)) \nabla I_1(w(x)) \cdot \\ \cdot |\nabla_{norm} I_2(x)| + \alpha(\phi_1(w(x)) - \phi_2(x)) \nabla\phi_1(w(x)) \cdot \\ \cdot (1 - |\nabla_{norm} I_2(x)|) - \beta\nabla^2 w(x) \tag{6}$$

with ∇^2 the Laplacian operator. By gradient descent with backtracking line search [23] we obtain the associated Euler-Lagrange equations, parameterizing the de-

scent direction by an artificial time $t \geq 0$:

$$\frac{\partial w}{\partial t} = -\frac{\delta \Phi}{\delta w}. \tag{7}$$

The temporal and spatial gradients are approximated by finite difference methods.

Again, this model only uses the simplest intensity term, and the simplest geometric term. Our emphasis is in how to combine the two. One can conceive ways in which this approach can be extended to more complex functionals, an issue we discuss in Sect. 3.

2.5. Probabilistic interpretation

In formal terms, our goal can be stated as seeking the maximum a-posteriori probability of a diffemorphic warping, that is

$$\hat{w} \doteq \arg\inf_w \log p(w|I_1, I_2) = \\ = \arg\inf_w \log p(I_1, I_2|w)p(w). \tag{8}$$

The second term, $\log p(w)$, can be easily recognized, in the model (5), as the generic regularizer $\int_D \frac{1}{2}|\nabla w(x)|^2 dx$. 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 [18] could be employed, in principle, so what we need to compute is $p(I_1|I_2, w)$. Now, again in purely formal terms, we could represent the probability of matching bilateral regions as $P(B_1 \leftrightarrow B_2)$; then what we wish to compute is

$$p(I_1|I_2, w) = p(I_1|I_2, w, B_1 \leftrightarrow B_2)P(B_1 \leftrightarrow B_2) + \\ + p(I_1|I_2, w, U \leftrightarrow \emptyset)(1 - P(B_1 \leftrightarrow B_2)). \tag{9}$$

Now, this is just formal notation. The difficulty comes in when we try to write explicitly the probabilities above, because the condition $B_1 \leftrightarrow B_2$ is specific to each point $x \in D$, so again 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 Markov-Chain 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

$$\prod_{B_1 \leftrightarrow B_2} e^{-\frac{1}{2}|I_1(w(x)) - I_2(x)|^2} \prod_{U \leftrightarrow \emptyset} e^{-\frac{1}{2}|\phi_1(w(x)) - \phi_2(x)|^2} \tag{10}$$

which, modulo technicalities, should converge to (4), the "stiff version" of our functional (5).

3. Extensions

While the approach described in the previous section can be applied directly to simple synthetic phantoms such as those in Fig. 1, application to real multi-modal images requires a more powerful model.

Specifically, rather than $\mathbb{L}^2$, we use mutual information [21] between the deformed template $I_1 \circ w$ and the target I_2, denoted by

$$\Phi_{MI}(w) = \int \log \frac{p(I_1, I_2|w)}{p(I_1|w)p(I_2)} \, dP(I_1, I_2|w). \tag{11}$$

Estimation for the joint image intensity distribution $p(I_1, I_2|w)$ is carried on the region of overlap A of both images by using 2-D Parzen windowing with Gaussian kernel G_σ:

$$p(I_1(w(x)), I_2(x)|w) = \frac{1}{|A|} \int_A G_\sigma(I_1(w(x)) - \\ - I_1(w(y)), I_2(x) - I_2(y)) \, dy. \tag{12}$$

It is equivalent to convolving the joint intensity histogram with a discrete approximation of G_σ. The joint histograms of $I_1 \circ w$ and I_2 within their region of overlap are constructed by binning the corresponding intensity pairs $(I_1(w(x)), I_2(x))$, and the marginal histograms are obtained by integrating over rows and columns, respectively.

Substituting (12) into (11) and rearranging following [4, 9], yields the first variation of Φ_{MI}:

$$\frac{\delta \Phi_{MI}}{\delta w} = -\frac{1}{|A|} \cdot \\ \cdot \left[G_\sigma * \frac{\partial L_w^{I_1, I_2}}{\partial I_1} \right] (I_1(w(x)), I_2(x)) \nabla I_1(w(x)) \tag{13}$$

with $*$ the convolution operator, and $L_w^{I_1, I_2}$ given by

$$L_w^{I_1, I_2} = 1 + \log \frac{p(I_1, I_2|w)}{p(I_1|w)p(I_2)}. \tag{14}$$

In place of an $\mathbb{L}^2$ regularizer, we use a fluid model [5] where the deformation velocity $v(x, t)$ is governed by the simplified Navier-Stokes equations

