

# Reliability-Driven, Spatially-Adaptive Regularization for Deformable Registration

Lisa Tang<sup>1</sup>, Ghassan Hamarneh<sup>1</sup>, and Rafeef Abugharbieh<sup>2</sup>

<sup>1</sup> Medical Image Analysis Lab., School Computing Science, Simon Fraser University  
{hamarneh,lisat}@cs.sfu.ca

<sup>2</sup> Biomedical Signal and Image Computing Lab.,  
Department of Electrical and Computer Engineering, University of British Columbia  
rafeef@ece.ubc.ca

**Abstract.** We propose a reliability measure that identifies informative image cues useful for registration, and present a novel, data-driven approach to spatially adapt regularization to the local image content via use of the proposed measure. We illustrate the generality of this adaptive regularization approach within a powerful discrete optimization framework and present various ways to construct a spatially varying regularization weight based on the proposed measure. We evaluate our approach within the registration process using synthetic experiments and demonstrate its utility in real applications. As our results demonstrate, our approach yielded higher registration accuracy than non-adaptive approaches and the proposed reliability measure performed robustly even in the presences of noise and intensity inhomogeneity.

## 1 Introduction

The goal of deformable image registration is to recover a local transformation,  $T$ , that best aligns two images. Generally, the problem involves minimizing a weighted sum of two penalty terms [1,2,3,4], e.g.:

$$\hat{T} = \arg \min_T \mathcal{D}(F, T \circ M) + \alpha \mathcal{R}(T) \quad (1)$$

where  $\mathcal{D}$  denotes a data term, which measures how well  $T$  aligns two images,  $F$  and  $M$ ,  $\mathcal{R}$  denotes a regularization term that ensures  $T$  maintains certain smoothness properties (e.g. being continuous or homeomorphic), and  $\alpha$  is a weight that balances these two terms.

Literature in medical image registration has generally focused on the development of either image metrics [5], regularization models [6,7], or optimization algorithms [1,3], the three main ingredients of a registration method. In this paper, we attempt to address all three aspects of the problem, with a particular focus on adapting the regularization of deformations to local image content.

Commonly known as *inhomogeneous deformability* [6], deformable registration incorporating adaptive regularization allows one to obtain smooth deformations in some parts of an anatomy and highly varying ones in other regions. Inhomogeneous deformability is particularly useful when we wish to model the various

types of motions different tissues undergo, so that, for example, soft tissues can deform elastically while hard bones are restricted to move only rigidly. Most researchers to date achieve adaptive regularization by incorporating anatomical models. An example is the variable-elasticity registration algorithm of Davatzikos [7], where anatomical segmentations were used in an elastic registration algorithm to adjust the amount of regularization according to tissue types. Regions that were labeled as soft tissues (e.g. skin and fat) were regularized less than regions that were identified as hard tissues or rigid structures (e.g. bone). Similar works include [8], where Rexilius et al. derived the inhomogeneous elasticity parameters probabilistically using a statistical atlas, and [9], where Kabus et al. performed elastic registrations and suppressed smoothness constraints at the interfaces between identified structures via a spatially-dependent weight. In [10], Pitiot and Guimond developed a geometrically-driven regularization for a block-matching algorithm where the regularization of displacement vectors were localized to regions that were fitted to the geometry of the anatomy.

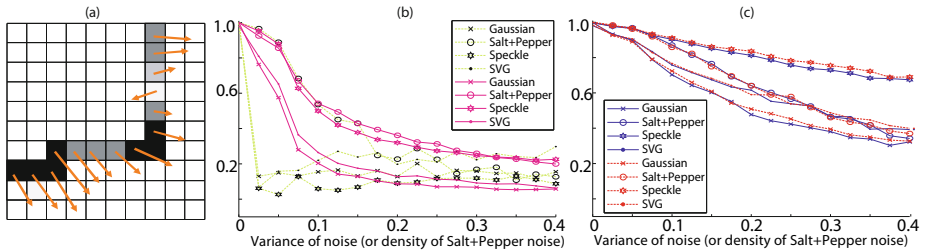
Several adaptive regularization methods have also been incorporated into registration frameworks that formulate image matching and regularization as separate processes, as opposed to minimizing the combined data and regularization energies directly. In [11], Stefanescu et al. incorporated an adaptive regularization approach into a compositional demons algorithm, which iteratively composed a correction field and a tentative displacement field. They proposed to regularize the deformation field by smoothing it with a variable Gaussian kernel whose size depended on a scalar field, which encoded the expected amount of deformation and was estimated from a region-based segmentation of the anatomy. The authors also proposed to filter the correction field using a measure called *local confidence* to estimate local intensity variance. In the template-matching framework of [12], after each iteration of deformation estimation, Suarez et al. smoothed the estimated field with a variable Gaussian kernel whose size was determined by a scalar measure of local structure.

