# Three-Dimensional Spine Model Reconstruction Using One-Class SVM Regularization

Fabian Lecron, Jonathan Boisvert, Saïd Mahmoudi, Hubert Labelle, Mohammed Benjelloun

***Abstract*—Statistical shape models have become essential for medical image registration or segmentation and are used in many biomedical applications. These models are often based on Gaussian distributions learned from a training set. We propose in this paper a shape model which does not rely on the estimation of a Gaussian distribution, but on similarities computed with a kernel function. Our model takes advantage of the One-Class Support Vector Machine to do so. In this context, we propose in this paper a method for reconstructing the spine of scoliotic patients using OCSVM regularization. Current state-of-art methods use conventional statistical shape models and the reconstruction is commonly processed by minimizing a Mahalanobis distance. Nevertheless, when a shape differs significantly from the statistical model, the associated Mahalanobis distance often overstates the need for statistical regularization. We show that OCSVM regularization is more robust and is less sensitive to weak landmarks definition and is hardly influenced by the presence of outliers in the training data. The proposed OCSVM model applied to 3D spine reconstruction was evaluated on real patient data and results showed that our approach allows precise reconstruction.**

***Index Terms*—3D Reconstruction, One-Class SVM, Spine, Scoliosis.**

## I. Introduction

STATISTICAL shape models have become essential for medical image registration or segmentation. They derive their effectiveness from a training set by providing specific information about the context of the application. More precisely, the shape is constrained so that, based on observations, the model can only synthesize plausible instances of an object.

These models were presented by Cootes *et al.* in [1], in the context of the Active Shape Model segmentation approach. Since then, they have been applied in numerous biomedical applications and we refer the reader to [2] for a detailed review on the subject. In these applications, the main idea is to define a parametric point distribution model using Principal Component Analysis (PCA). Other strategies exist to address the construction of a shape model. In [3], Tsai *et al.* represented a shape with a signed distance function instead of a collection of landmarks. Each shape in the sample is therefore seen as a level set function. PCA applied on a sample of level set functions leads to a statistical shape model. This particular shape representation does not require point correspondences between the shapes of the sample and it better handles topological changes. In [4], El-Baz *et al.* built shape prior by performing PCA on distance vectors described by the square distances from control points on the boundary to the model centroid. Recently, Khalifa *et al.* proposed a segmentation framework using a level set-based deformable model constrained with a probabilistic shape prior [5]. In order to build it, a collection of images are coaligned by rigid transformations maximizing their mutual information. The shape prior is then defined by an independent random field for the training images and specified by object and background probabilities on each pixel.

Applications making use of statistical shape models are usually based on Gaussian distributions learned from a training set. However, it is often difficult to get sufficient number of training cases to well represent the distribution and actual data distributions are often heavy-tailed. This is especially true for biomedical applications since pathological cases are relatively common in the clinics, but can deviate from statistical model in extreme ways (see [6] for example).

In some cases, this discrepency between the assumed distribution and reality could prevent the registration or the segmentation method based on statistical shape models from converging to a solution. An interesting solution to this issue is to consider the training set as a sub-set of the shape space without assuming hypothesis on the statistical distribution of the global shape space. We thus propose in this paper a shape model which does not rely on Gaussian distributions, but on similarities computed with a kernel function.

The originality of our model is to take advantage of One-Class Support Vector Machine (OCSVM) to define a statistical shape model and use it for 3D model reconstruction. Related works have used OCSVM for medical image analysis in the past with different goals. For instance, tumor segmentation from MRI based on OCSVM was proposed in [7]. In this application, OCSVM was used to separate tumor data from non-tumor data. Wang *et al.* described in [8] a framework for 3D reconstruction of head MRI. The tissues are first segmented and extracted from medical images. Then, this information is used as feature data input for OCSVM. An hypershpere enclosing these data is obtained by minimizing its radius. The 3D reconstruction is deducted from the support vectors related to this hypersphere. A deformable model using OCSVM was also proposed by Chang in [9]. A succession of three steps defines the model. First, information about edge orientation is extracted. Then, an algorithm based on the cross-correlation theorem is developed for shape detection. Finally, OCSVM is



F. Lecron, S. Mahmoudi, and M. Benjelloun are with the Computer Science Department, Faculty of Engineering, University of Mons, Belgium (e-mail: Fabian.Lecron@umons.ac.be)

J. Boisvert is with the National Research Council, Information and Communications Technologies, Canada (e-mail:Jonathan.Boisvert@nrc-cnrc.gc.ca)

H. Labelle is with the Sainte-Justine Hospital, Montréal, Canada

only used to detect outliers among the detected shapes.

OCSVM was also used for face detection in [10]. Jin *et al.* described an approach divided into three steps. First, face candidates are located with a skin-color detector. These face candidates are then refined so that good candidates match some shape features based on grayscale rules and gradient rules. Finally, OCSVM is used as a classifier only keeping faces among the remaining candidates. The interest of their statistical model based on OCSVM is to discriminate faces from outliers. Our approach is very different since the OCSVM model proposed in the present work is part of the process of shape extraction and is not a post-processing step as in [10]. In this way, OCSVM model is not a binary process that accepts or rejects a shape but it brings more nuance and guides the deformation of the shape. In other words, the support vectors of the OCSVM are directly used to define a statistical shape model. The idea is to define an hyperplane holding in its positive side a set of shapes. This hyperplane, computed with the help of a kernel function and defined using its support vectors, is actually used to build the model. Therefore, we call it *kernel-based* shape model. The extraction of a shape from a given image is conducted by optimizing the distance between the deformed shape and the hyperplane. This way, there is no need to rely on predefined distribution.

The three-dimensional representation of the spine proved of great interest in biomedical science. Thereby, the 3D analysis of the vertebral column provides relevant information for the diagnosis and the treatment of three-dimensional deformities such as scoliosis. Part of the literature has focused on the reconstruction of the spine from bi-planar radiographs. The advantage of X-ray imaging is to avoid exposure of the patient to high radiations. Moreover, this imaging technique reduces the acquisition time and allows the patient to be in a standing position. We propose in this paper a 3D spine reconstruction method from bi-planar radiographs using OCSVM regularization.

