

# Biomedical Information from a National Collection of Spine X-rays: Film to Content-based Retrieval

L. R. Long<sup>\*a</sup>, S. Antani<sup>a</sup>, D. J. Lee<sup>b</sup>, D. M. Krainak<sup>c</sup>, G. R. Thoma<sup>a</sup>

<sup>a</sup>National Library of Medicine, Bethesda, MD 20894; <sup>b</sup>Brigham Young University, Provo, UT 84602; <sup>c</sup>Northwestern University, Evanston, IL 60208

## ABSTRACT

We summarize research and development for the extraction and distribution of biomedical information from a collection of 17,000 spine x-ray images collected by the second National Health and Nutrition Examination Survey (NHANES II). We present a history of the technical milestones of this work, including the data collection as film, digitization, quality control, archiving technology, database organization, medical expert content evaluation, and Web data distribution. We conclude by presenting our current work in content-based image retrieval (CBIR) to exploit the information content of these images directly by using image processing. We provide an overview and current research results from this CBIR work, which includes: extensive segmentation research, focusing on Active Shape Modeling and Active Contour methods; alternative techniques for shape representation, including invariant moments, simple polygon approximation, and Fourier descriptors; neural network classification of shapes into biomedical categories, such as “anterior osteophytes present/not present”; and the implementation of a prototype CBIR system for the vertebrae that supports hybrid text/image queries using MATLAB and the MySQL relational database system.

**Keywords:** x-ray, multimedia database, Web, CBIR, content-based image retrieval

## 1. INTRODUCTION

*Medicine has emerged in the late twentieth-century as an information-intensive practice. Human memory, anecdote, and folklore can no longer train and equip medical practitioners to do their best possible job nor can they serve the interest of the patient.*<sup>1</sup>

These words of C.G. Chute of the Mayo Clinic may be taken as an underlying theme for our research and development work in the field of biomedical multimedia databases.

The evolution of biomedical information systems over the past two decades may be broadly described as:

- making the information system electronic, and
- making the information system intelligent.

### 1.1 The electronic information system

In this phase the information sources are computerized: electronic data is created from paper and film; data is archived and organized consistently with a query retrieval scheme; data dissemination methods are created for user access; all of these steps are integrated into a computer-controlled system for which software is created. Methods and technologies that have been critical for this phase include electronic image capture by film scanning, text conversion by Optical Character Recognition (OCR), relational or object-oriented databases, modern software development languages such as C++ and Java, and, especially, the technology and widespread usage of the World Wide Web for information dissemination. This phase is characterized by the use of technologies that have gained widespread acceptance, have a record of success, and may be considered largely known. For example, if we want to build a text and biomedical image database, and make the data retrievable over the Web, we know how to approach the problem. In this paper we provide a retrospective of the work we have done in this phase in making the information in two nationwide health surveys

---

\* [long@nlm.nih.gov](mailto:long@nlm.nih.gov); phone 1 301 435-3208; fax 1 402-0341; <http://archive.nlm.nih.gov>

available in electronic form through a multimedia database accessible over the Internet, by the Web-based Medical Information Retrieval System (WebMIRS).

### **1.2 The intelligent information system**

The information sources, which have been converted to electronic form, are exploited by automated and semi-automated methods, and the non-textual part of the information sources assumes a major role of interest: “intelligent” algorithms index the information sources by mathematical characteristics beyond the capacity of conventional human indexing; algorithms classify the data into categories at various semantic levels; non-textual query methods are incorporated; data are grouped according to “similarity” or “fuzzy membership” in particular categories of interest; navigating or browsing data in “similar” categories, relative to an input example, is supported. Methods and technologies critical for this phase include image segmentation, pattern recognition, data clustering, and neural networks. Contributions from these and other closely-related methods and technologies have created a research area of sufficient interest and scope to have its own name: Content-Based Image Retrieval (CBIR). Making the system intelligent is characterized by the use of technologies that are still emerging, whose success is restricted to limited domains, and which may be considered to be largely unknown. We are in the research phase in understanding how to build a successful CBIR system for a collection of biomedical images of significant size—regardless of the anatomy or image modality. We provide the results of the work we have done in this phase toward building a CBIR system for the health survey text and image data.

### **1.3 Comment**

While it is useful to think of the evolution of biomedical multimedia information systems in the past two decades in these phases, this description, like any broad categorization, is inexact. For example, there is significant research work continuing in biomedical image compression, which we would place in the first phase of making the information electronic. Therefore it is not accurate to think of the work of that phase as consisting of merely applying methods that are completely known. On the other hand, for the second phase, CBIR has made notable headway with a few implementations such as the QBIC<sup>2</sup> system that has been applied to various image types, including art reproductions, U.S. postage stamps, and trademarks. Therefore, intelligent information systems are not a completely uncharted territory. However, there remain numerous research challenges in this research phase, and these are the areas of our focus.

## **2. THE MULTIMEDIA DATABASE: WEBMIRS**

### **2.1 Overview**

WebMIRS is an R&D biomedical multimedia database system that is currently providing access to large parts of the data collected in two U.S. nationwide health surveys, including the x-ray images collected in one of the surveys. WebMIRS has both U.S. and non-U.S. users in the academic and corporate worlds, as well as in hospitals and medical centers, civilian government and the military, and users who access the system for personal use. Self-reported user interest areas includes image processing, epidemiology, information technology, education, rheumatology, and “general research”. In this section we give a history of the evolution of the WebMIRS system from the point of collection of the system data to the current time.

### **2.2 Data collection, digitization, archive**

The data of the WebMIRS system consists of major parts of the Second- and Third National Health and Nutrition Examination Surveys (NHANES II and NHANES III). The goals of the NHANES surveys include estimating prevalence of selected diseases, monitoring disease trends, monitoring trends in risk behaviors and environmental exposures, analyzing risk factors for selected diseases, studying the relationship among diet, nutrition, and health, exploring emerging public health issues, and establishing a national probability sample of genetic material for future genetic testing<sup>3</sup>. In this paper we discuss primarily the NHANES II data, since WebMIRS incorporates image data from that survey.

#### **2.2.1 NHANES II data**

NHANES II was conducted 1976-1980 and included participants aged 6 months to 74 years. For the NHANES II survey, the records contain information for approximately 20,000 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, approximately 10,000 cervical spine and 7,000 lumbar spine x-ray films were collected on survey participants aged 25-74. No x-rays were taken on pregnant women, and no lumbar x-rays were taken on women under 50. The pathologies of interest on these x-rays were osteoarthritis and degenerative disc disease<sup>4</sup>. This data was collected by the Mobile Examination Centers (MECs) of the National Center for Health Statistics (NCHS), part of the Centers for Disease Control and Prevention (CDC). The MECs, consisting of four trailer trucks connected with walkways and organized into medical stations, are shown in Figure 1, along with an x-ray station used to collect NHANES II spine films. The film was digitized using a Lumisys scanner at 146 dpi and stored on a combination of CD-ROM and magneto-optical (MO) media; at a later stage of the work, all of the CD-ROM images were converted to magneto-optical (MO) format, so that a complete collection of data, which totaled 140 GB, was available on this media. This media was initially accessed through a 144 platter Hewlett-Packard MO jukebox. A significant disadvantage of this storage was that not all of the 140 GB of survey data could be contained within the jukebox system, so that some data remained offline and required manual intervention to access. At a later stage of system development all data was migrated to a Sun A5000 RAID system, where it currently resides and is completely accessible online.

