Statistical shape model of the spine fitting study: impact of clipping the latent representation

Manon Ansart[a], Thierry Cresson[a], Benjamin Aubert[a,b], Jacques de Guise[a], and Carlos Vázquez[a]

[a]Laboratoire de recherche en imagerie et orthopédie (LIO), École de technologie supérieure, CHUM Research Centre, Montréal, Québec, Canada
[b]EOS Imaging Inc., 6525 boulevard Saint-Laurent, Suite 200, H2S 3S2, Montréal, Québec, Canada

ABSTRACT

Statistical shape models (SSMs) represent the distribution of labeled points across a training set of shapes. The standard practice for SSMs based on principal component analysis (PCA) is to use clipping, thresholding the latent representation so that all shapes lie within 3 standard deviations of the mean. This practice precludes the representation of shapes that are not well represented by the training set, constraining the model to realistic solutions, but making it impossible to work with shapes at the edges of the statistical population. In this study, we investigate the impact of clipping in a PCA-based SSM and whether using L2 regularization is a good replacement for clipping in the context of the automatic 2D to 3D reconstruction of the spine. We first show that using L2 regularization is equivalent to using a probabilistic PCA with two error variables, accounting for the suppression of the least important principal components and for the fact that the training set cannot perfectly represent all shapes at test time. Secondly, we use two data sets of 1746 and 768 patients with adolescent idiopathic scoliosis to study the effect of regularization, for different regularization weights and with or without clipping, for removing landmark detection errors using a simulated noise or a reconstruction pipeline. In both sets of experiments, we show that regularization removes noise in a way similar to clipping without preventing the reconstruction of out-of-distribution shapes, leading to outputs closer to ground truth, demonstrating that using a regularized SSM should be preferred to clipping.

Keywords: Satistical shape model, principal component analysis, clipping, scoliosis, spine, vertebrae, automatic reconstruction

1. INTRODUCTION

Statistical shape models (SSM) learn the typical distribution of labeled points from a training data set. They are used to bring statistical information when detecting shapes in images, for segmentation tasks and 3D shape reconstruction from 2D images, such as in 1 for the fully automated reconstruction of the 3D scoliotic spine.

The first statistical shape model, proposed in 2, relies on an average shape and deformation modes computed using principal component analysis (PCA). This original method proposed to set limits on the latent representation, as parameters are assumed to follow a gaussian distribution and 99.7% of the data set should lie within three standard deviations (std) of the average shape. This procedure known as "clipping" does not allow the reconstruction of shapes that are poorly represented by the training set. Since then, different PCA-based SSMs have been proposed to represent vertebrae models,[3,4,5] the whole spine[1,6,7] or the rib cage[8] for example. Several studies used variations of the PCA, using hierachical models,[9] kernel PCA,[5] a mixture of probabilistic PCAs.[8]

Further author information: (Send correspondence to M.A.)
M.A.: manon.ansart@u-bourgogne.fr
T.C.: thierry.cresson@etsmtl.ca
B.A.: baubert@eos-imaging.com
J.G.: jacques.deguise@etsmtl.ca
C.V.: carlos.vazquez@etsmtl.ca

Medical Imaging 2022: Image Processing, edited by Olivier Colliot, Ivana Išgum, Proc. of SPIE Vol. 12032, 1203212 · © 2022 SPIE
1605-7422 · doi: 10.1117/12.2606202

Despite the number of variations proposed around the original method, clipping remains a standard practice for 3D reconstruction,[10] but to our knowledge its impact has never been evaluated.

In 1 the authors proposed to use a regularization term which penalizes large values for coordinates in the latent space. This regularization has an effect similar to clipping, with a major difference: it attributes a higher cost to large values without making them impossible. This regularization therefore has the potential to constrain the SSM outputs to realistic shapes in a way similar to clipping while allowing the reconstruction of spine shapes poorly represented in the training set. No theoretical justification has however been proposed for this regularization, and it has never been used without clipping.

In this study, we investigate the impact of clipping in a PCA-based SSM of the spine. We use the model proposed in 1, and investigate whether the added regularization term can replace the clipping procedure, allowing to reconstruct shapes that are not well represented in the training set while correcting errors in landmark detection. Our contribution is two-fold. Firstly, we propose a theoretical justification for the PCA-based SSM with regularization proposed in 1, using a probabilistic PCA with two error variables, accounting for the suppression of the least important principal components and for the fact that the training set cannot perfectly represent all shapes at test time. Secondly, we investigate the impact of clipping this SSM in the context of the automatic 3D reconstruction of the scoliotic spine, using a simulated noise and using the reconstruction pipeline described in 1, on data sets of 768 and 1746 patients.

