

Non-rigid 2D-3D Medical Image Registration using Markov Random Fields

Enzo Ferrante, Nikos Paragios

▶ **To cite this version:**

Enzo Ferrante, Nikos Paragios. Non-rigid 2D-3D Medical Image Registration using Markov Random Fields. 16th International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI 2013, Sep 2013, Nagoya, Japan. hal-00855662v1

HAL Id: hal-00855662

https://hal.science/hal-00855662v1

Submitted on 30 Aug 2013 (v1), last revised 7 Nov 2013 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Non-rigid 2D-3D Medical Image Registration using Markov Random Fields

Enzo Ferrante and Nikos Paragios

Center for Visual Computing, Ecole Centrale de Paris, France
GALEN Team, INRIA Saclay, France
{enzo.ferrante,nikos.paragios}@ecp.fr

Abstract. The aim of this paper is to propose a novel mapping algorithm between 2D images and a 3D volume seeking simultaneously a linear plane transformation and an in-plane dense deformation. We adopt a metric free locally over-parametrized graphical model that combines linear and deformable parameters within a coupled formulation on a 5-dimensional space. Image similarity is encoded in singleton terms, while geometric linear consistency of the solution (common/single plane) and in-plane deformations smoothness are modeled in a pair-wise term. The robustness of the method and its promising results with respect to the state of the art demonstrate the extreme potential of this approach.

Keywords: 2D-3D registration, medical imaging, markov random fields, discrete optimization.

1 Introduction

2D-3D image registration is an important problem in medical imaging and it can be applied in multiple medical procedures. Depending on the technology used to capture the 2D image, it can be a projective (e.g. X-Ray) or sliced (e.g. Ultrasound) image; in this work we focused on sliced images. Image guided surgeries, as laparoscopic or endoscopic [1], and brain surgeries [2] use such images. In those surgeries, pre-operative 3D images (e.g. Computed Tomography (CT) or Magnetic Resonance Images (MRI)) and intra-operative 2D images are used to guide surgeons during the procedure. 2D-3D registration plays a key role in this process because it allows doctors to guide surgery using the 3D pre-operative high resolution annotated data. Tissue shift, as well as breathing and heart motion during the surgery, causes elastic deformation in the images and makes the registration process an extremely challenging problem.

Several methods to deal with slice-to-volume registration have been proposed. [1] proposes a method to register endoscopic and laparoscopic ultrasound (US) images with pre-operative computed tomography volumes that potentially could work in real time. It is based on a new phase correlation technique called LEPART and it manages only rigid registration. [3] presents a flexible framework for intensity based slice-to-volume nonrigid registration algorithms used to register histological section images to human brain MRI. A feature based method that performs slice to volume registration is presented in [4]. It uses several slices in order to improve the quality of the results. [5] tracks intra-operative MRI slices of prostate images with a pre-operative MRI volume.

This monomodal registration is designed to provide patient tracking information for prostate biopsy performed under MR guidance.

Discrete optimization of Markov Random Fields (MRFs) has been widely used to solve the problem of non-rigid image registration in recent years [6, 7]. However, to the best of our knowledge, most of those works focus on 2D-2D or 3D-3D registration instead of 2D-3D registration. [8] presents a method based on MRFs to perform 2D-3D registration, but it estimates only rigid transformations and works with projective images.

The main contributions of this paper consists of a local pair-wise method to register 2D and 3D images using MRFs and discrete optimization techniques capable of capturing the plane and the in-plane dense deformations. It is intensity based and independent of the metric that is being used, so it can be adapted to different image modalities.

The remainder of the paper is organized as follows: the MRF based formulation of 2D-3D deformable registration is described in Section 2. Section 3 provides the validation tests and results, while Section 4 concludes our paper and provides some ideas on relevant future directions.

