Non-rigid Registration using Discrete MRFs: Application to Thoracic CT Images

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

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

Abstract. We recently introduced labeling of discrete Markov random fields (MRFs) as an attractive approach for non-rigid image registration. Our MRF framework makes use of recent advances in discrete optimization, is efficient in terms of computation time, and provides great flexibility. Any similarity measure can be encoded right away, since no differentiation is needed. In this work, we investigate the performance of our framework in a challenging scenario: the registration of thoracic CT images. In order to assess the potential of the discrete MRF setting, we employ the simplest registration objective function based on intensity differences. The registration is fully-automatic, constant parameters are used throughout the experiments, we omit the use of the available segmentations, and (except for linear pre-alignment) no pre-processing of the data is performed. Despite the simplicity of our experimental setup, we are able to obtain accurate registration results for most of the data in a very efficient manner. Our registration software is freely available[1].

1 Non-rigid Registration using Discrete MRFs

Non-rigid image registration is an important task in computer vision and medical imaging applications. Given two images I and J, one seeks a transformation T which aligns corresponding objects visible in the images. This task is commonly solved by posing an energy minimization problem where the objective function is a sum of a matching criterion S and a regularization term R,

$$\hat{T} = \arg\min_{T} \; S(I, J \circ T) + \alpha \, R(T) \quad . \tag{1}$$

Here, α is a weighting factor controlling the influence of the regularization term. In case of non-rigid registration, the transformation is often defined as the identity transformation plus a dense displacement field D. The new location of an image point p is then computed by

$$T(p) = p + D(p) \quad . \tag{2}$$

Recently, labeling of discrete Markov random fields has become an attractive approach for non-rigid image registration [1–4]. In the following, we will briefly introduce our general framework, originally presented in [5].

[1] http://www.mrf-registration.net/

1.1 Discrete Markov Random Fields

Let us consider a Markov random field $\mathbf{X}$ which is a set of n random variables $X_i \in \mathbf{X}$. Each variable can take a value $x_i \in L$ where L is a discrete set of events. The events are commonly referred to as labels, and $X_i = x_i$ is a label assignment of variable X_i. If every variable is assigned a label, this is what we call a *labeling* of the field denoted by $\mathbf{X} = \mathbf{x}$ with $\mathbf{x} = (x_1, ..., x_n)$. Sometimes the labeling is also referred to as the *configuration* or *realization* of the field. We further introduce the concept of neighborhood. Neighboring random variables are conditionally dependent, and by defining a neighborhood system we can encode spatial interactions and contextual constraints between variables. Neighborhood relationship is defined via so called *cliques*. A clique is a subset of variables $C \subseteq \mathbf{X}$ and variables within the same clique are neighbors. The set containing all cliques is denoted by $\mathbf{C}$. We further define the *order* of a field as the size of the maximal clique minus one. A first-order field contains cliques of size up to two, a second-order field has cliques of size up to three, and so on.

As a consequence of the famous theorem by Hammersley and Clifford [6], we can define the energy of a labeling as

$$E(\mathbf{x}) = \sum_{C \in \mathbf{C}} \psi_C(\mathbf{x}_C) \quad . \tag{3}$$

The functions ψ_C are the so called *potential functions*, where one function is defined per clique. The energy of an MRF is simply the sum over clique potentials. The definition of these functions is an essential part of modeling a problem as a random field. The potentials are unrestricted real-valued functions which evaluate the quality of sub-labelings $\mathbf{x}_C$ in terms of energies. The lower the energy the more likely the labeling.

In this work, we will focus on first-order fields for which very efficient optimization methods are available. The energy of an first-order MRF is simply

$$E(\mathbf{x}) = \sum_{i} \psi_i(x_i) + \sum_{(i,j)} \psi_{ij}(x_i, x_j) \quad . \tag{4}$$

Of course, we are interested in the configuration with the lowest energy (which is equivalent to the *maximum a posteriori* (MAP) estimate which is the labeling with highest probability). Mathematically, we are interested in the following minimization problem

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} E(\mathbf{x}) \quad . \tag{5}$$

Now, the interesting task is how to model a real world problem, such as the non-rigid registration, in terms of a discrete random field formulation. This task involves four steps: (i) identification of the role of random variables, (ii) definition of the discrete label space, (iii) derivation of the energy function, and (iv) selection of an appropriate optimization strategy.

