

Markov Random Field Optimization for Intensity-based 2D-3D Registration

Darko Zikic¹, Ben Glocker¹, Oliver Kutter¹, Martin Groher¹,
Nikos Komodakis², Ali Khamene³, Nikos Paragios^{4,5}, and Nassir Navab¹

¹ Computer Aided Medical Procedures (CAMP), Technische Universität München, Germany

² Computer Science Department, University of Crete, Greece

³ Siemens Corporate Research (SCR), Princeton, NJ, USA

⁴ Laboratoire MAS, Ecole Centrale Paris, Chatenay-Malabry, France

⁵ Equipe GALEN, INRIA Saclay - Ile-de-France, Orsay, France.

ABSTRACT

We propose a Markov Random Field (MRF) formulation for the intensity-based N-view 2D-3D registration problem. The transformation aligning the 3D volume to the 2D views is estimated by iterative updates obtained by discrete optimization of the proposed MRF model. We employ a pairwise MRF model with a fully connected graph in which the nodes represent the parameter updates and the edges encode the image similarity costs resulting from variations of the values of adjacent nodes. A label space refinement strategy is employed to achieve sub-millimeter accuracy. The evaluation on real and synthetic data and comparison to state-of-the-art method demonstrates the potential of our approach.

Keywords: 2D-3D Registration, Markov Random Fields, Discrete Optimization

1. INTRODUCTION

Relating 3D to 2D medical images has received lots of attention in the past years. Radiation therapy,¹ guidance of surgical robots,² or catheter navigation in vascular images³ are just a few examples where a successful registration of pre- and intra-operative data can ease medical procedures. Algorithms based on anatomical features,² image intensities,⁴ hybrid approaches,⁵ or methods using 3D reconstruction⁶ have been proposed, where a focus has been laid on the assessment of similarity between 2D and 3D data.

Less work can be found that directly addresses the optimization strategy. Feature-based approaches usually use gradient-based optimization techniques such as the Newton method, whereas intensity-based techniques rely on optimization techniques that do not require the computation of the cost function gradient, such as the Powell, or the Nelder–Mead Simplex method. All of these approaches are local, so that a good initialization is required for the 2D-3D registration to converge to the right solution. In⁷ a sequential Monte-Carlo sampling method tries to subsequently create hypotheses, which are directed by particle filters to find the most probable solution.

In the last years, the solution of computer vision problems by Markov Random Fields (MRFs)^{8,9} and discrete optimization has become increasingly popular. Different medical imaging problems such as segmentation¹⁰ or non-linear registration¹¹ have been treated by this approach. Recent advances^{12–14} in discrete optimization methods make this approach increasingly attractive. MRF-based optimization has, to the best of our knowledge, not yet been applied to the problem of 2D-3D registration. In this work, we demonstrate how the 2D-3D registration can be formulated by MRFs and solved by discrete optimization. The results demonstrate that the proposed method can compete with the state-of-the-art in terms of robustness and quality.

2. METHOD

The task of 2D-3D registration is to recover a rigid 3D transformation T which aligns the coordinate frame of the 3D volume I to the coordinate system of the 2D imaging devices, which generate the 2D images J_k^* . The intensity-based approach

*In state-of-the-art imaging systems calibration information can be assumed to be given and image distortion is absent due to flat-panel detector technology.

Algorithm 1: Basic MRF-based 2D-3D Registration

```

1: set initial parameters  $p^l$ ;
2: until convergence do repeat
3:    $\Delta p := \text{optimizeMRF}(I, J_1, \dots, J_n, L, p^l)$ ; // Sec. 2.2
4:    $p^l := p^l + \Delta p$ ;
5:   refine_label_ranges; // Sec. 2.3
6: end

```

Figure 1: Outline of the basic algorithm for 2D-3D registration by MRF optimization.

for modeling this problem is to optimize a suitable similarity measure ξ between the projections of the transformed 3D image $P_k(I \circ T)$ and the 2D images J_k , resulting in following optimization problem

$$T_{\text{opt}} = \arg \min_T \frac{1}{n} \sum_{k=1}^n \xi(P_k(I \circ T), J_k) \quad , \quad (1)$$

where n is the number of available 2D images and T_{opt} is the optimal rigid transformation. A wide choice of similarity measures can be employed.⁴

The transformation is parameterized by parameters p , which is denoted by T_p . If the parameters are expressed in terms of the initial estimate p^l and an update Δp , we can restate the optimization problem as estimation of the optimal parameter update, that is

$$\Delta p_{\text{opt}} = \arg \min_{\Delta p} \frac{1}{n} \sum_{k=1}^n \xi(P_k(I \circ T_{p^l + \Delta p}), J_k) \quad , \quad (2)$$

such that $T_{\text{opt}} = T_{p^l + \Delta p_{\text{opt}}}$.

