Dense Registration with Deformation Priors

Ben Glocker[1,2], Nikos Komodakis[3]
Nassir Navab[2], Georgios Tziritas[3], Nikos Paragios[1,4]

[1] Laboratoire MAS, Ecole Centrale Paris, Chatenay-Malabry, France
[2] Computer Aided Medical Procedures (CAMP), Technische Universität München, Germany
[3] Computer Science Department, University of Crete, Greece
[4] Equipe GALEN, INRIA Saclay - Ile-de-France, Orsay, France
`glocker@in.tum.de`

Abstract. In this paper we propose a novel approach to define task-driven regularization constraints in deformable image registration using learned deformation priors. Our method consists of representing deformation through a set of control points and an interpolation strategy. Then, using a training set of images and the corresponding deformations we seek for a weakly connected graph on the control points where edges define the prior knowledge on the deformation. This graph is obtained using a clustering technique which models the co-dependencies between the displacements of the control points. The resulting classification is used to encode regularization constraints through connections between cluster centers and cluster elements. Additionally, the global structure of the deformation is encoded through a fully connected graph on the cluster centers. Then, registration of a new pair of images consists of displacing the set of control points where on top of conventional image correspondence costs, we introduce costs that are based on the relative deformation of two control points with respect to the learned deformation. The resulting paradigm is implemented using a discrete Markov Random Field which is optimized using efficient linear programming. Promising experimental results on synthetic and real data demonstrate the potential of our approach.

1 Introduction

Image deformation estimation consists of recovering a transformation which aligns two images. Optical flow estimation [1] or fusion of medical images [2] are prominent applications in the fields of computer vision and medical image analysis. Starting from the pioneering formulation of the visual preservation constraint [3], it has been an active problem for almost two decades. The spatial transformation to be recovered establishes correspondences between the two images according to some similarity measure. The task of registration often involves three aspects, the transformation model, the similarity measure, and the optimization strategy used to estimate the transformation parameters.

Variational techniques [4], statistical methods [5] and more recently discrete optimization approaches [6] were considered to address this task. The main challenge to be addressed is related with the ill-posedness definition of the problem. In the most general case, one has to determine a vector of variables from a single constraint, while at the same time signals can be non-linearly related with an unknown spatial transformation (registration). The use of regularization techniques [7] is often used as a soft

prior to deal with the above limitation which often deteriorates estimation along region boundaries.

In medical imaging, if one neglects the global component of the transformation, and since observed signals measure information on anatomical structures, the notion of repetitive behavior is present, for example the deformation of the heart during the cardiac cycle. The use of prior models can be an excellent choice to address the ill-posedness of the registration task, in particular when considering intra-modal registration of challenging, emerging imaging modalities (functional MRI, diffusion tensor imaging, ultrasound, ...), or in case of inter-modal registration for pairs of images where the similarity metric is ill-defined. A lot of work has been done in modeling variations in shape in a more global sense using principal component analysis (PCA). Active shape models (ASMs) [8] have been successfully and widely used in automatic segmentation and shape matching. However, despite enormous investment on model-free deformable registration, one can observe limited work on registration with deformation priors.

The main challenge of dense registration with deformation priors is with regard to the dimensionality of the problem. Opposite to conventional learning problems, in the most general case, learning the deformations corresponding to the entire image domain requires a huge number of training examples. One can overcome this limitation through a rough dimensionality reduction of the deformation fields [9], but then the performance of the model is compromised in terms of ability to capture local deformations. An alternative to such an approach is to consider a rather dense sampling and determine dependencies between deformations. If such a task can be addressed, then one can model prior knowledge on the deformation through decomposition into several local parts where each exhibits similar behavior. This will introduce two novel priors, one that encodes local dependencies and one that accounts for the global structure of the deformation.

In this paper, we propose a novel approach to deformable image registration with priors on the deformation. The deformation field is represented using a set of control points and an interpolation strategy. Using a training set, we perform clustering on the determined deformations according to the displacements of the control points. The aim is to determine pairs with statistical correlation in terms of deformation behavior. Recent advances from linear programming are used to address the clustering task. The outcome of the process is integrated within a discrete Markov Random Field (MRF) [10] approach for dense image registration. Conventional regularization constraints – e.g. penalizing the gradients of the displacement field – are replaced in order to encode co-dependencies between the cluster centers and the corresponding cluster elements. The global structure of the learned deformation is modeled through a fully connected graph on the cluster centers. The resulting graph structure involves a very small number of connections and can be learned from a rather small training set. Experimental results using these two novel deformation priors demonstrate the potential of our approach.