One simple way to reconstruct the spine from radiographs is to manually identify anatomical landmarks on the images and match their 3D positions [11]. These methods require to manually locate multiple points per vertebra and are thus very time-consuming. Mitulescu *et al.* proposed to also consider landmarks that are only visible in one radiograph [12]. In order to reduce the number of landmarks involved, parametric vertebra models were proposed by Pomero *et al.* in [13]. They formulated the reconstruction as statistical inferences based on the relationship between geometric descriptors of a single vertebra. This way, four points per vertebra are needed for the reconstruction. Later, Humbert *et al.* used the parametric model of Pomero *et al.* and combined it with a parametric model of the whole spine to reduce the identification of landmarks in a first level of reconstruction [14]. Nevertheless, considerable user-interaction is required for further fine reconstructions of the spine. Other approaches were designed to use the information contained in the images. In [15], Kadoury *et al.* used a 3D statistical model regulated with 2D image level set functionals to improve the accuracy of an initial reconstruction only based on a statistical model.

Articulated statistical models of the spine were used in [16] by Moura *et al.*. They have represented the position and the orientation of a given vertebra as rigid geometric transformations from the vertebra to the others along the spine. This model is then inferred with a spline representing the spinal centerline. Recently, Boisvert *et al.* proposed to constrain a statistical shape model of the spine [17]. A multilevel statistical shape model of the spine was then proposed in [18], [19]. In those methods involving a statistical model, the reconstruction is processed by minimizing the Mahalanobis distance between a shape and the model distribution. When a shape differs strongly from the statistical model, the associated Mahalanobis distance is large. As a consequence, the optimization hardly converges to an acceptable solution. A typical example is when a patient adopts a particular position which differs from the model, *i.e.* bending position [6].

The main contribution of this paper is to propose a formulation of a kernel-based shape model. This novel framework for modeling shapes has been considered for representing a spine model in the context of the 3D spine reconstruction from bi-planar radiographs. More precisely, we present a model of the spine based on OCSVM. The reconstruction algorithm is then formulated as a program optimizing the correspondence score of a given shape with respect to the hyperplane defined by the OCSVM approach. This score is expressed so that it can deal with uncommon shapes robustly. These theoretical aspects are described in section II. Section III provides results about 3D reconstruction of the spine and shows that our kernel-based shape model provides better results than actual state-of-art approaches. Finally, section IV concludes the paper and gives future prospects.

## II. Methods

### A. One-Class Support Vector Machine

Support Vector Machine (SVM) approach [20] is mainly applied in classification or regression problems. The principle of SVM is to separate a set of data into two distinctive classes. A variant, called One-Class SVM (OCSVM), was proposed in [21] to deal with the detection of outliers. The general idea is define a region of the space using support vectors that includes most of the training points, but not all, and maximizes the classifier's margin. Thus, separating "normal" instances from outliers.

Let us consider a set of data points $x_k \in \mathbb{R}^n$ and a function $\Phi : \mathbb{R}^n \rightarrow \mathcal{F}$ transposing them into a higher-dimensional space, often called *feature space*. Let $w \in \mathcal{F}$ and $b \in \mathbb{R}$ be the couple of variables defining an hyperplane. The purpose of OCSVM is to determine the hyperplane that holds most of the data points in its positive part in such a way that $\langle w, \Phi(x_i)\rangle - b > 0$.

The couple $(w, b)$ can be determined by solving the quadratic programming problem:

$$\begin{cases} \min\limits_{w,b} & \frac{1}{2}\|w\|^2 + \frac{1}{\nu m}\sum_{i=1}^{m}\xi_i - b \\ s.t. & \langle w, \Phi(x_i)\rangle \geq b - \xi_i, \forall i = 1, \dots m \cdot \\ & \xi_i \geq 0, \forall i = 1, \dots m. \end{cases} \quad (1)$$

Two variables are introduced in the equation (1). First, $\xi = (\xi_1, \xi_2, \dots, \xi_m)$ is called slack-variable whose positive

value penalizes the objective function. Second, $\nu$ stands for the fraction of data allowed to be in the negative side of the hyperplane.

The problem (1) is formulated at equation (2) in a dual form, where Lagrangian coefficients $\alpha$ are introduced.

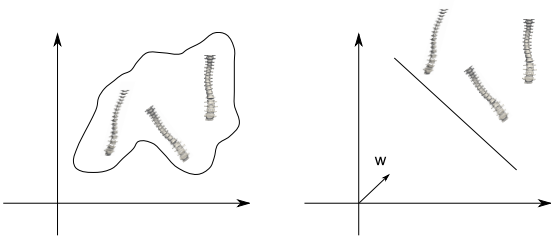
$$\begin{cases} \underset{\alpha}{max} & \widetilde{L}(\alpha) = -\frac{1}{2}\sum_{i,j=1}^{m}\alpha_i\alpha_j\left\langle\Phi(x_i),\Phi(x_j)\right\rangle \\ s.t. & 0 \leq \alpha_i \leq \frac{1}{\nu m}, \forall i = 1,\ldots,m \\ & \sum_{i=1}^{m}\alpha_i = 1. \end{cases} \tag{2}$$

The formulation of the problem (2) shows that the function $\Phi(x)$ is not needed explicitly. Only the inner products $\langle\Phi(x_i),\Phi(x_j)\rangle$ are used for the optimization. Let $K$ be a symmetric matrix consisting of the inner products $\langle\Phi(x_i),\Phi(x_j)\rangle$. It can be shown that any positive semidefinite matrix could be used as a kernel matrix $K$, since the positive semidefiniteness ensures the existence of the function $\Phi(x)$.

Finally, once the parameters of the hyperplane are computed, any new point $x \in \mathbb{R}^n$ can be classified by determining which side of the hyperplane $x$ belongs to.

### B. One-Class SVM Shape Model

The idea is to define an hyperplane holding in its positive side a set of 3D shapes. Fig. 1 illustrates our purpose when the shape represents a spine. On the left, a set of 3D spine models are represented in a given space. A kernel function allows to transpose these 3D models into a higher-dimensional space (on the right). They are separated with the rest of the feature space by an hyperplane.




