Automatic Localisation of Vertebrae in DXA Images using Random Forest Regression Voting

Paul A. Bromiley[1], Judith E. Adams[2], and Timothy F. Cootes[1]

[1] Imaging Sciences Research Group, University of Manchester, UK.
{paul.bromiley, timothy.f.cootes}@manchester.ac.uk
[2] Radiology & Manchester Academic Health Science Centre, Central Manchester University Hospitals NHS Foundation Trust, UK. judith.adams@manchester.ac.uk

Abstract. We describe a method for automatic detection and localisation of vertebrae in clinical images that was designed to avoid making a-priori assumptions of how many vertebrae are visible. Multiple Random Forest regressors were trained to identify vertebral end-plates, providing estimates of both the location and pose of the vertebrae. The highest-weighted responses from each model were combined using a Hough-style voting array. A graphical approach was then used to extract contiguous sets of detections representing neighbouring vertebrae, by finding a path linking modes of high weight, subject to pose constraints. The method was evaluated on 320 lateral dual-energy X-ray absorptiometry spinal images with a high prevalence of osteoporotic vertebral fractures, and detected 92% of the vertebrae between T7 and L4 with a mean localisation error of 2.36mm. When used to initialise a Constrained Local Model segmentation of the vertebrae, the method increased the incidence of fit failures from 1.5% to 2.1% compared to manual initialisation, and produced no difference in fracture classification using a simple classifier.

1 Introduction

Osteoporosis is a common skeletal disorder defined by a reduction in bone mineral density (BMD) resulting in a T-score of <2.5 (i.e. more than 2.5 standard deviations below the mean in young adults), measured using dual energy X-ray absorptiometry (DXA) [16]. It significantly increases the risk of fractures, most commonly occurring in the hips, wrists or vertebrae. Approximately 40% of postmenopausal Caucasian women are affected, increasing their lifetime risk of fragility fractures to as much as 40% [16]. Osteoporosis therefore presents a significant public health problem for an ageing population. However, between 30%-60% of vertebral fractures may be asymptomatic and only about one third of those present on images come to clinical attention; they are frequently not noted by radiologists, not entered into medical notes, and do not lead to preventative treatments [6]. Many of these cases involve images acquired for purposes other than assessment for the presence of vertebral fractures, so identification may be opportunistic. However, a recent multi-centre, multinational prospective study [7] found a false negative rate of 34% for reporting vertebral fractures from lateral radiographs of the thoracolumbar spine. The potential utility of computer-aided vertebral fracture identification systems is therefore considerable. Modern clinical imaging is primarily

digital, with images acquired in Digital Imaging and Communications in Medicine (DICOM) format and stored on a Picture Archiving and Communication System (PACS). A system that could interface with a PACS, use information from the DICOM header to select images that cover a portion of the spine, automatically segment vertebrae, and detect any abnormal shape would therefore be particularly valuable.

Several authors have investigated the use of methods based on statistical shape models to segment vertebrae in both radiographs, e.g. [4], and DXA images e.g. [17]. In recent work [2], the Random Forest Regression Voting Constrained Local Model (RFRV-CLM) [14] was applied to DXA images. This approach uses Random Forest (RF) [1] regressors to predict the locations of landmarks that delineate the vertebrae. Each RF predicts one landmark, and fitting is performed subject to a constraint provided by a global shape model. The RFRV-CLM was shown to be more robust, i.e. suffered from a smaller number of fitting failures on more severely fractured vertebrae, than the Active Appearance Model [5], an important consideration in clinical tasks where pathological cases are most significant. High-resolution RFRV-CLMs require initialisation relatively close to the location of the structure being segmented. The use of multi-stage, coarse-to-fine models can reduce the required initialisation accuracy. However, fully automatic application of such models still requires an initial estimate of the location of each vertebra. This is problematic due to the repetitive nature of vertebrae and the extensive shape differences between normal and pathological anatomy.

A variety of methods that detect potential vertebral candidates have been proposed; see [10] for references. However, most assume that the number of vertebrae visible in the image is known a-priori. This is a significant limitation, particularly when using midline sagittal reformatted images from CT scans performed for various clinical indicators and not specifically acquired to view the spine, since the region of the body imaged can vary significantly. DXA images acquired for vertebral fracture assessment (VFA) typically cover the anatomy from T4 to L4. However, confounding bony structures, adipose tissue in the abdomen, or variation in the field of view can result in vertebrae being obscured or omitted, leading to fit failures with models that assume a certain number of vertebrae are present and can be accurately located. Several authors have proposed solutions to the more general problem of vertebra localisation where the number present is not known a-priori. [12] described a framework for fully automatic localisation, identification and segmentation of vertebrae in CT images, producing a triangulated surface mesh and a level label for all visible vertebrae. However, a complex chain of processing steps was required. [10] used RF regressors trained on features from arbitrary, unregistered CT images to predict the locations of all visible vertebral centroids simultaneously, followed by a refinement stage using a probabilistic graphical model. Later work [11] reported that this approach had problems with images featuring pathology or very narrow fields of view, and instead used RF classifiers to generate a probabilistic label map for vertebral centroids, combined with a shape and appearance model to remove false positives.

