CHAPITRE 6

ARTICLE 3: MODIFIED LARGE MARGIN NEAREST NEIGHBOR METRIC LEARNING FOR REGRESSION

K. C. Assi[a,b,*] , H. Labelle[b] and F. Cheriet[a,b], *Member, IEEE*

(a) *École Polytechnique de Montréal, P.O. Box 6097, Succursale Centre-ville, Montréal, Québec, Canada H3C 3A7*

(b) *Sainte-Justine Hospital Research Center, 3175 Côte-Sainte-Catherine, Montréal, Québec, Canada H3T 1C5*

6.1 Abstract

The main objective of this letter is to formulate a new approach of learning a Mahalanobis distance metric for nearest neighbor regression from a training sample set. We propose a modified version of the large margin nearest neighbor metric learning method to deal with regression problems. As an application, the prediction of post-operative trunk 3D shapes in scoliosis surgery using nearest neighbor regression is described. Accuracy of the proposed method is quantitatively evaluated through experiments on real medical data.

6.2 Introduction

The k-nearest neighbors (k-NN) rule [70] is one of the oldest and simplest methods in statistical prediction. Nearest neighbor regression consists in assigning to a new data point the response of the most similar in a dataset [73]. In k-NN regression, the output variable is predicted as a weighted average of the k nearest observations in a dataset, where the neighborhood is defined in terms of a chosen distance metric. Applications of k-NN methods range from computer vision, image retrieval and classification, to face recognition [108], speech recognition [109], human activity recognition and pose estimation, text analysis, and wireless sensor networks [110].

One of the key point in nearest neighbor based methods is to define a distance measure in the input space to identify nearest neighbors, and this mostly depends on the domain application. The default distance metric often used is the Euclidean distance. However ideally, each application requires a specific adapted distance metric since nearest neighbor methods have been demonstrated to have improved performance when used with a learned appropriate distance metric from a sample examples. One of the most learned metric is the Mahalanobis distance, and one of the most widely used Mahalanobis distance learning methods for k-NN

classification is the large margin nearest neighbor (LMNN) proposed by Weinberger et al. [75, 77]. Other metric learning methods for classification have also been proposed by different researchers such as adaptive metric nearest neighbor (ADAMENN) [78], and discriminant adaptive nearest neighbor (DANN) [79]. However these approaches have the disadvantage of requiring more than one parameter to be tuned, which make them less attractive compared to LMNN. The LMNN method has since been extended to other distances, for example in the χ^2-LMNN [111] where the χ^2 histogram distance is used in place of the Mahalanobis distance. Unfortunately most of the metric learning for k-NN approaches are essentially designed for classification problems [112]. Although nearest neighbors regression play an important role in statistical prediction [72], to the best of our knowledge, very few metric learning methods were proposed for nearest neighbor regression problems. The k-NN regression gives fairly similar performance as linear regression with respect to the average RMSEs in some applications, and it would be interesting to design an appropriate metric learning algorithm for k-NN regression.

In this letter, we propose a metric learning method for k-NN regression. We extend the LMNN method proposed in [75, 77] to the case of nearest neighbor regression. Although based on the same framework, however, our method introduces new features to deal with the specific case of regression, which otherwise can not be addressed in the classical LMNN. We then apply our model to the nearest neighbor prediction of postoperative 3D trunk shapes of scoliotic patients.

6.3 Modified large margin nearest neighbor metric learning

Let $D_n = \{z_i = (\mathbf{x}_i, \mathbf{y}_i) : \mathbf{x}_i \in \mathcal{X}, \mathbf{y}_i \in \mathcal{Y}, i = 1, ..., n\}$ be a dataset where $\mathcal{X}$ and $\mathcal{Y}$ are some metric spaces. We will refer to $\mathcal{X}$ and $\mathcal{Y}$ as the input and response space, respectively. The desired properties of distance metrics for regression are expressed as follows. Intuitively, in order to reduce regression prediction error, one may wish that two inputs $\mathbf{x}_i$ and $\mathbf{x}_j$ are close one to another in the input space if their respective responses $\mathbf{y}_i$ and $\mathbf{y}_j$ are also close one to another in the response space. More specifically, we would like, for any triplet of pairs $(\mathbf{x}_i, \mathbf{y}_i)$, $(\mathbf{x}_j, \mathbf{y}_j)$, $(\mathbf{x}_l, \mathbf{y}_l)$, if $\mathbf{x}_i$ is much closer to $\mathbf{x}_j$ than to $\mathbf{x}_l$ with respect to a distance $\delta_{\mathcal{X}}$ defined on $\mathcal{X}$ then it is likely $\mathbf{y}_i$ is much closer to $\mathbf{y}_j$ than to $\mathbf{y}_l$ with respect to a chosen error distance $\delta_{\mathcal{Y}}$ in the response space. In this case, the proximity order of the triplet is preserved.

6.3.1 Intuition of our modified LMNN

We base our model on the following intuitions to insure an accurate nearest neighbor regression: (i) each training input $\mathbf{x}_i$ and its k nearest neighbors should preserve proximity order,
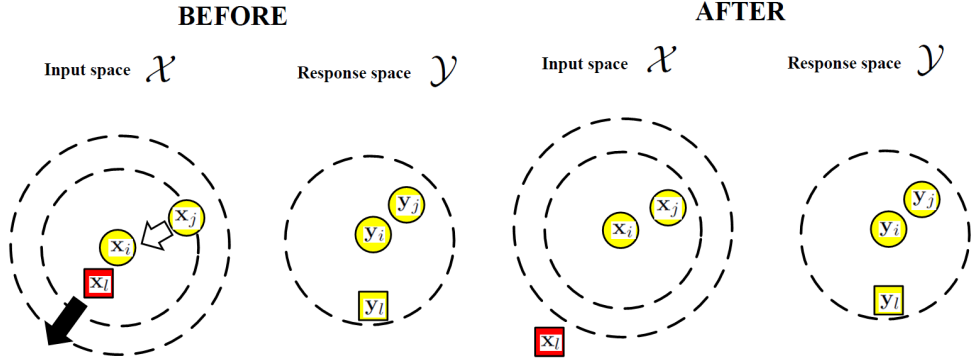
BEFORE **AFTER**

Figure 6.1 Illustration of the intuition behind the modified LMNN metric learning for regression. The point $\mathbf{x}_j$ is referred to as the target neighbor. The point $\mathbf{x}_l$ is referred to as an impostor since it violates the proximity order preservation (in this case). The training consists in finding a learned metric $\delta_{\mathcal{X}}$ such that: (1) the target neighbor is pushed closer to the input query point within a smaller radius after training, (2) impostor is pushed outside the smaller radius domain by a finite margin.

