# Automatic Spine Curvature Estimation from X-ray Images of a Mouse Model

Omar Al Okashi, Hongbo Du and Hisham Al-Assam
Department of Applied Computing
University of Buckingham,
Buckingham, MK18 1EG, UK
{hisham.al-assam@buckingham.ac.uk}

## Abstract

Automatic segmentation and quantification of skeletal structures has a variety of applications for biological research. Although solutions for good quality X-ray images of human skeletal structures are in existence in recent years, automatic solutions working on poor quality X-ray images of mice are rare. This paper proposes a fully automatic solution for spine segmentation and curvature quantification from X-ray images of mice. The proposed solution consists of three stages, namely preparation of the region of interest, spine segmentation, and spine curvature quantification, aiming to overcome technical difficulties in processing the X-ray images. We examined six different automatic measurements for quantifying the spine curvature through tests on a sample data set of 100 images. The experimental results show that some of the automatic measures are very close to and consistent with the best manual measurement results by annotators. The test results also demonstrate the effectiveness of the curvature quantification produced by the proposed solution in distinguishing abnormally shaped spines from the normal ones with accuracy up to 98.6%.

**Keywords —** Spine, X-ray, segmentation, curvature, classification

## 1   Introduction

The Mouse Genetics Project (White, et al., 2013) was initiated to systematically knock-out mammalian genes and screen for a broad range of resulting traits. One element of the phenotyping pipeline is the thorough evaluation of the skeleton through systematic X-ray imaging of individual animals and manual observations of variation. Considering the number of images generated by the high-throughput screening process, an automated approach for annotating and triaging X-ray images will reduce processing time and costs on human resources, and eliminate possibilities of human error and inconsistency in manual measurement.

Accurate identification of any perturbation in skeletal structure is of a significant biological relevance. Spinal curvature, for example, is one of important criteria for spine abnormality that associated with different types of spine diseases (Lonstein, 1999). The quantitative evaluation of the

curvature serves as the first step toward identifying spinal diseases like scoliosis, a back condition that causes the spine severely curving towards the left or the right side (Kenny, 2012). Measurement of scoliosis is usually extracted from the dorso-ventral view of medical images (Anderson, 2007). Tomaž et al (Tomaž, et al., 2009) reviewed five main proposals in quantifying the spinal curvature either as an angle or as an index from 2D dorso-ventral X-ray images ( (Cobb, 1948), (Diab, et al., 1995), (Chen, et al., 2007), (Ferguson, 1930) and (Greenspan, et al., n.d.)), as illustrated in Fig.1. With all the five methods being applied to the same spine model, all the angle-based measures as shown in Fig. 1A, 1B and 1C are fundamentally the same, and the angle-based measure illustrated in Fig.1D is in fact the complement of the other angle-based measures. The difference lies in the method of computation. Greenspan et al (Greenspan, et al., n.d.) proposed a measure of the spine curvature as an index rather than an angle as shown in Fig. 1E. Among all five proposed measures, Cobb angle was formally adopted as the *de facto* standard quantification of scoliosis by Scoliotic Research Society (SRS) in 1966 (Tomaž, et al., 2009).
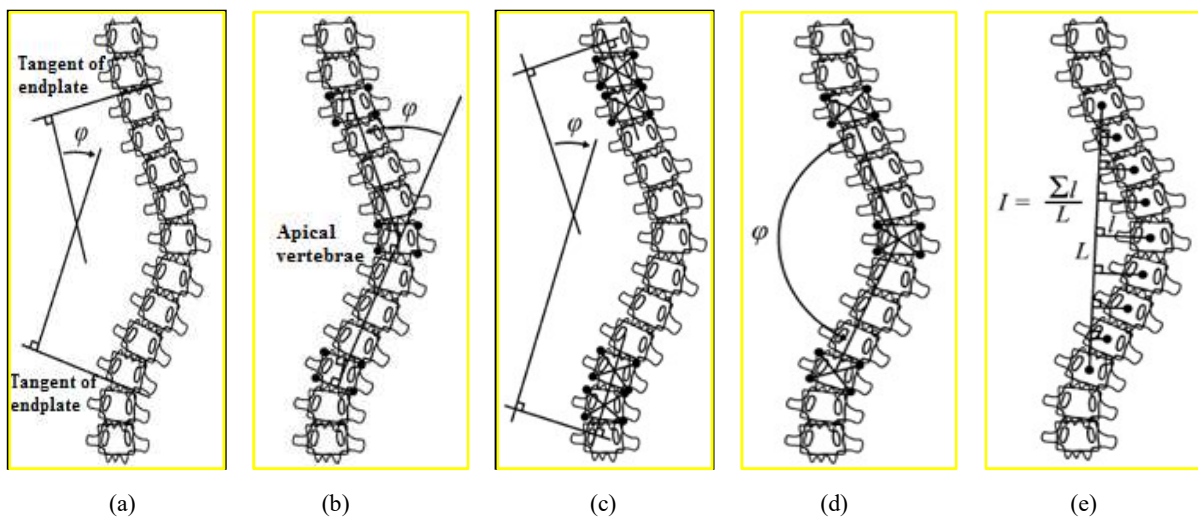


|     |     |     |     |     |
| --- | --- | --- | --- | --- |
| (a) | (b) | (c) | (d) | (e) |

Fig. 1 Methods of spinal curvature estimation. Adapted from (Tomaž, et al., 2009).

(a) Cobb method: estimating the curvature as the angle $\varphi$ between the two tangents to the upper and lower endplates of the upper and lower end vertebra respectively; (b) Diab et al method: estimating the curvature as the angle $\varphi$ that is formed between the intersection of lines extended from both end vertebrae and intersect at the centre of apical vertebrae. The centre of each vertebra is determined by the intersection of lines orthogonal to the upper and lower endplates; (c) Centroid method: estimating the curvature as the angle $\varphi$ between the lines passed through the two upper and through the two lower vertebral centroids. The centre of each vertebra (centroid) was found as the intersection point of two lines extended between the opposite corners of the vertebrae; (d) Fergusson method: estimating the curvature as the angle $\varphi$ between the two lines that extended from centres of the end vertebrae to the centre of the apical vertebra. (e) Greenspan index: estimating the curvature as the index ($I = \frac{\Sigma l}{L}$) where $L$ is the length of vertical line extended between centres of upper and lower end vertebrae and $l$ is the length of each orthogonal line between the centre of each vertebrae in spine curve and the vertical line.

Using any methods for accurate quantification of spine curvature manually, as depicted in Fig. 1, requires accurate locations of end and apical vertebrae, their centres and/or endplates. The small size nature of mouse's spine compared to human's one affects image quality-ray images generated

by the MGP tend to show unclear and blurry borders of the vertebrae, which makes it difficult even for domain experts to precisely locate the borders. This means that manual measurement of spine curvature will encounter the problem known as *inter-observer* and *intra-observer* variations: the measurement results may be different when the same spine is measured by different observers or by the same observer at different times of measurement. An automated computer-based solution will overcome this problem by always giving a deterministic result. However, such solutions still face difficulties in accurate segmentation of vertebrae and identification of the end and apical vertebrae. Finding a novel solution to overcome the difficulties is of high importance.