We describe a method for detecting vertebrae in clinical images, based on RF regressors, that is intended to be robust to obscured or non-present vertebrae. The key idea is to use a set of relatively non-specific regressors trained to predict the location and pose of lower vertebral end-plates. Each regressor detects multiple vertebrae in a query im-

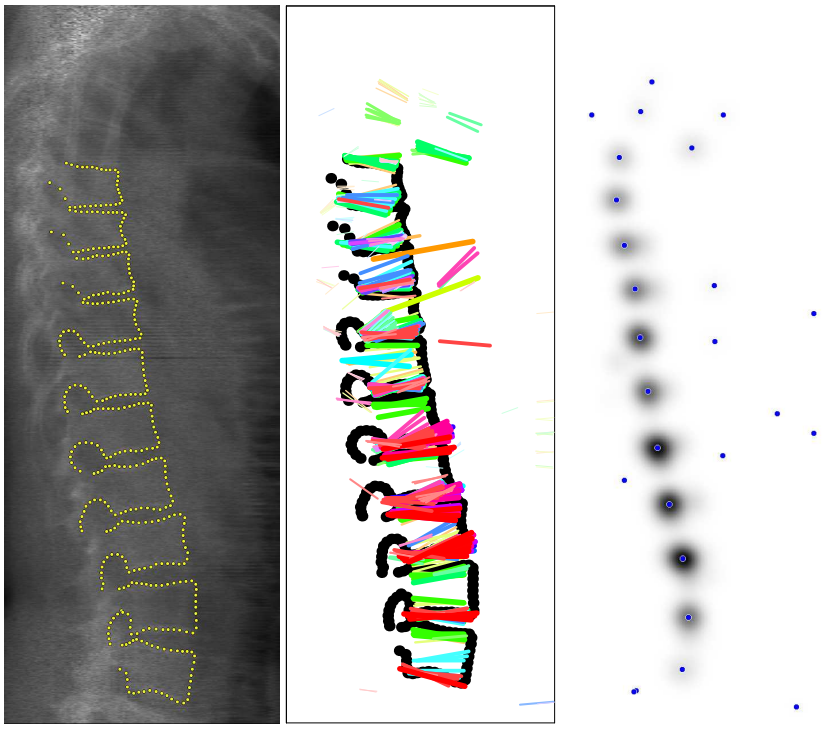

Fig. 1: (Left) Example lateral DXA spinal image with manual annotations of 405 points on T7-L4. (Middle) Top 100 detections from each of the 10 lower end-plate detectors, overlaid on the manual annotations. The hue indicates the vertebral level on which the detector was trained, from T7 (green) to L4 (red), and the saturation and line thickness indicate the weighting. (Right) The smoothed Hough voting array produced from the posterior point of the end-plate detections, and the detected modes.

age, and so does not require a specific vertebra to be present and visible. The results are combined using a Hough-style voting array, and the modes of the smoothed array represent potential detections. A graphical approach is then used to find the highest-weighted path through the modes subject to pose constraints. An evaluation on 320 DXA VFA images with a high prevalence of osteoporotic fractures is described in Section 3.

2 Method

The algorithm described here was based on our own implementation of Hough Forests [8], and used a set of RF regressors, each trained on a different vertebral level, to predict the offset to a distinctive portion of that vertebra given local patches of image features. Figures 1 and 2 show the various stages of the algorithm. Training data consisted of a set of lateral DXA spinal images $\mathbf{I}$ with manual annotations $\mathbf{x}_l$ of N points $l = 1...N$ on