1.2 Bridging the Gap

Let us start by discussing the role of the random variables. The ultimate goal of non-rigid registration is the recovery of an optimal non-linear transformation T. This is expressed in Equation (1). The transformation is fully defined by the identity mapping and a dense displacement field D (cf. Equation (2)). So, what we need to determine are the displacement vectors $D(p)$ for every image point p. Intuitively, we could introduce a random variable for every image point. Then, the labels x_i correspond to displacements. However, such an approach is doomed in practice. Considering a 3D registration scenario with moderately sized images, let's say 256^3 voxels, this would result in a random field with more than 16.7 millions of variables. Optimization of such huge fields is quite inefficient. We propose to reduce the dimensionality of the problem by introducing a deformation grid with a sparse set of control points and an interpolation strategy. We introduce a random variable for every control point, and the labels correspond to the control point displacements. The dense displacement field is then defined as the linear combination of the control point displacements

$$D(p) = \sum_i \omega_i(p)\, x_i \quad . \tag{6}$$

Here, $\omega_i(p)$ corresponds to the interpolation or weighting function. In this work, we employ free form deformations (FFDs) based on cubic B-splines [7, 8]. In common FFDs the control points are uniformly distributed over the image domain. The displacement of a control point has only local influence on the displacement field, and thanks to the cubic basis functions the resulting deformation is guaranteed to be C^2 continuous.

Let us now derive the MRF energy function for the registration problem. As already mentioned, in this work we are focusing on first-order MRFs, so the energy consists of *unary* potential functions ψ_i and *pairwise* potential functions ψ_{ij}. The unaries correspond to the energy for assigning a label to a random variable, independently of all other variables. In contrast, the pairwise terms evaluate the energy for a simultaneous label assignment to two variables. We encode the matching criterion S from Equation (1) locally distributed on the unary terms. For the sum of absolute intensity differences (SAD), which is the matching criterion we consider in this work, the unary terms are defined as

$$\psi_i(x_i) = \sum_{p \in \Omega_i} \hat{\omega}_i(p)\, |I(p + D(p) + x_i) - J(p)| \quad . \tag{7}$$

Here, Ω_i defines a local domain centered at the control point i. Only image points p within this domain are considered for the local matching term. We define this local domain to compass only the immediate neighboring cells of the regular control grid. The function $\hat{\omega}_i$ is again a weighting function. Image points closer to a control point will have a larger contribution to the energy than points further away. This is similar to the weighting function for the control grid earlier introduced (cf. Equation (6)). In contrast to the cubic B-spline function for the

deformation, here we commonly use a simple linear weighting function. Note, that the matching term already considers the displacement field D. This is later needed in the hierarchical, iterative procedure.

In fact, the use of FFDs implicitly guarantees smooth deformations. However, additional explicit regularization can be beneficial, in particular in image areas with less visible structure. We employ a simple neighborhood system on the random variables, where every variable forms pairwise cliques with its immediate neighbors (i.e. 6-connected neighborhood in 3D). We encode the regularization term R from Equation (1) locally distributed on the pairwise terms. Here we consider a simple penalty term, which favors similar displacements between neighboring control points and which is an approximation for penalizing the first derivatives of the displacement field

$$\psi_{ij}(x_i, x_j) = \alpha \frac{\|(d_i + x_i) - (d_j + x_j)\|}{\|i - j\|} \quad . \tag{8}$$

Again, we also consider displacements d_i and d_j which might have been occurred in previous iterations. The factor α controls the influence of the regularization.

Both, our matching term in Equation (7) and the regularization term in Equation (8) are extremely simple and easy to implement. Still, we will later see that we are able to obtain quite good registration results with this formulation. Of course, our framework allows to consider much more sophisticated energies. Different matching criteria, regularization terms, learned neighborhood systems and deformation priors have been presented in [1, 9, 10].

1.3 Refinement Strategy

Two of four steps are still missing in our formulation. So far we did not discuss the definition of the label space and the selection of the optimization strategy. We continue with the set of labels. Remember, the labels correspond to displacements and thus the set L defines our solution space for the registration. Of course, the random variables are in fact of continuous nature and their range of values is actually $\mathcal{R}^3$ (in 3D registration). So, how can we efficiently sample this space and form a set of discrete labels? It is important to find a good compromise. On the one hand, a small number of labels allows fast optimization. On the other hand, a too sparse sampling of the displacement space may lead to inaccurate registration.

To this end, we propose an iterative refinement strategy which allows us to keep the number of labels small, but also to achieve high-accurate results. We employ a sparse sampling of the displacement space with a fixed number of samples k. We uniformly sample along the positive and negative directions of the three coordinate axes up to a maximum value $d_{\max}$. The total number of labels is then $|L| = 6 \cdot k + 1$ including the zero-displacement. The registration is then performed iteratively. In every iteration, we try to determine the optimal update for the displacement field by solving a discrete labeling on the current set of labels and w.r.t. to the current deformation. After each iteration, the update

is added to the displacement field and the label space is refined by rescaling the displacements in L by a factor $0 < s < 1$. The whole procedure is sketched in pseudo-code in the following.