In summary, adaptive regularization as formulated in all of the methods highlighted above have some inherent problems. The methods of [6,7,8,9,10] all required prior information (e.g. manual segmentation), making them impractical, especially when anatomical models are laborious to prepare, and difficult to define in cases containing pathologies. On the other hand, the data-driven regularization approaches of [11,12], where regularization depended on image gradients or local structures, suffered from sensitivity to noise. Unfortunately, the reliability and effectiveness of regularization become questionable when regularization depended on such noise-sensitive functions.

To address these issues, we propose a data-driven, spatially-adaptive regularization approach that is robust to noise and does not require prior information (e.g. segmentation or knowledge of material properties). As image forces computed in uncertain regions (e.g. suffering from high noise level, or bounded by weak or missing boundaries) should play a smaller role in estimating the transformation solution, we propose a robust *reliability* measure that analyzes local noise levels and image structures from which we derive data-driven, adaptive

regularization. In our reliability-based regularization, the amount of regularization is increased in regions that have high noise level, or decreased otherwise. Similarly, data-derived forces computed in highly structured and uncorrupted regions should have low regularization because they are derived from reliable and discriminative information. Fig. 1a provides an illustration of these concepts. By accounting for local noise levels when measuring image cues, our proposed reliability measure can better distinguish reliable regions from unreliable, noise-corrupted regions more robustly than previously proposed measures (Fig. 1b-c).

To the best of our knowledge, this is the first work in the context of registration that examines reliability of image data as a means to balance the data and regularization terms in an image-dependent manner. We are also the first to incorporate spatially varying regularization into two Markov Random Field (MRF)-based registration frameworks. Formulating deformable registration as MRF-energy minimization has become a growing trend due to the recent developments of efficient solvers [13,14,15,16]. With these solvers, volumetric, multi-resolution registration can be achieved in minutes, a significant speedup over previous methods [17]. While efficient, the regularization adopted in these frameworks had always been controlled globally (across the whole image) by a scalar weight [3,2,1]. Determination of this weight also required empirical experiments, which were often done through a rather ad-hoc and tedious process. For that reason, as another part of our major contribution, we will illustrate how our



**Fig. 1.** (a) An illustrative example of a deformation field obtained from a non-parametric registration. Displacement vectors computed at the boundary of an object are shown as arrows. Gray values at each pixel along this boundary indicate the measured reliability (dark gray indicates high reliability). With our proposed reliability measure, the outlier vector (one pointing to left) will be regularized more than the rest, while those that are reliable and located along salient structures (e.g. corners) will be regularized relatively less, subject to the influence of the data term. (b) and (c) a measure’s sensitivity to noise, which was defined as the correlation between the measure computed before and after noise corruption. Each curve shows the effect of one noise type: Gaussian, speckle, Salt+Pepper, and spatially varying Gaussian (SVG). Sensitivity of (b) local confidence [11] (magenta) and local structure [12] (black), and (c) our reliability measure. In (c), two different parameter settings (discussed in Sec. 2.2) were used to evaluate the performance of our measure. Under all noise types, local confidence and local structure were more sensitive to noise than our proposed measure.

reliability-based regularization can be easily incorporated into these frameworks to improve their accuracies.

In the past, few researchers have proposed the use of measures of certainty or saliency of image data to improve registration, e.g. [5]. However, these measures were solely used to improve the fidelity of the data term and not for regularization, as we propose in this work. In dealing with noisy images, Paquin et al. [18] proposed multi-scale decomposition of images to iteratively register the obtained decomposed components. Their method, however, did not employ adaptive regularization either. Accordingly, our contributions are: 1) adapting image regularization according to local image information using an image-derived measure, as opposed to use of prior information; 2) improving registration robustness by enhancing the fidelity of the data term using this measure; 3) introducing the novel use of a spatially varying weight to a discrete optimization-based registration framework; 4) proposing and examining various techniques to encode the proposed reliability-based regularization; and 5) validating the overall method with synthetic and real data.

## 2 Methods

### 2.1 Deformable Registration via MRF-Minimization

Let  $F$  and  $M$  be discrete representations of a fixed and moving image, respectively, in a domain  $\Omega \subset \mathbb{R}^d$ , where  $d$  is the image dimension, i.e.  $F : \Omega \rightarrow \mathbb{R}$ ,  $M : \Omega \rightarrow \mathbb{R}$ . Our goal is to recover a displacement field  $T$ ,  $T : \Omega \rightarrow \mathbb{R}^d$ , that maps each pixel location  $\mathbf{x} = (x^1, x^2, \dots, x^d)$  in  $M$  to  $F$  by minimizing the energy in Eq. 1. To formulate deformable registration as an MRF optimization, the pixel coordinates of  $M$  are usually converted to a graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where vertex  $p \in \mathcal{V}$  represents a spatial coordinate  $\mathbf{x}_p$  and the set of edges  $(p, q) \in \mathcal{E}$  describe a 4-neighbourhood system of the image grid of  $M$ . Next, the deformation space (e.g.  $\mathbb{R}^2$  in 2D) is discretized into a finite set of translations  $\mathcal{L}$  of size  $L$ , where each element represents a translation vector  $t_i$ , i.e.  $\mathcal{L} = \{t_1, \dots, t_L\}$ . We then seek to label each  $p$  with a label in  $\mathcal{L}$  to obtain  $T$  that minimizes:

