Landmarking and Feature Localization in Spine X-rays

L. Rodney Long, National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894

E-mail: long@nlm.nih.gov

George R. Thoma, National Library of Medicine, 8600 Rockville Pike, Bethesda, MD 20894

E-mail: thoma@nlm.nih.gov

Abstract

The general problem of developing algorithms for the automated or computer-assisted indexing of images by structural contents is a significant research challenge. This is particularly so in the case of biomedical images, where the structures of interest are commonly irregular, overlapping, and partially occluded. Examples are the images created by digitizing film x-rays of the human cervical and lumbar spines. We have begun work toward the indexing of 17,000 such spine images for features of interest in the osteoarthritis and vertebral morphometry research communities. This work requires the segmentation of the images into vertebral structures with sufficient accuracy to distinguish pathology on the basis of shape, labeling of the segmented structures by proper anatomical name, and classification of the segmented, labeled structures into groups corresponding to high level semantic features of interest, using training data provided by biomedical experts. In this paper, we provide a technical characterization of the cervical spine images and the biomedical features of interest, describe the evolving technical approach for the segmentation and indexing problem, and provide results of algorithms to acquire basic landmark data and localization of spine regions in the images.

Keywords

Segmentation, digitized x-ray, biomedical, classification, cervical spine, lumbar spine, ASM, image indexing, NLM, NIH, NIAMS, NHANES

1. Problem

At the Lister Hill National Center for Biomedical Communications, a research and development division of the National Library of Medicine, we are developing a prototype multimedia database system to provide access to text and related x-ray images over the World Wide Web. This *WebMIRS (Web-based Medical Information Retrieval System)*[1-2] will allow access to databases containing text and images and will allow database query by standard Structured Query Language (SQL), by image content, or by a combination of the two. The WebMIRS results screen is illustrated in Figure 1. This beta-level system is capable of retrieving text and reduced resolution cervical spine and lumbar spine x-ray images. However, except for image display, the current WebMIRS is very similar to many other databases that are text-only: no image content information, such as quantitative measures of vertebrae, or classifications of the vertebrae for pathology, are currently available in the database; further, all queries are made with GUI-assisted, standard SQL. No query by image example is currently possible. Toward building these more advanced capabilities, we are addressing fundamental problems in the required image processing and pattern classification.

1.1 Current WebMIRS operation

In the current WebMIRS system, the user manipulates GUI tools to create a query such as, "Search for all records for people over the age of 60 who reported chronic back pain. Return the age, race, sex, and age at pain onset for these people." In response, the system returns a display of values for these four fields for all matching records, plus a display of the associated x-ray images.

1.2 Future WebMIRS operation

A future WebMIRS system is envisioned to have additional capabilities to support queries such as the following:

> Example 1: "Search for all records for people over the age of 60 who reported chronic back pain. Return the age, race, sex, age at pain onset, and ratio of anterior/posterior vertebral heights for the L4 vertebra."

Example 2: "Search for all records for people over the age of 60 who reported chronic back pain and who have an *L4 vertebra with shape resembling this one*. Return the age, race, sex, age at pain onset, and ratio of anterior/posterior vertebral heights for the L4 vertebra."

The requirements for a system capable of queries as exemplified by example 1 differ from those exemplified by example 2. In example 1, the query is conventional; the image content data is simply text in relational tables. (However, the costs in time and money of acquiring this data by the manual labor of medical experts are prohibitive. The implication is that only by means of an automated or sufficiently cost-effective computer-assisted system will the image content data be successfully acquired. Hence, even to populate our RDBMS tables with this type of data, research into algorithms to segment and identify anatomy and derive measurements meaningful to the end user are needed.)

In example 2, the query itself is non-conventional. The system is required to accept as input not just SQL, but an example image, in addition. The database tables contain *feature descriptors* for each image in the database. The feature descriptors for an image consist of data derived from that image that characterize the image contents in a manner that allows for retrieval based on visual features meaningful to the user. In this case, the program has an additional requirement for a record to match the input query: the feature descriptors for the record being compared must satisfy a similarity requirement relative to the input example image. For an example 2 system to become operational, basic problems in indexing image contents, and deriving feature descriptors, must be solved in order to populate the RDBMS tables, plus the system must be enabled to accept a fundamentally different type of query; in addition, the system must incorporate a concept of similarity measurement for judging the degree to which an image in the database resembles an input example image.

Our eventual goal is to have a system that will support not only example 1 but also example 2 queries. Toward this goal we are pursuing research into image processing techniques that will support the hierarchical segmentation of the images, first into anatomically-related regions at a level of gross detail, then a fine level segmentation of the spine area into individual vertebrae. This segmentation stage will be followed by a stage of identification of the

anatomy within the spine by specific vertebra number, i.e., for the cervical spine, an individual vertebra will be identified as one of C1 through C7. Finally, the features of interest within the segmented, labeled spine anatomy will be classified. For example, the disc spacing for each pair of vertebrae will be classified as "normal" or "abnormal": a result might be, "C5-6 disc spacing is abnormal".

The challenges in building biomedical image databases are many and complex, and it is beyond the scope of this paper to address them completely. A comprehensive overview has been authored by Tagare[3].

2. Goals, approach, and related work

Our goals are to index the images by specific features of interest identified by biomedical subject matter experts. Primarily, these are, for the cervical spines, anterior osteophytes, disc space narrowing, and subluxation; for the lumbar spines, these are anterior osteophytes and disc space narrowing. These features are of interest to the osteoarthritis community and were identified in two workshops[4] conducted at the National Institutes of Health (NIH) and sponsored by the National Institute of Arthrtitis and Musculoskeletal and Skin Diseases (NIAMS). Additional features include basic dimensionality and inter-vertebral geometry measures such as anterior, posterior, and mid-body heights for each vertebra, and intervertebral spacing measures. These measures are used within the vertebral morphometry research community[5-9] and/or are closely related to the features of interest to the osteoarthritis community.

In our approach we conceive the problem as (1) a hierarchical analysis and segmentation of the images, beginning at gross level ("blob" level) features defined by grayscale and connectivity characteristics, and proceeding through deformable template segmentation of vertebrae using statistical anatomical models; (2) classification and labeling of the segmented anatomy by structural name; and (3) classification ("indexing") of the segmented, labeled anatomy according to pathology or high-level semantic feature of interest.

To our knowledge, no comprehensive solution to this automated or semi-automated indexing problem has been achieved for a large collection of digitized spine x-rays, although progress on sub-problems has been reported in the

technical literature by several researchers, either for digitized film of the spine or closely related image modalities for the spine or the hand. An early approach to localizing the spinal canal, intervertebral discs, and other lower spine features in magnetic resonance images (MRI) was reported by Chwialkowski[10], who achieved good results in modeling the spinal canal curvature as a second-order polynomial. Later, a fundamental and comprehensive treatment of the whole field of Active Shape Modeling (ASM), which has given technical direction to a number of research efforts, including our own, was provided by Cootes[11]. A semi-automated implementation of the ASM approach has been achieved for the vertebral segmentation of lumbar spine dual x-ray absorptiometry (DEXA) images[12]; in this implementation, the user must manually identify two "anchor points" for placing a template on the target image; the template then deforms according to the ASM algorithm, maintaining invariant location of the anchor points, which are placed at the top and bottom of a column of vertebrae. Gardner[13] also developed a semi-automated system, based on active contour (snake) modeling of the vertebrae, which, in particular, operated on digitized lumbar spine x-ray film. In this system, points on the vertebral boundaries are specified by the user, with assistance from the system in point placement. The selected boundary points then become constraints on the active contour that is automatically fit to the vertebra boundary. This process is carried out vertebra by vertebra. The same author has pointed out[14] the potential problems in taking dimensional measurements from vertebrae on x-rays, due to the projective nature of the imaging modality, which results in overlapping edges and concealed boundaries. A comprehensive treatment of x-ray segmentation as applied to hand radiographs has been given by Efford[15], and additional technical papers are provided in the references[16-18].

Previous direct work on the segmentation for this particular x-ray collection has occurred, though it is at an early stage. We have previously reported[19-20] promising segmentation work for small test sets of these images, using manually-acquired vertebral boundary data sampled from 15-20 images and deformed to fit image data by an implementation of the ASM algorithm. Work done independently by Sari-Sarraf[21] using similar techniques and tools also appeared to give successful results to a first-order level in a significant number of cases. It should be noted that both in our work and the work of Sari-Sarraf a number of problems were observed, and technical issues that require resolution were raised. Among the most outstanding of these are the need for a good method for initializing ASM to segment the vertebrae (although Zamora[22], and we, in this paper, have done work toward that goal), the need to investigate the effects of modeling the image grayscale distribution in the neighborhood of

vertebrae boundary points with a mixed Gaussian probability distribution function, and the need to understand nonconvergence of the algorithm for certain images even when the template initial conditions are set near known truth.

The final indexing and classification step for our images requires taking the segmented images along with expert training data, and labeling the anatomy by structural name, as well as dividing the labeled structures into classes of normal or abnormal for the conditions of interest. It also generates all of the geometry measures desired from the data. Work toward this step that applies radius of curvature criteria to segmented vertebrae boundaries has been reported by Stanley[23], who investigated features and classification algorithms for the computer-assisted indexing of cervical spine x-rays for normality/abnormality with respect to anterior osteophytes. Stanley's work concentrates on the feature selection/classification problem; the vertebra boundary determination is done by fitting a B spline to a set of manually-selected boundary points (chosen with the aid of Kirsch edge detection and a small number of points placed by medical expert). He measured radius of curvature along the anterior part of the vertebra boundary, and examined a total of 31 features, including radius of curvature and its first and second derivatives calculated at the boundary point of minimum radius of curvature, and at boundary points in the neighborhood of this minimum radius of curvature point; additional gradient-based features were used, including maximum, minimum, mean, and standard deviation of gradient at each of the boundary points on the vertebra anterior. Stanley reported classification results using a back propagation neural network, K-means classification, a quadratic discriminant classifier, and Learning Vector Quantization 3. Of the four classifiers, the neural network produced best results with 71% correctly classified vertebra on a test set of 35 vertebrae (trained on 83 vertebrae).

3. Characterization of the images

3.1 Global characteristics

The 17,000 images in our collection consist of approximately 10,000 cervical spine, and approximately 7,000 lumbar spine images. These lateral view images were collected as part of the second National Health and Nutrition

Examination Survey [24-25] (NHANES II) in film form, and were digitized at 146 dpi, 12 bits per pixel, with a Lumisys laser scanner. The resulting image dimensions are 1463x1755 for the cervical spine images, and 2048x2487 for the lumbar spine images. No formal study of the bit level significance in these images was conducted, but informal analysis of the lower bit planes in these images, plus estimations of bit level significance in similar images in the published literature [26-28] strongly suggest that the information content in the images beyond the most significant 8 bits is likely to be extremely small. In this paper, we report on work with 8-bit images of the cervical spine, only. These images are 1462x1755x8 (the modification in image width from 1463 to 1462 was done strictly for convenience of software in dealing with an even number of pixels in a line). A summary of the overall characteristics of a sample of these images is given in Table 1. By displaying a sample of the cervical spine images, it is straightforward to arrive at the hypothesis that, at a gross level, the images appear to contain two bright regions, corresponding approximately to the skull and shoulder, and at least one consistently dark background region, corresponding to the region back of the skull. This observation became the basis for an analysis of basic landmarks in these images that is presented later in this paper.

3.2 Spine region

For a small subset (550) of the images we have acquired coordinate data for key points on and around the vertebrae. This data was collected under supervision of a board certified radiologist with expertise in bone x-rays. Up to 9 data points were collected per vertebra, as illustrated in Figure 2, with points 1-6 corresponding to the standard 6 points commonly collected in the field of vertebral morphometry, point 7 corresponding to the anterior midpoint of the vertebra, and points 8 and 9 corresponding to the maximum protrusion of osteophytes on the anterior top and bottom, respectively, of the vertebral body. (An *osteophyte* is a "bony outgrowth or protuberance"[29].) These data points were collected for all of the vertebrae with sufficiently visible boundaries. Typically, for the cervical spine images, these included the vertebrae boundaries from the bottom of C2 through C6, though in a few cases C7 and T1 were also visible. For the lumbar spine boundaries, the collected data typically spanned L1-L5, although in a few cases parts of the boundaries for the thoracic T12 and sacral S1 were collected. Additionally, a few special points were collected: for the cervical spine, these included a point marking the approximate center of gravity of the C1 vertebra. The nomenclature for the vertebrae naming and the overall spine anatomy are illustrated in Figure 3.