2. MATERIALS AND METHODS

2.1 Theoretical proposition

In this section, we propose a theoretical justification for the addition of a regularization term to the PCA-based SSM model proposed in 1. We first detail the solution proposed in 1 before showing that it is equivalent to maximising the posterior distribution of a probabilistic PCA model given partial data.

2.1.1 Existing model

Let the data set be composed of n_r shapes, each one represented by landmarks contained in a row vector of size n_c. The data set can be represented as a matrix $X \in \mathbb{R}^{n_r \times n_c}$. When fitting the model to a new shape, only a subset of the landmarks are observed and used as targets for the model. In 1, the authors propose to minimize the distance between the target observed landmarks of a shape x_o and the SSM output for coordinate i by using the objective function

$$\hat{m} = \arg\min_{m} \left(\sum_{i=1}^{n_c} w_i (x_{o,i} - (\overline{x} + Bm)_i)^2 + \frac{\beta}{2} \sum_{k=1}^{|m|} m_k^2 \right) \tag{1}$$

B contains the representation modes, m the latent representation of the shape and $\overline{x}$ the mean shape model. This equation is equivalent to minimizing the squared distance between the observed landmarks and the PCA outputs while penalizing deviation from the mean shape using a L2 penalty,[11] with a regularization weight β. In addition, 1 clips the latent representation, ensuring that the output is within 3 standard deviation of the mean shape by thresholding m to a minimum of -3 and a maximum of 3. Assigning a weight of $w_i = 1$ to all landmarks the solution to equation 1 is :

$$\hat{m} = (B_o^T B_o + \beta^2 I)^{-1} B_o^T (x_o - \overline{x}_o) \tag{2}$$

where B_o contains the rows of B corresponding to the observed landmarks. We show here that this model is equivalent to maximising the posterior distribution of a probabilistic PCA model given partial data, as proposed in 12, while accounting for two kinds of loss of information.

2.1.2 Maximum likelihood estimation for probabilistic PCA

Following 8, we use a probabilistic PCA model and define each shape x as a linear transformation of the latent variable m: $x = \overline{x} + Bm + \epsilon$. The latent variable m and the noise ϵ are assumed to follow gaussian distributions, such that $m \sim \mathcal{N}(0, I_d)$ (I_d being the identity matrix of size $d \times d$) and $\epsilon \sim \mathcal{N}(0, \sigma^2 I_{n_c})$ (I_{n_c} being the identity matrix of size $n_c \times n_c$). As a result, x also follows a gaussian distribution[13] defined as $x \sim \mathcal{N}(\overline{x}, BB^T + \sigma^2 I_{n_c})$.

The parameters B, $\overline{x}$ and σ can be chosen so as to maximize the likelihood.[13] Noting $C = BB^T + \sigma^2 I_{n_c}$ and Σ the covariance matrix of X, the log-likelihood can be written as :

$$\begin{aligned} ln(p(X \mid \overline{x}, B, \sigma^2)) &= -\frac{n_r}{2} ln(2\pi) - \frac{n_r}{2} ln|C| - \frac{1}{2} \sum_{i=1}^{n_r} (x - \overline{x})^T C^{-1} (x - \overline{x}) \\ &= -\frac{n_r}{2} ln(2\pi) - \frac{n_r}{2} ln|C| - Tr(C^{-1}\Sigma) \end{aligned} \tag{3}$$

where $Tr(A)$ is the trace of matrix A.

The maximum likelihood estimation of the parameters is given by $B = U_d(\Lambda_d - \sigma^2 I)^{\frac{1}{2}}$, $\sigma^2 = \frac{1}{n_c - d} \sum_{j=d+1}^{n_c} \lambda_j$, $\hat{m}_i = (\Lambda_d - \sigma^2 I)^{\frac{1}{2}} U_d^T (x - \overline{x})$, where $U_d \in \mathbb{R}^{n_c \times d}$ contains the first d eigenvectors of the covariance matrix of X, sorted by decreasing eigenvalues, and Λ_d contains the corresponding eigenvalues λ_j.

2.1.3 Handling observed and missing landmarks

Although all coordinates are known for all shapes in the training set, at test time only a subset of landmarks detected in the images, x_o are observed. The other landmarks are estimated using the statistical shape model, so as to minimize the distance on the observed landmarks. However, because the training set is not fully representative of the test set, the desired shape, using the observed landmarks, might not be perfectly represented by a linear combination of shapes in the training set. Following 12, we represent this variability by introducing a second slack variable ϵ_2 concerning only the observed landmarks: $x_o = \overline{x}_o + B_o m + \epsilon_o + \epsilon_2$, where B_o contains the rows of B corresponding to the observed entries, and ϵ_o contains the corresponding elements of ϵ.

