3D Registration of Articulated Spine Models Using Markov Random Fields

Samuel Kadoury, Nikolaos Paragios

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

3D Registration of Articulated Spine Models Using Markov Random Fields

Samuel Kadoury — Nikos Paragios

N° 0364

Avril 2009

Rapport technique

3D Registration of Articulated Spine Models Using Markov Random Fields

Samuel Kadoury*, Nikos Paragios*

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 through statistical embedding and image segmentation. 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 (MRF) 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.

Key-words: Registration, physical modeling, image segmentation, articulated 3D spine model, Markov Random Field

* Laboratoire MAS, Ecole Centrale de Paris, Grande Voie des Vignes, 92295 Chatenay-Malabry, France

Recalage 3D de Modèles Articulés de la Colonne Vertébrale à partir de Markov Random Fields

Résumé : Ce papier présente une méthode d'inférence d'un modèle personnalisé de la colonne vertébrale en 3D à partir de données CT acquises dans un contexte de chirurgies correctives du rachis. Une reconstruction 3D précise à partir d'images radiographiques standards est obtenue avant l'opération afin d'offrir la géométrie des vertèbres par une modélisation statistique et une segmentation d'image. Le résultat de l'opération est exploitée comme base pour dériver un modèle articulé de la colonne représentée par une série consécutive d'articulations intervertébrales relatives aux paramètres de rotation et de translation (6 degrés de liberté). L'inférence du modèle est effectuée par rapport aux paramètres qui sont intégrés et interconnectés dans un Markov Random Field (MRF). Des valeurs potentiels unitaires et binômes mesurent respectivement le lien entre des données images (surface ou volume) avec les paramètres du modèle, et les contraintes géométriques entre les vertèbres. L'optimisation des paramètres dans un contexte multi-modale est effectuée par une approche de programmation linéaire et par dualité. Nous présentons des résultats prometteurs pour le recalage d'images à partir de données simulées et réels dans l'objectif d'une fusion d'image pour l'assistance chirurgicale.

Mots-clés : Recalage d'images, modélisation physique, segmentation, modèle 3D articulé de la colonne vertébrale, Markov Random Field

Contents

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 of the spine segment with screws and arthrodesis (bone fusion) of the intervertebral articulations. This long procedure can be very 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],[3], or 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 vertebrae remain 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 [4]. For example in [5, 6, 7], 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 Point Distribution Models (PDMs) [8] 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 [9] by incorporating knowledge-based intervertebral constraints. These 3D intervertebral transformations were transposed in [10] 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. Our approach integrates several advantages by generating a personalized 3D model from biplanar X-rays using minimal interaction, and proposing an image-based registration which avoids CT image segmentation, is modular (encodes different data-terms), 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. This makes the fusion update of the preoperative models feasible for real-time guidance procedures. 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. The paper is organized as follows. Sections 2 and 3 presents the method in terms of image and geometric-driven 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 $I_{i=\{1,2\}}$ 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 works of [11], and summarized in Fig. 1a. The 3D spine centerline $C_i(u)$, obtained from quadratic curves extracted from the images is first embedded onto a 3D database containing 732 scoliotic spines (M) to predict an initial spine, modeled by 17 vertebrae (12 thoracic, 5 lumbar), 6 points per vertebra (4 pedicle tips and 2 endplate midpoints). To map the high-dimensional 3D curve assumed to lie on a nonlinear manifold into a low-dimensional subspace, we first determine the manifold reconstruction weights W to reconstruct point i from it's K neighbors, and then determine the global internal coordinates of Y by solving:

$$\Phi(Y) = \sum_{i=1}^{M} \left\| Y_i - \sum_{j=1}^{K} W_{ij} Y_j \right\|^2. \tag{1}$$

The projection point Y_{new} is used to generate an appropriately scaled model from an analytical method based on nonlinear regression using a Radial Basis Function kernel function f to perform the inverse mapping such that $X_{\text{preop}} = [f_1(Y_{\text{new}}), ..., f_D(Y_{\text{new}})]$ with $X_{\text{preop}} = (s_1, s_2, \ldots, s_{17})$, where s_i is a vertebra model defined by $s_i = (p_1, p_2, ..., p_6)$, and $p_i \in \Re^3$ is a 3D vertebral landmark.