2 2D-3D Non-rigid Registration Using MRFs

The problem of non-rigid 2D-3D image registration can be formulated as an optimization problem. Given a 2D source image I and a 3D target volume J, we seek the 2D-2D in-plane local deformation field $\hat{T}_D$ and the plane $\hat{\pi}[J]$ (i.e. a bi-dimensional slice from the volume J) which in the most general case minimize the following objective function:

$$\hat{T}_D, \hat{\pi} = \underset{T_D, \pi}{\operatorname{argmin}} \mathcal{M}(I \circ T_D(\boldsymbol{x}), \pi[J](\boldsymbol{x})) + \mathcal{R}(T_D, \pi), \tag{1}$$

where $\mathcal{M}$ represents the data term and $\mathcal{R}$ the regularization term. The data term $\mathcal{M}$ measures the matching quality between the deformed 2D source image and the corresponding 3D slice. The regularization term $\mathcal{R}$ imposes certain constraints on the final solution that can be used to render the problem well posed and imposes certain expected geometric properties on the deformation field. The plane, $\hat{\pi}$, that minimizes the equation indicates the location of the 3D volume slice that is most similar to the deformed source image. The deformation field $\hat{T}_D$ represents the in-plane deformations that must be applied to the source image in order to minimize the energy function.

Our MRF based formulation of the 2D-3D non-rigid registration problem consists of an undirected pair-wise graph $G =< V, E >$ super-imposed to the 2D image with a set of nodes V and a set of cliques E. The nodes (a regular lattice) are interpreted as control points of a bi-dimensional quasi-planar grid that models at the same time the in-plane deformations and the current position of the 2D image into the 3D volume. In order to represent the in-plane deformations, the grid is interpreted as a Free Form Deformation model (FFD) where each control point has local influence on the deformation.

Such a coupled estimation problem can be expressed using graphical models of varying complexity. The most natural one is to consider a two layer graph, one modeling the global linear mapping as done in [8] and another modeling the in-plane deformation as done in [6] where interconnection between them will produce coupling

and global consistency on the obtained solution. Despite the theoretical soundness of such an approach, it is not suitable simply because the linear mapping is a global variable. This is due to the fact that global linear mapping variables are to be associated with all control nodes resulting in a densely connected graph, while at the same time the parameter space is high-dimensional and the associated variables are not bounded resulting in rather complex discrete sampling requirements. The aforementioned limitations make such an approach problematic in terms of the quality of the obtained solution (non-submodular terms) and inefficient from a computational view point.

We propose to overcome these limitations through a local pair-wise over-parameterized graphical model. In our formulation, the energy is formed by data terms $\boldsymbol{g} = \{g_i(\cdot)\}$ (unary potentials) associated to each graph vertex and regularization terms $\boldsymbol{f} = \{f_{ij}(\cdot,\cdot)\}$ (pairwise potentials) associated to the edges. The first ones are typically used for encoding some sort of data likelihood, whereas the later ones act as regularizers and thus play an important role in obtaining high-quality results [6]. The minimization energy problem in the context of a discrete MRF is thus defined as:

$$MRF(\boldsymbol{g}, \boldsymbol{f}) = \min \sum_{i \in V} g_i(\boldsymbol{u_i}) + \sum_{(i,j) \in E} f_{ij}(\boldsymbol{u_i}, \boldsymbol{u_j}), \tag{2}$$

where $\boldsymbol{u_i}, \boldsymbol{u_j} \in L$ are the labels assigned to the vertices $v_i, v_j \in V$ respectively.

Vertices $v_i \in V$ are moved by assigning them different labels $\boldsymbol{u_i} \in L$ (where L is the label space) until an optimal position is found. To reach such an optimal position, we need to define an energy term that will be minimized using an optimization algorithm. We adopt the FastPD algorithm [9] for the optimization of the aforementioned MRF due to good trade-off between complexity and quality of the obtained minimum in the context of non-submodular MRFs (this is our case due to the definition of the pair-wise terms).

Label Space It includes all the possible values that a vertex label can take to deform the graph. Our label space L consists of 5-tuples $\boldsymbol{u} = (d_x, d_y, d_z, \phi, \theta)$, where the first three parameters define a displacement vector $\boldsymbol{d_i}$ in the cartesian coordinate system, and the angles (ϕ, θ) define a vector $\boldsymbol{N_i}$ over a unit sphere, expressed using spherical coordinates. Let us say we have a control point $\boldsymbol{p_i^t} = (p_{xi}^t, p_{yi}^t, p_{zi}^t)$ at optimization step t and we assign the label $\boldsymbol{u_i} = (d_{xi}, d_{yi}, d_{zi}, \phi_i, \theta_i)$ to this point. So, the new point position at optimization step $t+1$ will be calculated using the displacement vector resulting in $\boldsymbol{p_i^{t+1}} = (p_{xi}^t + d_{xi}, p_{yi}^t + d_{yi}, p_{zi}^t + d_{zi})$. Additionally, we define a plane π_i containing the displaced control point $\boldsymbol{p_i^{t+1}}$ and whose unit normal vector (expressed in spherical coordinates and with constant radius $r = 1$) is $\boldsymbol{N_i} = (\phi_i, \theta_i)$. One of the most important constraints to be considered is that our graph should have a quasi-planar structure, i.e. it should be similar to a plane; the plane π_i associated to every control point $\boldsymbol{p_i}$ will be used by the energy term to take this constraint into account.