### **2.2.2 NHANES III data**

NHANES III was conducted 1988-1994, included participants as young as 2 months and had no upper age limit. The NHANES III WebMIRS database consists of approximately 30,000 records with 72 fields/record. The NHANES III text database at present contains data such as age, height, weight, race, sex, blood pressure, body mass index, summary results from the physician's examination of the upper and lower extremities, dietary change questionnaire results, and approximately 30 laboratory measurements including the results of serum cholesterol and plasma glucose tests<sup>5</sup>. Hand and knee x-ray films were collected by the NHANES III survey, but NCHS has not released this data for public use, so the WebMIRS NHANES III database contains no images. However, NCHS determined that the hand films should be digitized for archive and restricted research purposes. To support this effort, we developed a second retrieval/display/database system to support a multiple-reader study to determine the best spatial resolution level to use for the NHANES III hand x-ray images; two radiologists participated, using a set of 50 hand films that were digitized at each of three different resolution levels (50, 100, and 150 micrometers, respectively). The results, published in the *Journal of Digital Imaging*<sup>6</sup>, were implemented by NCHS in the subsequent digitization of these images. Images from this digitization have been used in subsequent research into the automated assessment of hand/wrist radiographs for arthritis, by the use of neural networks<sup>7</sup>.

### **2.3 Biomedical review**

Two workshops were convened at NIH under the sponsorship of the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) to obtain expert consensus on the question: "What radiological findings can be interpreted from the NHANES II spine images with a high level of inter- and intra-expert repeatability?" The consensus was that the biomedical features that may be repeatably interpreted from these images are anterior osteophytes, disc space narrowing, and spondylolisthesis, for the cervical spine; and anterior osteophytes, disc space narrowing, and spondylolisthesis, for the lumbar spine. These features were selected from a larger list of candidate features that were identified as "highly interesting" to researchers, but not susceptible of repeatable inter/intra-observer interpretation. C-spine and L-spine anatomy are illustrated in Figure 2. A sample of radiological features observable in the x-rays is shown in Figure 3. The candidate and final features from the NIH workshops are given in Table 1.

### **2.4 Quality control**

During the digitization process, a first level quality control was put in place by the scanning operators, who viewed a reduced resolution version of the scanned image, along with its histogram, to minimize errors due to bad film positioning on the scanner or to equipment failure or degradation. A second level quality control was conducted by NCHS survey data specialists, who viewed the images at 1Kx1K spatial resolution on PC monitors and verified that personal identifier codes were removed from the images, that image orientation was correct, and that the image contrast was not excessively light or dark. Anomalous cases were flagged and redigitized. A final, third level of quality control was then carried out by a medical expert in radiology, who reviewed images and films in a side-by-side display arrangement, using an E-Systems Megascan monitor to display the digital images and a lightbox to display the film. The medical expert followed a protocol that required him to check whether the digital images were of comparable visual quality as the film, in an overall sense, and whether the final features identified by the NIH workshops were visually observable in the digital images. The expert reviewed 2,051 image/film pairs for overall visual quality and scored 1,625 digital images

as being the same quality as the film, 400 digital images worse than the film, and 26 digital images better than the film. A total of 14,820 digital images were reviewed in standalone fashion, and the percentages where the pertinent features were judged to be observable were as follows: anterior osteophytes: 94.2%; subluxation (or spondylolisthesis in L-spine): 94.1%; and disc space narrowing: 93.7%.

## **2.5 System development**

### **2.5.1 WebMIRS**

#### **2.5.1.1 Overview**

Our initial work to create a multimedia database system was a Sun platform dependent application called the Medical Information Retrieval System<sup>8</sup> (MIRS). With the advent of World Wide Web technology, though, the opportunity arose to eliminate or greatly reduce platform dependencies in the system by adopting Web technology as our client/server model: the MIRS client was in effect replaced by the user's Web browser. Key design goals of WebMIRS were to (1) operate on any hardware platform which supports a standard Web browser and (2) require *only* a standard Web browser for user access.

The first WebMIRS<sup>9</sup> system, built using Java JDK 1.0.2, provided access to a small number of NHANES II records and associated x-ray images, and was successfully demonstrated at the 1996 Radiological Society of North America convention in Chicago, where it accessed an NHANES II database at NLM over a dedicated ATM high-speed link. WebMIRS was demonstrated at the Centers for Disease Control and Prevention (CDC) Data User's Conference in July 1997, and at the American College of Rheumatology meeting in November 1997, both in Washington, DC. In each case, it accessed NHANES databases at NLM over a T-1 Internet connection. Since these early demonstrations, WebMIRS has evolved into a highly functional multimedia database system that provides health survey text and image data across the Web. The current WebMIRS has features that include query building with a point-and-click interface, integrated text/image display, and capability to save query results to the user's local hard disk. An example WebMIRS SQL query, in English, is "Find the records of all survey participants aged 60 or older who have had severe back pain on most days for at least two weeks." When the results of the query are returned, images are presented at the top of the screen in a sliding window, while the text for the matching records is showing in tabular form at the bottom (See Figure 4a.) The user may change the current image, which is always highlighted in a red rectangle (not visible in the black and white illustration), by clicking on any record within the text data at the bottom. The WebMIRS NHANES II database also contains quantitative data for a subset of 550 of the images; this data consists of 9-point boundary landmarks placed on the vertebrae in these images under supervision of a board-certified radiologist. (See Figure 1, right.) This quantitative vertebral data may also be used for WebMIRS query and retrieval. Figure 4b gives an example of the results when the query of Figure 4a is repeated, with the additional input: "Restrict the query to the 550 records having the radiologist landmarks, and return the cervical spine anterior/posterior height ratios for the records matching the query." The radiologist landmarks are also displayed on the images in this example. The present WebMIRS system is version 1.0.10 and is deployed as a Java application using Sun's Java Web Start 1.0.1 technology. Most WebMIRS users operate on the PC platform.

#### **2.5.1.2 Compression**

At the current time, the x-ray images available through WebMIRS are reduced in spatial resolution by a factor of four both horizontally and vertically. However, work is underway to create a lossy-compressed version of the x-ray images based on Wavelet Transform technology and to make this Wavelet-compressed version available through WebMIRS. The particular method being researched is called hybrid multiscale vector quantization (HMVQ)<sup>10,11</sup>, and incorporates both vector and scalar quantization.

#### **2.5.1.3 Extending WebMIRS**

The WebMIRS system has attracted interest from a number of researchers interested in applying similar technology to making their own text and image databases usable over the Web. We are currently analyzing the WebMIRS architecture to address two perceived additional needs within the research community: (1) capability for users to record information to the database(s) from distributed sites, and (2) incorporation of a practical degree of flexibility in the database schema to allow the database to evolve beyond the table/field organization that was implemented at time of the initial creation of the database.

### **2.5.2 Digital atlas**

A Digital Atlas of the Cervical and Lumbar Spine was developed, using a subset of these images which were interpreted to consensus by a panel of three medical experts convened by the NIAMS. This Atlas, containing representative spine images at various grades of severity for each of the workshop-identified features, is publicly available from the Web and on CD-ROM. The Atlas is Java-based and has image processing functions for contrast enhancement, and the capability to add and annotate a user's private collection of images.