This crude statistical 3D model is refined with an individual scoliotic vertebra segmentation approach by extending 2D geodesic active regions in 3D, in order to evolve prior deformable 3D surfaces by level sets optimization. An atlas of vertebral meshes $S_i = \{\text{x}_{i1}, ..., \text{x}_{iN}\}$ with triangles x_j are initially positioned and oriented from their respective 6 precise landmarks s_i composing X_{preop}. The surface evolution is then regulated by the gradient map and image intensity distributions [12], where $E_{\text{RAG}} = \alpha E_{\text{CAG}}(S) + (1 - \alpha) E_{\text{R}}(S)$ is the energy function with the edge and region-based components controlled by α are defined as:

$$E_{\text{CAG}} = \sum_{i=1}^{2} \oint_{S_i} \frac{1}{1 + |\nabla I_i(\text{u}_i)|^\alpha} d\text{u}_i \tag{2}$$

$$E_{\text{R}} = -\sum_{i=1}^{2} \iint_{\Pi_i(S_i)} \log(p_R(I_i(\text{u}_i))) d\text{u}_i \tag{3}$$

with Π_i as the perspective projection parameters, and p_R is a Gaussian distribution. The projected silhouettes of the morphed 3D models would therefore match the 2D information on the biplanar X-rays in the image domain u, replicating the specifics of a particular scoliotic deformity. 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.
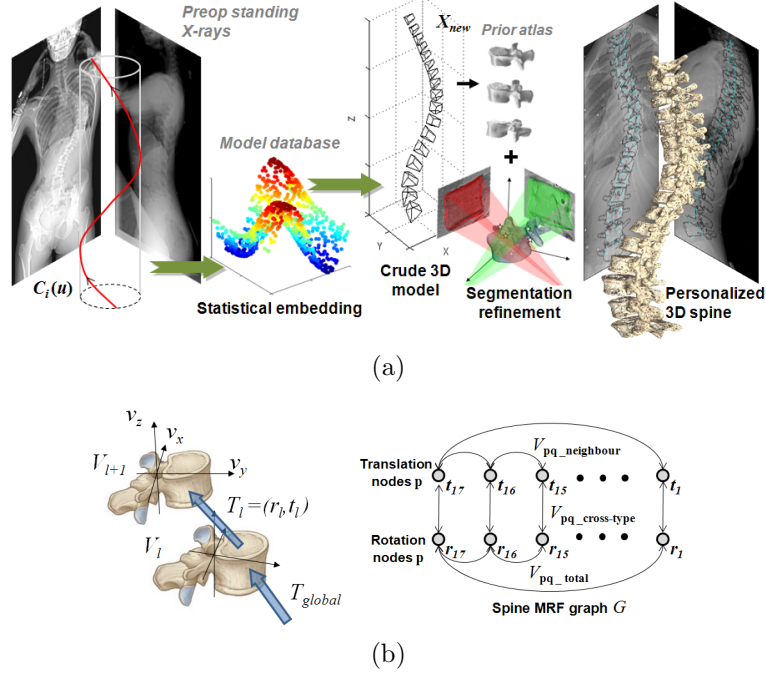
(a)

(b)

Figure 1: (a) Personalized spine 3D reconstruction from preop X-rays. (b) Articulated spine in an MRF graph, integrating three types of constrained pairwise potentials.

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 $A = [T_1, T_2, \ldots, T_N]$ as illustrated in Fig. 1b. To perform global anatomical modeling of the spine, we convert A into an absolute representation $A_{\text{absolute}} = [T_1, T_1 \circ T_2, \ldots, T_1 \circ T_2 \circ \ldots \circ T_N]$ using recursive compositions ($\circ$ is the operator of composition). 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 $x, y, t \in \Re^3$. 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 registration as a Markov Random Field (MRF) optimization where a set of labels $L = \{l^1, \ldots, l^i\}$ defined in the quantized space $\Theta = \{\mathbf{d}^1, \ldots, \mathbf{d}^i\}$ 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 source and target models is optimal for all vertebrae. The form of the MRF model is:

$$E_{\text{total}} = \sum_{\mathbf{p} \in G} V_{\mathbf{p}}(l_{\mathbf{p}}) + \sum_{\mathbf{p} \in G} \sum_{\mathbf{q} \in \mathcal{N}(\mathbf{p})} V_{\mathbf{pq}}(l_{\mathbf{p}}, l_{\mathbf{q}}) \tag{4}$$

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:

$$V_{\mathbf{p}}(l_{\mathbf{p}}) = \int_{\Omega} \eta_{\text{data}}(g(x), S_i(T_i + \mathbf{d}^{\alpha})) dT. \tag{5}$$