$$E(T) = \sum_{p \in \mathcal{V}} \psi_i(p, t_i) + \alpha \sum_{(p,q) \in \mathcal{E}; i,j \leq L} \psi_{ij}(p, q, t_i, t_j) \quad (2)$$

where  $\psi_i$  denotes the cost of assigning  $t_i$  to  $p$ ,  $\psi_{ij}$  denotes the cost of assigning  $t_i$  to  $p$  and  $t_j$  to its neighbour  $q$ , and  $\alpha$  denotes a global weight between the two terms. Essentially,  $\psi_i$  and  $\psi_{ij}$  correspond to the data term  $\mathcal{D}$  and regularization term  $\mathcal{R}$  in Eq. 1, respectively. This energy can then be minimized via combinatorial optimization algorithms of [13,14,16].

In this work, we propose the use of spatially adaptive regularization by replacing  $\alpha$  with a function that is spatially dependent on pairwise neighbours, i.e.  $\lambda(p, q)$ . Thus, if the data term is, for example, based on the absolute difference (AD) between image intensities and  $\psi_{ij}$  is a distance-based metric, the energy minimization becomes:

$$\hat{T} = \arg \min_T \sum_{p \in \mathcal{V}} \frac{1}{I_{max}} |F(p + t_i(p)) - M(p)| + \sum_{(p,q) \in \mathcal{E}} \frac{\lambda(p,q)}{4 \times 2 d_{max}} \|t_i(p) - t_j(q)\| \quad (3)$$

where  $I_{max}$  denotes the maximum intensity difference between  $F$  and  $M$ ,  $t_i(p)$  denotes translating  $\mathbf{x}_p$  with the  $i$ -th label in  $\mathcal{L}$  and  $d_{max}$  denotes the maximum displacement allowed (as set by the user)<sup>1</sup>.

## 2.2 Adaptive Regularization via Measure of Image Reliability

As motivated earlier, regularization of  $T$  should be adapted according to the local image content of  $M$ . Here, we advocate adaptive regularization based on local noise levels and image structures. When images are corrupted with spatially varying noise levels, the amount of regularization should be increased in noise-corrupted regions so that the local registration of these regions can be better driven by their more reliable neighbours. Regularization should also be decreased in regions with high signal-to-noise ratio and meaningful local structures because their local content has sufficient discriminatory information to provide reliable motion estimates.

To adaptively regularize  $T$ , we propose a *reliability measure*,  $R(\mathbf{x})$ , that analyzes two types of cues of an image  $I$ : data fidelity and local image structure. Here, we define the fidelity of a pixel as a function of noise levels and edge evidence. The local noise levels,  $N(\mathbf{x})$ , is estimated using the spectral flatness measure defined in [19]. To estimate local edge strength,  $G(\mathbf{x})$ , we compute the image gradient and employ the noise-gating strategy of [19] to dampen strong responses that might have been provoked by noise, yielding:

$$G_{gated}(\mathbf{x}) = |\nabla I(\mathbf{x})|(1 - N(\mathbf{x}))^{\alpha_G} \quad (4)$$

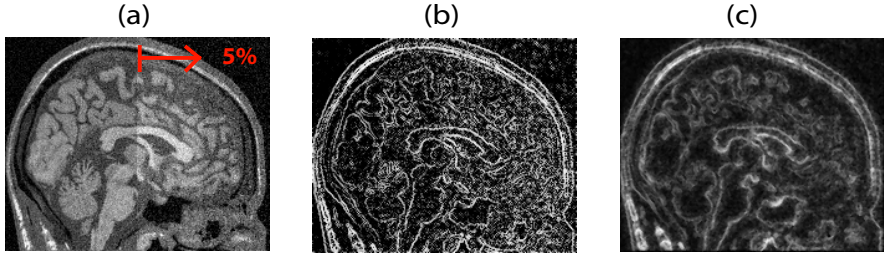
where the scalar  $\alpha_G$  controls the level of noise-gating on the image gradient.

In estimating the local image structures,  $\kappa(\mathbf{x})$ , we first smooth  $I$  using a Gaussian kernel of size  $\sigma$ . The local curvature of the smoothed image, denoted as  $I_\sigma$ , is then computed in scale space as shown in [20]:

$$\kappa(\mathbf{x}; \sigma) = (I_{y,\sigma}^2 I_{xx,\sigma} - 2I_{x,\sigma} I_{y,\sigma} I_{xy,\sigma} + I_{x,\sigma}^2 I_{yy,\sigma}) (\sqrt{I_{x,\sigma}^2 + I_{y,\sigma}^2})^{-1} \quad (5)$$

where  $I_{x,\sigma}$  and  $I_{y,\sigma}$  denote the image derivatives of  $I_\sigma$  along  $x$  and  $y$ , respectively, and  $I_{yy,\sigma}$  denote the second-order image derivatives of  $I_\sigma$  along  $y$ , etc. We then use the normalized scale coordinates of [21] to address the scale-selection problem. Specifically, to compare curvature values across different scales, we scale-normalize  $\kappa(\mathbf{x}; \sigma)$ , yielding  $\kappa_n(\mathbf{x}) = \max_\sigma \sigma^3 \kappa(\mathbf{x}; \sigma)$ . Finally, to weaken the strong responses at non-structural and noisy regions, we also apply noise-gating, yielding:

<sup>1</sup> Since each node of a 4-neighbourhood grid system contributes to  $\psi_{ij}$  four times and contributes to  $\psi_i$  once, the constant 4 is used to equalize the contribution of  $\psi_{ij}$  and  $\psi_i$  in the total energy cost.  $2d_{max}$  is the maximum difference between two assigned displacement vectors.



**Fig. 2.** (a) An image from BrainWeb [23] corrupted with spatially varying Gaussian white noise, 2% and 5% noise level on left and right side of image, respectively. (b) Its local confidence measure [11], and (c) our proposed reliability measure, which has successfully identified important image cues (e.g. local curvatures) despite of noise-corruption. Conversely, (b) assigned higher emphasis to the local image gradients located on the more noise-corrupted, right side of the image.

$$\kappa_{gated}(\mathbf{x}) = \kappa_n(\mathbf{x})(1 - N(\mathbf{x}))^{\alpha_C} \quad (6)$$

where the scalar  $\alpha_C$  controls the level of noise-gating on the local curvatures.

With  $N(\mathbf{x})$ ,  $G_{gated}(\mathbf{x})$ , and  $\kappa_{gated}(\mathbf{x})$  estimated, the reliability measure  $R$  at  $\mathbf{x}$  is then computed as:

$$R(\mathbf{x}) = G_{gated}(\mathbf{x})^{(1 - \kappa_{gated}(\mathbf{x}))} \quad (7)$$

where the exponential term is adopted as a cue-gating strategy as proposed in [22] to suppress gradient information in highly textured regions.  $R$  is subsequently normalized to a range of [0,1]. Fig. 2c shows the reliability map computed from an image corrupted with spatially varying noise. Observing Fig. 2c, we note that despite of data corruption, the change in  $R$  is small, indicating that it is relatively insensitivity to noise. In general, as presented in Fig. 2b-c,  $R$  is much more robust to noise than some of the measures proposed in the literature, e.g. [11]. Details of how parameters of  $R$  were chosen will be discussed in Sec. 2.6.

We next present how the proposed reliability measure is used in MRF-based registration, augmented with adaptive regularization. For this task, we compute  $R$  over  $F$  and  $M$  to create reliability maps  $R_F$  and  $R_M$ , respectively.

### 2.3 Reliability Encoded as Edge-Weights for Regularization

In the context of MRF-based registration,  $\lambda(p, q)$  is the regularization weight assigned to two neighbouring pixels  $p$  and  $q$ . The higher the value of  $\lambda(p, q)$ , the higher the required coherence between the deformations at  $p$  and  $q$ . To incorporate adaptive regularization into such a framework, we propose to determine  $\lambda(p, q)$  based on  $R_M(\mathbf{x}_p)$  and  $R_M(\mathbf{x}_q)$ . To this end, three encoding schemes have been conceived for determining  $\lambda(p, q)$  from  $R_M$ .

**1. Continuous scheme (CONT).** This scheme enforces minimal regularization if both  $\mathbf{x}_p$  and  $\mathbf{x}_q$  have the highest reliability scores, or enforces maximal

regularization otherwise. It does so by defining the weight function as:  $\lambda(p, q) = \exp(-R_M(\mathbf{x}_p)R_M(\mathbf{x}_q))$ .

**2. Clustered scheme (CLUST).** This scheme follows the ideas presented in [24] (for matching of stereo images) where regularization weights are quantized. It computes  $\lambda(p, q)$  as in CONT but further clusters the set of weights into  $K$  values using K-means. As shown in our experiments, the smaller the size of the set of all possible edge weights, the faster the convergence of MRF-optimization due to fewer number of ‘graph reparameterizations’ [16] needed to minimize the MRF energy.

**3. Discrete scheme (DISCR).** Based on CLUST, DISC incorporates the use of intensity cues for adapting  $\lambda$  (referred to as ‘contextual information’ in the segmentation work of [16]) and assumes that a change in intensity values indicates presence of boundary between different tissues types. It examines the noise-gated edge strength  $G(\mathbf{x})$  given in Sec. 2.2 and assigns  $\lambda(p, q)$  to one of four discrete values  $w = \{w_1, w_2, w_3, w_4\}$ , where  $w_1 < w_2 < w_3 < w_4$ , as similarly done in [16]. Specifically, if both  $p$  and  $q$  are identified as reliable pixels, as determined by a threshold  $\tau_{rely}$ , and they both have high edge strengths, i.e.  $G(\mathbf{x}_p)$  and  $G(\mathbf{x}_q)$  are both greater than a threshold  $\tau_{edge}$ , then we assume that  $p$  and  $q$  belong to different tissue types and assign the lowest regularization weight possible, i.e.  $\lambda(p, q) = w_1$ . If both pixels are reliable, but their edge strengths are lower than  $\tau_{edge}$ , then we set  $\lambda(p, q) = w_4$ , hence assuming that they belong to same tissue type. However, if one of the two pixels or both pixels are unreliable, we assign intermediate weights, depending on their noise-gated local edge strengths: we set  $\lambda(p, q) = w_3$  if their edge strengths are greater than  $\tau_{edge}$ , or  $\lambda(p, q) = w_2$  otherwise. Finally, to minimize the number of free variables introduced in this scheme, we parameterize  $w$  with the variable  $\mu$  and set  $w = \{w_1 = \mu, w_2 = 2\mu, w_3 = 3\mu, w_4 = 4\mu\}$ .