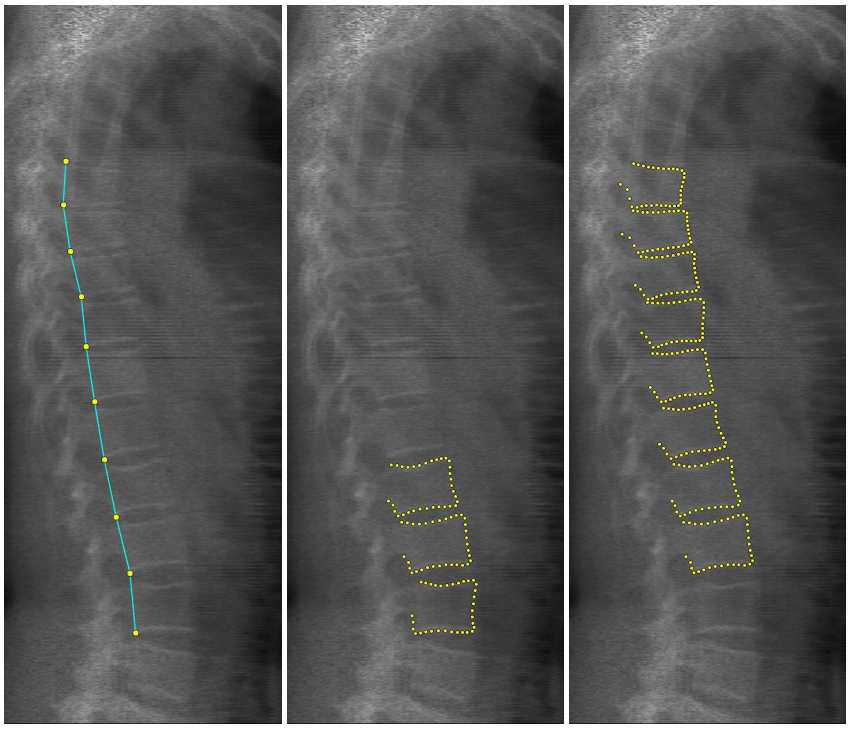

Fig. 2: (Left) The highest-weight set of linked modes from the Hough array, overlaid on the original image (see Fig. 1). (Middle) Example fitted 99-point RFRV-CLM triplet model initialised with the three lowest automatic detections. (Right) Concatenated points from the central vertebra in all fitted RFRV-CLM triplet models initialised from the automatic vertebra detections.

each, outlining the vertebrae (see Fig. 1). The two end points of the curve that delineated the lower end plate were extracted. The lower end-plate was used as, of the four sides of the vertebra visible in the lateral view, it tends to exhibit the smallest changes in size and pose when osteoporotic fractures are present. The two reference points were used to calculate the parameters θ of a similarity registration that transformed the image into a standardised reference frame, such that the reference points were transformed to specific coordinates. The image was then resampled into this frame by applying $\mathbf{I}_r(m,n) = \mathbf{I}(T_\theta^{-1}(m,n))$, where (m,n) specify pixel coordinates. A scaling parameter w_{frame} set the reference frame width in pixels, allowing variation of the resolution.

A set of random displacements $\mathbf{d}_j$ was then generated by sampling from a uniform distribution in the range $[-d_{max}, +d_{max}]$ in x and y. Patches of image data of area w_{patch}^2 were extracted from the reference frame images at these displacements, and features $\mathbf{f}_j$ extracted from them. Haar-like features were used [18], as they have proven effective for a range of applications and can be calculated efficiently from integral images.

The process was repeated with random perturbations in scale and orientation to make the detector locally pose-invariant. The free parameters were chosen to set w_{patch} to twice the length of the lower end-plate and the reference points to relative coordinates $(0.75, 0.25)$ and $(0.75, 0.75)$ in an undisplaced patch i.e. the patch size was equivalent to an entire vertebra plus a border of (approximately, given that vertebrae are not square) one quarter of the end-plate length around each boundary. The resolution was set to 1mm per pixel, and d_{max} was set to 0.4 of the patch size, so that the patches always covered more of the target vertebra than its neighbours. The pose variation was set to 0.1 radians, the resolution of the search (see below).

A RF was then constructed; each tree was trained on a bootstrap sample of pairs $\{(\mathbf{f}_j, \mathbf{d}_j)\}$ from the training data using a standard, greedy approach. At each node, a random set of n_{feat} features was chosen by sampling from a uniform distribution of range $1...N_{MRS}$ and using the result as skip sizes through the feature list. A feature f_i and threshold t that best split the data into two compact groups were chosen by minimising

$$G_T(t) = G(\{\mathbf{d}_i : f_i < t\}) + G(\{\mathbf{d}_i : f_i \geq t\}) \quad where \quad G(\{\mathbf{d}_i\}) = N_s log|\boldsymbol{\Sigma}| \qquad (1)$$

$G(S)$ is an entropy measure; N_s is the number of samples and $\boldsymbol{\Sigma}$ is their covariance matrix. Splitting terminated at either a maximum depth, D_{max}, or a minimum number of samples, N_{min}. The process was repeated to generate a forest of size n_{trees}. Free parameters of the RFs (N_{MRS}, D_{max}, N_{min}, n_{trees}) were set to the values given in [2].

