A HIERARCHICAL STATISTICAL MODELING APPROACH FOR THE UNSUPERVISED 3D RECONSTRUCTION OF THE SCOLIOTIC SPINE

S. Benameur[†♯∗], *M. Mignotte*[†♯], *S. Parent*[†‡], *H. Labelle*[‡], *W. Skalli*[♮] *and J.A. De Guise*[†‡∗]

[†] Laboratoire de recherche en imagerie et orthopédie, Centre hospitalier Universitaire de Montréal, Canada
[‡] Laboratoire d'imagerie en scoliose 3D, Centre de recherche, Hôpital Sainte-Justine, Montréal, Canada
[♯] Laboratoire de vision et modélisation géométrique, DIRO, Université de Montréal, Canada
[♮] Laboratoire de biomécanique, École nationale supérieure d'arts et métiers, Paris, France
[∗] École de technologie supérieure, Montréal, Canada
E-MAIL : BENAMEUS@IRO.UMONTREAL.CA

ABSTRACT

In this paper, we propose a new and accurate 3D reconstruction technique for the scoliotic spine from a pair planar and conventional radiographic images (postero-anterior and lateral). The proposed model uses *a priori* hierarchical global knowledge, both on the geometric structure of the whole spine and of each vertebra. More precisely, it relies on the specification of two 3D templates. The first, a rough geometric template on which rigid admissible deformations are defined, is used to ensure a crude registration of the whole spine. 3D reconstruction is then refined for each vertebra, by a template on which non-linear admissible global deformations are modeled, with statistical modal analysis of the pathological deformations observed on a representative scoliotic vertebra population. This unsupervised coarse-to-fine 3D reconstruction procedure is stated as a double energy function minimization problems efficiently solved with a stochastic optimization algorithm. The proposed method, tested on several pairs of biplanar radiographic images with scoliotic deformities, is comparable in terms of accuracy with the classical CT-scan technique while being unsupervised and requiring a lower amount of radiation for the patient.

1. INTRODUCTION

Scoliosis is a 3D deformation of the natural curve of the spinal column, including rotations and vertebral deformations. To analyze the 3D characteristics of these deformations, which can be useful for the design, evaluation and improvement of orthopedic or surgical operations, several 3D reconstruction methods have been developed.

Among these methods, computerized tomography 3D reconstruction methods (e.g., X-rays or magnetic resonance) provide accurate 3D information of the human anatomy. However, the high level of radiation received by patients, the large quantity of data to be acquired and processed, and the cost of these methods make them less functional. Also, these medical imaging techniques require that the patient be in a lying position, which is incompatible with many diagnostic protocols evaluating scoliosis.

To circumvent this difficulty, an original 3D biplanar statistical reconstruction method was proposed in [1] which uses only a pair of planar and conventional (postero-anterior (I_{PA}) and lateral (I_{LAT})) radiographic images (and consequently ensures a lower amount of radiation) for patients in a standing position. Nevertheless, the above-mentioned technique remains widely supervised and requires knowledge of the six anatomical points (namely, the center of the superior and inferior end-plates, the upper and lower extremities of both pedicles) to initialize 3D reconstruction process of each vertebra of the spine.

To overcome this problem of supervision, we propose in this paper a 3D statistical reconstruction method based on the likelihood introduced in [1] but using hierarchical global *a priori* knowledge on the geometric structure of both the whole spine and each vertebra. More precisely, the hierarchical model we propose relies on the specification of two 3D templates. The first, a rough geometric template on which rigid admissible deformations are defined, is used to fit its (postero-anterior and lateral) projections with the preliminary segmented contours of each cubic approximation of each vertebra body of the spine on the two calibrated radiographic views. It ensures crude registration of the whole spine and gives a rough position and orientation of each vertebra. 3D reconstruction is then refined by a second template which takes into account *a priori* global knowledge on the geometric structure of each vertebra. This geometric knowledge is efficiently captured by a statistical deformable template integrating a set of admissible deformations, expressed by the first modes of variation in Karhunen-Loeve (KL) expansion, of the pathological deformations observed on a representative scoliotic vertebra population.

This paper is organized as follows. Section 2 presents the hierarchical prior model used in our coarse-to-fine reconstruction method. Section 3 briefly recalls the likelihood model and the 3D/2D registration strategy introduced in [1]. Section 4 presents energy function minimization related to coarse-to-fine registrations and the stochastic optimization procedure used to estimate optimal reconstruction. In Section 5, we show some 3D reconstruction results. Finally, the conclusion is given in Section 6.