Unary Potentials The formulation of the unary potentials that we propose is independent of the similarity measure. It is calculated for each control point given any intensity based metric δ capable of measuring the similarity between two bi-dimensional images

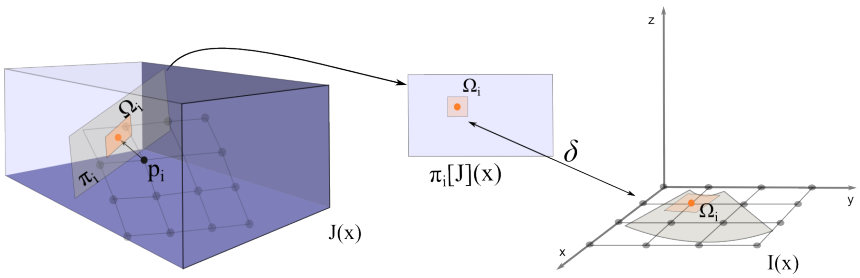


Fig. 1. Data term (Unary potential). The points $\boldsymbol{x} \in \Omega_i$ are used to calculate the unary potential. $\pi[J](\boldsymbol{x})$ returns the intensity of the point in the 2D slice corresponding to the plane π_i in the 3D image, whereas $I(\boldsymbol{x})$ returns the 2D image intensity. δ represents the similarity measure.

(e.g sum of absolute differences, mutual information, normalized cross correlation). This calculation is done for each control point $\boldsymbol{p_i}$, using its associated plane π_i in the target image J and the source 2D image I. An oriented patch Ω_i over the plane π_i (centered in $\boldsymbol{p_i}$) is extracted from the volume J, so that the metric δ can be calculated between that patch and the corresponding area over the source 2D image (see Figure 1):

$$g_i(\boldsymbol{u_i}) = \int_{\Omega_i} \delta(I(\boldsymbol{x}) - \pi_i[J](\boldsymbol{x}))d\boldsymbol{x}. \tag{3}$$

In monomodal scenarios, where two images of the same modality are compared, the simplest and the most used similarity measure is the Sum of Absolute Differences (SAD). In multimodal scenarios, where different modalities are compared (e.g. CT with US images), statistical similarity measures such as Mutual Information (MI) are generally used since we can not assume that corresponding objects have the same intensities in the two images. So, depending on the type of image that we want to register, the framework can be adapted using any similarity measure defined over two bi-dimensional images.

Pairwise Potentials Generally, these terms are used to encode the regularization of the displacement field. In our formulation, the pairwise potentials are defined by two terms: the first one (F_1) controls the grid deformation assuming that it is a plane, whereas the second one (F_2) maintains the plane structure of the mesh. Those terms are weighted by a coefficient $\alpha \in [0, 1]$ resulting in $f_{ij}(\boldsymbol{u_i}, \boldsymbol{u_j}) = \alpha F_1(\boldsymbol{u_i}, \boldsymbol{u_j}) + (1 - \alpha)F_2(\boldsymbol{u_i}, \boldsymbol{u_j})$.

The in-plane grid deformation is thus controlled using a distance preserving approach: it tries to preserve the original distance between the control points of the grid. Since this metric is based on the euclidean distance between the points, it assumes that they are coplanar. So, the equation that regularizes the in-plane deformations is:

$$F_1(\boldsymbol{u_i}, \boldsymbol{u_j}) = 1 - \frac{||\ (\boldsymbol{p_i} + \boldsymbol{d_i}) - (\boldsymbol{p_j} + \boldsymbol{d_j})\ ||}{||\ (\boldsymbol{p_{o,i}}) - (\boldsymbol{p_{o,j}})\ ||}, \tag{4}$$

where $p_{o,i}$ and $p_{o,j}$ are the original positions of the control points. Regarding the plane preservation regularization term, it penalizes the average distance between the control

