

# Surface/Volume-Based Articulated 3D Spine Inference through Markov Random Fields\*

Samuel Kadoury and Nikos Paragios

Laboratoire MAS, Ecole Centrale de Paris, France  
{samuel.kadoury,nikos.paragios}@ecp.fr  
GALEN Group, INRIA Saclay, Ile-de-France, France

**Abstract.** This paper presents a method towards inferring personalized 3D spine models to intraoperative CT data acquired for corrective spinal surgery. An accurate 3D reconstruction from standard X-rays is obtained before surgery to provide the geometry of vertebrae. The outcome of this procedure is used as basis to derive an articulated spine model that is represented by consecutive sets of intervertebral articulations relative to rotation and translation parameters (6 degrees of freedom). Inference with respect to the model parameters is then performed using an integrated and interconnected Markov Random Field graph that involves singleton and pairwise costs. Singleton potentials measure the support from the data (surface or image-based) with respect to the model parameters, while pairwise constraints encode geometrical dependencies between vertebrae. Optimization of model parameters in a multi-modal context is achieved using efficient linear programming and duality. We show successful image registration results from simulated and real data experiments aimed for image-guidance fusion.

## 1 Introduction

Spinal deformity pathologies such as idiopathic scoliosis are complex three-dimensional (3D) deformations of the trunk, described as a lateral deviation of the spine combined with asymmetric deformation of the vertebrae. Surgical treatment usually involves correction of the scoliotic curves with preshaped metal rods anchored in the vertebrae with screws and arthrodesis (bone fusion) of the intervertebral articulations. This long procedure can be complex since it requires high level of precision for inserting pedicle screws through the spinal canal [1,2].

With recent advances in medical imaging enabling CT acquisitions during the surgical procedure, real-time fusion of anatomical structures obtained from various modalities becomes feasible. It offers the unique advantage to visualize anatomy during intervention and localize anatomical regions without segmenting operative images. By fusing the 3D volume images such as CT, C-arm CT [2] or

---

\* This work was partially supported by an FQRNT grant. The authors would like to thank Philippe Labelle from Sainte-Justine Hospital for data processing.

MR with an accurate preoperative model, the surgeon can see the position and orientation of the instrumentation tools on precise anatomical models in real time. In this work, we take advantage of a personalized preoperative 3D model which reflects the detailed geometry of the patient’s spine from standard biplanar X-rays. While the morphology of each vertebra remains identical between initial exam and surgery, intervertebral orientation and translation vary substantially.

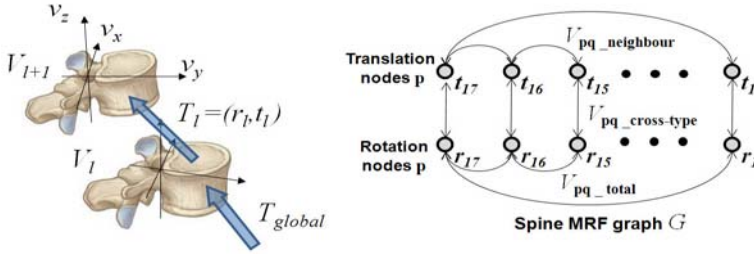
Registration of intraoperative fluoroscopic images and preoperative CT/MR images has been proposed to aid interventional and surgical orthopedic procedures [3]. For example in [4,5], 3D models obtained from CT or MR were registered to 2D X-ray and fluoroscopic images using gradient amplitudes for optimizing the correspondence of single bone structures. Similar objective functions using surface normals from statistical PDMs [6] were applied for the femur. In spine registration however, one important drawback is that each vertebra is treated individually instead of as a global shape. An articulated model may allow to account for the global geometrical representation [7] by incorporating knowledge-based intervertebral constraints. These 3D intervertebral transformations were transposed in [8] to accomplish the segmentation of the spinal cord from CT images, but multi-modal registration has yet to be solved. Optimization is also based on gradient-descent, prone to non-linearity and local minimums. These methods require segmentation of 3D data or fluoroscopic image, which itself is a challenging problem and has a direct impact on registration accuracy.

