33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Interactive 3D Reconstruction of the Spine from Radiographs Using a Statistical Shape Model and Second-Order Cone Programming

Jonathan Boisvert and Daniel C. Moura

Abstract— **Three-dimensional models of the spine are commonly used to diagnose, to treat, and to study spinal deformities. Creating these models is however time-consuming and, therefore, expensive. We propose in this paper a reconstruction method that finds the most likely 3D reconstruction given a maximal error bound on a limited set of landmark locations supplied by the user. This problem can be solved using second-order cone programming, leading to a globally convergent method that is considerably faster than currently available methods. A user can, with our current implementation, interactively modify the landmark locations and receive instantaneous feedback on the effect of those changes on the 3D reconstruction instead of blindly selecting landmarks. The proposed method was validated on a set of 53 patients who had adolescent idiopathic scoliosis using real and synthetic tests. Test results showed that the proposed method is considerably faster than currents methods (about forty times faster), is extremely flexible, and offers comparable accuracy.**

I. INTRODUCTION

Three-dimensional models of the spine are now commonly used to diagnose, study and plan appropriate treatment of spinal deformities. These models are most commonly reconstructed from radiographs since this image modality allows patients to stand up during the examination, is inexpensive, and is widely available. However, anatomical structures such as lungs, ribs, pelvis, and vertebrae are superimposed, which makes human intervention needed.

For almost twenty-years, the most common method to reconstruct the spine was to manually identify landmarks in two or more calibrated radiographs and to later triangulate the position of those landmarks [1]. A qualified technician had to typically identify between 136 and 850 landmarks. The method was, of course, considered expensive and time-consuming.

The time and effort needed to obtain good reconstructions was indeed limiting clinical use of those models as well as their prevalence in biomechanics research. It was thus proposed to reduce the user-input to the four corners of the vertebral bodies as seen on the radiographs [2]. The remaining anatomical landmarks were then inferred based on a statistical model of the shape of individual vertebrae. This method speeded up the reconstruction process significantly, but each vertebrae of interest had to be manually handled.

To alleviate this problem, Dumas et al. [3] proposed to compute the position and scale of the vertebrae using an interpolation technique. This meant that not all vertebrae needed to be manually adjusted. However, vertebral deformations were left out.

J. Boisvert is with the Institute for Information Technology of Canada's National Research Council, 1200 Montreal Road, Ottawa, Canada jonathan.boisvert@nrc-cnrc.gc.ca

D. Moura is with Universidade do Porto, Faculdade de Engenharia, Departamento de Engenharia Informática, daniel.moura@fe.up.pt

A method based on statistical inference performed on an articulated model representing the whole spine (including both global and local deformations) was later proposed by Boisvert et al. [4]. The method was able to successfully reconstruct complete spine models even when a limited number of 3D landmarks were available. However, the execution time associated with the method was incompatible with clinical constraints.

Humbert et al. [5] combined simpler statistical models of the vertebral shapes and inter-vertebrae shape dependency with features such as spinal centerline and selected landmarks. The authors report reconstruction times around three minutes for severe cases of scoliosis. Moura et al. [6] later proposed a method where interaction was reduced to identifying the spinal centerline with splines, which were then used to estimate the most likely articulated model representing the spine.

A few methods were also designed to directly use the information contained in the radiographs (see [7], [8], [9] for example), which can increase the repeatability and accuracy of the reconstructions. However, those methods all rely heavily on user-supplied initialization, which is usually given by one of the previously described methods.

Unfortunately, none of the current methods is fast enough to provide instantaneous feedback as the user is working with the reconstruction software. Most of the current methods take from several seconds to a few minutes to run. The underlying reason is that most reconstruction methods rely on non-linear optimization methods. Those optimization methods are prone to find local minimums instead of the global one and do not (in general) have reliable execution times.

We propose in this manuscript a novel fast 3D reconstruction method that is based on a convex optimization problem. More specifically, reconstructing the spine is formulated as a second-order cone programming problem, which leads to a method that is considerably faster than current methods. The resulting 3D reconstruction is obtained with arbitrary fixed precision in polynomial time without any initialization. We also propose a new user-interaction model that takes advantage of the performances of the proposed method to provide real-time feedback to the user.

II. METHOD

The proposed method aims at minimizing the distance from a reconstructed 3D spine model to a prior distribution of

3D spine models while constraining the differences between the reconstructed model and the user input to remain within an acceptable range (which can be specified by the user). The next subsections provide more details about the statistical modeling, the error computations, and the integration of these two aspects into a second-order cone program (SOCP).

