Inter and Intra-Modal Deformable Registration: Continuous Deformations Meet Efficient Optimal Linear Programming

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

[1] Chair for Computer Aided Medical Procedures & Augmented Reality
Technische Universität München
{glocker,navab}@in.tum.de
[2] Computer Science Department, University of Crete
{komod,tziritas}@csd.uoc.gr
[3] Mathématiques Appliquées aux Systèmes, Ecole Centrale Paris
nikos.paragios@ecp.fr

Abstract. In this paper we propose a novel non-rigid volume registration based on discrete labeling and linear programming. The proposed framework reformulates registration as a minimal path extraction in a weighted graph. The space of solutions is represented using a set of a labels which are assigned to predefined displacements. The graph topology corresponds to a superimposed regular grid onto the volume. Links between neighborhood control points introduce smoothness, while links between the graph nodes and the labels (end-nodes) measure the cost induced to the objective function through the selection of a particular deformation for a given control point once projected to the entire volume domain. Higher order polynomials are used to express the volume deformation from the ones of the control points. Efficient linear programming that can guarantee the optimal solution up to (a user-defined) bound is considered to recover the optimal registration parameters. Therefore, the method is gradient free, can encode various similarity metrics (simple changes on the graph construction), can guarantee a globally sub-optimal solution and is computational tractable. Experimental validation using simulated data with known deformation, as well as manually segmented data demonstrate the extreme potentials of our approach.

Key words: Discrete Optimization, Deformable Registration, Linear Programming

1 Introduction

Deformable registration is one of the most challenging problems in medical imaging. The problem consists of recovering a local transformation that aligns two signals that have a non-linear relationship often unknown. Several methods exist in the literature where specific metrics are designed to account for this non-linearity and optimize the transformation that brings together these two signals. This optimization is often sub-optimal due the non convexity of the designed cost functions. The aim of our approach

is to overcome both limitations present in all registration methods. Dependency on the similarity metric selection, as well as to the initial conditions.

Local image alignment is often performed according to geometric or photometric criteria. Landmark-based methods [1] are a classic example of geometric-driven registration. In such a setting, a number of anatomical key points [2]/structures (segmented values) are identified both in the source and the target image and a transformation that aims to minimize the Euclidean distance between these structures is to be recovered. The main limitation of these methods related to the selection and extraction of landmarks, while their main strength is the simplicity of the optimization process.

Iconic registration methods seek for "visual" correspondences between the source and the target image. Such a problem is tractable when one seeks registration for images from the same modality due to an explicit photometric correspondence of the image intensities. Sum of squared differences [3], sum of absolute differences [3], cross correlation [3] or distances on subspaces that involve both appearance and geometry (intensities, curvature, higher order image moments) [4] have been considered. On the other hand it becomes more challenging when seeking transformations between different modalities where a non-linear transformation often relates them. Non-linear metrics have often been used [5] like normalized mutual information [6], kulback-leiber divergence [7] and correlation ratio [8] are some of the metrics used to define similarity between different modalities.

In this paper we propose a novel technique that can either be used for inter or intra modal image registration. Towards satisfying smoothness of the deformation field and reducing the dimensionality of the problem we represent deformation through Free Form Deformations. Our method reformulates registration as an MRF optimization where a set of labels is associated with a set of deformations, and one seeks to attribute a label to each control point such that once the corresponding deformation has been applied, the similarity metric between the source and the target is maximal for all voxels. The optimization procedure is independent from the graph construction, and therefore any similarity metric can be used.

The reminder of this paper is organized as follows; In section 2 we introduce the proposed registration framework, while in section 3 we discuss the optimization aspects. Implementation and experimental validation are part of section 4. The last section concludes our paper.

2 Deformable Registration in a Discrete Setting

In order to introduce the concept of our approach, we consider (without loss of generality) the 2D image domain. Let us consider a source $f : [1,N] \times [1,M] \rightarrow \mathcal{R}^n$ and a target image g. In general, these images are related with a non linear transformation as well as a non-linear relation between intensities, that is

$$g(\mathbf{x}) = h \circ f(\mathcal{T}(\mathbf{x})) \tag{1}$$

where $\mathcal{T}$ is the transformation and h is a non-linear operator explaining the changes of appearance between them. The most common way to formulate the registration problem, is through the definition of a distance between the source and the target image that

is to be minimized in the entire domain Ω, or

$$E(\mathcal{T}) = \iint_{\Omega} \rho(g(\mathbf{x}), h \circ f(\mathcal{T}(\mathbf{x}))d\mathbf{x} \tag{2}$$

where ρ is a similarity metric used to determine meaningful correspondence. Since in most of the cases the non-linear transformation relating the two images is not known, the selection of similarity metric ρ explicitly or implicitly accounts for this non-linearity, or

$$E(\mathcal{T}) = \iint_{\Omega} \rho_h(g(\mathbf{x}), f(\mathcal{T}(\mathbf{x}))d\mathbf{x} \tag{3}$$