The remainder of this paper is organized as follows: in Section 2 we present the construction of the deformation priors, while in Section 3 the registration paradigm is presented. Experimental results and validation are part of Section 4, while the last Section concludes the paper.

2 Learning and Construction of Deformation Priors

Let us consider a set of N training examples (pairs of images). The corresponding deformation fields $\{D_1, D_2, \ldots, D_N\}$ are assumed to be known (e.g. obtained through dense registration). Considering a transformation model such as Free Form Deformations (FFD) [11], a deformation field D can be efficiently represented through a set of M control points $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_M\}$ defined on a regular grid where each control point $\mathbf{p}_i$ is associated with a displacement vector $\mathbf{d}_i$. Then, the displacement $D(\mathbf{x})$ of any pixel $\mathbf{x}$ in the image domain can be determined through interpolation between control point displacements, or

$$D(\mathbf{x}) = \sum_{i=1}^{M} \eta(\mathbf{x}) \mathbf{d}_i \ , \tag{1}$$

with $\eta(\cdot)$ being the interpolation function (often based on B-spline basis functions [12]).

Modeling prior knowledge based on observed deformations aims at determining a probability density function (pdf) $\psi(D)$. Such a pdf could then be incorporated into the registration procedure and hopefully would improve the estimation of the deformation. In order to construct a compact representation of such a prior, we assume that correlations exist between the behavior of control points. We further assume that we can separate the control points into two groups – the masters and the slaves. The master control points are the ones that encode the most important information, while the slaves are the ones which can be determined to some extent from the master ones. This can be viewed as a clustering problem where cluster centers will correspond to master control points and cluster elements to the slave ones. Each of the slaves will be attributed to a single cluster according to their statistical dependency, while one should reduce the number of retained clusters to a minimum. Such a clustering task involves the following unknown variables: (i) number of clusters K, (ii) identity of the cluster centers $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K\}$, and (iii) assignment of elements to clusters $A = \{a_1, a_2, \ldots, a_M\}$ with $a_i \in [1, K]$. Additionally, let κ be a multivariate probability distribution determined from the N observed displacement vectors of a control point $\mathbf{p}$. Then, the statistical dependency between two control points $\mathbf{p}_i$ and $\mathbf{p}_j$ can be obtained by a distance function $\xi(\kappa_i, \kappa_j)$. From a mathematical perspective, we try to minimize the following function

$$\min_{K,C,A} \sum_{i=1}^{M} \xi(\kappa_{\mathbf{p}_i}, \kappa_{\mathbf{c}_{a_i}}) + \sum_{k=1}^{K} f(\mathbf{c}_k) \ , \tag{2}$$

where $f(\cdot)$ is a cluster penalty term which avoids the trivial solution of choosing all elements as cluster centers.

A common drawback of many popular clustering techniques (such as the K-means algorithm) is that they need to be given the number of clusters K beforehand (which simplifies Eq. (2)). This is, however, problematic as this number is very often not known in advance. Another very bad symptom of many clustering techniques is that they are particularly sensitive to initialization. To address these issues, we make use of a recently proposed clustering method based on linear programming [13] which automatically estimates the optimal number of clusters and works independent from the initialization. Due to the limited space we refer the reader to the given reference which also discusses

the right choice of the cluster penalty term. However, other clustering approaches might be considered as well.

Still, a crucial part is the choice of an appropriate distance function $\xi(\cdot,\cdot)$. It can be for instance the Kullback-Leibler divergence [14] or the Bhattacharyya divergence [15] which is considered in this paper. The Bhattacharyya measure is defined as

$$B(\kappa_i, \kappa_j) = \int_{-\infty}^{\infty} \sqrt{\kappa_i(\mathbf{x})\kappa_j(\mathbf{x})} d\mathbf{x} \ , \tag{3}$$

which satisfies the properties $0 \le B(\kappa_i, \kappa_j) \le 1$, $B(\kappa_i, \kappa_j) = B(\kappa_j, \kappa_i)$ and $B(\kappa_i, \kappa_j) = 1$ if and only if $\kappa_i = \kappa_j$. The corresponding Bhattacharyya divergence is defined as $\xi_B(\kappa_i, \kappa_j) = -\log B(\kappa_i, \kappa_j)$. For Gaussian distributions the Bhattacharyya divergence has a closed form expression, while for more complex distributions such as Gaussian mixture models (GMMs) closed form expressions do not exist and sampling strategies [16] have to be used.