The increasing use of digital medical images requiring expert interpretation has given rise to the need for convenient online digital reference tools, to assist in producing interpretations that conform to recognized standards. We developed the Atlas<sup>12</sup> in collaboration with NIAMS and NCHS to fill a perceived need for such reference data for osteoarthritis in the cervical and lumbar spine, especially since a standard reference<sup>13</sup> of photographs of these features is out of print and difficult to obtain. Important features of the Atlas include: (1) presentation of standard reference images for a subject area (osteoarthritis of the cervical and lumbar spines) not previously addressed by digital atlases, to our knowledge; (2) single/multiple image display; (3) image processing for contrast enhancement; and (4) capability to add user-provided images, without code modifications, and to annotate these images graphically and with text. Color and grayscale images may be added in JPEG, TIFF, PNG, GIF or flat file formats, in color or grayscale. An example of an Atlas display is given in Figure 5. In this example, four Atlas images illustrate anterior osteophytes with varying degrees of severity. The Atlas is currently available for downloading from the CEB Website, or as a CD.

### **2.5.3 X-ray archive**

All 17,000 NHANES II x-ray images have been made publicly available through an FTP archive that is publicized on the CEB website. Users have accessed the images for use in medical research, medical education, image processing, compression, display work, database research, art and illustration, and physiology/kinematics studies. Users have reported employing the images in four Ph.D. theses, including two published recently<sup>14,15</sup> where extensive use has been made of the spine images for segmentation work. To view these images in full spatial and 12-bit grayscale resolution, CEB has developed a Java image viewer that is publicly available from the same site. 550 of these images have been converted to standard TIFF 8-bit form and made publicly available also, along with the 9-point radiologist marks.

## **2.6 New opportunities**

Although medical experts made the consensus decision that only the biomedical features marked with (\*) in Table 1 are repeatably interpretable in the x-ray films, image processing of the digital images raises the possibility that some of the other significant features of Table 1 can be observed. An unsharp masking technique that we have used appears to give good visual detail of posterior osteophytes, for example, and the range of visible spine anatomy is also extendable by this technique. For example, thoracic vertebrae T10-T12 become visible on some lumbar images, as well as lower cervical spine vertebrae (C6-C7) that are otherwise not visually perceptible, or only vaguely perceptible. An example of a lumbar spine image processed this way is shown in Figure 6.

## **2.7 Non-spine x-ray use**

Are the x-ray images useful for other than spine data? We see evidence that exploitation of the images for purposes other than spine information may be possible. Researchers in Spain<sup>16</sup>, for example, have developed image processing algorithms for the automatic localization of landmarks within the skull and the extraction of geometric measurements derived from these landmarks. This process has application to the practice of orthodontics. The process, when carried out manually, can take 10-15 minutes per image, so it is an area where image processing can potentially have a practical impact.

## **3. CONTENT-BASED IMAGE RETRIEVAL (CBIR)**

### **3.1 CBIR overview**

Our current research is directed toward extending and developing methods for computer-assisted indexing and retrieval of these images by image content. For indexing, extensive segmentation work and investigation of alternative methods of representing shape has been carried out, as well as neural network experiments in classifying images as normal/abnormal for anterior osteophytes and disc space narrowing. For retrieval, comparative recall work across several shape representations and matching methods has begun, and a prototype CBIR system has been created.

## **3.2 Indexing for CBIR**

### **3.2.1 Segmentation**

A large part of our segmentation research has concentrated on Active Shape Modeling (ASM), although we have not ignored other approaches, in particular the active contour approach. The ASM formulation that we have followed is that described by Cootes and Taylor<sup>17</sup> of the University of Manchester. For CEB segmentation work, ASM has represented a significant advance beyond heuristic, edge detection methods which have yielded very little promise of success in segmenting irregular, noisy images, into the domain of model-based, statistical, deformable template methods.

#### **3.2.1.1 Motivating work**

Researchers associated with Cootes carried out successful segmentation on images of the lower thoracic and upper lumbar spine (T7-L4) acquired by dual x-ray absorptiometry (DXA), as reported by Smyth<sup>18</sup>. Smyth used a single shape template to model the 10 T7-L4 vertebrae in 78 DXA images acquired from females aged 44-80. To segment, the user was required to manually initialize the ASM search by anchoring the template with three manually-placed points, after which the algorithm would deform and move the template to seek to the vertebrae. An example of the results reported by Smyth: successful ASM convergence, as compared to human performance, on 94% of the L4 vertebrae. This work provided the most complete published results of successful semi-automated segmentation we had seen for spine images acquired with a modality similar (although likely of better contrast quality) to our x-ray images. An additional significant publication for us was the work of Gardner<sup>19</sup> in using active contour modes to obtain spine x-ray segmentation at a useful level in an interactive system for digitized lumbar spine image.

#### **3.2.1.2 ASM segmentation**

ASM operates with two models of the objects to be segmented: (1) a shape model, which characterizes the shapes of the individual objects, and (2) a grayscale model, which characterizes the expected image grayscale along the boundaries of the objects. Both models are statistical in nature and are derived from sample images that are assumed to represent the target images. The shape model is used to provide an initial template for the objects in a target image and to provide constraints on the range of shapes to which the deformed template may converge. The grayscale model is used to drive the deformation of the template from its initial pose (position, orientation, and shape) to a pose that is an optimal fit to the image grayscale data, subject to the constraints of the shape model.

This work has continued with collaborators at Texas Tech University. Initial research<sup>20,21</sup> was carried out using 80-point vertebral models that were built by hand from each of 40 C-spine images, spanning from the bottom of C2 through C6; leave-one-out testing was done to evaluate the performance of ASM segmentation on each of these same images. (I.e. the shape and grayscale models used for a particular ASM run were constructed from the statistics of the shape and grayscale in 39 images, excluding the particular image being segmented. This process was carried out sequentially for each of the 40 images.) This work was preliminary and focussed on the dependency of ASM performance on initial pose of the template; it addressed the issue of ASM segmentation accuracy to a coarse level. ASM performance was measured by computing the mean squared error (MSE) between the converged ASM boundary and the manually-created "truth" boundary. The 40 tests were run twice: in the first run, the initial template pose was taken as the average of the pose in each of the 39 images that contributed to the shape and grayscale models used on the test; in the second run, the pose was derived from an algorithm that estimated spine region location and orientation through line integral processing. The results showed that in 35 of the 40 cases, the MSE was reduced when the second method of pose selection was used: the average MSE was reduced by a factor of about four. For the five cases where the MSE was not reduced when the second method of pose selection was used, tests were again run where the initial template pose was obtained by manually placing the template near "truth": the protocol followed was to manually match C4 from the template with "truth" C4, the C4 vertebrae being the approximate center vertebrae in the template that was used. Results of these five tests showed improved ASM performance in the manual placement, relative to the other two methods, in four of the five cases. Conclusions of this work were that (1) significant improvement in ASM performance on these images is feasible with an automated template placement method and (2) some anomalous cases exist where even a good initial pose will not assure good ASM performance. As part of this work, the Landmark Tool was developed in MATLAB to collect boundary landmarks; this tool provides on-screen profiles of both grayscale and gradient values along line segments normal to the estimated boundary, at selected landmark points. Using these, the boundaries may be marked while being guided not only by visual edge information, but also by measured mathematical properties along the

visual edge. The Landmark Tool is shown in Figure 7. Our project has also developed its own implementation of ASM, which is shown in Figure 8.