A. Statistical Model

There are many ways to build a statistical model of the spine. It could be done using an *ad hoc* parametric model [5], using articulated modeling [4] or even with the help of a non-linear dimension reduction technique such as locally linear embedding [9]. These methods allow compact representation of the anatomical variability seen in patients. However, the reconstruction of a 3D spine model from a lower-dimension space then becomes a non-linear operation that cannot be addressed by SOCP. We therefore selected a simpler multi-variate approach.

Let N be the number of points considered per vertebrae and M be the number of vertebrae. A spine model X can thus be represented either as a set of NM points $x_i \in \Re^3$ with $1 \leq i \leq NM$ or as a single column-vector $X \in \Re^{3NM}$.

Given a set of k previously reconstructed models, one can compute the associated mean $\mu = \frac{1}{k}\sum_{j=1\ldots k} X_j$ and covariance matrix $\Sigma = \frac{1}{k-1}\sum_{j=1\ldots k}(X_j-\mu)(X_j-\mu)^T$.

To reduce the dimensionality of the model, it is possible to apply principal components analysis (PCA). This type of analysis successively decomposes the original space into orthogonal dimensions (or components) along which the remaining projected variance is maximized. It yields a matrix $A^{(X)}$ that can be used to compute a new spine model X_{new} based on a vector of PCA weights $\alpha^{(X)}$ as $X_{new} = A^{(X)}\alpha^{(X)} + \mu$ and a vector σ which contains the variances of these weights. The same matrix $A^{(X)}$ can also be used to compute the PCA weights of a new shape X_{new} as $\alpha^{(X)} = A^{(X)T}(X_{new} - \mu)$. To reduce the dimensionality one simply needs to consider the first n components of $\alpha^{(X)}$ and treat the remaining weights as being equal to zero (the components are ordered by decreasing variance).

The Mahalanobis distance is a widely used metric to determine the level of similarity of a multivariate sample to a known distribution. It only requires the mean and covariance matrix of the prior distribution and is given by:

$$D_M = \sqrt{(X-\mu)^T \Sigma^{-1} (X-\mu)}. \tag{1}$$

This distance will be used to ensure that the reconstruction that is the most similar to the prior distribution will be returned when more than one could explain the user input. This happens, for example, when the user identifies a few points to reconstruct a complete spine model.

Second-order cone programming solvers do, however, expect to operate with linear expressions or their norm. We therefore need to express the Mahalanobis distance as:

$$\begin{aligned} D_M^2 &= (X-\mu)^T \Sigma^{-1} (X-\mu) \\ &= \left(L^T(X-\mu)\right)^T L^T (X-\mu) \\ &= \|L^T(X-\mu)\|_2^2 \end{aligned} \tag{2}$$

where L is the result of a Cholesky decomposition of Σ^{-1}.

B. Landmarks Projection Error

X-ray image formation process can be modeled using a pinhole camera. The projection of a 3D point (x_i) to image coordinates (u_i^j) is thus given by the standard equation:

$$\begin{pmatrix} u_i^j \\ 1 \end{pmatrix} \propto P^j \begin{pmatrix} x_i \\ 1 \end{pmatrix} \tag{3}$$

where P^j is the projection matrix associated with the j^{th} radiograph. The squared error between the projection of a 3D point x_i and its empirical measure on the radiograph $\tilde{u}_i^j$ can then be written as:

$$\|u_i^j - \tilde{u}_i^j\|_2 = \frac{1}{P_3^j(x_i,1)^T} \left\| \begin{pmatrix} P_1^j - P_3^j \tilde{u}_{i,x}^j \\ P_2^j - P_3^j \tilde{u}_{i,y}^j \end{pmatrix} \begin{pmatrix} x_i \\ 1 \end{pmatrix} \right\|_2, \tag{4}$$

where P_i^j is the i^{th} line of the projection matrix P^j.