(a) A set of spines represented in a shape space

(b) Hyperplane defined in the feature space holding the set in its positive side

Fig. 1: Conceptual view of a spine shape model based on OCSVM.

In order to develop a shape reconstruction algorithm, we need to define a score that quantifies the level of correspondence between a shape and the training data. Thus, let us consider a shape characterized by the vector $x_i$. The distance from $x_i$ to the hyperplane is given by:

$$d = \frac{(\langle w,\Phi(x_i)\rangle - b)}{\|w\|}. \tag{3}$$

Unfortunately, $w$ potentially belongs to an infinite-dimensional space, which makes the distance hard to use in practice. However, since $w = \sum_{i=1}^{m}\alpha_i\Phi(x_i)$, we can formulate an unnormalized correspondence score $s$ :

$$s = \sum_{j=1}^{m}\alpha_j\left\langle\Phi(x_j),\Phi(x_i)\right\rangle. \tag{4}$$

The relation (4) now relies only the variables $\alpha_i$, which are obtained by the resolution of the problem (2), and on the kernel $K$ which is used to determine the $m$ inner products $\langle\Phi(x_j),\Phi(x_i)\rangle$. This makes it a good candidate to use in practice.

### C. Spine Shape Representation

In the previous section, we showed that the vector $x$ can stand for a 3D model of the spine. But how to represent the spine in three dimensions? The current section provides two answers to this question.

*1) Landmark Representation:* A common way to represent a 3D shape is to use landmarks. These points of reference can be selected based on geometrical features or in accordance with the application. In our case, an expert selected 6 anatomical landmarks to represent one vertebra. These points are the center of inferior and superior endplates (2 points), and the inferior and superior extremities of pedicles (4 points). If we consider the vertebræ T1 to L5, the shape of a spine is defined by 102 landmarks. Since the shape is in a three-dimensional space, $x$ is a vector of size 306. More formally, the vector $x$ can be written:

$$x = \left(p_1^{abs}, p_2^{abs}, \ldots, p_k^{abs}, \ldots, p_n^{abs}\right), \tag{5}$$

where $n$ is the number of landmarks and $p_i^{abs}$ are the three-dimensional absolute coordinates of the $i^{th}$ landmark.

*2) Articulated Representation:* Using landmarks to define the shape of the spine has the advantage of being simple and it allows the use of conventional statistics. However, the spine has a natural articulated structure. A convenient way to take into account the articulations between the vertebræ was proposed by Boisvert *et al.* in [22]. In their work, the spine is represented by a vector of intervertebral rigid transformations and local anatomical landmarks. These articulated models have shown great interest in several studies (*e.g.* [16], [23]).

The relative rigid transformation $T_i$ related to a given vertebra $i$ is determined by the rigid transformation related to the previous vertebra $i-1$. In order to also take into account the shape of vertebra, anatomical landmarks, represented in the local system of reference of the corresponding vertebra, are used to define the articulated model. As a result, the vector $x$ is written:

$$x = \left(T_1, T_2, \ldots, T_n, p_{1,1}, p_{1,2}, \ldots, p_{i,j}, \ldots, p_{n,m}\right), \tag{6}$$

where $n$ is the number of vertebræ, $m$ is the number of landmarks per vertebra, and $p_{i,j}$ are the three-dimensional coordinates of the $j^{th}$ landmark of the vertebra $i$ expressed in the local vertebra system of reference.

### D. Reconstruction Algorithm

The reconstruction algorithm that we propose in this work aims at matching a 3D shape of the spine with multiple views. In our case, two views are available: a posteroanterior radiograph (PA) and a lateral radiograph (LAT). Two parameters are required to be optimized during the reconstruction process. First, the Euclidean distance between the projection of a 3D point $x_i$ belonging to the spine and its theoretical location on the radiograph has to be minimized. Let us consider that the spine is comprised of $N$ vertebræ, that $M$ is the number of anatomical landmarks per vertebra and that $K$ is the number of considered views. Let us note $p^{2D}_{i,j,k}$, a point located by the user on the radiograph, with $0 < i \leq N$, $0 < j \leq M$ and $0 < k \leq K$. If the projection of a 3D point is represented by $\hat{p}^{2D}_{i,j,k}$, the sum of squared errors ($SSE$) of reprojection is computed according to:

$$SSE = \sum_i \sum_j \sum_k \left\| p^{2D}_{i,j,k} - \hat{p}^{2D}_{i,j,k} \right\|^2, \qquad (7)$$

where $p^{2D}_{i,j,k}$ only concerns points that have been located by the user.

The solution resulting from the minimization of the relation (7) needs to conform with the shape model defined in section II-B. We have seen that the equation (4) represents a correspondance score. The greater the score, the better. As a result, $s$ has to be maximized and the complete optimization problem for the 3D reconstruction is formulated as:

$$\min \sum_i \sum_j \sum_k \left\| p^{2D}_{i,j,k} - \hat{p}^{2D}_{i,j,k} \right\|^2 + \beta \left( \frac{1}{s} \right)^2, \qquad (8)$$

where $\beta$ is a weight controlling the importance of the model during the optimization process. The problem (8) is a non-linear least-squares problem which can be solved by Levenberg-Marquardt algorithm.

In order to reduce the number of variables involved in the problem (8), a principal component analysis is operated on the training set. Any spine shape can be approximated according to $x = \bar{x} + \phi b$, where $\phi$ is the matrix of principal components and $b$ is a vector of principal scores. The goal of the optimization process is then to find the value of $b$ which minimizes the relation (8).

### E. Kernel Choice for Reconstruction

The score $s$ defined at equation (4), which represents how a shape is convenient with the model, relies on a kernel function. The choice of a good kernel is therefore important and depends on the application. In this section, we define two kernels in the context of the 3D reconstruction of the spine. These kernels will be assessed in the experimental section of the paper.