In our case, T_p is parameterized by a vector $p = [t_x, t_y, t_z, \phi_x, \phi_y, \phi_z]$, where $[t_x, t_y, t_z]$ is the translation, and the rotation R is parameterized by Euler angles ϕ_x, ϕ_y, ϕ_z as $R = R_{\phi_x} R_{\phi_y} R_{\phi_z}$.

Formulating Equation (2) as an MRF problem shares the advantages of the state-of-the-art optimization methods for the 2D-3D problem such as Simplex or Powell's method,¹⁵ that the similarity measure is easily interchangeable, since no derivatives of the measure are required. Furthermore, discrete optimization methods allow for a large search range.

With the above parameterization, we can summarize the proposed approach, compare also Figure 1. We perform a series of iterations, where in each iteration updates to the current transformation parameters are found. The updates are determined by discrete optimization of the MRF model for 2D-3D registration, which is introduced in Section 2.2. Since the optimization is performed using a discrete set of label values as possible solutions, in each iteration a refinement of the label range is performed in order to achieve sub-millimeter registration accuracy. Our approach is generic since no specific assumptions are made on the similarity measure, the number of 2D views, or the imaging device setup.

The remainder of this section is organized as follows. After briefly introducing the basic concepts of MRF modeling in Section 2.1, we present how the 2D-3D registration problem can be mapped to the general MRF framework in Section 2.2. Section 2.3 describes the discretization strategy for our MRF model.

2.1 Markov Random Fields

Markov Random Field optimization is a common approach for parameter estimation. Given a set of parameters, one can define a graph $\mathcal{G} = (\mathcal{V}, \mathcal{C})$ consisting of a set of nodes \mathcal{V} (one node per parameter) and a set of cliques \mathcal{C} (where each clique is a subset of \mathcal{V}). Assuming that each node v_i takes a label x_i from a discrete set \mathcal{L} , the task becomes to find the optimal labeling \mathbf{x} which minimizes

$$E_{\text{mrf}}(\mathbf{x}) = \sum_{c \in \mathcal{C}} \psi_c(\mathbf{x}_c) \quad , \quad (3)$$

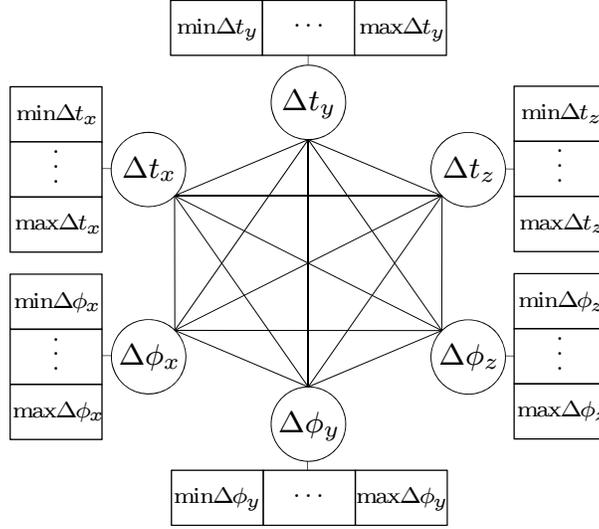


Figure 2: Topology of the MRF Model. The transformation parameter updates are represented by the nodes. The edges encode the image similarity costs associated with the variation of the labeling of the adjacent nodes. The MRF determines the parameter updates by optimizing the sum of the edge costs.

which is a sum of *clique potentials* $\psi_c(\cdot)$ determining the costs of certain label assignments and \mathbf{x}_c is the vector of labels assigned to the parameter subset c .

The most common MRF model used in computer vision tasks (e.g. segmentation) is the second-order (pairwise) model containing at most cliques of size two. Many efficient algorithms have been proposed^{12–14} to solve the inference problem for this special case. Since the optimum of the similarity measure cannot be determined by optimizing the single parameters independently, the 2D-3D registration problem actually requires a sixth-order clique model. However, higher-order cliques are less commonly used since fast optimization methods are currently not available. The next section shows how our approach circumvents this difficulty by approximating the solution of the 2D-3D registration by a second-order clique model.