The data term η_{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.(4) are the pairwise potentials representing the smoothness term between vertebrae connected in the MRF (Fig. 1b). 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:

$$V_{\mathbf{pq}}(l_{\mathbf{p}}, l_{\mathbf{q}}) = \begin{cases} \lambda_{\mathbf{pq}} \|(T_{\mathbf{p}}^{\text{pre}} \times \mathbf{d}^{l_{\mathbf{p}}}) - (T_{\mathbf{q}}^{\text{pre}} \times \mathbf{d}^{l_{\mathbf{q}}})\|^2, & \text{if } \mathbf{p} \in l \text{ and } \mathbf{q} \in l+1 \\ \lambda_{\mathbf{pq}} (\|\mathbf{d}_{r_z}^{l_p} + \mathbf{d}_{r_y}^{l_p}\| - \|\mathbf{d}_{t_x}^{l_p} + \mathbf{d}_{t_z}^{l_p}\|), & \text{if } \mathbf{p} \in \Re^t \text{ and } \mathbf{q} \in \Re^R \\ \lambda_{\mathbf{pq}} |(T_{\mathbf{p}}^{\text{pre}} - T_{\mathbf{q}}^{\text{pre}}) - (\mathbf{d}^{l_{\mathbf{p}}} - \mathbf{d}^{l_{\mathbf{q}}})|, & \text{if } \mathbf{p} \equiv T_{17} \text{ and } \mathbf{q} \equiv T_1. \end{cases} \tag{6}$$

where $\lambda_{\mathbf{pq}}$ plays the role of a weighting factor defined in the spatial domain.

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 [13] [1] 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-operative CT volume images.

---

[1] Details of authors implementation : http://www.csd.uoc.gr/ komod/FastPD/

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 only displacements along the main axis was selected, with $6N + 1$ labels in 3D (N is the sampling rate). The smoothness term was set at $\lambda_{\mathbf{pq}} = 0.4$. 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 cadaver specimens was used to construct the 3D preoperative model. Models were segmented using a connecting cube algorithm. The same six precise anatomical landmarks were added on each model by an expert operator. The atlas is divided into 3 levels of polygonal mesh catalogues of increasing complexity (Fig. 2), to adopt the widely used multi-resolution registration approach.

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-operative 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 $\eta_{\mathrm{RX}} = |S_i \bigcap X_{\mathrm{intra}}|/|S_i \bigcup X_{\mathrm{intra}}|$, which represents the volume intersection between the source S_i and target model X_{intra}.

- **Volume/CT Inference Support**: the singleton data term potential defined as $\eta_{\mathrm{CT}} = \sum_{\mathbf{x}_{ij} \in S_i} (\gamma^2 + \gamma\|\nabla CT(\mathbf{x}_{ij})\|)/(\gamma^2 + \|\nabla CT(\mathbf{x}_{ij})\|^2)$ attracts mesh triangles to target high-intensity voxels in the gradient CT volume without segmentation. The term γ 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 deformations (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. Fig. 2 illustrates the near perfect alignment of the MRF approach with constrained articulations, while gradient-descend may cause vertebra collisions.

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	**0.59**
Angular (T_R) error (deg)	0.37	0.34	0.39	0.61	0.92	0.44	**0.51**
3D MED error - MRF method (mm)	0.37	0.57	0.12	0.45	0.89	0.52	**0.48**
3D MED error - Grad. desc. (mm)	7.33	7.94	6.34	8.79	9.15	9.10	**8.11**

Figure 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.

4.2 Validation by Comparison of Intra-operative 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 0.91 and 0.94 (Fig. 3b shows box plots), while the mean distances were of 2.25 ± 0.46 and 2.42 ± 0.87 mm. While these results seem promising and confirm the ability to compensate the shape-pose changes,
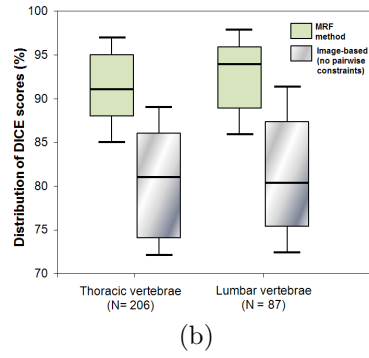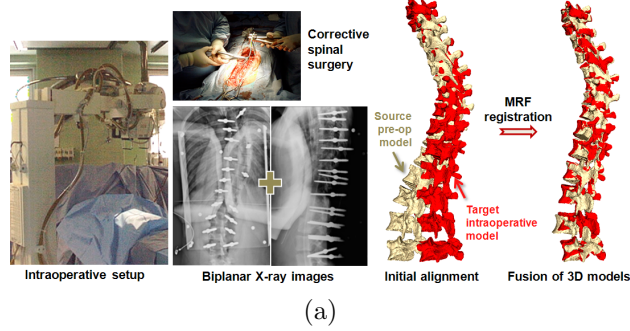
(a)

