3D Spine Reconstruction of Postoperative Patients from Multi-level Manifold Ensembles

Samuel Kadoury[1,2], Hubert Labelle[2], and Stefan Parent[2]

[1] MEDICAL, Polytechnique Montreal, Montreal, QC, Canada
samuel.kadoury@polymtl.ca
[2] Sainte-Justine Hospital Research Center, Montreal, QC, Canada

Abstract. The quantitative assessment of surgical outcomes using personalized anatomical models is an essential task for the treatment of spinal deformities such as adolescent idiopathic scoliosis. However an accurate 3D reconstruction of the spine from postoperative X-ray images remains challenging due to presence of instrumentation (metallic rods and screws) occluding vertebrae on the spine. In this paper, we formulate the reconstruction problem as an optimization over a manifold of articulated spine shapes learned from pathological training data. The manifold itself is represented using a novel data structure, a multi-level manifold ensemble, which contains links between nodes in a single hierarchical structure, as well as links between different hierarchies, representing overlapping partitions. We show that this data structure allows both efficient localization and navigation on the manifold, for on-the-fly building of local nonlinear models (manifold charting). Our reconstruction framework was tested on pre- and postoperative X-ray datasets from patients who underwent spinal surgery. Compared to manual ground-truth, our method achieves a 3D reconstruction accuracy of 2.37 ± 0.85mm for postoperative spine models and can deal with severe cases of scoliosis.

1 Introduction

Spinal deformity pathologies such as adolescent idiopathic scoliosis are three-dimensional (3D) deformations of the trunk, described as a lateral deviation of the spine combined with asymmetric deformation of the vertebrae. Modalities such as MRI or CT are limited for pre- and postoperative assessment since they require the patient to be lying in a prone position during acquisition. For this reason, biplanar radiography is still the imaging technique which is most frequently used for the 3D clinical assessment of spinal deformities, as it allows acquiring the image in a natural standing posture with very little radiation.

Several attempts were made to reconstruct the spine from biplanar X-rays. *A priori* knowledge of the vertebral shapes was used with morphologic descriptors to estimate the geometrical spine model and refined with projected silhouettes [1]. This approach was improved by adding inference-based adjustments to obtain an accurate estimate of the vertebra's orientation and 3D location [2]. Humbert et al. [3] proposed to evaluate a parametric model based on the spinal centerline,

P. Golland et al. (Eds.): MICCAI 2014, Part III, LNCS 8675, pp. 361–368, 2014.

Moura et al. [4] inferred an articulated model of the spine based on splines, and Boisvert et al. [5] formulated the estimation of the spine shape as a second-order cone program. Kadoury et al. [6] introduced a statistical and image-based approach to reconstruct the spine using geometrical models. However, all of these methods were designed for the reconstruction from preoperative X-rays.

Surgical treatment involves correcting the scoliotic curves with pre-shaped metal rods anchored on the vertebrae with screws. Few methods have focused on the 3D reconstruction from instrumented spines, even though postoperative 3D analysis is crucial to assess a treatment's efficacy. In [7], a multilevel statistical model was proposed to reconstruct the spine in postoperative patients. Still, the method remains highly supervised and assumes a linear statistical distribution of the underlying pathological variations. In contrast, manifold learning is based on the premise that data are often of artificially high dimension and can be embedded in a lower dimensional space. However most global approaches (LLE, ISOMAP) fail to adequately model closed distributions by unwrapping the manifold, thus changing the intrinsic topology. On the other hand, graph-search approaches preserve the manifold, avoiding losing continuity during the embedding [8]. This could add robustness to the inference process of unseen cases from a low-dimensional manifold embedded in the ambient space.

We propose a novel biplanar 3D reconstruction method of the instrumented spine, where the shape reconstruction task is formulated as optimizing an energy function over the manifold of uninstrumented spine shapes. The energy function is designed such that it is robust to instrumentation, leading to solutions which fit inside the domain of valid anatomical configurations. The manifold is learned from a training set of visual hull reconstructions from spine X-ray images. To this manifold we attach a vector field of generating parameters for articulated poses and shapes. Our contributions are two-fold. First, we propose an approach to build local charts whenever it is required on the manifold, approximating the tangent space around a point and maximizing the accuracy of the nonlinear approximation. Our approach combines the construction of a neighbourhood graph and learning the manifold onto this graph, to effectively infer new models. Second, we propose a spine reconstruction method from X-ray images, where the anatomy of the spine is occluded by instrumentation (screws and rods).