$$\mu \nabla^2 v + (\mu + \lambda) \nabla (\nabla \cdot v) + f = 0 \tag{15}$$

with μ and λ viscosity constants. Here f is the force field which drives the warping w in the appropriate direction. It is derived from image information, usually set

to the first variation. In the combined model, we have

$$f(x, w(x)) = -\frac{\delta \Phi_{MI}}{\delta w} - \frac{\delta \Phi_{geom}}{\delta w} =$$
$$= \frac{1}{|A|} \left[G_\sigma * \frac{\partial L_w^{I_1, I_2}}{\partial I_1} \right] (I_1(w(x)), I_2(x)) \nabla I_1(w(x)) \cdot$$
$$\cdot P(x \in B_2) - \alpha * (\phi_1(w(x)) - \phi_2(x)) \nabla \phi_1(w(x)) \cdot$$
$$\cdot (1 - P(x \in B_2)). \quad (16)$$

The deformation $u(x) = x - w(x)$ is successively accumulated through

$$\frac{\partial u}{\partial t} = v - (v \cdot \nabla) u. \quad (17)$$

This extension is rather straightforward, the only significant change being the expression of the weight $P(x \in B_2)$. Since mutual information between two images is related to pixel locations through intensity distributions, we need to not only consider the intensity gradients, but also the spatial gradients of intensity distributions. A direct extension of the argument for (5) leads to multiplying the target intensity gradients by the spatial gradient of a Parzen window-based joint density estimator from the sample histograms of the two images, i.e.,

$$P(x \in B_2) = |\nabla_{norm} I_2(x)| \cdot$$
$$\cdot \left| \left[G_\sigma * \frac{\partial L_w^{I_1, I_2}}{\partial I_2} \right]_{norm} (I_1(w(x)), I_2(x)) \right|. \quad (18)$$

4. Experiments

In this section we report a subset of the experiments we have conducted to validate the model proposed. As discussed in Sect. 2.1, ground-truth cannot be established for correspondence transfer since *there is no correspondence for unilateral regions*, so validation is performed subjectively by domain experts, and ultimate performance will hinge on how our approach is integrated with shape priors and other high-level information.

Fig. 2 shows the results of applying our approach to the simple synthetic phantoms shown in Fig. 1. It is patent that our approach has the desirable property of not making the unilateral regions disappear, and at the same time of properly deforming bilateral regions.

In Fig. 3 we illustrate our approach on real data. In order to do so, we must use a more elaborate discrepancy measure, as discussed in Sect. 3. In the top row we show the results obtained using only an intensity term where discrepancy is measured using mutual information. As it can be seen, the deformation grid is rather

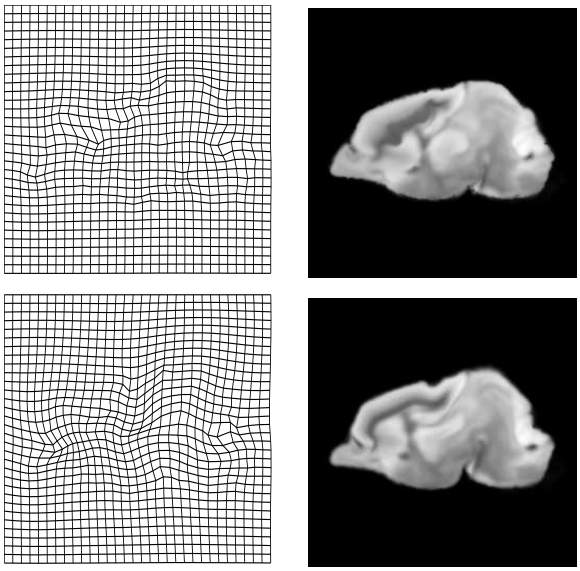

Figure 3. *Comparison with Mutual Information: (top) warping and registration using mutual information, and (bottom) using our approach. Notice that the dark region in the template is mapped in a highly irregular fashion to the target, and the its geometric structure is not preserved. Our combined model, displayed in the bottom part of the image, shows a more plausible deformation field, with unilateral regions being smoothly mapped into the anatomical template, and bilateral regions correctly deformed.*

irregular, and in particular the dark region of the left becomes "turbulent" in a way that is not compatible with high-level knowledge of anatomy. Our approach (bottom), on the other hand, shows that unilateral regions are mapped smoothly while bilateral features are deformed consistently. Additional examples on different test data are reported in Fig. 4.

