Indexing of image content in spine x-rays

L. Rodney Long, G. R. Thoma

National Library of Medicine, Bethesda, MD 20894

ABSTRACT

This paper presents work toward indexing the image content in a collection of 17,000 cervical spine and lumbar spine images, for purposes of public dissemination by such systems as the Web-based Medical Information System (WebMIRS). These images were collected as part of a national health survey and to date no radiological or quantitative content has been derived from the images, except for our work, described in this paper. Practical considerations, primarily of labor cost, make the job of deriving radiological interpretations or quantitative anatomical measures by manual methods very difficult. For this reason, the acquisition of content information by automated means, or even by semi-automated means, which require human interaction, but significantly reduce the required labor, are very important. This field is not in an advanced state of development, and the results we present are necessarily work in progress.

Keywords: biomedical image analysis, NLM, NIH, digital x-ray image, Chow-Kaneko algorithm

1. INTRODUCTION

At the Lister Hill National Center for Biomedical Communications, a research and development division of the National Library of Medicine, we are building a biomedical information resource consisting of digitized x-ray images and associated national health survey data. We have described the architecture and characteristics of this system in previous papers [1-2]. This resource is called the *Web-based Medical Information Retrieval System*, or WebMIRS. This paper discusses our current work toward the long-range goal of directly exploiting the image contents in our system. In a future WebMIRS system, we envision that we will have in our databases not only *text and raw image data*, but *quantitative anatomical feature information* that is derived directly from the images. We further envision the system to have capability to retrieve images based on image characteristics, either alone or in conjunction with text descriptions associated with the images.

The current version of WebMIRS lets the user create a query using a Graphical User Interface and get the results displayed as either a table of database text records or as a display of images and related text. **Figure 1** shows a sample output screen from WebMIRS. WebMIRS is an Internet database application, written as a Java applet, and usable by means of a standard Web browser. Information on WebMIRS is available on our Web site: http://archive.nlm.nih.gov.

As detailed in the WebMIRS papers referenced above, our archive consists of records for two databases, corresponding to the survey data collected in the second and third National Health and Nutrition Examination Surveys (NHANES), conducted by the National Center for Health Statistics. For the NHANES II survey, the records contain information for approximately 20,000 survey participants. Each record contains about two thousand data points, including demographic information, answers to health questionnaires, anthropometric information, and the results of a physician's examination. In addition, many of the participants were x-rayed: approximately 10,000 cervical spine and 7,000 lumbar spine x-rays were collected. The WebMIRS system makes the alphanumeric data available for query in a relational database system, and, at the user's option, also returns the image data for display. Image results returned are raw images only; no quantitative or even descriptive information about the images themselves is stored in the database.

The images in the NHANES II survey were collected primarily for the study of osteoarthritis and degenerative disc disease. Biomedical features of interest to researchers in these areas have been identified by two workshops conducted at NIH, and consist of anterior osteophytes, disc space narrowing, and subluxation. A database containing quantitative and descriptive image information that allows intelligent search for images that show varying degrees of these features is expected to be of research use for the osetoarthritis and related communities. In addition, published research work[3-4] in the field of vertebral morphometry suggests that a database with quantitative measures of vertebral dimensions may be useful for purposes such as studies in occurrence of spinal fracture and estimation of normative values for vertebral size.

In order to derive such information directly from the images, detailed analysis and measurement of each image is required. Due to the prohibitive cost of manually carrying out such a process on 17,000 images, we are conducting research into methods for automating or semi-automating the process. In this paper, we describe our initial work toward this goal, which consists of (1) developing an algorithm to find basic landmarks in the images by fixing a coordinate system in the anatomy itself, (2) implementing an adaptive grayscale thresholding method for isolating the anatomical components of the images, and (3) planned work for locating boundaries of individual vertebrae for the purpose of measuring geometrical dimensions, inter-vertebral spacing, and shape characteristics.