(ii) for a given triplet $(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_l)$, training inputs $\mathbf{x}_l$ that violate proximity order should be widely separated from $\mathbf{x}_i$ in such a way that proximity order is restored. Borrowing from the same terminology as in [75], for an input $\mathbf{x}_i$ with response $\mathbf{y}_i$ and target neighbor[1] $\mathbf{x}_j$, we call an impostor any input $\mathbf{x}_l$ with response $\mathbf{y}_l$ such that

$$\begin{cases} \delta_{\mathcal{Y}}(\mathbf{y}_i, \mathbf{y}_l) & > \delta_{\mathcal{Y}}(\mathbf{y}_i, \mathbf{y}_j) \\ \delta_{\mathcal{X}}(\mathbf{x}_i, \mathbf{x}_l) & \leq \delta_{\mathcal{X}}(\mathbf{x}_i, \mathbf{x}_j) + \epsilon, \end{cases} \tag{6.1}$$

where $\epsilon > 0$ is the margin. Specifically, an impostor $\mathbf{x}_l$ is any input violating proximity order and that invades the perimeter within a ϵ-margin defined by any target neighbor $\mathbf{x}_j$ of the input $\mathbf{x}_i$. We aim to learn a linear transformation of the input space such that the training inputs satisfy the above mentioned properties. Figure 6.1 illustrates the main idea behind our modified LMNN metric learning for regression. This is a regression oriented adaptation of the idealized error reduction scenario of the classical LMNN [75]. It shows how nearest neighbor regression errors in the original input space are corrected by learning an appropriate linear transformation. Before learning, a training input has both target neighbor and impostor in its local neighborhood. During learning, the impostor is pushed outside the perimeter established by the target neighbor. After learning, the mapped inputs points are such that there exists a finite margin between the perimeter and the impostor, and proximity order in both input and response spaces is restored. From the way they are presented, these

[1]Target neighbors are selected by using prior knowledge (if available) or by simply computing the k nearest neighbors using Euclidean distance [75].

ideas can be cast into the framework of the large margin nearest neighbor and be stated as two competing terms in our model's loss function, where one term penalizes large distances between nearby inputs that preserve proximity order, while the other term penalizes small distances between inputs which violate proximity order.

6.3.2 Proximity order preservation indicator

The key point of our metric learning approach for regression problems is the introduction of the proximity order preservation concept. Let us define the proximity order function $F_{\mathcal{M},\delta_{\mathcal{M}}}$ for a metric space $\mathcal{M}$ equipped with the distance $\delta_{\mathcal{M}}$, as for $\tau_{ijl} = (u_i, u_j, u_l) \in \mathcal{M}^3$,

$$F_{\mathcal{M},\delta_{\mathcal{M}}}(\tau_{ijl}) = \delta_{\mathcal{M}}(u_i, u_j) - \delta_{\mathcal{M}}(u_i, u_l).$$

For the triplets of pairs of input and response $((\mathbf{x}_i, \mathbf{y}_i), (\mathbf{x}_j, \mathbf{y}_j), (\mathbf{x}_l, \mathbf{y}_l))$, we denote $\mathbf{t}_{ijl} = (\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_l)$ and $\mathbf{t}^*_{ijl} = (\mathbf{y}_i, \mathbf{y}_j, \mathbf{y}_l)$ as the triplets in $\mathcal{X}$ and $\mathcal{Y}$, respectively. Let us define the triplets labeling function Π as:

$$\Pi_{\mathcal{Y},\delta_{\mathcal{Y}}}(\mathbf{t}) := \begin{cases} 1 & \text{if } F_{\mathcal{Y},\delta_{\mathcal{Y}}}(\mathbf{t}^*) \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.2}$$

The function $\Pi_{\mathcal{Y},\delta_{\mathcal{Y}}}$ assigns 0/1 labels to triplets $(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_l)$ based on a chosen distance $\delta_{\mathcal{Y}}$ used for measuring errors in the response space $\mathcal{Y}$. Let $\mathcal{C}_{\delta_{\mathcal{X}}}$ be a function on triplets of $\mathcal{X}$ associated to distance $\delta_{\mathcal{X}}$ defined as :

$$\mathcal{C}_{\delta_{\mathcal{X}}}(\mathbf{t}) := \begin{cases} 1 & \text{if } F_{\mathcal{X},\delta_{\mathcal{X}}}(\mathbf{t}) \geq 0 \\ 0 & \text{otherwise .} \end{cases} \tag{6.3}$$

We introduce the proximity order preservation indicator ν_{ijl}, as

$$\nu_{ijl} = \nu(\mathbf{t}_{ijl}) = \mathcal{C}_{\delta_{\mathcal{X}}^{(0)}}(\mathbf{t}_{ijl}) \Pi_{\mathcal{Y},\delta_{\mathcal{Y}}}(\mathbf{t}_{ijl}). \tag{6.4}$$

Note that ν_{ijl} is equal to 1 only if the proximity order of the triplet is preserved. A k-NN regression using a distance $\delta_{\mathcal{X}}$ with a lower rate of proximity order violation is likely to have a better accuracy performance (See Appendix, Proposition 1 and Corrolary 1). We use this to define a cost term that penalizes small distances between input points which violate the proximity order.

6.3.3 Loss function

Given a training set of features $\mathbf{x}_i$ along with their response $\mathbf{y}_i$ with $(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^{\dim_1} \times \mathbb{R}^{\dim_2}$, $(i = 1, \ldots, n)$, we are interested in learning a Mahalanobis distance metric parameterized by a linear transformation $\mathbf{L}$, i.e.

$$D_{\mathbf{L}}(\mathbf{x}_i, \mathbf{x}_j) = \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j)\|^2 = (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{L}^\top \mathbf{L}(\mathbf{x}_i - \mathbf{x}_j), \tag{6.5}$$

which allows an accurate nearest neighbor regression. The parameter matrix $\mathbf{L}$ is to be chosen such as to minimize the distance between the mappings of a vector and its k nearest neighbors

$$\mathcal{E}_1(\mathbf{L}) = \sum_{ij} \eta_{ij} \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j)\|^2, \tag{6.6}$$

where η_{ij} is given by

$$\eta_{ij} = \begin{cases} 1 & \text{if } j \text{ is a target neighbor of } i \\ 0 & \text{otherwise.} \end{cases}$$