The slack variable ϵ_2 also follows a gaussian distribution $\epsilon_2 \sim \mathcal{N}(0, \sigma_2^2 I_{n_c})$, therefore $\xi = \epsilon_o + \epsilon_2$ follows a gaussian distribution defined by $\xi \sim \mathcal{N}(0, \sigma_\xi^2 I_{n_c})$, with $\sigma_\xi = \sigma + \sigma_2$. Therefore, $x_o \mid m \sim \mathcal{N}(\overline{x}_o + B_o m, \sigma_\xi^2 I_q)$.

We note $M = B_o^T B_o + \sigma_\xi^2 I$. Following 14, the conditional probability distribution of the latent variable m knowing x_o is $m \mid x_o \sim \mathcal{N}(M^{-1} B_o^T (x_o - \overline{x}_o), \sigma^2 M^{-1})$. The latent vector maximising this distribution is therefore

$$\hat{m} = (B_o^T B_o + \sigma_\xi^2 I)^{-1} B_o^T (x_o - \overline{x}_o) \tag{4}$$

This solution is equivalent to the one proposed in 1 and described in equation 2, with $\beta = \sigma_\xi^2$.

2.1.4 Interpretation

Two slack variables are introduced in this model. The first one, ϵ, accounts for the loss of information due to the use of the d first eigenvectors only. The second one, ϵ_2, accounts for the representation of shapes that do not belong to the training set. The combination of these slack variables explain the addition of the regularization term in equation 1.

2.2 Data set

Two data sets are used to perform two sets of experiments, described in section 2.3. Firstly, the AIS (adolescent idiopatic scoliosis) data set is composed of 768 AIS patients from the Sainte-Justine mother and child university hospital center, affiliated with the University of Montréal in Canada. Secondly, the mixed data set combines adults and AIS patients in a training set of 1536 patients and an evaluation set of 210 patients. All data sets were retrospectively collected after ethical approvals from ethical committees of the Montreal University Hospital Center (CHUM) and the École de technologie supérieure (ÉTS), Montréal, Canada. They contain two low dose x-ray images (in sagittal and frontal views) acquired with the EOS system.[15] 3D spine reconstruction was

performed for each pair of images using the semi-automated method described in 16. For the evaluation set, reconstruction was performed by an expert before being corrected twice. The 3D spine is represented using a parametric spine model,[1] resulting in 1124 parameters including the 3D position of the pedicles, endplates and centers of each vertebra. As in 1 the origin is the middle of the two femoral head, the z axis is the axis joining the origin to the topmost vertebral center, the y axis is the axis joining the origin to the left femoral head and the x axis is perpendicular to the z and y axis and and pointing towards the front of the subject.

2.3 Experiments

Two sets of experiments were performed. In a first set of experiments, we evaluate the impact of clipping in a controlled simulated setting, by using the SSM to correct a simulated noise on vertebral centers. In a second set of experiments we evaluate the impact of clipping the SSM in a real pipeline, which aims at reconstructing a 3D model of the spine from 2D images. Both sets of experiments are described in more details bellow.

For all experiments, the SSM was evaluated by studying the 3D distance between estimated landmarks and true values for vertebral centers, pedicles and endplate centers. When PCA is performed, 99.999% of explained variance is kept.

2.3.1 Using simulated errors

The first set of experiments aims at evaluating the capacity of the SSM to correct simulated errors on vertebral centers by regularizing the global spine shape. This set of experiments was performed on the AIS data set by adding a simulated noise on vertebral centers for all vertebrae.

The noise added to the vertebral centers used as predictors was simulated using a gaussian mixture of two components, representing a standard gaussian noise (standard deviation of 3 mm across all vertebral levels) applied to all coordinates, and an aberrant noise (20 mm with a probability of $p = \frac{2}{46}$) on the x and y components only. As sagittal and frontal views are combined, detection error in z rarely happens,[1] so no aberrant noise was added on the z component. The resulting probability function for the simulated noise of the x and y components is:

$$P(e) = (1-p)\mathcal{N}(0,\sigma) + \frac{p}{2}\mathcal{N}(20,\sigma) + \frac{p}{2}\mathcal{N}(-20,\sigma) \tag{5}$$

As different errors can be expected for each coordinate and vertebra levels, different σ are used. The standard deviation were computed so that they are proportional to the global root mean square error reported in 1, with a standard deviation of 3 mm across all axes and vertebral levels. The combination of the standard and aberrant errors leads to an average 3D residual of 4.9 mm (standard deviation of 5 mm).