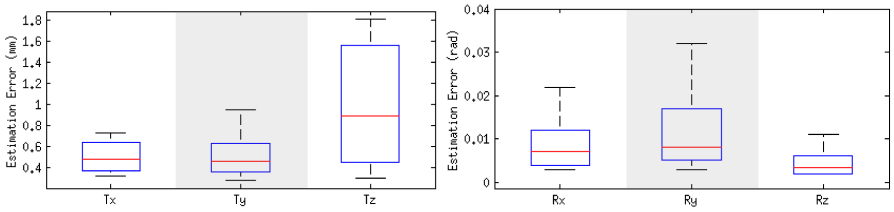


Fig. 2. Average of absolute differences between the ground truth and estimated plane translation (T_x, T_y and T_z) and rotation (R_x, R_y and R_z) parameters for 10 sequences of 20 images each one. The average error is less than $0.013rad$ for rotations and less than 1mm for translations.

points and the plane corresponding to the neighboring one. The aim is to maintain the quasi-planar structure of the grid. Given that the distance between a point and a plane is zero when the point is contained in the plane, this term will be minimum when both of the control points are on the same plane. The term F_2 is then defined using the distance between a point $\boldsymbol{p} = (p_x, p_y, p_z)$ and a plane π given by $D_\pi(\boldsymbol{p})$. So, we calculate the average of $D_{\boldsymbol{\pi_j}}(\boldsymbol{p_i} + \boldsymbol{d_i})$ and $D_{\boldsymbol{\pi_i}}(\boldsymbol{p_j} + \boldsymbol{d_j})$:

$$F_2(\boldsymbol{u_i}, \boldsymbol{u_j}) = \frac{1}{2}(D_{\pi_j}(\boldsymbol{p_i} + \boldsymbol{d_i}) + D_{\pi_i}(\boldsymbol{p_j} + \boldsymbol{d_j})). \tag{5}$$

3 Validation & Results Discussion

A dataset was created in order to validate both the resulting 2D-2D deformation field and the final plane location using a temporal series of 3D heart MRI. The monomodal dataset consists of a temporal series of twenty 2D slices, I_i, each one extracted from its corresponding volume M_i. Starting from a random initial translation $T_0 = (T_{x_0}, T_{y_0}, T_{z_0})$ and rotation $R_0 = (R_{x_0}, R_{y_0}, R_{z_0})$, we extracted a 2D slice I_0 from the initial volume M_0. Gaussian noise was added to every parameter in order to generate the position used to extract the next slice from the next volume. We used $\sigma_r = 3^\circ$ for the rotation and $\sigma_t = 5mm$ for the translation parameters. Those parameters generate maximum distances of around 25mm between the current and its succeeding plane. In that way, we generated a series that corresponds to a trajectory into the volume. Since the series consists of temporally spaced volumes of the heart, there are local deformations between them due to the heartbeat.

We tested it over 10 sequences of 20 images to validate the estimated plane locations, giving a total of 200 registration cases, using SAD similarity measure. The MRI resolution was $192 \times 192 \times 11$ voxels and the voxel size was $1.25 \times 1.25 \times 8mm^3$. The resulting position of the slice I_i was used to initialize the slice I_{i+1}. As shown in Figure 2, the average error is less than 0.013rad (0.74°) for rotation and less than 1mm for translation parameters. Given that the image resolution in z axis is lower than in x and y, we can recognize a bigger error in the estimated translation for z coordinate.

In order to validate the in-plane deformations, we created a set of manual segmentations $\mathcal{S}$ of the left endocardium from a set of 20 slices (extracted from the volumes M_i).

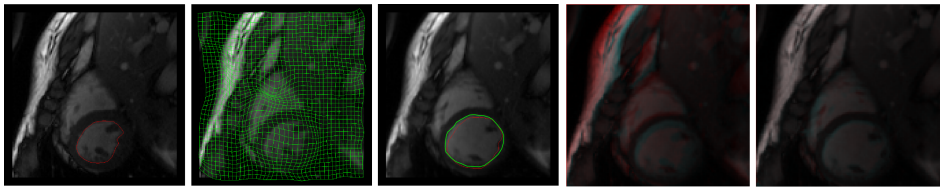


Fig. 3. One of the slices used in the validation test of the in-plane deformations: (a) Source image. The red line is the initial segmentation of the left endocardium. (b) Deformation field obtained after the registration process. (c) Slice corresponding to the estimated plane and extracted from the target volume M_0. The red line corresponds to the deformed initial segmentation and the green line is the ground truth. (d) Overlapping between source image (red) and target initial plane (cyan). (e) Overlapping between deformed source image (red) and estimated plane (cyan).