2 Method

Given a set of calibrated biplanar X-rays, a visual hull reconstruction of the global spine shape is obtained from the silhouettes of the anterior portion of the spine extracted from both images (frontal and lateral) using a Hessian filtering technique [9]. Our aim is to create a multi-level manifold ensemble from a training set of visual hulls (Fig. 1), so to efficiently navigate towards the best match during the spine reconstruction process for an unseen set of X-rays.

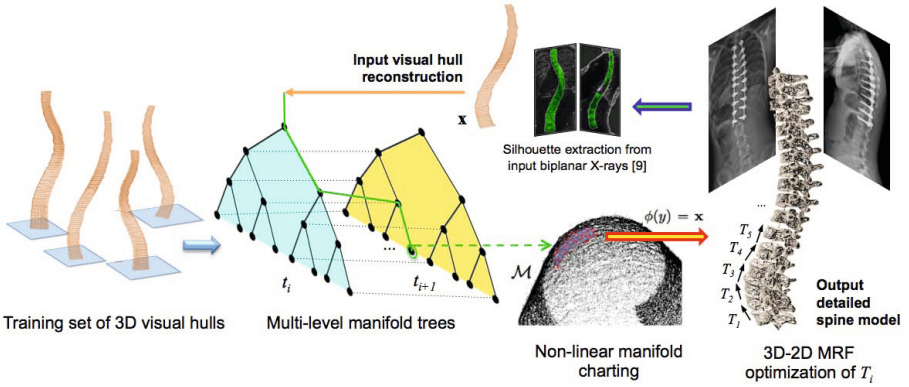


Fig. 1. Illustration of the multi-level manifold ensembles for the 3D reconstruction of postoperative spines. The data structure allows both vertical and horizontal displacements between tree nodes given an input visual hull reconstruction of the spine.

2.1 Building the Multi-level Manifold Ensemble

The first step is to learn an ensemble of randomized space partioning trees that are connected between each other. Essentially, the trees can be viewed as defining an adaptive grid on the ambient space, in a manner similar to k-d trees. The ensemble is a set, $\mathcal{T}$, of binary trees $t_i \in \mathcal{T}$ which hierarchically partitions the ambient data space $\mathbb{R}^D$. We train each tree with the same dataset $\mathcal{X} = \{\mathbf{x}_i\}$, $\mathbf{x}_i \in \mathbb{R}^D$ of global spine visual hulls of dimension D. In our application, we assume the samples $\mathbf{x}_i$ to lie on a d-dimensional manifold $\mathcal{M}$ embedded in $\mathbb{R}^D$ with $d < D$. The parameters $\Theta_j = (\theta_j, \tau_j)$ for each node j define a separating hyperplane in the ambient space $\mathbb{R}^D$ by its unit normal vector $\theta_j \in \mathbb{R}^D$ and a threshold $\tau_j \in \mathbb{R}$, set as the average of the two projected samples with the largest distance on the manifold. The data assigned to each node, $\mathcal{X}_j$, is partitioned into two subsets: $\mathcal{X}_j^L$ and $\mathcal{X}_j^R$, depending on the value of the split function $h(\mathbf{x}, \Theta_j) \in \{0, 1\}$. The split functions take the form:

$$h(\mathbf{x}, \Theta_j) = \mathrm{I}(\mathbf{x}^T \theta_j > \tau_j) \tag{1}$$

where I is the indicator function that penalizes points further away from the manifolds. The set $\mathcal{X}_j^L$ contains samples $\mathbf{x} \in \mathcal{X}$ for which $h(\mathbf{x}, \Theta_j) = 0$, the set $\mathcal{X}_j^R$ contains those for which $h(\mathbf{x}, \Theta_j) = 1$. To find Θ_j, we sample a random subset, $\mathcal{D}_j \subset \mathcal{X}_j$, sample a point $\mathbf{x}_k \in \mathcal{D}_j$ and find the most distant point in $\mathcal{D}_j$:

$$\mathbf{x}_l = \underset{\mathbf{x} \in \mathcal{D}_j}{\operatorname{argmax}} \|\mathbf{x}_l - \mathbf{x}_k\|. \tag{2}$$

The normal θ_j to our hyperplane is the unit length vector between these two points: $\theta_j = (\mathbf{x}_l - \mathbf{x}_k)/(\|\mathbf{x}_l - \mathbf{x}_k\|)$.