We may characterize the content that we want from the images as being (1) basic orientation landmarks and directions, grounded in the anatomy of the image itself; (2) geometrical and positional data specific to the vertebrae, including counts of vertebrae present, location of vertebrae, identification of vertebrae, and coordinate location of reference points on vertebrae boundaries; and (3) data closely related to pathology in the spine, including presence of osteophytes (or sharpening of vertebral corners, sometimes accompanied by presence of isolated bony masses near the vertebrae corners), disc space measurements, and subluxation (or slippage of vertebrae out of normal alignment). These three categories may be characterized as a rough progression from the lower to higher levels of semantic content in the images. Steps (1) and (2) are necessary for the eventual computer-assisted extraction of what might be termed the high-level, biomedical content of step (3). It is this last, highest level of semantic content that is expected to be of greatest interest and value to the biomedical research community. Our work is concerned with making progress toward the goal of getting all of the images indexed at these various levels by making the best combination of machine capability and human effort. To date, most of our work has focused on the image processing aspects, and we continue that direction in this paper.

2. APPROACH

In previous work[5] we have given an overview of the indexing problem for these images and also have demonstrated[6] some initial success in getting basic orientation information from cervical spine images by an algorithm that automatically places an orthogonal coordinate system with one axis U at the base of the skull and the perpendicular axis V positioned to intersect this axis at the point O (coordinate system origin) of greatest bone density in the spine (as defined by brightest grayscale concentration). In this paper we present work that improves the results obtained in [6] by implementing an adaptive grayscale thresholding technique, and we discuss plans to extend our work by using the derived U/V coordinate system to count and locate individual vertebrae in the cervical spine. The adaptive thresholding technique, by improving the segmentation of the images over previously-used methods, improves this process.

Review of the anatomy-based coordinate system. The anatomy-based coordinate system described above is located using an algorithm with these basic steps: (1) Apply grayscale thresholding to the image using a threshold that preserves the essential shape of the skull base in the jaw region and in the region of the skull posterior; (2) in the thresholded image, locate the skull base region at a coarse level, and measure slopes of tangent lines at points expected to be in the jaw region, and also for the points expected to be in the lower skull posterior region; (3) find the pair of points, one chosen from jaw region, one chosen from skull region, that have the closest agreement in slope of tangent line; (4) take these points as defining the coordinate system axis U along the skull base; (5) determine a subinterval R of the U axis which contains the intersection of U with the spine; (we define this region R as the subinterval of the U axis bounded by certain characteristic grayscale patterns--two dark areas marking the spine boundaries, separated by a continuous bright area corresponding to the spine itself); and (6) search within the U axis/spine intersection region R for the point O of maximum brightness; this point defines the origin of the U/V coordinate system and the intersection of the orthogonal axis V with the U axis. Steps in calculating the U/V axes are illustrated in **Figure 2**. **Figure 2A** shows a grayscale cervical spine x-ray image. **Figure 2B** is the same image, with the boundaries for a grayscale threshold superimposed on the original image. It is the boundaries of this thresholded image that are used in the computation of the U axis. **Figure 2C** shows the U axis that has been computed. Finally**, Figure 2D** shows the V axis added orthogonally to the U axis, and intersecting the U axis at the point of greatest grayscale value in the line segment determined by the U axis/spine intersection.

This anatomy-based coordinate system (ABCS) algorithm is described in detail in [6], which discusses test results, failure modes, and planned improvements. Most current effort with this algorithm is oriented toward improving robustness of the grayscale thresholding, and for handling cases where expected gross anatomical features (e.g. jaw) are not present because of the positioning of the subject in the image. Of 25 test images selected at random from our collection, 4 fail to exhibit a clear jaw profile, which the ABCS algorithm requires. This is a significant fraction of cases, which the current ABCS algorithm cannot handle. In addition, cases occur where the grayscale threshold used to separate out the skull base shape from the rest

of the image does not preserve the essential shape well. This latter case is illustrated by **Figures 3A** and **3B**. For this particular image, the simple approach (i.e., thresholding by image global mean value) used to grayscale threshold the image failed to yield a meaningful contour in the jaw area; the result is a badly located set of U/V axes. To address this problem, we have implemented an adaptive grayscale thresholding algorithm based on work by Chow and Kaneko[7-8].