To detect vertebrae in a query image, a grid of points covering the image was defined; the resolution of this grid was set to $3mm$. Each of the RF end-plate detectors was applied at each grid location. The required Haar-like features were extracted and passed into the RF, which output a prediction of the displacement to the reference points. The process was repeated with a range of angle and scale variation. Since RF searching is fast, optimisation of free parameters was avoided by using large search ranges: -0.8 to 0.8 radians in steps of 0.1, and scales from 0.1 to 4 in rational/integer steps. For each detector, the predictions were used to vote into a Hough array. Each vote was weighted by the determinant of the covariance matrix of the samples that reached the relevant RF leaf node. The array was then smoothed using a Gaussian kernel of standard deviation (s.d.) 1mm, to allow mode detection using a 9-way maximum, and the modes detected using exhaustive search. Each mode was a potential end-plate detection.

Vertebrae are repeating structures with considerable similarity in shape between neighbours. Therefore, each RF end-plate detector tended to locate multiple vertebrae across a considerable range of the spine (see Fig. 1). This provided robustness to variations in exactly which vertebrae were visible in the image. A second stage of Hough voting was performed to combine the results from the multiple models and estimate the location and pose of all detected vertebrae[1]. The highest-weighted N_{modes} modes from each detector were used to cast votes into a single Hough array. Each vote was weighted by the length of the line subtended by the two reference points. This had two effects. First, it acted as a shape prior; vertebrae are the largest structures in spinal images that might be delineated by parallel edges, so the weighting aided in elimination of false detections on ribs, humerus etc. Second, it weighted the vote according to the amount

[1]This was performed as a separate step for implementation reasons. We have not yet investigated the possibility of performing the search for all RF regressors using a combined array.

of information it contained about the vertebral pose. The array was then smoothed with a Gaussian kernel of s.d. σ_{comb}, and modes were detected. Estimating the location and pose simultaneously would require a four-dimensional array, which would be inefficient in terms of memory requirements, and would require a large N_{modes} to avoid problems with sparsity. Therefore, a two stage approach was adopted. The posterior-most point of the detection was used to vote into a 2D array. For each detected mode in this array, all votes within $3\sigma_{comb}$ were extracted, and their anterior-most points were used to vote into a second 2D array. This was again smoothed using a Gaussian kernel of s.d. σ_{comb}, and the main mode was detected. σ_{comb} was set to 5mm, and N_{modes} was set to 100.

The result of the combined Hough voting was a single set of N_m potential lower end-plate detections (see Fig. 1), containing false positives. To initialise a subsequent appearance model, an ordered set of points with no missing vertebrae or false detections was required, and so a graphical method was used to extract the highest-weighted subset of modes subject to pose constraints. This could be performed by optimising over all possible paths through the graph. In practice, it was found to be sufficient to start from the highest-weighted mode in the posterior-point Hough array. The main mode of the corresponding anterior-point Hough array gave an estimate of the vertebral pose. The normals to the vector between these modes specified the local superior and inferior directions of the spine, and a path was extracted by searching in each direction using

$$\underset{i=1\ldots N_m, i\neq c}{\arg\min} ||p_i - p_c|| \quad s.t. \quad w_i > w_t, \quad \frac{rot(a_c - p_c, \pm\pi/2).(p_i - p_c)}{|rot(a_c - p_c, \pm\pi/2)||(p_i - p_c)|} > \cos\theta_t \qquad (2)$$

where p_i is the location of mode i in the posterior-point Hough array, a_i the corresponding anterior point, c specifies the current mode, w_i is the weight of mode i, w_t and θ_t are thresholds, and $\pm$ specifies the search direction. θ_t was set to twice the angular resolution of the RF search, and w_t was set to the 2σ value of a Gaussian distribution with a s.d. of σ_{comb} and a mean scaled to the height of the highest-weighted mode in the posterior-point Hough array. This eliminated any statistically insignificant modes whilst accounting both for the smoothing applied and the typical width of the distribution of votes contributing to the modes. The result of the search was a set of ordered points that specified the posterior-most points on the lower end-plates; an example is given in Fig. 2. The remaining orientation ambiguity was resolved by assuming that the superior-to-inferior axis of the body corresponded more closely to the direction of increasing image y-coordinate.

The extracted path might have missing detections, posing a problem for initialisation of subsequent shape and appearance models. Therefore, a second stage of searching through the posterior-point array, equivalent to the use of an adaptive threshold, was included. The lengths of the links in the path were compared to normative vertebral heights learned from the training data, assuming that the lowest detection was on L4, and any heights more than 1.5 times that expected were referred for the second search. A threshold w_t' was calculated using the same procedure as w_t, but the 3σ point of the distribution. The values in the posterior-point voting array along the link were analysed at a set of rational steps (1/2, 1/3 and 2/3 etc.). The set, if any, that had the highest proportional weight, and individual weights higher than w_t', was added to the path. This procedure was sub-optimal as it assumed a vertebral level assignment and, in future work, we intend to replace it with a robust, shape-model based procedure.