The second step learns the graphs, composed of a set of tree nodes, $\mathcal{V}$, and a set of directed edges, $\mathcal{E}$. Here, we denote nodes as $v_i^t \in \mathcal{V}$ with t the tree, and

i the node index respectively, while an edge set is composed of all parent-child tree edges $\mathcal{E}_t$ along with edges between trees $\mathcal{E}_{s,t}$. A set of edges is defined as:

$$\mathcal{E} = (\bigcup_{t \in \mathcal{T}} \mathcal{E}_t) \cup (\bigcup_{(s,t) \in \mathcal{T}} \mathcal{E}_{s,t}). \tag{3}$$

The idea here is that two nodes v_i^s and v_j^t in trees s and t are connected if the regions they define intersect. Exact computation of these intersections is too expensive in high dimensions, even in the case of linear splits. Instead, we use the data samples to estimate intersections, and connect nodes v_i^s and v_j^t if the intersection of their sample sets $\mathcal{D}_i^s$ and $\mathcal{D}_j^s$ is non-empty. The weight $w_{(v_i^s, v_j^t)}$ of each directed edge is set proportional to $\mathcal{D}_i^s \cap \mathcal{D}_j^s$.

2.2 Optimization on the Manifold

Once the multi-level manifolds are trained, our aim, given a test case, is to optimize an energy function f defined on points that lie on a manifold $\mathcal{M}$. We first find an initial solution by traversing the ensemble, both horizontally and vertically (within and between trees), using an optimized k-d search with aligned trees. Upon reaching a leaf node, we query the local neighborhood with the horizontal connections and build a nonlinear chart on $\mathcal{M}$.

Generating an Initial Solution. In order to minimize an energy function f with the multi-level manifold structure, the first step consists of generating a set of initial points on the manifold. These potential nodes will provide a pool of candidates which will be close to the global solution. Our aim is not to obtain a single point, but rather a group of points on the manifold used later to chart a non-linear map. To overcome the overall complexity of the multi-level trees and facilitate searches to find the nearest neighbour to a query, we employ a strategy where prior to searching the trees, binary trees $t_i \in \mathcal{T}$ are aligned by their principal components via a PCA approach. Data is therefore split up in the tree by hyperplanes perpendicular to the principal axes. To achieve this, the datasets are first translated so that their centroids coincide at the origin. Then, we construct the matrix $U = \sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^{\mathrm{T}}$, with N the size of t_i. The eigenvectors of U are the principal axes of the dataset, and the eigenvalues are referred to as the principal moments which are used to rotate the trees $t_i \in \mathcal{T}$.

Given a query vector, a descent down the trees leads to a single leaf node. The data point associated with the node v_j^t is the first candidate for the nearest neighbour. However this point will not necessarily be the nearest neighbour to the query vector; it must be followed by a process of horizontal and vertical moves, in which other cells are searched for better candidates. The recommended method is priority search [10] in which the cells are searched in the order of their distance from the query point. This may be accomplished efficiently using a priority tree for ordering the cells, in which the cells are numbered in the order of their distance from the query vector. The search terminates when there are no more cells within the distance defined by the best point found so far.

At the end of the process, the minimizing cost function yields a set of leaf averages $\bar{\mathbf{x}}_j^t$ which are used as points to create a non-linear chart on the manifold $\mathcal{M}$. Hence, this feature represents a clear advantage to methods using graph-building approaches, since charts are computed on the fly over the training data.