Use of the U/V coordinate system in locating, identifying, and indexing spine anatomy. To make an interpretation of the image contents, basic orientation within the image anatomy is required: we need a beginning frame of reference for making progress in analyzing the geometry of the anatomical structures that we expect to find. The U/V coordinate frame, with origin fixed within the upper spine area, and with orientation fixed relative to the skull base, is hypothesized to provide sufficient precision to allow the image region containing the vertebrae to be identified. Our first approach to getting characteristics of the vertebrae themselves is to initially estimate dimensions and location of a rectangular region expected to contain the vertebrae, using a priori information manually derived from test images. Specifically, we compute the U/V coordinate system for these images, manually identify a rectangular region which contains the vertebrae, but which is no larger than is necessary.

The next step toward indexing the image using characteristics of the spine vertebrae consists of analyzing this region to determine the boundaries of the vertebrae contained within it, and, once vertebral boundaries are obtained, to derive measurements from the vertebral shapes, which are relevant to our indexing goals. For example, an indexing measure of interest is spacing between vertebrae, which is a quantity used to assess disk degeneration within the spine. Other measures of interest include the anterior and posterior vertebral heights: the ratio of these values has been used to assess fracture within the spine. Some features will require shape analysis of the vertebral boundaries; these include detection of osteophytes—sharpening of the corners of the vertebrae; these are expected to require careful derivative analysis of the vertebral boundaries, and the fidelity with which the boundary has been derived will be critical for getting useful results. This last point highlights one of the greatest difficulties in image processing of the spine images: in many cases, the vertebral boundaries are ill-defined in the sense of appearing visually fuzzy or indistinct.

An approach that we are pursuing for analyzing regions for presence of vertebrae and for identifying boundaries of these vertebrae uses a technique of edge-searching which combines (1) estimation of the expected range of tangent angle values, measured in the U/V coordinate system, for anterior vertebrae edges; (2) gradient thresholding with high threshold to compute a pixel set expected to contain parts of the anterior vertebrae edges, (3) reduction of this set to only those pixels lying on the anterior vertebrae edges by using the tangent angle criteria computed above, (4) gradient thresholding with low threshold to compute a pixel set expected to contain most of the vertebral edge pixels (but also extraneous edges and lines within and external to the vertebrae), and, finally, (5) reduction of the pixel set computed in (4) to only those pixels lying on vertebral edges, by constraining the pixels in the reduced set to lie on continuous curves with those computed for the anterior edges in (3).

3. RESULTS

Results are given in this section for the initial testing we have done on our implementation of an algorithm for adaptive thresholding. Our implementation of this algorithm of Chow and Kaneko may be summarized in the steps below. We have followed the original published approach closely, except as noted.

(1) Divide the image into overlapping subimages. Our images are cervical spine images of dimension 335x408 (reduced one-quarter scale from the original spatial resolution with which they were digitized.) We used 64x64 subimage blocks with 50% overlap for the processing, yielding a 10x12 array of 64x64 subimages for processing.

(2) Within each subimage, model the distribution of the grayscale values z as a mixture of two Gaussian density functions. The model used is given by the standard formula

$$f(z;p1,p2,m1,m2,s1,s2) = p1 * g1(z;m1,s1) + p2 * g2(z;m2,s2),$$

where g1() and g2() are Gaussian with (mean, standard deviation) of (m1,s1) and (m2,s2), respectively, and p1, p2 are the proportions of g1() and g2() which contribute to the final density. p1 and p2 satisfy the relationship: p1 + p2 = 1, so that the model has only five independent parameters.

(3) Estimate the five independent parameters of this mixed-Gaussian density. Let h(z) be the observed histogram for a 64x64 block. We use a conjugate gradient method to minimize the objective function J(), defined as the sum-square-differences between the observed histogram values and the values predicted by the model f():

$$J(p1,p2,m1,m2,s1,s2) = SUM\ [f(z;p1,p2,m1,m2,s1,s2)-h(z)]**2,$$

where the summation is taken over all grayscale values z that are observed in the 64x64 block; J() is minimized as a function of the parameters (p1,p2,m1,m2,s1,s2). The outputs of the minimization process are estimates for p1 (from which p2 is computed), m1, s1, m2, and s2. Prior to doing this estimation, we do a variance test to determine the reasonability of modeling the grayscale values in the subimage as two distinct probability distributions. If very low grayscale variance was found, we assume that such a model is not reasonable for the subimage. Chow and Kaneko implemented an additional test for likelihood of being a mixed Gaussian density by analyzing the ratio of peaks to valleys in the grayscale distribution; we have not implemented this feature at this point,