In this paper, we propose a framework for registering preoperative 3D articulated spine models in a standing position to lying intraoperative 3D CT images. The general approach is described as follows. We first use a personalized 3D spine reconstructed from biplanar X-rays to derive an articulated model represented with intervertebral transformations. We then formulate inference through a Markov Random Field (MRF) model, proposing an image-based registration which avoids CT image segmentation, is computational efficient (few seconds) and with known optimality bounds. The optimization integrates prior knowledge to constrain the adjustment of intervertebral links between neighboring objects of the articulated model with pairwise potentials, as well as modular image data-terms. One of the applications is to help surgeons treat complicated deformity cases by fusing high-resolution preoperative models for increased accuracy of pedicle screw insertion, reducing surgery time. Sections 2 and 3 presents the method in terms of geometric modeling and MRF inference. Experiments are showed in Section 4, with a discussion in Section 5.

## 2 Personalized 3D Reconstruction of Articulated Spines

### 2.1 Preoperative Spine 3D Reconstruction

From calibrated coronal and sagittal X-ray images of the patient’s spine, the personalized 3D model is achieved by means of a reconstruction method merging statistical and image-based models based on the work of [9]. The patient’s spine centerline is first embedded onto a 3D database containing scoliotic spines to predict an initial model with 17 vertebrae (12 thoracic, 5 lumbar), 6 points



**Fig. 1.** Articulated spine in an MRF graph, integrating three types of constrained pairwise potentials

per vertebra (4 pedicle tips and 2 endplate midpoints). This crude statistical 3D model is then refined with an individual scoliotic vertebra segmentation approach by extending 2D geodesic active regions in 3D, in order to evolve prior deformable 3D surfaces. An atlas of vertebral meshes  $S_i = \{x_{i1}, \dots, x_{iN}\}$  with triangles  $x_j$  are initially positioned and oriented from their respective 6 precise landmarks  $s_i$  composing  $X_{preop}$ , and the surface is evolved so that the projected silhouettes of the morphed 3D models would therefore match the 2D information on the biplanar X-rays. At the end of process, the 3D landmark coordinates  $s_i$  and corresponding polygonal vertebral meshes  $S_i$  are optimal with regards to statistical distribution and image correspondences.

### 2.2 Articulated Spine Model

The 3D landmarks  $s_i$  obtained in the previous section are used to rigidly register each vertebra to its upper neighbor, and the resulting rigid transforms are optimized in the registration problem. Hence, the spine is represented by a vector of local intervertebral rigid transformations  $A = [T_1, T_2, \dots, T_N]$  as illustrated in Fig. 1. To perform global anatomical modeling of the spine, we convert  $A$  into an absolute representation  $A_{absolute} = [T_1, T_1 \circ T_2, \dots, T_1 \circ T_2 \circ \dots \circ T_N]$  using recursive compositions. The transformations are expressed in the local coordinate system of the lower vertebra, defined by vectors  $v_x, v_z$  and  $v_y = v_x \times v_z$ , where  $v_x$  and  $v_z$  are the vectors linking pedicle and endplate midpoints respectively. Center of transformation is located at the midpoint of all 4 pedicle tips. The rigid transformations described in this paper are the combination of a rotation matrix  $R$  and a translation vector  $t$ . We formulate the rigid transformation  $T = \{R, t\}$  of a vertebral mesh triangle as  $y = Rx + t$  where  $x, y, t \in \mathbb{R}^3$ . Composition is given by  $T_1 \circ T_2 = \{R_1 R_2, R_1 t_2 + t_1\}$ , while inversion as  $T^{-1} = \{R^T, -R^T t\}$ .