This expert-identified data is very useful for model building for the spine, for use as reference data for evaluating performance of algorithms to localize spine structures [21-22], and for setting parameters within algorithms that analyze the spine. For 46 of the cervical spine images, we computed the statistics of basic distance measures both within the vertebrae and for intervertebral quantities. Figure 2 illustrates the geometry for the distance measurements taken for the intervertebral and within-vertebra quantities, respectively. Table 2a gives the statistics for the intervertebral quantities: for C1/C2 these numbers characterize the distance from the C1 approximate center of gravity to the C2 bottom midpoint; for each of the other entries the table values correspond to the distances between neighboring pairs of points in adjacent vertebrae. These distances might be considered in some sense as first-order approximations to the intervertebral spacing; however, note that the term "disc spacing" is a high level semantic description applied by biomedical experts, and that we are not attempting in this paper to equate that description to these measurements. Table 2b gives the statistics for the within-vertebra quantities. Note that deriving dimensional data from these coordinates requires choosing among different computational options, based on the likely use of the computed result. For example, vertebral "height" may be computed as the distance between points 3 and 6 on a vertebra, but this essentially treats the vertebra as a rectangle, with the corresponding error in what we may conceive as the "true" height of the vertebra. A more sophisticated approach would be to fit straight lines to the top and bottom of the vertebra, and to measure height as the distance between these lines along a perpendicular passing through point 3. (In the system described by Gardner[13], the user marks the points 1 and 3, then the system automatically places point 4 so that it lies on a perpendicular to the line determined by points 1 and 3.) Note also that the last two columns contain distances related to the extent of the protrusion of anterior osteophytes. These values correspond to simple distance measurements between points 3 and 8, and between points 6 and 9, for top and bottom osteophytes, respectively, and may not be "good" characterizations of the extent of the osteophyte protruberance. For example, this extent might be better measured as distance along a perpendicular dropped from point 8 (point 9) to a straight line fit to the three points (points 3, 7, and 6) that lie on the front of the vertebra.

The vertebrae grayscale characteristics are complex. It is easy to find examples in the images of vertebrae with interior regions having grayscale values similar to those in regions that are neighboring, but external to, the vertebral body. This is illustrated in Figures 4a and 4b. In the surface plot shown in Figure 4b, the prominent body near the

center of the image is C4, with parts of C3 and C5 shown at its top and bottom, respectively. The C4 grayscale distribution for the interior region near the center of the vertebral body closely resembles the distribution in the region neighboring, but external to, C4, that lies within the C3/C4 disc space. A sample of grayscale in the two regions yielded: $\mu = 197.4$, $\sigma = 2.2$, for the interior region; and $\mu = 198.1$, $\sigma = 1.6$, for the external region, within the C3/C4 disc space. Hence, segmentation methods for the vertebrae that rely heavily on grayscale value as a discriminator of vertebra/non-vertebra regions tend not to be robust.

3.3 Biomedical features of interest: anterior osteophytes, an example

One of the important biomedical features desirable for indexing the cervical spine images is presence/absence of anterior osteophytes. Figures 5-7 illustrate this feature as a localized shape characteristic not susceptible to analysis by global indexing methods such as global histogram or global shape analysis available in existing image database systems [30-31]. Figure 5 shows four cervical spine images with, first, no osteophytes, then the presence of progressively more severe osteophytes. The grading of osteophyte presence and severity in these images was done by consensus of three rheumatologists with expertise in interpreting spine x-rays for features related to osteoarthritis in a project sponsored by the National Institute of Arthritis and Musculoskeletal and Skin Diseases. This project resulted in the creation of a digital atlas of the cervical and lumbar spines[32]. Figure 5 illustrates the fact that, at the global image level, the available visual information is not of much use in detecting osteophytes. Figure 6 shows subimages cropped from the Figure 5 x-rays where the local shape characteristics of the vertebrae are more apparent. In Figure 5 the progression from a (normal) smooth rounded corner to an irregular, strongly distorted corner may be seen on the lower anterior corner of the central vertebra in each subimage. In Figures 7a-d the grayscale in these subimages has been plotted as a surface plot viewed from an elevation of 90 degrees (directly above); the vertebral boundaries have been traced by hand to clearly indicate the osteophyte shapes and extents.

4. Multiresolution analysis overview

At full spatial resolution, the cervical spine images exhibit a variability of image characteristics that significantly complicates analysis and segmentation of the contents. The most successful approaches that we have used require

statistical or integrative techniques, or similar techniques that analyze the image contents by grouping pixels as regions, lines or curves, rather than individual points. An alternative approach is to begin the image analysis at a lower spatial resolution version of the image, produced by blurring and subsampling the original. Multiresolution (or "multiscale") approaches are widely employed in boundary detection problems within the image processing community. Examples are the work of Cootes[11] in ASM segmentation, and of Liang[33] and Mignotte[34], both for the detection of artery boundaries in ultrasound images.

As the resolution of the cervical spine image decreases, the visual distinctions among the individual vertebrae are lost, as well as the spinous process anatomy at the back of the vertebrae, and all distinctive anatomical features within the skull area (teeth, sinus area, orbits of the eyes, etc.). At a very low resolution the predominant visual features of the cervical spine images are the bright spots in the image corresponding to the approximate skull and shoulder regions, and the distinctively dark region corresponding to the background behind the skull. As shown below, it is possible to segment these regions in a very low resolution version of the image, and to map from the centers of gravity of these regions to points in the full resolution image that correspond to landmarks in the near skull, shoulder and background regions in a reliable manner. These landmarks can then (presumably) be used to refine knowledge of the image anatomy in the full resolution image.

5. Results and discussion

All algorithm work in this paper used 1462x1755x8 cervical spine images with grayscale values normalized to lie between 0 and 1, corresponding to the minimum (maximum) grayscale value in the original image.

5.1 Finding basic landmarks in the images

The landmark algorithm that we have developed takes as input one cervical spine image I and outputs three pairs of (x,y) coordinates labeled SK,SH, and BG, which correspond to the approximate locations of skull, shoulder, and back-of-skull background in the image I. The algorithm is heuristic in nature and relies on the observation that

skull, shoulder and back-of-skull background regions appear to be preserved as recognizable bright or dark "blobs", even when the images are heavily smeared and subsampled. Steps in the algorithm are described below:

(1) Compute B=G(I). Produce the blurred image B by filtering the input image I so that the pixels remaining after subsequent subsampling represent the grayscale characteristics of regions, rather than individual pixels, in the original image. For the filtering, we experimented with Gaussian filters with a variety of parameters (filter size and sigma) by viewing the filtered images and noting whether the image had been largely reduced to the grayscale blob level, and chose a 100x100 filter with sigma of 100 to produce the severe blurring that we desired.

(2) Compute SB=S(B). Produce the subsampled, blurred image SB by using a subsampling process S on the blurred image B to reduce it in spatial dimensions to a size easy to analyze. One of the areas that we wanted to explore was, whether significant information could be obtained from very low resolution versions of the images. With this motivation, we used a subsampling factor of 2^8, so that the SB matrix is only of size 6x7, (as compared to the original image size of 1462x1755). Figures 8a and 8b show an example of an original image I and the resulted subsampled blurred image SB. In Figure 8b the two bright regions corresponding to skull and shoulder, and the dark background region behind the skull, can be observed.

(3) Identify regions of interest R_{s1}, R_{s2}, and R_{bg} in the subsampled blurred SB image, and compute landmarks within each of these regions. The regions R_{s1} and R_{s2} correspond to the two brightest blobs in SB, which we expect to correspond to the skull and shoulder regions. At this step we do not determine specifically which of the R_{s1} and R_{s2} regions is skull and which is shoulder. The region R_{bg} corresponds to the dark image background behind and, in some images, on top of the skull. The identification of R_{s1} is accomplished as follows: (a) find the brightest grayscale value g_{b1} in the image SB; (b) using g_{b1} as a seed, grow a region containing g_{b1} by iteratively examining the 8-neighbors of pixels already in the region, and adding to the region any of these 8-neighbors having a grayscale value within a tolerance of g_{b1}; the resulting connected region is R_{s1}. The region R_{s2} is similarly found by finding the second brightest grayscale value g_{b2} in SB that lies in a region disjoint from R_{s1} (i.e. no 8-neighbors of R_{s1} boundary pixels lie in this region). Just as for R_{s1}, the g_{b2} pixel is used to seed a region-growing operation that results in an 8-connected region of pixels R_{s2} that have grayscale values within a tolerance. The tolerances used to define connectivity of R_{s1} and R_{s2} were

found experimentally: for R_{s1}, the tolerance g_{b1}-0.05*dr was used, and for R_{s2}, the tolerance was g_{b2}-0.05*dr, where dr is the dynamic range (max – min grayscale value) in the image SB. While determining useful tolerance values that hold across a range of images can be at best problematic in the full resolution images, we were able to determine, with relatively few trials of varying the tolerance values and viewing the resulting connected regions, that the above values result in connected regions that satisfactorily correspond to skull and shoulder. The background region R_{bg} was determined by finding all pixels with grayscale value below a tolerance empirically determined by visual checks to produce satisfactory connected background regions across a range of images. (An absolute grayscale tolerance of 0.1 was used.) For each of the regions R_{s1}, R_{s2}, and R_{bg} the corresponding centers of gravity L_1, L_2, and L_{bg} were computed as landmarks.

(4) Classify the landmarks L_1 and L_2 as corresponding to approximate skull or shoulder regions. These landmarks were classified using four different methods. All classification was done in the low-resolution images (though one method uses a resolution one step finer than the SB images). For each method the landmarks were labeled "SK" or "SH", the labels were mapped to corresponding (x,y) coordinates on the full resolution image, and the resulting labeled images were viewed. The background landmarks were also labeled "BG" and similarly mapped to the full resolution images. We displayed the full resolution, labeled images and judged acceptability of the labeling. We considered a label "acceptable" if it clearly lay within the boundaries of the corresponding region on the full resolution image; in addition, we considered a skull or shoulder label acceptable if the resulting skull-shoulder line segment lay reasonably close in position and orientation to the spine, so that it appeared to be sufficiently accurate to position and orient a spine region template for initializing a search for the spine region. The last criterion is subjective: the real validation of the labeling algorithm comes in the application of the results to the problem of locating the spine region. Figures 9a-c show examples of the labeling. Figure 9c illustrates a case of the "SK" landmark being placed off the skull region, but we still judged its placement acceptable for getting an approximate spine position and orientation by using the line passing through "SH" and "SK" as a reference, along with the knowledge of the placement of the background landmark "BG". Classifying the L_1, L_2 landmarks as skull or shoulder: four methods were applied, and the resulting classifications for L_1, L_2 were compared for correctness. The first method used to classify the L_1 and L_2 landmarks is based on the observation that, for many full resolution images, the brighter pixels occur in the shoulder region. The second and third methods are based on the

observation that in the full resolution images, the grayscale variance in the shoulder region is almost always low as compared to the skull grayscale variance: these methods use variance as a discriminator. (Of course, we are operating on very low-resolution images where the bright and dark pixels lying near object boundaries in the full resolution images have been smeared together.) The fourth method uses the location of the background label BG to infer shoulder location. Details of these four methods used to classify L_1 and L_2 are as follows: (a) consider the brighter point of L_1, L_2 to be shoulder; (b) consider the point of L_1, L_2 with smaller grayscale variance in neighboring SB pixels to be shoulder; (c) the same basic idea as (b), but use the images at one finer level of resolution, to get a better estimate of variance in a region; each pixel in the image SB corresponds to a set of four pixels in the next highest resolution image (increased by a factor of 2 in both the x and y directions); map the SB regions R_{s1} and R_{s2} (which have centers of gravity L_1 and L_2, respectively) into the corresponding regions in the next higher resolution image; find the variance of each of these sets of pixels, and consider L_1 to be shoulder if the region corresponding to R_{s1} in the higher resolution image has the smaller grayscale variance, as compared to the variance of the region corresponding to R_{s2}; else consider L_2 to be shoulder; (d) for each image, use the background region R_{bg} to determine the "background corner": the image corner closest to the background region; once the background corner is found, determine the "shoulder corner" of the image: the image corner expected to be closest to the shoulder region; finally, consider L_1 to be shoulder if it lies closest to the shoulder corner; else, consider L_2 to be shoulder.

Results of the automatic landmarking are given in Table 3 for a set of 48 cervical spine images, and show that the classifications obtained by methods (a) and (d) yielded the identical results of 46 acceptable skull classifications and 48 acceptable shoulder classifications. Method (c) yielded 46 acceptable skull classifications and 47 acceptable shoulder classifications. Method (b) relies on measuring the variance of the image SB in a 3x3 region centered at the candidate skull (shoulder) point, and the relatively poor performance of this method is not surprising, considering the severe blurring and subsampling used to produce SB. What is somewhat surprising is that, as shown by the performance of method (c), sufficient variance information is preserved in the next highest resolution image (of size only 12x14) to discriminate between skull and shoulder. It should be noted that method (a) relies on comparing the differences between grayscale values; in some cases the differences between these values may be quite small. A similar comment may be made about the comparison of the variances in method (c). This suggests

that in a larger test set, some failures may occur with these methods due to outlying cases where the skull grayscale (variance) is in fact greater than (less than) the shoulder grayscale (variance) in the low resolution images. Method (d), however, only relies on the suppositions that we can properly classify the background behind the skull, that we can identify the corresponding image corner closest to this background region, and, using this information, that we can identify the image corner where the shoulder is expected to lie. The shoulder point is then classified as the candidate point closest in Euclidean distance to this "shoulder corner" of the image. In all of the observed cases, the background region, background corner, and shoulder corner were reliably identified, and the differences in the Euclidean distance between each candidate point and the shoulder corner are much larger than the differences used to discriminate between the candidate points in methods (a) and (c). Hence, we might expect method (d) to prove to be the most reliable method of classifying the landmarks. The two failure cases in skull labelling for methods (a), (c) and (d) were for the same two images. Both of these images show strong leakage of light along the bottom border of the original x-ray film, which in each case results in one of the candidate landmarks which should have been placed in the skull region, being placed along the bottom border. The fact that method (a) was able to correctly label the shoulder landmark in both of these problem images, and that method (c) correctly labeled the shoulder landmark in one of these images, can be attributed to chance, since only method (d) has a rational basis for identifying the shoulder landmark in such a case. The brightness and size of these borders suggest that a method to automatically detect and remove them should be achievable and straightforward. One of these failure cases is shown in Figures 10a-b, where Figure 10a shows the labeling obtained when the classification method (c) was used, and Figure 10b shows the labeling when method (d) was used. In each case, there is no acceptably placed skull landmark, due to the influence of the bright border region. Method (d) however is still able to properly label one of the landmarks as "SH" (shoulder) since, for shoulder landmarking, method (d) only relies on finding the landmark closest to the "shoulder corner" of the image.