(4) Compute a threshold to separate the two densities based on a maximum likelihood approach, which uses the parameters estimated for the mixed-Gaussian density. Specifically, the threshold t was computed by solving

$$A*t**2 + B*t +C = 0$$

for t, where

$$A=(1/s1**2) – (1/s2**2),\ B=2*((m2/s2**2)-(m1/s1**2)),\ \text{and}$$

$$C=(m1**2/s1**2)-(m2**2/s2**2)+2*\ln(p2/p1).$$

(5) For each subimage for which the estimation process was successful, a threshold t was obtained. For the subimages where the variance was too low to justify the mixed-Gaussian model, or subimages where the estimation process failed to converge, a null threshold value was temporarily assigned.

(6) The thresholds obtained by the estimation process are then used to interpolate thresholds to the entire set of 64x64 subimages and, finally, this grid of thresholds, one per 64x64 subimage, with the overlapping subimages covering the entire image, is used to interpolate a threshold value for each point in the entire image.

Example results are shown in the figures. **Figure 4A** shows a cervical spine image input which was first thresholded using the image global mean, for comparison with the output from the Chow-Kaneko processing. **Figures 4B** and **4C** show the respective results of these two methods. For this particular image, the thresholding produced by the global mean thresholding was actually adequate for location of the U/V axis system (note the obvious jaw and posterior skull boundaries in **Figure 4B**). However, in many images, the jaw contour in particular is very hard to distinguish using global thresholding. Also, note the improvement in accuracy of the posterior skull contour in **Figure 4C**, as compared to **4B**. Finally, even though it is not of importance for U/V coordinate system locating, note the much greater detail in spine anatomy of **Figure 4C**, as compared to **4B**. This suggests a possible future role for this thresholding algorithm as an aid in segmenting the vertebrae themselves.

Figure 5 demonstrates the increased vertebral detail possible in the adaptive thresholding with an example subimage taken from **Figure 4**. In **Figure 5A**, a 64x64 subimage taken from the region of the C2 vertebrae is shown, along with the appearance of the same subimage after global thresholding (**Figure 5B**) and by Chow-Kaneko thresholding (**Figure 5C**). **Figure 5C** shows clear separation of two vertebrae, while **5B** does not even reveal the presence of vertebrae.

Initial results given by the Chow-Kaneko algorithm appear quite promising, but our work is still at a preliminary stage. Some of the problems we are dealing with include apparent sensitivity of the convergence to good initial parameter estimates. It is not clear at this stage whether the approach taken in the original work is adequate for having the minimization algorithm converge for the subimages in these digitized x-rays. We have incorporated the same method used by Chow and Kaneko for the computation of the a priori parameter values. This method is as follows: the histogram for the 64x64 subimage is divided into two parts, one part consisting of the grayscale values in the subimage larger than the mean grayscale value for the subimage, and the other part, of the grayscale values in the subimage smaller than this mean grayscale value. Within these two regions of the histogram, grayscale mean values (m1,m2) and standard deviations (s1,s2) are computed. Then, the parameter pairs (m1,s1) and (m2,s2) are used as a priori values for the Gaussian curves in the estimation process. A priori values for the fractional coefficients for the two curves are set as p1=p2=0.5.

It should be noted that convergence of the minimization problem for this choice of a priori parameter values is not assured, although, in our experience on the cervical spine images, improvement in the parameter values, as measured by reduction in the magnitude of the objective function, is usually attainable, but there is no assurance that the result is the desired optimal