One of the most common kernels used in practice is the Radial Basis Function (RBF) kernel defined by:

$$K(x, y) = \exp\left( -\frac{\|x - y\|}{2\sigma^2} \right). \qquad (9)$$

In addition to this baseline kernel, we propose a specific kernel based on a Mahalanobis distance between two shapes. We define this Mahalanobis kernel according to:

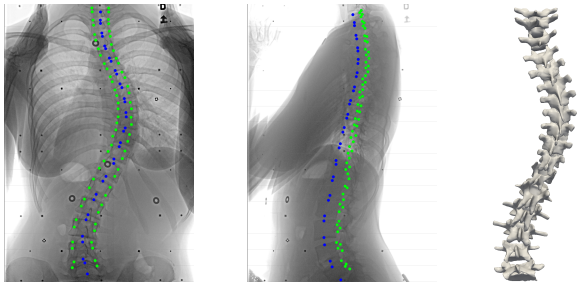
$$K(x, y) = \exp\left( -\frac{(x - y)^T \Sigma^{-1} (x - y)}{\sigma} \right), \qquad (10)$$

where $\Sigma$ is the variance-covariance matrix related to our sample of 3D spine shapes.

## III. Results and Discussion

In this section, we aim at showing the interest of the formulation of our kernel-based shape model. As previously stated, this model is evaluated in the context of a biomedical application: the 3D reconstruction of the spine from bi-planar radiographs.

To validate the reconstruction method, we used a sample of 100 scoliotic patients randomly chosen without any constraint to build the training set and a sample of 25 randomly chosen post-operative patients and 20 randomly chosen patients with severe scoliosis. These data were provided by the Sainte-Justine Hospital of Montréal, Canada. We have considered 17 vertebræ, *i.e.* T1 to L5. Each vertebra is represented by 6 points of reference, *i.e.* the center of inferior and superior endplates, and the inferior and superior extremities of pedicles. All these landmarks have been previously digitized in a three-dimensional space following a reference method [11]. The post-operative cases were used to assess the accuracy of the reconstruction approach. The reason for having chosen post-operative patients is related to the specific deformation of the spine when surgical instrumentation is present to redress the spine. Indeed, the result is often neither a normal spinal curve nor a scoliotic one. Therefore, it is difficult to capture these specific deformations with a classical statistical model.



(a) PA View (b) LAT View (c) 3D Reconstruction

Fig. 2: Anatomical landmarks manually identified on bi-planar radiographs (plates are in blue and pedicles in green).

### A. Kernel Validity

The main parameter influencing the reconstruction is the choice of the kernel. Indeed, we have shown that the measure of $s$ is only possible with a kernel $K(x, y)$. Such a function can be considered as a similarity measure between two objects $x$ and $y$ in the feature space. In our case, the objective is to distinguish 3D spine shapes. As a reminder, two representations of the spine are considered in this study: a landmark

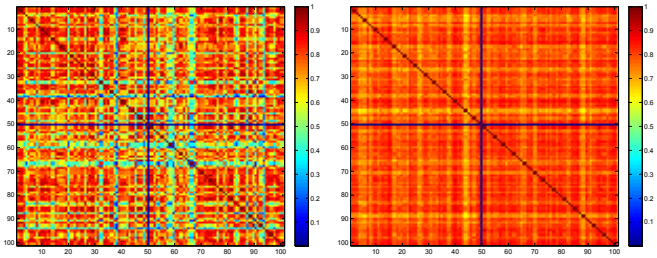
representation and an articulated one. The influence of two kernels is evaluated in this section: an RBF and a Mahalanobis kernel.

In order to demonstrate the validity of these kernels, we have computed the similarities between the 100 items of the sample. To be valid, we expect a kernel to provide high similarities between these 100 spines. Furthermore, we have artificially generated an unrealistic spine by greatly shifting several vertebræ so that the result is not anatomically possible. In this case, we expect that the kernel will clearly distinguish this unrealistic spine from other. In the following, the sample used to compute the covariance matrix $\Sigma$ (for the Mahalanobis kernel) is different from the sample used to evaluate the similarities.

We first propose to discuss the validity of the RBF and the Mahalanobis kernels in the case of a landmark representation of the spine. The similarity between the 100 spines represented by landmarks and the unrealistic spine was computed both for the RBF and the Mahalanobis kernels. The corresponding similarity matrices are presented at Fig. 3. RBF and Mahalanobis kernels depend on a parameter, namely $\sigma$. The greater $\sigma$ is, the higher the similarity between shapes is. Here, the objective is not to find the best value of $\sigma$ but to show that the RBF and the Mahalanobis kernels are valid to represent the similarity between scoliotic spines. In the similarity matrices of Fig. 3, we note that the elements of the diagonal are all equal to 1 since the similarity between identical spines is maximum. Then, the similarities related to the unrealistic spine are represented at the $50^{th}$ row and the $50^{th}$ column. The similarity matrix associated to the RBF kernel shows that the unrealistic spine is clearly seen as an outlier. The same conclusion can be drawn if we observe the similarity matrix associated to the Mahalanobis kernel. However, there is a difference between RBF and Mahalanobis kernels. For the RBF kernel, the range of similarity is larger than for the Mahalanobis kernel. The fact that the Mahalanobis kernel takes into account the correlation between variables seems to provide less disparity in the similarities. Globally, the RBF and Mahalanobis kernels appear to be relevant to represent similarities of 3D spine shapes in the case of a landmark representation of the spine.

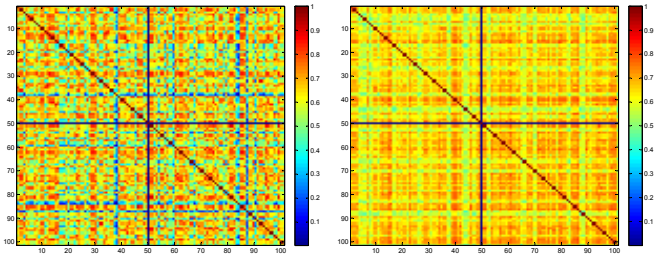
To discuss the validity of these kernels in the case of an articulated representation of the spine, we computed the similarity between the 100 articulated spines of our sample and the unrealistic articulated model. These results are presented at Fig. 4. In this illustration, the unrealistic spine is represented at the $50^{th}$ row and the $50^{th}$ column. As above, the results show that the 100 articulated models of the sample are seen as similar while the unrealistic spine is seen as an outlier. This fact is valid for both the RBF and the Mahalanobis kernels. Once again, the difference between the kernels relies on the range of similarity. This range is larger for the RBF kernel. Finally, Fig. 4 shows that the RBF and Mahalanobis kernels appear to be relevant to represent similarities of 3D spine shapes in the case of an articulated representation of the spine.