The parameter $\mathbf{L}$ should also allow that the distance of $\mathbf{x}_i$ from a target neighbor $\mathbf{x}_j$ be less than its distance from an incorrect neighbor $\mathbf{x}_l$ (referred to as an impostor). To this end, we use ν_{ijl} in (6.4) and consider minimizing a hinge loss over triplets of input vectors

$$\mathcal{E}_2(\mathbf{L}) = \sum_{ijl} \eta_{ij}(1 - \nu_{ijl}) \left[\epsilon + \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_j)\|^2 - \|\mathbf{L}(\mathbf{x}_i - \mathbf{x}_l)\|^2 \right]_+, \tag{6.7}$$

where $\left[z \right]_+ = \max \left\{ z, 0 \right\}$, and ϵ is the margin. We end up with a total cost function $\mathcal{E}(\mathbf{L})$ that combines the two competing terms $\mathcal{E}_1$ and $\mathcal{E}_2$ using a weight parameter $\mu \in (0, 1)$:

$$\mathcal{E}(\mathbf{L}) = (1 - \mu)\mathcal{E}_1(\mathbf{L}) + \mu\mathcal{E}_2(\mathbf{L}). \tag{6.8}$$

Since $\mathbf{M} = \mathbf{L}^\top \mathbf{L} \succeq 0$, and by introducing a nonnegative slack variables ξ_{ijl} for each triplet, the minimization of the cost function $\mathcal{E}(\mathbf{L})$ can be formulated as a convex semidefinite pro-

gramming problem (SDP),

$$\begin{aligned}
\underset{\mathbf{M},\xi}{\text{minimize}} \quad & (1-\mu)\sum \eta_{ij}(\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{M}(\mathbf{x}_i - \mathbf{x}_j) \\
& + \mu \sum_{ijl} \eta_{ij}(1 - \nu_{ijl})\xi_{ijl} \\
\text{subject to} \quad & (\mathbf{x}_i - \mathbf{x}_l)^\top \mathbf{M}(\mathbf{x}_i - \mathbf{x}_l) \\
& - (\mathbf{x}_i - \mathbf{x}_j)^\top \mathbf{M}(\mathbf{x}_i - \mathbf{x}_j) \geq \epsilon - \xi_{ijl}, \\
& \xi_{ijl} \geq 0, \\
& \mathbf{M} \succeq 0.
\end{aligned} \tag{6.9}$$

The SDP problem (6.9) resulting from our formulation has the same form as the one proposed by Weinberger *et al.* [75] for their LMNN. However there are some differences in the terms involved, in particular the introduction of the proximity order preservation indicator ν_{ijl} (6.4) in the loss function. Equation (6.9) addresses nearest neighbor regression problems, whereas [75] aimed to solve only nearest neighbor classification problems.

6.4   Results

Applications to prediction of scoliotic trunk 3D shapes after spine surgery using nearest neighbor (NN) regression were conducted. For the experiments, some characteristic feature curves of the human scoliotic trunk surface topography were considered. The so-called back valley curves were extracted on preoperative and postoperative trunk surfaces. The shape of this feature curve almost follows that of the spine, but is also influenced by the muscle surrounding the spine and supporting the trunk in the upright posture. The upper and lower end points of the curve are anatomical landmarks corresponding to the spinous process of C7 (seventh cervical vertebra) and L5 (the fifth lumbar vertebra), respectively. This feature curve has the advantage of capturing the changing taking place in the back from the preoperative to the postoperative state. It is indeed the feature curve on the trunk surface whose shape change is the most directly influenced by a spine surgery instrumentation. Figure 6.2 (Left) shows the back valley curve along the spinous processes of a scoliotic patient in the preoperative (red) and the postoperative (blue) status overlayed on the patient postoperative trunk mesh. The associated deformation field is depicted on Figure 6.2 (Right). A dataset of 141 pairs of scoliotic shapes data, from teenagers aged between $11 - 18$ years old, is considered. Our proposed Mahalanobis distance metric learning is evaluated using the leave-one-out cross-validation procedure. Each sample point is composed of a pair (the preoperative and postoperative shapes). The predicted curve can then be compared with
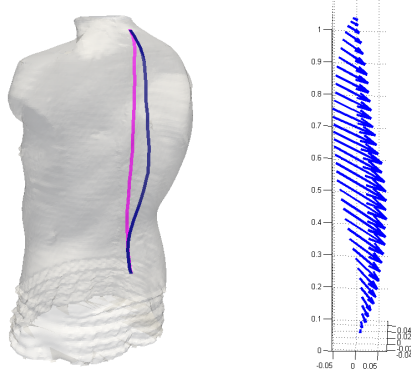
Figure 6.2 Left: Overlayed back valley curves (preoperative (blue) and postoperative (red)) on a patient surface mesh. Right: Displacement vector field along the back valley curve from preoperative to postoperative state.

the actual postoperative feature curve and a shape prediction error can be computed. To evaluate our results, we compute the prediction error in terms of the normalized root mean square prediction error, which has the advantage of allowing prediction error to be measured on the same scale for all observations. Quarter polar plots are used to visualize pointwise prediction errors along the back valley curve. The radius represents errors values while the angles θ (or the points along the arc) correspond each to the location of points along the back valley curve, where $\theta = 0$ is the bottom endpoint and $\theta = 90$ is the upper endpoint. The error unit is relative to the span of the spine, which is set to 1 after a common rigid registration of all trunk shapes data. Distances with better performances have errors graph closer to the origin point of the polar plot.

The proposed learned metric is compared with two other metrics: the default Euclidean distance and an arbitrary Mahalanobis metric which is defined by a random semidefinite positive matrix. The quarter polar plots of the mean errors are presented in Figure 6.3. The mean errors for the learned Mahalanobis are less than the ones for the Euclidean and arbitrary Mahalanobis metrics, all along the back valley, with a maximum mean error (0.043) attained around the mid-level of the trunk for the learned Mahalanobis and a maximum of 0.055, 0.056 for the Euclidean and arbitrary Mahalanobis respectively attained around the mid-lumbar level. A maximum error difference between the learned Mahalanobis and the Euclidean distance is found around the mid-lumbar level (MLL). The mean prediction error at the MLL level decreased significantly (0.016) between NN prediction using the Euclidean metric (Mean $\pm$ SD: 0.0531 ± 0.037, $N = 141$) and prediction using learned Mahalanobis metric (Mean $\pm$ SD: 0.0371 ± 0.0213, $N = 141$), (two-sample t-test assuming equal variance [113], p < 0.001). It appears that the NN regression using our learned Mahalanobis metric
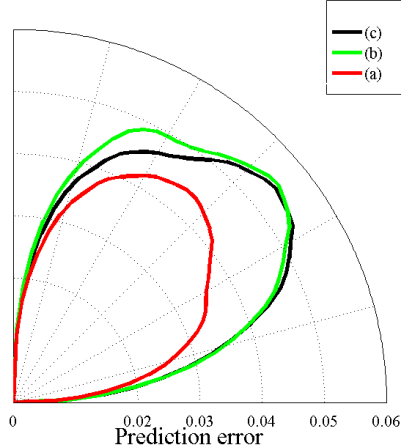
Figure 6.3 Quarter polar plots of the mean pointwise prediction errors distribution along the back valley curve of $N = 141$ scoliotic patients. (a) Red: Learned Mahalanobis metric, (b) Green: Arbitrary Mahalanobis metric, (c) Black: Euclidean metric.

has the lower errors, on average, for the nearest neighbor regression prediction.