minimizing parameter set. As is typical of many nonlinear minimization problems, convergence to a solution within the feasible range of values is dependent on good starting values. We have done some testing on both synthetic and real image data to try to understand the degree of dependence on good a priori parameter values for this problem. **Figure 6** shows one result, using synthetic data. **Figure 6A** shows a mixed-Gaussian probability density with "truth" parameters (.4,160,15,190,10) which we synthesized to model the "truth" grayscale distribution for an image subregion which has grayscale values lying in the range (125,215). To each value on this curve we added 10% Gaussian random noise, also shown in **Figure 6A**. This noisy data was then used to model for histogram data collected from an image subregion. We then solved for the best set of parameters which fit a mixed-Gaussian distribution to this noisy data, using 10% perturbations added to each of the components of the truth parameter set to obtain a priori parameters, i.e. the a priori values used were 1.1 * the truth parameters. The resulting a priori mixed-Gaussian curve is shown in **Figure 6B**. **Figure 6C** shows the solution curve obtained by minimizing J() for a set of minimizing parameters. The solution parameters obtained were (.41,160.4,15.6,190.2,9.8). This example gives a basically good result; however, pushing the a priori error to 20% results in a nonfeasible solution to the problem. More extensive testing is needed, but this may give some indication of the error tolerance in the initial guess at the solution. In **Figure 6D** we show the grayscale histogram for the 64x64 image subregion of **Figure 5A**. (This data lies in the grayscale range (125,215) and in fact was used to choose the range for the synthesized data discussed above.) Superimposed on this histogram is the a priori mixed-Gaussian distribution computed from a priori parameter values, which were calculated from the subregion histogram by using the Chow-Kaneko approach, as described above. In **Figure 6E** the subregion histogram data is again shown, with the solution mixed-Gaussian curve obtained by minimizing the objective function J() defined above.

It should be noted that solving the minimization problem is relatively expensive in computer time: carrying out the complete Chow-Kaneko algorithm for one of our test images, including solving for the parameters which minimize the objective function for each 64x64 subimage, takes on the order of 1 or 2 minutes on a 266 MHz PC.

For implementation of the above algorithm, we have used the MATLAB m-file language, under MATLAB 5.2. For the conjugate gradient minimization, we used the E04DGF routine of the Numerical Algorithms Group (NAG), as implemented in the MATLAB NAG Foundation Toolbox.

4. CONCLUSION

Our approach to indexing digitized x-ray images of the cervical spine for content relevant to researchers in the field of spine disease and injury is to first obtain a basic geometrical orientation within the anatomical features contained within the image, by locating an orthogonal U/V coordinate system with origin within the spine, and with one axis tangent to the base of the skull. This coordinate system is then used as the basic reference frame for estimating the subregion of the image containing the vertebrae, and for estimating the orientation of the vertebrae. We are approaching the problem of determining vertebral boundaries within this subregion using gradient thresholding to detect vertebral edges, and expected edge orientation to eliminate spurious points not on the vertebral edges. Our current work in progress consists of continued refinement of the U/V coordinate system algorithm for improved performance, and in developing and evaluating the algorithm to locate and identify vertebral boundaries. A key concern in our work is to assess the degree to which computerized methods may assist in the accurate derivation of such x-ray image content and to create a concept of human-machine interaction that exploits the strengths of both the human participant and computer capability for solution of this particular problem.

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 IV*, San Jose, CA, January 28-30, 1998, SPIE vol. 3312, pp. 392-403.
3. Rosol MS, Cohen GL, Halpern EF, Chew FS, Kattapuram SV, Palmer WE, Dupuy DE, Rosenthal DI. Vertebral morphometry derived from digital images. *American Journal of Roentgenology*, Vol. 167, December 1996, pp. 1545-1549.
4. Hedlund LR, Gallagher JC. Vertebral morphometry in diagnosis of spinal fractures. *Bone and Mineral*, Vol. 5, 1988, pp. 59-67.

5. Long LR, Thoma GR. Image query and indexing for digital x-rays. Proceedings of SPIE Storage and Retrieval for Image and Video Databases VII, San Jose, California, January 26-29, 1999, pp. 12-21.
6. Long LR, Thoma GR. Segmentation and feature extraction of cervical spine x-ray images. Proceedings of SPIE Medical Imaging 1999: Image Processing. Vol. 3661, San Diego, CA, February 20-26, 1999, pp. 1037-1046.
7. Chow CK, Kaneko T. Automated Boundary Detection of the Left Ventricle from Cineangiograms. *Computers and Biomedical Research*. Vol. 5, Num. 4. August 1972, pp. 388-410.
8. Chow CK, Kaneko T. Boundary detection and volume determination of the left ventricle from a cineangiogram. Computers in Biology and Medicine. Vol. 3, 1973, pp. 13-26.