Since these kernels can well represent the similarities between the shapes, the parameter $\nu$ of the OCSVM model has a weak influence in our application. As remainder, this parameter represents the fraction of data which are allowed to be in the negative side of the hyperplane. Here, outliers were not present in the sample of 100 spines. In that case, a value of $\nu$ equal to 1 is equivalent to use all the spines to build the hyperplane. In case of presence of outliers, the value of $\nu$ can be adapted.



(a) RBF kernel - $\sigma = 600$ (b) Mahalanobis kernel - $\sigma = 50$

Fig. 3: Similarity matrices - Landmark representation.



(a) RBF kernel - $\sigma = 60$ (b) Mahalanobis kernel - $\sigma = 35$

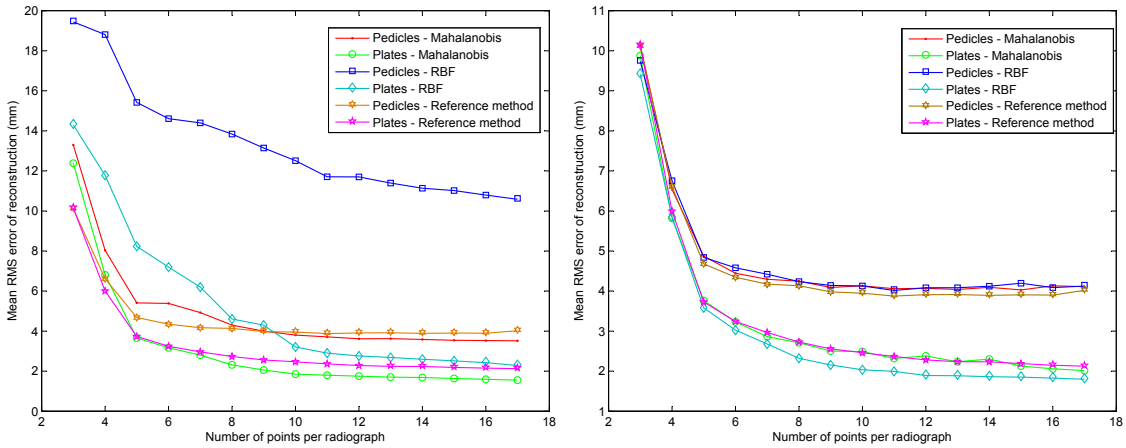
Fig. 4: Similarity matrices - Articulated representation.

### B. Reconstruction Experiments

In order to process a reconstruction, two radiographs are presented to the user. Some points of reference are then pointed out by the operator so that the optimization leads to the final solution, *i.e.* the 3D spine shape of the patient. Mean RMS error has been evaluated as a function of the number of points per radiograph for the 25 post-operative patients. We distinguish the reconstruction of the plates and the pedicles. In the following, the parameter $\nu$ has been fixed to $0.8$. Since this parameter has an important role, we discuss its influence and its interest at the end of section III-E.

*1) Landmark Representation:* We propose at Fig. 5a the mean RMS reconstruction error related to a representation of the spine with landmarks. The RBF and Mahalanobis kernels have been distinguished. The results related to a state-of-art method [17] using a classical statistical shape model is also present at Fig. 5a. For all the methods, we observe that the reconstruction of the plates is better than for the pedicles. Logically, we notice that the error decreases as the number of points per radiograph increases.

An important difference between the RBF and Mahalanobis kernels appears at Fig. 5a. If we consider 17 control points, the mean error of reconstruction related to the RBF kernel is equal to $10.59mm$ and to $2.28mm$ for, respectively, the pedicles and the plates. In comparison, the mean error of




(a) Landmark representation (b) Articulated representation

Fig. 5: Evolution of the mean RMS 3D reconstruction error as a function of number of control points per radiograph.

reconstruction related to the Mahalanobis kernel is equal to $3.51mm$ and to $1.55mm$ for, respectively, the pedicles and the plates. These results show that the RBF kernel is not working in the context of a representation of the spine with three-dimensional landmarks. It seems that the vertebræ are reconstructed independently when the RBF kernel is used. The link between the vertebræ is not taken into account in the model. Conversely, the Mahalanobis kernel allows to compute the similarities between spine shapes by using the correlation between the variables (which are the coordinates of the anatomical landmarks).

Finally, we observe that both for the plates and the pedicles, the Mahalanobis kernel offers globally better results than the classical statistical shape model when a reasonable number of control points are provided, showing the interest of using a kernel-based shape model.

*2) Articulated Representation:* The mean RMS reconstruction error related to an articulated representation of the spine is illustrated at Fig. 5b. The RBF and Mahalanobis kernels have been distinguished, as well as a state-of-art method using a classical statistical shape model. The same preliminary trends than for the landmark representation can be drawn here. For both of the kernels, we observe that the reconstruction of the plates is better than for the pedicles. We also notice that the error decreases as the number of points per radiograph increases.

Nevertheless, the difference between the RBF and Mahalanobis kernels is less contrasted. If we observe Fig. 5b, curves related to RBF kernel are nearly identical to curves related to Mahalanobis kernel. Let us consider 17 control points per radiograph. The mean RMS error of reconstruction related to the RBF kernel is equal to $4.12mm$ and to $1.80mm$ for, respectively, the pedicles and the plates. The same values for the Mahalanobis kernel are, respectively, $4.11mm$ and $2.01mm$. Previously, the RBF kernel provided poor results because the link between the vertebræ were not processed adequately. In an articulated representation of the spine, this link is always integrated in the model, whatever the kernel. This fact explains the differences between Fig. 5a and Fig. 5b.