2.1 Continuous Domain

Since we are interested in local registration, let us introduce a deformation grid $\mathcal{G} : [1,K] \times [1,L]$ (usually $K \ll M$ and $L \ll N$) super-imposed to the image (no particular assumption is made on the grid resolution). The central idea of our approach is to deform the grid (with a 2D displacement vector $\mathbf{d}_p$ for each control point) such that the underlying image structures are perfectly aligned. Without loss of generality one can assume that the transformation of an image pixel $\mathbf{x}$ can be expressed using a linear or non-linear combination of the grid points, or

$$\mathcal{T}(\mathbf{x}) = \mathbf{x} + \mathcal{D}(\mathbf{x}) \quad \text{with} \quad \mathcal{D}(\mathbf{x}) = \sum_{p \in \mathcal{G}} \eta(|\mathbf{x} - \mathbf{p}|)\, \mathbf{d}_p \tag{4}$$

where $\eta(\cdot)$ is the weighting function measuring the contribution of the control point p to the displacement field $\mathcal{D}$. The position of point p is denoted as $\mathbf{p}$. In such a theoretical setting without loss of generality we consider Free Form Deformations (FFD) based on cubic B-splines as a transformation model. FFD are successfully applied in non-rigid image registration [9, 10]. Deformation of an object is achieved by manipulating an underlying mesh of uniformly spaced control points. The displacement field for a two-dimensional FFD based on cubic B-Splines is defined as

$$\mathcal{D}(\mathbf{x}) = \sum_{l=0}^{3} \sum_{m=0}^{3} B_l(u) B_m(v)\, \mathbf{d}_{i+l,j+m} \tag{5}$$

where $i = \lfloor x/K \rfloor - 1$, $j = \lfloor y/L \rfloor - 1$, $u = x/K - \lfloor x/K \rfloor$, and $v = y/L - \lfloor y/L \rfloor$ and where B_l represents the lth basis function of the B-Spline. The three-dimensional version is defined straightforward.

By defining the registration problem based on such a deformation model we can now rewrite the criterion earlier introduced,

$$E_{\text{data}}(\mathcal{T}) = \sum_{p \in \mathcal{G}} \iint_{\Omega} \eta^{-1}(|\mathbf{x} - \mathbf{p}|) \cdot \rho_h(g(\mathbf{x}), f(\mathcal{T}(\mathbf{x})))d\mathbf{x}. \tag{6}$$

where $\eta^{-1}(\cdot)$ is the inverse projection for the contribution to the objective of the image pixel $\mathbf{x}$ according to the influence of the control point p.

Such a term will guarantee photometric correspondence between the two images. Hence, this term is also called the data term. The transformation due to the interpolation

inherits some implicit smoothness properties. However, in order to avoid folding of the deformation grid, one can consider a smoothness term on the grid domain, or

$$E_{\text{smooth}}(\mathcal{T}) = \sum_{p \in \mathcal{G}} \phi(|\nabla_{\mathcal{G}} \mathbf{d}_p|) \tag{7}$$

with ϕ being a smoothness penalty function for instance penalizing the first derivatives of the grid deformation. The complete term associated with the registration problem is then defined as the sum of the data and smoothness term, or

$$E_{\text{total}} = E_{\text{data}} + E_{\text{smooth}}. \tag{8}$$

The most common way to obtain the transformation parameters is through the use of a gradient-descent method in an iterative approach. Thus given an initial guess, one updates the estimate according to the following formula $\left[\mathcal{T}^m = \mathcal{T}^{m-1} - \delta t \frac{E_{\text{total}}}{\partial T}\right]$. Such a process involves the derivative of the similarity metric with respect to the transformation parameters and therefore it is model and criterion dependent. Slight modifications on the cost function could lead to a different derivative and require novel numerical approximation methods.