In Fig. 5 we show some representative experiments where our approach fails to yield a meaningful correspondence. As it can be seen, the template is deformed locally, so the fold visible in the target is not correctly mapped. This kind of behavior is to be expected since our algorithm is based on a data discrepancy term alone, which encodes bottom-up, low-level information and is oblivious to any knowledge of the anatomy or physics of the underlying structures. This can be obviated by taking high-level prior information into account, an issue that is beyond our scope in this paper, where considerable efforts are undergoing in the medical imaging community.

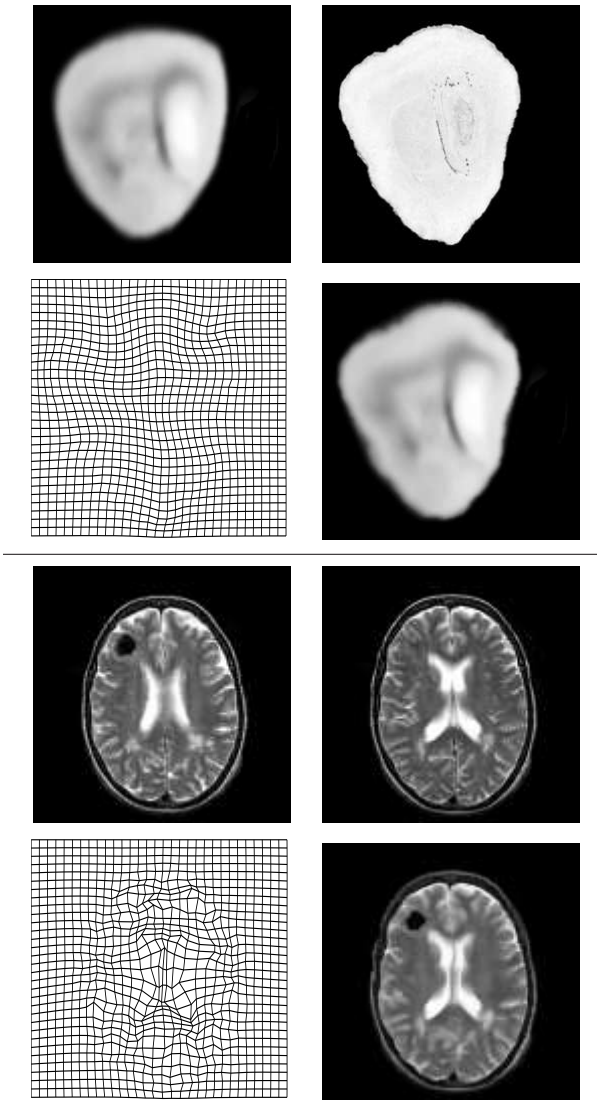

Figure 4. *Some more representative examples where our combined model yields subjectively successful correspondence of unilateral regions. The template (top-left), and target (top-right) are mapped one onto the other by a deformation field (bottom-left) that yields the deformed template (bottom right).*

4.1. Validation and discussion

As we have discussed in Sect. 2.1, no ground truth can be available for the problem we address. That is, unless high-level prior knowledge is brought to bear that our algorithm does not exploit. Ultimately, our algorithm will have to be complemented with shape priors, similarly to what done by [13]. For now, however, we limit ourselves to subjective evaluation of our results by expert anatomists. At this stage, cross-validation and other statistical tests cannot be performed because there is no secondary task (*e.g.* classification) for which we could have ground truth. It is possible that these will become available in the future (*e.g.* post-mortem studies of each individual subject), but that would not be practical. Our hope is that our model, integrated with suitable shape priors, can help science by automating multi-modal registration tasks. Our experiments show that including an explicit model of the correspondence hypothesis for bilateral versus unilateral regions there are improvements over both traditional intensity-based registration, as well as over mutual information-based approaches.

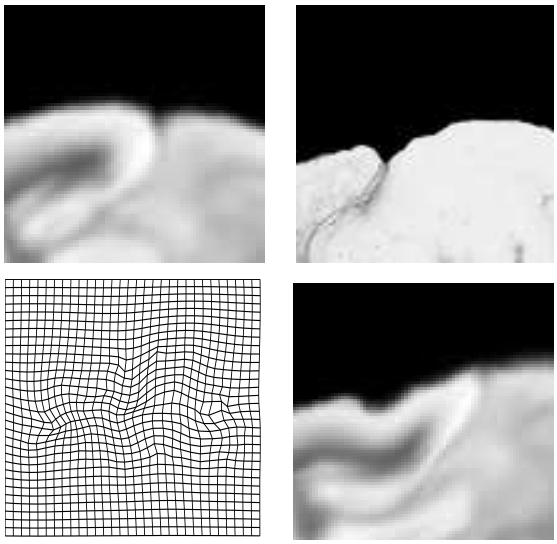

Figure 5. *Example of limitations of our model. Fine-scale geometric features (*e.g. *the small gap on the top right portion of the atlas) are not mapped correctly because our model only uses low-level information and is not cognizant of anatomical structure and constraints. It is clear that prior knowledge of the geometry of the underlying anatomy has to be enforced to achieve the reliability and precision of human experts.*

Acknowledgements

This work was supported by AFOSR FA9550-06-1-0138. We thank Gregorio Guidi and Andrea Vedaldi for their comments and suggestions, and Erh-Fang Lee and Ivo Dinov for providing data and feedback.