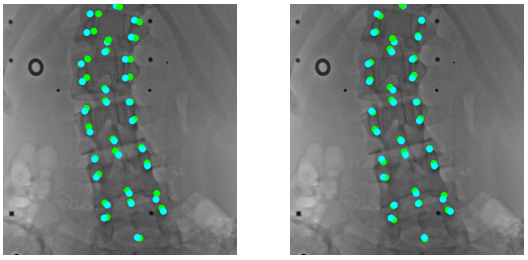
If we compare RBF and Mahalanobis kernels with a classical statistical shape model, we observe that for the reconstruction of pedicles, the results are quite similar. However, the RBF kernel outperforms the Mahalanobis kernel and the statistical shape model in the case of plates reconstruction.

*3) Comparison Between the Reference Method and the OCSVM Method:* If we consider 17 control points at Fig. 5a, the OCSVM method with Mahalanobis kernel improves the precision over the reference method by $0.55mm$ for the plates and $0.51mm$ for the pedicles. To illustrate the effect of this improvement, we propose at Fig. 6 a visual comparison between the OCSVM method (landmark representation with Mahalanobis kernel) and the reference method on lumbar vertebræ L1 to L5 in the PA view. The points in cyan are the anatomical landmarks identified by an expert while the points in green are the projections of the 3D reconstructed anatomical landmarks. The main difference between the two figures is concerned by the pedicles of the vertebræ L4 and L5. For these vertebræ, the precision is obviously better for the OCSVM method. The corresponding 3D anatomical landmarks reconstructed with the OCSVM method are obtained with a precision greater than $1mm$. A paired t-test was also performed to see if the improvements for the plates and the pedicles are statistically significant. Both for the plates and the pedicles, the t-test allows to reject at 0.01 significance level the fact that the distributions of the RMS errors are identical (p-values are lower than 0.01).

### C. Execution Time

Our reconstruction approach is assessed here with respect to the execution time. We report at Table I average computation times for the two different representations of the spine. The RBF and the Mahalanobis kernels are also distinguished. The tests were performed on an Intel Core 2 Duo 2.53 GHz.

The order of magnitude of our reconstruction procedure is between $2.91s$ (articulated representation and RBF kernel) and $7.95s$ (landmark representation and Mahalanobis kernel). The Mahalanobis kernel is always characterized by greater



(a) Reference Method (b) OCSVM Method

Fig. 6: Qualitative comparison between the reference method and the OCSVM method. 3D reconstructed models are projected on the PA view. Manually identified landmarks are in cyan and the landmarks obtained from the reconstruction method are in green.

computation times. This result is logical due to the presence of the covariance matrix in the mathematical formulation of the kernel. Nevertheless, the execution times associated to our approach are in the same order of magnitude, *i.e.* a few seconds.

| Method | Spine Repres. | Kernel | Execution Time (s) |
|---|---|---|---|
| OCSVM | Landmarks | RBF | 5.01 |
| | | Mahalanobis | 7.95 |
| | Articulated | RBF | 2.91 |
| | | Mahalanobis | 7.66 |
| Ref. Meth. | Landmarks | / | 0.15 |

TABLE I: Execution time for the proposed approach.

### D. Sensitivity

The reconstruction process has to be initiated by pointing out some landmarks on the radiographs. This operation is operator-dependent and prone to errors. As a consequence, it is interesting to analyze how does the OCSVM model react when the control points are not well located. To do so, a noise was randomly generated following a uniform distribution and added to the coordinates of 17 control points marked by an operator. The evolution of the mean RMS error of reconstruction with the standard deviation of the applied noise is provided at Fig. 8. A distinction is made between the plates and the pedicles and we have also evaluated the evolution of the error for a classical statistical shape model with the method described in [17]. One can observe at Fig. 8 that the two errors related to the statistical model (one curve for the plates, one for the pedicles) increase dramatically. It means that the precision is highly penalized by errors on the location of control points. On the contrary, the OCSVM models are very robust and the error evolves slowly with the standard deviation of the noise.

In order to visualize the possible variations of landmarks location, we propose at Fig. 7 the representation of two landmarks related to the plates (in green). The red points are obtained by adding an error of 10 pixels to the x- and y-coordinates of the landmarks.

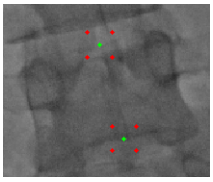


Fig. 7: Limits of variation (red points) of the location of two landmarks (green points).

The effect of location errors is depicted at Fig. 10a (reference method) and Fig. 10b (OCSVM method with landmark representation and Mahalanobis kernel). The PA view was chosen for better visualization of the landmarks. The robustness of OCSVM model is obvious on these illustrations. According to [24], the maximum error of landmarks location by an expert is $3mm$, representing 7.5 pixels on our images. A noise following a uniform distribution from -7.5 to +7.5 pixels was chosen for this illustration.

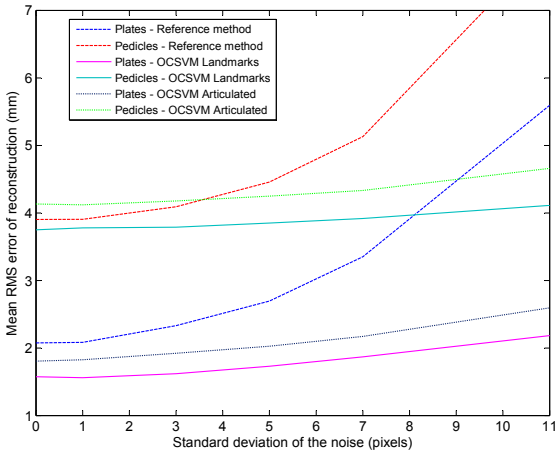



Fig. 8: Effect of the simulated noise on the reconstruction of the plates and the pedicles. Comparison with a classical statistical shape model.