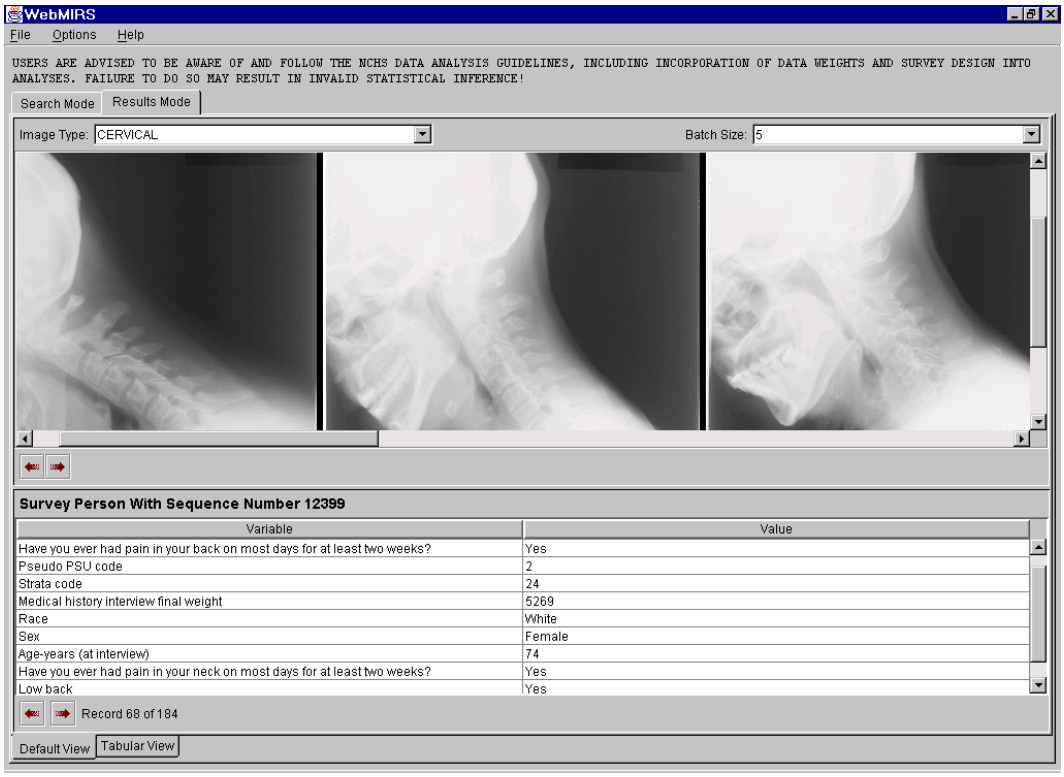
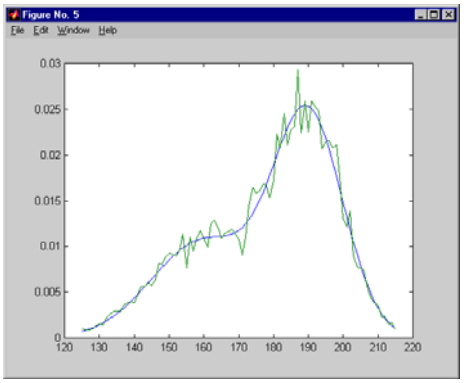
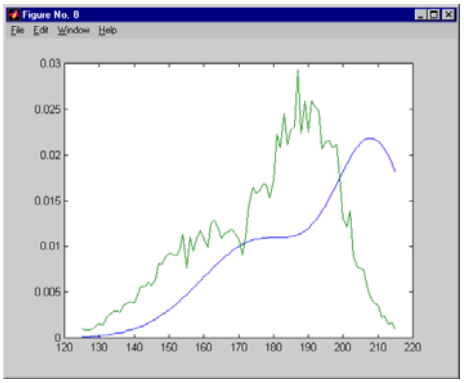


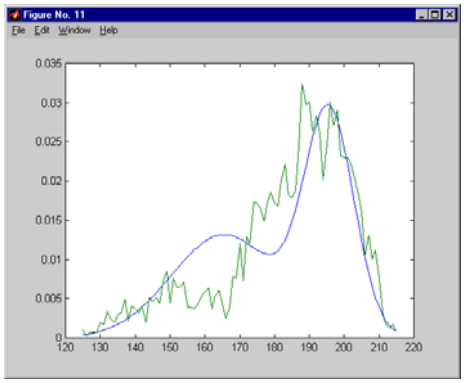
Figure 1. WebMIRS output screen, resulting from a database query to *find all records for people over 60 who had low back pain for at least a two-week period*. The text at bottom gives field values for one of the people whose image is shown in the middle at the top of the screen.