## 2.4 Truncation of the Unary Term Based on Reliability (DTrunc)

Since the reliability measure readily reflects the quality of local image content, we can improve the fidelity of the data term by lowering the influence of the unary costs of all unreliable pixels via truncation [16]. If the pixel values  $M(\mathbf{x}_p)$  and  $F(\mathbf{x}_p + t_i(\mathbf{x}_p))$  are reliable, then the unary cost computed at the corresponding locations are also reliable. If either of the pixel values is considered unreliable, then we can enforce higher regularization at  $p$  by modifying the corresponding unary costs of  $p$ . In other words, if  $R_M(\mathbf{x}_p)$  and  $R_F(t_i(\mathbf{x}_p) + \mathbf{x}_p)$  are both above  $\tau_{rely}$ , then we leave  $\psi_i(t_i, p)$  unmodified; otherwise, we assign  $\psi_i(t_i, p) = \eta$ . We shall denote this truncation strategy as DTrunc.

## 2.5 Implementation Details

To illustrate the generality of our adaptive regularization approach, we implemented two versions of MRF-based deformable registration with spatially adaptive regularization to compute  $\hat{T}$ . The first, denoted as DENSE, follows the

construction of Tang and Chung in [2] and computes  $\hat{T}$  explicitly as described in Sec. 2.1. The second, denoted as BSPLINE, employs a B-Spline transform model to compute  $\hat{T}$  as done in [1,3]. BSPLINE registration minimizes a similar energy function<sup>2</sup> as DENSE, but  $\hat{T}$  is produced from a B-Spline transform grid. Thus, in 2D, each  $p$  represents a control point in a grid of size  $C_x \times C_y$ , with grid spacing  $\delta$ , and we seek to label the set of control points with a label in  $\mathcal{L}$  (Sec. 2.1). Displacement at  $\mathbf{x}$  is then computed using a cubic B-Spline model as similarly done in [1,3].

Both DENSE and BSPLINE were implemented in MATLAB R2009b (Mathworks Natwick, MA) and employed the FastPD software [14]. We used DENSE if a registration trial involved  $d_{max} \leq 8$  pixels or BSPLINE otherwise, although the two algorithms may be used sequentially (employ BSPLINE to pre-align images and refine the alignment using DENSE). To illustrate the robustness of our proposed regularization method to noise and to intensity inhomogeneity, we derive the data term from the absolute difference image similarity metric (Sec. 3), which normally fails under noise and intensity inhomogeneity corruptions.

## 2.6 Summary of Parameters

The parameters involved in our algorithm are  $\tau_{edge}$ ,  $\tau_{rely}$ ,  $\mu$ ,  $K$ ,  $\eta$ ,  $\alpha_G$ ,  $\alpha_C$ ,  $d_{max}$ ,  $d_{res}$ , and  $\delta$ . We emphasize that most of these parameters are only related to the efficiency of MRF optimization and does not relate to our proposed regularization approach. For DISCR, based on emperical evaluation, we chose  $\mu = \{0.8, 1.5\}$  and computed  $\tau_{rely}$  as the 25<sup>th</sup> percentile of  $R$ ,  $\tau_{edge}$  as the 75<sup>th</sup> percentile of  $G$ , and  $\eta$  to be the mean of all reliable unary costs. Preliminary results did not indicate sensitivity to these values, thanks to the robustness of  $R$ . This is demonstrated in 1c which shows that the sensitivity of  $R$  to different noise corruptions is minimal (the blue curves in the figure represent results using parameters  $\alpha_G = .8$  and  $\alpha_C = .4$  and those in red represent results using  $\alpha_G = .4$  and  $\alpha_C = .1$ ). For CLUST, we set  $K = 10$  or if the clustering algorithm that we employed did not converge, we rerun the clustering algorithm with  $K$  increased by one. The value of  $d_{res}$  was restricted by  $d_{max}$  so that at most 250 labels were used ( $L \leq 250$  best balances between optimization time and the optimal accuracy of  $\hat{T}$ ) and  $d_{max}$  was set to half the diagonal length of the image.

## 3 Results

We performed validation of our method with both synthetic deformations and real clinical data. Our test data comprised of a pair of magnetic resonance imaging (MRI) brain slices from BrainWeb [23], and a set of 18 sagittal brain slices from the Internet Brain Segmentation Repository (IBSR)<sup>3</sup>.