5.2 Using the landmarks to locate the spine

The immediate use of the landmarks is to localize the spine region. We have proceeded by using the landmarks to estimate important geometrical characteristics of this region of interest, for the eventual purpose of vertebrae segmentation, labeling, and classification.

5.2.1 Motivating the search for image surface curves in the spine region

Figure 11a is a surface plot of the Figure 8a image (reduced in spatial resolution horizontally and vertically by a factor of 64, for the purpose of looking at the large-scale surface characteristics). Figure 11a illustrates some of the characteristics of the surface corresponding to the image of the spine region: (1) the spine is a ridge in the data visually distinguishable from the background, but with less prominence than either the skull or shoulder regions and (2) for most of its extent, the spine has a visually distinguishable slope; this slope is observable in Figure 11a on the front of the spine; on the back of the spine (not observable in this figure), the slope is even more prominent. The slope on the front side tends to be less distinguishable for the upper part of the spine, where image data for the jaw is in close proximity, but it is quite prominent at the lower front of the spine, where the image data frequently has an apparent "valley" that shows up on the image (Figure 8a) as a prominent dark area in the throat region. In some images, instead of this "valley" region (where the surface has local minima), the surface data merely appears to approximately flatten in this region, without exhibiting local minima. These observations, repeated over some dozens of images, suggested that four curves, denoted C_1-C_4, respectively, as candidates for detection in these images. (Our terminology C_1, C_2, etc. for these curves should not be confused with the vertebral designations C1, C2, etc. for the cervical spine used earlier.) These curves are (1) C_1: a boundary curve for the back edge of the spine, separating spine from the very dark background; Visual inspection of images suggested that the C_1 curve is a very good candidate for a detectable feature, owing to the high contrast, running essentially the entire length of the spine, between the back-of-spine background, and the spine itself. It should be noted, of course, that detecting C_1 in the smeared images yields a curve that would not be expected to mark a significant part of the boundary of any anatomical object in the original unsmeared images, since the original images show the protruding, separated bodies of the spinous processes on the back of the spine, while these bodies form a largely homogeneous, bright mass in the smeared images. At best, C_1 as detected in the smeared images could be expected to be an envelope curve that is approximately tangent to the extremities of the spinous processes; (2) C_2: a curve following the prominent ridge within the spine itself; like C_1, C_2 appears to be a good candidate for detection, owing to the visually recognizable ridge characteristics in the observed spines, usually extending the length of the spine, but particularly visible in the region of the lower vertebrae. Note that one definition for curvature of the spine might be taken as the curvature

defined by a curve that is fit to the midpoints of the top and bottom of each vertebral body. Unfortunately, these midpoints do not have any prominent associated visual characteristics, either in grayscale or in shape, and are very poor candidates for detection until a later stage of processing when the spine anatomy is already known at a finer grained level. However, the grayscale ridge points, i.e., the points on the C_2 curve, appear to lie on the vertebral faces, or near the right edges of the vertebral faces, and hence C_2 might be conjectured to have approximately the same curvature as a curve fit to the vertebral upper and lower midpoints. Thus the curvature of C_2 might be conjectured to give a reasonable approximation to the spine curvature; (3) C_3: a boundary curve following the front edge of the spine, separating the spine from non-spine bright tissue, or from background; and (4) C_4: a curve lying in the prominent valley in front of the spine or, when this valley is not present, lying in the approximately flat region in this area. Curves C_3 and C_4 are more problematic, since neither the front edges of the vertebrae nor the dark valley in front of the vertebrae are prominent features along the entire spine length. In an original resolution image, it is frequently observed that the front vertebral edge adjoins a region of tissue that appears to be a relatively bright, narrow strip that merges with this edge. Ideally, we would be able to detect a curve C_3 that would mark the boundary between the vertebrae and this tissue; and, in fact, this is a goal of the final stage of the segmentation of these vertebrae. However, in the smeared images, the vertebrae front edges and the background tissue are no longer distinct. C_3 for these images may be expected to be a curve that follows the general shape of the front of the spine, with best agreement with that shape for the lower spine vertebrae area. Likewise, the valley used to define C_4 in the smeared images yields only an approximation of the valley curve in the original image, and, again, the greatest agreement with the general shape of this curve in the original images is expected in the lower vertebrae regions.

Examples of these curves, denoted C_1-C_4, respectively, and hand-drawn for illustration, are shown in Figure 11b, superimposed over a smeared version of the Figure 8a image. Capability to compute this curve data robustly would provide a means of estimating spine region "side" boundaries (C_1 and C_3/C_4) and of estimating a vertebrae-traversing curve running within the body of the spine by using C_2, or a curve computed from some subset of the $\{C_i\}$.

5.2.2 Detecting the curves

We formulated the general problem of the detection of these C_1-C_4 curves as an optimization problem, as follows. Denote the image we are searching by I=I(x,y); then, for each i, let C_i be parameterized by t, so that $C_i=(x_i(t),y_i(t))$. Intuitively, we expect that each C_i is constrained to lie within a particular subregion of the image; for example, C_2 by definition lies along the brightest grayscale within the spine region and therefore within the body of the spine. Conceptually, this may be formalized by defining for each C_i a feasible search region D_i that constrains the location of the solution curve C_i; i.e. C_i must satisfy $C_i \subset D_i \subset I$. In the most general formulation, we would find the C_i by minimizing an objective function $J(C_1,C_2,C_3,C_4)$, where the form of J would include coupling between the C_i, and also include constraints to insure that the solution curves C_i have reasonable geometries for curves representing spine boundaries. The objective would be to minimize J over all curves C_i satisfying $C_i \subset D_i$.

To achieve a simpler and implementable result, we posed the more limited problem of finding the optimal C_i individually for each i, by formulating separate objective functions, each with simple form and each decoupled from the others. That is, for each i, we defined an objective function $J_i=J_i(C_i)=J_i((x_i(t),y_i(t)))$. Note that for a given curve C_i in the (x,y) plane, we may denote the corresponding image grayscale profile by $I(C_i)=I(C_i(x_i(t),y_i(t))$, and the corresponding image grayscale gradient, evaluated along the curve C_i, by $\nabla I(C_i)=\nabla I(C_i(x_i(t),y_i(t)))$. Then, for each i, we seek to minimize the corresponding objective function, i.e., we seek a solution curve $\hat{C}_i$ such that $J_i(\hat{C}_i) \le J_i(C_i)$ for all $C_i \subset D_i$ that satisfy reasonable constraints on the C_i properties.

For the individual i, we defined these specific objective functions:

$$J_1(C_1) = -\int |\nabla I(C_1)| + \lambda \int ds \quad (1)$$

$$J_2(C_2) = -\int I(C_2) + \lambda \int ds \quad (2)$$

$$J_3(C_3) = -\int |\nabla I(C_3)| + \lambda \int ds \quad (3)$$

$$J_4(C_4) = \int |\nabla I(C_4)| + I(C_4) + \lambda \int ds \quad (4)$$

where ds is arc length along C_i, and for each i, C_i is constrained to lie in D_i. It should be noted that our description of the constraints on the C_i omits features that should be incorporated into a more complete model: for example, the

C_i should be approximately the same lengths, hence, the begin and end points for each C_i should be constrained to lie within pre-specified regions.

5.2.3 Motivating the choice of integrands

The choice of integrands in equations (1)-(4) was made with both theoretical considerations and experimental observations. Figure 12 illustrates the expected grayscale characteristics along normals to the spine axis (i.e. approximate normals to the solution curves C_i). From equations (1)-(4) it can be seen that, for the cases where we are seeking edges (i=1,3), then, along these normals, we maximize the absolute gradients; for the cases where we are seeking a ridge (i=2) we maximize the grayscale value itself. For case i=4, we found that for some images, there is a true "valley" (local minimum) in the D_4 region; for these cases, minimizing the grayscale itself ($I(C_4)$) worked well; in other cases, however, lines transverse to the spine that pass through D_4 have monotonically decreasing grayscale, with only an approximate point of inflection instead of a valley. For these cases, having the absolute gradient also in the integrand was necessary. For the final formulation, we integrated the sum of absolute gradient and grayscale. In each of the cases, we penalize this minimization by the length of the minimizing curve, in order to avoid solutions where the curve length grows arbitrarily large. The multiplier λ allows for different weighting of the two terms, for example, in order to compensate for different scales of measurement being used for curve length and for gradient measurements.

5.2.4 Setup for the discrete problem

We denote the discretized versions of C_i, D_i, and J_i with a superscript asterisk (*) on each term. For solving for each of the C_i*, the general approach is the same: (1) determine a (non-rectangular, in general) grid D_i* which will be the discrete version of the feasible solution region; and (2) minimize the objective function J_i*(.) over all paths on this grid D_i*. The minimizing path is the desired solution C_i*. For step 1, the method used is heuristic and requires the use of experimentally determined parameters to determine D_i* based on expected characteristics of the image grayscale. This is illustrated in Figure 13 for D_1* and discussed in detail in Section 5.2.5. For step 2, the method of dynamic programming is used; this enables the global minimization of J_i* to be achieved over the grid D_i* within a

reasonable amount of computation. The solutions were computed on the heavily smeared images, but at full spatial resolution (1462x1755). The images were smeared with a 100x100 Gaussian filter with standard deviation set to 100. The J_i parameter λ of equations (1)-(4) was set to 0.05 to compensate for the different measurement scales between gradient and distance in the images. For each image, a boundary area of 60 pixels was used on all four image sides (left, right, top, bottom); pixels that were within this boundary limit of an edge were not processed. This was done to avoid the frequent problems encountered by the presence of very bright pixels due to light leakage near the edges of the images.

5.2.5 Determining the D_i*

The determination of each of the D_i* (i=1,4) uses a common geometrical framework. First, we use the landmarks results of Section 5.1 to define a line segment S taken to be the tentative "spine axis". At this stage of the processing, this initial "spine axis" is only expected to lie near enough to the actual spine, with a similar enough orientation, so that lines transverse to S will cut the actual spine at an approximately perpendicular angle. We obtain S by connecting the points defined by the "skull" and "shoulder" labels (S_{sk} and S_{sh}, respectively) of the Section 5.1 outputs.

The transverse lines M_k (again, see Figure 13) were then defined with the following constraints: (1) M_k is normal to S for all k; (2) M_1 intersects S at $S_{sk;}$ (3) M_{kmax} intersects S at $S_{sh;}$ and (4) the M_k are uniformly spaced along S. After some experimentation, k_{max} was taken to be 20, but no determination was made of an optimal value. The M_k were taken to span the entire width of the image, except for the image border areas noted above.

In Figure 13, the lines M_k are seen to be normal to the "spine axis" S that connects the skull and shoulder landmarks. To determine any of the D_i*, we proceed through the lines M_k one by one, determining which points belong to D_i* on the current line before continuing to the next line. When operating on a particular line, the process is the same for each of the D_i*: (1) determine bounds on the interval on the current line expected to contains points of D_i*, and then (2) sample this interval for the points to actually collect for D_i*. Figure 13 illustrates the particular case of determining the grid for D_1* (the search region for C_1*). For this specific example, we proceed along a particular

line M_k: first we search the grayscale profile along this line, proceeding from the back-of-skull region toward the front-of-skull region, and, in informal terms, "look for the first occurrence of a large increase in grayscale that occurs over a large spatial span"; i.e., we want to find the approximate point where the grayscale begins climbing rapidly from the very dark background values to the relatively bright spine values as we move along the line from the back-of-skull region toward the front-of-skull region. More precisely, we find the first point **p** such that

$$I(\mathbf{p} - \Delta_1 * \mathbf{u}) - I(\mathbf{p}) >= f * r$$

where

p = p(x,y) = point we are searching for along a particular M_k,

I(**p**) = grayscale value of the image I at point **p**,

u = unit vector parallel to the M_k segments, pointing toward back-of-skull region; note that we know the orientation of u from the skull background landmark BG of Section 5.1;

Δ_1 = half-length of interval around **p** used to define the interval containing D_1* points on a particular M_k,

r = grayscale dynamic range along a particular M_k,

f = threshold control factor, $0 < f <= 1$; together, f and r define a threshold for detecting **p**;

The detected point **p** is taken to be a rough indicator or landmark of the center of the interval of points that belong to the D_1* region and that lie on the current M_k line; we then sample an interval on the current M_k, centered around this landmark point, to get the points from this line for D_1*. In the figure, the sampling interval is determined by the vectors $\mathbf{b_1}$ and $\mathbf{b_2}$. This process is repeated for each of the M_k lines. In general, the resulting set of points forms a non-rectangular grid. This grid of points is taken to be D_1*, the discretized search region for the solution curve C_1*.