2. HIERARCHICAL PRIOR MODEL

2.1. Crude prior model of the spine

To ensure crude registration of the whole spine and to estimate a rough position $t = (t_x, t_y, t_z)$, scale k and orientation α of each vertebra, we first consider a crude *a priori* geometric model for the whole spine.

This model relies on a set of cubic templates, roughly representing each vertebral body and stacked on top of one another to form the spinal column (cf. Fig. 4). Each cubic template (defined by a set of control point vectors of the cubic representation) associated with each vertebral level has its scale, orientation and position constrained within a restricted domain whose center is given by knowledge of the (previously) estimated parameters of the cubic template which is located below. For example, if the registration of the whole crude spine is made from bottom to top, then the position $t^l = (t^l_x, t^l_y, t^l_z)$, scale vector $k^l = (k^l_x, k^l_y, k^l_z)$ and orientation $\alpha^l = (\alpha^l_x, \alpha^l_y, \alpha^l_z)$ at (vertebral) level l, given the rough parameter estimation of vertebra below the level l (i.e., $l-1$), are restrained within the domain defined by,

$$
\begin{aligned}
t^{l-1}_x + k^{l-1}_x - \Delta t^l_x \leq \quad & t^l_x \quad \leq t^{l-1}_x + k^{l-1}_x + \Delta t^l_x, \\
t^{l-1}_y - \Delta t^l_y \leq \quad & t^l_y \quad \leq t^{l-1}_y + \Delta t^l_y, \\
t^{l-1}_z - \Delta t^l_z \leq \quad & t^l_z \quad \leq t^{l-1}_z + \Delta t^l_z, \\
k^{l-1} - \Delta k^l \leq \quad & k^l \quad \leq k^{l-1} + \Delta k^l, \\
\alpha^{l-1} - \Delta \alpha^l \leq \quad & \alpha^l \quad \leq \alpha^{l-1} + \Delta \alpha^l,
\end{aligned}
$$

where α is the rotation vector (ensuring the rotation in x, y or z axes). Δt_x, Δt_y, Δt_z, Δk and $\Delta\alpha$ are given by statistical knowledge on the scoliotic deformation of the spine [7][8]. A global configuration of the deformable spine model is thus described by 9 rigid transformation parameters for each cubic template associated with each vertebral level.

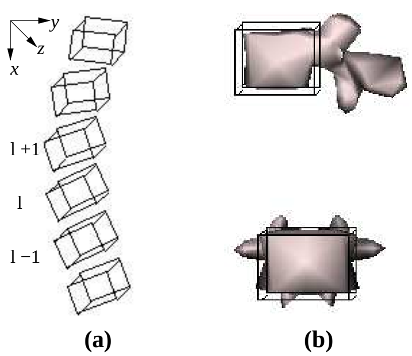

FIG. 1 – *Crude prior model of the spine.* **(a)** *Deformable model of the whole spine,* **(b)** *cubic template representation associated with each vertebra.*

2.2. Fine prior model of each vertebra

Our *a priori* knowledge model relies also on the description of each vertebra by a 3D deformable template (i.e., a vector $s \in R^{3n}$ of *n* control points) which incorporates statistical knowledge about its geometrical structure and its pathological variability. The deformations of this template are expressed by the first modes of variation in KL expansion of pathological deformations observed on a representative training scoliotic vertebra population [1]. This can be done by using principal component analysis (PCA), i.e., by computing the covariance matrix C of shapes $\{s_i\}$. The main deformation modes of the template model s are then described by the eigenvectors ϕ of C, with the largest eigenvalues λ (cf. Fig. 2). The globally deformed template is defined by,

$$s \;=\; M(k,\alpha)[\bar{s} + \Phi b] + T, \tag{1}$$

where,

- T and $M(k,\alpha)$ account for rigid deformations of the template, (T is a global translation vector, and $M(k,\alpha)$ performs a rotation (in the x, y or z axes) and a scaling by k).
- $\Phi = (\phi_1, \ldots, \phi_m)$ is the matrix of the first m eigenvectors associated with the m largest eigenvalues and $\mathbf{b} = (b_1, \ldots, b_m)^T$ is a vector containing the weights for these m deformation modes.

A global configuration of the deformable vertebra template is thus described by $7 + m$ parameters corresponding to rigid transformations and m modal weights b_j. To penalize the deviation of the deformed template from the mean shape (and not the affine transformations), we define the energy term of this prior model by,

$$E_p(s) \;=\; \frac{1}{2}\sum_{i=1}^{m} \frac{b_i^2}{\lambda_i}, \tag{2}$$

which is close to the Mahalanobis distance [1].