Algorithm 1: Non-rigid Registration using Discrete MRFs

input : Images I, J
output: Transformation T

```
T ← initializeTransformation() ;
L ← initializeLabelSpace() ;
repeat
    x ← computeLabeling(I,J,T,L) ;
    T ← updateTransformation(T,x) ;
    L ← refineLabelSpace() ;
until convergence;
```

We should note, that of course the update of the transformation is only performed if the energy has decreased. The convergence criterion can be either based on the change of the energy – if the change is very little and the label space has been refined sufficiently often, we can stop – or we simply set a maximum number of iterations. The latter on has been used in this work. For the MRF optimization, we use a recently proposed method called FastPD [11, 12] which we also used in [5, 1, 9, 10]. Due to the limited space, we refer the reader to the given references for more details about the algorithm.

2 Experiments on Empire Database

Our experiments are part of the MICCAI 2010 Grand Challenge "Evaluation of Methods for Pulmonary Image Registration" (EMPIRE). The EMPIRE database consists of 20 pairs of thoracic CT images. A full description of the data and the evaluation of the registration results can be found in [13] and on the website[2] of the EMPIRE study.

In the following, we will describe our experimental setup and the setting of the internal parameters of our registration method. All parameters have been empirically determined by running a few test registrations on one of the image pairs and visual inspection of the results. The parameters were then fixed throughout the experiments. Our registration is fully-automatic and does not require any pre-processing of the image data. The maintainers of the database provide segmentations of the lung areas for all images. We decided to omit the use of these segmentations and thus do not rely on any segmentation algorithm.

[2] http://empire10.isi.uu.nl/

2.1 Linear Pre-Alignment

Prior to the non-rigid registration, we employ a linear alignment using a Downhill-Simplex optimization and the SAD matching criterion. This step roughly recovers a transformation containing translation, rotation, and anisotropic scaling. To speed up the computation, random sampling with 10% of the voxels is used here. We also use a common Gaussian image pyramid with four resolution levels where only the two coarsest levels are considered in the pre-alignment step. In our image pyramid, we do not perform any downsampling along a dimension where the number of voxels would drop below 32.

2.2 Non-rigid Alignment

The final non-rigid alignment is performed with a hierarchical setting. Again, we use a Gaussian image pyramid, this time with five resolution levels and the preservation of the 32 voxel limit. Only the four coarsest levels are considered for the registration to speed up the computation. Simultaneously, we use a four level FFD control grid, starting with a control point spacing of approximately 95mm and ending with a spacing of about 10mm. On each grid level, we perform 5 iterations, which means we compute 5 discrete labelings each corresponding to an update of the displacement field. After these 5 iterations we continue the registration on the next higher resolution level, both for the control grid and the image pyramid. The labels space is defined as follows: we initialize the maximum displacement $d_{\max}$ on each level with 0.4 times the control point spacing. The number of samples k is set to 5, so in every iteration we have in total 31 labels (including the zero-displacement). After each iteration the displacements in L are scaled by a factor $s = 0.67$. Finally, the weighting parameter α is set to 20.

2.3 Results

After registration of all 20 pairs, we submitted our displacement fields to the maintainers of the EMPIRE study who did the evaluation remotely. Our results are summarized in Table 1. More details are given on the EMPIRE website under the results section, where our method can found under the name *drop*. We make several observations from which we conclude that our method is performing very well on most of the datasets. Observation one, the lung boundaries are almost perfectly registered in all datasets. Observation two, the alignment of major fissures is very good in 13 of 20 datasets. In the other 7 (datasets 1,7,8,11,14,18,20) the alignment exhibits larger errors. Observation three, also the alignment of the landmarks is very accurate. 11 of 20 datasets exhibit an average error of less than 2mm, in 8 datasets it is even below 1mm. This is quite remarkable, considering that we did not perform the registration on the full image resolution. Again, the before mentioned 7 datasets exhibit the largest errors in the landmark category. The last observation is that in all cases, our displacement fields are free of any singularities such as folding or tearing. This is a pleasant behavior of our (simple) regularization term. We should note, that in this work we did not employ any geometrical constraints on the FFDs to guarantee bijective deformations.

All our experiments were carried out on an Intel® Core™2 Duo 2.16GHz processor and our software is implemented in C++. The time taken for the pre-alignment step is between 3 and 30 seconds. For the non-rigid part we measured running times between 20 and 300 seconds, depending on the size of the images (which is varying approximately between 5 and 70 millions of voxels).