Similar methods were used to determine the grids D_2*-D_4*. A summary of the heuristics for finding the D_i*, including parameter values used, was given in previous work[35].

5.2.6 Solving for the C_i*

The method for solving for the C_i* by minimizing the J_i* is illustrated in Figures 14a-b, for J_1*. As previously noted, we implement D_1 as the non-rectangular grid D_1* of nodes placed on the lines M_k that are normal to the

approximate "spine axis". Each node has an assigned value $g_{k,r}$ (gradient magnitude) and transition cost $e_{k,r,s}$ (Euclidean distance) for moving from node (k,r) to node (k+1,s). To every path C consisting of an ordered set of nodes

$(g_{1,j1}, g_{2,j2},\ldots,g_{kmax,jkmax})$, one per line, beginning at M_1 and ending and M_{kmax}, we assign an associated cost:

$$J_1^*(C) = -\sum_{i=1}^{k\ max} g_{i,j_i} + \lambda \sum_{i=1}^{k\ max\ -1} e_{i,j_i,j_{i+1}}$$

J_1* is the objective function that we minimize over all possible paths on the grid D_1*. The illustration in Figure 14a suggests the combinatorial expense of directly computing J_1* for all paths. By casting this as a dynamic programming problem, however, the computations are made feasible within practical computation time. We take the minimizing path

$$C_1^* = \arg\min J_1^*(C)$$

to be the discrete, computable solution corresponding to the ideal spine region curve C_1.

It should be noted that some of the grids D_i* have a dependence on the optimal solution points for other D_j*, j≠i. For example, on any line, the points in D_2* must lie "to the left" (toward the front of the spine) of the optimal point for D_1* on that line. This is a way of specifying the observable fact that the spine ridge point on a line must lie "to the left" of the back-of-spine edge point and means, in practical terms, that we must solve for the back-of-spine curve C_1* before we solve for the spine ridge curve C_2*, for example. The illustration in Figure 14b shows a particular solution D_1* that has resulted from applying dynamic programming to find the path that minimizes J_1* among all possible paths on the D_1* grid .

5.2.7 Results of computing the C_i*

We computed the C_i curves for all 46 of the 48 test images for which the basic landmark data of Section 5.1 was obtainable. Figures 15a-b show examples of the results of the computation of the curves, overlaid on the original resolution images (although the curves were computed from the smeared image data). We evaluated the results by displaying and informally inspecting each of the four curves for each of the 46 images, and by using the C_2* curves

to estimate orientations of the spine which we compared to an independent data source. From the visual inspection of the results, we concluded that, when overlaid on the original unsmeared images, (1) the C1* curves tended to fall along the faint boundary separating the non-bony tissue along the back of the neck from the x-ray background; (2) the C_2* curves strongly tended to follow the spine curvature, and tended to lie along the rightmost edge of the vertebral faces; (3) the C_3* curves tended to lie along the vertebrae front edges, for the lower spine vertebrae, but sometimes fell close to the C_2* curves for the upper vertebrae; (4) similarly, the C_4* curves tended to fall in the dark tissue area in the throat area for the lower vertebrae, but also tended to fall close to the C_2* curves for the upper vertebrae. Overall, the visual inspection supported the hypothesis that the C_2* curves were the most predictably correlated with the spine anatomy over the length of the cervical spine. To obtain a quantifiable performance measure, we computed linear fits to each of the C_2* curves. Figure 16 illustrates linear fits computed for all of the C_i* curves of Figure 15a. For each of the 46 images in our test set we used the expert-collected (x,y) coordinates of Section 3.2 (illustrated in Figure 2) for comparison data. For each image we took the top and bottom midpoints (points 2 and 5) for each vertebra, and used them to fit a straight line to the spine for that image. This procedure gave two straight lines for each image, one that was fit to our C_2* curve, and one that was fit fit to the manually-collected point 2 and 5 coordinates. For each of the two straight lines, the slope angle, which may be taken as an estimate of spine orientation, was computed and the absolute difference taken. The results are plotted in Figure 17 and show, that, with one exception, the difference in spine orientations between the two methods was less than 15 degrees. Areas for improvement in this work include not only possible improvement in the heuristics for computing the D_i*, but perhaps also improvement in the integrands used in the modeling of the objective functions J_i of equations (1)-(4), in particular for the case J_4, where the image grayscale characteristics are more variable and less understood than in the other cases.

Our next steps will be to evaluate the use of the obtained C_i data for spine template placement, vertebrae counting, and obtaining intervertebral boundary marks. All algorithms were implemented in MATLAB 5.3 and executed on a 400 MHz PC.

6 Evaluation

Evaluation of technology for the content-based retrieval of biomedical images entails the special consideration of the need for the involvement of subject matter experts, both in the evaluation of the retrieval precision and in the evaluation of the efficiency of the retrieval as compared to conventional methods. Further complications are the inter- and intra-observer variabilities that occur in the execution of any protocol of evaluations done by humans; this is particularly true in the medical domain. In addition, prior to the evaluation of the final retrieval system, evaluations of component subsystems may be required: the classifiers for the various biomedical features of anterior osteophytes, disc space narrowing, and subluxation may be considered individual subsystems which will each require their own evaluation with the collaboration of medical experts. For our work to date, as well as the work of Stanley and Zamora, evaluation has been done by using reference data collected by one expert radiologist. More comprehensive evaluation of component subsystems and of the final retrieval system, is expected to require the assistance of multiple medical experts and a carefully-designed study that will allow comparison of the computer-assisted results with the results generated by humans, taking into account both inter- and intra-observer variances.

7 Emerging multimedia data standards and content based medical image retrieval

Much interest has been generated in the multimedia community by the emerging MPEG-7 standard. MPEG-7, formally called "Multimedia Content Description Interface", is intended to be a "standardized description of various types of multimedia information"[36], including "graphics, still images, video, film, music, speech, sounds, text and any other relevant AV medium"[36]. MPEG-7 is intended to provide a description of multimedia information that facilitates rapid search by content. Although a thorough critique of the MPEG-7 work as it impacts the human and software accessibility of digitized spine x-rays is beyond the scope of this paper, we may make some observations, based on the published MPEG-7 documents. First, the MPEG-7 work explicity takes into account the need for ready interchange and access to biomedical images among clinicians and researchers[37]. Second, support for key capabilities such as spatial searching, including limited region-of-interest searching, shape representation, and multiple levels of image representation (such as multiple spatial resolution) are supported; both 2-D and 3-D content representations are supported[36,37]. Third, the MPEG-7 work is generic in the sense that it is intended to provide a broad framework of content description, but does not create a standard for the representation of any particular

application's data. MPEG-7 includes both a standard set of Descriptors for describing features, and a Descriptor Definition Language (DDL) for creating user-defined Descriptors, beyond the standard set. Relationships among the features described are defined by a Description Schema[36]. A key issue for creators of biomedical image applications is whether the standard Descriptors, the DDL, and the Description Schema are comprehensive enough to allow for effective representation of the variable and irregular shapes, object geometries, ill-defined and/or occluded boundaries, and vagaries of content interpretation. For maximum interoperability in use of spine x-rays among user groups wanting to do content-based retrieval, it is desirable that the standard features defined in MPEG-7 include the features important for this particular image type, subject to the need in MPEG to maintain broad-based applicability. Current MPEG-7 features defined for images include grid layout for enabling rectangular subimage descriptions, bounding box definition for making local regions with arbitrary orientations addressable, and shape description using both region-based and curvature scale space representations[38].

8 Summary

The problem of automating the indexing and retrieval of a large collection of digititized x-ray film for biomedical features of interest requires the integration of robust methods for obtaining basic landmarks in the images, segmenting to the individual vertebra level with good enough fidelity to preserve the distinctive shapes of irregular corner features, and classification of the vertebra by anatomical label and by degree of pathology. All this must be accomplished within a projective data modality where edge boundaries may be ambiguous, and a data quality that commonly allows for edge segments to be cloudy or even not visible. We have shown progress in the fundamental initial tasks of obtaining basic landmark data in the images, using this to estimate first-order features of the spine, and (in previous work) by ourselves and others it has been shown that the prospects for obtaining automated or computer-assisted solutions to this indexing and retrieval task warrant continued effort and investigation.

Many of the resources in this paper, including the WebMIRS system, the Digital Atlas of the Spine, the cervical and lumbar spine images, and the radiologist coordinate data, are available via the Web at NLM Communications Engineering Branch Web site: http://archive.nlm.nih.gov

Acknowledgments

The authors acknowledge the very helpful ideas, examples, criticism, and code that is continuing to be provided by Dr. Hemant Tagare of Yale University in the area of using deformable templates to find structures in x-ray images. We also acknowledge the work of Matthew Freedman, M.D. and Dr. Ben Lo of Georgetown University in providing significant biomedical and quantitative data from 600 digitized x-ray images for the continuing research into developing effective image processing techniques for automated image analysis for biomedical applications. Finally, we are grateful to David Blum of the National Library of Medicine Audiovisual Program Development Branch and to branch chief Joe Fitzgerald, for the anatomical artwork used in Figure 3.

References

1. Long LR, Pillemer SR, Lawrence RC, Goh G-H, Neve L, Thoma GR. World Wide Web platform-independent access to biomedical text/image databases. *Proceedings of SPIE Medical Imaging 1998: PACS Design and Evaluation: Engineering and Clinical Issues,* SPIE Vol. 3339, San Diego, CA, February 21-26, 1998, pp. 52-63.
2. Long LR, Pillemer SR, Lawrence RC, Goh G-H, Neve L, Thoma GR. WebMIRS: Web-based Medical Information Retrieval System. *Proceedings of SPIE Storage and Retrieval for Image and Video Databases VI*, SPIE vol. 3312, San Jose, CA, January 24-30, 1998, pp. 392-403.
3. Tagare HD, Jaffe CC, Duncan J. Medical image databases: a content-based retrieval approach. *JAMIA*, 1997(4), pp. 184-198.
4. Digitized Radiographic Images: Challenges and Opportunies. Chaired by: Stanley R. Pillemer, M.D. June 2-3, 1993, Bethesda, MD, jointly sponsored by the National Institute of Arthristis and Musculoskeletal and Skin Diseases (NIAMS), the National Center for Health Statistics, and the National Library of Medicine. Digitized Radiographic Images. Chaired by Stanley R. Pillemer, M.D., Reva C. Lawrence, M.P.H.. May 9, 1995, Bethesda, MD, sponsored by NIAMS.

5. Rosol MS, Cohen GL, Halpern EF, Chew FS, Kattapuram SV, Palmer WE, Dupuy DE, Rosenthal. Vertebral morphometry derived from digital images. *American Journal of Roentgenology*, Vol. 167, December 1996, pp. 1545-1549.
6. Gallagher JC, Hedlund LR, Stoner S, Meeger C. Vertebral morphometry: normative data. Bone and Mineral, Vol 4., 1988, pp. 189-196.
7. Hedlund LR, Gallagher JC. Vertebral morphometry in diagnosis of spinal fractures. *Bone and Mineral*, Vol. 5, 1988, pp. 59-67.120
8. Ross PD, Davis JW, Epstein RS, Wasnich RD. Ability of vertebral dimensions from a single radiograph to identify fractures. *Calcified Tissue International*, Vol. 51, 1992, pp. 95-99.
9. Smith-Bindman R, Cummings, SR, Steiger P, Genant HK. A comparison of morphometric definitions of vertebral fracture. *Journal of Bone and Mineral Research*, Vol. 6, No. 1, 1991, pp. 25-34.
10. Chwialkowski MP, Shile PE, Pfeifer D, Parkey RW, Peshock RM. Automated localization and identification of lower spinal anatomy in magnetic resonance images. Computers and Biomedical Research, vol. 24, 1991, 99-117.
11. Cootes TF, Taylor CJ. Statistical models of appearance for computer vision. Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, September 13, 1999. Draft report, available at http://www.wiau.man.ac.uk.
12. Smyth PP, Taylor CJ, Adams JE. Vertebral shape: automatic measurements with active shape models. Radiology, May 1999, vol. 211, no. 2, pp. 571-578.
13. Gardner JC, Heyano SL, Yaffe LG, von Ingersleben G, Chesnut CH III. A semi-automated computerized system for fracture assessment of spinal x-ray films. Proceedings of SPIE Medical Imaging 1996: Image Processing, SPIE vol. 2710, February 12-15, 1996, Newport Beach, CA, pp. 996-1008.
14. Gardner JC, Yaffe LG, Johansen JM, von Ingersleben G, Chesnut CH III. Problems with six point vertebral morphometry. *Proceedings of SPIE Medical Imaging 1998: Image Processing*, SPIE vol. 3338, San Diego, CA, February 23-26, 1998, pp. 107-117.
15. Efford ND. Knowledge-based segmentation and feature analysis of hand-wrist radiographs. University of Leeds School of Computer Studies Report 94.31, October 1994. Available at: s.ac.uk/comir/research/hands/hands.html

16. Pietka E, Kaabi L, Kuo ML, Huang HK. Feature extraction in carpal-bone analysis. *IEEE Transactions on Medical Imaging*, vol. 12, no. 1, March 1993, pp. 44-49.