### 3 Intraoperative Spine Inference from Images with MRFs

Our method reformulates inference as an MRF optimization where a set of labels  $L = \{l^1, \dots, l^i\}$  defined in the quantized space  $\Theta = \{\mathbf{d}^1, \dots, \mathbf{d}^i\}$  is associated with the set of vertebral transformations  $T$  represented by nodes  $\mathbf{p}$ . One seeks

to attribute a label to each node of graph  $G$  such that once the corresponding deformation has been applied, the MRF energy measure between the source and target models is optimal for all vertebrae:

$$E_{\text{total}} = \sum_{\mathbf{p} \in G} V_{\mathbf{p}}(l_{\mathbf{p}}) + \sum_{\mathbf{p} \in G} \sum_{\mathbf{q} \in \mathcal{N}(\mathbf{p})} V_{\mathbf{pq}}(l_{\mathbf{p}}, l_{\mathbf{q}}) \quad (1)$$

where  $V_{\mathbf{p}}(\cdot)$  are the unary potentials representing the image data term, which can be defined independently from the target imaging modality  $g(x)$  such that:

$$V_{\mathbf{p}}(l_{\mathbf{p}}) = \int_{\Omega} \eta_{\text{data}}(g(x), S_i(T_i + \mathbf{d}^{\alpha})) dT. \quad (2)$$

The data term  $\eta_{\text{data}}$  seeks to minimize the distance between the multi-modal images. We will discuss the choice of these costs in the next section where two different applications are considered. The right hand side of Eq.(1) are the pairwise potentials representing the smoothness term between vertebrae connected in the MRF (Fig. 1) and help to constrain the vertebrae main directions in the optimization process. Three classes of pairwise neighborhoods  $\mathcal{N}$  are defined in this problem: neighboring nodes between levels  $l$  and  $l + 1$  measuring the deviation from the initial pose; deformation magnitudes between interconnected translation and rotation nodes; and consistency in length of the segment. These smoothness terms are described below, with  $\lambda_{\mathbf{pq}}$  used as a weighting factor:

$$V_{\mathbf{pq}}(l_{\mathbf{p}}, l_{\mathbf{q}}) = \begin{cases} \lambda_{\mathbf{pq}} \|(T_{\mathbf{p}}^{\text{pre}} \times \mathbf{d}^{l_{\mathbf{p}}}) - (T_{\mathbf{q}}^{\text{pre}} \times \mathbf{d}^{l_{\mathbf{q}}})\|^2, & \text{if } \mathbf{p} \in l \text{ and } \mathbf{q} \in l + 1 \\ \lambda_{\mathbf{pq}} (\|\mathbf{d}_{r_z}^{l_{\mathbf{p}}} + \mathbf{d}_{r_y}^{l_{\mathbf{p}}}\| - \|\mathbf{d}_{t_x}^{l_{\mathbf{p}}} + \mathbf{d}_{t_z}^{l_{\mathbf{p}}}\|), & \text{if } \mathbf{p} \in \mathbb{R}^t \text{ and } \mathbf{q} \in \mathbb{R}^R \\ \lambda_{\mathbf{pq}} |(T_{\mathbf{p}}^{\text{pre}} - T_{\mathbf{q}}^{\text{pre}}) - (\mathbf{d}^{l_{\mathbf{p}}} - \mathbf{d}^{l_{\mathbf{q}}})|, & \text{if } \mathbf{p} \equiv T_{17} \text{ and } \mathbf{q} \equiv T_1. \end{cases} \quad (3)$$

The optimization strategy for the resulting MRF is based on a primal-dual principle where we seek to assign the optimal labels  $L$  to each translation and rotation node  $\mathbf{p}$  of the linked vertebrae, so that the total energy of the graph is minimum. We apply a recently proposed method called FastPD [10]<sup>1</sup> which can efficiently solve the registration problem in a discrete domain by formulating the duality theory in linear programming. The advantage of such an approach lies in its generality, efficient computational speed, and guarantees the global optimum without the condition of linearity. Two types of inter-modality inferences are explored: 3D surface reconstructed X-ray, and intra-op CT volume images.

## 4 Experimental Validation

While validating image registration is not a straightforward problem and ground truth data in medical applications is often not available, we assessed the methods performance using both synthetic and real deformations from datasets obtained in scoliosis clinics. To explore the solution space, sparse sampling considering

<sup>1</sup> Details of authors implementation: <http://www.csd.uoc.gr/~komod/FastPD/>

only displacements along the 6 main axis was selected, resulting in  $6N + 1$  labels in 3D ( $N$  is the sampling rate). The smoothness term was set at  $\lambda_{pq} = 0.4$ . Tests were performed in C++ on a 2.8 GHz Intel P4 processor and 2 GB DDR memory.