The noisy vertebral centers were the only predictors in the SSM, which was used to estimate the other spine parameters (providing pedicle and endplate landmarks) and denoise the vertebral centers. Different SSMs were considered, by varying the value of the regularization weight β between $\beta = 10^{-5}$ and $\beta = 1.0$, with and without clipping. 10-fold cross-validation was used to get an estimation on the whole data set.

2.3.2 Using the reconstruction pipeline

The second set of experiments was performed in actual conditions, on the mixed data set of adolescent and adult patients, using the automatic reconstruction method of the 3D spine from a pair of x-ray images proposed in 1.

This method combines the detection of vertebral landmarks (vertebral centers in a first stage, pedicles and endplate centers in a second stage), performed iteratively for each vertebra, with a correction by the spine SSM, performed after each vertebra detection. The whole spine is initialized and updated after each vertebra detection using the SSM. The landmark detection is performed using a convolutional neural network (CNN), which computes the optimal landmark displacements given the previous positions. In this method, the purpose of the SSM is two-fold: firstly, the SSM gives the position of landmarks which have not yet been detected and the value of spine parameters which are not computed from the image; secondly, the SSM corrects the position of landmarks that have already been detected, taking into account the global spine shape so as to output realistic spines. Four sets of parameters were considered: SSM with clipping, using $\beta = 0.01$ (original method[1]) and $\beta = 0.2$; and SSM without clipping, using $\beta = 0.01$ and $\beta = 0.1$. The full pipeline was trained on the training set of 1536 subjects, and the evaluation of 210 subjects was used to evaluate the SSM.

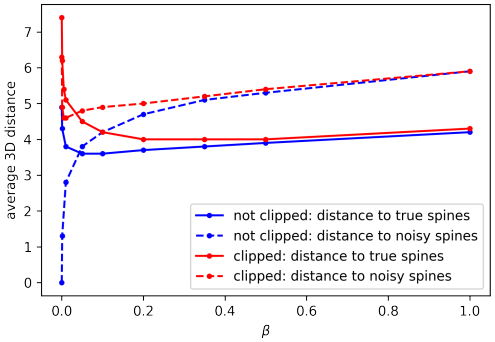

Figure 1: Effect of the regularization weight β, with and without clipping, on the AIS data set with simulated errors. 3D distance (mm) is averaged over all vertebral centers over 768 patients after cross-validation.

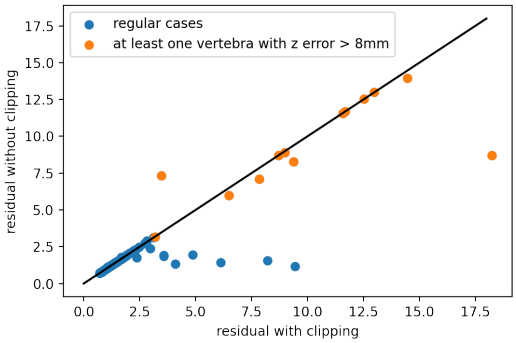

Figure 2: Mean 3D residual (mm) across vertebral centers for each of the 210 subjects in the mixed evaluation set, with and without clipping, ($\beta = 0.01$). The diagonal line represents equality with and without clipping, points bellow the line represent subjects for which the 3D residual is lower without clipping.

3. RESULTS

3.1 Using simulated error

Parameters	β	0.01	0.1	0.01	0.2
	clipping	no	no	yes	yes
3D residuals (mm): mean (std)	vertebral center	3.8 (3.3)	3.6 (2.7)	5.1 (5.5)	4 (3.3)
	superior endplate	4.1 (3.5)	3.8 (2.9)	5.3 (5.6)	4.1 (3.4)
	inferior endplate	4.4 (3.9)	4 (3.2)	5.5 (5.5)	4.2 (3.4)
	left pedicle	5.9 (4)	4.9 (3.3)	6.5 (5.4)	4.9 (3.5)
	right pedicle	6 (4.2)	4.9 (3.4)	6.4 (5.2)	4.8 (3.4)

Table 1: Simulated noise: 3D distance (in mm) between output and true spine on several vertebra landmarks, across all subjects and vertebrae, for different regularization weights β, with and without clipping. Measures are given as mean (standard deviation).

By varying the regularization weight β in the SSM model, with and without clipping, we can better understand the effect of regularization and clipping as well as the interaction between them. Figure 1 shows the impact of these parameters on the distance between the outputs and the noisy spines, used as predictors for the SSM, as well as the distance between the outputs and the true spines, which should ideally be minimized.