We cast our metric learning for nearest neighbor regression as the SDP problem (6.9), which has the same form as the classical LMNN formulation. However, it is important to mention the main differences. Weinberger et *al.*'s LMNN was aimed to solve k-NN classification problems and for that the second term of their loss function contains an indicator (y_{il}) which expresses whether or not target neighbors have the same label (i.e $y_{il} = 1$ if and only if $y_i = y_l$ and, $y_{il} = 0$ otherwise). This indicator is no longer relevant in the case of a k-NN regression problem since the response space is continuous, and had to be replaced by some other appropriate relevant indicator. Here in our formulation, we make use of a triplet indicator ν_{ijl} which expresses whether or not proximity order, under the distance used, between target neighbors is preserved from the input space to the response space. This indicator is fixed during the learning process for a given training set, and has the advantage of allowing us to keep the same form of the metric learning SDP as LMNN in [75], while at the same time allowing to deal with regression. Our proposed modified LMNN metric learning method produces a Mahalanobis metric that outperforms the Euclidean metric for nearest neighbor regression on scoliotic trunk 3D shapes data.

6.5   Conclusion

In this letter, we have presented a new metric learning method for regression. The proposed method is an extension of the large margin nearest neighbor metric learning method to tackle nearest neighbor regression problems. It has been successfully applied to the prediction of

$3D$ trunk shapes data, and the learned metric has significantely improved the accuracy of the prediction compared to the default Euclidean metric. The modified LMNN described here could also be applied to general functional nearest neighbor regression.

APPENDIX: Risk of a distance metric for nearest neighbor regression

In nearest neighbor regression, the response estimate $\tilde{\mathbf{y}} \in \mathcal{Y}$ of a new data point $\mathbf{x} \in \mathcal{X}$, is given by

$$\tilde{\mathbf{y}} = r_{\mathrm{NN}_{\delta_{\mathcal{X}}}}(\mathbf{x}) = \mathbf{y}_{k:\mathbf{x}_k = NN_{\delta_{\mathcal{X}}}(\mathbf{x})}, \tag{6.10}$$

where $NN_{\delta_{\mathcal{X}}}(\mathbf{x})$ denotes the nearest neighbor of $\mathbf{x}$ in $D_n^{\mathcal{X}} = \{\mathbf{x}_i : (\mathbf{x}_i, \mathbf{y}_i) \in D_n, i = 1, ..., n\}$ with respect to the distance $\delta_{\mathcal{X}}$. Let $(X_0, Y_0) \in D_n$ such that X_0 is the true nearest neighbor of X in $D_n^{\mathcal{X}}$ with respect to $\delta_{\mathcal{X}}$. We define the error $\mathrm{err}_{\mathrm{NN}_{\delta_{\mathcal{X}}}}(X, Y) = \delta_{\mathcal{Y}}(Y, r_{\mathrm{NN}_{\delta_{\mathcal{X}}}}(X)) = \|Y - r_{\mathrm{NN}_{\delta_{\mathcal{X}}}}(X)\|$. We are interested in selecting a distance $\delta_{\mathcal{X}}$ to reduce the expected error $\mathbb{E}[\mathrm{err}_{\mathrm{NN}_{\delta_{\mathcal{X}}}}]$. From the initial dataset D_n, one can derive a dataset $\mathcal{D}$ of $\{0, 1\}$-labeled triplets

$$\begin{aligned} \mathcal{D} = \{s = &(\mathbf{t}_{ijl}, \lambda_{ijl}) \in \mathcal{X}^3 \times \{0, 1\}, \\ &\mathbf{t}_{ijl} = (\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_l), \lambda_{ijl} = \Pi_{\mathcal{Y}, \delta_{\mathcal{Y}}}(\mathbf{t}_{ijl})\}. \end{aligned} \tag{6.11}$$

When $\mathcal{C}_{\delta_{\mathcal{X}}}(\mathbf{t}) \neq \lambda$, a violation of the proximity order occurs. The loss function for the distance $\delta_{\mathcal{X}}$ is defined for $s = (\mathbf{t}, \lambda) \in \mathcal{X}^3 \times \{0, 1\}$ as

$$\ell(\delta_{\mathcal{X}}, s) = \mathbf{1}_{\{\mathcal{C}_{\delta_{\mathcal{X}}}(\mathbf{t}) \neq \lambda\}} = 1 - \nu_{\mathbf{t}}, \tag{6.12}$$

where $\nu_{\mathbf{t}} = \mathcal{C}_{\delta_{\mathcal{X}}}(\mathbf{t}) \Pi_{\mathcal{Y}, \delta_{\mathcal{Y}}}(\mathbf{t})$ is obtained from (6.4). The risk of a distance $\delta_{\mathcal{X}}$ is then defined as the expected loss :

$$R(\delta_{\mathcal{X}}) = \mathbb{E}[\ell] = \mathbb{P}_{s = (\mathbf{t}, \lambda) \sim P'}\{\mathcal{C}_{\delta_{\mathcal{X}}}(\mathbf{t}) \neq \lambda\}. \tag{6.13}$$

The key property of distances supporting our metric learning approach is: the lower the proximity order violation rate, the better the nearest neighbor regression. Let us illustrate this idea on a single triplet. Consider the three points sample set $D_3 = \{(\mathbf{x}_0, \mathbf{y}_0), (\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2)\}$, and let $s_0 = (\mathbf{t}_{012}, \lambda_{012})$, with $\mathbf{t}_{012} = (\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2)$ and $\lambda_{012} = 1$. Without loss of generality, let us choose $(\mathbf{x}_0, \mathbf{y}_0)$ as the test point. Suppose that $\delta_{\mathcal{X}}^{(1)}$ and $\delta_{\mathcal{X}}^{(2)}$ are two distances in $\mathcal{X}$ such that $\ell(\delta_{\mathcal{X}}^{(1)}, s_0) = 0$ and $\ell(\delta_{\mathcal{X}}^{(2)}, s_0) = 1$, i.e., $\ell(\delta_{\mathcal{X}}^{(1)}, s_0) < \ell(\delta_{\mathcal{X}}^{(2)}, s_0)$. In this case, $r_{\mathrm{NN}_{\delta_{\mathcal{X}}^{(1)}}}(\mathbf{x}_0) = \mathbf{y}_1$, $r_{\mathrm{NN}_{\delta_{\mathcal{X}}^{(2)}}}(\mathbf{x}_0) = \mathbf{y}_2$ and $\mathrm{err}_{\mathrm{NN}_{\delta_{\mathcal{X}}^{(1)}}}(\mathbf{x}_0, \mathbf{y}_0) = \|\mathbf{y}_0 - \mathbf{y}_1\| < \|\mathbf{y}_0 - \mathbf{y}_2\| = \mathrm{err}_{\mathrm{NN}_{\delta_{\mathcal{X}}^{(2)}}}(\mathbf{x}_0, \mathbf{y}_0)$. This property is formally stated as:

Proposition 1. *Let $\delta_{\mathcal{X}}^{(1)}$ and $\delta_{\mathcal{X}}^{(2)}$ be two distances in $\mathcal{X}$ with risk $R(\delta_{\mathcal{X}}^{(1)})$ and $R(\delta_{\mathcal{X}}^{(2)})$ respec-*

tively, such that $R(\delta_{\mathcal{X}}^{(1)}) < R(\delta_{\mathcal{X}}^{(2)})$. Then

$$\mathbb{E}[\text{err}_{\text{NN}_{\delta_{\mathcal{X}}^{(1)}}}] < \mathbb{E}[\text{err}_{\text{NN}_{\delta_{\mathcal{X}}^{(2)}}}]. \tag{6.14}$$

The best distance $\delta_{\mathcal{X}}$ in $\mathcal{X}$ is the one with minimum risk. The minimum risk distance δ^* is

$$\delta^* = \arg\inf_{\delta_{\mathcal{X}}} R(\delta_{\mathcal{X}}). \tag{6.15}$$

and

Corollary 1. *For all $\delta_{\mathcal{X}} \in D$,*

$$\mathbb{E}[\text{err}_{\text{NN}_{\delta^*}}] \leq \mathbb{E}[\text{err}_{\text{NN}_{\delta_{\mathcal{X}}}}]. \tag{6.16}$$

Minimizing the distance risk may allow us to design a metric learning method for regression.

RÉFÉRENCES

[1] S. Kadoury, F. Cheriet, C. Laporte, and H. Labelle. A versatile 3D reconstruction system of the spine and pelvis for clinical assessment of spinal deformities. *Medical and Biological Engineering and Computing*, 45(6) :591–602, 2007.

[2] W. Mollemans, F. Schutyser, N. Nadjmi, F. Maes, and P. Suetens. Predicting soft tissue deformations for maxillofacial surgery planning system : from computational strategies to a complete clinical validation. *Medical Image Analysis*, 11(3) :282–301, 2007.

[3] C.-E. Aubin, H. Labelle, and O. C. Ciolofan. Variability of spinal instrumentation configurations in adolescent idiopathic scoliosis. *European Spine Journal*, 16(1) :57–64, 2007.

[4] C.-E. Aubin, H. Labelle, C. Chevrefils, G. Desroches, J. Clin, and A. B. Eng. Preoperative planning simulator for spinal deformity surgeries. *Spine*, 33(20) :2143–2152, 2008.

[5] V. J. Raso, E. Lou, D. L. Hill, J. K. Mahood, M. J. Moreau, and N. G. Durdle. Trunk distorsion in adolescent idiopathic scoliosis. *J.Pediatr. Orthop.*, 18 :222–226, 1998.

[6] C. Denoel, M. F. I. Aguirre, G. Bianco, P. H. Mahaudens, R. Vanwijck, S. Garson, R. Sinna, and A. Debrun. Idiopathic scoliosis and breast asymmetry. *Journal of Plastic Reconstructive and Aesthetic Surgery*, 62 :1303–1308, 2009.

[7] T. G. Lowe, M. Edgar, J. Y. Margulies, N. H. Miller, V. J. Raso, K. A. Reinker, and C. H. Rivard. Etiology of idiopathic scoliosis : current trends in research. *J Bone Joint Surg Am*, 82-A :1157–1168, 2000.

[8] E. J. Rogala, D. S. Drummond, and J. Gurr. Scoliosis : incidence and natural history. a prospective epidemiological study. *J Bone Joint Surg Am*, 60 :173–176, 1978.

[9] J.R. Cobb. Outline for the study of scoliosis. *Am. Acad. Orthop. Surg. Instruct. Lect.*, 5 :261–275, 1948.

[10] S. Negrini. Bracing adolescent idiopathic scoliosis today. *Disabil Rehabil Assist Technol*, 3 :107–111, 2008.

[11] R.R. Betz and H. Shufflebarger. Anterior versus posterior instrumentation for the correction of thoracic idiopathic scoliosis. *Spine*, 26(9) :1095–1100, 2001.

[12] P. Papin, H. Labelle, S. Delorme, C.E. Aubin, J.A. De Guise, and J. Dansereau. Long-term three-timensional changes of the spine after posterior spinal instrumentation and fusion in adolescent idiopathic scoliosis. *European Spine Journal*, 8(1) :16–21, 1999.

[13] S. Delorme, P. Violas, J. Dansereau, J. de Guise, C.-E. Aubin, and H. Labelle. Preoperative and early postoperative three-dimensional changes of the rib cage after posterior instrumentation in adolescent idiopathic scoliosis. *European Spine Journal*, 10(2) :101–107, 2001.

[14] M. Asher, S. M. Lai, D. Burton, and B. Manna. Maintenance of trunk deformity correction following posterior instrumentation and arthrodesis for idiopathic scoliosis. *Spine*, 29 :1782–1788, 2004.

[15] R. K. Pratt, J. K. Webb, R. G. Burwell, and A. A. Cole. Changes in surface and radiographic deformity after universal spine system for right thoracic adolescent idiopathic scoliosis : is rib-hump reassertion a mechanical problem of the thoracic cage rather than an effect of relative anterior spinal overgrowth ? *Spine*, 26 :1778–1787, 2001.

[16] U. Willers, E. E. Transfeldt, and R. Hedlund. The segmental effect of cotrel-dubousset instrumentation on vertebral rotation, rib hump and the thoracic cage in idiopathic scoliosis. *European Spine Journal*, 5(6) :387–393, 1996.

[17] T. R. Haher, A. Merola, R. I. Zipnick, J. Gorup, D. Mannor, and J. Orchowski. Meta-analysis of surgical outcome in adolescent idiopathic scoliosis. a 35-year english literature review of 11,000 patients. *Spine*, 20(14) :1575–1584, 1995.

[18] A. E. Geissele, J. W. Ogilvie, M. Cohen, and D. S. Bradford. Thoracoplasty for the treatment of rib prominence in thoracic scoliosis. *Spine*, 19(14) :1636–1642, 1994.

[19] K.H. Bridwell. Surgical treatment of idiopathic adolescent scoliosis. *Spine*, 24(24) :2607–2616, 1999.

[20] 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 and Biological Engineering and Computing*, 35(6) :611–618, 1997.

[21] D. Périé, C. E. Aubin, M. Lacroix, Y. Lafon, and H. Labelle. Biomechanical modelling of orthotic treatment of the scoliotic spine including a detailed representation of the brace-torso interface. *Medical and Biological Engineering and Computing*, 42 :339–344, 2004.