2.2 MRFs for 2D-3D Registration

The key idea of our approach is to estimate the parameters of the rigid 3D transformation through MRF inference on a common pairwise model. In order to define the dependencies between the transformation parameters, we introduce a fully-connected graph $\mathcal{G}^* = (\mathcal{V}, \mathcal{E})$ where the nodes correspond to the transformation parameters ($|\mathcal{V}| = 6$) and the set of edges \mathcal{E} contains all possible parameter pairs. This allows us to model the global consistency of the parameter estimation while efficient optimization methods for pairwise models can be used. The resulting MRF topology is illustrated in Figure 2.

So the MRF cost function for our model is defined as a sum over pairwise potentials

$$E_{\text{mrf}}(\mathbf{x}) = \sum_{(v_i, v_j) \in \mathcal{E}} \psi_{ij}(x_i, x_j) , \quad (4)$$

where $\psi_{ij}(x_i, x_j)$ determines the cost for assigning labels x_i and x_j to the pair (v_i, v_j) . In our case, the potentials determine the cost of a simultaneous variation of the transformation parameters Δp_i and Δp_j (denoted as Δp_{ij}) with respect to the image similarity measure

$$\psi_{ij}(x_i, x_j) = \frac{1}{n} \sum_{k=1}^n \xi(P_k(I \circ T_{p' + \Delta p_{ij}}), J_k) . \quad (5)$$