Nonlinear Charting. All the leaf averages $\bar{\mathbf{x}}_j^t$ define a neighbourhood in the local nonlinear charting step of the manifold $\mathcal{M}$, finding a mapping function around a query. The seeds are expanded to their neighborhood using a random walk of horizontal moves. The walk carries on until a given number of nodes W has been reached. The parameter W controls the local chart size and must therefore be chosen carefully, depending on a given optimization problem. We propose to use the nodes reached by the walk as samples for training $\phi(y)$, which provides the transformation from a nonlinear chart near $\mathbf{x}$ on the manifold $\mathcal{M}$ at $y \in \mathbb{R}^d$ to the ambient space, such that $\phi(y) = \mathbf{x}$ where y contains the non-linear dimensionally components. We use the basis of the tangent plane to estimate $\phi(y)$ in d-dimensional space, where in differential geometry, the tangent space is used to compute a local chart. The mapping of $\phi(y) = \mathbf{x}$ estimates the relationship between the D-space and manifold $\mathcal{M}$ as a joint distribution. The non-linear chart should follow a conditional expectation which captures the trend within a local neighborhood of the manifold. Gaussian kernels G estimate densities in the conditional expectation setting [11]:

$$\phi(y_i) = \underset{\mathbf{x}_i}{\operatorname{argmin}} \frac{\sum_{j \in \mathcal{N}(i)} G(y_i, y_j) \|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sum_{j \in \mathcal{N}(i)} G(y_i, y_j)} \tag{4}$$

which integrates the distance $\|\mathbf{x}_i - \mathbf{x}_j\|^2$, that: 1) acts as the similarity metric, 2) avoids giving more weight to the neighbors that are further away within a the neighbourhood $\mathcal{N}(i)$ and, 3) updates $\phi(y_i)$ using the closest neighbors of point y_i in the manifold space. This constrains the regression to be valid for all the neighbourhood of the manifold chart around y_i as it preserves locality in $\mathbf{x}_i$. The chart provides a non-linear parameterization of the space, where outside this range, a new chart is recomputed around the new initial solution.

2.3 3D Reconstruction of Postoperative Patients

We formulate the 3D reconstruction process of postoperative spines, where the anatomy is partly obscured with rods and screws on the X-ray images, by an optimization over the manifold of visual hulls from uninstrumented spines. The energy term we seek to optimize is asymmetric, meaning that the solution that is obtained from the manifold must lie inside the visual hull of the instrumented input. Our model is learned from visual hull shapes (vectorized as $\mathbf{x}$) which are smoothed and subsampled. The function f used to navigate throughout the manifold and find the closest points to the query $\mathbf{x}$ is formulated as:

$$f(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{|\mathbf{x}|} k \circ |(x_i - y_i)(1 + \mathrm{I}(x_i < y_i)\beta)| \tag{5}$$

which defines the dissimilarity between input visual hull $\mathbf{x}$ and points $\mathbf{y}$ on the manifold. Here, $k(\cdot)$ is a kernel function reducing the influence of outliers, and $\mathrm{I}(\cdot)$ is the indicator function. The function f induces a penalty of β for points falling outside the input reconstruction $\mathbf{x}$ i.e. the uninstrumented visual hull must lie within the input shape. We compute initial solutions on the manifold, followed by nonlinear charting using Eq. (4) over a neighbourhood from which a geometrical spine model is built.

Once the closest node on the manifold is identified, we revert to the ambient space for the final refinements brought to the detailed spine model, by adjusting the parametric spine model to its projection on the biplanar X-rays. We build the parametric body model from the vector field of generating parameters attached to the manifold neighborhood found previously. This models the spine as an array of local intervertebral rigid transformations $A = [T_1, T_2, \ldots, T_m]$, with m the number of intervertebral transformations and $T_i = \{R, t\}$ a rigid transform. We propose to use a Markov Random Field (MRF) to refine A, which was obtained from the manifold. In our formulation, the energy includes data term $\mathbf{g} = \{g_i(\cdot)\}$ (unary potentials) associated to each vertex of the MRF and a regularization term $\mathbf{f} = \{f_{ij}(\cdot,\cdot)\}$ (pairwise potentials) associated to the edges. The first encodes the geodesic active contours of the projected 3D vertebral meshes with the X-ray gradient images, whereas the later ones act as a regularizer by restraining large transformation differences between two vertebral levels (e.g. T_4 and T_5), introducing coherency in the minimization of the discrete MRF:

$$MRF(\mathbf{g}, \mathbf{f}) = \min \sum_{i=1}^{m} g_i(\mathbf{u}_i) + \sum_{i=1}^{m} \sum_{j \in \mathcal{N}(i)} f_{ij}(\mathbf{u}_i, \mathbf{u}_j). \tag{6}$$