2.1   Deformation Prior

The outcome of the clustering consists of a set of cluster centers $C \subseteq P$ and disjunct clusters of control points $\{P_1, P_2, \ldots, P_K\}$ with $\bigcup_{k=1}^{K} P_k = P$. Then, in order to capture the local dependencies within a cluster, we consider the pairwise probability distributions on the relative deformation between cluster elements $\mathbf{p} \in P_k$ and the corresponding cluster center $\mathbf{c}_k \in P_k, C$, or

$$\psi_k^{\text{local}}(D) = \prod_{\mathbf{p},\mathbf{c}_k \in P_k, \mathbf{c}_k \in C, \mathbf{p} \neq \mathbf{c}_k} \kappa_{\mathbf{p}\mathbf{c}_k}(\|\mathbf{d}_\mathbf{p} - \mathbf{d}_{\mathbf{c}_k}\|) \ . \tag{4}$$

Additionally, in order to capture the global structure of the learned deformation, we consider the distributions between cluster centers

$$\psi^{\text{global}}(D) = \prod_{\mathbf{c}_i,\mathbf{c}_j \in C, i \neq j} \kappa_{\mathbf{c}_i\mathbf{c}_j}(\|\mathbf{d}_{\mathbf{c}_i} - \mathbf{d}_{\mathbf{c}_j}\|) \ . \tag{5}$$

The pairwise distributions for the relative deformation of two control points are estimated from the training data, once the clustering is computed.

We can now approximate the overall pdf $\psi(\mathcal{D})$ by combining the above terms into a prior on the deformation, or

$$\psi(D) = \psi^{\text{global}}(D) \prod_{k=1}^{K} \psi_k^{\text{local}}(D) \tag{6}$$

One should note that such a representation is invariant to global translation. Implementing this prior into a registration framework constraints the space of feasible deformations. This could be of great benefit in cases where the similarity measure deteriorates due to noise or data corruption. Also, conventional soft priors – e.g. penalizing the gradients of the displacement field – might result in oversmoothed displacement fields which could also be overcome by incorporating the above priors.

Throughout this paper, we represent the probability density functions by GMMs. Other representations can be considered, however the main advantage of GMMs is their compact representation for complex distributions while being efficient in terms of evaluation and implementation. The density estimation from the training data is performed through the EM algorithm [17] while the optimal number of Gaussians is determined in a brute force manner. To this end, we evaluate the minimum description length (MDL) for 1 to 5 Gaussians and keep the GMM with the lowest MDL.

3   MRF Registration with Deformation Priors

In order to prove the concept of using the proposed deformation priors for image registration, we implemented the proposed paradigm in a dense registration framework based on discrete MRFs [18]. In our case, the main advantage of such a framework is that MRFs naturally allow for encoding dependencies between pairs of variables (here, the FFD control points). We show that simple changes in the MRF topology and the replacement of conventional regularization costs are sufficient to enhance the registration algorithm with the learned deformation priors.

3.1   Dense Registration through MRF Labeling

We will briefly recall the principles of the intensity-based registration framework described in [18]. Considering the common approach of energy minimization for the registration of two images I and J, or

$$D^* = \arg\min_D \int_\Omega \phi\left(I(\mathbf{x}), J(\mathbf{x} + D(\mathbf{x}))\right) d\mathbf{x} \tag{7}$$

one seeks for recovering the optimal deformation D^* w.r.t. a similarity measure ϕ. Considering the FFD transformation model defined Eq. (1), we can define the objective function based on the control points as

$$E_{\text{data}}(D) = \sum_{\mathbf{p} \in P} \int_\Omega \hat{\eta}(\mathbf{x}) \cdot \phi\left(I(\mathbf{x}), J(\mathbf{x} + D(\mathbf{x}))\right) d\mathbf{x}. \tag{8}$$

where $\hat{\eta}(\cdot)$ is a weighting function determining the influence of a pixel $\mathbf{x}$ to the local similarity at control point $\mathbf{p}$. Different approaches for the weighting function can be considered (cp. [18]), depending on the nature of the similarity measure.

The key idea in this approach is now to reformulate the registration problem as a *discrete labeling problem*. Based on the previous assumptions, the control points P of the deformation grid are considered as a set of discrete variables. Additionally, a discrete set of labels $L = \{l_1, ..., l_i\}$ corresponding to a quantized version of the deformation space $\Theta = \{\mathbf{d}_1, ..., \mathbf{d}_i\}$ is introduced. A label assignment $l_{\mathbf{p}}$ to a grid node $\mathbf{p}$ is associated with displacing the node by the corresponding vector $\mathbf{d}_{l_{\mathbf{p}}}$. Once a label is assigned to every node we obtain a discrete labeling $\mathbf{l}$. A popular and efficient model for representing such discrete labeling problems are second-order MRFs [19]:

$$E_{\text{mrf}}(\mathbf{l}) = \sum_{\mathbf{p} \in P} V_{\mathbf{p}}(l_{\mathbf{p}}) + \lambda \sum_{(\mathbf{p}, \mathbf{q}) \in S} V_{\mathbf{pq}}(l_{\mathbf{p}}, l_{\mathbf{q}}) \ , \tag{9}$$

where $V_{\mathbf{p}}(\cdot)$ are the unary potentials representing the data term, $V_{\mathbf{pq}}(\cdot,\cdot)$ are the pairwise potentials representing dependencies between neighboring nodes, and S defines the neighborhood system through a set of edges. Additionally, λ acts as a weighting factor controlling the influence of the pairwise term. We define the unary potentials (in iteration t) according to our data term as

$$V_{\mathbf{p}}(l_{\mathbf{p}}) = \int_{\Omega} \hat{\eta}(\mathbf{x}) \cdot \phi\left(I(\mathbf{x}), J(\mathbf{x} + D^{t-1}(\mathbf{x}) + \mathbf{d}_{l_{\mathbf{p}}})\right) d\mathbf{x} \ . \tag{10}$$

The pairwise potentials can encode penalty costs for assigning different labels to connected nodes. The FFD transformation model already inherits implicit smoothness. Additionally, one can consider explicit regularization constraints on the grid domain. A common, but rather heuristic, approach is to consider regularization on the squared differences of displacement vectors – i.e. an approximation of penalizing the gradients of the deformation field. This can be defined as

$$V_{\mathbf{pq}}(l_{\mathbf{p}}, l_{\mathbf{q}}) = \left((\mathbf{d}_{\mathbf{p}}^{t-1} + \mathbf{d}_{l_{\mathbf{p}}}) - (\mathbf{d}_{\mathbf{q}}^{t-1} + \mathbf{d}_{l_{\mathbf{q}}})\right)^2 \ , \tag{11}$$

where $\mathbf{d}_{\mathbf{p}}^{t-1}$ and $\mathbf{d}_{\mathbf{q}}^{t-1}$ are the accumulated displacements for the control points $\mathbf{p}$ and $\mathbf{q}$ in iteration t.

Many optimization algorithms exist for efficiently solving discrete labeling problems in form of an MRF. We use a recently proposed method called FastPD [20] which is also used in [6, 18]. Due to the limited space, we refer the reader to the given references for more details about the algorithm.

3.2   Local and Global Prior Costs

Conventional regularization techniques as introduced in Eq. (11) can be seen as soft priors which are used to deal with the ill-posedness of the registration problem. Still, these approaches may deteriorate the estimation along region boundaries or result in oversmoothed displacement fields, since there is no relationship between the penalty function based on the grid topology and the actual local dependencies of object deformations. To overcome this limitation, we propose a novel regularization term based on the results of the control point clustering. One can claim that variations on the deformation of nodes belonging to one cluster should be penalized w.r.t. to the relative deformation learned from the training data while discontinuities between nodes of different clusters should be explicitly allowed (even desired). We propose to remove the conventional regularization term which simply imposes smoothness for neighboring nodes. Instead, we define an intra-cluster regularization that accounts for the local dependencies of control points. For each cluster k, we connect its elements $\mathbf{p} \in P_k$ with the cluster center $\mathbf{c}_k$ and define edge penalty costs derived from Eq. (4) as:

$$V_{\mathbf{pc}_k}(l_{\mathbf{p}}, l_{\mathbf{c}_k}) = -\log\left(\kappa_{\mathbf{pc}_k}(\|(\mathbf{d}_{\mathbf{p}}^{t-1} + \mathbf{d}_{l_{\mathbf{p}}}) - (\mathbf{d}_{\mathbf{c}_k}^{t-1} + \mathbf{d}_{l_{\mathbf{c}_k}})\|)\right) \ . \tag{12}$$

The advantage of this cluster-based regularization is on one hand its direct relation to the learned local dependencies on the level of control points, and on the other hand, it allows for discontinuities between neighboring control points belonging to different

clusters. Additional to the cluster-based regularization, we propose a second prior term that accounts for the global structure of the deformation. Using the concept of pairwise densities defined in Eq. (5), we can impose a global prior cost on the desired deformations by introducing connections between all cluster centers (fully connected graph), or

$$V_{\mathbf{c}_i \mathbf{c}_j}(l_{\mathbf{c}_i}, l_{\mathbf{c}_j}) = -\log\left(\kappa_{\mathbf{c}_i \mathbf{c}_j}(\|(\mathbf{d}_{\mathbf{c}_i}^{t-1} + \mathbf{d}_{l_{\mathbf{c}_i}}) - (\mathbf{d}_{\mathbf{c}_j}^{t-1} + \mathbf{d}_{l_{\mathbf{c}_j}})\|)\right) \ . \tag{13}$$