Our ASM research has continued with work to understand reasons for poor performance in certain cases, to further characterize performance of the algorithm on L-spine as well as C-spine images, to further explore options for initialization, and to develop solutions. A study was undertaken to assess the performance of the shape model and the grayscale model independently. In this study, 80-point C2-C6 shape and grayscale models were built from each of 100 C-spine images. The ASM shape model behavior for these images was analyzed by studying one key assumption used in the algorithm: the individual components of the  $b$ -vector, used to generate shapes  $z$  by the equation:  $z = z_0 + \Phi b$ , where  $z_0$  is the mean shape in the images being sampled, and  $\Phi$  is a matrix of principal components (called “modes of variation”) in its columns. ASM relies on the assumption that the individual components  $b_i$  each have a Gaussian distribution. In the work done, it was determined that 98% of the shape variance in these 100 images could be represented by 43 modes of variation, so the particular  $b$  studied had 43 components, and  $\Phi$  had 43 columns, each corresponding to a basis shape or mode or variation. For each  $z^j$  of the 100 shapes in the images,  $b^j$  was calculated by  $b^j = \Phi^T (z^j - z_0)$ ; then, for each fixed  $i$ , the Gaussian property for the set  $\{b_i^j : j = 1, 100\}$  was checked with the Kolmogorov-Smirnoff (K-S) goodness-of-fit test and the Mardia test of normality (which checks both skewness and kurtosis of the distribution). All modes passed the K-S test; all but five of the modes passed both Mardia checks. To further understand whether the departures from the Gaussian are likely to have any significant impact, the distribution of each  $b_i$  was made to conform closely to a Gaussian distribution by methodically making small changes in the values of the  $b_i$ 's. The shapes reconstructed from these new values of the  $b$  vectors were observed to closely resemble the shapes from the unmodified  $b$ 's. Hence, even though some departures from a Gaussian distribution had been found, they only had small significance for shape reconstruction in these images. This essentially positive result for the shape model then placed the focus on the grayscale model. The behavior of the grayscale model in ASM processing of these images was studied by (1) computing the “localization error” for ASM in these images as compared to images of two types different from the x-rays: ears and a spine phantom; and by (2) computing and inspecting grayscale profiles for each of the 80 landmark points used, across the 100 images. The localization error was the MSE distance between the converged and “truth” shapes when the algorithm was run after it was initialized to the actual truth value for the shape: ideally, the grayscale model should not move the template under this condition. However, the localization error was seen for the x-rays to be substantially greater than for either the ear images or the spine phantom. When the grayscale profiles were computed and graphed for each landmark, it was observed that for the ear and phantom images, the mean profiles curves had easily discernable shape characteristics at each landmark, implying that it is meaningful to speak of the average behavior of a grayscale profile for a landmark in those images. However, for the x-ray images, the profiles were essentially flat, for many of the landmark points, implying that the behavior of the profile was close to zero-mean noise. The conclusion reached was that, for the x-ray images, the grayscale model did not provide a reliable method for moving the templates in the ASM algorithm. A second difficulty encountered was the problem of modeling vertebrae with significant shape distortions in the corners due to the presence of anterior osteophytes. Normal vertebrae exhibit smoothly rounded corners; vertebrae with osteophytes show distorted corners, with the shape and severity of the distortion being highly variable. For practical use, our system needs to be able to detect the full range of corner distortion due to osteophytes, yet it is difficult or impossible a priori to be sure that we have included in our shape model images that represent all of the shape distortions that we may encounter. In addition, the high variability in shape of these corner distortions makes the collection of consistent, corresponding boundary landmarks problematic. This suggests the need for an augmentation to the ASM segmentation that is less model-driven and more image-data-driven in these local areas.

### 3.2.1.3 Active contours

This work has been carried out in collaboration with Dr. Hemant Tagare of Yale University. The particular approach taken is to (1) create a model of the object to be segmented in the form of a shape template; (2) interactively position this template close to the object to be segmented; (3) automatically generate a search grid in the neighborhood of this template's boundary; and (4) minimize an energy functional over all possible feasible curves within this grid. Strong points of this approach have been: (1) it incorporates a priori information (with the shape template) into the standard

active contour method; (2) optimizing over a search grid guarantees that a global minimum (relative to the grid) is computed for the energy functional; and (3) the incorporation of dynamic programming makes the optimization step very fast. Disadvantages have been that (1) the method is highly local in nature and thus strongly dependent on good initial placement of the template and (2) creation of the search grid has been done by drawing line segments orthogonal to the template at various points on the template boundary; at sharp corners of the template, this may result in line segments crossing and segmentations that are not simple closed curves. With both ASM and active contours, it has also been observed that vertebra boundaries will sometimes be missed because they are close to stronger edges formed by tissue/background interfaces.

#### 3.2.1.4 Unified segmentation

A comprehensive approach to segmenting the vertebrae is being developed that addresses the problems of (1) template pose initialization, (2) grayscale model information weakness for these images, and (3) difficulty in shape modeling for the vertebral corners.

For template pose initialization, we have focused on the Generalized Hough Transform<sup>23</sup> (GHT). The Generalized Hough Transform is an algorithm inspired by the standard Hough Transform, which is commonly used to detect straight lines in images. The generalization of this algorithm allows the detection of arbitrary shapes. The GHT as applied to the spine x-rays operates on a template of the vertebrae and finds the best match in the image to the template while varying scale, position, and orientation. The matching is effectively an exhaustive search of the entire image, using the Hough “bin counting” method. Because the GHT is not a deformable template method, the resulting match cannot be expected to agree with the vertebrae in fine detail: the matching accuracy is limited by that obtainable by modifying the position, scale, and orientation of the input template. Nevertheless, initial tests obtained by this technique have shown<sup>24</sup> robust capability to match a vertebral template to the “truth” vertebral pose, up to possible shifts of the template along the spine by integral numbers of vertebrae. For example, a template created from vertebrae C2-C5 might be matched by the GHT to C3-C6 instead. Because the matching method is an exhaustive search, the method as currently implemented is very time consuming. Current collaborative research is investigating methods to increase the accuracy of the matching and to reduce the computation time.

To deal with the problem of the weak informational content of the grayscale for these images, logic for a new grayscale model has been incorporated into the ASM algorithm. Rather than operating on the original grayscale image, this new logic first creates an edge-enhanced image by (1) applying unsharp masking to the original image to enhance edges; (2) thresholding to produce an edge-only image; and (3) blending the edge-only image with the original image by averaging; then the ASM grayscale model is constructed in the conventional way, operating on these blended images. In a test of 100 NHANES II x-ray images, it has been shown that the ASM convergence accuracy obtained with this new grayscale model is clearly improved over the old grayscale model, and in fact is comparable to the accuracy obtained when ASM is applied to a computer-generated vertebra phantom<sup>22</sup>.