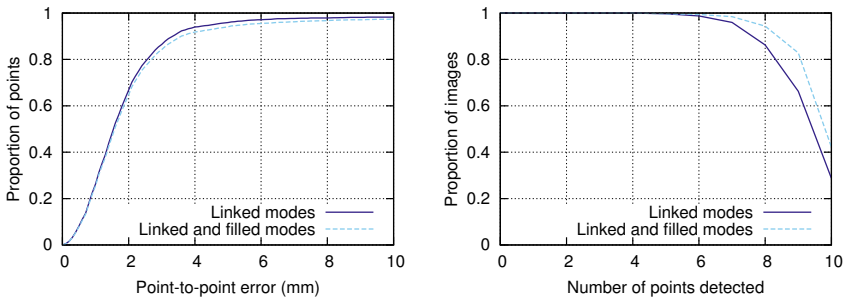

Fig. 3: Cumulative distribution function (CDF) of the P-to-P errors (left) and number of points detected (right) for the 320 DXA images, after path extraction and missing point detection.

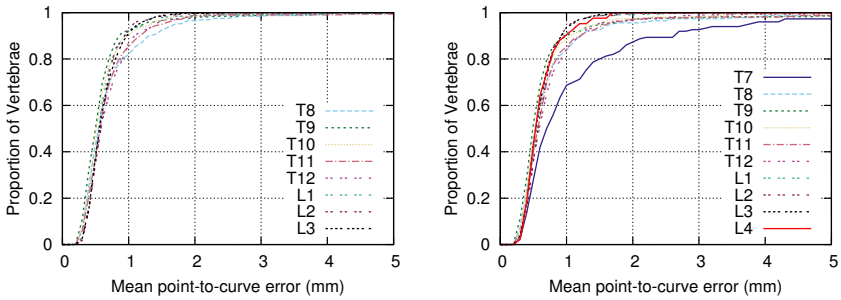

Fig. 4: CDFs of the mean P-to-C errors of the 33 points on the central vertebra in each RFRV-CLM triplet model for the 320 DXA images, for each vertebral level, for manual (left) and automatic (right) initialisation.

3 Evaluation

The method was evaluated on 320 DXA VFA images scanned on various Hologic (Bedford MA) scanners[2]. Manual annotations of 405 landmarks were available for each image, covering the vertebrae from T7 to L4 (see Fig. 1). Each of these vertebrae was also classified by an expert radiologist into one of five groups (normal, deformed but not fractured, and grade 1, 2 and 3 fractures according to the Genant definitions [9]).

The automatic initialisation algorithm was applied in a leave-1/4-out procedure, using 3/4 of the data to train the detectors and then applying them to the remaining 1/4. Figure 3 shows the accuracy of the detected lower end-plate points, and the number of such points detected, after the extraction of the linked path and after the additional

[2]44 patients from a previous study [15]; 80 female subjects in an epidemiological study of a UK cohort born in 1946; 196 females attending a local clinic for DXA BMD measurement, for whom the referring physician had requested VFA (approved by the local ethics committee).

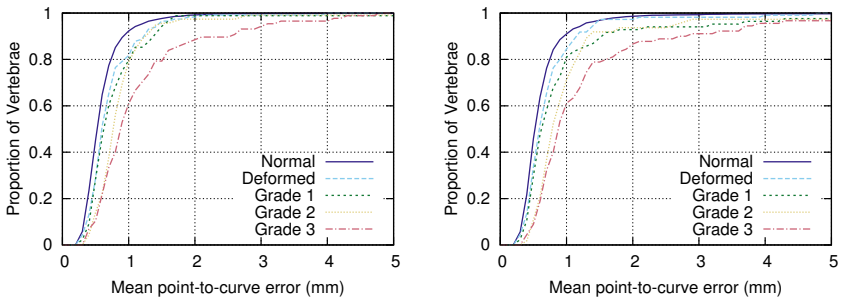

Fig. 5: CDFs of the mean P-to-C errors of the 33 points on the central vertebra in each RFRV-CLM triplet model for the 320 DXA images, for each vertebral status, for manual (left) and automatic (right) initialisation.