It has been demonstrated that limiting this error to e_{max} while optimizing for the point position is a second-order cone constraint [10]. Provided that $P_3^j \begin{pmatrix} x_i & 1 \end{pmatrix}^T \geq 0$ (*i.e.* the point is in front of the X-ray source and not behind it), this constraint can be expressed in standard form as:

$$\left\| \begin{pmatrix} P_1^j - P_3^j \tilde{u}_{i,x}^j \\ P_2^j - P_3^j \tilde{u}_{i,y}^j \end{pmatrix} \begin{pmatrix} x_i \\ 1 \end{pmatrix} \right\|_2 \leq e_{max} P_3^j \begin{pmatrix} x_i & 1 \end{pmatrix}^T \tag{5}$$

C. Optimization Programs

We minimize the Mahalanobis distance while constraining the solution to result in a projection error smaller than e_{max} using the following program:

$$\begin{aligned} \underset{X,t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \left\|L^T(X-\mu)\right\|_2 \leq t \\ & \left\| \begin{pmatrix} P_1^j - P_3^j \tilde{u}_{i,x}^j \\ P_2^j - P_3^j \tilde{u}_{i,y}^j \end{pmatrix} \begin{pmatrix} x_i \\ 1 \end{pmatrix} \right\|_2 \\ & \leq e_{max} P_3^j \begin{pmatrix} x_i & 1 \end{pmatrix}^T, \end{aligned} \tag{6}$$

which is a standard SOCP problem constructed by combining Eq. 2, Eq. 5, and the auxiliary variable t. This formulation leads to a rather large number of variables. To further improve computational performance, it is possible to reduce the number of variables by optimizing PCA weights instead of point's coordinates. The optimization program then becomes:

$$\begin{aligned} \underset{\alpha,t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \left\|\text{diag}(1/\sqrt{\sigma})\alpha\right\|_2 \leq t \\ & \left\| \begin{pmatrix} P_1^j - P_3^j \tilde{u}_{i,x}^j \\ P_2^j - P_3^j \tilde{u}_{i,y}^j \end{pmatrix} \begin{pmatrix} A_i\alpha + \mu_i \\ 1 \end{pmatrix} \right\|_2 \\ & \leq e_{max} P_3^j \begin{pmatrix} A_i\alpha + \mu_i & 1 \end{pmatrix}^T, \end{aligned} \tag{7}$$

where A_i and μ_i designate the lines of A and μ associated with point x_i.

Several SOCP solvers are currently available on the market (both free and commercial). We choose to use SeDuMi [11], which is available as Matlab sub-routines and is sufficiently efficient for providing real-time reconstructions with the proposed method.

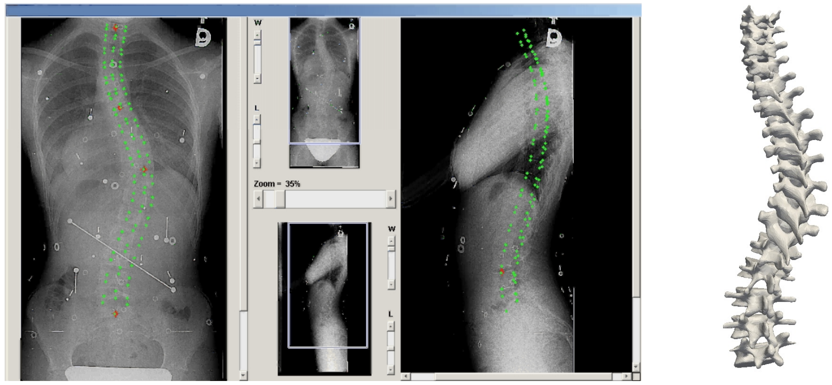

Fig. 1. Example of typical use of the proposed method. Left: User interface showing a posterior-anterior and a lateral radiograph, as well as landmarks identified by the user in red, and projections of selected anatomical landmarks from the reconstructed 3D model to the radiographs in green. Right: Corresponding three-dimensional reconstruction of the spine (reconstructed using the proposed method).

D. User Interface

The user is first presented with the two radiographs of the patient and then identifies anatomical landmarks by clicking on them. A new 3D reconstruction is produced each time a landmark is added or modified (see Figure 1). By default, the first landmark is labeled as the center of the inferior end-plate of the last lumbar vertebrae (although the user can easily specify any landmark). Subsequent landmarks are labeled according to the closest landmark in the current projection of the 3D reconstruction. Because the reconstruction process is very efficient, the user can also modify landmarks' positions by dragging them with the mouse. Three-dimensional reconstructions are then produced and displayed as the landmark is moved. The user, therefore, has full interactive control on the three-dimensional reconstruction as he adds or adjusts the control points.