To improve the segmentation of the vertebral corners, in light of the difficulties encountered in modeling these highly-variable, finer detailed structures in the ASM shape model, an augmentation of the ASM segmentation is being researched. This Deformable Model (DM) method is based on active contour concepts that seek a contour along which the sum of external energy and internal energy terms are minimized. A template with landmark points is used to define the initial contour. In our formulation, the external energy is taken as the Mahalanobis distance between the mean grayscale profile at a landmark point and the observed grayscale profile, summed along all landmark points on the contour; the internal energy term is composed of a proximity term, which measures distance between the current contour and the initial template, and a smoothness term, which measures the distance between each landmark point and the template, relative to the corresponding distance for the adjacent landmark point. Each term is weighted with an empirically set value. Deformation of the initial template occurs by finding a contour that minimizes the above energy terms through exhaustive search of a grid of potential contour points. This grid is created to lie in the neighborhood of the initial template. The placement of the initial template is done so that the template lies in the neighborhood of the vertebra corner, as determined by convergence of the ASM algorithm. Initial tests of this approach have yielded mixed results; in tests using 100 NHANES II C-spine images, no overall improvement in segmentation accuracy was observed, when a DM step was added to GHT+ASM for segmentation. However, a similar test with 100 L-spine images showed that acceptable segmentations (i.e., errors were as given by the following paragraph) were improved from 47% to 49%

with the addition of the DM step<sup>15</sup>. Work is continuing to investigate DM for the accurate segmentation of vertebral corners.

For the performance tests of these segmentation approaches, “ground truth” was taken as hand segmentations done by engineers, which used the radiologist 9-point marks as guides, and used the Landmark Tool for observing local grayscale and gradient profiles during point placement. For the C-spines, templates consisted of 80 landmark points and covered the vertebral range from the bottom of C2 through C6. For the L-spines, templates consisted of 200 landmark points and covered L1-L5. Segmentation error was calculated as the root sum square of the landmark point distances between the converged shape and the “ground truth”. Overall results were that, for the C-spines, GHT+ASM+DM achieved an error less than 3 mm in 75% of the cases; for the L-spines, GHT+ASM+DM improvement achieved an error of less than 6.4 mm in 49% of the cases<sup>15,22</sup>. Work is continuing on the improvement of this unified segmentation approach.

### 3.2.2 Indexing by shape

The next step after shape segmentation in a CBIR system is representing the shape boundary information. The dense boundary points extracted as (x,y) coordinates in the image space need to be represented in a form suitable for archiving, indexing, and similarity matching. For this they are reduced to a small set of meaningful representative points by a shape representation algorithm. This coarse boundary and a binary image representation of the vertebra are used to find meaningful shape features that are invariant to translation, rotation, scaling and starting-point shift.

Most shape-based CBIR methods, to date, have been applied to art, trademark databases, fish images, silhouettes of tools, etc. From the literature, we observe that most shape representation methods use global shape characteristics for indexing; i.e. the final shape representation is controlled by the distribution of all the boundary points in the image space. Such an approach may not be suitable for biomedical shapes due to (i) high similarity across anatomical shapes, such as vertebrae (ii) loss of subtle differences in boundary representation which could be indicative of certain pathology, and (iii) inadequacy of shape representation methods for supporting region-localized queries. The methods currently selected include invariant moments, as defined by Hu<sup>25</sup>, scale-space filtering<sup>26</sup>, polygon approximation<sup>27</sup>, and Fourier descriptors<sup>28</sup>.

While these methods are able to determine overall shape similarity, they are found to approximate for subtle pathology in biomedical images that may be critical for the information provided by the system. We have conducted an evaluation on the recall performance of the above methods on vertebra shapes<sup>29</sup>. It was found that the best method, polygon approximation, achieves 58.44% precision and 57.57% recall rates. This is fairly unimpressive. However, on closer analysis we determined that the low scores are due to incorrect ranking of similar shapes. We have used this analysis to develop a new shape representation algorithm that retains the pathology in the vertebra shapes<sup>30</sup>. Concepts of polygon approximation are illustrated in Figure 9; the basic descriptor is turn angle versus cumulative curve length, normalized by total curve length.

Most techniques, such as global shape properties or scale-space filtering, lose or fail to detect local details. As the result of this shortfall, the number of retrieved images is so high that the retrieval result is sometimes meaningless. To retrieve a small number of best matched images, shape representation and similarity measurement techniques must distinguish shapes with even minor variations. The main challenges of indexing by shape are to define a shape representation method that is invariant with respect to rotation, translation, scaling, and the curve starting point shift, and to measure the similarity among the shapes. A polygon curve evolution technique has been developed for smoothing polygon curves and reducing the number of data points while preserving the significant pathology of the shape. The x and y coordinates of the simplified boundary points can then be converted into a bend angle versus normalized curvature length function to represent the curve. Shape data similarity can be calculated by finding the  $l_2$ -norm of the Fourier descriptors between two shapes being compared<sup>30</sup>.

### 3.2.3 Biomedical feature classification

Automated or computer-assisted classification of biomedical features is potentially of great significance for future biomedical information systems. With effective biomedical feature classification tools, the biomedical indexer of the future will be able to efficiently add important classification information such as normal/abnormal for osteophytes, disc

space narrowing, etc. We have conducted research in feature classification in collaboration with Dr. Joe Stanley of the University of Missouri.

### **3.2.3.1 Anterior osteophytes**

The approach to the classification research was as follows: obtain “truth” classifications for spine vertebrae from medical experts; obtain segmentations for spine vertebrae; classify the vertebrae using automated techniques; compare the automated results against the expert “truth”. All classifications of the vertebrae were carried out using artificial neural network technology. For the anterior osteophytes of the cervical spine, a total of 704 vertebrae were used, with 352 “truth” classified as normal and 352, as abnormal, by a board-certified radiologist. For a vertebra shape, 32 features were derived for use in classifying the shape as normal/abnormal for anterior osteophytes. These features included radius of curvature and gradient measures along the shape, and a mathematical morphology feature that measures how much the shape protrudes from its average local neighborhood. (See Figure 10, left.) The vertebrae data was divided into training, validation and test sets, following standard practice in development of neural network classifiers: the weights in the network were iteratively modified to make the network classifications approach the “truth” classifications for the training data; this training phase was terminated when the classification on the validation data reached a minimum; then the network was used to classify the test data, and the network performance was taken as the rate of correct classification achieved on this data. On the test data, an overall agreement score of 85% was achieved, as compared to the given “truth”. Similarly, for the anterior osteophytes of the lumbar spine, a total of 782 vertebrae were used, with 391 “truth” classified as normal and 391, as abnormal, by the same board-certified radiologist. The same feature set and test procedure was used for the lumbar spine as for the cervical spine. An overall agreement score of 71% was achieved, as compared to the given “truth” for the lumbar spine. It was noted that the poorer performance of the network for the lumbar spine was perhaps due in part to the lower contrast of these images and the resulting ambiguities in segmentation.

### **3.2.3.2 Disc space narrowing**

In order to investigate classification techniques for disc space narrowing, a reference set of “truth” readings was obtained from collaborating radiologists at Phelps County Medical Center in Rolla, Missouri. Using the CEB-developed Digital Atlas of the Cervical and Lumbar Spine as a standard, 50 cervical spine images were interpreted for disc space narrowing. Specifically, levels C3/4, C4/5, C5/6, and C6/7 were interpreted in each of these images. Two radiologists carried out the interpretation independently. Similarly, 50 lumbar images were interpreted for narrowing, at L3/4, L4/5, and L5/S1, by one radiologist. An algorithm for assessing disc spacing was developed. The algorithm operates on an image region containing two adjacent, segmented vertebrae (and the space between them). The algorithm computes a “vertebrae separator”, a curve with points lying equidistant between the adjacent vertebrae boundaries; for each point on this separator, the “distance to a vertebra” is taken to be the Euclidean distance to the closest neighboring point on one of the vertebra. Figure 10 (right) illustrates the separator created by this algorithm.