First looking at the results without clipping (in blue), we can see that a regularization weight close to 0 leads to no noise correction, resulting in an output equal to the noisy input and a large output to true spine distance. As β increases, the outputs get closer to the true spines and farther from the noisy ones, showing the expected noise correction. The distance to true spine reaches a minimum for a regularization weight of 0.05. Increasing β above this value leads to a slight increase in the distance to true spine, showing that the regularization results in a loss of information as well a noise correction. The improvement on noise correction is however still greater than the loss of information, until larger values of β (around 0.3).
Secondly, looking at the results with clipping (in red), we can see that a regularization weight close to 0 leads to a large distance both to true and noisy spines, showing that shape information is not well represented. As β increases up to 0.01, the output gets closer to the true spine, but the distance to noisy spine also decreases. Clipping prevents the representation of global shapes in the test set by thresholding some latent coordinates. As β increases, high values are penalized, encouraging spreading across all modes, limiting the impact of thresholding the latent representation and hence of clipping. For values higher than 0.01, increasing β leads to effects similar

to those observed without clipping, first removing noise, then removing shape information as well when β is too large. The loss of shape information however happens at 0.05, which is larger than without clipping, probably because increasing regularization also mitigates the negative impacts of clipping more strongly. The distance to true input obtained with clipping is however never smaller than the distance to true input obtained without clipping.

Table 1 shows the distance between the SSM output and true spine for vertebral centers, pedicles and endplate centers, with and without clipping, for different values of β. Together with figure 1, these results show that clipping has a negative impact, regardless of the choice of regularization weight, as the model without clipping removes the simulated noise (albeit not completely), while loosing less information, leading to outputs closer to the true spines for all β values.

The choice of β highly depends on the noise magnitude, which was arbitrarily set here. Values from 0.05 to 0.2 seem reasonable choices, both with and without clipping, although validating this parameter properly would require access to the true noisy inputs, corresponding to the vertebral center detection in the images by the CNN. This information is not available, but experiments using the reconstruction pipeline allow more insight.

3.2 Using the reconstruction pipeline

Parameters	β	0.01	0.1	0.01	0.2
	clipping	no	no	yes	yes
3D residuals (mm): mean (std)	vertebral center	1.9 (3.2)	2.3 (3.1)	2.1 (3.7)	2.8 (3.7)
	superior endplate	2.1 (3.2)	2.5 (3.1)	2.2 (3.6)	2.9 (3.7)
	inferior endplate	2.1 (3.3)	2.5 (3.2)	2.3 (3.7)	2.9 (3.8)
	left pedicle	2.7 (3.2)	3.2 (3.1)	2.9 (3.5)	3.6 (3.6)
	right pedicle	2.7 (3.1)	3.1 (3.1)	2.9 (3.5)	3.6 (3.6)

Table 2: Reconstruction pipeline: 3D distance (in mm) between output and true spine on several vertebra landmarks, across all subjects and vertebrae, for different regularization weights β, with and without clipping. Measures are given as mean (standard deviation).

Table 2 shows the 3D residuals obtained on vertebra landmarks, with and without clipping, for different β values. For $\beta = 0.01$, the table shows that the results are better on average without clipping (1.9 mm for vertebral centers, standard deviation of 3.2 mm) than with clipping (2.1 mm, standard deviation of 3.7 mm), for all studied landmarks.

Although the average improvement is small, study of 3D residuals for each subject shows the impact of clipping more clearly. Figure 2 shows the average 3D residual across vertebral centers for each subject, with and without clipping, for $\beta = 0.01$. Most points are on or close to the diagonal line, representing subjects for whom the 3D residual is similar with and without clipping. Subjects lying exactly on this diagonal are subjects for which clipping has no effect, because no coordinates are saturated in the latent representation. Subjects at the bottom of the figure, bellow the diagonal, are subjects for whom the SSM without clipping yields better results than with clipping. These are subjects with unusual spine shapes, for which some coordinates are saturated. Figure 3 shows an example of such subjects. Figure 2 also shows about 12 subjects who have high residuals with clipping and for whom the residual stays high when removing clipping. Further examination of these subjects shows that the computed global spine curve is right, but the vertebrae are not well placed along the spinal curve. In particular, some vertebrae seem to not be detected at all, leading to a shift in a series of vertebrae which are identified as the next one. An example is provided in Figure 4. Subjects with at least one vertebral center with a z absolute error of more than 8.6 mm (minimum difference in the z coordinate between 2 adjacent vertebrae in the data set) are colored in orange in Figure 2. This figure shows that all the subjects with a high error that is not corrected when removing clipping belong to this category. Notably, the only subject with an error smaller with clipping than without belongs to this category. This issue is probably due to an error in the CNN detection which is not modeled by our simulated noise. Given that all the subjects that still have a high error without clipping belong to this category, this issue seems of particular importance and requires further study.