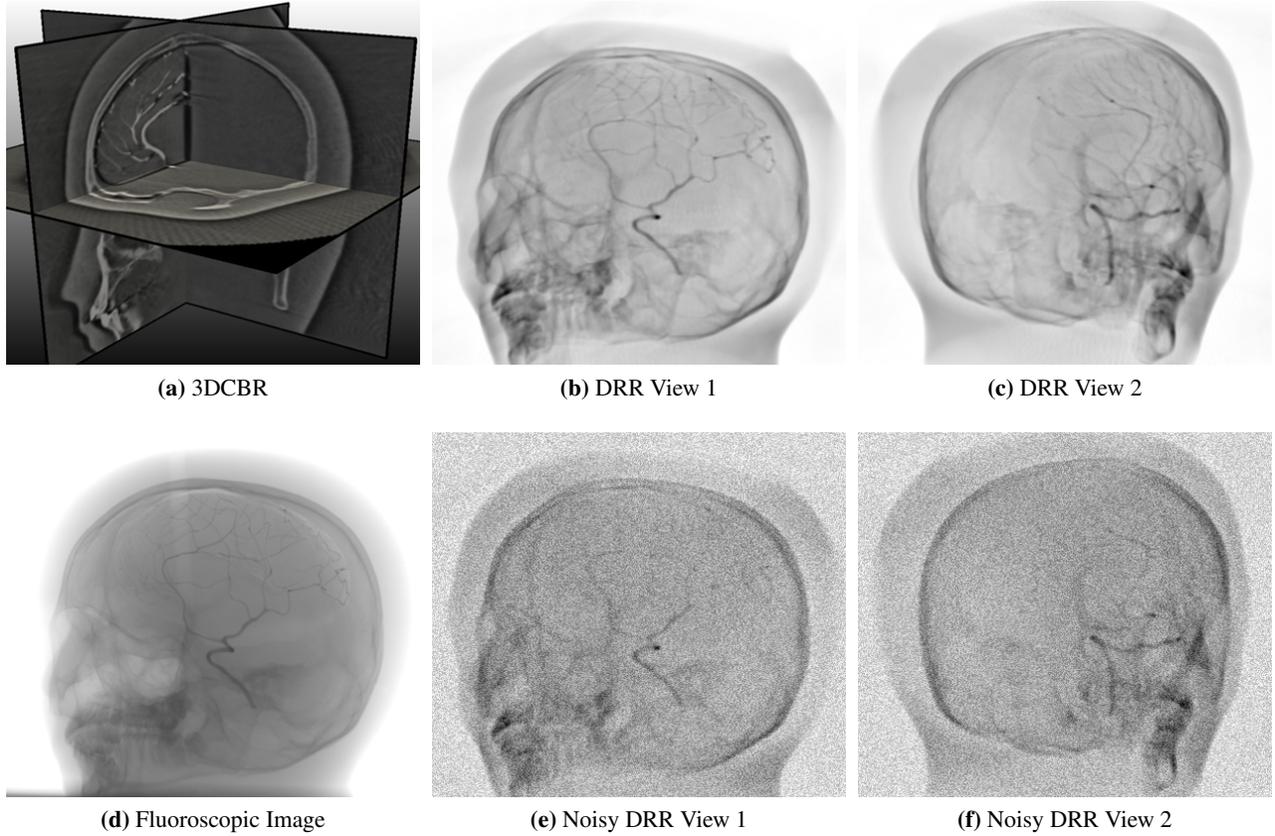


Figure 3: Visualization of the test data: (a) 3D cone beam reconstruction (3DCBR), which is used in all experiments; (b,c) Exemplary DRRs which are computed in the iterations of the algorithm; (d) Fluoroscopic image used as target in the 1-view test; (e,f) Exemplary DRRs with 20% uniform noise used as targets in the 2-view tests.

Based on equation Equation (5), we finally define the MRF cost function for the 2D-3D registration problem as

$$E_{2D3D}(\mathbf{x}) = \sum_{(v_i, v_j) \in \mathcal{E}} \frac{1}{n} \sum_{k=1}^n \xi(P_k(I \circ T_{p' + \Delta p_{ij}}), J_k) . \quad (6)$$

For minimizing Equation (6), we use the recently proposed FastPD method[†].¹⁴ Due to the limited space, we refer the reader to the given reference for more details about the optimization algorithm. The following section describes the discretization of the parameter space.

2.3 Parameter Space Discretization

A crucial part of the proposed method is the discretization of the parameter search space which involves the definition of the label space \mathcal{L} . On the one hand, one would like to keep the number of labels small for efficient minimization of Equation (6). On the other hand, an undersampling of the parameter search space might result in inaccurate registration results. To this end, we employ a successive refinement strategy for computing MRF labelings, which results in refined parameter updates. For each parameter, we define a value range which is uniformly sampled in order to generate a discrete subset \mathcal{L}_i . In iteration k , we rescale the ranges by a factor α^k ($\alpha < 1$) which defines the new possible values for the parameter updates. The iterative label space refinement allows us to keep the number of labels quite small and we can start with a large parameter range, while being able to achieve sub-millimeter registration accuracy.

[†] Available on <http://www.csd.uoc.gr/~komod/FastPD/>.

3. EVALUATION

In this section we test the MRF-based optimization approach. We assess the performance of the proposed method by comparing it to the Simplex optimization method. To this end, exactly the same framework and same input data is used for both methods in all experiments with only difference being the optimization module. In order to perform the evaluation in a controlled setting, we conduct the 2-view tests on real data, but with synthetically created projections. This way, the choice of the similarity measure plays a smaller role and we can compare the performance of the optimization methods. This also has the advantage that we can test on a large number of different views with known ground truth, rather than performing the tests on only a few real views by randomly disturbing the initialization. The robustness and applicability to real settings is demonstrated in experiments by adding noise to the projections. Furthermore, we perform our algorithm in a 1-view test scenario, in which a real fluoroscopic image is used.

3.1 General Setting

As input data for the experiments we use a 3D cone beam reconstruction (3DCBR) of a phantom head with a cerebral vessel structure, computed from a single sweep of a monoplanar stationary C-arm with flat-panel detector (Siemens Axiom Artis dTA). As 2D input we use a fluoroscopic image obtained by the same device, with ground truth transformation obtained by feature-based registration, and verified by careful inspection. An overview of the data is given in Figure 3. Internal camera parameters are inferred from calibration data given by the X-ray device.

The 2D projections required inside the algorithm and as synthetic targets are computed as digitally reconstructed radiographs (DRRs). The DRRs are generated from the volume data by GPU accelerated ray-casting, implemented in OpenGL Shading Language (GLSL). The intensity for each detector pixel is computed by the line integral along the ray from the source to the detector. A conversion operator is used to remap the intensity values to X-ray energies.

All tests are performed with a set of random offset poses. The poses are generated randomly from a parameter range of up to 30mm for translations and rotation angles of up to 80° for the 2-view test, and 15mm and 40° for the 1-view test. We evaluate the results by the Target Registration Error (TRE). The TRE is computed as the mean of the distances between a set of points transformed by the ground truth transformation and the same point set transformed by the resulting transformation. The point set is created by regularly sampling a 20cm cube centered at the phantom head with a resolution of $10 \times 10 \times 10$.