2.2 Discrete Domain

Let us now consider a discrete set of labels $\mathcal{L} = \{u^1, ..., u^i\}$ corresponding to a quantized version of the deformation space $\times = \{\mathbf{d}^1, ..., \mathbf{d}^i\}$. A label assignment u_p to a grid node p is associated with displacing the node by the corresponding vector $\mathbf{d}^{u_p}$. The image transformation associated with a certain discrete labeling u becomes

$$\mathcal{D}(\mathbf{x}) = \sum_{p \in \mathcal{G}} \eta(|\mathbf{x} - \mathbf{p}|)\, \mathbf{d}^{u_p}. \tag{9}$$

One can reformulate the registration as a discrete optimization problem, that is assign individual labels u_p to the grid nodes such that

$$E_{\text{data}}(u) = \sum_{p \in \mathcal{G}} \iint_{\Omega} \eta^{-1}(|\mathbf{x} - \mathbf{p}|) \rho_h(g(\mathbf{x}), f(\mathcal{T}(\mathbf{x}))) d\mathbf{x} \approx \sum_{p \in \mathcal{G}} V_p(u_p) \tag{10}$$

where $V_p(\cdot)$ represents a local similarity metric. There is a main issue coming along when using MRF-based optimization methods for our proposed setting. Here, the singleton potential functions $V_p(\cdot)$ are not independent, thus the defined data term can only be approximated. Hence, we pre-compute the $|\mathcal{L}| \times |\mathcal{G}|$ data term look-up table for a given image pair by simple shift operators. The entry for node p and labels u_p is determined by

$$V_p(u_p) = \iint_{\Omega} \eta^{-1}(|\mathbf{x} - \mathbf{p}|) \rho_h(g(\mathbf{x}), f(\mathbf{x} + \mathbf{d}^{u_p})) d\mathbf{x}. \tag{11}$$

The computation of such functions is very fast and straightforward. Any similarity metric can be simply plugged in this formulation without any changes or adaptations. Since

the metrics are only considered on the image domain and no further analytical differentiation is needed, our approach is extremely flexible. Due to the approximation of the data term we allow to improve the estimation by successive optimizations resulting in a series of cost functions, or

$$E^t_{\text{data}}(u) = \sum_{p\in\mathcal{G}} \iint_{\Omega} \eta^{-1}(|\mathbf{x}-\mathbf{p}|)\rho_h(g(\mathbf{x}), f(\mathcal{T}^{t-1}(\mathbf{x}) + \mathbf{d}^{u_p}))d\mathbf{x}. \tag{12}$$

We should note, that from the optimization point of view we achieve (quasi) optimal solutions for the discrete labeling in every cycle. However, we can achieve a higher accuracy by successive cycles using the previous transformation $\mathcal{T}^{t-1}$.

The number of labels and their range play a significant role to the registration process. It is clear that setting the number of labels to infinity will converge to the continuous formulation which though it is intractable from computational perspective. However, the fact that we perform several cycles to improve the accuracy of the deformation allows us to keep the set of labels quite small.

The next aspect to be addressed, is the definition of the smoothness term in the label domain. One can express distances between the deformation vectors using difference between labels if a ranking has been considered within the definition of the label set, or

$$E_{\text{smooth}}(u) = \sum_{p,q\in\mathcal{E}} V_{pq}(u_p, u_q) \tag{13}$$

where $\mathcal{E}$ represents the neighborhood system associated with the deformation grid $\mathcal{G}$. For the distance $V_{pq}(\cdot,\cdot)$ we consider a simple piecewise smoothness truncated term based on the euclidean geometric distances between the deformations corresponding to the assigned labels:

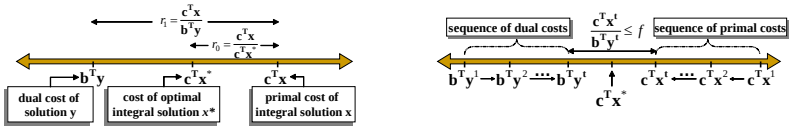
$$V_{pq}(u_p, u_q) = \lambda_{pq} \min\left(|\mathbf{d}^{u_p} - \mathbf{d}^{u_q}|, T\right) \tag{14}$$

with T being the maximum penalty and λ_{pq} being a (spacial varying) weighting to control the influence of the prior term. Basically, this is a discrete approximation of the smoothness term defined in equation 7 extended by the piecewise property. Such a smoothness term together with the data term allows to convert the problem of image registration into the form of a Markov Random Field (MRF) [11] in a discrete domain, or

$$E_{\text{total}}(u) = \sum_{p\in\mathcal{G}} V_p(u_p) + \sum_{p,q\in\mathcal{E}} V_{pq}(u_p, u_q). \tag{15}$$

3 MRF optimization based on Linear Programming

For optimizing the resulting MRF, we seek to assign a label $u_p \in \mathcal{L}$ to each node $p \in \mathcal{G}$, so that the MRF energy in (15) is minimized. To this end, a recently proposed method, called Fast-PD, will be used [12]. This is an optimization technique, which builds upon principles drawn from the duality theory of linear programming in order to efficiently derive almost optimal solutions for a very wide class of NP-hard MRFs. When applied to the image registration task, this technique thus offers a series of important advantages compared to prior art (see section 3.2).