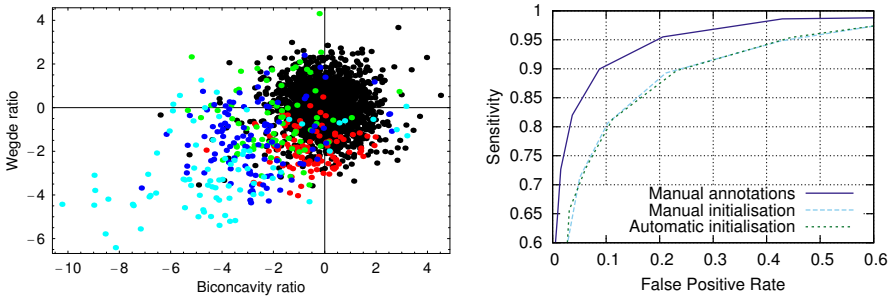

Fig. 6: (Left) The whitened biconcavity and wedge ratios used by the simple classifier (black=normal; red=deformed; green,blue,cyan=grade 1,2,3 fractures, respectively). (Right) ROC curves for classification of the detected vertebrae using the manual annotations, and RFRV-CLM annotations with manual and automatic initialisation.

search to detect missing points in the path. A vertebral level was manually assigned to each detection, by finding the closest manual posterior lower end-plate point. Errors are given as the Euclidean distance, in millimetres, between the detected point and the corresponding manual annotation (point-to-point, or P-to-P, error). Additional vertebrae were detected above T7 in many images, and below L4 in some. Since manual annotations were not available for these, detections above the centroid of T7, and below the centroid of L4 plus its height, were removed from the analysis. The proportions of the vertebrae detected at each level are given in Table 1. There is a reduction in the proportion of detections above T10 and for L4, where the vertebrae are obscured by other bony structures. However, over 90% of the vertebrae were detected at other levels. The additional search to detect missing points resulted in no significant accuracy loss, but a 4.4% rise in detected points. Nine out of the ten target points were found in 85% of the images, and at least six were found in 99% of the images. RF searching was inefficient

Table 1: Statistics of the mean P-to-C errors on each vertebra after RFRV-CLM fitting, using manual and automatic initialisation. Column five gives the %age of vertebrae at each level that had an end-plate detection and, in brackets, a central vertebra from a RFRV-CLM fit assigned to them. Fitting of L4 required L5 to be detected; this is rarely present in DXA VFA images, hence the low percentage of L4 fits.

	Manual Initialisation			Automatic Initialisation			
Vertebral Level	Median (mm)	Mean (mm)	Errors $> 2mm$ (%)	Detected (%)	Median (mm)	Mean (mm)	Errors $> 2mm$ (%)
T7	-	-	-	64.69 (46.87)	0.70	1.19	12.7
T8	0.54	0.72	3.44	87.81 (71.21)	0.53	0.79	4.39
T9	0.49	0.60	2.19	95.63 (88.75)	0.48	0.69	3.17
T10	0.51	0.60	1.56	98.75 (95.63)	0.51	0.67	2.61
T11	0.56	0.70	1.25	100.0 (98.12)	0.55	0.78	2.86
T12	0.59	0.70	1.88	100.0 (99.38)	0.58	0.75	3.14
L1	0.56	0.64	1.25	100.0 (99.38)	0.54	0.62	0.94
L2	0.57	0.63	0.31	100.0 (98.75)	0.57	0.62	0.32
L3	0.56	0.63	0.31	96.88 (79.38)	0.53	0.60	0.0
L4	-	-	-	77.19 (13.13)	0.53	0.60	0.0

by design (see Sec. 2) and took 16.7s per regressor per image on a machine with two Xeon X5670 processors using a single core; the Hough voting based combination of the results and extraction of the linked modes took 1.24s per image.

The automatically detected lower end-plate points were then used to initialise the fitting of a Random Forest Regression Voting Constrained Local Model (RFRV-CLM) covering a triplet of neighbouring vertebrae, as described in [2], and the results compared to those achieved using manual initialisation of the same points. This model covered 99 points (33 on each of the vertebrae in the triplet). For the automatic initialisation, the RFRV-CLM was applied to all triplets of neighbouring initialisation points. The 33 points on the central vertebra were then extracted; no use was made of the overlap of neighbouring models. A vertebral level was manually assigned by comparing the centroid of these points and of the manual annotations for each vertebra. Additional vertebrae above T7 were again removed from the analysis, and errors were calculated for each point as the minimum Euclidean distance to a piecewise linear curve through the manual annotations (point-to-curve or P-to-C error), to compensate for the aperture problem with dense annotations on extended edges. The mean error over each fitted vertebra was then calculated. The same metric was applied to the manually initialised fits and, for consistency, only those triplets for which all three of the initialisation points were available were included. Therefore, results on T7 and L4 were not available for manual initialisation. Example results from a single triplet model fit, and the concatenated central vertebra points from all triplets, are shown in Fig. 2.

Figures 4 and 5 show CDFs of the mean P-to-C error on manually and automatically initialised RFRV-CLM fits for each vertebral level, and for each vertebral status. Tables

Table 2: Statistics of the mean P-to-C errors on each vertebra after RFRV-CLM fitting, using manual and automatic initialisation.