17. Weiler F, Vogelsang. Model-based segmentation of hand radiographs. *Proceedings of SPIE Medical Imaging 1998: Image Processing*, SPIE vol. 3338, San Diego, CA, February 23-26, 1998, pp. 673-682.

18. Duryea J, Jiang Y, Countryman, Genant HK. A fully automated algorithm for the segmentation of the middle and proximal phalanges of digitized hand radiographs. *Proceedings of SPIE Medical Imaging 1998: Image Processing*, SPIE vol. 3338, San Diego, CA, February 23-26, 1998, pp. 1036-1046.

19. Long LR, Thoma GR. Use of shape models to search digitized spine x-rays. Proceedings of IEEE Computer Based Medical Systems 2000, Houston, TX, June 23-24, 2000, pp. 255-260.

20. Long LR, Thoma GR. Segmentation and image navigation in digitized spine x-rays. Proceedings of SPIE Medical Imaging 2000: Image Processing. Vol. 3979, San Diego, CA, February 12-18, 2000, pp. 169-179.

21. Sari-Sarraf H, Mitra S, Zamora G, Tezmol A. Customized active shape models for segmentation of cervical and lumbar spine vertebrae, Texas Tech University College of Electrical Engineering technical report, available at http://www.cvial.ttu.edu/~sarraf

22. Zamora G, Sari-Sarraf H, Mitra S. Estimation of orientation of cervical vertebrae for segmentation with active shape models. Proceedings of SPIE Medical Imaging 2001: Image Processing, San Diego, CA, February 17-23, 2001.

23. Stanley RJ, Long R. A radius of curvature approach to cervical spine vertebra image analysis. Proceedings of the 38th Annual Rocky Mountain Bioengineering Symposium, Copper Mountain, Colorado, April 20-22, 2001.

24. Plan and operation of the second National Health and Nutrition Examination Survey, 1976-1980, DHHS Pub. No. (PHS) 81-1317, Vital and Health Statistics Series 1, No. 15, Public Health Service, Stock Number 017-022-00752-5, U.S. Government Printing Office, Washington, DC 20402.

25. National Health and Nutrition Examination Survey, 1976-80 (NHANES II), Public Use Data Tape Documentation, Physician's Examination, Tape Number 5302, National Center for Health Statistics, Hyattsville, MD.

26. Chuang K-S, Liu BJ, Huang HK, Yonekawa H, Sankaran A. Noise contnet analysis in clinical digital images. Radiographics, vol. 14, no. 2, March 1994, pp. 397-405.

27. Chuang K-S, Huang HK. Assessment of noise in a digital image using the join-count statistic and the Moran test. Physics in Medicine and Biology [check title], vol. 37, no. 2, 1992, pp. 357-369.

28. Lo SC, Krasner B, Num SK. Noise impact on error-free image compression. IEEE Transactions on Medical Imaging, vol. 9, 1990, pp. 202-205.

29. *Stedman's Medical Dictionary*, 26th edition, Williams & Wilkins, Baltimore, MD, 1995.

30. Niblack W, Zhu X, Hafner JL, Breuel T, Ponceleon D, Petkovic D, Flickner M, Upfal E, Nin SI, Sull S, Dom B, Yeo B-L, Srinivasan S, Zivkovic D, Penner M. Updates to the QBIC system, Proceedings of SPIE Storage and Retrieval for Image and Video Databases VI, vol. 3312, 28-30 January 1998, San Jose, CA, 150-161.

31. Niblack W, Barber R, Equitz W, Flickner M, Glasman E, Petkovic D, Yanker P. The QBIC project: querying images by content using color, texture, and shape, Proceedings of SPIE Storage and Retrieval for Image and Video Databases 1993, February 1993, vol. 1908, 173-187.

32. Long LR, Pillemer S, Goh G-H, Berman LE, Neve L,Thoma GR, Premkumar A, Ostchega Y, Lawrence R, Altman RD, Lane NE, Scott WW, Jr. A digital atlas for spinal x-rays. *Proceedings of SPIE Medical Imaging 1997: PACS Design and Evaluation: Engineering and Clinical Issues*, SPIE vol. 3035, pp. 586-594, Newport Beach, CA, February 22-28, 1997.

33. Liang Q, Wendelhag I, Wikstrand J, Bustavsson. A multiscale dynamic programming procedure for boundary detection in ultrasonic artery images. IEEE Transactions on Medical Imaging, vol. 19, no. 2, February 2000, pp.127-142.

34. Mignotte M, Meunier J. A multiscale optimization approach for the dynamic contour-based boundary detection issue. Computerized Medical Imaging and Graphics, vol. 25, no. 3, 2001, pp. 265-275.

35. Long LR, Thoma GR. Identification and classification of spine vertebrae by automated methods. Proceedings of SPIE Medical Imaging 2001: Image Processing, San Diego, CA, February 17-23, 2001.120

36. MPEG-7 Context, Objectives and Technical Roadmap. International Organisation for Standardisation, July 1999, Vancouver. http://www.darmstadi.emd.de/mobile/MPEG7/Documents/W2861.htm

37. MPEG-7 Applications Document v.10. International Organisation for Standardisation, January 2001, Pisa. http://www.darmstadt.gmd.de/mobile/MPEG7/Documents/W3934.htm

38. Overview of the MPEG-7 standard. International Organisation for Standarisation, March 2001, Singapore. http://www.cselt.stet.it/mpeg/standards/mpeg-7/mpeg-7.htm

Figure Captions

Figure 1. WebMIRS results screen.

Figure 2. Data points were expert-collected for points 1-9. Example geometries for distances shown in Table 2 are given also.

Figure 3. Nomenclature for vertebrae.

Figure 4a. A C4 vertebra and neighboring grayscale regions. Some of the grayscale interior to C4 is distributed similarly to the grayscale external to the C4 vertebral body, in the C3/C4 disc space.

Figure 4b. Surface plot of the grayscale in the Figure 4a image, illustrating the similarity in interior and exterior grayscale for C4.

Figure 5. Cervical spine images with graded anterior osteophytes. The severity progresses from 0 (normal) to 3 (most abnormal) from left to right in the images.

Figure 6. Vertebrae showing the osteophytes on lower anterior corner of center vertebra (grade 0 is normal—no osteophyte).

Figures 7a-d. Surface plots of the vertebrae with boundary outlines

Figure 8a. Cervical spine image

Figure 8b. Cervical spine image heavily smeared and subsampled

Figures 9a-c. Results of basic landmark labeling

Figures 10a-b. Errors in labeling due to heavy light leakage. (a) shows labeling results using method (a); (b) shows results using method (d), which gives correct shoulder labeling.

Figure 11a. Surface plot of a cervical spine x-ray.

Figure 11b. Conceptual curves C_1-C_4, drawn on the smeared image of Figure 11a.

Figure 12. For points on any particular C_i, the neighboring grayscale curves along intersecting line segments that are normal to the spine axis have the expected characteristics shown above. Note that the above curves are not the C_i themselves, but the expected grayscale characteristics along line segments that intersect points on the C_i.

Figure 13. Illustration of heuristic for determining D_1*, the discretized search region for C_1*.

Figure 14a. The grid D_1* for searching for C_1*, with node and transition costs illustrated.

Figure 14b. Illustration of a discrete solution C_1* on the grid D_1*.

Figures 15a-b. Example C_1-C_4 solution curves (two different images).

Figure 16. Example of linear fits from C_1-C_4 solution curves

Figure 17. Comparison of spine angles derived from C_2 to independent estimates

Table Captions

Table 1. Global cervical spine image characteristics

Table 2a. Cervical spine interveterbal characteristics

Table 2b. Cervical spine within-vertebral characteristics

Table 3. Results of classifying landmarks by four methods

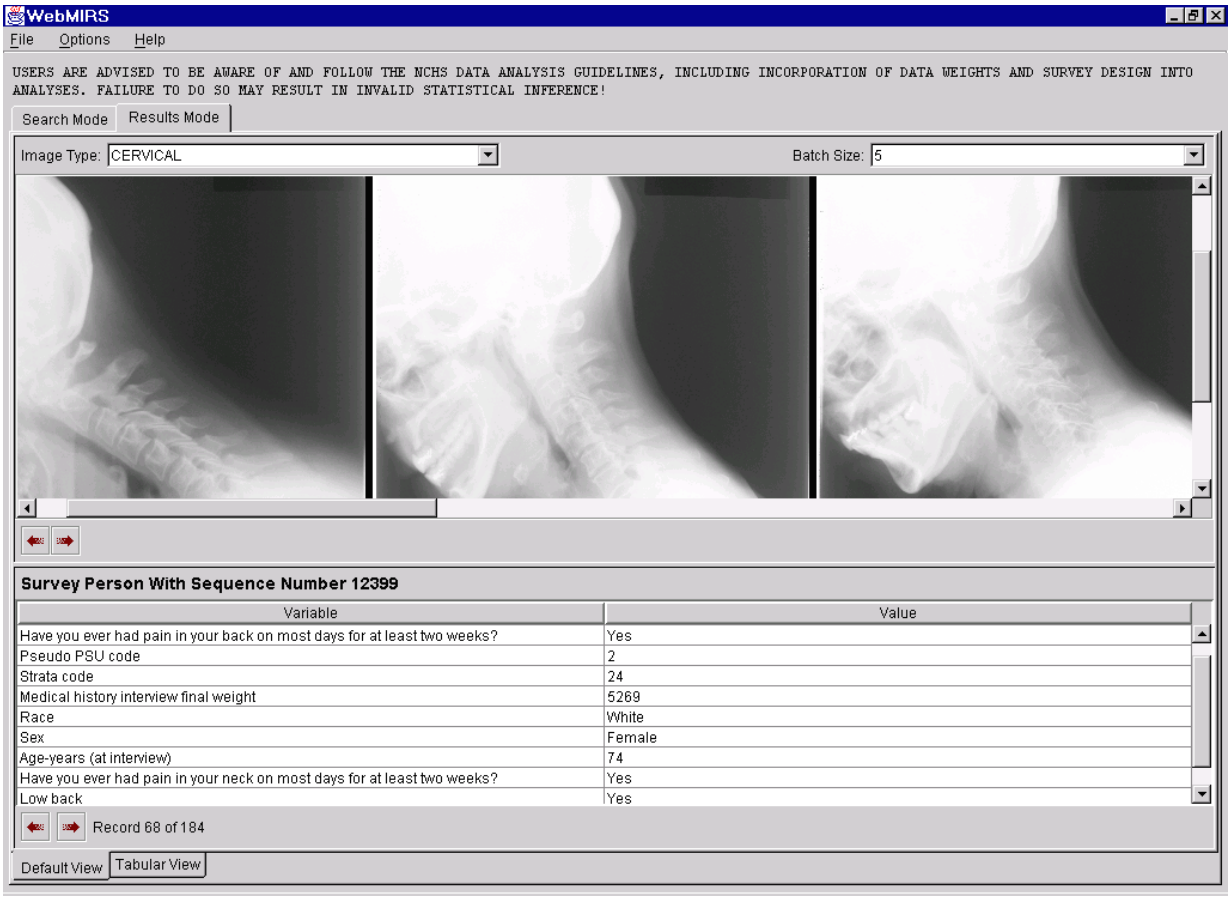


Figure 1. WebMIRS results screen.