Several studies in developing automatic and semi-automatic solutions for spine curvature quantification in human X-ray images have been conducted, including vertebra segmentation through masking followed by locating vertebra centre or endplate (Samuvel, et al., 2012), using Hough transform to detect orientations of end vertebrae (Zhang, et al., 2009) or slopes of the spine (Kundu, et al., 2012), or curve fitting the spine with Charged Particle Model (Sardjono, et al., 2013), before finding the best fitting central spine curve line and then the Cobb angle or an index is calculated . However, the existing solutions suffer from the following problems. First, X-ray images of human spine structures tend to be of a better quality than the X-ray images of mouse spine structures obtained from the MGP, and hence many of the existing solutions may not be applicable in coping with the poor quality images. We believe that the main cause of the problem is the small size of the spine mouse that makes the clarity of vertebrae and endplate very poor, especially in the upper part of the chest area. Fig. 2 shows an example X-ray image of a mouse spine structure (Fig. 2A) in comparison with that of the human spine structure (Fig. 2B). The red boxes zoom into the upper parts of the chest areas of mouse and human respectively. It is clear that the level of contrast between the vertebrae and the backgrounds for the mouse image is much lower than that for the human image. Second, most of the existing solutions are semi-automatic with a certain degree of manual intervention at the start of the process. This type of intervention is not feasible in processing a large number of images in a high throughput situation. Besides, all existing solutions only produce one type of spine curvature measure either as the Cobb angle or as an index.
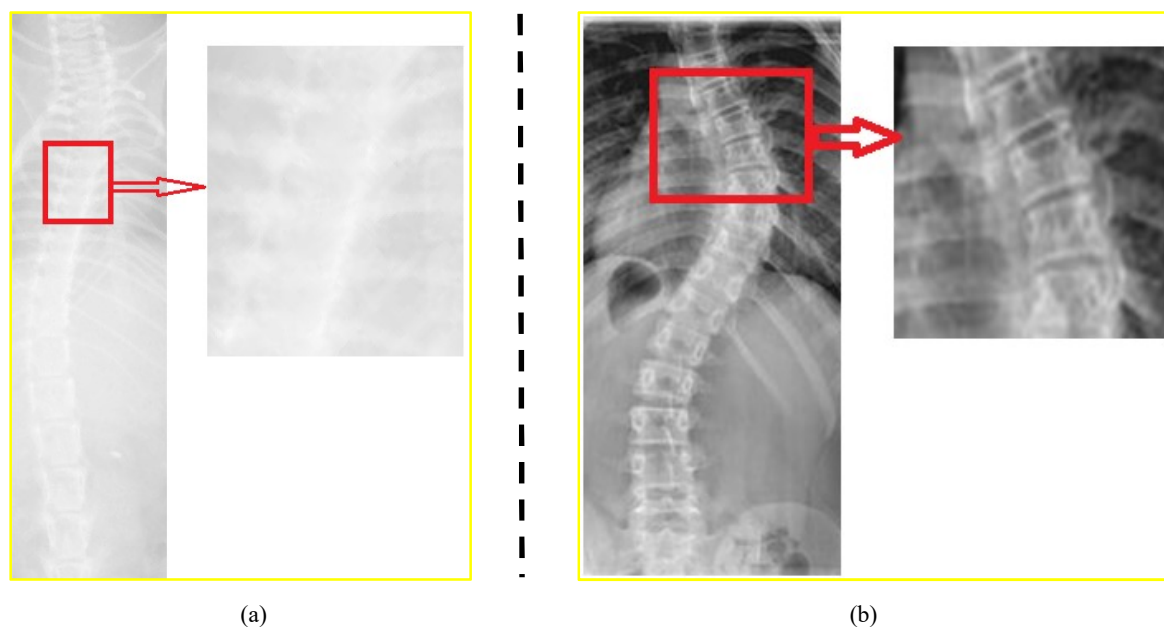
(a)                                    (b)

Fig. 2 Clarity of mouse spine X-ray image versus human spine X-ray image.

(a) Mouse spine X-ray image shows unclear vertebrae and endplate especially in the zoom in area, whereas (b) Human spine X-ray image shows clear vertebrae and endplate.

This paper proposes a fully-automated solution to segment the whole spine in X-ray images of mice and then quantify its curvature with both angle-based and index-based measurements. The fully automatic process aims at reducing time and cost for high throughput situations on one hand, and producing consistent quantification measurements without the inter- and intra-observer variations on the other. To our knowledge, the proposed solution is the first attempt that deals with fully automatic spine segmentation and curvature quantification for mice X-ray images. The automatic solution must meet the following technical challenges. First, the solution must be capable of cropping the Region of Interest (ROI) automatically from full-body input images of mice. Second, it must make accurate segmentation of the spine by overcoming problems related to low contrast within the ROI and blurred spine borders. Finally, the solution must accurately locate important landmarks, i.e. the end and the apical vertebrae for precise curvature quantification. Besides presenting a spine curvature estimation solution, the paper also examines the effectiveness of angle-based and index-based measures in determining normal from abnormal spines. We investigate several different variants of Cobb angles and indices. To evaluate the performance of the proposed solution and various measures of quantification, we conducted several tests on a sample data set of 100 X-ray images obtained from Wellcome Trust Sanger Institute. The test results show that the proposed method produces sensible spine curvature measurements which in turn lead to effective separation of normal and abnormal spines with high level of accuracy.

4

## 2    Materials and Methods

The proposed solution consists of three main steps. The process starts with pre-processing that includes cropping and enhancing the ROI of the spine region. This is followed by segmenting the spine from the ribs and other artefacts within the ROI. The method finally derives angle and index measurement metrics from the segmented spine image. The following sub-sections explain the materials used in this paper and also each step of solutions in more details.

### 2.1    Materials

100 full body X-ray images of mice in the dorso-ventral view were randomly selected (one image per mouse) from a collection of 6000+ images of mice obtained from Wellcome Trust Sanger Institute. Fig.3A shows a sample of dorso-ventral mouse X-ray image. The dataset consists of 75 images of wild type (normal) mice and 25 images of mutant mice with gene knocked-out known to be linked with some form of scoliosis which defined and labelled by domain experts at the Wellcome Trust Sanger Institute. Among the 25 images of mutant mice with scoliosis, there are 22 images of the Thoracic Scoliosis cases, 2 images of the Lumbar Scoliosis cases, and 1 image of Thoracolumbar Scoliosis.

To produce manual Cobb angle measurements for evaluation purposes, the manual process as explained in Section 1 was followed by one expert observer ($A_1$) and two inexperienced observers ($A_2$, $A_3$). All 100 X-ray images were measured for observing inter-observer variations. In addition, $A_1$ performed the measurements 4 times for observing any intra-observer variations. The average of the four measurement results obtained from $A_1$ was then taken as referred manual measurement in comparison with the manual measurements by the other two observers as well as the automatic measurements by the proposed solution.