An atlas of 17 generic prior vertebra models obtained from serial CT-scan reconstruction of a cadaver specimen was used to construct the 3D preoperative model. Models were segmented using a connecting cube algorithm [11]. The same six precise anatomical landmarks were added on each model by an expert. The atlas is divided into 3 levels of polygonal mesh catalogues of increasing complexity, to adopt the widely used multi-resolution registration approach where coarse-to-fine geometrical models are applied for optimal convergence.

The method was evaluated with three experiments: (a) simulate synthetic deformations on preoperative spines for ground truth data comparison; (b) evaluate intra-modal registration accuracy on 20 cases with pre- and intra-op 3D X-ray models; and (c) test multi-modal image registration using 12 CT datasets. The data term in (a) and (b) was based on the geometric distance between the reconstructed spine and the inferred one, while in (c) it measures the strength of the edges over the triangles corresponding to the inferred spine.

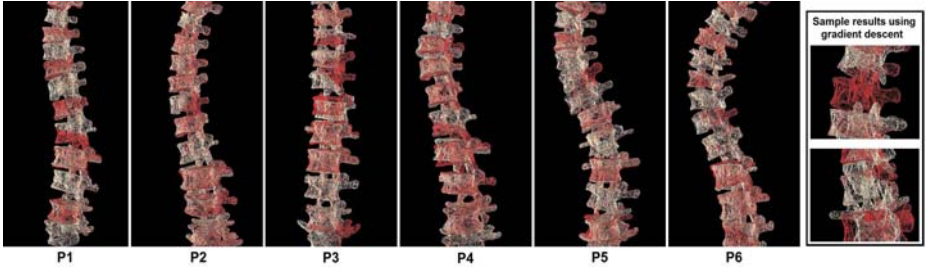
- **Geometric Inference Support:** the singleton data term potential is defined as  $\eta_{RX} = |S_i \cap X_{\text{intra}}| / |S_i \cup X_{\text{intra}}|$ , which represents the volume intersection between the source  $S_i$  and target model  $X_{\text{intra}}$ .
- **Volume/CT Inference Support:** the singleton data term potential defined as  $\eta_{CT} = \sum_{x_{ij} \in S_i} (\gamma^2 + \gamma \|\nabla CT(x_{ij})\|) / (\gamma^2 + \|\nabla CT(x_{ij})\|^2)$  attracts mesh triangles to target high-intensity voxels in the gradient CT volume without segmentation. The term  $\gamma$  is defined as a dampening factor.

#### 4.1 Ground Truth Validation Using Synthetic Deformations

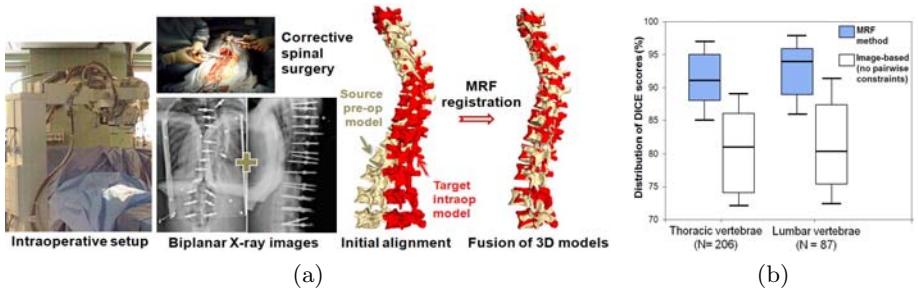
The first experiment consisted of taking six baseline scoliotic patients exhibiting different types of mild curvatures (15 - 50 deg), and simulating target models by applying synthetic deformations to the spine replicating variations observed intraoperatively. Uniformly distributed random noise (mean 0, SD 2 mm) was added to the target models. In Table 1, we present average translation and rotation errors to ground truth data for all six patients. Direct correspondences of mesh vertices between source and target spines were used to compute the Euclidean distance error, compared to an image gradient-descent method which is equivalent to an optimization without any pairwise constraints. The average 3D error was improved by 7.6 mm compared to gradient-descent. This confirms the advantage of integrating global anatomical coherence of the articulated object during registration instead of straightforward optimization techniques which are sensitive to large deformations and poor initialization. Fig. 2 illustrates good model alignment of the MRF approach with constrained articulations, while gradient-descent may cause vertebra collisions and overlapping.