### E. Discussion

In section III-B, we showed that our reconstruction method provides better or similar results than the actual state-of-art method (using a statistical shape model) in the case of post-operative patients. The execution times presented in section III-C were also obtained from these postoperative patients. In order to discuss the results and to compare our approach with more methods of the literature, we have performed a 3D reconstruction for 20 patients with a severe scoliosis and no surgical instrumentation. The severity of the scoliosis is measured by the Cobb angle. In our case, the Cobb angle of the 20 patients ranges from $44°$ to $70°$. In the literature, three recent methods proposed results computed on about 20 patients with moderate or severe scoliosis and no instrumentation. Since experiments have been performed on different data sets, direct comparison is not possible. However, it seems interesting to compare orders of magnitude between the methods. In this context, approaches proposed by Moura *et al.* [16], Kadoury

*et al.* [15] and Humbert *et al.* [14] are characterized by RMS errors between 2 and $4mm$. Our OCSVM model provides RMS errors between 1.57 and $4.13mm$ and appears to be competitive with these approaches. The order of magnitude for the computing time of the fastest approaches [14], [16] is a few seconds, just as our approach based on a OCSVM model.

In the present work, two representations of the spine are proposed: a landmark representation and an articulated representation. This could be seen as a limitation in some biomedical applications where it is difficult to mark the shape of an organ with reference points. Nevertheless, our OCSVM model only requires to represent a shape with a vector. This vector could be a vector of distance maps such as described in [4].

The last aspect of this section is to discuss the interest of the parameter $\nu$ influencing the OCSVM model. As a reminder, $\nu$ is an image of the fraction of data allowed to be in the negative side of the hyperplane. Therefore, the OCSVM model accepts that outliers can be present in the training data. This characteristic is very relevant in our case since in biomedical engineering, the presence of outliers in the data is not rare.

Human experts notice and correct most mistakes that arise during data collection and processing, but some mistakes invariably go undetected and can lead to outliers in the data used to produced statistical models. In our application, outliers can, for instance, be the result of mislabeling of points or vertebrae leading to 3D models that pass visual inspections but which would falsely increase the observed variability in a conventional statistical model. They can also be the result of completely or partially occluded features in one of the radiographs. Furthermore, the shear quantity of data and repetitiveness of the manipulations may cause human observers to commit mistakes they would not make on a single isolated case.

Nevertheless, classical statistical shape models are not able to detect the presence of outliers. Indeed, we propose at Fig. 9 the evolution of the mean RMS error of reconstruction when $5\%$ of outliers are present in the training data.

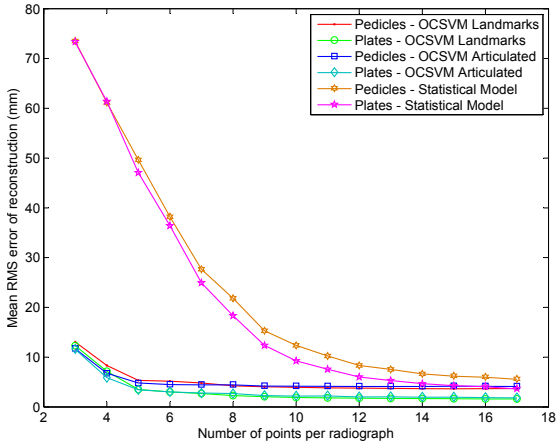
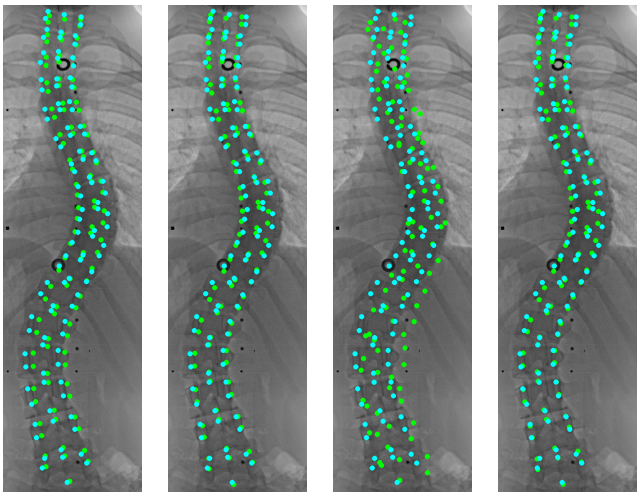



Fig. 9: Evolution of the mean RMS 3D reconstruction error as a function of number of control points per radiograph in presence of outliers in the training data.

These outliers have been generated artificially and replace good items in the data. In reality, extreme outliers can be detected by a decent human observer. However, this detection can be very time-consuming given the size of a sample. The idea is here to propose an automatic method detecting these outliers. The resulting errors are averaged for the 45 patients (severe cases and post-operative patients). The OCSVM model is evaluated with regard to the two representations of the data: the landmark representation (with Mahalanobis kernel) and the articulated representation (with RBF kernel). The parameter $\nu$ of the OCSVM model has been fixed to $0.8$ so that it can accept no more than $20\%$ of outliers. The results are quite instructive. We observe that classical statistical shape models are not able to deal with the presence of outliers. The model badly explains the variables and the precision is greatly reduced. On the contrary, the OCSVM model accepts that no more than $20\%$ of items in the sample belong to the negative side of the hyperplane so that the outliers do not influence the results. At the very least, the influence on the precision is minimal (a few hundredth of millimeters). Let us mention that for the Mahalanobis kernel, the matrix $\Sigma$ is firstly computed with the original training data, then computed again without the detected outliers. The RBF kernel do not require any modification.

The effect of the presence of outliers in the sample is illustrated at Fig. 10c and Fig. 10d for, respectively the reference method and the OCSVM method (landmark representation with Mahalanobis kernel). The advantage of OCSVM model is visually obvious.



(a) Ref. Meth. (b) OCSVM (c) Ref. Meth. (d) OCSVM

Fig. 10: Qualitative comparison between the reference method and the OCSVM method: (a) and (b) are characterized by simulated noise applied on control points ($\pm 7.5$ pixels) - (c) and (d) are characterized by the presence of outliers in the sample. 3D reconstructed models are projected on the PA view. Manually identified landmarks are in cyan and the landmarks obtained from the reconstruction method are in green.