Here, $\mathbf{u}_i, \mathbf{u}_j \in L$ are the labels assigned to transformations $T_i, T_j \in A$ respectively, with transformation T_i being moved by assigning them different labels $\mathbf{u}_i \in L$ (L is the label space) until an optimal configuration is found. To reach such an optimal configuration, we define an energy term that will be minimized using an optimization algorithm. We adopt FastPD (Primal-Dual) to optimize the aforementioned non-submodular MRF.

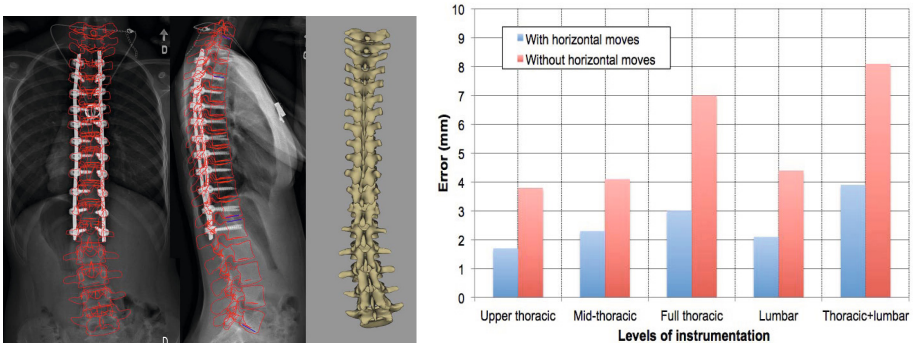
3 Experiments

We tested the multi-level manifold structure for the 3D reconstruction of spine models with instrumentation, by handling occlusions from X-rays. The manifold was built from 843 scoliotic spines demonstrating several types of deformities. For each spine in the dataset, both a visual hull reconstruction and a high-resolution geometrical model was obtained from calibrated biplanar X-rays. The models include 12 thoracic and 5 lumbar vertebrae, each represented by 6 landmarks (4 pedicle extremities and 2 endplate center points), yielding a total 102 landmarks per model. An atlas of 17 generic prior vertebra mesh models obtained from serial CT-scans were fitted to the landmark-based models using FFD.

We first evaluate the performance on preoperative patients with scoliosis. Ten pairs of preoperative biplanar X-rays taken from patients with mild deformities

Table 1. Reconstruction results of pre- and postoperative spines. Results of the proposed multi-level manifold (MLM) are compared to standard PCA and a recently proposed approach using locally linear embeddings (LLE) [6].

	Preoperative X-rays ($n = 10$)						Postoperative X-rays ($n = 20$)					
dim.	2D RMS difference			3D RMS difference			2D RMS difference			3D RMS difference		
	PCA	LLE	**MLM**	PCA	LLE	**MLM**	PCA	LLE	**MLM**	PCA	LLE	**MLM**
1	4.77	2.72	**2.05**	5.04	2.97	**2.39**	6.37	3.44	**2.38**	6.72	3.86	**2.59**
5	4.18	2.53	**1.98**	4.52	2.85	**2.22**	5.74	3.31	**2.21**	6.16	3.80	**2.46**
10	3.71	2.07	**1.90**	4.10	2.33	**2.09**	5.53	3.18	**1.96**	5.95	3.68	**2.34**
15	3.64	2.02	**1.93**	4.04	2.28	**2.15**	5.52	3.14	**2.04**	5.87	3.63	**2.39**
20	3.59	2.10	**1.95**	4.01	2.39	**2.20**	5.47	3.22	**2.07**	5.85	3.72	**2.41**



(a) (b)

Fig. 2. (a) Sample 3D reconstruction result from postoperative frontal and lateral X-rays, yielding high-resolution geometrical models. (b) Errors with 5 different instrumentation configurations, comparing optimization with and without horizontal moves.