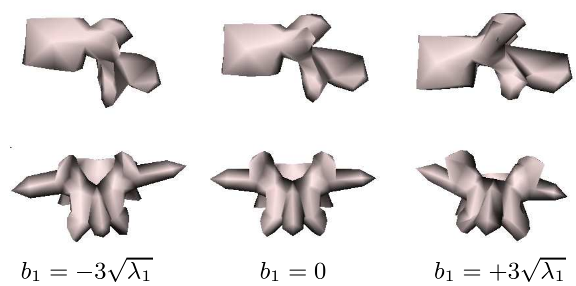

FIG. 2 – *Fine prior model of each vertebra. Two deformed shapes obtained by applying ± 3 standard deviations of the first deformation mode to the mean shape of $L3$ vertebra (middle row) and from the sagital (top row) and coronal views (bottom row).*

This low parametric representation for each vertebra level, along with the crude parametric representation of the whole spinal column constitutes our global hierarchical *a priori* model that will be used to rightly constrain the ill-posed nature of our proposed coarse-to-fine 3D reconstruction method.

3. LIKELIHOOD MODEL

As proposed in [1], the likelihood model is expressed by a measure of similarity between the external contour of the (postero-anterior and the lateral) projections of the 3D deformed template and a directional edge potential field estimated on the two radiographic views. This likelihood energy term is defined by,

$$E_l(s, I_{\text{PA}}, I_{\text{LAT}}) = -\frac{1}{n_{\text{PA}}}\sum_{\Gamma_{\text{PA}}} \Psi_{\text{PA}}(x,y) - \frac{1}{n_{\text{LAT}}}\sum_{\Gamma_{\text{LAT}}} \Psi_{\text{LAT}}(x,y), \tag{3}$$

where the summation of the first and second term of E_l is overall the n_{PA} and n_{LAT} points of the external contour of the respectively lateral and postero-anterior perspective projections of the deformed template on the two pre-computed edge potential fields of each radiographic image.

To compute the edge potential field Ψ associated with each radiographic view, we first use a Canny edge detector with the unsupervised technique proposed in [2]. Then, Ψ is defined as in [4] by,

$$\Psi(x,y) \;=\; \exp\left(-\frac{\sqrt{\xi_x^2+\xi_y^2}}{\tau}\right)\left|\cos\big(\gamma(x,y)\big)\right|,$$

where $\xi = (\xi_x, \xi_y)$ is the displacement to the nearest edge point in the image, and τ is a smoothing factor which controls the degree of smoothness of this potential field. $\gamma(x, y)$ is the angle between the tangent of the nearest edge and the tangent direction of the contour at (x, y) (cf. Fig. 3).

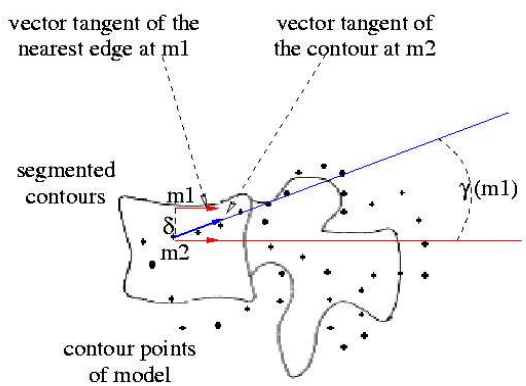

FIG. 3 – *Directional component used in the directional edge potential field* $\Psi(x, y)$.

4. COARSE-TO-FINE OPTIMIZATION STRATEGY

This unsupervised coarse-to-fine 3D reconstruction procedure is stated as a double energy function minimization problems, namely,

$$E(s,\theta) \;=\; E_l(s, I_{\text{PA}}, I_{\text{LAT}}) + \beta E_p(s), \tag{4}$$

where E_l is the likelihood energy term and E_p is the prior energy term (or the regularization term). β is a factor to control the balance between the two energy components.

• This optimization problem is first used with the template presented in Subsection 2.1 for the registration of the whole spine. In this crude reconstruction step, there is no prior energy term $E_p = 0$ since we do not use non-linear deformations. The ill-posed reconstruction problem is nevertheless constrained by the low parametric representation of the whole spine. This step requires identification the center of the inferior end-plate of the lowest vertebra on the two radiographic views of the spine with a simple graphic interface. User interaction is limited to simply placing this point on the two radiographic images. The 3D coordinates of this landmark were obtained by direct linear transformation (DLT) [5]. This crude registration estimates the rough position $t = (t_x, t_y, t_z)$, scale k and orientation α of each vertebra.