<sup>2</sup> The data term is modified slightly, rather than computing a similarity metric at each point, we compute the metric over a  $3 \times 3$  neighbourhood of each control point.  $R(\mathbf{x}_p)$  is computed as a distance-weighted sum of  $R$  over the neighborhood of  $\mathbf{x}_p$ .

<sup>3</sup> <http://www.cma.mgh.harvard.edu/ibsr/>



**Table 1. Synthetic experiments involving random TPS-warped BrainWeb images.** Shown are registration accuracies obtained by different schemes, under various types of noise corruption and intensity inhomogeneity (IIH). Accuracy is computed as the average MED obtained over all trials. U denotes best uniform regularization. Bolded numbers indicate better performance. Note how the proposed data term truncation strategy improved registration results in 8 out of 12 cases, irrespective of noise type and level of intensity inhomogeneity. In general, results obtained with adaptive regularization had higher accuracies than those obtained with uniform regularization.

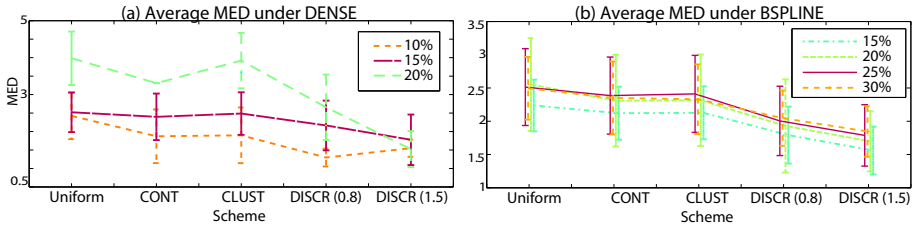
IIH	Under random speckle noise				Under random Gaussian noise			
	U	U+DTrunc	DISCR	DISCR+DTrunc	U	U+DTrunc	DISCR	DISCR+DTrunc
6%	4.69	<b>3.93</b>	2.85	<b>2.76</b>	4.23	<b>3.81</b>	2.85	<b>2.39</b>
18%	3.49	<b>3.43</b>	2.70	2.96	3.61	<b>3.20</b>	2.55	<b>2.49</b>
30%	<b>4.89</b>	4.99	<b>2.75</b>	2.96	<b>4.64</b>	4.78	3.37	<b>2.60</b>

We first studied the effects of employing DTrunc, the procedure described in Sec. 2.4. Following the approach of [25], we simulated groundtruth data by applying random thin-plate-spline (TPS) warps, denoted as  $T_{GT}$ , to an image from BrainWeb, with  $T_{GT}$  containing maximum displacement of 5 pixels in each dimension, yielding 20 sets of  $F$ ,  $M$  and  $T_{GT}$  (10 such trials generated for each BrainWeb image). We then corrupted the images by adding intensity inhomogeneity at different levels and added Gaussian or speckle noise (at 5%), as similarly done in [15,18]. We subsequently performed DENSE registration with and without DTrunc. We discretized the deformation space with  $d_{max} = 6$ ,  $d_{res} \geq 0.5$  in the x and y dimensions.

Registration accuracy of each registration trial was computed as the mean Euclidean distance (MED) of every pixel between  $T_{GT}$  and  $\hat{T}$ , the one recovered from registration. For all tests on uniform regularization, we repeated trials with different values of  $\alpha$  in  $\{0.05, 0.1, \dots, 0.95\}$  and the trial with the lowest MED was selected as the final result. The obtained results are summarized in Table 1, which indicates that the registration errors obtained with DTrunc are generally lower than those obtained without DTrunc.

We next compared the registration performance of our proposed reliability-based regularization using the encoding schemes outlined in Sec. 2.3. We generated a set of groundtruth data as described before, but the maximum displacement introduced by the TPS-warps was set to 8 pixels and the generated images were corrupted with spatially varying Gaussian noise (additive Gaussian noise patterns of different variances at random locations). We then performed DENSE registration to register the images with  $d_{max} = 10$  and  $d_{res} = 1.2$ .

Fig. 3a presents a quantitative comparison between our proposed schemes against uniform regularization. Results demonstrate that the use of the proposed reliability measure for adaptive regularization using DISCR can reduce MED by as much as 2.45 pixels. We also examined the accuracy of  $\hat{T}$  by examining the Euclidean distance between  $T_{GT}$  and  $\hat{T}$  on a per-pixel basis as shown in Fig. 4. Clearly, adaptive regularization (Fig. 4) allowed for much more accurate deformations than those generated with uniform regularization. As the proposed reliability measure was able to identify unreliable regions and the amount of regularization



**Fig. 3.** Results of registration under different levels of spatially varying noise applied to BrainWeb images. Plots show the average MED obtained when registration was incorporated with adaptive regularization. Uniform denotes registration without reliability-based adaptive regularization; CONT, DISCR, CLUST each denotes one of our proposed schemes as presented in Sec. 2.3. For DISCR, the number in brackets are the values of parameter  $\mu$ . Results involving (a) DENSE to recover TPS warps ( $d_{max} = 8$ ) and (b) BSPLINE to recover B-Spline warps ( $d_{max} = 12$ ). The differences in trend were due to differences in registration schemes, types of warp used, and discretization levels. Reliability-based regularization encoded with DISCR generally yielded low MED.