Each slice was registered with the volume M_0 starting from a random position around the ground truth. Positions were generated adding gaussian noise with $\sigma_r = 4.5^\circ$ and $\sigma_t = 5mm$ to every translation (T_x, T_y, T_z) and rotation (R_x, R_y, R_z) parameters respectively. Those parameters generate maximum distances of about 25mm between the initialization and the ground truth. The estimated deformation field T_{D_i} was applied to the corresponding initial segmentation $s_i \in \mathcal{S}$ and it was compared with the ground truth using DICE coefficient. The average DICE coefficient for the 20 test cases was 0.93 and the average distance between the initial and the estimated parameters was $\hat{R} = (0.011, 0.007, 0.003)rad, \hat{T} = (0.503, 0.302, 0.578)mm$, showing that our method can capture in-plane deformations at the same time as it looks for the optimal plane location (see Figure 3). In all these cases (both plane estimation and in-plane deformation tests) we use 13122 labels, $\alpha = 0.9$, 3 levels of grid refinement and final grid resolution of 16×16 nodes; the execution time was about 4min for every case. Another dataset was used to test our approach over multimodal registration. Since it was performed over images of different modalities, we used Mutual Information as similarity measure instead of SAD. The dataset consists of a preoperative brain MRI volume (voxel size of $0.5 \times 0.5 \times 0.5mm^3$ and resolution of $394 \times 466 \times 378$ voxels) and 6 series of 10 US images extracted from the patient 01 of the database MNI BITE presented in [2]. The intraoperative US images were acquired using the prototype neuronavigation system IBIS NeuroNav. We generated 6 different sequences of 10 2D US images of the brain ventricles, with a size of $48 \times 38mm$ and resolution of $0.3 \times 0.3mm$. The ventricles were manually segmented in both modalities. The estimated position of the slice n was used to initialize the registration process of slice $n + 1$. Slice 0 was initialized in a position near to the Ground Truth using the rigid transformation provided together with the dataset. The DICE coefficient and Contour Mean Distance (CMD) were calculated before and after registration. Figure 4 summarizes the average DICE and CMD coefficients for every series. It shows that the DICE increases after the registration process an average of 0.05 while CMD decreases an average of $0.4mm$. Note that average DICE coefficients are always greater thanc 0.7. Given that we are dealing with highly challenging images of low resolution being heavily corrupted from speckle,

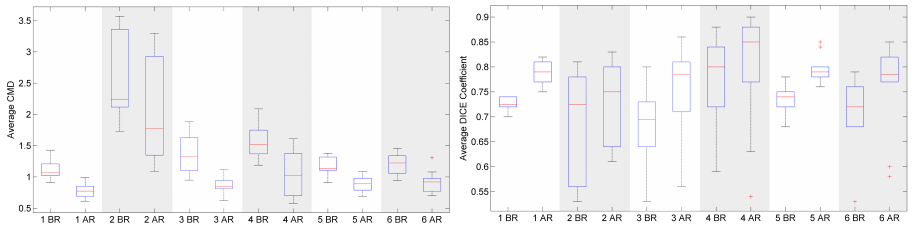


Fig. 4. Average DICE Coefficient (a) and Average CMD (Contour Mean Distance) (b) of all the slices, for every brain series, before (BR) and after (AR) registration process. The average DICE increment after registration of all the series is 0.05 and the average CMD decrement is 0.4mm.

those results are extremely promising. In all the registration cases an initial grid size of $4mm$, 6174 labels, $\alpha = 0.9$, 3 grid levels and 4 iterations of the optimization process were used, giving an average time of around 10min per case. It is also important to analyze results presented by other authors performing slice-to-volume registration. [5] worked with monomodal registration of prostate MRI images and reported average target registration errors below 1mm. [1] tackled the problem of multimodal registration (US and CT images) reporting results with an error of 1.56 ± 0.78mm, using initializations with uniformly random shifts in the range -5 to 5mm. Both of them model only the rigid transformations without taking into account the in-plane deformations. Our method achieved results below 1mm in case of monomodal registration and it maintained the DICE coefficient greater than 0.7 in case of multimodal registration, but was able to deal with the in-plane deformations.