**Table 1.** Ground truth errors from 6 synthetic deformation models, with a 3D mean Euclidean distance (MED) comparison of spine models to a gradient-descent approach

Measures / Subject	P1	P2	P3	P4	P5	P6	Average
Translation ( $T_t$ ) error (mm)	0.41	0.48	0.44	0.76	1.10	0.38	<b>0.59</b>
Angular ( $T_R$ ) error (deg)	0.37	0.34	0.39	0.61	0.92	0.44	<b>0.51</b>
3D MED error - MRF method (mm)	0.37	0.57	0.12	0.45	0.89	0.52	<b>0.48</b>
3D MED error - Grad. desc. (mm)	7.33	7.94	6.34	8.79	9.15	9.10	<b>8.11</b>



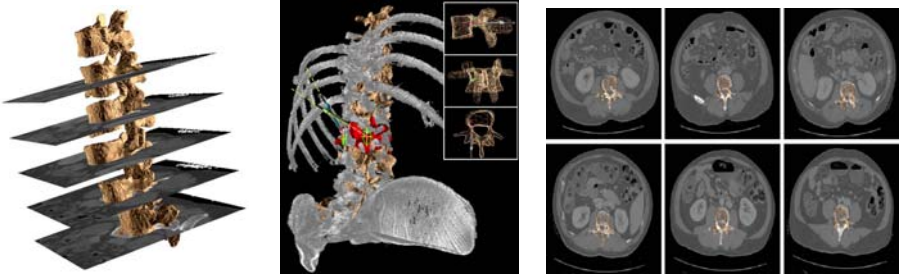
**Fig. 2.** Ground truth evaluation of multi-level MRF method using synthetic deformations on 6 typical scoliotic cases (target in red). Results show the importance of pairwise intervertebral links in the registration process compared to gradient descent.



**Fig. 3.** (a) Operating room configuration for acquiring biplanar reconstructive X-rays. (b) Box-whisker diagrams of DICE scores for the 20 operative patients.

## 4.2 Validation by Comparison of Intra-op Reconstructed X-Rays

In the next experiment, registration accuracy was determined *in-vivo* in a surgical context using intraoperative 3D models generated from X-ray images. In addition to the patient’s preoperative model, the 3D reconstruction was also obtained from X-rays taken during surgery in a setup illustrated in Fig. 3a. A set of 20 operative patients with corresponding pre- and intraoperative biplanar X-rays were selected for this experiment. We compared the average point-to-surface distances and DICE scores between source and target mesh models for all 20 patients and all vertebral levels. For thoracic and lumbar regions respectively, DICE scores were



**Fig. 4.** Visual inspection of registration results. From left to right. Global alignment of preop model with CT images. Fused 3D model for guidance of pedicle screw insertion. Series of CT slices with corresponding geometrical vertebral models.

0.91 and 0.94 (Fig. 3b shows box plots), while the mean distances were of  $2.25 \pm 0.46$  and  $2.42 \pm 0.87$  mm. While these results seem promising and confirm the ability to compensate the shape-pose changes, discrepancies can be explained from the intensity and slight shape variations between both acquisitions, which may influence the statistical shape instantiation.

### 4.3 Validation through Multi-modal Model Registration

We finally performed multi-modal medical image registration using the articulated MRF method. Data consists of 12 separate CT volumes of the lumbar and main thoracic regions obtained from different patients ( $512 \times 512 \times 251$ , resolution:  $0.8 \times 0.8$  mm, thickness: 1–2 mm), acquired for operative planing purposes. Preoperative X-rays of patients were obtained for initial 3D reconstruction. The CT data was manually annotated with 3D landmarks, corresponding to left and right pedicle tips as well as midpoints of the vertebral body. A coarse initialization is performed using a customized user interface to roughly align both models. Registration is performed to automatically align the CT dataset with  $\gamma = 0.05$  and segmentation error is estimated by measuring the average distance with the manually segmented landmarks. Table 2 presents the quantitative evaluation of this experiment with 3D landmark differences, final energy term and registration time. Results for vertebral pedicle landmark errors are  $1.62 \pm 0.57$  mm, which is promising for the required accuracy of surgical screw insertion. Visual registration results of the 3D model with CT is shown in Fig. 4, demonstrating the multi-modal alignment where one could observe accurate superposition of geometrical models on selected CT slices.