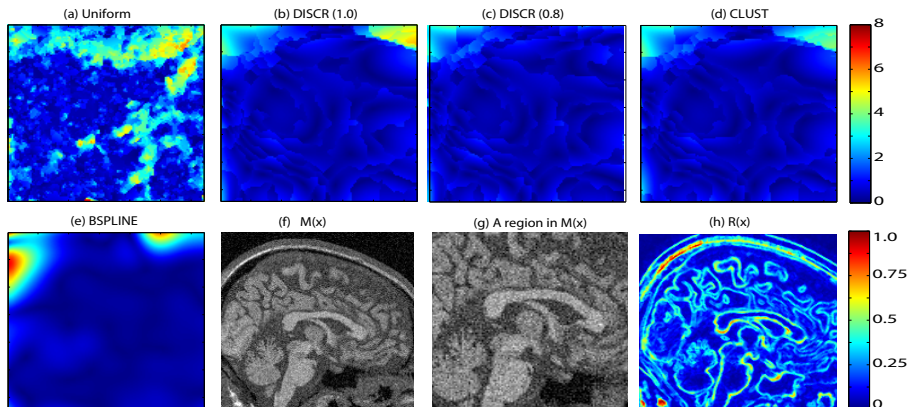
was increased in these regions accordingly, errors at these unreliable regions were relatively lower than those obtained with uniform regularization.

We next tested BSPLINE registration incorporated with adaptive regularization. We applied synthetic B-Spline warps containing maximum displacement of 12 pixels and evaluated results by comparing the known warps with the recovered ones. Results from this experiment are summarized in Fig. 3b, which shows that reliability-based regularization using DISCR had reduced MED by as much as 0.90 pixels.

Finally, we performed pairwise registrations on the clinical brain MR images from the IBSR dataset. Segmentations of these images are also available (each slice was segmented into 3 structures) so registration validation could be done with segmentation-based measures (as performed in [26]), which allows us to examine registration accuracy in relation to the anatomical structures of interest. In absence of pathologies, more accurate registration will result in better overlap of corresponding regions.

A total of 153 DENSE registrations (all 18 images with repetitions) were done per registration scheme, e.g. uniform regularization and reliability-based regularization encoded with CONT, CLUST, and DISCR (we selected  $\mu = 1.5$  based on evaluation in Fig. 3). For each registration, we applied  $\hat{T}$  to the label field of  $M$ , using nearest-neighbour interpolation and compared the warped segmentation to the label field of  $F$ . After registration, we computed target overlap (TO) and distance error (DE) as defined in [26]. The obtained results were as follows, which are reported in (TO,DE) pairs, with TO expressed in fraction and DE expressed in pixels: uniform regularization obtained (0.712,7.82), while reliability-based regularization with CONT obtained (0.758,6.46), CLUST obtained (0.766,6.89); and DISCR obtained (0.768,7.12).

Overall, results from synthetic experiments have demonstrated that the use of reliability-based, adaptive regularization can recover TPS and B-Spline warps more accurately than those recovered with non-adaptive approaches: up to 2.45 and 0.90 improvement in MED for BSPLINE and DENSE registration, respectively. This is true regardless of the severity of intensity inhomogeneity and noise types, e.g. (spatially varying) Gaussian and speckle. The reliability-based data term truncation strategy was also shown to improve registration results in general (adaptive and uniform regularization). Results from our segmentation-based evaluation also showed that our proposed method increased TO by 0.056 ( $> 5\%$ ).



**Fig. 4.** Results obtained from a DENSE registration trial. Sub-figures (a)-(e) show the Euclidean distance between  $T_{CT}(\mathbf{x})$  and  $\hat{T}(\mathbf{x})$ , which was obtained using one of the following registration schemes: (a) uniform regularization; reliability-based regularization encoded with (b) DISCR ( $\mu = 1.0$ ), (c) DISCR ( $\mu = 0.8$ ), and (d) CLUST ( $K = 10$ ). Although the input image (f) was corrupted with 5% Gaussian noise,  $R$  was capable of capturing important image cues as shown in (h). An enlarged view of a region in (f) is shown in (g). Our proposed approach (b-d) gave more accurate results than (a) uniform regularization. Incorrect deformation estimation occurred mostly in background regions where discriminative information was lacking (note the upper left and right corners of (b-e)). Obtained displacement fields were also smoother than the one obtained in (a) as reflected by the abrupt changes in colour. (e) A result of BSPLINE registration ( $R$  encoded with CONT).