## **3.3 CBIR prototype**

We are currently building a prototype CBIR system as a testbed for performance comparisons among the shape representation and similarity methods, and a demonstration tool for soliciting comments from the biomedical community. This tool incorporates indexing capability, with both manual and computer-assisted vertebral segmentation based on Active Contours, and retrieval using text, query by image example, and query by sketch. Segmented data is stored in an XML format, using data tags that are intended to provide useful semantic interpretations of the file contents. For example, tags are implemented to allow data users to distinguish among the level of expertise of the segmentors who produce the vertebrae segmentations stored. This allows us to accumulate segmentation data at the engineering level, for our continuing technical research, and at the medical expert level, as opportunities arise for medical experts to participate in the segmentation. Vertebral shapes are stored using each of the shape representation methods given above, to provide a framework for comparative evaluation among the various methods for retrieval performance. This prototype system is implemented using MATLAB and uses the MySQL DBMS for storage of the text data. Details of this prototype CBIR system can be found in Antani’s paper<sup>31</sup>. Both the indexing and retrieval capabilities of the prototype are illustrated in Figure 11.

## 4. SUMMARY

We have presented a retrospective view of our work in building a biomedical multimedia database with national health survey data, our ongoing research in content-based image indexing and retrieval, and our efforts to develop a biomedical information system based on CBIR. Our focus is on (a) developing robust algorithms for localizing and identifying anatomy relevant for that image class and relevant to the indexing goals, (b) developing algorithms for labeling the segmented anatomy based on its pathology, (c) developing a suitable indexing and similarity matching method for visual data, and (d) associating the text information on the imaged person, indexed separately, for query and retrieval along with the visual information. We are in the process of building such a system which includes a biomedical image database that will support content-based image retrieval in combination with queries on the text data, and intelligently fuse the results. The work in progress is a natural step after our development of WebMIRS, the Web-based Medical Information Retrieval System that provides access to the National Health and Nutrition Examination Survey (NHANES) text and image data. We continue to enhance a prototype content-based retrieval system that allows shape-based retrieval by example image and user sketch of vertebrae on the spine x-ray images while also providing access to associated text information.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the help of Lew Berman and Ami Ostchega of NCHS, Dr. Matthew Freedman, M.D., and Dr. Ben Lo, of Georgetown University Medical Center, Molly Dickens of Texas Tech University, Leif Neve, and Mike Bopf, Aquilent Corporation, Jim Goh, Reva Lawrence, NIAMS/NIH, Dr. Stanley Pillemer, M.D., NIDR/NIH, Dr. Alhalya Premkumar, M.D., CC/NIH, Dr. Daniel Abodeely, M.D. and Dr. Greg Cizek, M.D., both of the Phelps County (Missouri) Medical Center, Dr. Roy D. Altman, M.D., of the University of Miami, Dr. Nancy Lane, M.D., of the University of California at San Francisco, and Dr. W.W. Scott Jr., M.D., of Johns Hopkins University.

## REFERENCES

1. Chute CG. Clinical data retrieval and analysis: I've seen a case like that before, *Annals of the New York Academy of Sciences*, **670**, 1992, 133-140.
2. 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*, **3312**, San Jose, CA, January 28-30, 1998, 150-161.
3. National Center for Health Statistics Web site: <http://www.cdc.gov/nchs/about/major/nhanes/goals.htm>
4. *Plan and Operation of the second National Health and Nutrition Examination Survey 1976-80*, Programs and Collection Procedures, Series 1, No. 15, DHHS Publication No. (PHS) 81-1317, National Center for Health Statistics, Hyattsville, MD, July 1981, 4.
5. *Plan and operation of the Third National Health and Nutrition Examination Survey 1988-94*, National Center for Health Statistics. *Vital Health Stat* 1(32), 1994
6. Ostchega Y, Long LR, Hirsch R, Ma LD, Scott, Jr. WW, Thoma GR. Establishing the level of digitization for wrist and hand radiographs for the third National Health and Nutrition Examination Survey. *Journal of Digital Imaging*, **11**(3), August 1998, 116-120.
7. Duryea J, Zaim S, Wolfe F. Neural network based automated algorithm to identify joint locations on hand/wrist radiographs for arthritis assessment. *Medical Physics*, **29**(3), March 2002, 403-411.
8. Long LR, Berman LE, Thoma GR. A prototype client/server application for biomedical text/image retrieval on the Internet. *Proceedings of SPIE Storage and Retrieval for Still Image and Video Databases IV*, **2670**, San Jose, CA, February 1-2, 1996, 362-372.
9. 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*, **3312**, San Jose, CA, January 24-30, 1998, 392-403.
10. Yang S, Mitra S. Statistical and adaptive approaches for segmentation and vector source encoding of medical images. *Proceedings of SPIE Medical Imaging 2002: Image Processing*, **4684**, February 24-28, 2002, San Diego, CA, 371-382.

11. Yang S, Mitra S. Content-based vector coder for information retrieval. BISC-FLINT International Workshop, UC Berkley, Berkeley, CA, August 14-18, 2001.
12. 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*, **3035**, Newport Beach, CA, February 22-28, 1997, 586-594.
13. *The Epidemiology of Chronic Rheumatism*, vol. II of *Atlas of Standard Radiographs of Arthritis*. F. A. Davis Company, Philadelphia, PA. 1963.
14. Ghebreab S. *Strings and Necklaces: On Learning and Browsing Medical Image Segmentations*. Intelligent Sensory Information Systems Group, University of Amsterdam, Amsterdam, Netherlands, 2002.
15. Zamora G. *Dissertation for the degree of Doctor of Philosophy*. Department of Electrical Engineering, Texas Tech University, Lubbock, TX, December 2002.
16. Grau V, Alcaniz M, Juan MC, Monserrat C, Knoll C. Automatic localization of cephalometric landmarks. *Journal of Biomedical Informatics*, 2001, **34**(3):146-156.
17. Cootes TF, Taylor CJ. *Statistical models of appearance for computer vision*. Tech. rep., University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, Manchester, M12 9PT, U.K. 2001 February.
18. Smyth PP, Taylor CJ, Adams JE. Vertebral shape: Automatic measurement with active shape models. *Radiology*. 1999;**211**(2):571-578.
19. Gardner JC, Heyano SL, Yaffe LG, von Ingersleben G, Chesnut CHI. A semi-automated computerized system for fracture assessment of spinal x-ray films. *Proceedings of SPIE Medical Imaging: Image Processing*. 1996;**2710**:996-1008.
20. Sari-Sarraf H, Mitra S, Zamora G, Tezmoz A. *Customized active shape models for segmentation of cervical and lumbar spine vertebrae*, (contract technical report). 2000 December.
21. Zamora G, Sari-sarraf H, Mitra S, Long R. Estimation of orientation and position of cervical vertebrae for segmentation with active shape models. *Proceedings of SPIE Medical Imaging 2001: Image Processing*. **4322**, San Diego, CA, February 19-22, 2001, 378-387.
22. Zamora G, Sari-sarraf, Long R. Hierarchical segmentation of vertebrae from x-ray images. *Proceedings of SPIE Medical Imaging 2003: Image Processing*, **5032**, San Diego, CA, February 15-20, 2003 (forthcoming).
23. Ballard DH. Generalizing the Hough Transform to detect arbitrary shapes. *Pattern Recognition*, **13**(2), 1981, 111-122.
24. Tezmoz A, Sari-Sarraf H, Mitra S, Long R. Customized Hough transform for robust segmentation of the cervical vertebrae from x-ray images, *Southwest Symposium on Image Analysis and Interpretation (SSIAI 2002)*, April 7-9, 2002, Santa Fe, New Mexico, 224-228.
25. Hu MK. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8:179-187, 1962.
26. Del Bimbo A, Pala P. Shape indexing by multi-scale representation. *Image and Vision Computing*, 1999, **17**(3-4):245-261.
27. Latecki LJ, Lakämper R. Shape description and search for similar objects in image databases. In R. C. Veltkamp, H. Burkhardt, and H. P. Kriegel, editors, *State-of-the-Art in Content-Based Image and Video Retrieval*, volume 22 of *Computational Imaging and Vision*, Kluwer Academic Publishers, 2001, 69-96.
28. Zahn C, Roskie R. Fourier descriptors for plane closed curves. *IEEE Computer*, 1972,C-21(3):269-281.
29. Antani S, Long LR, Thoma GR, Lee DJ. Evaluation of shape indexing methods for content-based retrieval of x-ray images. *Proc. of IS&T/SPIE 15th Annual Symposium on Electronic Imaging, Storage and Retrieval for Media Databases*, **5032**, 2003.
30. Lee DJ, Antani S, Long LR. Similarity measurement using polygon curve representation and Fourier descriptors for shape based vertebral image retrieval. *Proceedings of SPIE International Symposium on Medical Imaging: Image Processing*, **5032**, 2003.
31. Antani S, Long LR, Thoma GR. A biomedical information system for combined content-based retrieval of spine x-ray images and associated text information, *3<sup>rd</sup> Indian Conference on Computer Vision, Graphics, and Image Processing (ICVGIP '02)*, Ahemdabad, India. December 16-18, 2002, 242-47.