The overall pdf derived in Eq. (6) can then be encoded by combining the two terms which leads to the proposed registration approach using learned deformation priors. The new energy function of the MRF registration is then defined as

$$E_{\mathrm{mrf}}(\mathbf{l}) = \underbrace{\sum_{\mathbf{p} \in P} V_{\mathbf{p}}(l_{\mathbf{p}})}_{\text{Data Costs}} + \lambda \underbrace{\left(\sum_{\mathbf{c}_i, \mathbf{c}_j \in C} V_{\mathbf{c}_i \mathbf{c}_j}(l_{\mathbf{c}_i}, l_{\mathbf{c}_j}) + \sum_{k=1}^{K} \sum_{\mathbf{p}, \mathbf{c}_k \in P_k} V_{\mathbf{p}\mathbf{c}_k}(l_{\mathbf{p}}, l_{\mathbf{c}_k}) \right)}_{\text{Deformation Prior Costs}} \ . \tag{14}$$

4   Experimental Validation

In our first experiment, we investigate the performance of the proposed approach in a synthetic scenario. On the one hand, this experiment should illustrate the single steps of our method from the training phase to the final application of image registration. On the other hand, we will compare its performance to conventional regularization techniques – i.e. penalizing the vector difference between neighboring control points [18]. The second experiment is on real data where the idea is to learn the deformation on images of good quality which are suitable for conventional registration. After the learning, the registration with deformation priors is used to register images of bad quality showing the same anatomy as in the training data.

4.1   Experiments on Synthetic Deformations

In the absence of a gold standard, evaluation in deformable registration settings is a challenging task. One option is to consider synthetic deformations where everything is known by construction. To this end, we generate a 2D image with a resolution of 128×96 showing a bone like structure and define a deformation grid of size 17×13 (see Fig. 1(e)). We separate the control points into nine clusters from which eight clusters are used to generate random deformations and the ninth cluster is without deformation. Random displacements of elements belonging to the same cluster are following a similar Gaussian distribution. In total we generate 500 deformations. Examples are shown in Fig. 1(a)-(d) where the dense displacement field is overlaid. We use 400 deformations for the training and the prior construction. A Gaussian density with zero mean and standard deviation one is assumed for the control points of the ninth cluster without deformation. All other densities are estimated from the training data. Based on the clustering result (see Fig. 1(f)) we define the new MRF topology using intra-cluster edges for the local prior as shown in Fig. 1(g) and for the global prior we define a fully connected graph on the cluster centers as shown in Fig. 1(h). For better visibility all connections within and to the ninth cluster are discarded from the drawings.
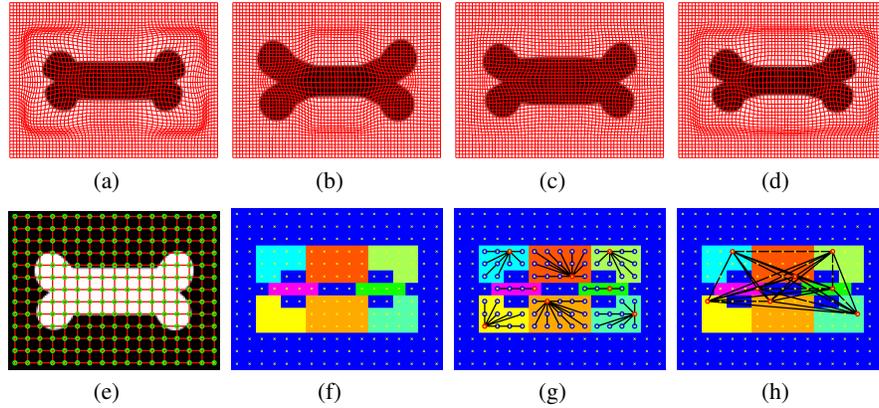
Fig. 1: Synthetic experiment: the upper row shows exemplary deformations and overlaid the corresponding dense displacement fields. The lower row illustrates the clustering (f) and the resulting MRF topology consisting of intra-cluster edges (g) and the fully connected graph on the cluster centers (h).