III. RESULTS

A database of 307 scoliotic patients was utilized to validate the proposed method; 254 were used to compute the mean, covariance, and PCA decomposition while the remaining 53 cases were used to assess the method's performance (mean age was 13.5 years and standard deviation was 2 years).

Six landmarks per vertebrae were used and all the patients' exams were previously reconstructed using a reference method [1]. Typical radiographs' image size was 1190 by 1959 pixels. All the experiments reported in this section were performed on a desktop computer equipped with 6 GB of RAM and an Intel Core(TM) i7 CPU cadenced at 3 GHz. A total of 40 principal components were used in the optimization programs (accounting for 99.9% of the observed variance).

In a series of experiments, we evaluated the 3D reconstruction error and execution time as function of the number of landmarks used as input. The anatomical landmarks used were the centers of the inferior end-plate of the vertebrae and were evenly distributed along the spine. Furthermore, e_{max} was set to six pixels and a uniform noise of six pixels was applied to all landmarks coordinates. Figure 2 and 3 summarizes the results of the experiment. As expected, the time needed to reconstruct a spine increases and the error decreases as input points are added. Reconstruction error decreases quite rapidly as the first few landmarks are added but improvements become small after approximatively six control points per radiographs. This could mean that the accuracy is limited by e_{max} after approximatively six points in this experiment. Using a smaller value could potentially improve the results, but would also be more demanding for the user. Execution time increases almost linearly and the 95^{th} percentile is approximatively 25% worse than the average reconstruction, which means the variations in the execution times are small enough not to affect the user's experience.

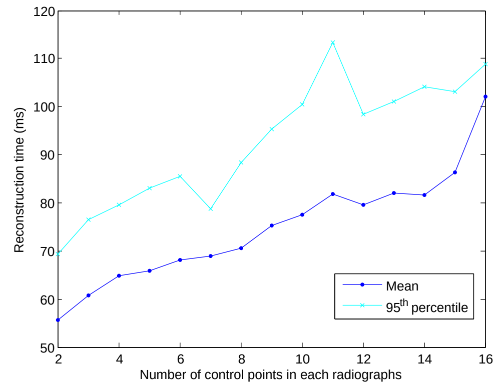

Fig. 2. Execution time (no user interaction included) as a function of the number of control points used to perform the reconstruction.

Table I compares the execution time of the proposed method (for seven control points in each radiograph) to other methods (as reported in the literature). The proposed method appears to be about forty to sixty times faster. However, the execution times were compiled from the literature and were thus not measured on the same computer, which means the actual speed-up may be slightly different.

We also performed an experiment with landmark coordinates provided by a qualified technician. For a total of 15 patients, we used seven inferior end-plates coordinates (on the two radiographs) to reconstruct a spine model using

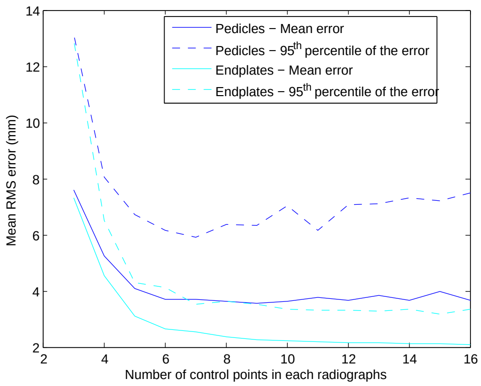

Fig. 3. Mean 3D error obtained with the proposed reconstruction method (with respect to a ground truth) as a function of the number of control points.

TABLE I

COMPARISON OF THE EXECUTION TIMES WITH PREVIOUS METHODS.

Method	Execution time
Humbert et al. [5]	4000ms
Moura et al. [6]	3000ms
Proposed method	**70ms**

the proposed method and compared it with a spine model reconstructed with a reference model [1] (which needs 102 manually identified landmarks for each radiograph). The parameter e_{max} was set to 6 pixels for PA radiographs and 12 pixels for lateral radiographs. A mean absolute difference of $3.9mm$ for the end-plates and $4.6mm$ for the pedicles was obtained. These numbers are slightly higher than what was presented in Figure 3, however in this case there is no ground truth. These differences are a combination of the error associated with the reference method and the error associated with the proposed method. Moreover, the landmark identification error is in all likelihood not strictly uniformly distributed.