Figure 1: Mobile Examination Center (MEC) of the NCHS

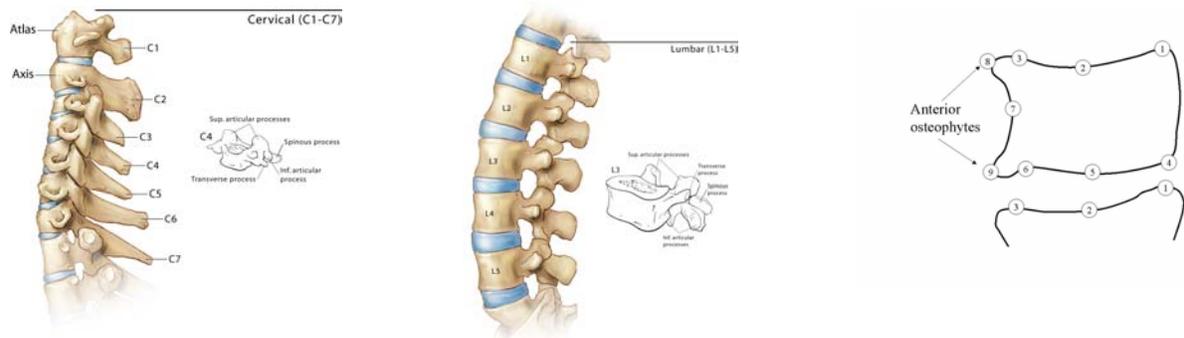


Figure 2. (Left to right) Cervical spine, lumbar spine, 9-point radiologist marks, with anterior osteophytes shown.

Anterior osteophytes*
Disc space narrowing*
Spondylolisthesis (L-spine)*
Subluxation (C-spine)*
Posterior osteophytes
Plate erosion/sclerosis
Vacuum phenomenon
Abnormalities (if any noted)
Ankylosing spondylitis
Apophyseal OA
Congenital/Developmental Disease (specify)
Degenerative Disc Disease
DISH
Evidence of surgery (level)
Fracture: specify level
Infection
Disc calcification
Neuropathic spine
Osteopenia
Paget's disease: specify level
Rheumatoid arthritis
Spondyloarthropathy
Spondylosis Deformans (spurs)
Anterior ligamentous calcification
Congenital fusion (level)
Tumor (level)
Other (specify)

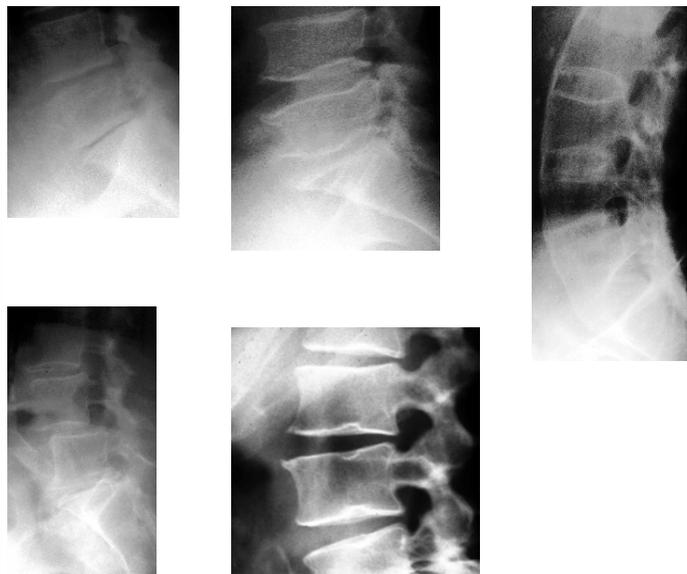


Figure 3. Sample of radiological features observable in NHANES II x-rays. Clockwise from upper left: disc space narrowing, osteophytes, fusion/biconcavity, Schmorl's nodes, dislocation

Table 1. Candidate biomedical features for the NHANES II x-rays. A (\*) indicates the final feature set of interest from the NIH workshops.

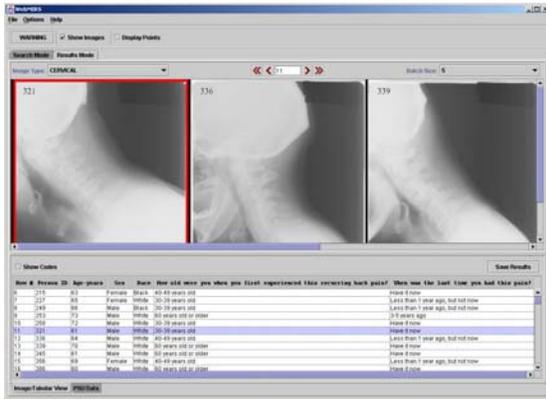


Figure 4a. WebMIRS query results.



Figure 4b. Same as Figure 1a, but with C-spine anterior/posterior height ratios returned also.