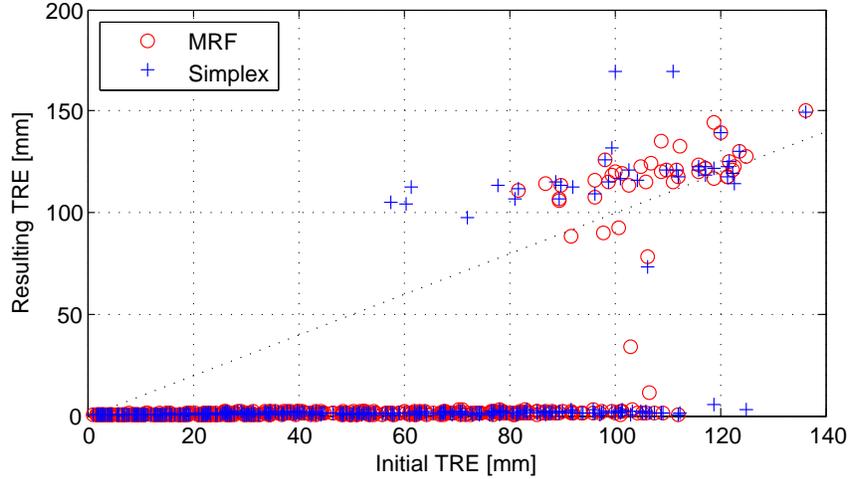
The effect of the choice of the similarity measures does not present a focus of this work. As a similarity measure we use the local version of the Normalized Cross Correlation (NCC).⁴ We also performed tests with the Gradient Difference measure,⁴ with very similar results. The evaluation of the similarity measure is implemented on the CPU.

The registrations are performed on a standard multi-scale strategy using a Gaussian image pyramid. In all experiments, the initial search space ranges for the MRF method are ± 50 mm for the translations and $\pm 90^\circ$ for the rotations while the ranges are discretized using 7 sampling steps (i.e. $|\mathcal{L}_i| = 7$). The label ranges are successively refined in each iteration (cf. Section 2.3) with $\alpha = 0.66$. Different parameter values are tested for both the number of sampling steps (3-10) and the refinement factor (0.5-0.9), resulting in very similar registration results. For these settings, the average registration run takes about 4 minutes.

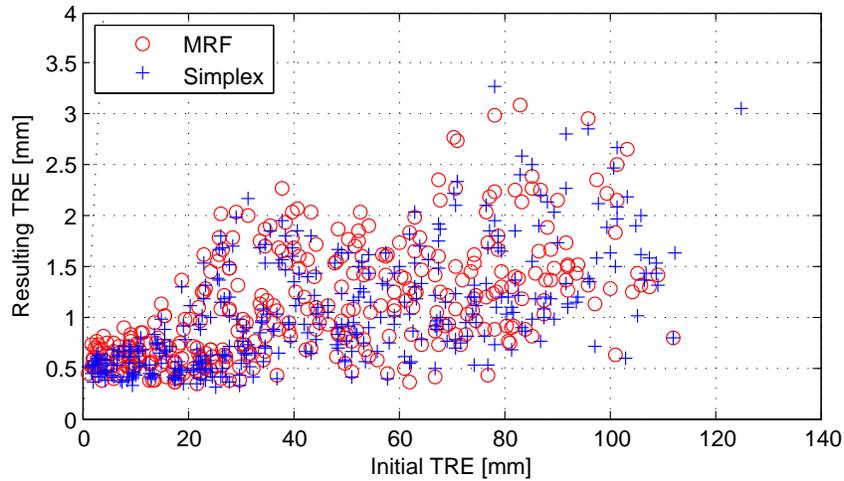
For the Simplex method, the step sizes for the six parameters are set to the same values as the search space ranges of the MRF method, and were shown to give good results.

3.2 2-View Test with DRRs as Target Image

We carry out the 2-view test by performing 400 runs. In each run, a new pair of orthogonal DRRs J_1, J_2 is generated by adding a random pose displacement to an initial pose of the 3DCBR image. The initial pose is chosen such that one projection produces an anteroposterior view of the 3DCBR image; the second image orthogonal to the first yields a sagittal view. Internal camera parameters of the projection matrix have been chosen from a calibration of an interventional C-arm to resemble a clinical scenario. In order to test the robustness of the approach and to approximate a realistic scenario, uniform noise in the range of 20% of the intensities is added to all generated DRRs, cf. Figs. 3e, 3f. The same test was also performed without noise, yielding comparable results. The tests show that both approaches perform well, with a small number of failed registrations only for very large initial misalignments. For more details please compare also Figure 4.