IV. DISCUSSION AND CONCLUSION

The proposed method is fast, provides good accuracy, and is intuitive for the user. However, the user needs to select an acceptable value for the maximum allowable error (e_{max}). If this value is too low, then the optimization may not be feasible and no 3D reconstruction will be generated. On the other hand, if the value is too large, then the reconstruction method does not take full advantage of the user input. It would be possible to integrate e_{max} in the optimization process by applying a bisection procedure [10]. The execution time would, however, increase significantly.

Nevertheless, experimenting with different values of e_{max} is easy (it only implies selecting a different value in a drop-down menu); so different values can be tried by the user. We currently use one global e_{max} value for all the landmarks in a radiograph. One possible improvement to the method might be to allow the user to adjust e_{max} for each point (perhaps by dragging a circular target sign around the landmark). This would acknowledge the fact that certain landmarks are easy to locate accurately while others are more challenging.

In summary, we proposed a novel reconstruction method that creates 3D spine models from radiographs. The reconstruction problem was expressed as a convex problem that was solved using second-order cone programming, leading to a large improvement in execution time. This improvement then allowed for a different user interaction paradigm where 3D spine reconstructions are generated in real-time while users add information. A preliminary validation of the method was performed and indicates that its accuracy is at least comparable to current reconstruction methods. Experiments performed also lead us to believe that the flexibility, the efficiency and the user-friendliness of the proposed reconstruction method are a step toward a more pervasive use of 3D reconstructions in spinal deformity care.

V. ACKNOWLEDGMENTS

The authors would like to express their gratitude to Dr. Hubert Labelle and Montreal's St.-Justine Hospital personnel for providing access to existing 3D reconstructions.

REFERENCES

[1] B. Andre, J. Dansereau, and H. Labelle, "Effect of radiographic landmark identification errors on the accuracy of three-dimensional reconstruction of the human spine," *Medical and Biological Engineering and Computing*, vol. 30, no. 6, pp. 569–575, 1992.

[2] V. Pomero, D. Mitton, S. Laporte, J. A. de Guise, and W. Skalli, "Fast accurate stereoradiographic 3D-reconstruction of the spine using a combined geometric and statistic model," *Clinical Biomechanics*, vol. 19, no. 3, pp. 240–247, Mar. 2004.

[3] R. Dumas, B. Blanchard, R. Carlier, C. de Loubresse, J. Le Huec, C. Marty, M. Moinard, and J. Vital, "A semi-automated method using interpolation and optimisation for the 3D reconstruction of the spine from bi-planar radiography: a precision and accuracy study," *Medical and Biological Engineering and Computing*, vol. 46, no. 1, pp. 85–92, 2008.

[4] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, and N. Ayache, "Articulated spine models for 3-D reconstruction from partial radiographic data," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 11, pp. 2565–2574, Nov. 2008.

[5] L. Humbert, J. D. Guise, B. Aubert, B. Godbout, and W. Skalli, "3D reconstruction of the spine from biplanar x-rays using parametric models based on transversal and longitudinal inferences," *Medical Engineering & Physics*, vol. 31, no. 6, pp. 681–687, 2009.

[6] D. C. Moura, J. Boisvert, J. G. Barbosa, H. Labelle, and J. M. R. S. Tavares, "Fast 3D reconstruction of the spine using a deformable articulated model," *Medical Engineering & Physics*, 2011 (in press).

[7] T. Cresson, R. Chav, D. Branchaud, L. Humbert, B. Godbout, B. Aubert, W. Skalli, and J. De Guise, "Coupling 2D/3D registration method and statistical model to perform 3D reconstruction from partial x-rays images data," in *IEEE Engineering in Medicine and Biology Society*, 2009, pp. 1008–1011.

[8] S. Benameur, M. Mignotte, S. Parent, H. Labelle, W. Skalli, and J. de Guise, "3D/2D registration and segmentation of scoliotic vertebrae using statistical models," *Computerized Medical Imaging and Graphics*, vol. 27, no. 5, pp. 321–337, 2003.

[9] S. Kadoury, F. Cheriet, and H. Labelle, "Personalized X-Ray 3D Reconstruction of the Scoliotic Spine From Hybrid Statistical and Image-Based Models." *IEEE Trans. Med. Imaging*, vol. 28, no. 9, pp. 1422–1435, 2009.

[10] F. Kahl and R. Hartley, "Multiple-view geometry under the L_∞-norm," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 9, pp. 1603–1617, 2008.

[11] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization methods and software*, vol. 11, no. 1, pp. 625–653, 1999.