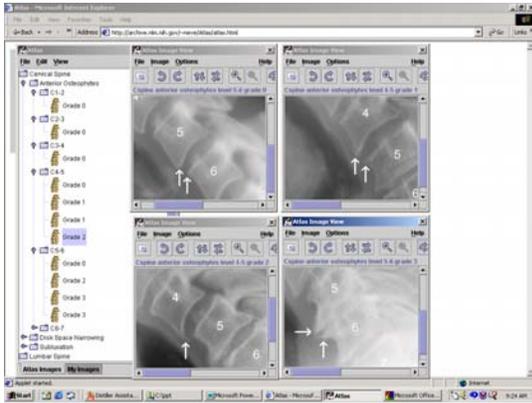


Figure 5. Digital Atlas of the Cervical and Lumbar Spine

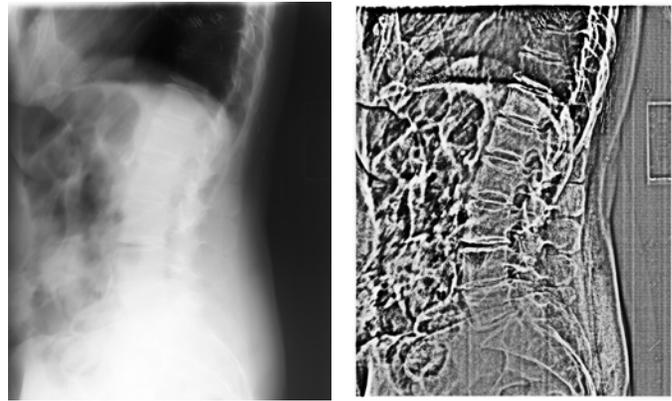


Figure 6. Lumbar spine image before and after unsharp masking. After the processing, posterior boundaries are visible, plus vertebrae from T10 down through the sacrum.

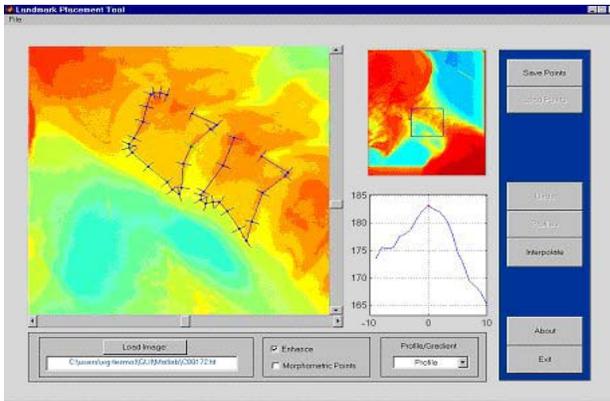


Figure 7. Landmark Tool.

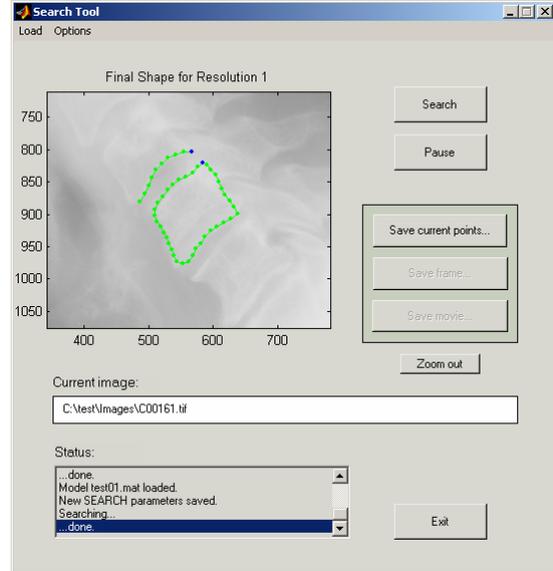


Figure 8. Active Shape Modeling tool.

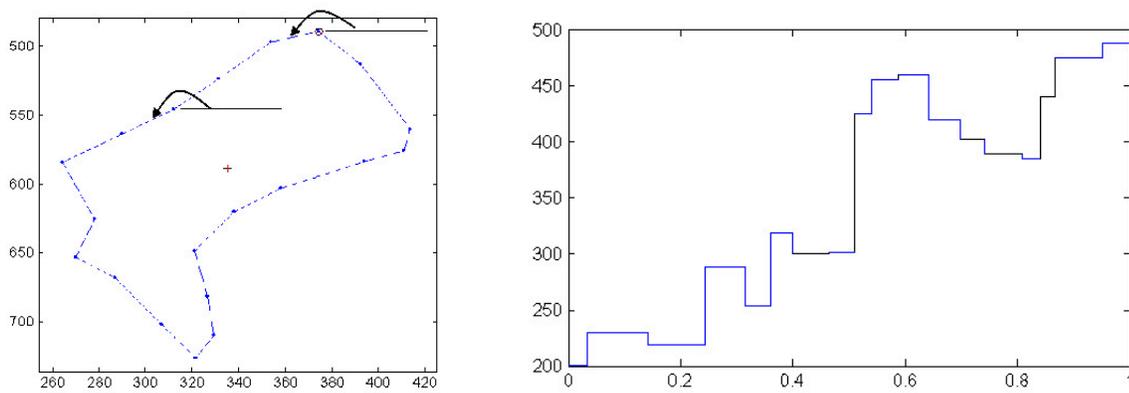


Figure 9. Polygon approximation shape description. At left, polygon approximation to vertebra is shown, with turn angle illustrated. At right is plot of turn angle versus normalized curve length.



Figure 10. Some features used in the neural network classification. For osteophytes, morphological smoothing was done to compute one feature (left). At right the separator used to compute intervertebral distances for disc space narrowing is shown.

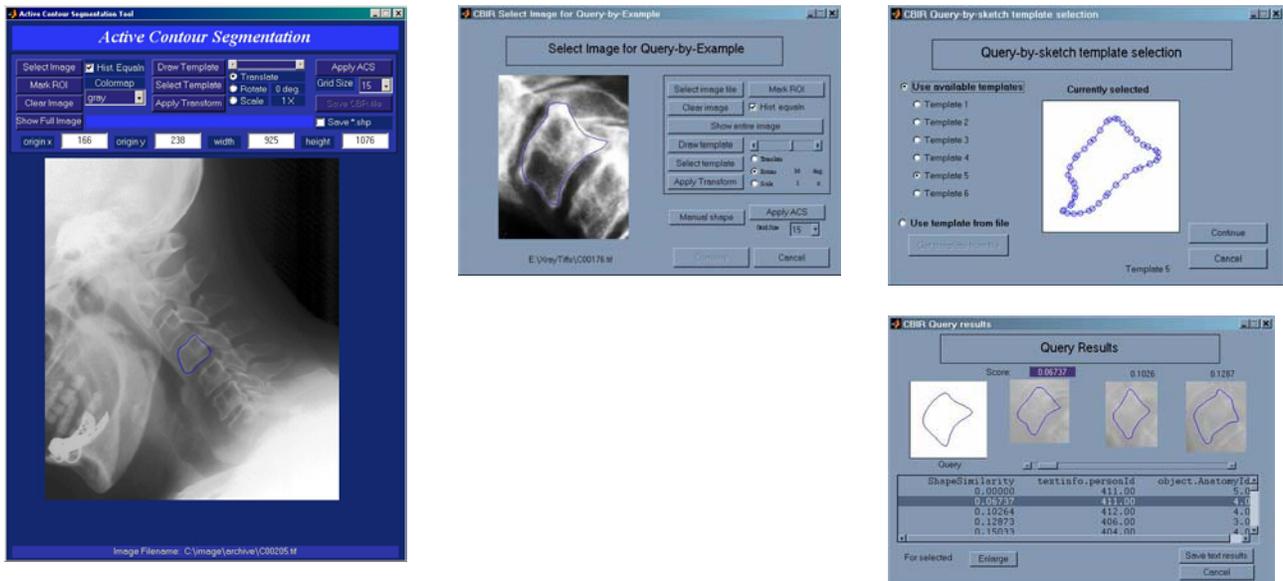


Figure 11. The prototype CBIR system. At left is the segmentation screen, for indexing. Center is the query-by-example screen; top right is query-by-sketch, and bottom right shows ranks results of one query-by-sketch.