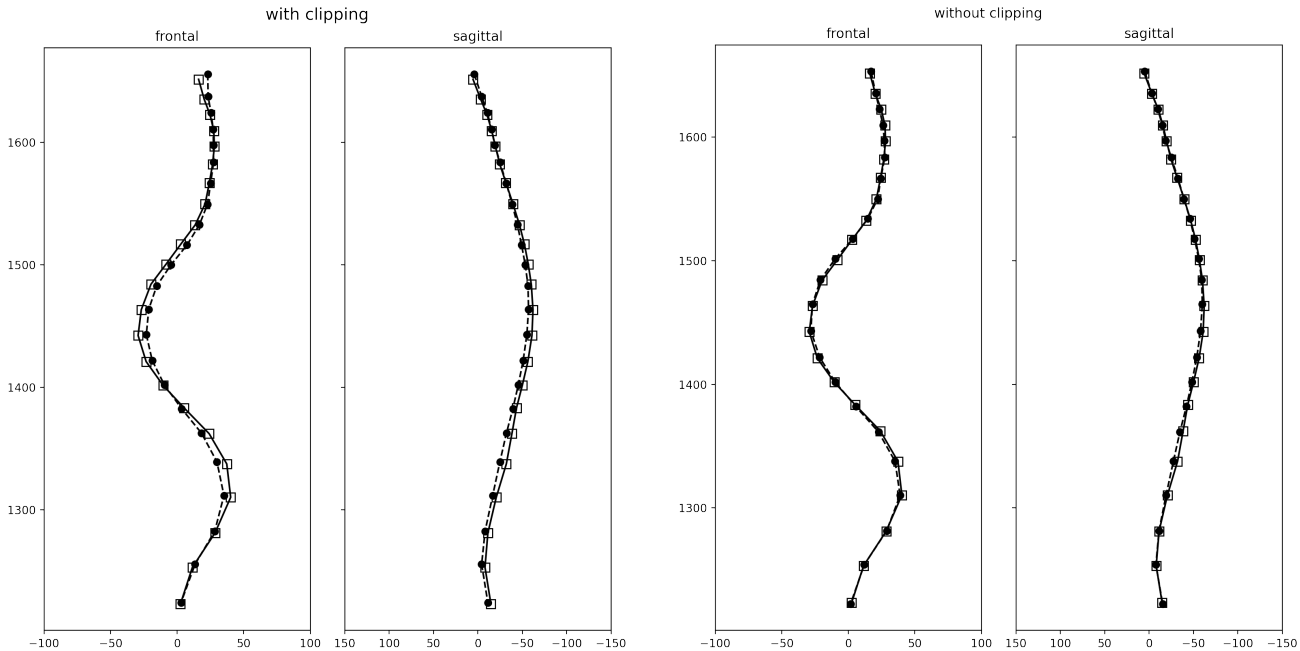

Figure 3: Example of a subject for which the reconstruction is worse using clipping (left, average 3D residual across all vertebral centers of 4.9 mm) than without clipping (right, average 3D residual across all vertebral centers of 2.9 mm) for $\beta = 0.01$. For the true spine vertebral centers are displayed using squares linked by dashed lines. For the output vertebral centers are displayed using dots linked by plain lines.

Regarding the regularization weight, Table 2 shows that increasing β leads to worse results, with clipping (3D residual of 2.1 mm (std=3.7) for $\beta = 0.01$ and of 2.8 mm (std=3.7) for $\beta = 0.2$) and without (3D residual of 1.9 mm (std=3.2) for $\beta = 0.01$ and of 2.3 mm (std=3.1) for $\beta = 0.1$).

4. DISCUSSION AND CONCLUSION

We studied the impact of clipping in a SSM fit process from constraints that were simulated or detected in actual X-ray images. Both sets of experiments show that clipping limits model performance, by not allowing spine shapes that are not well represented in the training set, and therefore making the model depend heavily on the choice of training set. The regularizing effect of clipping, which corrects landmark detection errors, is not necessary when using a regularization term in the objective function used to fit the SSM on partial data. This regularization term penalizes deviation from the mean shape, attributing a higher cost to shapes farther from the average shape. We showed that it is equivalent to using a slack variable accounting for the non-representativity of the training set, which attributes a higher cost to shapes not represented in the training set without making the reconstruction of these shapes impossible. We showed that this regularization term lessens the negative impact of clipping, but it also makes clipping unnecessary by correcting landmark detection errors. Because using the regularized SSM model also allows the reconstruction of out-of-distribution spine shapes, resulting in better reconstruction performance, it should be used instead of and without clipping. It is to be noted that these conclusions are only valid for fully automatic reconstruction. When the user is allowed to edit the model by manually moving landmarks, clipping restricts manual modifications to valid shapes. If clipping is removed, another method for limiting user modifications has to be used.