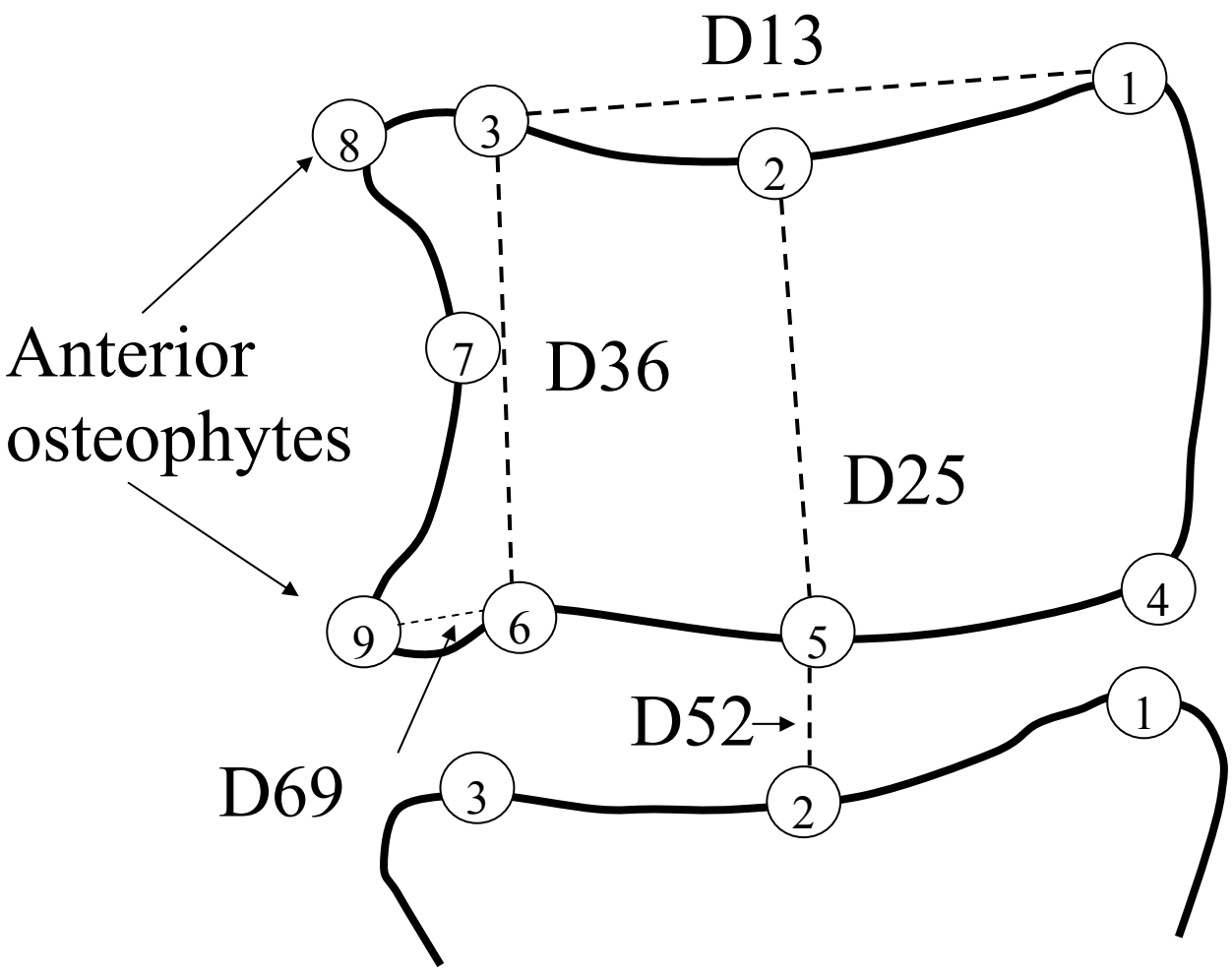


Figure 2. Data points were expert-collected for points 1-9. Example geometries for distances shown in Table 2 are given also.

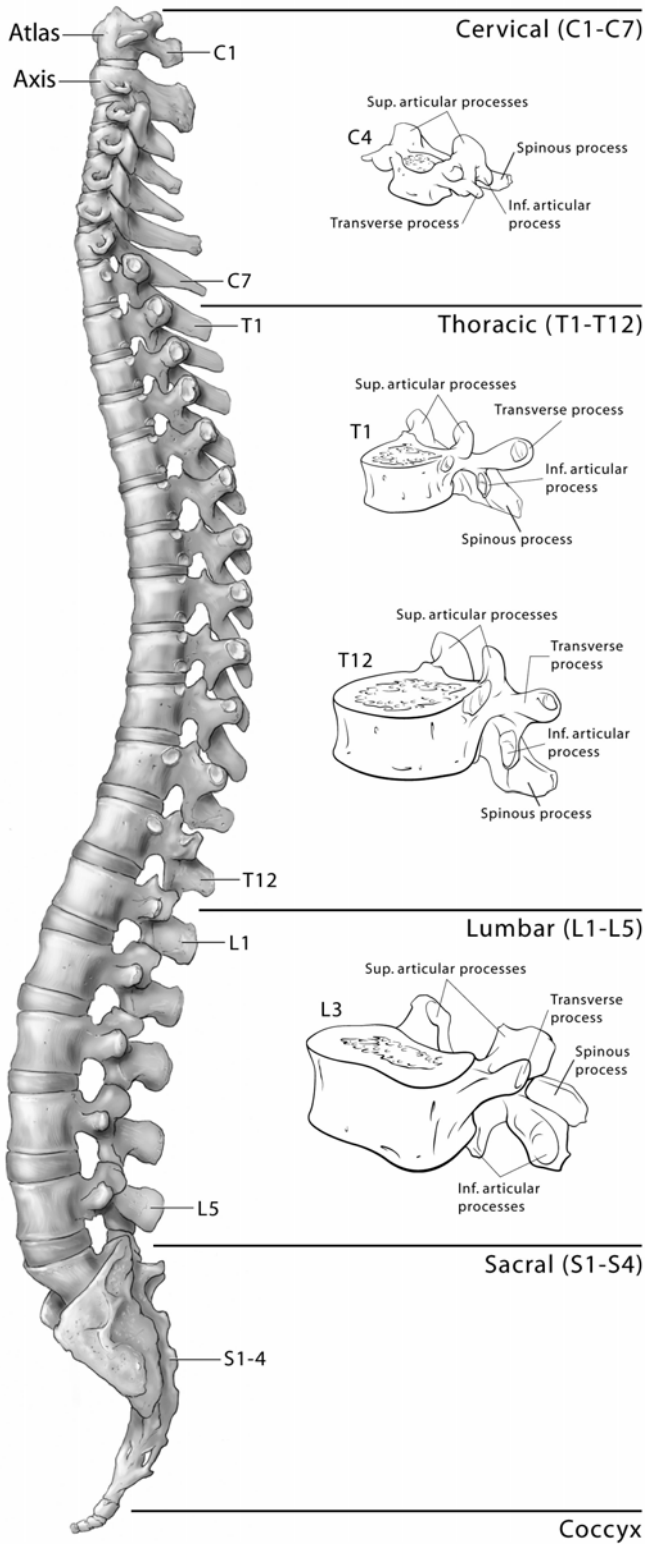


Figure 3. Nomenclature for vertebrae.

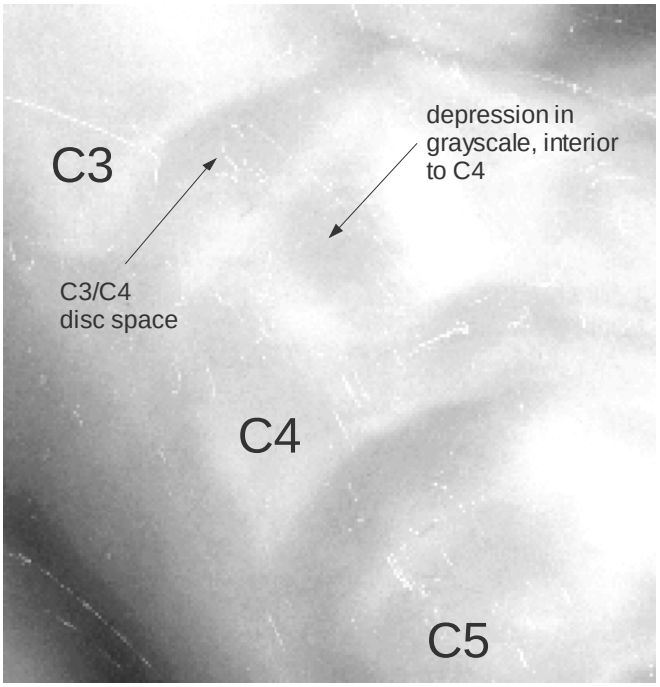


Figure 4a. A C4 vertebra and neighboring grayscale regions. Some of the grayscale interior to C4 is distributed similarly to the grayscale external to the C4 vertebral body, in the C3/C4 disc space.

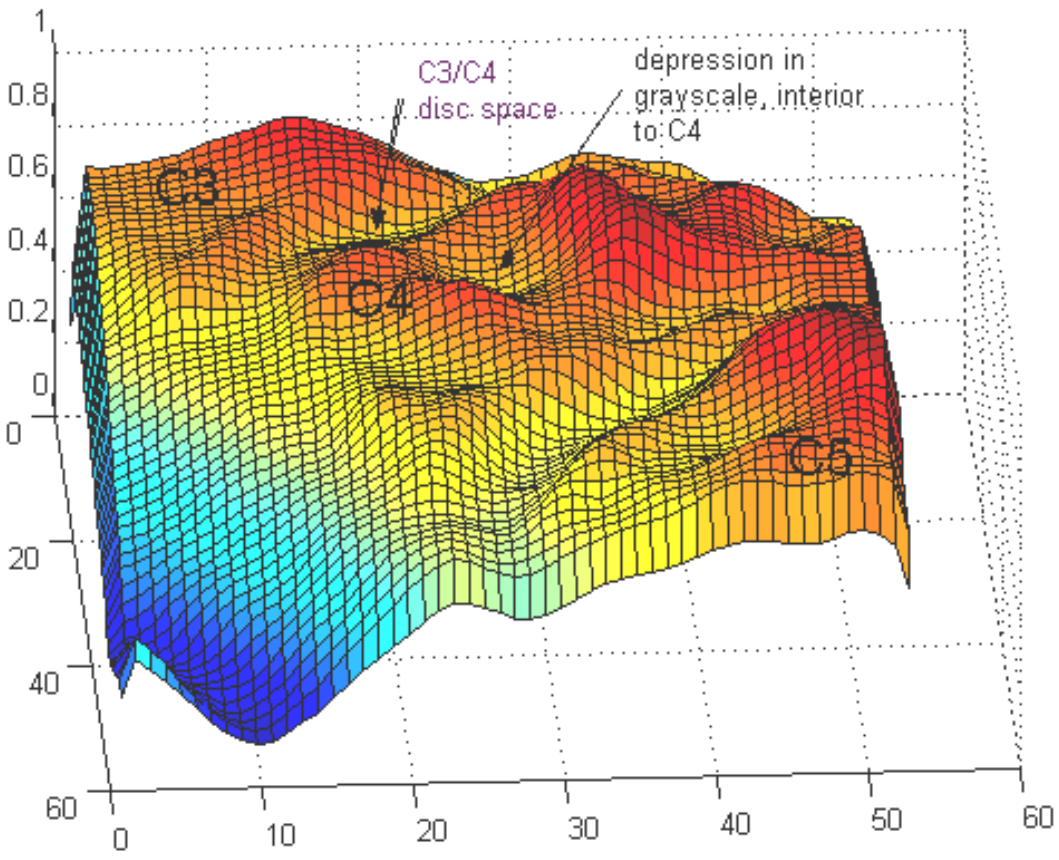


Figure 4b. Surface plot of the grayscale in the Figure 4a image, illustrating the similarity in interior and exterior grayscale for C4.

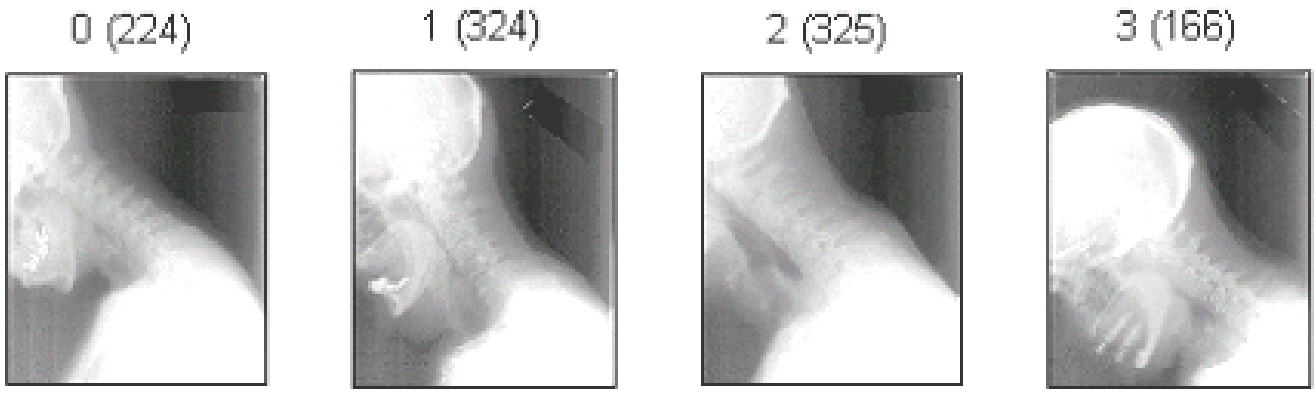


Figure 5. Cervical spine images with graded anterior osteophytes. The severity progresses from 0 (normal) to 3 (most abnormal) from left to right in the images.