References

[1] M. A. Audette, F. P. Ferrie, and T. M. Peters. An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 4:201–217, 2000. 3

[2] M. F. Beg, M. I. Miller, A. Trouve, and L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *IJCV*, 61:139–157, 2005. 2

[3] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. Simultaneous structure and texture image inpainting. *IEEE Trans. Image Processing*, 12:882– 889, 2003. 2, 3

[4] C. Chefd'Hotel, G. Hermosillo, and O. Faugeras. A variational approach to multi-modal image matching. In *IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 21–28, 2001. 5

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

[6] A. C. S. Chung, W. M. Wells III, A. Norbash, and W. E. L. Grimson. Multi-modal image registration by minimising kullback-leibler distance. In *MICCAI*, pages 564–571, 2002. 4

[7] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *PAMI*, 24(5):603–619, 2002. 4

[8] C. M. Cyr, A. F. Kamal, T. B. Sebastian, and B. B. Kimia. 2d-3d registration based on shape matching. In *Mathematical Methods in Biomedical Image Analysis*, pages 198–203, 2000. 1

[9] 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. 5

[10] U. Grenander, A. Srivastava, and S. Saini. Characterization of biological growth using iterated diffeomorphisms. *IEEE Int. Symp. Biomedical Imaging: Nano to Macro*, pages 1136–1139, 2006. 2, 4

[11] B. K. Horn. *Robot Vision*. MIT Press, 1986. 4

[12] D. P. Huttenlocher, G. A. Klanderman, and W. A. Rucklidge. Comparing images using the hausdorff distance. *PAMI*, 15:850–856, 1992. 4

[13] M. E. Leventon, W. E. L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *CVPR*, volume 1, pages 316–323, 2000. 7

[14] T. Lindeberg. Principles for automatic scale selection. Technical report, Royal Institute of Tehnology, Stockholm, Computational Vision and Active Perception laboratory, 1998. 4

[15] 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. 1, 4

[16] M. I. Miller. Computational anatomy: shape, growth, and atrophy comparison via diffeomorphisms. *NeuroImage*, 23:S19–S33, 2004. 1, 2, 4

[17] M. I. Miller and L. Younes. Group actions, homeomorphisms and matching: a general framework. *IJCV*, 41:61–84, 2001. 4

[18] D. Mumford and B. Gidas. Stochastic models for generic images. *Communications in Pure and Applied Mathematics*, 54(1):85–111, 2001. 5

[19] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on the hamilton-jacobi formulation. *Computational Physics*, 79:12–49, 1988. 3

[20] N. Paragios, M. Rousson, and V. Ramesh. Matching distance fuctions: a shape-to-area variational approach for global-to-local registration. In *ECCV*, pages 775–789, 2002. 1

[21] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viegever. Image registration by maximization of combined mutual information and gradient information. *IEEE Trans. Medical Imaging*, 19:809–814, 2000. 5

[22] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viegever. Mutual information based registration of medical images: a survey. *IEEE Trans. Medical Imaging*, 22:986–1004, 2003. 1

[23] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1988. 4

[24] A. Rangarajan, H. Chui, and J. S. Duncan. Rigid point feature registration using mutual information. *Medical Image Analysis*, 3:425–440, 1999. 1, 4

[25] M. Rousson and N. Paragios. Shape priors for level set representations. In *ECCV*, pages 78–92, 2002. 1

[26] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica*, 60:259–268, 1992. 1, 4

[27] R. Stefanescu, O. Commowick, G. Malandain, P.-Y. Bondiau, N. Ayache, and X. Pennec. Non-rigid atlas to subject registration with pathologies for conformal radiotherapy. In *MICCAI*, pages 704–711, 2004. 1

[28] B. C. Vemuri, J. Ye, Y. Chen, and C. M. Leonard. Image registration via level-set motion: Applications to atlas-based segmentation. *Medical Image Analysis*, 7:1–20, 2003. 4

[29] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *IJCV*, 24:137–154, 1997. 4

[30] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000, 2003. 1