## IV. Conclusion

In this paper, we have described a novel way of modeling shapes, based on OCSVM. In our approach, the model is defined by the hyperplane holding in its positive side a sample of shapes. Then, this model was used in the context of the 3D reconstruction of the spine from bi-planar radiographs. Two representations of the spine were considered: a classical representation with anatomical landmarks and an articulated representation of the spine. In order to define the model, two kernel functions were also evaluated: an RBF kernel and a Mahalanobis kernel. Results showed that our approach was efficient for cases of patients with severe scoliosis and for post-operative patients with surgical instrumentation. The RBF kernel is to be preferred for the articulated representation while the Mahalanobis kernel provided better results for the landmark representation of the spine. Moreover, we have demonstrated that our model is hardly influenced by the presence of outliers in the training data, in contrast to classical statistical shape models. This aspect makes the OCSVM model a great alternative to classical statistical shape models in applications where the presence of outliers in the training data is possible.

Our new kernel-based shape model has proven to be effective with two three-dimensional representations of the spine. However, we could use it for other biomedical shapes in two or three dimensions. The choice of the kernel is application-specific and depends on the shape of interest.

Finally, we think that modeling shapes with kernels opens new possibilities in numerous biomedical applications.

## Acknowledgment

The first author performed part of the work at École Polytechnique de Montréal and thanks the Minister of Scientific Research of the French Community of Belgium, and the Fonds de Recherche du Québec - Nature et Technologies (FRQNT), for granting his visits. We also would like to express our gratitude to Prof. Farida Cheriet, head of the Laboratoire d'imagerie et de vision 4D (LIV4D) from École Polytechnique de Montréal for her precious help in this work. Finally, the authors would like to thank the anonymous reviewers for their valuable comments and thoughtful suggestions to improve the quality of the paper.

## References

[1] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active Shape Models-Their Training and Application," *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[2] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543–563, 2009.

[3] A. Tsai, J. Yezzi, A., W. Wells, C. Tempany, D. Tucker, A. Fan, W. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 137–154, 2003.

[4] A. El-Baz and G. Gimel'farb, "Robust medical images segmentation using learned shape and appearance models," in *MICCAI*, ser. LNCS, 2009, vol. 5761, pp. 281–288.

[5] F. Khalifa, G. Beache, G. Gimel'farb, G. Giridharan, and A. El-Baz, "Accurate automatic analysis of cardiac cine images," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 445–455, 2012.

[6] J. Novosad, F. Cheriet, Y. Petit, and H. Labelle, "3d reconstruction of the spine from a single x-ray and prior vertebrae models," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 9, pp. 1628–1639, 2004.

[7] J. Zhang, K.-K. Ma, M.-H. Er, and V. Chong, "Tumor Segmentation from Magnetic Resonance Imaging by Learning via one-class support vector machine," in *International Workshop on Advanced Image Technology*, 2004, pp. 207–211.

[8] L. Wang, G. Xu, L. Guo, X. Liu, and S. Yang, "3D Reconstruction of Head MRI Based on One Class Support Vector Machine with Immune Algorithm," in *29th Annual International Conference of the IEEE EMBS*, 2007, pp. 6015–6018.

[9] C.-C. Chang, "Deformable Shape Finding With Models Based on Kernel Methods," *IEEE Transactions on Image Processing*, vol. 15, no. 9, pp. 2743–2754, 2006.

[10] H. Jin, Q. Liu, H. Lu, and X. Tong, "Face detection using one-class svm in color images," in *7th International Conference on Signal Processing*, vol. 2, 2004, pp. 1431–1434.

[11] C. E. Aubin, J. Dansereau, F. Parent, H. Labelle, and J. A. de Guise, "Morphometric evaluations of personalised 3D reconstructions and geometric models of the human spine," *Medical & Biological Engineering & Computing*, vol. 35, no. 6, pp. 611–618, 1997.

[12] A. Mitulescu, I. Semaan, J. A. De Guise, P. Leborgne, C. Adamsbaum, and W. Skalli, "Validation of the non-stereo corresponding points stereoradiographic 3D reconstruction technique," *Medical & biological engineering & computing*, vol. 39, no. 2, pp. 152–158, 2001.

[13] 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, 2004.

[14] L. Humbert, J. A. De 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.

[15] S. Kadoury, F. Cheriet, and H. Labelle, "Personalized x-ray 3-d reconstruction of the scoliotic spine from hybrid statistical and image-based models," *IEEE Transactions on Medical Imaging*, vol. 28, no. 9, pp. 1422 –1435, 2009.

[16] D. C. Moura, J. Boisvert, J. G. Barbosa, H. Labelle, and J. M. R. S. Tavares, "Fast 3D reconstruction of the spine from biplanar radiographs using a deformable articulated model," *Medical Engineering & Physics*, vol. 33, pp. 924–933, 2011.

[17] J. Boisvert and D. C. Moura, "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*, 2011, pp. 5726–5729.

[18] F. Lecron, J. Boisvert, M. Benjelloun, H. Labelle, and S. Mahmoudi, "Multilevel statistical shape models: a new framework for modeling hierarchical structures," in *IEEE International Symposium on Biomedical Imaging*, 2012, pp. 1284–1287.

[19] F. Lecron, J. Boisvert, S. Mahmoudi, H. Labelle, and M. Benjelloun, "Fast 3D Spine Reconstruction of Postoperative Patients Using a Multilevel Statistical Model," in *MICCAI*, ser. LNCS, 2012, vol. 7511, pp. 446–453.

[20] V. N. Vapnik, *The nature of statistical learning theory*. Springer-Verlag New York, 1995.

[21] B. Schölkopf, J. C. Platt, J. C. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[22] J. Boisvert, F. Cheriet, X. Pennec, H. Labelle, and N. Ayache, "Articulated Spine Models for 3-D Reconstruction from Partial Radiographic Data," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 11, pp. 2565–2574, 2008.

[23] R. Harmouche, F. Cheriet, H. Labelle, and J. Dansereau, "3d registration of mr and x-ray spine images using an articulated model," *Computerized Medical Imaging and Graphics*, vol. 36, no. 5, pp. 410 – 418, 2012.

[24] P. Cunha, D. C. Moura, and J. G. Barbosa, "Pedicle detection in planar radiographs based on image descriptors," ser. LNCS, vol. 7325, 2012, pp. 278–285.