(a) TRE for 2-view test with noise



(b) TRE of successful runs

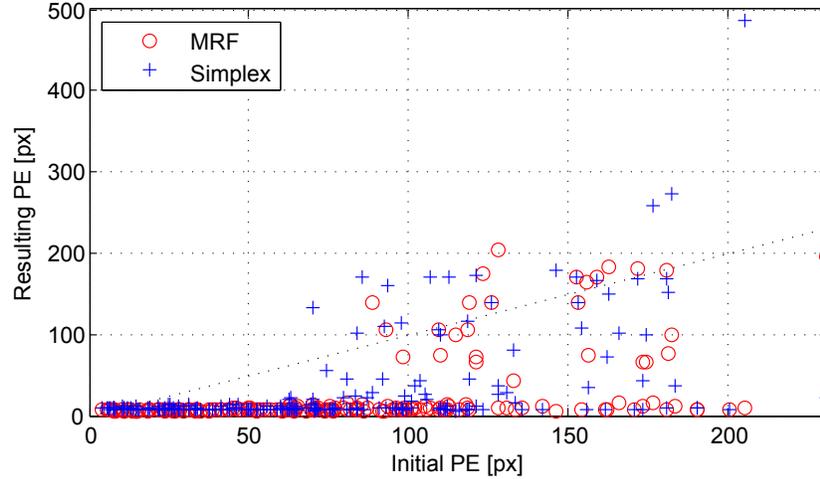
Figure 4: Results of 400 random 2-view test runs with 20% noise. The x and y-coordinates of the graph points represent the TRE before and after registration. The diagonal is the line of no improvement. (b) shows a zoom of the area of successful test runs.

3.3 1-View Test with Fluoroscopic Image as Target

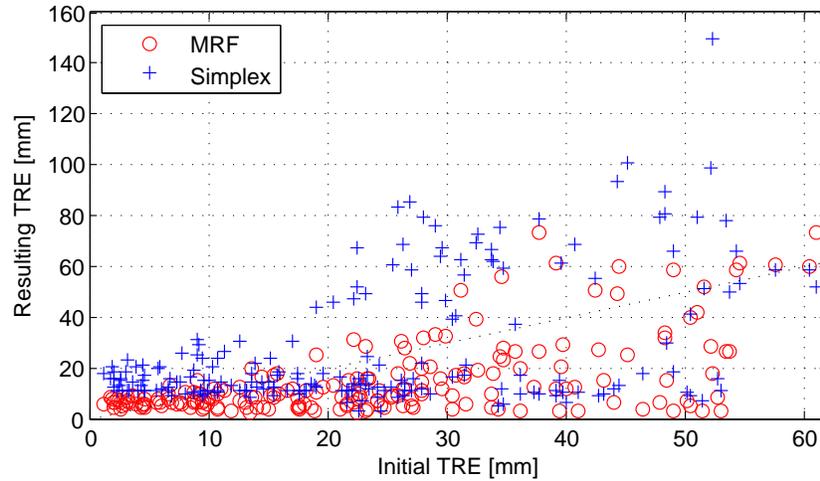
The second series of experiments is performed in a 1-view scenario by registering the 3DCBR volume to a real fluoroscopic image. Here, we use 200 random poses as initial offsets to the 3D transformation. Additionally to the TRE, in this experiment we also evaluate the Projection Error (PE), which is computed by projecting the transformed 3D samples used for TRE computation to the image and computing their mean distance. The PE measures the error visible in the image, such that it is not very sensitive to misalignments in the viewing direction. By comparing the TRE and PE measures, we observe that while the PE is approximately the same for both approaches, the TRE is slightly better for the MRF-based approach. This suggests that the MRF optimization is able to better recover the translation along the viewing direction. For details please see Figure 5.