(a) The primal-dual principle (b) The primal-dual schema

Fig. 1: (a) By weak duality, the optimal cost $\mathbf{c}^T\mathbf{x}^*$ will lie between the costs $\mathbf{b}^T\mathbf{y}$ and $\mathbf{c}^T\mathbf{x}$ of any pair $(\mathbf{x},\mathbf{y})$ of integral-primal and dual feasible solutions. Therefore, if $\mathbf{b}^T\mathbf{y}$ and $\mathbf{c}^T\mathbf{x}$ are close enough (e.g. their ratio r_1 is $\leq f$), so are $\mathbf{c}^T\mathbf{x}^*$ and $\mathbf{c}^T\mathbf{x}$ (e.g. their ratio r_0 is $\leq f$ as well), thus proving that $\mathbf{x}$ is an f-approximation to $\mathbf{x}^*$. **(b)** According to the primal-dual schema, dual and integral-primal feasible solutions make local improvements to each other, until the final costs $\mathbf{b}^T\mathbf{y}^t$, $\mathbf{c}^T\mathbf{x}^t$ are close enough (e.g. their ratio is $\leq f$). We can then apply the primal-dual principle (as in Fig. (a)) and thus conclude that $\mathbf{x}^t$ is an f-approximation to $\mathbf{x}^*$.

For more details about the Fast-PD algorithm, the reader is referred to [12, 13]. Here, we will just provide a brief, high level description of the basic driving force behind that algorithm. This driving force will consist of the *primal-dual schema*, which is a well-known technique in the Linear Programming literature.

3.1 The primal-dual schema for MRF optimization

To understand how the primal-dual schema works in general, we will need to consider the following pair of primal and dual Linear Programs (LPs):

$$\begin{array}{lll} \text{PRIMAL:} & \min \mathbf{c}^T\mathbf{x} & \text{DUAL:} \quad \max \mathbf{b}^T\mathbf{y} \\ & \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} & \text{s.t.} \quad \mathbf{A}^T\mathbf{y} \leq \mathbf{c} \end{array} \tag{16}$$

Here $\mathbf{A}$ represents a coefficient matrix, while $\mathbf{b},\mathbf{c}$ are coefficient vectors. Also, $\mathbf{x}$, $\mathbf{y}$ represent the vectors of primal and dual variables respectively. We seek an optimal solution to the primal program, but with the extra constraint of $\mathbf{x}$ being integral. Due to this integrality requirement, this problem is in general NP-hard and so we need to settle with estimating approximate solutions. A primal-dual f-approximation algorithm achieves that by use of the following principle (illustrated also in Fig. 1(a)):

Primal-Dual Principle 1 *If* $\mathbf{x}$ *and* $\mathbf{y}$ *are integral-primal and dual feasible solutions having a primal-dual gap less than* f*, i.e.:*

$$\mathbf{c}^T\mathbf{x} \leq f \cdot \mathbf{b}^T\mathbf{y}, \tag{17}$$

then $\mathbf{x}$ *is an* f*-approximation to the optimal integral solution* $\mathbf{x}^*$*, i.e.* $\mathbf{c}^T\mathbf{x}^* \leq \mathbf{c}^T\mathbf{x} \leq f \cdot \mathbf{c}^T\mathbf{x}^*$

Based on the above principle, that lies at the heart of any primal-dual technique, the following iterative schema can be used for deriving an f-approximate solution (this schema is also illustrated graphically in Fig. 1(b)):

Primal-Dual Schema 1 *Keep generating pairs of integral-primal, dual solutions* $\{(\mathbf{x}^k,\mathbf{y}^k)\}_{k=1}^t$*, until the elements* $\mathbf{x}^t$*,* $\mathbf{y}^t$ *of the last pair are both feasible and have a primal-dual gap which is less than* f*, i.e. condition* (17) *holds true.*

In order to apply the above schema to MRF optimization, it suffices that we cast the MRF optimization problem as an equivalent integer program. To this end, the following integer programming formulation of (15) has been used as the primal problem:

$$\min \sum_{p\in\mathcal{G}}\sum_{l\in\mathcal{L}} V_p(l)x_p(l) + \sum_{(p,q)\in\mathcal{E}}\sum_{l,l'\in\mathcal{L}} V_{pq}(l,l')x_{pq}(l,l') \tag{18}$$

$$\text{s.t.} \sum_l x_p(l) = 1 \qquad \forall\, p\in\mathcal{G} \tag{19}$$

$$\sum_l x_{pq}(l,l') = x_q(l') \qquad \forall\, l'\in\mathcal{L},\ (p,q)\in\mathcal{E} \tag{20}$$

$$\sum_{l'} x_{pq}(l,l') = x_p(l) \qquad \forall\, l\in\mathcal{L},\ (p,q)\in\mathcal{E} \tag{21}$$