### 2.2    Pre-Processing

The pre-processing step prepares the image for accurate spine segmentation. This step involves three operations: (a) aligning the mouse skeleton in a given image, (b) cropping the ROI, and (c) to de-noising and enhancing the cropped ROI. It is realised that the mouse may be positioned in different orientations when the X-ray image is taken. For instance, many images from the MGP show placement positions of mice around 45 degrees as shown in Fig. 3A. The alignment operation is intended to position the mouse body to vertical upright in order to obtain accurate cropping of ROI. This is achieved by an initial rotation of the mouse body by certain degrees (e.g. -45 degrees for the MGP images), followed by automatically locating two key points, P and Q as illustrated in Fig. 3B. Using the coordinates of the top point of the mouse's body, P and Q are respectively

estimated at the middle of the head and the middle of the body. Then, the mouse body position is further rotated by the angle between the vertical line passing Q and the line formed by P and Q as shown in Fig. 3B. The final aligned image is shown in Fig. 3C.


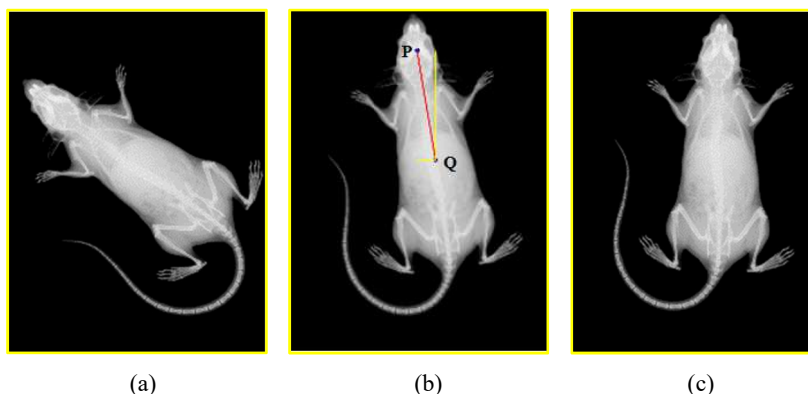
(a)                    (b)                    (c)

Fig. 3 Alignment of mouse body.

(a) Input mouse X-ray image of full body dorso-ventral view before alignment, (b) P and Q key alignment points and the line between them annotated on the initially rotated mouse body, and (c) Mouse body after rotating it by the obtained angle of rotation.

Once the mouse body skeleton is aligned to vertical, we adopt the method developed in our previous work (Al Okashi, et al., 2014) to crop the ROI from the aligned image. The method works as follows. First, a limb-and-skull mask similar to that in Fig. 4A is produced by applying the High Emphasis-Butterworth High Pass Filter followed by the region grow segmentation (Gonzalez & Woods, 2008). After that, five reference points, as highlighted by the yellow circles in Fig. 4A, were used to estimate the ROI as illustrated by a red rectangle in Fig. 4B. Further details of this method can be found in (Al Okashi, et al., 2014).
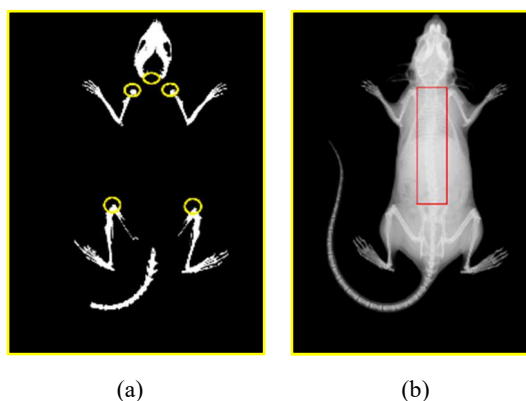


(a)                    (b)

Fig. 4 Cropping of spine ROI from aligned mouse body.

(a) Located reference points annotated on limb-and-skull mask; (b) Borders of spine ROI.

As shown in Fig. 5A, the cropped ROI may have very poor contrast that may affect the accuracy of spine segmentation. To solve this problem, the cropped ROI was further enhanced through the following sequence of operations:

1. Enhancing image contrast by applying the Contrast Limited Adaptive Histogram Equalization (ADHE) (Pratt, 2001) twice on the cropped ROI. The effect of this operation is shown in Fig. 5B.

2. Highlighting high contrast bone structures and eliminating low contrast irrelevant structures in the enhanced ROI resulted from operation 1 by applying the Gamma transformation (Gonzalez & Woods, 2008) as illustrated in Fig. 5C.

3. Further refining the image by applying wavelet-based fusion (Huang, 2010) between the output images of operations 1 and 2 above. The enhanced images from operation 1 tend have non-spine structures and the resulting image from operation 2 tend to have missing parts in the contour of the spine. Wavelet-based fusion can remedy the above problems as shown in Fig. 5D.



<div align="center">(a)        (b)        (c)        (d)</div>
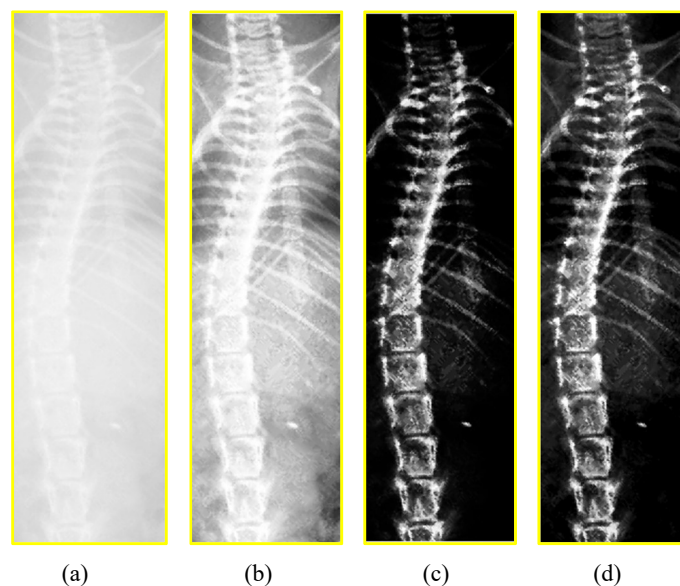
Fig. 5 Enhancing the cropped ROI of spine.

(a) Original cropped spine ROI before enhancements, (b) After applying ADHE, (c) After applying Gamma transformation, and (d) After applying Wavelet fusion between images in (b) and (c).

## 2.3 Spine Segmentation

The goal of spine segmentation is to obtain accurate and smooth outlines of the spine contour. This step involves a two-stage process: initial segmentation and refinement. The initial spine segmentation starts by applying the Otsu method (Otsu, 1979) to the resulting image of the pre-processing step. The Otsu method chooses a threshold to binarise an image by minimizing the intra-class variance of the black and white pixels as shown in Fig. 6A. To produce an initial spine mask and remove irrelevant structures, a series of morphological operations were applied to the binary image. These operations involve closing and opening the binary objects within the image so that pixels for spinal objects were connected and unwanted object structures were removed. The final spine mask is as shown in Fig. 6B.