4. CONCLUSION AND DISCUSSION

In this paper we present an MRF model for the optimization of the intensity-based 2D-3D registration problem. To our knowledge, this is the first time that MRFs are considered in this scenario. Our approach does not require gradient computation of the cost function and allows an intuitive control of the search ranges for the single transformation parameters



(a) Projection Error (PE)



(b) Target Registration Error (TRE)

Figure 5: Results of the 200 random 1-view test runs with fluoroscopic image. In (a), the projection error (PE) is shown. (b) illustrates the TRE. While PE is approximately the same for both approaches, the TRE is slightly better for the MRF-based approach. Since in contrast to TRE, PE measures only the error visible in the image, this suggests that the MRF optimization better recovered the translation along the viewing direction.

while being robust to the setting of the algorithm parameters (number of sampling steps, label space refinement factor). The focus of this work is on the optimization method - for specific applications, many possible improvements can be considered in more detail, such as an optimized DRR generation, or selecting a dedicated similarity measure for a specific application. The evaluation demonstrates that our MRF-based approach can compete with currently employed methods, and is robust to noise and initialization. Due to the active development in discrete optimization, we see a strong potential of the proposed method, since it allows to transfer the advances in MRF optimization to 2D-3D registration, e.g. the integration of future fast methods for higher-order clique models. Further work includes the investigation of alternative MRF topologies, and modification of edge costs to incorporate prior knowledge on the relevance of certain transformation parameters, e.g. the translation in viewing direction for the 1-view problem.

REFERENCES

- [1] Khamene, A., Bloch, P., Wein, W., Svatos, M., and Sauer, F., “Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy,” *IEEE Transactions on Medical Imaging (TMI)* **10** (2006). 1

- [2] Gueziec, A., Kazanzides, P., Williamson, B., and Taylor, R. H., "Anatomy-based registration of ct-scan and intraoperative x-ray images for guiding a surgical robot," *IEEE Transactions on Medical Imaging (TMI)* **17** (Oct. 1998). [1](#)
- [3] Groher, M., Bender, F., Hoffmann, R., and Navab, N., "Segmentation-driven 2D-3D registration for abdominal catheter interventions," in [*Proc. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*], (2007). [1](#)
- [4] Penney, G. P., Weese, J., Little, J. A., Desmedt, P., Hill, D. L. G., and Hawkes, D. J., "A comparison of similarity measures for use in 2D-3D medical image registration," *IEEE Transactions on Medical Imaging (TMI)* **17**(4) (1998). [1](#), [2](#), [5](#)
- [5] Livyatan, H., Yaniv, Z., and Joskowicz, L., "Gradient-based 2-d/3-d rigid registration of fluoroscopic x-ray to ct," *IEEE Transactions on Medical Imaging (TMI)* **22** (Nov. 2003). [1](#)
- [6] Markelj, P., Tomazevic, D., Pernus, F., and Likar, B., "Robust gradient-based 3-D/2-D registration of CT and MR to x-ray images," *IEEE Transactions on Medical Imaging (TMI)* **27** (Dec. 2008). [1](#)
- [7] Florin, C., Williams, J., Khamene, A., and Paragios, N., "Registration of 3D angiographic and X-ray images using sequential monte carlo sampling," in [*Computer Vision for Biomedical Image Applications*], **3765** (2005). [1](#)
- [8] Geman, S. and Geman, D., "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," in [*IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*], (1984). [1](#)
- [9] Li, S. Z., [*Markov random field modeling in image analysis*], Springer-Verlag New York, Inc. (2001). [1](#)
- [10] Boykov, Y. and Jolly, M.-P., "Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images," in [*Proc. International Conference of Computer Vision (ICCV)*], (2001). [1](#)
- [11] Glocker, B., Komodakis, N., Tziritas, G., Navab, N., and Paragios, N., "Dense image registration through mrfs and efficient linear programming," *Medical Image Analysis (MedIA)* (2008). [1](#)
- [12] Boykov, Y., Veksler, O., and Zabih, R., "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* **23** (Nov. 2001). [1](#), [3](#)
- [13] Kolmogorov, V., "Convergent tree-reweighted message passing for energy minimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* (2006). [1](#), [3](#)
- [14] Komodakis, N., Tziritas, G., and Paragios, N., "Fast, approximately optimal solutions for single and dynamic mrfs," in [*Proc. Conference on Computer Vision and Pattern Recognition (CVPR)*], (2007). [1](#), [3](#), [4](#)
- [15] Press, W., Teukolsky, S., Vetterling, W., and Flannery, B., [*Numerical Recipes in C*], Cambridge University Press, New York (1993). [2](#)