		Manual Initialisation			Automatic Initialisation			
Vertebra Status	%age of Sample	Median (mm)	Mean (mm)	Errors $> 2mm$ (%)	Detected (%)	Median (mm)	Mean (mm)	Errors $> 2mm$ (%)
Normal	84.0	0.52	0.60	0.80	78.7 ± 0.8	0.52	0.63	1.38
Deformed	4.6	0.59	0.71	1.74	78.6 ± 3.4	0.58	0.73	2.73
Grade 1	3.5	0.61	0.79	1.14	83.2 ± 4.0	0.64	0.95	7.14
Grade 2	4.5	0.76	0.88	2.63	85.4 ± 3.6	0.77	1.06	6.30
Grade 3	3.5	0.87	1.16	11.49	85.7 ± 4.0	0.88	1.33	13.33

1 and 2 provide statistics derived from these curves, including the percentage of vertebrae with mean P-to-C errors above 2mm, which gives an indication of the proportion of fit failures. There is little difference between the medians or means for either status or level, indicating that the accuracy of the RFRV-CLM fitting is largely independent of the initialisation as long as it is within the capture range of the model. The percentage of fitting failures on all detected vertebrae between T8 and L3 increased from 1.5% for manual initialisation to 2.1% for automatic, but the increase was larger for fractured vertebrae. However, using the 2mm threshold, at least 86.7% of vertebrae were successfully fitted, regardless of status. Importantly, there was no statistically significant variation in the proportion of vertebrae detected with vertebral status i.e. no evidence of bias against detecting fractured or non-fractured vertebrae.

To evaluate the importance of the differences in segmentation accuracy, a simple classifier based on the standard six-point morphometry technique was applied to the results from both manually and automatically initialised RFRV-CLM fits. The anterior H_a, middle H_m, and posterior H_p heights of each vertebra were extracted. A predicted posterior height $H_{p'}$ of each vertebra was also calculated by taking the posterior heights of the closest four annotated vertebrae, multiplying by ratios of vertebral heights in normative data obtained from [13], and taking the maximum of the four values (on the basis that fractures reduce vertebral height). This process used the vertebral level assignment derived from the manual annotations for the automatically initialised results. Three ratios were then calculated: the biconcavity ratio H_m/H_p, the wedge ratio H_a/H_p, and the crush ratio $H_p/H_{p'}$. The data were whitened by subtracting the median and dividing by the square root of the covariance matrix (estimated using the median absolute deviation). The resulting data for the automatically initialised annotations are shown in Fig. 6, showing clear separation between the normal and fractured vertebrae and, in particular, that deformed vertebrae are displaced from the normal class along a different vector through the space than the fractured vertebrae. However, since only a simple classifier was intended, a threshold was applied to the Euclidean distance between the origin and the point defined by the three height ratios, to classify the vertebrae into non-fractured (normal and deformed) and fractured (grade 1, 2 and 3) classes. Figure 6 shows ROC curves produced by varying the threshold. The simple classifier achieves 90% sensitiv-

ity at a false positive rate of 10% when applied to the manual annotations. Using the manually initialised annotations reduced this to 80%, largely due to errors on grade 1 fractures, where shape changes can be quite subtle. However, there was no significant difference in classifier accuracy between manual and automatic initialisation.

4 Conclusion

This paper has described a method for automatic detection and localisation of vertebrae in clinical images. The algorithm relies on a set of RF regressors trained to predict the location of vertebral lower end-plates. We have demonstrated that the individual RFs are not specific to the vertebrae they were trained on but instead, due to the similarity of neighbouring vertebrae, respond over a considerable range of the spine. This provides robustness to vertebrae that are obscured or not present. The use of multiple models, and Hough voting to combine their results, provides robustness to fit failures [3]. The failure of any one RF to detect vertebrae will result in responses scattered throughout the voting array. Only locations that result in strong responses from multiple models will result in significant modes. These are detected, and a graphical method applied to find a path through the detections, subject to pose constraints. This can be used to initialise a high-resolution appearance model that provides an accurate segmentation.

The method was evaluated on 320 DXA VFA images with a high prevalence of osteoporotic fractures. Other authors who have studied this problem [10] [11] [12] have used CT images that, since the modality is used for a wider range of clinical purposes, show a much larger variation in the region of the spine being imaged and the number of vertebrae visible. The difference in the dimensionality of the the images prevents any comparison of localisation accuracy. However, we note that the proportions of vertebrae detected in [10] and [11] are similar to the detection rates presented here. Instead, our evaluation focused on using the automatic annotations, and the equivalent manually annotated points, to initialise a RFRV-CLM. This demonstrated no difference in accuracy when the initialisation was within the capture range of the RFRV-CLM. Automatic initialisation led to more fit failures on grade 1 fractures, where the shape change may be subtle. However, the increase was smaller on grade 2 and 3 fractures. A simple classifier was applied, and showed little difference in performance between RFRV-CLM segmentations using manual and automatic initialisation.