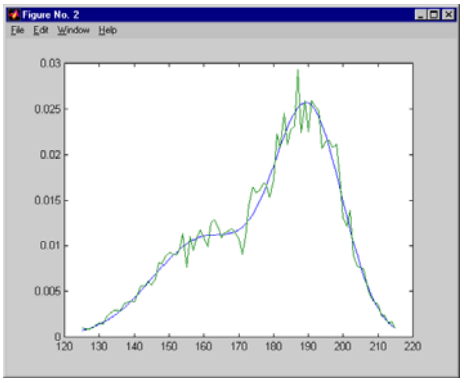
A – synthesized truth curve and data



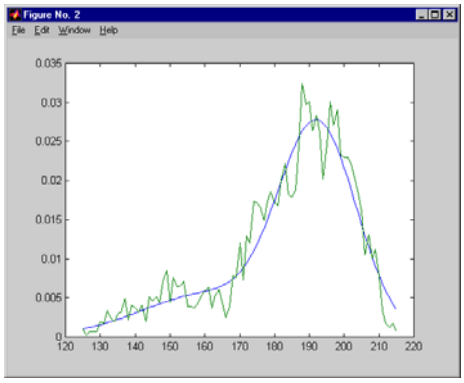
B – a priori curve (perturbed truth curve) and synthesized data



D – a priori curve and image 64x64 subregion histogram data



C – solution curve for synthesized data



E – solution curve for image 64x64 subregion histogram data

Figure 6. (A) A mixed-Gaussian density "truth curve" was synthesized, and Gaussian zero-mean random noise was added at each point (with sigma = 10% of that data point value). This noisy curve models the histogram within an image subregion. (B) A priori parameters for a mixed-Gaussian distribution were selected by perturbing the truth parameters from (A) additively by 10% for each parameter. (C) The noisy data from (A) and the a priori parameters from (B) were used to solve iteratively for the best parameter set to fit a mixed-Gaussian curve to the data. The solution parameters were used to plot the curve in (C). In (D) the actual histogram data for the 64x64 image subregion shown in Figure 5A is plotted, along with a mixed-Gaussian curve computed from a priori estimates of the curve's parameters. (E) shows the solution curve for the actual image data case, iterating to a solution from the a priori curve shown in (D).

Figure 4. (A) CSPINE image; (B) Image thresholded by global mean grayscale value; (C) Image thresholded adaptively by Chow-Kaneko algorithm.

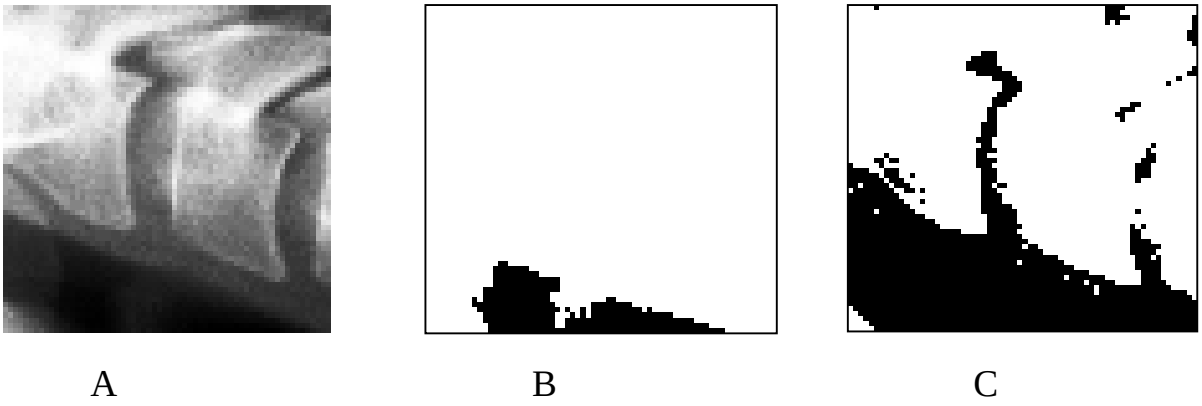


Figure 5. (A) 64x64 block cut from image 4A. (B) Corresponding block cut from the globally-thresholded image in 4B. (C) Corresponding block cut from the Chow-Kaneko thresholded image in 4C.