[22] J. Carrier, C.-E. Aubin, I. Villemure, and H. Labelle. Biomechanical modelling of growth modulation following rib shortening or lengthening in adolescent idiopathic scoliosis. *Medical and Biological Engineering and Computing*, 42(4) :541–548, 2004.

[23] Clin J., Aubin C.-E., Parent S., Ronsky J., and Labelle H. Biomechanical modeling of brace design. *Stud. Health Technol. Inform.*, 123 :255–260, 2006.

[24] J. Clin, C.-E. Aubin, and H. Labelle. Virtual prototyping of a brace design for the correction of scoliotic deformities. *Medical and Biological Engineering and Computing*, 45(5) :467–473, 2007.

[25] J. Clin, C.-E. Aubin, S. Parent, A. Sangole, and H. Labelle. Comparison of the biomechanical 3D efficiency of different brace designs for the treatment of scoliosis using a finite element model. *Eur. Spine J.*, 19(7) :1169–1178, 2010.

[26] J. Clin, C.-E. Aubin, S. Parent, and H. Labelle. A biomechanical study of the charleston brace for the treatment of scoliosis. *Spine*, Publish Ahead of Print, 2010.

[27] J. Clin, C.-E. Aubin, H. Labelle, and S. Parent. Immediate correction in brace treatment : how much is needed to obtain a long-term effectiveness ? In *8th Meeting of the International Research Society of Spinal Deformities (IRSSD 2010)*, Montréal, Québec, Canada, 2010.

[28] M. Beauséjour, C.E. Aubin, A.G. Feldman, and H. Labelle. Simulations de tests d'inflexion latérale à l'aide d'un modèle musculo-squelettique du tronc. *Annales de chirurgie*, 53 :742–750, 1999.

[29] O. Dionne, K. C. Assi, S. Grenier, H. Labelle, F. Guibault, and F. Cheriet. Simulation of the postoperative trunk appearance in scoliotic surgery. In *International Symposium On Biomedical Imaging, ISBI 2012*, pages 1208–1211, 2012.

[30] U. Meier, O. López, C. Monserrat, M. C. Juan, and M. Alcaniz. Real-time deformable models for surgery simulation : a survey. *Computer Methods and Programs Biomedicine*, 77(3) :183–197, 2005.

[31] R. M. Koch, M. H. Gross, F. R. Carls, D. F. von Büren, G. Fankhauser, and Y. I. H. Parish. Simulating facial surgery using finite element models. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 421–428. ACM, 1996.

[32] R. M. Koch, S. H. M. Roth, M. H. Gross, A. P. Zimmermann, and H. F. Sailor. A framework for facial surgery simulation. In *Proceedings of the 18th spring conference on Computer graphics*, pages 33–42, 2002.

[33] D. Terzopoulos and K. Waters. Physically-based facial modeling, analysis, and animation. *Journal of Visualization and Computer Animation*, 1(2) :73–80, 1990.

[34] E. Keeve, S. Girod, and B. Girod. Craniofacial surgery simulation. In *Visualization in Biomedical Computing*, pages 541–546. Springer : Berlin, 1996.

[35] W. Mollemans, F. Schutyser, J.V. Cleynenbreugel, and P. Suetens. Tetrahedral mass spring model for fast soft tissue deformation. In *Surgery Simulation and Soft Tissue Modeling IS4TM 2003, LNCS 2673*, pages 145–154. Springer : Berlin, 2003.

[36] M. Bro-Nielsen. Surgery simulation using fast finite elements. In *VBC '96 : Proceedings of the 4th International Conference on Visualization in Biomedical Computing*, pages 529–534. Springer-Verlag, 1996.

[37] S. Cotin, H. Delingette, and N. Ayache. Real-time elastic deformations of soft tissues for surgery simulation. *IEEE Transactions On Visualization and Computer Graphics*, 5(1) :62–73, 1999.

[38] Matthias Müller, Julie Dorsey, Leonard McMillan, Robert Jagnow, and Barbara Cutler. Stable real-time deformations. In *SCA '02 : Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation*, pages 49–54. ACM, 2002.

[39] Guillaume Picinbono, Herve Delingette, and Nicholas Ayache. Real-time large displacement elasticity for surgery simulation : Non-linear tensor-mass model. In *MICCAI '00 : Proceedings of the Third International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 643–652. Springer-Verlag, 2000.

[40] M. Chabanas, V. Luboz, and Y. Payan. Patient specific finite element model of the face soft tissues for computer-assisted maxillofacial surgery. *Medical Image Analysis*, 7(2) :131–151, 2003.

[41] Liesbet Roose, Wim De Maerteleire, Wouter Mollemans, and Paul Suetens. Validation of different soft tissue simulation methods for breast augmentation. *International Congress Series*, 1281 :485–490, 2005.

[42] R. Mahnken and E. Stein. A unified approach for parameter identification of inelastic material models in the frame of the finite element method. *Computer Methods in Applied Mechanics and Engineering*, 136(3-4) :225 – 258, 1996.

[43] R. Mahnken and E. Stein. Parameter identification for viscoplastic models based on analytical derivatives of a least-squares functional and stability investigations. *International Journal of Plasticity*, 12(4) :451 – 479, 1996.

[44] R. Mahnken and E. Stein. Parameter identification for finite deformation elasto-plasticity in principal directions. *Computer Methods in Applied Mechanics and Engineering*, 147(1-2) :17 – 39, 1997.

[45] T. Seibert, J. Lehn, S. Schwan, and F.G. Kollmann. Identification of material parameters for inelastic constitutive models : Stochastic simulations for the analysis of deviations. *Continuum Mechanics and Thermodynamics*, 12(2) :95 – 120, 2000.

[46] T. Harth, S. Schwan, J. Lehn, and F. G. Kollmann. Identification of material parameters for inelastic constitutive models : statistical analysis and design of experiments. *International Journal of Plasticity*, 20(8-9) :1403 – 1440, 2004.

[47] W. T. D'Arcy. *On Growth and Form*. Cambridge University Press, 1917.

[48] U. Grenander and M. I. Miller. Computational anatomy : an emerging discipline. *Q. Appl. Math.*, LVI(4) :617–694, 1998.

[49] Daniel P. Huttenlocher, Gregory A. Klanderman, Gregory A. Kl, and William J. Rucklidge. Comparing images using the hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15 :850–863, 1993.

[50] Guillaume Charpiat, Olivier Faugeras, and Renaud Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Found. Comput. Math.*, 5(1) :1–58, 2005.

[51] *Distance-Based Shape Statistics*, volume 5, 2006.

[52] D.G. Kendall. Shape manifolds, procrustean metrics, and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2) :81–121, 1984.