**Table 2.** Quantitative results from multi-modal registration using 12 CT datasets

Subject	1	2	3	4	5	6	7	8	9	10	11	12	Mean
3D landmark diff. (mm)	1.9	2.2	1.6	1.8	2.0	2.7	2.2	3.1	1.8	2.1	2.5	1.7	2.1
Registration time (sec)	3.8	4.5	4.8	3.3	3.7	4.1	3.5	3.5	3.2	4.2	3.8	2.9	4.4

## 5 Discussion and Future Work

We presented a method for registering preoperative images to intraoperative 3D data for spinal surgery applications. Compared to previous works, our method represents a personalized 3D spine model obtained from baseline X-rays with articulated intervertebral transformations for fast and accurate multi-modal inference through MRFs. We showed results obtained on data acquired in both X-ray and CT experiments, demonstrating good alignment for simulated and natural configurations. Simple modular data terms were able to achieve satisfactory results, although the use of alternative image data costs better capturing the spine properties with bone density could enhance the performance. Introducing prior knowledge with respect to the allowable geometric dependencies between the relative position of vertebrae is also a promising direction. Such a concept could be enhanced through a hierarchical decomposition of the spine using higher order cliques improving the accuracy and the precision of the results. By extending the framework to online cases using tracked dynamic CT, this can help surgeons improve screw insertion accuracy and reduce surgery time.

## References

1. Kim, Y., Lenke, L., Cheh, G., Riew, D.: Evaluation of pedicle screw placement in the deformed spine using intraoperative plain radiographs with computerized tomography. *Spine* 30, 2084–2088 (2005)
2. Lee, C., Kim, M., Ahn, Y., Kim, Y., Jeong, K., Lee, D.: Thoracic pedicle screw insertion in scoliosis using posteroanterior C-arm rotation method. *J. Spinal Disord. Tech.* 20, 66–71 (2007)
3. Foley, K., Simon, D., Rampersaud, Y.: Virtual fluoroscopy: computer-assisted fluoroscopic navigation. *Spine* 26, 347–351 (2001)
4. Livyatan, H., Yaniv, Z., Joskowicz, J.: Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. *IEEE Trans. Med. Imag.* 22, 1395–1406 (2003)
5. Markelj, P., Tomazevic, D., Pernus, F., Likar, B.: Robust gradient-based 3-D/2-D registration of CT and MR to X-ray images. *IEEE Trans. Med. Imag.* 27, 1704–1714 (2008)
6. Zheng, G., Dong, X.: Unsupervised reconstruction of a patient-specific surface model of a proximal femur from calibrated fluoroscopic images. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 834–841. Springer, Heidelberg (2007)
7. Boisvert, J., Cheriet, F., Pennec, X., Labelle, H., Ayache, N.: Geometric variability of the scoliotic spine using statistics on articulated shape models. *IEEE Trans. Med. Imag.* 27, 557–568 (2008)
8. Klinder, T., Wolz, R., Lorenz, C., Franz, A., Ostermann, J.: Spine segmentation using articulated shape models. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) *MICCAI 2008, Part I. LNCS*, vol. 5241, pp. 227–234. Springer, Heidelberg (2008)
9. Kadoury, S., Cheriet, F., Labelle, H.: Personalized X-ray 3D reconstruction of the scoliotic spine from statistical and image models. *IEEE Trans. Med. Imag.* (2009)
10. 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. *CVIU* 112, 14–29 (2008)
11. Lorensen, W., Cline, H.: Marching cubes: a high resolution 3-D surface construction algorithm. *Comput. Graph.* 4, 163–169 (1988)