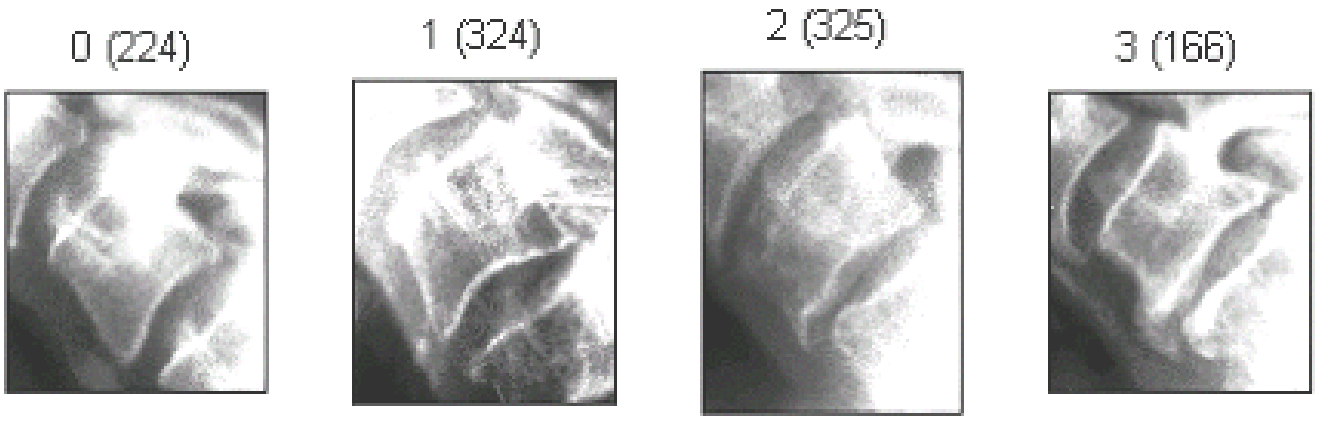
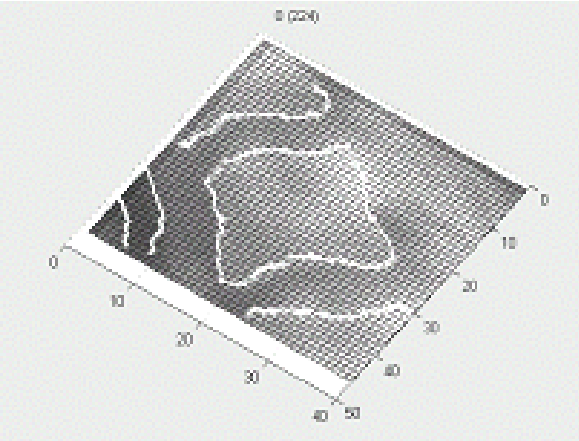
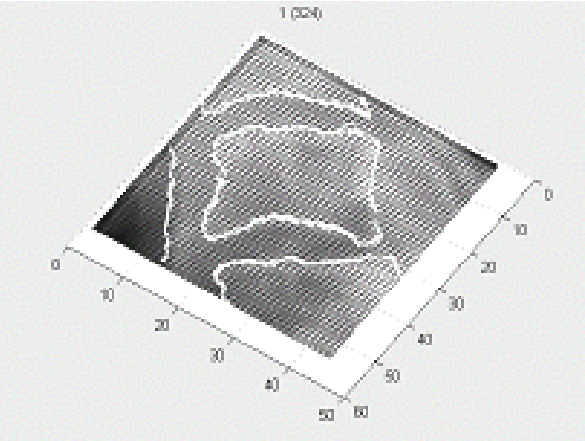
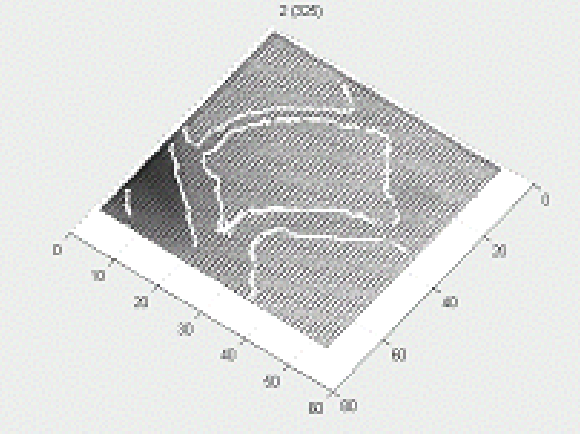
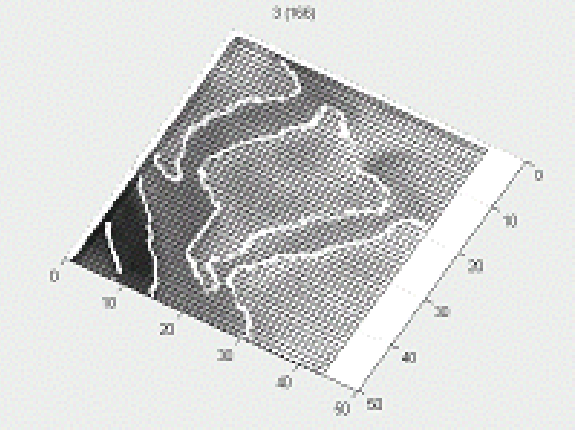


Figure 6. Vertebrae showing the osteophytes on lower anterior corner of center vertebra (grade 0 is normal—no osteophyte).









Figures 7a-d. Surface plots of the vertebrae with boundary outlines

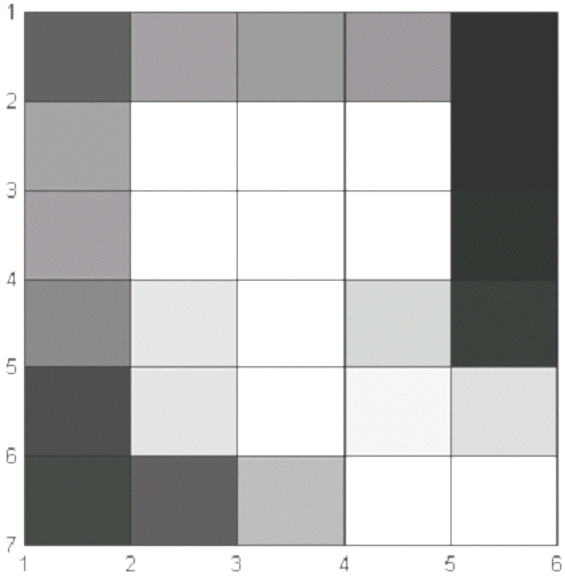
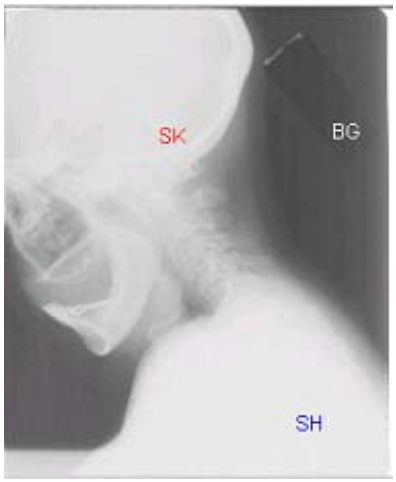
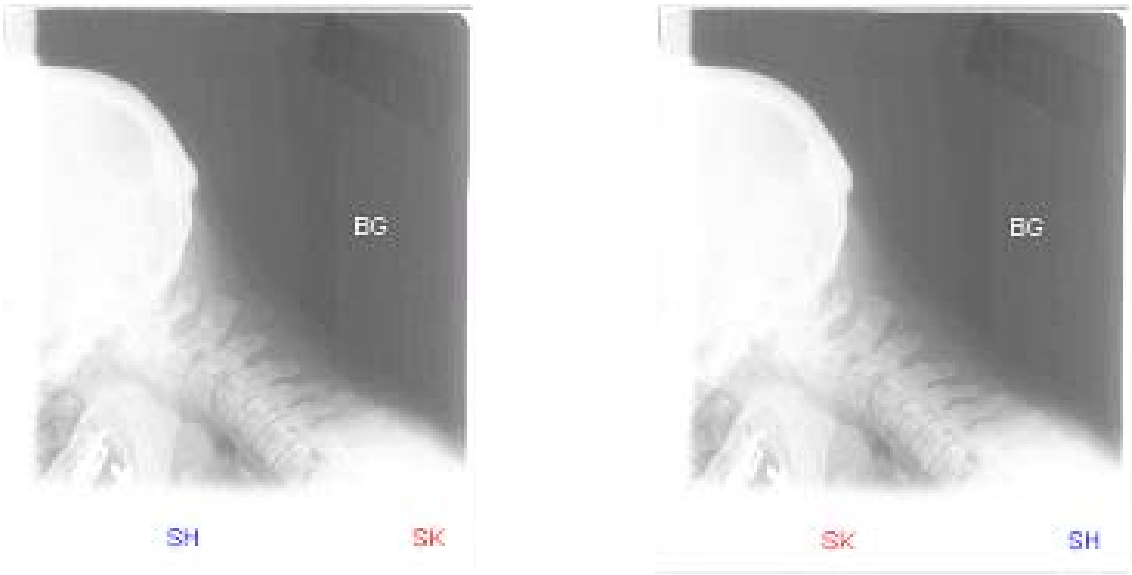


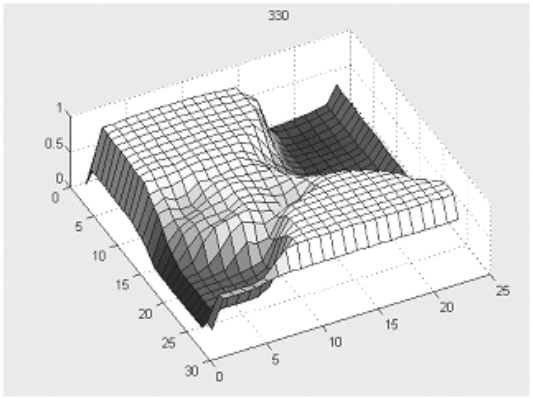
Figure 8a. Cervical spine image
Figure 8b. Cervical spine image heavily smeared and subsampled



Figures 9a-c. Results of basic landmark labeling



Figures 10a-b. Errors in labeling due to heavy light leakage. (a) shows labeling results using method (a); (b) shows results using method (d), which gives correct shoulder labeling.



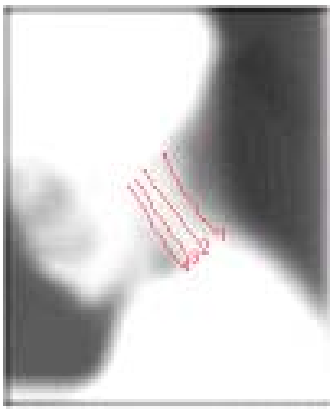


Figure 11a. Surface plot of a cervical spine x-ray.
Figure 11b. Conceptual curves C_1-C_4, drawn on the smeared image of Figure 11a.

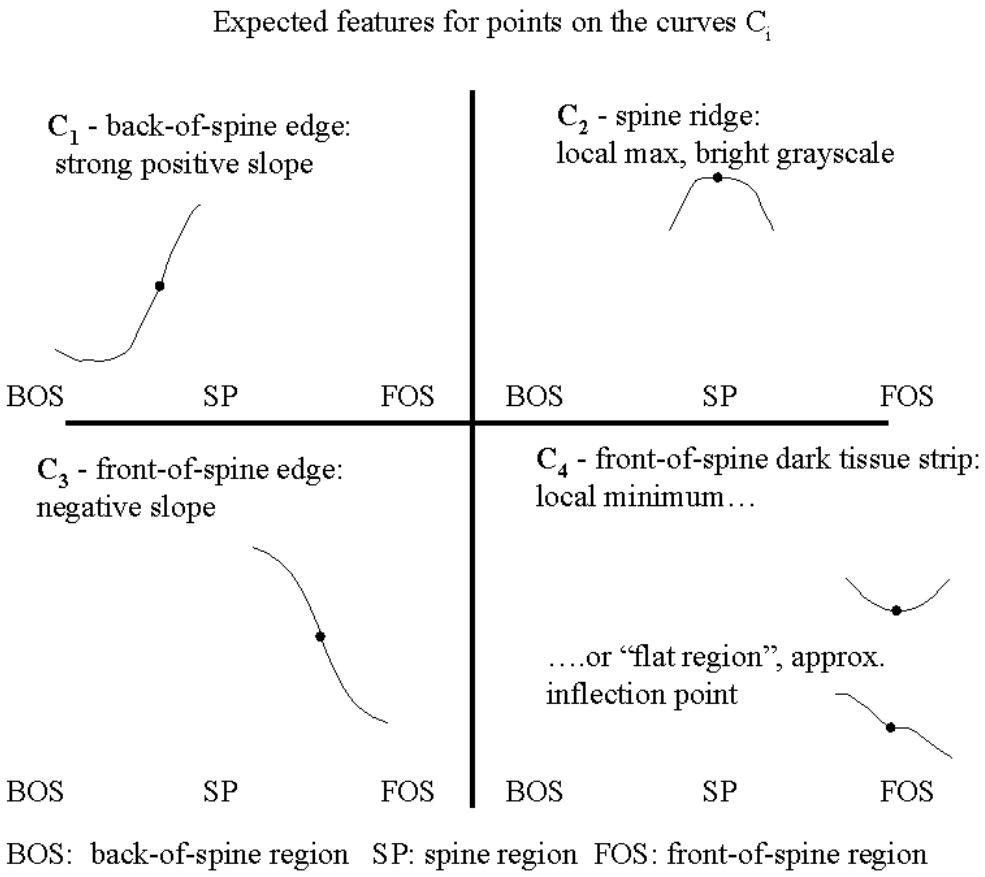


Figure 12. For points on any particular C_i, the neighboring grayscale curves along intersecting line segments that are normal to the spine axis have the expected characteristics shown above. Note that the above curves are not the C_i themselves, but the expected grayscale characteristics along line segments that intersect points on the C_i.