[53] F. L. Bookstein. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, 1(2) :181–242, 1986.

[54] Simone Ceolin, William A. P. Smith, and Edwin Hancock. Facial shape spaces from surface normals and geodesic distance. In *DICTA '07 : Proceedings of the 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications*, pages 416–423, Washington, DC, USA, 2007. IEEE Computer Society.

[55] Simone Ceolin, William Smith, and Edwin Hancock. Facial shape spaces from surface normals. In *Image Analysis and Recognition*, pages 955–965. 2008.

[56] C. R. Rao. Information and accuracy attainable in estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37 :81–91, 1945.

[57] *Shape analysis using the Fisher-Rao Riemannian metric : unifying shape representation and deformation*, 2006.

[58] Adrian M. Peter and Anand Rangarajan. Information geometry for landmark shape analysis : Unifying shape representation and deformation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31 :337–350, 2009.

[59] Simone Ceolin and Edwin R. Hancock. Using the fisher-rao metric to compute facial similarity. In *ICIAR (1)*, pages 384–393, 2010.

[60] J. Glaunes, A. Trouvé, and L. Younes. Diffeomorphic matching of distributions : A new approach for unlabelled point-sets and sub-manifolds matching. In *In CVPR*, pages 712–718, 2004.

[61] M. Vaillant and J. Glaunès. Surface matching via currents. In *Proceedings of Information Processing in Medical Imaging (IPMI 2005), number 3565 in Lecture Notes in Computer Science*, pages 381–392, 2005.

[62] J. Glaunès, A. Qiu, M. I. Miller, and L. Younes. Large deformation diffeomorphic metric curve mapping. *Int. J. Comput. Vision*, 80(3) :317–336, 2008.

[63] S. Durrleman, X. Pennec, A. Trouvé, and N. Ayache. Statistical models of sets of curves and surfaces based on currents. *Medical Image Analysis*, 13(5) :793–808, 2009. Includes Special Section on the 12th International Conference on Medical Imaging and Computer Assisted Intervention.

[64] N. Acir and C. Guzelis. Automatic spike detection in eeg by a two-stage procedure based on support vector machines. *Computers in Biology and Medicine*, 34(7) :561–575, 2004.

[65] G.-Z. Li, J. Yang, C.-Z. Ye, and D.-Y. Geng. Degree prediction of malignancy in brain glioma using support vector machines. *Comput. Biol. Med.*, 36(3) :313–325, 2006.

[66] C. Bergeron, F. Cheriet, J. Ronsky, R. Zernicke, and H. Labelle. Prediction of anterior scoliotic spinal curve from trunk surface using support vector regression. *Eng. Appl. Artificial Intell.*, 18(8) :973–983, 2005.

[67] J. Jaremko, P. Poncet, J. Ronsky, J. Harder, J. Harder, J. Dansereau, H. Labelle, and R. Zernicke. Genetic algorithm-neural network estimation of cobb angle from torso asymmetry in scoliosis. *J. Biomech. Eng.*, 124(5) :496–503, 2002.

[68] L. Ramirez, N.G. Durdle, V.J. Raso, and D.L. Hill. A support vector machines classifier to assess the severity of idiopathic scoliosis from surface topography. *IEEE Trans. Inf. Technol. Biomed.*, 10(1) :84–91, 2006.

[69] L. Seoud, M.M. Adankon, H. Labelle, J. Dansereau, and F. Cheriet. Prediction of scoliosis curve type based on the analysis of trunk surface topography. In *Biomedical Imaging : From Nano to Macro, 2010 IEEE International Symposium on*, pages 408–411, 2010.

[70] T.M. Cover and P.E. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, IT-13 :21–27, 1967.

[71] T. M. Cover. Estimation by the nearest neighbor rule. *IEEE Transactions on Information Theory*, IT-14(1) :50–55, 1968.

[72] L. Devroye, L. Gyorfi, A. Krzyzak, and G. Lugosi. On the strong universal consistency of nearest neighbor regression function estimates. *The Annals of Statistics*, 22 :1371–1385, 1994.

[73] S. R. Kulkarni and S. E. Posner. Rates of convergence of nearest neighbor estimation under arbitrary sampling. *IEEE Trans. Inf. Theory*, 41 :1028–1039, 1995.

[74] G. Biau, F. Cérou, and A. Guyader. Rates of convergence of the functional k-nearest neighbor estimate. *IEEE Transactions on Information Theory*, 56 :2034–2040, 2010.

[75] Kilian Weinberger, John Blitzer, and Lawrence Saul. Distance metric learning for large margin nearest neighbor classification. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Advances in Neural Information Processing Systems 18*, pages 1473–1480. MIT Press, Cambridge, MA, 2006.

[76] K. Weinberger and L. Saul. Fast solvers and efficient implementations for distance metric learning. pages 1160–1167, 2008.

[77] K.Q. Weinberger and L.K. Saul. Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 10.

[78] C. Domeniconi, J. Peng, and D. Gunopulos. Locally adaptive metric nearest neighbor classification. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(9) :1281–1285, 2002.

[79] T. Hastie and R. Tibshirani. Discriminant adaptive nearest neighbor classification. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(6) :607–615, 1996.

[80] Y. Lee, D. Terzopoulos, and K. Waters. Realistic modeling for facial animation. In *SIGGRAPH '95 : Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 55–62, New York, NY, USA, 1995. ACM.

[81] S. Sarni, A. Maciel, R. Boulic, and D. Thalmann. Evaluation and visualization of stress and strain on soft biological tissues in contact. In *SMI '04 : Proceedings of the Shape Modeling International 2004*, pages 255–262, Washington, DC, USA, 2004. IEEE Computer Society.

[82] S. Delorme, Y. Petit, J. A. de Guise, H. Labelle, C.-E. Aubin, and J. Dansereau. Assessment of the 3D reconstruction and high-resolution geometrical modeling of the human skeletal trunk from 2D radiographic images. *IEEE Transactions on Biomedical Engineering*, 50(8) :989–98, 2003.

[83] H. Si. Tetgen : A quality tetrahedral mesh generator and three-dimensional delaunay triangulator. http ://tetgen.berlios.de/.

[84] J. R. Shewchuk. Tetrahedral mesh generation by delaunay refinement. In *SCG '98 : Proceedings of the fourteenth annual symposium on Computational geometry*, pages 86–95. ACM, 1998.

[85] M. Teschner, B. Heidelberger, M. Müller, and M. Gross. A versatile and robust model for geometrically complex deformable solids. In *Proceedings of the Computer Graphics International*, pages 312–319. IEEE Computer Society, 2004.

[86] J. L. Jaremko, P. Poncet, J. Ronsky, J. Harder, and J. Dansereau. Indices of torso asymmetry related to spinal deformity in scoliosis. *Clinical Biomechanics*, 17(8) :559–568, 2002.