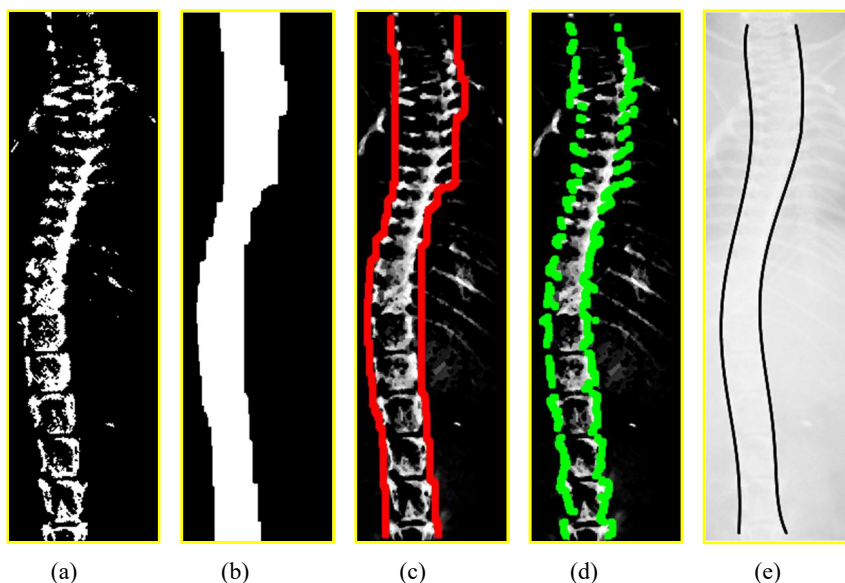|     |     |     |     |     |
|:---:|:---:|:---:|:---:|:---:|
| (a) | (b) | (c) | (d) | (e) |

Fig. 6 Spine segmentation.

(a) Result of thresholding enhanced spine ROI using Otsu method, (b) Spine mask after applying different binary morphological operations, (c) Edges of spine mask annotated on enhanced image of spine ROI, (d) Refined edges of spine using iterative process to find high intensity value on the spine ROI, and (e) Polynomial fitting result for refined edges of spine.

The spine mask image has a very rough estimation of the spine contour due to the thresholding and morphology operations applied at the previous stage. This means that the spine contour boundaries are imprecise, which affects the precision of the spine curvature measurement later. To further refine the spine contour, the following operations are proposed:

1. Minimising the noise around the spine and maximising the contrast between different tissues and structures. To achieve this, we firstly applied the minimum filter to the image in Fig. 5C followed by two grayscale morphology operations tophat and topbot (Gonzalez & Woods, 2008). Fig. 6C shows the further enhanced image.

2. Refining the spine borders. We first superimpose the border pixels of the spine mask in Fig. 6B onto the enhanced image from operation 1. It is clear, as shown in Fig. 6C, that many of those pixels are still not on the actual spine borders as the side-effect of the earlier morphological operations. We therefore refined the locations of the border pixels in the enhanced image from operation 1 by the following iterative process. For every pixel on the left spine border of the mask, we horizontally search towards the right for the first pixel with the maximum intensity value within a predefined distance. If such a pixel is found, it is taken as the new spine border pixel to replace the original border pixel from the spine mask; otherwise the original pixel of the mask is removed (i.e. the pixel intensity value is set to 0). We repeat the process for every border pixel on the right spine border, but instead of searching towards the right, the searching is towards the left. The output of the refinement stage is illustrated in Fig. 6D.

3. Modelling the curvature from the refined pixels' locations. It can be seen from Fig. 6D that there are some outlier points and gaps resulting from the previous operation. To find the best fitting smooth contour of the spine, we applied a polynomial fitting with degree 6 on the refined pixels on the both borders. The final spine contour lines (marked by black lines) are as shown in Fig. 6E.

## 2.4 Spine Curvature Quantification

The final step of the proposed method for quantifying spine curvature is based on two important elements: the Central Spinal Curve (CSC) line representing the actual spine shape and a Spinal Reference Model (SRM) representing a normal straight spine shape line. Our method quantifies spine curvature by comparing the CSC against the SRM. We use both spine borders located in the previous step to estimate the CSC, which compensates for minor variations caused by the left and right borders resulting from imprecise segmentation of spine on low quality images. The estimated CSC was obtained by taking half of the distance between every pair of points on the right and the left borders as illustrated in Fig. 7A.

### 2.4.1 Estimating Linear Spinal Reference Models

We acknowledge that there can be a number of ways of obtaining a SRM and will therefore propose two possible ways of building a linear SRM as follows:

1. $SRM_1$ as the straight line between the top point P and the bottom point Q of the CSC, as shown in Fig. 7B.

2. $SRM_2$ as the best-fitting straight line, $f(x) = a_0 + a_1 x$, among points on the CSC calculated by the least square method as follows:

$$a_1 = \frac{\left(n \sum xy - (\sum x)(\sum y)\right)}{\left(n \sum x^2 - (\sum x)^2\right)} \tag{1}$$

$$a_0 = \frac{\left(\sum y - a_1(\sum x)\right)}{n} \tag{2}$$

where x and y are the coordinates of each point on the CSC. An example $SRM_2$ is shown in Fig. 7C.

While $SRM_1$ seems a nature and simple way to estimate a reference model, it is heavily dependent on the accuracy of locating the two points P and Q. $SRM_2$, on the other hand, relies on all the points on CSC to estimate the reference model and hence should be more stable even the locations of P and Q are not very precise.

## 2.4.2 Locating the Landmark Points

Based on the CSC and the two possible SRMs, our proposed method further identifies five key landmarks as the intersection points between the CSC and the SRMs as follows.

1. The maximum peak (MP) point on the CSC. This is achieved by finding multiple local maximum peak points and further selecting from these points the one that has the maximum distance between the CSC and $SRM_1$, as shown in Fig. 7D.

2. The upper and lower points on $SRM_1$ and $SRM_2$ intersecting with the CSC. Those points were determined as the first intersection points with the CSC above and below the maximum peak point. Fig. 7E shows the intersection points $U_1$ and $L_1$ derived from the $SRM_1$, and the intersection points $U_2$ and $L_2$ derived from the $SRM_2$. There are subtle differences between the locations of intersection points depending on the reference model used.



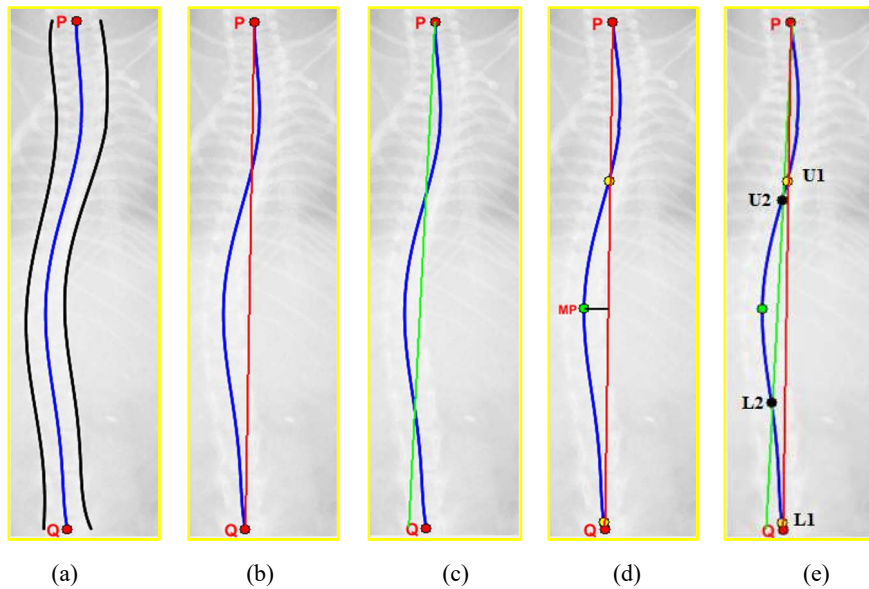(a)       (b)       (c)       (d)       (e)

Fig. 7 Locating CSC, SRMs and landmark points.