4 Conclusion

In this paper we proposed a novel method for deformable 2D-3D registration using a single shot optimization method that involves plane selection and in plane dense-deformation. This was achieved through an over-parameterized graphical model (5-dimensional representation) that is metric free, can cope with arbitrary deformation models and encodes different in-plane regularization constraints. Clinical validation using real scenarios and examples where 2D acquisitions have been simulated demonstrated the potentials of our method, proved its efficiency in terms of precision and, compared to other methods that tackle a similar problem (like [1] and [5]), seems to achieve state of the art results. The proposed formulation from theoretical view point inherits two limitations, one related with the dimensionality of the label space (that can be handled due to the limited 2D grid size), and a second related with the coplanarity constraint that is approximately imposed through the suggested over-parameterization. Both of them could be alleviated through the use of third-order potentials. Given a 3D label deformation space and triples of neighborhood control points, the plane and the in-plane deformation can be automatically determined. This will introduce certain additional complexity in terms of optimization that could be easily dealt with higher-order to pair-wise MRF reduction methods [10] or dual decomposition [11]. Such an approach

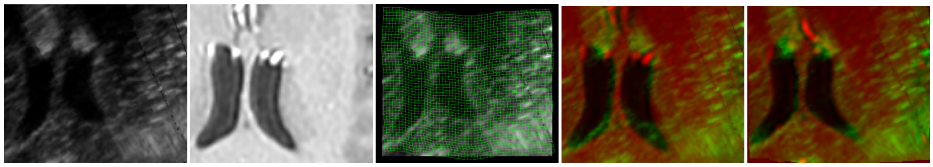


Fig. 5. Registration of a 2D US image and a MRI scan of the brain. (a) 2D US source image. (b) Slice extracted from the MRI corresponding to the initial position of the plane. (c) Deformation field obtained after the registration process. (d) Overlapping between images *a* (red) and *b* (green). (e) Overlapping between the deformed source image (red) and the MRI slice corresponding to the estimated plane (green).

will be more precise in terms of data/regularization term definition and of comparable complexity.

References

1. San José Estépar, R., Westin, C., Vosburgh, K.: Towards real time 2d to 3d registration for ultrasound-guided endoscopic and laparoscopic procedures. IJCARS **4**(6) (2009) 549–560
2. Mercier, L., Del Maestro, R.F., Petrecca, K., Araujo, D., Haegelen, C., Collins, D.L.: Online database of clinical mr and ultrasound images of brain tumors. Medical Physics **39** (2012)
3. Osechinskiy, S., Kruggel, F.: Slice-to-volume nonrigid registration of histological sections to mr images of the human brain. Anatomy Research International **2011** (2010)
4. Dalvi, R., Abugharbieh, R.: Fast feature based multi slice to volume registration using phase congruency. EMBS (2008) 5390–5393
5. Gill, S., Abolmaesumi, P., Vikal, S., Mousavi, P., Fichtinger, G.: Intraoperative prostate tracking with slice-to-volume registration in mri. SMIT (08 2008) 154–158
6. Glocker, B., Sotiras, A., Komodakis, N., Paragios, N.: Deformable medical image registration: setting the state of the art with discrete methods. Annu Rev Biomed Eng **13** (2011)
7. Mahapatra, D., Sun, Y.: Nonrigid registration of dynamic renal mr images using a saliency based mrf model. Metaxas, D., Axel, L., Fichtinger, G., Szkely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 771–779. Springer, Heidelberg (2008)
8. Zikic, D., Glocker, B., Kutter, O., et. al, M.G.: Linear intensity-based image registration by markov random fields and discrete optimization. Med Image Anal **14**(4) (2010) 550–562
9. Komodakis, N., Tziritas, G., Paragios, N.: Fast, approximately optimal solutions for single and dynamic mrfs. CVPR (2007) 1–8
10. Ishikawa, H.: Transformation of general binary mrf minimization to the first-order case. PAMI, IEEE Transactions on **33**(6) (2011) 1234–1249
11. Komodakis, N., Paragios, N.: Beyond pairwise energies: Efficient optimization for higher-order mrfs. CVPR (2009) 2985 –2992