[87] L. Seoud, F. Cheriet, H. Labelle, and J. Dansereau. A novel method for the 3d reconstruction of scoliotic ribs from frontal and lateral radiographs. *IEEE Transactions in biomedical engineering*, 58(5) :1135–1146, 2011.

[88] V. Pazos, F. Cheriet, J. Dansereau, Janet Ronsky, Ronald F. Zernicke, and Hubert Labelle. Reliability of trunk shape measurements based on 3-d surface reconstructions. *European Spine Journal*, 16(11) :1882–1891, 2007.

[89] R. Buchanan, J. G. Birch, A. A. Morton, and R. H. Browne. Do you see what I see ? Looking at scoliosis surgical outcomes through orthopedists' eyes. *Spine*, 28(24) :2700–2704, 2003.

[90] Goran Devedzic, Sasa Cukovic, Vanja Lukovic, Danijela Milosevic, K. Subburaj, and Tanja Lukovic. Scoliomedis : Web-oriented information system for idiopathic scoliosis visualization and monitoring. *Computer Methods and Programs in Biomedicine*, 108(2) :736–749, 2012.

[91] V. Pazos, F. Cheriet, H. Labelle, and J. Dansereau. 3d reconstruction and analysis of the whole trunk surface for non-invasive follow-up of scoliotic deformities. *Studies in health technology and informatics*, 91 :296–299, 2002.

[92] T.M.L. Shannon. *Dynamic Surface Topography and Its Application to the Evaluation of Adolescent Idiopathic Scoliosis.* PhD thesis, Oxford Brookes University, Oxford, U.K., september 2010.

[93] Peter O. Ajemba, Nelson G. Durdle, and V. James Raso. Characterizing torso shape deformity in scoliosis using structured splines models. *IEEE Trans. Biomed. Engineering*, 56(6) :1652–1662, 2009.

[94] S. Li, W. Zhou, Q. Yuan, S. Geng, and D. Cai. Feature extraction and recognition of ictal eeg using emd and svm. *Computers in Biology and Medicine*, 43(7) :807–816, 2013.

[95] A. Subasi. Classification of emg signals using pso optimized svm for diagnosis of neuromusculardisorders. *Computers in Biology and Medicine*, 43(5) :576–586, 2013.

[96] E. Comak, A. Arslan, and I. Türkoglu. A decision support system based on support vector machines for diagnosis of the heart valve diseases. *Computers in biology and Medicine*, 37(1) :21–27, 2007.

[97] S. Kiranyaz, T. Ince, J. Pulkkinen, M. Gabbouj, J. Ärje, S. Kärkkäinen, V. Tirronen, M. Juhola, T. Turpeinen, and K. Meissner. Classification and retrieval on macroinvertebrate image databases. *Computers in Biology and Medicine*, 41(7) :463–472, 2011.

[98] H. M. Muda, P. Saad, and R. M. Othman. Remote protein homology detection and fold recognition using two-layer support vector machine classifiers. *Computers in Biology and Medicine*, 41(8) :687–699, 2011.

[99] T. Hastie, R. Tibshirani, and Friedman. *The elements of statistical learning.*

[100] H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24 :417—-441, 1933.

[101] B. Schölkopf, A.J. Smola, and K.-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10(5) :1299–1319, 1998.

[102] X. He and P. Niyogi. Locality Preserving Projections. Cambridge, MA, 2004. MIT Press.

[103] X. He, D. Cai, S. Yan, and H.-J. Zhang. Neighborhood preserving embedding. In *Proceedings of IEEE International Conference on Computer Vision*, volume 2, pages 1208–1213. IEEE, 2005.

[104] H. Wold. Soft Modeling by Latent Variables ; the Nonlinear Iterative Partial Least Squares Approach. *Perspectives in Probability and Statistics. Papers in Honour of M. S. Bartlett*, 1975.

[105] S. Wold, H. Ruhe, H. Wold, and W.J. Dunn III. The collinearity problem in linear regression. the partial least squares (PLS) approach to generalized inverse. *SIAM Journal of Scientific and Statistical Computations*, 5(3) :735–743, 1984.

[106] V. Pazos, F. Cheriet, L. Song, H. Labelle, and J. Dansereau. Accuracy assessment of human trunk surface 3d reconstructions from an optical digitizing system. *Medical and Biological Engineering and Computing*, 43(1) :11–15, 2005.

[107] J.A.K. Suykens, V.T. Gestel, J. De Brabanter, B. De Moor, and J. Vandewalle. *Least Squares Support Vector Machines*. World Scientific, Singapore, 2002.

[108] S.-M. Huang and J.-F. Yang. Unitary regression classification with total minimum projection error for face recognition. *IEEE signal processing letters*, 20(5) :443–446, 2013.

[109] B.L. Pellom, R. Sarikaya, and J.H.L. Hansen. Fast likelihood computation techniques in nearest-neighbor based search for continuous speech recognition. *IEEE Signal Processing Letters*, 8(8) :221–224, 2001.

[110] S. Marano, V. Matta, and P. Willett. Nearest-neighbor distributed learning by ordered transmissions. *IEEE Transactions on Signal Processing*, 61(21) :221–224, 2013.

[111] Dor Kedem, Stephen Tyree, Kilian Weinberger, Fei Sha, and Gert Lanckriet. Non-linear metric learning. In P. Bartlett, F.C.N. Pereira, C.J.C. Burges, L. Bottou, and K.Q.

Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 2582–2590. 2012.

[112] B. Kulis. Metric learning : A survey. *Foundations and Trends in Machine Learning*, 5(4) :287–364, 2013.

[113] J. Devore and N. Farnum. *Applied Statistics for Engineers and Scientists*. Duxbury Press, 1999.

[114] C.-É. Aubin, Y. Petit, I.A.F. Stokes, F. Poulin, M. Gardner-Morse, and H. Labelle. Biomechanical modeling of posterior instrumentation of the scoliotic spine. *Computer Methods in Biomechanics and Biomedical Engineering*, 6(1) :27–32, 2003.

[115] S. Cotin, H. Delingette, and N. Ayache. A hybrid elastic model allowing real-time cutting, deformations and force-feedback for surgery training and simulation. *The Visual Computer*, 16 :437–452, 2000.

[116] 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*, 36(5) :410 – 418, 2012.

[117] K.M Cheung and K.D Luk. Prediction of correction of scoliosis with se of the fulcrum bending radiograph. *J. Bone Joint Surg. Am.*, 79 :1144–1150, 1997.

[118] P. Debanné, V. Pazos, H. Labelle, and F. Cheriet. Evaluation of reducibility of trunk asymmetry in lateral bending. In *8th Meeting of the International Research Society of Spinal Deformities (IRSSD 2010)*, Montréal, Québec, Canada, 2010.