• This optimization problem is then used with the template presented in Subsection 2.2 and the prior energy term given by Eq. (2). Optimization is made within a range of values around the rigid parameters roughly estimated by this first crude reconstruction step.

In our application, due to the complexity and number of variables involved in the energy function expressed by Eq. (4), we use the stochastic optimization algorithm recently proposed in [6] (see Algo. 1.). We use this algorithm because the adjustment of all internal parameters does not depend on the function to minimize, and convergence is asymptotically ensured [6]. Let us add that this optimization algorithm is also especially well suited to minimize complex energy function [3].

E/S Algorithm

$E(.)$	A real-valued l-variable function, defined on F, to be minimized
F	A finite discrete subset of the Cartesian product $\Pi_{j=1}^{l}[m_j, M_j]$
n	The size of the population
r	A real number $\in [0,1]$ called the radius of exploration
$\mathcal{N}(a)$	The neighborhood $\{b \in F$: for some j, $\lvert b_j - a_j\rvert \le r(M_j - m_j)$, $b_i = a_i$, $i \neq j\}$
$D = l/r$	The diameter of the exploration graph
θ	$\theta = (\theta_1, \ldots, \theta_n)$, an element of F^n
$\hat{\theta} \in F$	$\hat{\theta} = \arg\min_{\theta_i \in (\theta_1,\ldots,\theta_n)} E(\theta_i)$, i.e., the minimal point in θ with the lowest label
p	The probability of exploration
k	The iteration step

1. Initialization
Random initialization of $\theta = (\theta_1, \ldots, \theta_n) \in F^n$
$k \leftarrow 2$

2. Exploration/Selection
repeat

1. Compute $\hat{\theta}$; $\quad \hat{\theta} = \arg\min_{\theta_i \in \theta} E(\theta_i)$
2. Draw m according to the binomial law $b(n, p)$
   - For $i \le m$, replace θ_i by $\vartheta_i \in \mathcal{N}(\theta_i) \setminus \hat{\theta}$ according to the uniform distribution
   - For $i > m$, replace θ_i by $\hat{\theta}$
3. $k \leftarrow k + 1$ and $p = k^{-1/D}$

until *a stopping criterion is met*;

Algorithm 1: E/S optimization algorithm

5. EXPERIMENTAL RESULTS

In our application, we use the training base described in [9]. The first *m* eigenmodes $\{(\lambda_1, \phi_1), \ldots, (\lambda_m, \phi_m)\}$ are chosen to cover at least 90% of the population's variability. For the experiments, we have chosen $\beta = 0.02$ for the weighting factor penalizing the prior energy term with respect to external energy. We used the Canny edge detector to estimate the edge maps which are then used for estimation of the edge *potential fields* on the two radiographic views that will be applied in the likelihood energy term. In our application, sigma=1, mask size is 5x5, and the lower and upper thresholds are given by the unsupervised estimation technique proposed in [2]. We fix the size of population to 100 and the number of iterations to 50. The E/S algorithm takes about 30-40 generations to converge to the true solution for each vertebral level. In fact, the convergence rate can vary significantly depending on the complexity of the energy function $E(s, \theta)$ to the minimized (or the quality of the input images). Our 3D reconstruction method takes between 200 seconds and 260 seconds on a PC workstation 650MHz for the images presented here. Fig. 4 shows an example of 3D reconstruction of vertebra segment (L1/L2/L3/L4/L5) from two radiographic images (postero-anterior and lateral) of scoliotic spine.