(a) Estimated CSC in blue and right and left spine borders in black; (b) SRM1 as the red straight line between the top point P and the bottom point Q of the CSC; (c) SRM2 as the green best-fitting straight line among points on the CSC; (d) Maximum peak point (MP) annotated in green on CSC;(e) Points of intersection between CSC and SRM1 (U1 and L1 yellow points), and between CSC and SRM2 (U2 and L2 black points).

## 2.4.3 Quantifying Spine Curvature

The Cobb angle measurement of curvature replicates the manual process by using the upper and lower intersection points and the maximum peak point as the centres of the start, the end, and the apical vertebrae. Due to the use of two different SRMs, two angle measurements, known as $Angle_1$ and $Angle_2$, can be derived and are illustrated in Fig. 8A and 8B.

Besides the Cobb angle measurements, we also derived distance-based index measurements based on the difference between the CSC and the SRM used. A distance-based index can be obtained by using the Root Mean Square Error (RMSE) as shown below:

$$\text{RMSE} = \sqrt[2]{\frac{\left(\sum_{i=1}^{n}(X_i - XR_i)^2\right)}{n}} \qquad (3)$$

where $X_i$ represents the x-coordinate of the point i on the CSC, and $XR_i$ represents the x-coordinate of the horizontally projected point on the SRM, and n is the number of the points on the CSC. Depending on the chosen SRM and selection of the starting and ending points on the CSC, the RMSE value can be calculated in different ways, which results in a number of index measurements:

- SRM₁ based indices:
  - SRM₁_IndexA: RMSE between the whole CSC and SRM₁, as illustrated in Fig. 8C;
  - SRM₁_IndexB: RMSE between the segment of the CSC bounded by the key points U₁ and L₁, and SRM₁, as illustrated in Fig. 8D;
- SRM₂ based indices:
  - SRM₂_IndexA and SRM₂_IndexB: similar to the above indices but using SRM2 instead of SRM₁.

It can be argued that the accuracies of the angle based measurements are more dependent on the precision of locating the three key landmark points than the index based measurements. This may make the index-based measurements more consistent and stable than angle-based measurements.
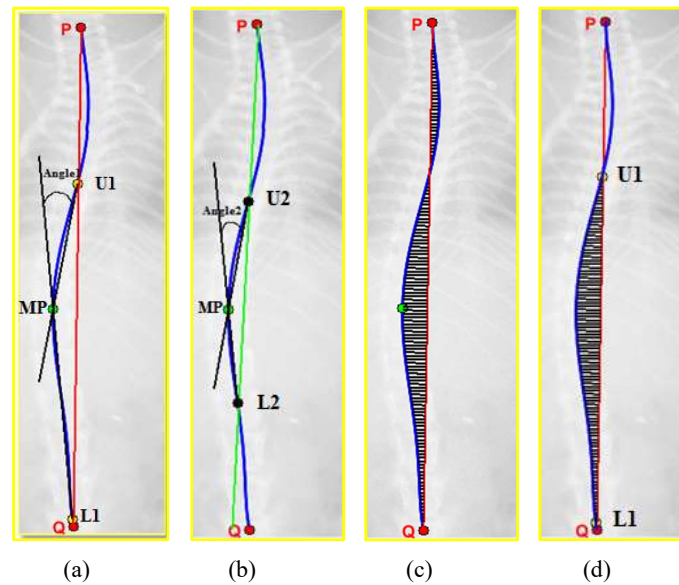


|  (a)  |  (b)  |  (c)  |  (d)  |

Fig. 8 Angle based and distance-based index measurements.

(a) Angle1 as the angle formed between line passes through U1 and MP points and line passes through L1 and MP points; (b) Angle2 as the angle formed between line passes through U2 and MP points and line passes through L2 and MP points; (c) SRM1_IndexA as the RMSE between the whole CSC and SRM1; (d) Comb_IndexA as the RMSE between the segment of the CSC bounded by the key points U1 and L1, and SRM1.

# 3    Results

To test the effectiveness of the proposed spine segmentation and curvature estimation method, two experiments were conducted. The first experiment aimed to evaluate the consistency and closeness of the automated angle-based measurement results compared with the manual measurement obtained by domain experts. The second experiment aimed to evaluate the effectiveness of all the automatic curvature measurements in accurately identifying cases of scoliosis from normal cases by training a simple classifier as detailed in section 3.2.

## 3.1    Manual versus Automatic Measurements

Before we compare the automatic measurements to the manual ones, it is necessary to establish the degree of variations in manual measurements. To examine closeness and correlation between different manual measures taken by $A_1$, $A_2$ and $A_3$ observers (as we described in Section 2.1), we calculated $R^2$, Absolute Error (MAE), and the Angle of Regression Line (ARL). To analyse the variations across different genotypes, Table 1 shows the evaluation results on the whole dataset, mutant only subset, and WT only subset separately.

*Table 1: Intra-Observer and Inter-Observer Variations in Manual Measurements of Spine Curvature Cobb Angles*

| Methods Used (Manual Measurements) | Data Set of All Mice | | | Data set of Mutant Mice | | Data Set of Wild Type Mice | |
|---|---|---|---|---|---|---|---|
| | $R^2$ | ARL | MAE | $R^2$ | ARL | $R^2$ | ARL |
| $A_1$ Trial Average | 0.92 | 44.4 | 1.20 | 0.82 | 39.78 | 0.32 | 34.12 |
| $A_1$ Trial Average vs $A_2$ Trial | 0.86 | 38.95 | 1.60 | 0.57 | 30.70 | 0.03 | 8.23 |
| $A_1$ Trial Average vs $A_3$ Trial | 0.85 | 42.03 | 1.69 | 0.74 | 35.88 | 0.15 | 12.65 |
| $A_2$ vs $A_3$ | 0.79 | 38.28 | 2.13 | 0.40 | 30.28 | 0.14 | 26.72 |

The test results in the table indeed indicate that inter- and intra-observer measurement variations do exist as shown by all the metrics used although most of the measurements from the three observers are highly correlated. As expected, the intra-observer variations are smaller than the inter-observer variations. The inter-observer variations are the highest for wild type mice where the spine should look almost straight and thus manually locating the landmarks would have a greater deal of randomness, particularly for inexperienced observers. There also seems a certain degree of consistent bias where inexperienced observers tend to measure the angles larger than those by the expert observer, specifically for the wild type mice, as indicated by consistently lower ARL values less than 45 degrees.