One drawback of the algorithm is that it implicitly assumes prior knowledge of an overall image scaling, in the form of the pixel size in mm, although this information is available in the DICOM header for digital clinical images. Another is the relatively simple technique used to extract the path through the modes of the voting array. An appearance model based technique, such as that described in [11], might prove more robust. Finally, we have not yet evaluated the technique on images showing wide variation in the number of visible vertebrae. These issues will be addressed in future work.

Acknowledgment

This publication presents independent research supported by the Health Innovation Challenge Fund (grant no. HICF-R7-414/WT100936), a parallel funding partnership between the Depart-

ment of Health and Wellcome Trust. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health or Wellcome Trust.

References

1. Breiman, L.: Random Forests. Machine Learning 45, 5–32 (2001)
2. Bromiley, P.A., Adams, J.E., Cootes, T.F.: Localisation of Vertebrae on DXA Images using Constrained Local Models with Random Forest Regression Voting. In: Recent Advances in Computational Methods and Clinical Applications for Spine Imaging. Lecture Notes in Computational Vision and Biomechanics, vol. 20, pp. 159–172. Springer (2015)
3. Bromiley, P.A., Schunke, A.C., Ragheb, H., Thacker, N.A., Tautz, D.: Semi-automatic landmark point annotation for geometric morphometrics. Front Zool 11(61), 1–21 (2014)
4. de Bruijne, M., Lund, M., Tanko, L., Pettersen, P., Nielsen, M.: Quantitative vertebral morphometry using neighbour-conditional shape models. Med Image Anal 11, 503–512 (2007)
5. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE TPAMI 23, 681–685 (2001)
6. Cummings, S.R., Melton, J.: Epidemiology and outcomes of osteoporotic fractures. The Lancet 359(9319), 1761–1767 (2002)
7. Delmas, P.D., van de Langerijt, L., Watts, N.B., Eastell, R., Genant, H.K., Grauer, A., Cahall, D.L.: Underdiagnosis of vertebral fractures is a worldwide problem: the IMPACT study. J Bone Miner Res 20(4), 557–563 (2005)
8. Gall, J., Lempitsky, V.: Class-specific Hough forests for object detection. In: Proc. CVPR. pp. 1022–1029. IEEE Computer Society (2009)
9. Genant, H.K., Wu, C.Y., Kuijk, C.V., Nevitt, M.C.: Vertebral fracture assessment using a semi-quantitative technique. J Bone Miner Res 8(9), 1137–1148 (1993)
10. Glocker, B., Feulner, J., Criminisi, A., Haynor, D.R., Konukolu, E.: Automatic localization and identification of vertebrae in arbitrary field-of-view CT scans. In: Proc. MICCAI 2012. LNCS, vol. 7512, pp. 590–598. Springer (2012)
11. Glocker, B., Zikic, D., Konukoglu, E., Haynor, D.R., Criminisi, A.: Vertebrae Localization in Pathological Spine CT via Dense Classification from Sparse Annotations. In: Proc. MICCAI 2013, LNCS, vol. 8150, pp. 262–270. Springer (2013)
12. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med Image Anal 13(3), 471–482 (2009)
13. Leidig-Bruckner, G., Minne, H.W.: The Spine Deformity Index (SDI); A New Approach to Quantifying Vertebral Crush Fractures in Patients with Osteoporosis. In: Vertebral Fracture in Osteoporosis, pp. 235–252. Osteoporosis Research Group, University of California (1995)
14. Lindner, C., Bromiley, P.A., Ionita, M., Cootes, T.F.: Robust and Accurate Shape Model Matching using Random Forest Regression-Voting. IEEE TPAMI p. In press (2015)
15. McCloskey, E., Selby, P., de Takats, D., Bernard, J., Davies, M., Robinson, J., Francis, R., Adams, J., Pande, K., Beneton, M., Jalava, T., Loyttyniemi, E., Kanis, J.A.: Effects of clodronate on vertebral fracture risk in osteoporosis: a 1-year interim analysis. Bone 28(3), 310–315 (2001)
16. Rachner, T.D., Khosla, S., Hofbauer, L.C.: Osteoporosis: now and the future. The Lancet 377(9773), 1276–1287 (2011)
17. Roberts, M.G., Cootes, T.F., Adams, J.E.: Automatic Location of Vertebrae on DXA Images Using Random Forest Regression. In: Proc. MICCAI 2012. LNCS, vol. 7512, pp. 361–368. Springer-Verlag, Berlin (2012)
18. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. CVPR. pp. 511–518. IEEE Computer Society (2001)