$$x_p(\cdot),\ x_{pq}(\cdot,\cdot)\in\{0,1\}$$

Here, in order to linearize the MRF energy, we have replaced the discrete variables u_p with the binary variables $x_p(\cdot)$ and $x_{pq}(\cdot,\cdot)$. More specifically, the $\{0,1\}$-variable $x_p(l)$ indicates that node p is assigned label l (i.e., $u_p = l$), while the $\{0,1\}$-variable $x_{pq}(l,l')$ indicates that vertices p,q are assigned labels l,l' respectively (i.e., $u_p = l$, $u_q = l'$). Furthermore, the constraints in (19) simply express the fact that each node must receive exactly one label, while constraints (20), (21) maintain consistency between variables $x_p(\cdot), x_q(\cdot)$ and variables $x_{pq}(\cdot,\cdot)$, in the sense that if $x_p(l) = 1$ and $x_q(l') = 1$ holds true, then these constraints force $x_{pq}(l,l') = 1$ to hold true as well (as desired).

The linear programming relaxation of the above integer program is then taken (by relaxing the binary constraints to $x_p(\cdot)\geq 0, x_{pq}(\cdot,\cdot)\geq 0$), and the dual of the resulting LP is used as our dual problem. The Fast-PD algorithm is then derived by applying the primal-dual schema to this pair of primal-dual LPs, while using $f{=}2\frac{d_{\max}}{d_{\min}}$[4] as the approximation factor in (17).

3.2 Advantages of the primal-dual approach

Fast-PD has many nice properties, which makes it a perfect candidate for our image registration task. In particular, it offers the following advantages: **1) Generality:** Fast-PD can handle a very wide class of MRFs, since it merely requires $V_{pq}(\cdot,\cdot)\geq 0$. Hence, by using Fast-PD, our image registration framework can automatically incorporate any similarity metric, as well as a very wide class of smoothness penalty functions. **2) Optimality:** Furthermore, Fast-PD can always guarantee that the generated solution will be an f-approximation to the true optimum (where $f{=}2\frac{d_{\max}}{d_{\min}}$). **3) Per-instance approximation factors:** In fact, besides the above worst-case approximation factor, Fast-PD can also continuously update a *per-instance* approximation factor during its execution. In practice, this factor drops to 1 very quickly, thus allowing the global optimum to be found up to a user/application bound. **4) Speed:** Finally, Fast-PD provides great computational efficiency, since it can reach an almost optimal solution very fast and in an efficient manner.

[4] $d_{\max} \equiv \max_{a\neq b} d(a,b),\ \ d_{\min} \equiv \min_{a\neq b} d(a,b)$

4 Implementation Details & Validation

4.1 Implementation Details

In order to prove our concept, we implemented a non-rigid image registration framework based on discrete optimization. We are using multi-level free-form deformations [14] together with a pyramidal image representation. The deformations are computed on each level in a course-to-fine manner. We define the set of labels for the finest pyramid by setting a minimum and maximum displacement and the number steps. Additionally, the displacements are scaled for the coarser levels to recover a larger deformations. In general, before running our algorithm, we rescale the image intensities of the source and target image to values between 0 and 1. Thus, the weighting of the prior term is less sensitive. In all experiments, we use an empirically determined $\lambda_{pq} = 0.0001$ equally for all grid nodes. In order to demonstrate the flexibility of our framework, we implemented a range of well-known similarity metrics, namely the Sum of Absolute Differences (SAD) [3], the Sum of Squared Differences (SSD) [3], the Normalized Cross Correlation (NCC) [3], the Normalized Mutual Information (NMI) [6], the Correlation Ratio (CR) [8], and the Sum of Absolute Differences plus image gradient information (SADG). The SADG metric involves an intensity-based and a geometric-based term. An additional weighting factor γ is used to control the influence of these two terms. The SADG metric is defined as

$$\begin{aligned}\rho(g(\mathbf{x}), f(\mathcal{T}(\mathbf{x}))) = (1-\gamma)|g(\mathbf{x}) - f(\mathcal{T}(\mathbf{x}))| + \\ + \gamma \arccos\left(\frac{\nabla g(\mathbf{x})}{|\nabla g(\mathbf{x})|} \cdot \frac{\nabla f(\mathcal{T}(\mathbf{x}))}{|\nabla f(\mathcal{T}(\mathbf{x}))|}\right).\end{aligned} \tag{22}$$

4.2 Validation Using Known and Unknown Deformations

In order to evaluate our framework we test our method on several data sets. In general, the evaluation and thus, validation of non-rigid image registration methods is a difficult task. Usually, ground truth data for real deformations, especially, in medical applications is not available. Therefore, we performed several experiments hopefully illustrating the great potentials of our approach.