## 4 Conclusions

We have proposed a data-driven, spatially-adaptive regularization approach for deformable registration. The amount of regularization enforced on the spatial transformation was dependent on an image reliability measure that we developed. As shown in our experiments, the reliability measure can also be used to improve the fidelity of the data term via reliability-based truncation. Different encoding schemes have been proposed to transform our measure to a

weight function for use in an MRF-based registration framework. Use of the DISCR encoding scheme achieved better and more consistent performance over other schemes. Evaluations of our method based on synthetic and real data have demonstrated that our regularization approach can greatly improve registration accuracy of existing MRF-based registration algorithms. Preliminary results on 3D test data already suggest improvement our approach brings, and thus we intend to conduct a thorough evaluation of our approach on full 3D clinical datasets. Future work will include multi-resolution implementation of the proposed method and incorporation of the reliability measure into region-based image similarity metrics to be applied to a wider class of registration problems.

## References

1. Glocker, et al.: Dense image registration through MRFs and efficient linear programming. *Med. Image Anal.* 12(6), 731–741 (2008)
2. Tang, T., Chung, A.: Non-rigid image registration using graph-cuts. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 916–924. Springer, Heidelberg (2007)
3. Kwon, D., et al.: Nonrigid image registration using dynamic higher-order MRF model. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) *ECCV 2008, Part I. LNCS*, vol. 5302, pp. 373–386. Springer, Heidelberg (2008)
4. Hellier, et al.: Hierarchical estimation of a dense deformation field for 3-d robust registration. *IEEE Trans. Med. Imaging* 20(5), 388–402 (2001)
5. Luan, et al.: Multimodality image registration by maximization of quantitative-qualitative measure of mutual information. *Pattern Recognit.* 41, 285–298 (2008)
6. Lester, et al.: Non-linear registration with the variable viscosity fluid algorithm. In: *IEEE IPMI*, pp. 238–251 (1999)
7. Davatzikos, C.: Spatial transformation and registration of brain images using elastically deformable models. *Comput. Vision Image Understand.* 66, 207–222 (1997)
8. Rexilius, J., Warfield, S., Guttman, C.: A novel nonrigid registration algorithm and applications. In: Niessen, W.J., Viergever, M.A. (eds.) *MICCAI 2001. LNCS*, vol. 2208, pp. 923–931. Springer, Heidelberg (2001)
9. Kabus, S.: *Multiple-Material Variational Image Registration*. PhD thesis, der Universität zu Lubeck (October 2006)
10. Pitiot, A., Guimond, A.: Geometrical regularization of displacement fields for histological image registration. *Med. Image Anal.* 12(1), 16–25 (2008)
11. Stefanescu, et al.: Grid powered nonlinear image registration with locally adaptive regularization. *Med. Image Anal.* 8(3), 325–342 (2004)
12. Suarez, et al.: Nonrigid registration using regularized matching weighted by local structure. In: *MICCAI*, pp. 581–589 (2003)
13. Ishikawa, H.: Exact optimization for markov random fields with convex priors. *IEEE Trans. Pattern Anal. Mach. Intell.* 25(10), 1333–1336 (2003)
14. Komodakis, et al.: Performance vs computational efficiency for optimizing single and dynamic MRFs. *Comput. Vision Image Understand.* 112(1), 14–29 (2008)
15. Shekhovtsov, A., Kovtun, I., Hlavac, V.: Efficient MRF deformation model for non-rigid image matching. *Comput. Vision Image Understand* 112, 91–99 (2008)
16. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* 23, 1222–1239 (2001)

17. Glocker, et al.: Dense image registration through MRFs and efficient linear programming. *Med. Image Anal.* 12(6), 731–741 (2008)
18. Paquin, D., Levy, D., Xing, L.: Multiscale deformable registration of noisy medical images. *Math. Biosc. Engin.* 5(1), 125–144 (2008)
19. Rao, J., et al.: Adaptive contextual energy parameterization for automated image segmentation. In: *Bebis, G. (ed.) ISVC 2009. LNCS, vol. 5875, pp. 1089–1100. Springer, Heidelberg (2009)*
20. Donias, M., Baylou, P., Keskes, N.: Curvature of oriented patterns: 2-d and 3-d estimation from differential geometry. In: *ICIP, vol. 70, pp. 236–240 (1998)*
21. Lindeberg, T.: On scale selection for differential operators. In: *Proc. 8th Scandinavian Conf. on Image Analysis, pp. 857–866 (1993)*
22. Leung, et al.: Contour and texture analysis for image segmentation. *Int. J. Comput. Vision* 43(1), 7–27 (2001)
23. Kwan, D., et al.: MRI simulation-based evaluation of image-processing and classification methods. *IEEE Trans. Med. Imaging* 18(11), 1085–1097 (1999)
24. Zhang, L., Seitz, S.: Estimating optimal parameters for MRF stereo from a single image pair. *IEEE Trans. Pattern Anal. Mach. Intell.* 29(1), 331–342 (2007)
25. Hamarneh, G., et al.: Simulation of ground-truth validation data via physically- and statistically-based warps. In: *Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 459–467. Springer, Heidelberg (2008)*
26. Klein, et al.: Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage* 46(3), 786–802 (2009)