A downside of the addition of the regularization term is the introduction of the new hyper-parameter, β, which depends on the input errors that the SSM should remove, and should therefore be validated for each pipeline and data set. Although regarding clipping the same conclusions can be drawn from the experiments performed on simulated errors and on the reconstruction pipeline, the same cannot be said for the choice of the regularization weight. This discrepancy can have several explanations. It can first be a sign of the sensitivity of the hyper-parameter, which could highly depend on the context and data. Secondly, the simulated error may not fully

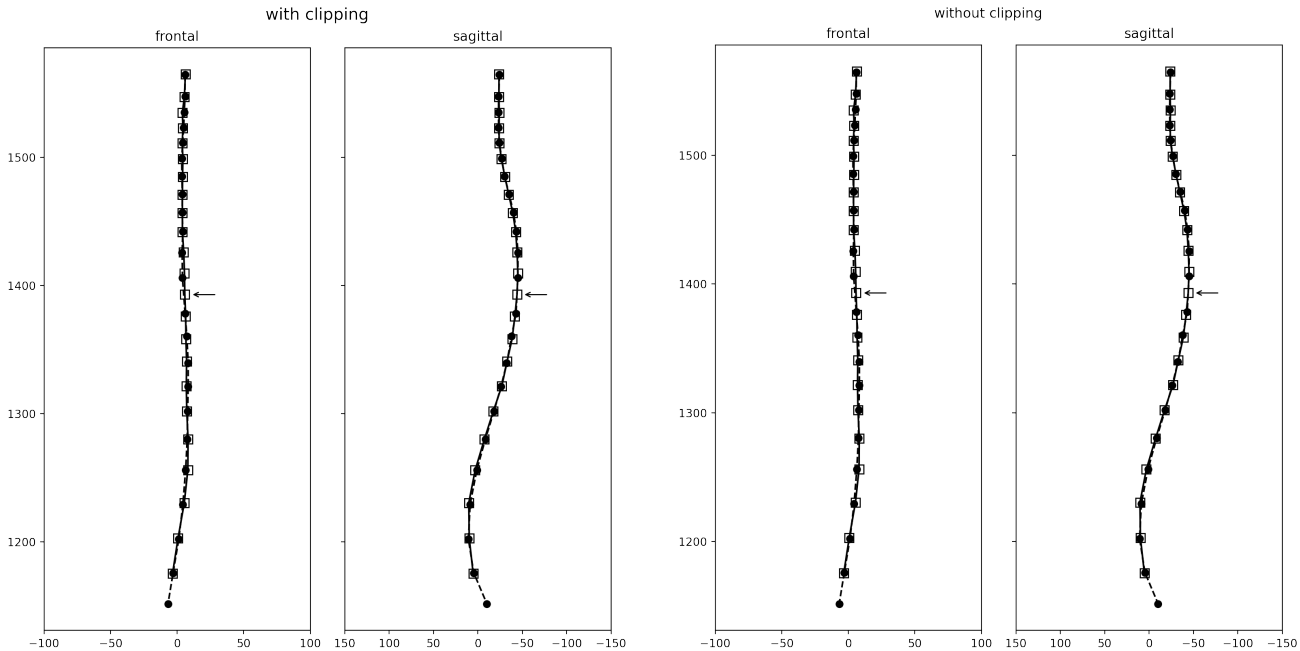

Figure 4: Example of a subject with a shift in the vertebra labelling, with clipping (left) and without clipping (right) for $\beta = 0.01$ (average 3D residual across all vertebral centers of 11.7 mm for both). For the true spine vertebral centers are displayed using squares linked by dashed lines. For the output vertebral centers are displayed using dots linked by plain lines. The shift is highlighted using arrows.

represent the detection errors made by the CNN. Improving error simulation would require more information on these detection errors. Lastly, besides this difference between real and simulated errors, the experimental conditions are very different: in the experiments with simulated errors the SSM is only used once, using all noisy vertebral centers as input. By contrast, in the reconstruction pipeline, the SSM is applied iteratively, after detection of each new vertebral center, and the process is repeated for pedicles and endplate centers after all vertebral centers are detected. The use of these landmarks in addition to vertebral centers during the second stage may also explain the lower error in the reconstruction pipeline compared to the experiments on simulated detection errors.