Realistic Synthetic Registration. The first two experiments are concerning the nature of the free choice of similarity metrics inherent in our framework. In order to evaluate the efficiency of different metrics we test our method on simulated realistic data. The target image is generated from the 2D MRI source image by randomly displaced deformation grid. Additionally, we added uniformly distributed noise up to 15 percent of the original target intensities. For the multi-modal experiment we use the inverse target image and squared intensities. The image resolution is 256x256. The registration is performed using a three-level image and grid pyramid. The range of the set of labels is from 0.25 to 5 pixels in 5 steps for the finest pyramid level defined on the 8 main directions (horizontal, vertical, and diagonal) leading to 41 labels in total (including the zero displacement). We perform 3 optimization cycles per pyramid level. The initial

Metric	AE Mean	AE Median	AE Std	MOD Mean	MOD Median	MOD Std
SSD	2.290	1.143	2.854	0.242	0.139	0.297
SADG $\gamma = 1.0$	1.957	1.077	2.350	0.227	0.136	0.349
SAD	1.220	0.675	1.653	0.123	0.071	0.194
SADG $\gamma = 0.75$	1.129	0.709	1.313	0.122	0.082	0.169
SADG $\gamma = 0.25$	1.046	0.603	1.307	0.104	0.067	0.142
SADG $\gamma = 0.5$	1.036	0.589	1.292	0.111	0.066	0.159
NMI	0.999	0.629	1.060	0.099	0.079	0.080
CR	0.927	0.536	1.116	0.089	0.068	0.092
NCC	0.765	0.402	1.082	0.070	0.047	0.076
CR	2.244	1.039	3.465	0.234	0.110	0.373
NMI	0.846	0.607	0.826	0.086	0.071	0.070

Table 1: Angular error (in degrees) and magnitude of difference error (in pixels) for the realistic synthetic image registration using different similarity metrics.

grid resolution is 6x6 increased to 11x11 and finally 21x21. One registration takes between 5-30 seconds depending on the similarity metric. The results are shown in Table 1. For the evaluation, two error metrics are considered, namely the angular error (AE) [15] and the magnitude of difference (MOD). We only consider the deformation field within a region of interest which is determined by the image mask shown in Fig. 2(d).

Automatic Cartilage Segmentation. Our third experiment is aiming at the registration accuracy. The medical application is similar to the one described in [16]. An automatic segmentation of the cartilage should be performed. Assuming that manual segmentations are available, one may create statistical models for an atlas-based segmentation procedure.

Image	OR	SD In	SD Out	HD
1	0.922	0.079	0.632	1.398
2	0.914	0.068	1.280	3.064
3	0.876	0.089	1.626	3.250
4	0.884	0.081	1.469	3.250
5	0.873	0.099	1.434	3.064
6	0.905	0.070	0.641	1.976

Table 2: Results for the cartilage segmentation experiment.

In our experiment, 7 data sets (256x256x20), all manually segmented by medical experts, are available. The MRI data was acquired for a follow-up experiment. Due to the intra-subject property and the limited number of data sets we simply selected one of it as a template segmentation. By deforming the template to the six other data sets and warping the corresponding segmentations, we are able to achieve a fully automatic segmentation in less than 80 seconds. We use a three-level pyramid, the SAD metric and a set of labels from 0.25 to 5 pixels in 5 steps in the six main directions ($\pm$x, $\pm$y, and $\pm$z) leading to 31 labels in total. We perform 5 optimization cycles per pyramid level. The segmentation results are then compared to the manual segmentations. With our method we achieve an average overlap ratio (OR) of $0.90(\pm 0.02)$, an average surface distance inside (SD In) of $0.08(\pm 0.01)$mm, average surface distance outside (SD Out) of $1.18(\pm 0.44)$mm, and an average Haussdorf distance (HD) of $2.67(\pm 0.79)$mm (which is less than the slice thickness of 3mm). (see also Table 2 and Fig. 3). The comparison of the segmentations is done using the tool[5] described in [17].