We compare the performance of the proposed method to conventional regularization as defined in Eq. (11) and used for instance in [18]. To this end, the source image (see Fig. 1(e)) is registered to different target images where either varying amount of Gaussian noise is added or the image data is heavily corrupted. We use the sum of squared differences (SSD) as the similarity measure. The SSD is very sensitive to noise and allows us to investigate the important role of regularization. The weighting factors λ are individually determined empirically for the conventional and the approach using learned priors such that smooth and reasonable deformations are obtained. In particular, the conventional method is sensitive to λ such that too less regularization leads the SSD measure to smear the source image while too much regularization constraints the deformation to rigid translations. Once a good compromise is found on exemplary images, the values are fixed throughout the experiments for both methods. Exemplary target images are shown in Fig. 3(a)-(d). After registration, we compare the original shape boundary of the target image with the warped boundary of the deformed source image and compute the symmetric average boundary distance (ABD). Visually, we observe that conventional regularization tends to result in oversmoothed displacement fields (see Fig. 3(m)) while the registration with deformation priors reflects the single clusters in the recovered deformations (see Fig. 3(q)). Visual registration results and the color encoded displacement fields are illustrated in Fig. 3(e)-(t). Quantitative measurements of the ABD after registration are summarized in Fig. 2. The overall ABD before registration is $2.16(\pm 0.46)$ pixels. The approach using learned priors clearly outperforms the conventional regularization. In particular under extreme conditions of noise and corruption, the priors on the deformation help to achieve meaningful registration results while the conventional approach fails.

Target	Conventional	Learned Priors
Noise σ=0.1	0.42 (± 0.11)	0.16 (± 0.06)
Noise σ=1.0	0.73 (± 0.18)	0.42 (± 0.15)
Noise σ=5.0	1.17 (± 0.26)	0.79 (± 0.25)
Corrupted	1.52 (± 0.49)	0.69 (± 0.12)
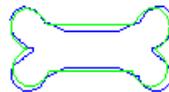
Fig. 2: Synthetic experiment: on the left, the average boundary distance (ABD) after registration using the conventional regularization and the approach using learned priors. Before registration the overall ABD is $2.16 (\pm 0.46)$ pixels. On the right, the initial alignment of source (in green) and an exemplary target image (in blue) also used in Fig. 3. For this example, the ABD is 1.77 pixels.

4.2   Experiments on Real Data

In our second experiment, we consider clinical data of a myopathy study. MRI scans showing the lower leg muscles are acquired from 23 subjects. Each scan consists of T1 and DTI images. Note, that from the DTI images only the B0 image is used. While the T1 images have a very good contrast, a high resolution ($256 \times 256 \times 20$), and a low signal-to-noise ratio, the DTI images are of low resolution ($64 \times 64 \times 20$) and rather bad quality where a lot of artifacts are present (see Fig 4). In a leave-one-out manner, we randomly select 22 subjects for the training where one subject is chosen as the reference. We perform a series of conventional registrations using only the T1 images where the reference is aligned with all other training images. The results of the 21 registrations are then used for the learning of the deformation priors. Since we consider a multi-resolution approach, we separately learn the deformation prior for each control grid level. For evaluation, we register the DTI image of the reference with the DTI of the subject which is not included in the training set. Here, we use the normalized cross correlation as the similarity measure. Exemplary visual results are shown in Fig. 4. For a better visual inspection of the results, we also warp the T1 images using the deformations obtained from the DTI registration, and visualize the difference images. A perfect alignment based on the challenging DTI images is not achieved. However, the results show that the deformation prior drives the registration towards the right solution where the similarity measure alone combined with conventional regularization fails to obtain a good alignment.

5   Discussion

In this paper we have proposed a novel approach to image registration using learned deformation priors. To this end, the deformation of the control points were considered as random variables and clustering on the statistical behavior of these variables was performed to determine their co-dependencies. Connections between cluster centers and attributed elements were used to impose regularization, while connections between cluster centers aimed to capture the global structure of the learned deformation. This prior was naturally considered within an MRF towards improving registration performance on challenging data as well as on sequences heavily corrupted by noise. The

distribution of control points and the number of clusters being retained are critical components of our process as well as the distance between variables. Non-uniform transformation models (such as NURBS) might be a more appropriate selection for the graph construction. Further investigation on the above mentioned components of the method could improve the performance. The use of this concept to address registration within modalities that do exhibit low signal to noise ratio is a natural extension of our approach. Ultrasound images, diffusion tensor imaging, or functional MRI are examples where conventional registration techniques might often fail to provide meaningful results. Therefore, we believe that our approach has great potential in such applications.