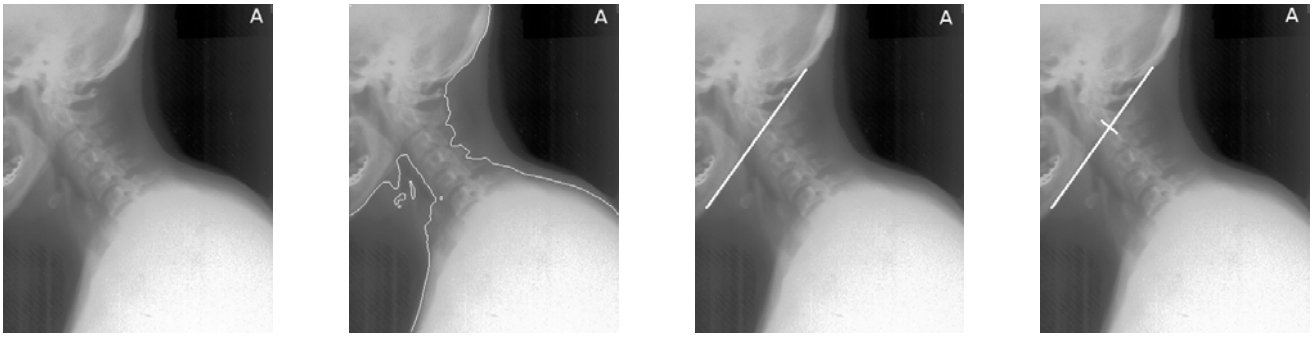


Figure 2. (A) Original cervical spine image with borders removed. (B) Foreground/background separation, by grayscale thresholding. The edge of the thresholded image is superimposed on the original image. (C) U axis fixed along the base of the skull. (D) U and V axes fixed. V is shown as the short axis orthogonal to U.

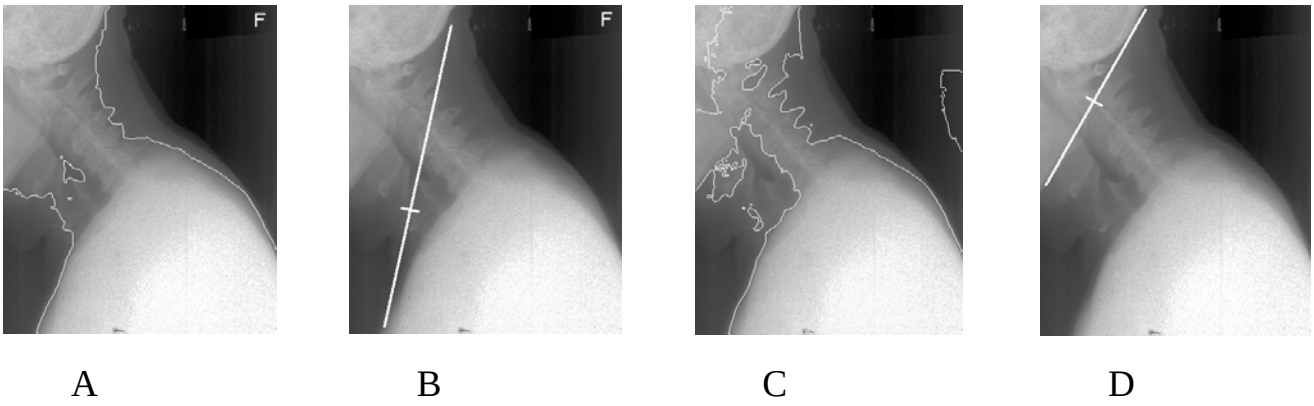


Figure 3. Illustration of effects of poor thresholding. (A) The jaw edge curve for this image is very inaccurate in following the jaw contour when global mean grayscale thresholding is used. (B) Results of the algorithm for (A). (C) The same image was adaptively thresholded using the Chow-Kaneko algorithm. Now, a fraction of the edge curve follows the jaw contour more reasonably. (D) Results of the algorithm on (C).