Although more about inter- and intra-observer variations can be further discussed, our interest lies in the closeness and consistency between the manual and automatic measurements. Again, the

average of the four measurement results obtained from $A_1$ was taken as the referred manual measurement in the comparison with the automatic measurements. Arguably, a good automatic measurement of the spine curvature that is close to the manual measurement results should produce a high value for $R^2$, a low value for MAE, and an ARL value close to 45 degrees. Table 2 shows the comparison results of automatic vs. manual angle-based measurements. Fig. 9 shows the scatter plots and the regression lines for reference manual measure versus $Angle_1$ and $Angle_2$.

*Table 2: Automatic vs. Manual Angle-based Measurements*

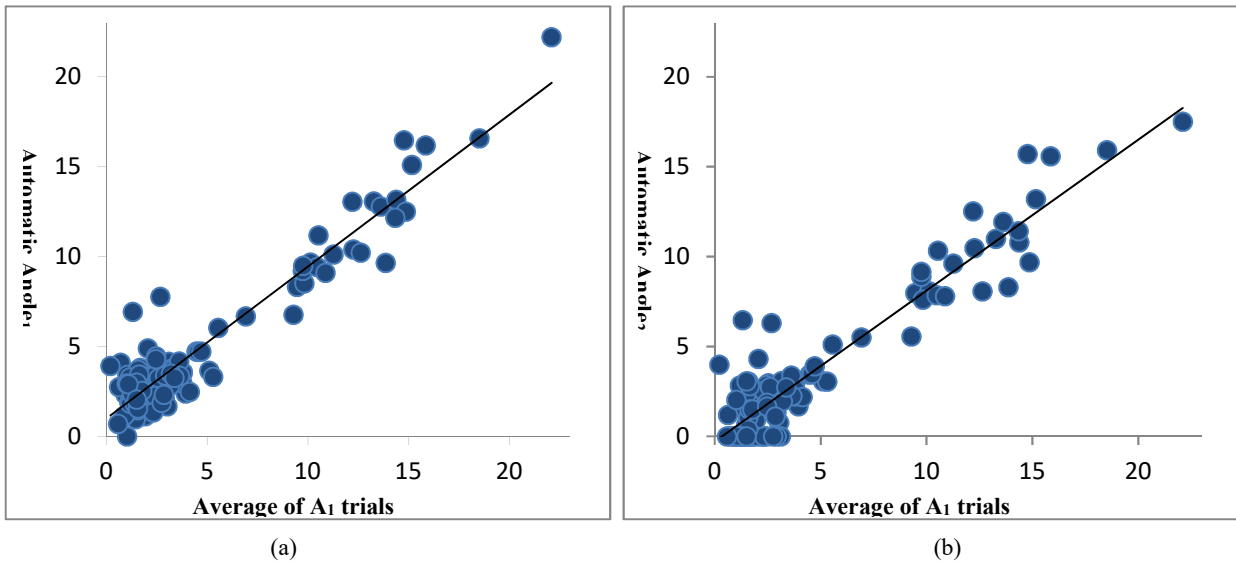| Methods Used (Automatic vs Manual Measurements) | Data Set of All Mice | | | Data set of Mutant Mice | | Data Set of Wild Type Mice | |
|---|---|---|---|---|---|---|---|
| | $R^2$ | ARL | MAE | $R^2$ | ARL | $R^2$ | ARL |
| $Angle_1$ vs. $A_1$ Trial Average | 0.91 | 40.20 | 1.12 | 0.87 | 45.42 | 0.17 | 24.50 |
| $Angle_2$ vs. $A_1$ Trial Average | 0.89 | 40.00 | 1.58 | 0.75 | 40.26 | 0.21 | 31.47 |



Fig. 9 Scatter plot for average of manual measurements of A1 versus Automatic measurement obtained by using Angle1 and Angle2. (a) Average of manual measures taken by A1 versus automatic measurement obtained by using Angle1; (b) Average of manual measures taken by A1 versus automatic measurement obtained by using Angle2.

From the $R^2$ and MAE values in the table and the scatter plots, it is evident that both automatic angle measurements are close to and consistent with the manual measurements by the expert observer. $Angle_1$ measurements are closer to the manual measurements than $Angle_2$ measurements for the whole data set and mutant mice, but not so for wild type mice. The ARL value of 45.42 indicates that $Angle_1$ measurements are very close to the manual measurements for mutant mice with scoliosis symptoms. Although the relative variations between the automatic and the manual measurements are large for the wild type mice, it should be recognised that the real differences between the measurements are small. In addition, the variations are comparable in terms of scale

with inter-observer variations shown in Table 1. Since the manual measurements of spine curvature index are not available, it is not feasible to compare the closeness of the automatic index measurements with manual measurement.

It should be realised that testing only the closeness of automatic measurements to the manual measurements is insufficient when there is no solid ground truth and when manual measurements have their own variations. The effectiveness of the automatic measurements must be judged when they are used in diagnosing scoliosis cases.

## 3.2 **Automatic Identification of Scoliosis Cases**

The aim here is to identify abnormal spines (Scoliosis) from the normal ones using the proposed angle and index based measurements. After obtaining the spine curvature estimation metrics from the X-ray images, we classified the image by applying the k-nearest neighbour (kNN) classifier (K = 1) (Cover & Heart, 1967) upon the extracted curvature estimation metrics. The training and testing of the proposed measurements were done by selecting a random sample of 20 WT and 20 mutant mice images followed by applying leave-one-out-cross-validation. This process was repeated 25 times and the average of accuracies of all rounds was reported. Table 3 shows the summary of the classification results for all proposed automatic measures.

*Table 3: Classification Results Using Automatic Spine Curvature Measures*

| Automatic Metric Used | True Positive | False Negative | True Negative | False Positive |
|:---:|:---:|:---:|:---:|:---:|
| $Angle_1$ | 100% | 0% | 96% | 4% |
| $Angle_2$ | 100% | 0% | 97.2% | 2.8% |
| $SRM_1\_IndexA$ | 98.8% | 1.2% | 89% | 11% |
| $SRM_1\_IndexB$ | 99.6% | 0.4% | 95.2% | 4.8% |
| $SRM_2\_IndexA$ | 99.8% | 0.2% | 95.4% | 4.6% |
| $SRM_2\_IndexB$ | 99.6% | 0.4% | 96.6% | 3.4% |

In general, the results in the table show that all proposed automatic measures perform well in classifying scoliosis cases but $SRM_1\_IndexA$ performed marginally worse than the others. For cases of normal spine, the automatic measures performed marginally worse than that for the scoliosis cases. Some normal spines are mistaken as scoliosis. Overall, $Angle_1$, $Angle_2$, and $SRM_2\_IndexB$ performed better than other measures. The results also indicate that measurements obtained from $SRM_2$ performed marginally better than $SRM_1$. This seems to confirm the earlier stated hypothesis of better promised performance of $SRM_2$ due to fact that $SRM_1$ is determined by only two points whereas the $SRM_2$ is determined by all points on the CSC. However, it is important to note that the

above results provide only a proof-of-concept. Since the proposed method was only tested on a relatively small data sample and the differences between the performances of different measures are only marginal, it is not feasible to draw a solid conclusion about the best measures. Both angle-based and index-based measures have their own merits.