References

1. Fleet, D., Weiss, Y.: Optical Flow Estimation. In: Handbook of Mathematical Models in Computer Vision. Springer (2006) 239–256
2. Hajnal, J., Hill, D.L.G., Hawkes, D.J., eds.: Medical Image Registration. CRC Press (2001)
3. Horn, B., Schunck, B.: Determining optical flow. Artificial Intelligence **17** (1981) 185–204
4. Bruhn, A., Weickert, J., Schnörr, C.: Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. International Journal of Computer Vision (IJCV) **61**(3) (2005)
5. Viola, P., Wells, W.M.: Alignment by maximization of mutual information. International Journal of Computer Vision (IJCV) **24**(2) (1997) 137–154
6. Glocker, B., Komodakis, N., Paragios, N., Tziritas, G., Navab, N.: Inter and intra-modal deformable registration: Continuous deformations meet efficient optimal linear programming. In: Information Processing in Medical Imaging (IPMI), Kerkrade, Netherlands (July 2007)
7. Tikhonov, A.: Ill-posed problems in natural sciences (1992) Coronet.
8. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models — their training and application. Computer Vision and Image Understanding (CVIU) **61**(1) (1995) 38–59
9. Black, M.J., Anandan, P.: The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. Computer Vision and Image Understanding (CVIU) **63**(1) (1996) 75–104
10. Geman, S., Geman, D.: Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. In: IEEE Transactions on Pattern Recognition and Machine Learning (PAMI). Volume 6. (1984)
11. Sederberg, T.W., Parry, S.R.: Free-form deformation of solid geometric models. In: SIGGRAPH, New York, NY, USA, ACM Press (1986)
12. Rueckert, D., Sonoda, L., Hayes, C., Hill, D., Leach, M., Hawkes, D.: Nonrigid registration using free-form deformations: application to breast mr images. IEEE Transactions on Medical Imaging (TMI) **18**(8) (1999) 712–721
13. Komodakis, N., Paragios, N., Tziritas, G.: Clustering via lp-based stabilities. In: Neural Information Processing Systems (NIPS). (2008)
14. Kullback, S.: Information Theory and Statistics. Dover Publications Inc., New York (1968)