(Cobb angle 15 - 30°) were used to quantify the 3D accuracy. This enables to assess the errors under normal settings, without any occlusion from instrumentation. For each case, differences between the proposed method and landmarks identified manually by a radiology expert were computed. Table 1 presents these results, demonstrating performances similar to state-of-the-art [6].

We then evaluated the method on twenty ($n = 20$) pairs of instrumented biplanar X-rays, where 3D spine reconstructions were generated by learning a manifold of visual hulls rendered from uninstrumented spines. Sample results are shown in Fig. 2(a). In Table 1, we compare results to both a PCA and a single spine manifold approach [6], prior to MRF optimization. Different values of d were tested in order to evaluate the effect of the manifold dimensionality. With $d = 10$, the mean difference in 3D landmark locations for all 20 patients was of 2.57 ± 0.73mm for thoracic vertebrae and 2.12 ± 0.67mm for lumbar vertebrae. To evaluate robustness with respect to different instrumentation strategies, we measured the reconstruction accuracy over five different types of surgical paradigms, based on instrumented levels (thoracic and lumbar). Fig. 2(b) shows these results, which also show the benefit of integrating horizontal moves in the ensemble search when different anatomical regions are occluded by metal rods.

4 Conclusion

In this paper, we proposed an unsupervised method to perform the 3D reconstruction of a spine geometry when surgical instrumentation is visible on biplanar X-rays. Our approach is based on multi-level manifold ensembles, which enable an efficient navigation on a low-dimensional domain to infer the closest match to a training set. Results show that this model allows reconstruction accuracies similar to gold-standard. Because the 3D reconstructions are obtained without any user supervision, the proposed approach could be transposed to clinical practice and used in the context of multi-centre evaluations of surgical practices.

References

1. Pomero, V., Mitton, D., Laporte, S., de Guise, J.A., Skalli, W.: Fast accurate stereoradiographic 3D-reconstruction of the spine using a combined geometric and statistic model. Clinical Biomech. 19, 240–247 (2004)
2. Dumas, R., Blanchard, B., Carlier, R., de Loubresse, C.G., Huec, J.C.L., Marty, C., Moinard, M., Vital, J.M.: A semi-automated method using interpolation and optimisation for the 3D reconstruction of the spine from bi-planar radiography: a precision and accuracy study. Med. Biol. Eng. Comput. 46, 85–92 (2008)
3. Humbert, L., de Guise, J., Aubert, B., Godbout, B., Skalli, W.: 3D reconstruction of the spine from biplanar X-rays using parametric models based on transversal and longitudinal inferences. Med. Eng. Phy. 31(6), 681–687 (2009)
4. Moura, D., Boisvert, J., Barbosa, J., Labelle, H., Tavares, J.: Fast 3D reconstruction of the spine from biplanar radiographs using a deformable articulated model. Med. Eng. Phy. 33, 924–933 (2011)
5. Boisvert, J., Moura, D.: Interactive 3D reconstruction of the spine from radiographs using a statistical shape model and second-order cone programming. In: 33rd Annual International Conference of the IEEE EMBS, pp. 5726–5729 (2011)
6. Kadoury, S., Cheriet, F., Labelle, H.: Personalized X-ray 3D reconstruction of the scoliotic spine from statistical and image models. IEEE Trans. Med. Imag. 28, 1422–1435 (2009)
7. Lecron, F., Boisvert, J., Mahmoudi, S., Labelle, H., Benjelloun, M.: Fast 3D spine reconstruction of postoperative patients using a multilevel statistical model. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part II. LNCS, vol. 7511, pp. 446–453. Springer, Heidelberg (2012)
8. Pitelis, N., Russell, C., Agapito, L.: Learning a manifold as an atlas. In: CVPR, pp. 1–8 (2013)
9. Kadoury, S., Cheriet, F., Labelle, H.: Segmentation of scoliotic spine silhouettes from enhanced biplanar X-rays using a prior knowledge Bayesian framework. In: Proc. ISBI, pp. 478–481 (2009)
10. Silpa-Anan, C., Hartley, R.: Optimised KD-trees for fast image descriptor matching. In: CVPR, pp. 1–8 (2008)
11. Davis, B., Fletcher, P., Bullitt, E., Joshi, S.: Population shape regression from random design data. In: ICCV, pp. 1–7 (2007)