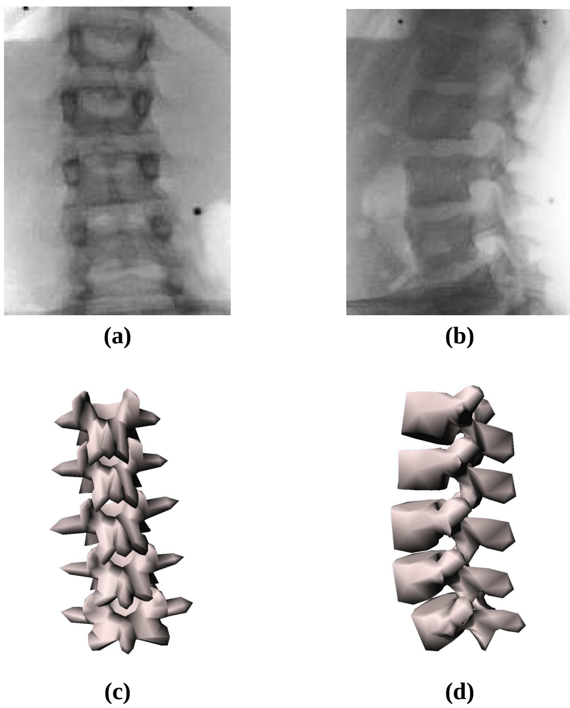

FIG. 4 – *3D reconstruction of vertebra segment of scoliotic spine.* **(a)** *Postero-anterior,* **(b)** *Lateral image (512x256 pixels),* **(c)** *and* **(d)** *Visualization of reconstruction vertebra segment (L1/L2/L3/L4/L5) from the sagittal and coronal views.*

6. CONCLUSION

We have presented an original coarse-to-fine approach for the 3D reconstruction of the scoliotic spine using contours extracted from biplanar radiographic images and *a priori* hierarchical global knowledge both on the geometrical structure of the whole spine and on the statistical structure of each vertebra. From an algorithmic point of view, the reconstruction problem is decomposed in two optimization problems of reduced complexity allowing us to drastically save computational effort and/or to provide accelerated convergence toward an improved estimate. In our application, these two optimization problems are efficiently solved with a stochastic optimization procedure. Let us note that the estimated global deformation parameters after reconstruction (i.e., the parameter vector *b*, setting the amplitude of each deformation mode of the scoliotic deformations) and the measures of morphometric parameters can then be used to quantify the scoliosis, its nature or to analyze the improvement of orthopedic or surgical corrections. This method has been tested on a number of pairs of radiographic images demonstrating its efficiency and robustness. The proposed scheme thus constitutes an alternative to CT-scan 3D reconstruction with the advantage of low irradiation and will be of great interest for 3D clinical applications and for reliable geometric models for finite element studies. The proposed method remains sufficiently general to be applied to other medical reconstruction problems (i.e., rib cage, pelvis, knee, etc.) for which the database of this anatomical structure is available (with two or more radiographic views). We now intent to improve the proposed method by refining the statistical model by local deformations.

Acknowledgements : The authors thank the Natural Sciences and Engineering Research Council of Canada, Canadian Foundation for Innovation, Valorisation recherche Québec and the Biospace Company, Paris, France, for supporting this study. The authors are also grateful to François Destrempes who has helped to implement the E/S algorithm.

7. REFERENCES

[1] S. Benameur, M. Mignotte, S. Parent, H. Labelle, W. Skalli, and J. De Guise. “3D Biplanar reconstruction of scoliotic vertebrae using statistical models”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR’2001*, Vol. 2, pp. 577-582, Kauai Marriott, Hawaii, USA, 2001.

[2] J. F. Canny. “A computational approach to edge detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.8, N$^\circ$6, pp.679-697, 1986.

[3] F. Destrempes, and M. Mignotte. “Unsupervised localization of shapes using statistical models”, *International Conference on Signal and Image Processing, IASTED*, pp.60-65, Kauai Marriott, Hawaii, USA, 2002

[4] A. K. Jain, Y. Zhong, and S. Lakshmanan. “Object matching using deformable templates”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.18, N$^\circ$3, pp.267-278, 1996.

[5] G. T. Marzan. “Rational design for close-range photogrammetry”, Ph.D. thesis, Department of Civil Engineering, University of Illinois at Urbana-Champaign, USA, 1976.

[6] F. Olivier. “Global optimization with exploration/selection algorithms and simulated annealing”, *The Annals of Applied Probability* , Vol.12, pp.248-271, 2002.

[7] M. M. Panjabi, K. Takata, V. Goel, D. Frederico, T. Oxland, J. Duranceau, M. Krag.“Thoracic human vertebrae : quantitative three-dimensional anatomy”, *Spine*, Vol.16, N$^\circ$8, pp.888-901, 1991.

[8] M. M. Panjabi, V. Goel, T. Oxland, K. Takata, J. Duranceau, M. Krag, M. Price.“Lumbar human vertebrae : quantitative three-dimensional anatomy”, *Spine*, Vol.17, N$^\circ$3, pp.299-306, 1992.

[9] S. Parent, H. Labelle, W. Skalli, Bruce Latimer, and J. A. De Guise. “Morphometric analysis of anatomic scoliotic specimens”, *Spine*, Vol.27, N$^\circ$21, pp.2305-2311, 2002.