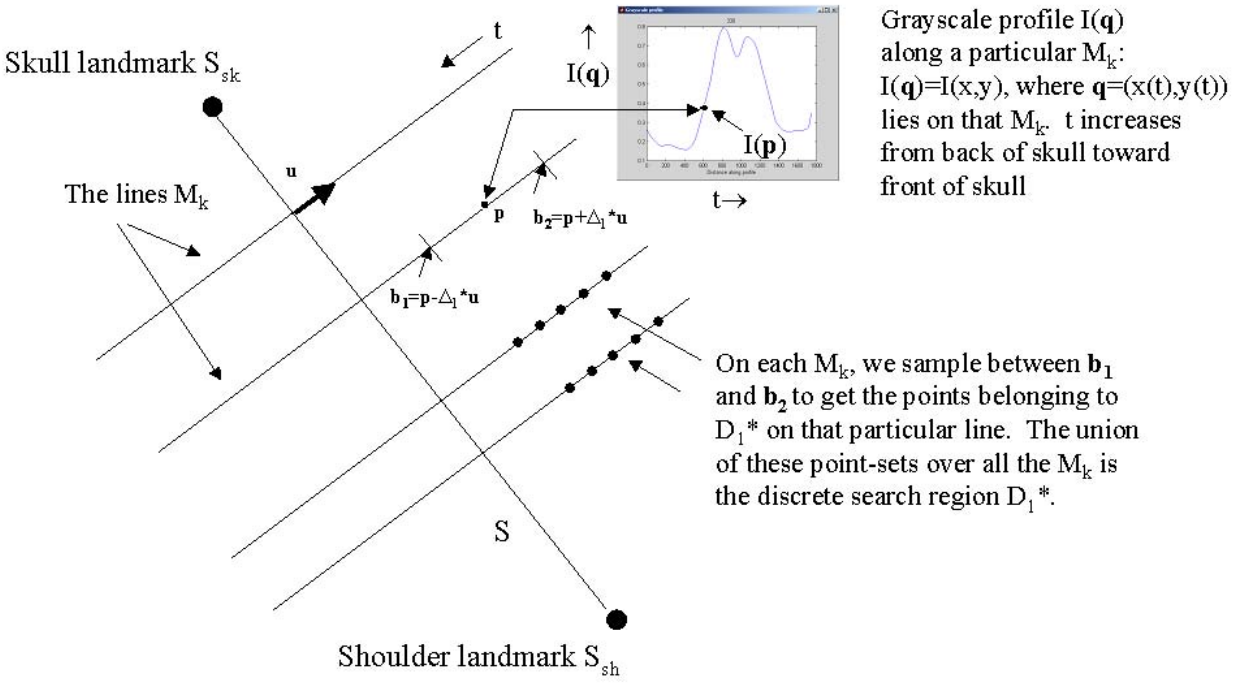


Figure 13. Illustration of heuristic for determining D_1*, the discretized search region for C_1*.

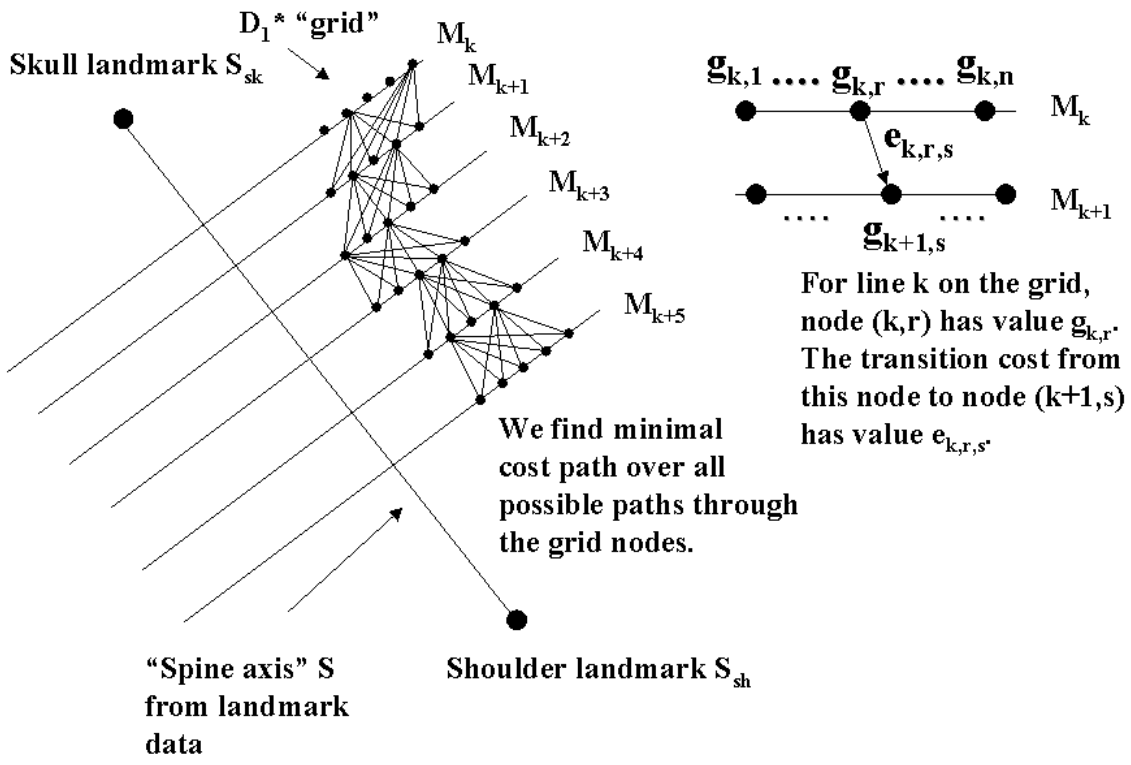


Figure 14a. The grid D_1* for searching for C_1*, with node and transition costs illustrated.

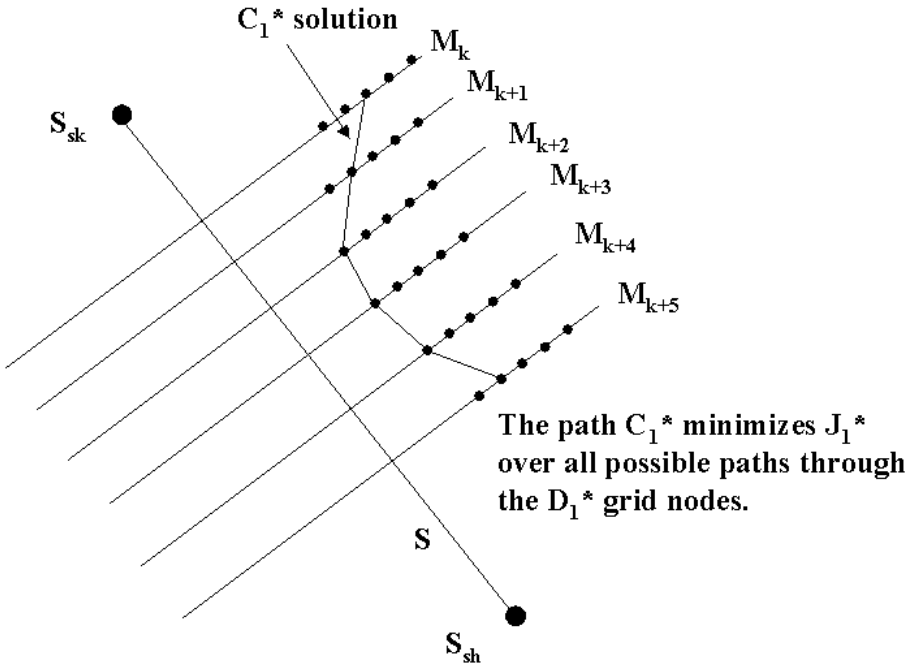
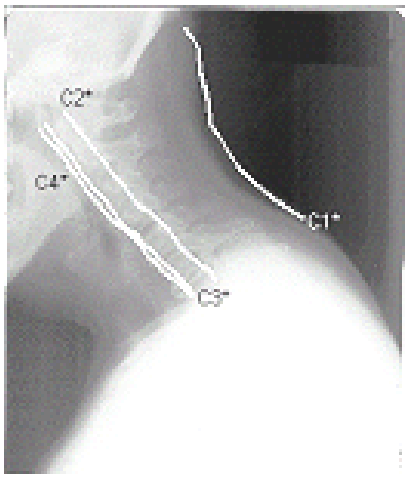
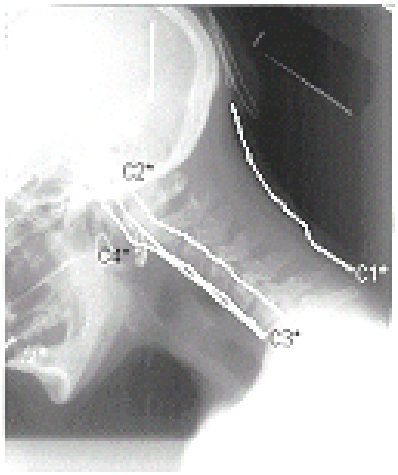


Figure 14b. Illustration of a discrete solution C_1* on the grid D_1*.





Figures 15a-b. Example C_1-C_4 solution curves (two different images).

Figure 16. Linear fits to the data of Figure 15a.

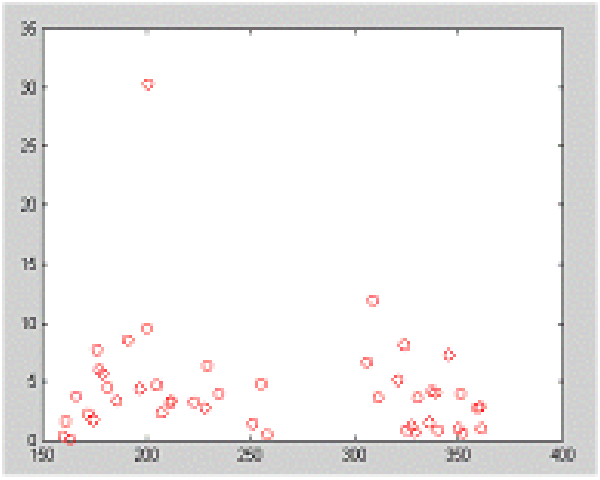


Figure 17. Differences (in degrees) between estimates from spine orientations derived by using (1) linear fits to the C2* curves vs. (2) linear fits to manually-collected coordinate data for the spines. Vertical axis: degrees; horizontal axis: image i.d. numbers.

WxHxD	1462x1755x8
Global grayscale	
Min	54
Max	255
Dynamic range	201
Mean	163.7
S.D.	62.3
Per image grayscale	
Mean of means	163.7
S.D. of means	9.5
Mean dynamic range	194.0
S.D. of dynamic range	5.1
Border area width	60 (nominal)

Table 1. Global cervical spine image characteristics.

Intervertebral distances (pixels)				
C1/C2	μ	204.0		
	σ	23.5		
	N	46		
		D63	D52	D41
C2/C3	μ	25.6	35.0	29.5
	σ	5.6	6.7	5.7
	N	46	46	46
C3/C4	μ	24.4	34.0	29.4
	σ	5.7	7.1	5.60
	N	48	48	48
C4/C5	μ	25.2	34.3	31.9
	σ	9.4	8.4	17.0
	N	47	47	47
C5/C6	μ	23.6	30.1	27.3
	σ	8.4	8.9	7.8
	N	45	45	45
C6/C7	μ	24.1	28.7	25.8
	σ	8.7	8.2	6.2
	N	36	36	35
C7/T1	μ	24.8	32.6	26.3
	σ	5.7	5.6	4.3
	N	12	12	12

Table 2a. Cervical spine interveterbal characteristics

Within-vertebra distances (pixels)								
		D13 (top)	D46 (bot)	D36 (ant)	D25 (mid)	D14 (pos)	Anterior osteophytes	
							D38 (top)	D69 (bot)
C2	μ	92.2	105.9	85.4	78.5	89.1	19.6	20.8
	σ	10.7	10.7	11.3	10.6	10.2	11.6	8.6
	N	48	48	48	48	48	2	7
C3	μ	97.6	107.0	80.9	74.8	85.5	20.2	20.2
	σ	11.5	12.8	10.2	8.6	9.5	4.6	5.1
	N	48	48	48	48	48	6	18
C4	μ	97.7	111.3	77.0	74.7	84.8	26.2	24.1
	σ	11.4	16.2	13.3	10.8	10.7	9.8	10.0
	N	47	47	47	47	47	12	36
C5	μ	105.2	114.4	75.50	74.7	79.6	25.8	19.5
	σ	13.7	12.1	10.2	8.8	9.4	14.0	6.8
	N	45	43	43	43	43	21	31
C6	μ	112.5	109.9	89.0	82.9	88.4	19.9	21.6
	σ	10.7	10.9	8.0	7.1	8.9	7.3	6.9
	N	35	27	26	27	27	14	4

Table 2b. Cervical spine within-vertebral characteristics

	Skull		Shoulder		Background	
Classification Method	Accept	Reject	Accept	Reject	Accept	Reject
a	46	2	48	0	48	0
b	32	16	34	14	48	0
c	46	2	47	1	48	0
d	46	2	48	0	48	0

Table 3. Results of classifying landmarks by four methods