(b)

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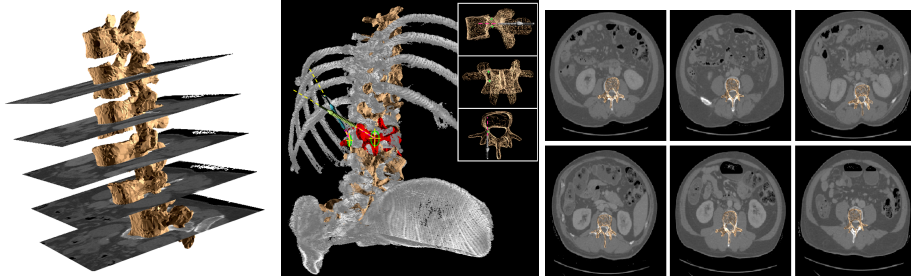

Figure 4: Visual inspection of registration results. From left to right. Global alignment of preop model with CT images. Fused 3D model for guidance of pedicle screw insertion. Series of CT slices with corresponding geometrical vertebral models.

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

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
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
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

and main thoracic regions obtained from different patients ($512 \times 512 \times 251$, resolution: 0.8×0.8 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 an 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 ± 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.

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 performs the automatic reconstruction of models from baseline X-rays using 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. This work can also be adapted to other preoperative modalities (MR) and for segmentation of skeletal spine structures. The use of alternative image costs better capturing the spine properties could greatly 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 case studies using tracked dynamic CT, this can ultimately help surgeons to learn the variations of spinal shape in complex corrective procedures, improve pedicle screw insertion accuracy and reduce surgery time.

References

[1] Kim, Y., Lenke, L., Cheh, G. and Riew, K.D.: Evaluation of pedicle screw placement in the deformed spine using intraoperative plain radiographs with CT. Spine **30** (2005) 1084-88.

[2] Lee, C., Kim, M., Ahn, Y., Kim, Y., Jeong, KI. and Lee, D.: Thoracic pedicle screw insertion in scoliosis using posteroanterior C-arm rotation method. J. Spinal Disord. Tech. **20** (2007) 66-71.

[3] Lauritsch, G., Boese, J., Wigstrom, L., Kemeth, H., and Fahrig, R.: Towards cardiac C-arm computed tomography. IEEE Trans. Med. Imag. **25** (2006) 922-34.

[4] Foley, K., Simon, D. and Rampersaud, Y.: Virtual fluoroscopy: computer-assisted fluoroscopic navigation. Spine **26** (2001) 347-51.

[5] Livyatan, H., Yaniv, Z. and Joskowicz, J.: Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. IEEE Trans. Med. Imag. **22** (2003) 1395-06.

[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 Trans. Med. Imag. **27** (2008) 1704-14.

[7] Tomazevic, D., Likar, B., Slivnik, T. and Pernus, F.: 3-D/2-D registration of CT and MR to X-ray images. IEEE Trans. Med. Imag. **22** (2003) 1407-16.

[8] Zheng, G. and Dong, X.: Unsupervised reconstruction of a patient-specific surface model of a proximal femur from calibrated fluoroscopic images. In: Proc. MICCAI. (2007) 834-41.

[9] Boisvert, J., Cheriet, F., Pennec, X., Labelle, H. and Ayache, N.: Geometric variability of the scoliotic spine using statistics on articulated shape models. IEEE Trans. Med. Imag. **27** (2008) 557-68.

[10] Klinder, T., Wolz, R., Lorenz, C., Franz, A. and Ostermann, J.: Spine segmentation using articulated shape models. In: Proc. MICCAI. (2008) 227-34.

[11] Kadoury, S., Cheriet, F. and Labelle, H.: Statistical image-based approach for the 3D reconstruction of the scoliotic spine from X-rays. In: Proc. ISBI. (2008) 660-63.

[12] Paragios, N. and Deriche, R.: Geodesic Active Regions: New paradigm to deal with frame partition problems in computer vision. Visual Comm. Image Repre. **13** (2002) 249-68.

[13] Komodakis, N., Tziritas, G. and Paragios, N.: Performance vs computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies. CVIU **112** (2008) 14-29.