ACKNOWLEDGMENTS

We thank Dr Christine Bakhous and Dr Manuela Pacheco from EOS Imaging Inc. for performing some of the experiments and providing their expertise regarding the existing models. This work is supported in part by the Collaborative Research and Development program from NSERC and in part by EOS Imaging Inc.

REFERENCES

[1] Aubert, B., Vazquez, C., Cresson, T., Parent, S., and de Guise, J. A., "Toward Automated 3D Spine Reconstruction from Biplanar Radiographs Using CNN for Statistical Spine Model Fitting," *IEEE Transactions on Medical Imaging* **38**, 2796–2806 (Dec. 2019).

[2] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J., "Training Models of Shape from Sets of Examples," in [*BMVC92*], Hogg, D. and Boyle, R., eds., 9–18, Springer, London (1992).

[3] Whitmarsh, T., Humbert, L., Del Río Barquero, L. M., Di Gregorio, S., and Frangi, A. F., "3D reconstruction of the lumbar vertebrae from anteroposterior and lateral dual-energy X-ray absorptiometry," *Medical Image Analysis* **17**, 475–487 (May 2013).

[4] Khallaghi, S., Mousavi, P., Gong, R. H., Gill, S., Boisvert, J., Fichtinger, G., Pichora, D., Borschneck, D., and Abolmaesumi, P., "Registration of a Statistical Shape Model of the Lumbar Spine to 3D Ultrasound

Images," in [*Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*], Jiang, T., Navab, N., Pluim, J. P. W., and Viergever, M. A., eds., *Lecture Notes in Computer Science*, 68–75, Springer, Berlin, Heidelberg (2010).

[5] Kirschner, M., Becker, M., and Wesarg, S., "3D Active Shape Model Segmentation with Nonlinear Shape Priors," in [*Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011*], Fichtinger, G., Martel, A., and Peters, T., eds., *Lecture Notes in Computer Science*, 492–499, Springer, Berlin, Heidelberg (2011).

[6] Korez, R., Aubert, B., Cresson, T., Parent, S., Vrtovec, T., de Guise, J., and Kadoury, S., "Sparse and multi-object pose+shape modeling of the three-dimensional scoliotic spine," in [*2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*], 225–228 (Apr. 2016). ISSN: 1945-8452.

[7] Boisvert, J., Cheriet, F., Pennec, X., and Ayache, N., "3D reconstruction of the human spine from radiograph(s) using a multi-body statistical model," 72612D (Feb. 2009).

[8] Benameur, S., Mignotte, M., Destrempes, F., and DeGuise, J., "Three-Dimensional Biplanar Reconstruction of Scoliotic Rib Cage Using the Estimation of a Mixture of Probabilistic Prior Models," *IEEE Transactions on Biomedical Engineering* **52**, 1713–1728 (Oct. 2005).

[9] Lecron, F., Boisvert, J., Mahmoudi, S., Labelle, H., and Benjelloun, M., "Fast 3D Spine Reconstruction of Postoperative Patients Using a Multilevel Statistical Model," in [*Medical Image Computing and Computer-Assisted Intervention – MICCAI 2012*], Ayache, N., Delingette, H., Golland, P., and Mori, K., eds., *Lecture Notes in Computer Science*, 446–453, Springer, Berlin, Heidelberg (2012).

[10] Heimann, T. and Meinzer, H.-P., "Statistical shape models for 3D medical image segmentation: A review," *Medical Image Analysis* **13**, 543–563 (Aug. 2009).

[11] Hoerl, A. E. and Kennard, R. W., "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics* **12**, 55–67 (Feb. 1970).

[12] Albrecht, T., Lüthi, M., Gerig, T., and Vetter, T., "Posterior shape models," *Medical Image Analysis* **17**, 959–973 (Dec. 2013).

[13] Tipping, M. E. and Bishop, C. M., "Probabilistic Principal Component Analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**(3), 611–622 (1999). _eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/1467-9868.00196.

[14] Bishop, C. M., [*Pattern recognition and machine learning*], Information science and statistics, Springer, New York (2006).

[15] Deschênes, S., Charron, G., Beaudoin, G., Labelle, H., Dubois, J., Miron, M.-C., and Parent, S., "Diagnostic imaging of spinal deformities: reducing patients radiation dose with a new slot-scanning X-ray imager," *Spine* **35**, 989–994 (Apr. 2010).

[16] Humbert, L., De Guise, J. A., Aubert, B., Godbout, B., and Skalli, W., "3D reconstruction of the spine from biplanar X-rays using parametric models based on transversal and longitudinal inferences," *Medical Engineering & Physics* **31**, 681–687 (July 2009).