[5] Available on http://www.ia.unc.edu/dev/download/valmet/

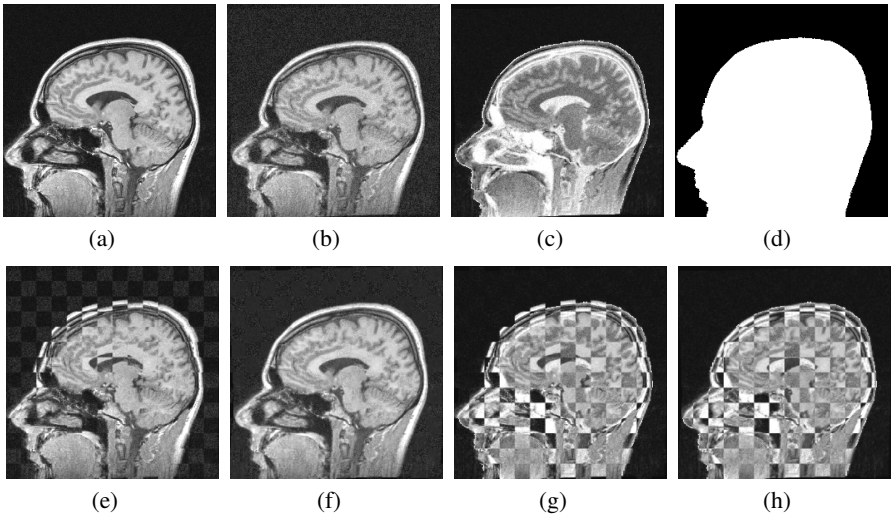
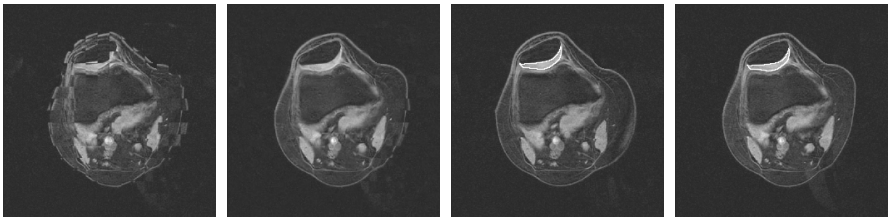


Fig. 2: Realistic synthetic data. **(a)** The source image, and **(b)** target image of the first (mono-modal) experiment. **(c)** Target image of the second (multi-modal) experiment. **(d)** Image mask used for error evaluation. **(e)** Checkerboard visualization before and **(f)** after registration using NCC for the mono-modal experiment. **(g)** Checkerboard visualization before and **(h)** after registration using NMI for the multi-modal experiment.

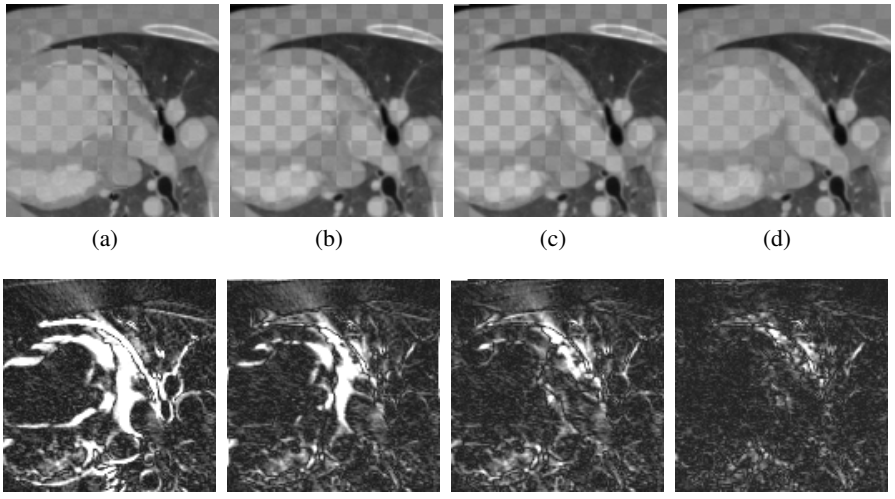
Comparison to State-of-the-art. Schnabel et al. [10] propose a non-rigid image registration method[6] based on B-Spline FFD together with a gradient-descent optimization. In order to obtain meaningful comparable results we try to set the registration parameters as similar as possible. Both algorithms are using the same deformation model and the SSD metric. We use a set of labels from 0.1 to 2 pixels in 5 steps and allow 20 optimization cycles. The test data are two CT volumes showing the heart of a pig. The image resolution is 128x128x88 with a voxel size of 0.848x0.848x1.25mm. Due to the heart beat a deformation of the shape is clearly visible. We run both methods on a deformation grid with 10mm control point spacing. Within the region of interest enclosing the heart and an average SSD error of 12278 before registration, we achieve an average SSD error of 3180, where the other method converges to a value of 3402. Also, by visual perception of the difference images we can achieve better results (see Fig. 4). Last but not least, the running time of our algorithm is less than 2 minutes in contrast to a running time of more than 2 hours for the other method (AMD Athlon64 2.21 GHz). We should note, that this experiment was not performed to obtain the best registration of the two data sets, but rather to compare the two algorithms. With our standard pyramidal approach we obtain a SSD error of 1233 by same running time of about 2 minutes.