While we believe that our registration results are quite convincing for most of the datasets, in particular when considering the simple objective function and the limit on the image resolution, we also tried to find out the reason for the 7 outliers. The 20 datasets exhibit quite different settings in the image acquisition. The 7 datasets where we perform worst are the only ones where one image has been taken in breath-hold inspiration and the other one in breath-hold expiration. Additionally, the inspiration scan was acquired with low dose CT and the expiration scan with ultra-low dose. Probably the larger amount of deformation from inspiration to expiration yields larger errors in the alignment, but also the varying intensities between the low dose and ultra-low dose scans might be a problem for the simple SAD matching criterion. This should be investigated in more detail in follow-up experiments with different internal parameters. A different matching criterion such as the correlation coefficient and higher resolutions of the images and FFD control grids might improve the overall accuracy of the registration.

References

1. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through MRFs and efficient linear programming. Medical Image Analysis **12** (2008) 731–741
2. Shekhovtsov, A., Kovtun, I., Hlavac, V.: Efficient MRF deformation model for non-rigid image matching. Computer Vision and Image Understanding **112** (2008) 91–99
3. Lee, K.J., Kwon, D., Yun, I.D., Lee, S.U.: Deformable 3d volume registration using efficient MRFs model with decomposed nodes. In: British Machine Vision Conference. (2008)
4. Kwon, D., Lee, K.J., Yun, I.D., Lee, S.U.: Nonrigid image registration using dynamic higher-order MRF model. In: European Conference on Computer Vision. (2008)
5. 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. (2007)
6. Hammersley, J., Clifford, P.: Markov fields on finite graphs and lattices. Unpublished Manuscript (1971)
7. Sederberg, T.W., Parry, S.R.: Free-form deformation of solid geometric models. ACM SIGGRAPH **20** (1986) 151–160
8. 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 **18** (1999) 712–721
9. Glocker, B., Komodakis, N., Navab, N., Tziritas, G., Paragios, N.: Dense registration with deformation priors. In: Information Processing in Medical Imaging. (2009)

	Lung Boundaries		Fissures		Landmarks		Singularities	
Scan Pair	**Score**	**Rank**	**Score**	**Rank**	**Score**	**Rank**	**Score**	**Rank**
01	0.02	17.00	5.55	23.00	6.75	23.00	0.00	11.50
02	0.00	11.00	0.00	15.00	0.46	15.00	0.00	12.50
03	0.00	11.00	0.00	28.00	0.57	19.00	0.00	12.00
04	0.00	21.00	0.00	16.50	2.83	26.00	0.00	14.00
05	0.00	13.00	0.00	16.00	0.06	19.00	0.00	13.50
06	0.00	16.00	0.00	28.00	0.40	18.00	0.00	14.00
07	0.00	8.00	6.72	26.00	7.48	27.00	0.00	10.00
08	0.00	11.00	5.70	30.00	3.75	28.00	0.00	12.50
09	0.00	16.00	0.01	26.00	0.67	20.00	0.00	13.00
10	0.01	20.00	0.00	15.00	3.93	23.00	0.00	13.50
11	0.05	17.00	5.46	31.00	2.90	29.00	0.00	11.50
12	0.00	10.00	0.00	13.50	0.78	25.00	0.00	14.50
13	0.00	13.00	0.10	18.00	1.20	22.00	0.00	13.00
14	0.01	8.00	7.71	24.00	8.94	26.00	0.00	9.50
15	0.00	8.00	0.00	27.00	0.68	16.00	0.00	12.50
16	0.00	9.00	0.12	22.50	1.58	24.00	0.00	13.50
17	0.00	6.50	0.07	31.00	1.21	25.00	0.00	14.00
18	0.00	8.00	7.18	27.00	6.13	25.00	0.00	10.50
19	0.00	14.00	0.00	12.00	0.49	11.00	0.00	14.50
20	0.00	13.00	10.93	29.00	9.77	25.00	0.00	10.50
Avg	0.00	12.52	2.48	22.92	3.03	22.30	0.00	12.52
Average Ranking Overall								17.56
Final Placement								19

Table 1. Results for each scan pair, per category and overall. Rankings and final placement are from a total of 34 competing algorithms.

10. Glocker, B., Komodakis, N., Paragios, N., Navab, N.: Approximated curvature penalty in non-rigid registration using pairwise MRFs. In: International Symposium on Visual Computing. (2009)
11. Komodakis, N., Tziritas, G., Paragios, N.: Fast, approximately optimal solutions for single and dynamic MRFs. In: Computer Vision and Pattern Recognition. (2007)
12. Komodakis, N., Tziritas, G., Paragios, N.: Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies. Computer Vision and Image Understanding **112** (2008) 14–29
13. Murphy, K., van Ginneken, B., Reinhardt, J., Kabus, S., Ding, K., Deng, X., Pluim, J.: Evaluation of methods for pulmonary image registration: The EMPIRE10 study. Grand Challenges in Medical Image Analysis (2010)