## 4   Discussions

A toolbox containing the proposed method for spine curvature quantification has been implemented using Matlab, and is now publicly available along with some test images[1]. The tool returns all six possible automatic measurements for a selected image as illustrated in the screenshot in Fig. 10.
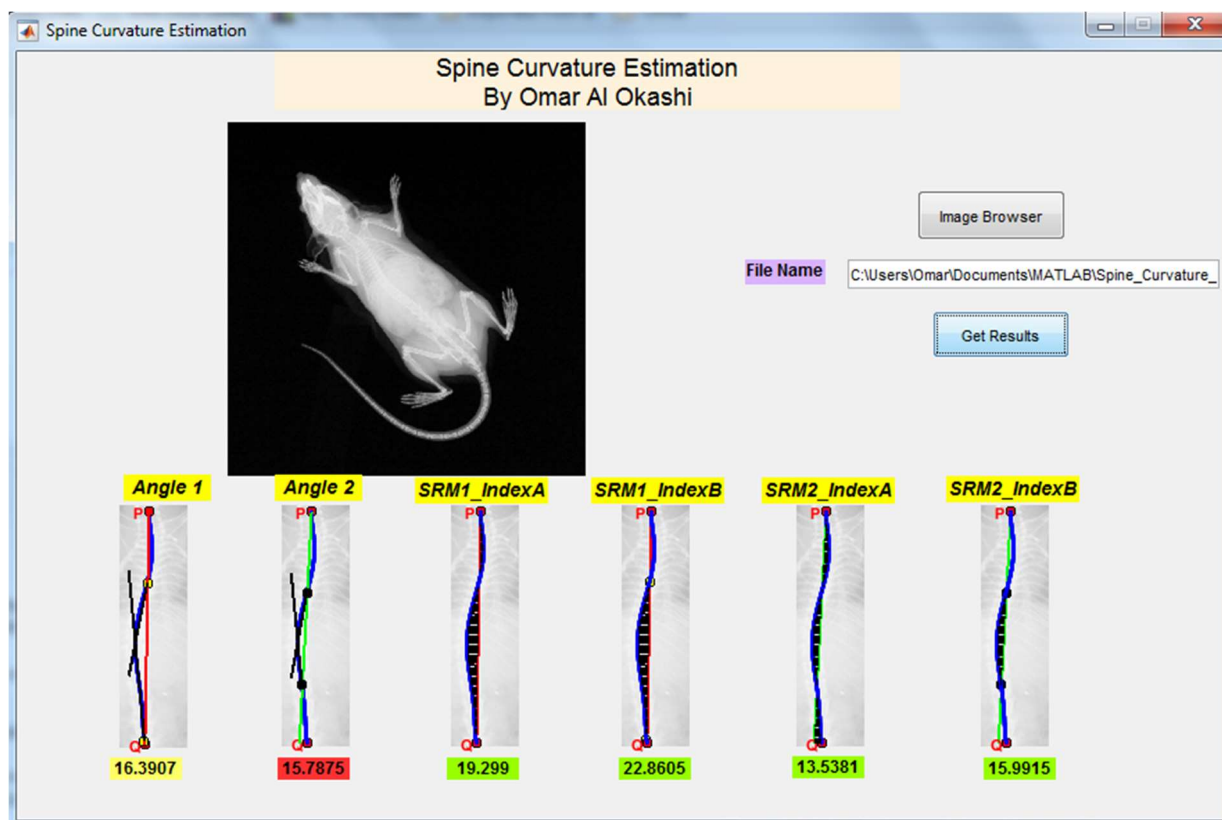


Fig. 10 Screenshot of developed Spine Curvature Estimation Matlab tool.

It is intriguing to know how well the classification results using the automatic measures would be when they are compared with the classification results using manual measurements. We applied the same kNN classifier on manual angle-based measurements of spine curvature. Table 4 below shows the classification results obtained from the manual measurements from the three observers. It

---

[1] http://www.buckingham.ac.uk/research/appliedcomputing/sce-tool

is clear that the classification results by the automatic angle-based measurements of spine curvature are comparable (better TP rate, but worse TN rate) to those from the manual measurements.

*Table 4: Classification Results Using Manual Measurements*

| Automatic Metric Used | True Positive | False Negative | True Negative | False Positive |
|:---:|:---:|:---:|:---:|:---:|
| $A_1$ Trial Average | 100% | 0% | 99.8% | 0.2% |
| $A_2$ Trial | 100% | 0% | 99.4% | 0.6% |
| $A_3$ Trial | 99.6% | 0.4% | 93.8% | 6.2% |

It is desirable to know the relationships between the automatic angle-based measures and the index-based measures. Table 5 shows the degrees of correlation between each of the 4 index-based measures with the $Angle_1$ measure. From the results, it is clear that the correlations between the angle-based measures and index-based measures are stronger for the whole data set and mutant subset, and weaker but still substantially for the wild type subset.

*Table 5: Relationships between Angle-based and Index-based Measures*

| Methods Used (Index Measurements vs $Angle_1$ Measurement) | Whole Data Set of Mice ($R^2$) | Data Set of Mutant Mice ($R^2$) | Data Set of Wild Type Mice ($R^2$) |
|:---:|:---:|:---:|:---:|
| $SRM_1$_IndexA | 0.85 | 0.72 | 0.34 |
| SRM1_IndexB | 0.91 | 0.81 | 0.50 |
| SRM2_IndexA | 0.93 | 0.80 | 0.58 |
| SRM2_IndexB | 0.90 | 0.84 | 0.47 |

A close examination of angle based and distance based measurements shows that when one measure misclassifies the image, the other one gets it right. This can be exploited by applying a simple decision-level fusion to achieve 98.6% accuracy of classifying the True Negative cases.

It is also intriguing to establish what would be the Cobb angle threshold to separate normal from abnormal spines when manual and automatic measures are used in classifying the cases. We used a One-Rule classifier (equivalent to a one level decision tree) on the known cases and obtained the Cobb angle threshold of 5.57° for the manual $A_1$ average measures. The classifier also established the Cobb angle threshold of 6.04° for the automatic $Angle_1$ measure and 5.10° for the automatic $Angle_2$ measure. Both thresholds for the automatic measures are close to that for the manual measure.

16

As a matter of interest, we investigated the applicability of the proposed method on the scoliosis X-ray images of humans after some minor adjustments. As we currently have no access to any sufficiently large datasets, we test our automatic method on a handful X-ray images of human spine as a proof of concept as illustrated in Fig. 11. We believe that our method could be a good solution for low quality human X-ray images especially where the vertebrae boarders are very difficult to be recognized.