[6] Available on http://wwwhomes.doc.ic.ac.uk/~dr/software/



(a) (b) (c) (d)

Fig. 3: Results for the cartilage segmentation of the first image. **(a)** Checkerboard visualization before and **(b)** after registration. **(c)** Template segmentation. **(d)** Warped template on top of the target image.



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 4: **(a)** Checkerboard visualization before registration, **(b)** after registration using the method in [10], and **(c)** after registration using our method **(d)** After registration using our approach with pyramidal settings. Same order for the difference images in **(e)**-**(h)**.

5 Discussion

In this paper we have proposed a novel framework to deformable image registration that bridges the gap between continuous deformations and optimal discrete optimization. Our method reformulates registration using a MRF definition, and recovers the optimal solution to the designed objective function through efficient linear programming. Towards capturing important deformations, we propose an incremental estimation of the deformation component. These objectives are met through a min cut problem defined over a graph with two terminal links. Graph edges introduce smoothness on the deformation field, while edges with the terminal links encode the image support for a given deformation hypothesis versus another. Therefore, the method is gradient free, can encode any similarity metric and can recover the optimal solution up to a bound.

In several applications, building anatomical atlases and models of variations between training examples is feasible. In such a context, one can consider a partial graph

where connections, as well as t-links hypotheses are determined according to the density of expected deformations. Such a direction will introduce prior knowledge in the registration process and will make the optimization step more efficient. Moreover, the use of shape and appearance models can be considered to perform segmentation through registration. Assuming a prior model that involves both geometry and texture, and given a new volume one can define/recover segmentation through the deformation of the model to the image that is a natural registration problem which can be optimally addressed from the proposed framework.

References

1. Hellier, P., Barillot, C.: Coupling dense and landmark-based approaches for nonrigid registration. IEEE Medical Imaging **22**(2) (Feb. 2003)
2. Pennec, X., Ayache, N., Thirion, J.P.: Landmark-based registration using features identified through differential geometry. In: Handbook of medical imaging. (2000) 499–513
3. Hajnal, J., Hill, D.L.G., Hawkes, D.J., eds.: Medical Image Registration. CRC Press (2001)
4. Davatzikos, C., Prince, J., Bryan, R.: Image registration based on boundary mapping (1996)
5. Hermosillo, G., Chefd'hotel, C., Faugeras, O.: Variational methods for multimodal image matching. Int. J. Comput. Vision **50**(3) (2002) 329–343
6. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P.: Multimodality image registration by maximization of mutual information. IEEE Medical Imaging **16**(2) (April 1997) 187–198
7. Zollei, L., Fisher, J., Wells, W.: An Introduction to Statistical Methods of Medical Image Registration. In: Handbook of Mathematical Models in Computer Vision. Springer (2005)
8. Roche, A., Malandain, G., Pennec, X., Ayache, N.: The correlation ratio as a new similarity measure for multimodal image registration. In: Medical Image Computing and Computer-Assisted Intervention. (1998)
9. Rohlfing, T., Maurer, C.R., J., Bluemke, D., Jacobs, M.: Volume-preserving nonrigid registration of mr breast images using free-form deformation with an incompressibility constraint. IEEE Medical Imaging **22**(6) (2003)
10. Schnabel, J.A., Rueckert, D., Quist, M., Blackall, J.M., Castellano-Smith, A.D., Hartkens, T., Penney, G.P., Hall, W.A., Liu, H., Truwit, C.L., Gerritsen, F.A., Hill, D.L.G., Hawkes, D.J.: A generic framework for non-rigid registration based on non-uniform multi-level free-form deformations. In: Medical Image Computing and Computer-Assisted Intervention. (2001)
11. Li, S.Z.: Markov random field modeling in image analysis. Springer-Verlag New York, Inc. (2001)
12. Komodakis, N., Tziritas, G., Paragios, N.: Fast, approximately optimal solutions for single and dynamic mrfs. In: Computer Vision and Pattern Recognition. (2007)
13. Komodakis, N., Tziritas, G.: A new framework for approximate labeling via graph cuts. In: IEEE International Conference on Computer Vision. (2005)
14. Forsey, D.R., Bartels, R.H.: Hierarchical b-spline refinement. SIGGRAPH Comput. Graph. **22**(4) (1988) 205–212
15. Fleet, D.J.: Measurement of Image Velocity. Kluwer Academic Publishers, Norwell, MA, USA (1992)
16. Folkesson, J., Dam, E., Olsen, O.F., Pettersen, P., Christiansen, C.: Automatic segmentation of the articular cartilage in knee mri using a hierarchical multi-class classification scheme. In: Medical Image Computing and Computer-Assisted Intervention. (October 2005)
17. Gerig, G., Jomier, M., Chakos, M.: Valmet: A new validation tool for assessing and improving 3d object segmentations. In: Medical Image Computing and Computer-Assisted Intervention. (2001)