15. Bhattacharyya, A.: On a measure of divergence between two statistical populations defined by probability distributions. Bull. Calcutta Math. Soc. **35** (1943) 99–109
16. Olsen, P., Hershey, J.: Bhattacharyya error and divergence using variational importance sampling. In: Interspeech, Antwerp, Belgium (August 2007)
17. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer (2006)
18. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through mrfs and efficient linear programming. Medical Image Analysis **12**(6) (2008)
19. Li, S.Z.: Markov random field modeling in image analysis. Springer (2001)
20. Komodakis, N., Tziritas, G., Paragios, N.: Fast, approximately optimal solutions for single and dynamic mrfs. In: Computer Vision and Pattern Recognition (CVPR). (2007)
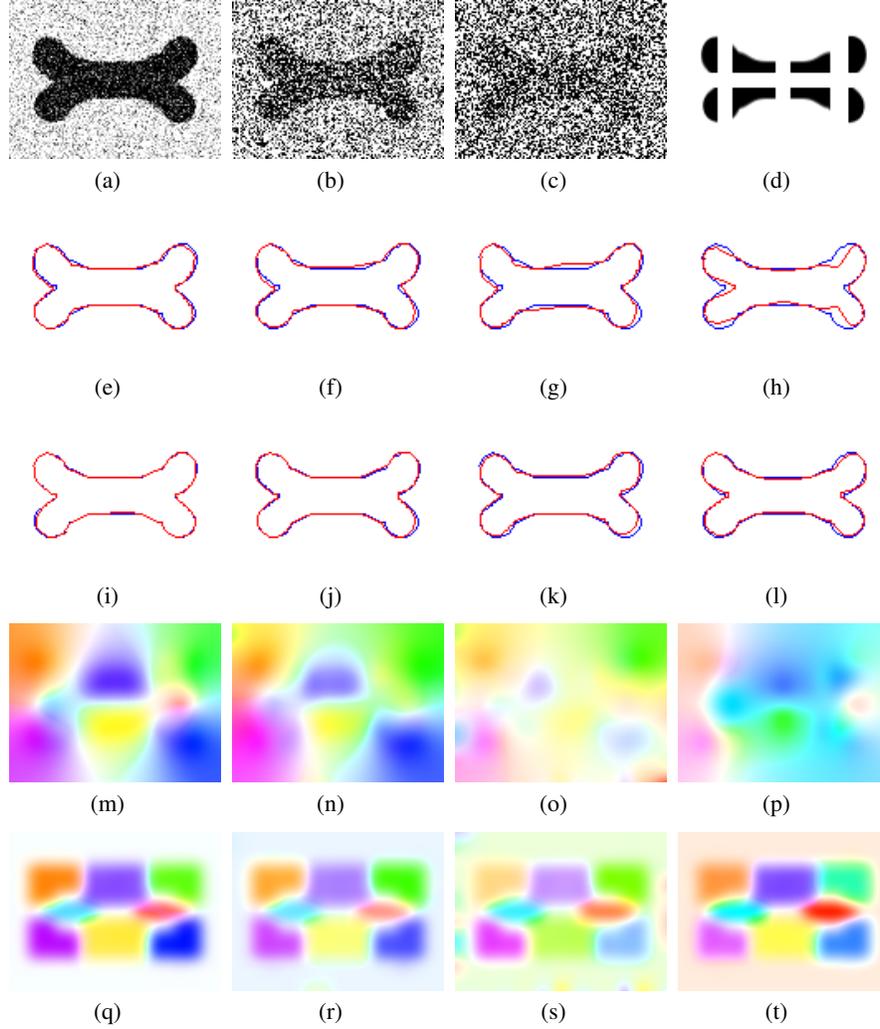
Fig. 3: Synthetic experiment: the first row illustrates exemplary target images with varying amounts of noise and corruption. The second row shows the registration results for the conventional registration while the third row shows the results for using learned priors. The target shape boundary is drawn in blue, the shape boundary of the deformed source image is drawn in red. The initial alignment before registration is illustrated in Fig. 2. The fourth and fifth row show the color encoded displacement fields for both methods. The color represents the angle, the intensity the magnitude of the displacement vectors. Note, how well the clustering topology is reflected in all fields for registration with deformation priors (compare Fig. 1(f)).
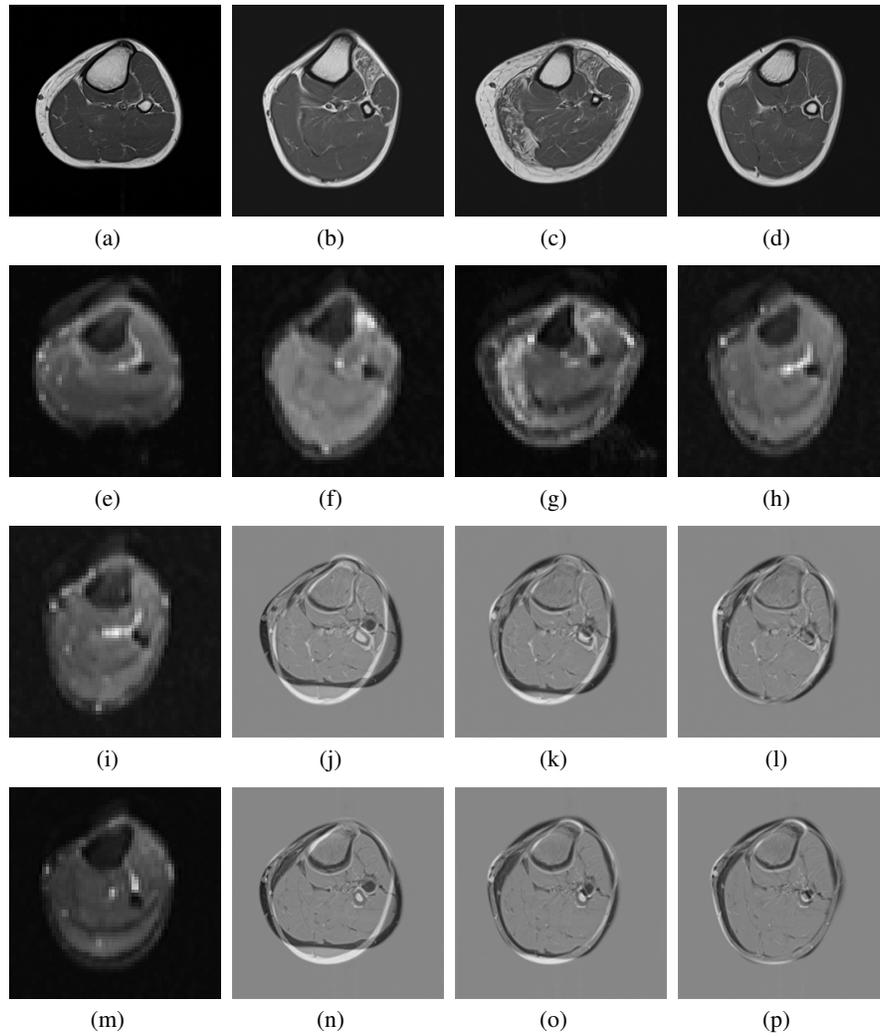
Fig. 4: Real data experiment: Upper row shows exemplary T1 images from the training set. Middle row shows the corresponding B0 DTI images. The two last rows show from left to right the target image, the difference image for the initial alignment, the alignment after conventional registration, and the result for the registration with deformation priors when using (e) as the source image. Note, for better visual inspection of the registration results, the difference images (j)-(l) are computed from the warped T1 images using the deformations obtained from the DTI registrations.