(a)                                     (b)                                     (c)
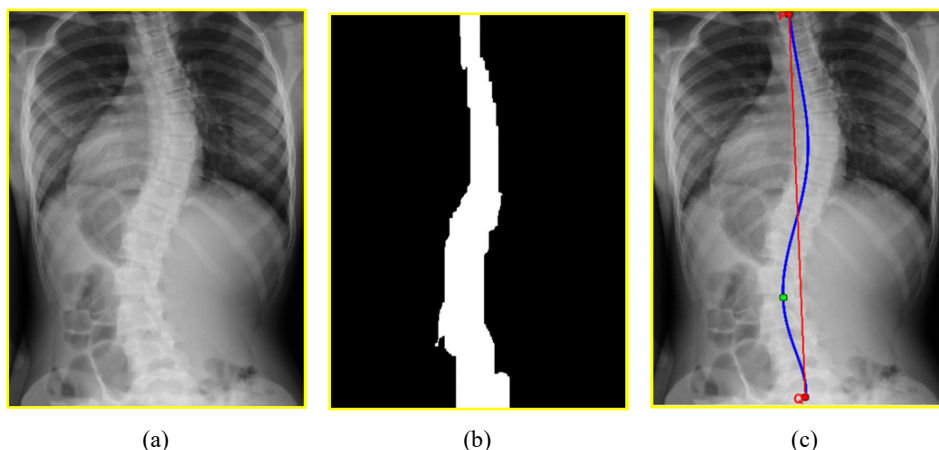
Fig. 11 Applying proposed method on X-ray image of human spine.

(a) X-ray image of human spine; (b) Spine mask; (c) Extracted spine curve (blue) and SRM1 (red).

# 5    Conclusion

Spine curvature estimation is essential to evaluate and quantify the degree of spine deformity in diseases such as scoliosis. In this paper, we proposed an automatic method for spine curvature estimation from X-ray images of mice using different angle and index based measurements. The performance on a small sample showed very good results in either of closeness to manual measurement or classification evaluation, clearly demonstrating the promises and potentials of the proposed method. The decision level fusion of angle based measures with index-based measures could further enhance the classification results. As a side-product of our work, we produced a software framework for the proposed method and made it available for researchers in the community for testing purposes.

Our future work includes extending the test of our method on more images, improving the efficiency of the proposed method, and developing a hybrid curvature measure that combines the strength of both angle-based and index-based measures.

## Acknowledgement

## References

Al Okashi, O., Du, H., Selway, J.L., Lelliott, C., Maguire, S., Melvin, D., Langlands, K. and Al-Assam, H., 2014. Automatic Ribs Segmentation and Counting From Mouse X-ray Images. In MIUA (pp. 143-148)

Al-Helo, S., Alomari, R. S., Chaudhary, V. & Al-Zoubi, M. B., 2011, August. Segmentation of lumbar vertebrae from clinical CT using active shape models and GVF-snake. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 8033-8036). IEEE.

Anderson, S. M., 2007. Spinal Curves and Scoliosis. *Radiologic technology,* 79(1), pp. 44-56.

Chen, . Y. L., Wen, J. C. & Chiou, W. K. , 2007. An alternative method for measuring scoliosis curvature. *Orthopedics ,* 30(10).

Cobb, J. R., 1948. Outline for the study of scoliosis. *Instr Course Lect 5*, pp. 261-275.

Cover, T. & Heart, P., 1967. Nearest neighbor pattern classification. *Information Theory,* 13(1), pp. 21-27.

Diab, K. M., Sevastik, J. A., Hedlund, R. & Suliman, I. A., 1995. Accuracy and applicability of measurement of the scoliotic angle at the frontal plane by Cobb's method, by Ferguson's method and by a new method. *European Spine Journal,* 4(4), pp. 291-295.

Ribeiro, E.A., Nogueira-Barbosa, M.H., Rangayyan, R.M. and Azevedo-Marques, P.M., 2010, August. Detection of vertebral plateaus in lateral lumbar spinal X-ray images with Gabor filters. In 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology (pp. 4052-4055). IEEE

Ferguson, A. B., 1930. The study and treatment of scoliosis. *South Med J,* 23(2), pp. 116-120.

Gonzalez, R. C. & Woods, R. E., 2008. Digital Image Processing. 3rd ed. :Pearson International Edition.

Greenspan, A., Pugh, J. W., Norman, A. & Norman, R. S., n.d. Scoliotic index: a comparative evaluation of methods for the measurement of scoliosis. *Bulletin of the Hospital for Joint Diseases,* Volume 39(2), pp. 117-125.

Huang, S.G., 2010. Wavelet for image fusion. Graduate Institute of Communication Engineering and Department of Electrical Engineering, National Taiwan University.

Janumala, T. & Ramesh, . K. . B., 2015. A Review of Different Analysing Methods to Detect Spinal Deformity. *International Journal of Advanced Research in Computer and Communication Engineering,* 4(8), pp. 352-355.

Kenny, T., 2012. *Patient.co.uk.* [Online]
Available at: http://www.patient.co.uk/health/Scoliosis-(Curvature-of-the-Spine).htm
[Accessed 15 08 2016].

Kundu, R., Chakrabarti, A. and Lenka, P.K., 2012. Cobb angle measurement of scoliosis with reduced variability. arXiv preprint arXiv:1211.5355.

Lonstein, J. E., 1999. CONGENITAL SPINE DEFORMITIES : Scoliosis, Kyphosis, and Lordosis. *Orthopedic Clinics of North America,* 30(3), pp. 387-405.

Moura, D.C., Correia, M.V., Barbosa, J.G., Reis, A.M., Laranjeira, M. and Gomes, E., 2006. Automatic vertebra detection in x-ray images. In CompIMAGE (pp. 307-312).

Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man and Cybernetics,* 9(1), pp. 62-66.

Pratt, W. K., 2001. *Digital Image Processing.* 3rd ed. :Wiley.

Samuvel, B., Thomas, V., Mini, M.G. and Kumar, R., 2012, August. A mask based segmentation algorithm for automatic measurement of Cobb angle from scoliosis x-ray image. In Advances in Computing and Communications (ICACC), 2012 International Conference on (pp. 110-113). IEEE.

Sardjono, T.A., Wilkinson, M.H., Veldhuizen, A.G., van Ooijen, P.M., Purnama, K.E. and Verkerke, G.J., 2013. Automatic Cobb angle determination from radiographic images. Spine, 38(20), pp.E1256-E1262.

Tang, F.H., Chan, L.W., Lau, H.P., Tsui, P.Y. and Cheung, C.W., 2008. Computer-generated index for evaluation of idiopathic scoliosis in digital chest images: a comparison with digital measurement. Journal of digital imaging, 21(1), pp.113-120.

Tomaž, V., Pernuš, . F. & Likar, B., 2009. A review of methods for quantitative evaluation of spinal curvature. *Vrtovec, Tomaž, Franjo Pernuš, and Boštjan Likar. "A review of methods European Spine Journal,* 18(5), pp. 593-607.

White, J.K., Gerdin, A.K., Karp, N.A., Ryder, E., Buljan, M., Bussell, J.N., Salisbury, J., Clare, S., Ingham, N.J., Podrini, C. and Houghton, R., 2013. Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. Cell, 154(2), pp.452-464.

Zhang, J., Lou, E., Le, L.H., Hill, D.L., Raso, J.V. and Wang, Y., 2009. Automatic Cobb measurement of scoliosis based on fuzzy Hough transform with vertebral shape prior. Journal of digital imaging, 22(5), pp.463-472.