ELSEVIER

Computerized Medical Imaging and Graphics 29 (2005) 171–193

Computerized Medical Imaging and Graphics

www.elsevier.com/locate/compmedimag

Image informatics at a national research center

L. Rodney Long*, Sameer K. Antani, George R. Thoma

Department of Health and Human Services, Communications Engineering Branch, US National Library of Medicine, Lister Hill National Center for Biomedical Communications, National Institutes of Health, 8600 Rockville Pike, Bldg 38A MS 55, Bethesda, MD 20894, USA

Received 1 April 2004; revised 25 August 2004; accepted 30 September 2004

Abstract

Image informatics at the Communications Engineering Branch of the Lister Hill National Center for Biomedical Communications (LHNCBC), an R&D division of the National Library of Medicine (NLM), includes document and biomedical images. In both domains, research into computer-assisted methods for information extraction, and the implementation of prototype systems incorporating such methods, is central to our mission. Current document image research focuses on extracting bibliographic data from scanned journal articles. Current biomedical imaging work focuses on content-based image retrieval (CBIR) and related problems in segmentation, indexing, and classifying collections of images of the spine and of the uterine cervix.

Keywords: Content-based image retrieval; Document image; Spine; X-ray; Uterine cervix

1. Introduction

As a creator and provider of biomedical information, including images, the National Library of Medicine, through its R&D divisions (particularly the LHNCBC), has several image informatics projects. Both document imaging and biomedical imaging are of interest, especially the automatic extraction of bibliographic information from the former for populating MEDLINE, and the recognition of shape, color, and texture in the latter for image indexing purposes.

In this paper we provide a broad overview of image informatics work done by our branch within the LHNCBC. We first briefly discuss work in the area of document imaging and provide a summary of major projects and results (Section 2). In subsequent sections, we transition to

Acronyms and terms AAM, active appearance models; ACS, active contour segmentation; AO, anterior osteophytes; API, application program interface; ASM, active shape models; ATM, asynchronous transfer mode; BMT, boundary marking tool; CBIR, content-based image retrieval; CCD, charge coupled device; CDC, Centers for Disease Control and Prevention; CEB, Communications Engineering Branch; C*n* ($n=1,\ldots,7$), cervical spine vertebra number *n*; Cspine, cervical spine; DIAU, document image analysis and understanding; DM, deformable model; DocMorph, NLM file conversion application; DocView, NLM application to aid Internet transmission of scanned documents; FTP, file transfer protocol; GHT, generalized Hough transform; GIF, graphic interchange format; HMVQ, hybrid multiscale vector quantization; HPV, human papillomavirus; JDBC, Java database connectivity; LBG, Linde–Buzo–Gray (compression method); LHNCBC, Lister Hill National Center for Biomedical Communications; L*n* ($n=1,\ldots,5$), lumbar spine vertebra number *n*; Lspine, lumbar spine; MARS, medical article records system (document-oriented); MIRS, medical information retrieval system (biomedical image-oriented); MDT, multimedia database tool; MECs, mobile examination centers; MEDLINE, NLM bibliographic database containing over 12 million biomedical references; MIRS, medical information retrieval system; MPPE, minimum point-to-point error; MyMorph, NLM file conversion application, enhancement to Docmorph; MySQL, a widely-used open source database management system; NCHS, National Center for Health Statistics; NCI, National Cancer Institute; NHANES, National Health and Nutrition Examination Survey; NIAMS, National Institute for Arthrititis and Muscuoloskeletal and Skin Diseases; NIH, National Institutes of Health; NLM, National Llibrary of Medicine; OCR, optical character recognition; PDF, portable document format; PNG, portable network graphics; PSM, partial shape matching; RAID, redundant arrays of inexpensive disks; *R*-table, table description of an object's shape used in generalized Hough transform; SAIL, system for automated interlibrary loan; SCJ, squamocolumnar junction; SQL, structured query language; TIFF, tagged image file format; WebMARS, web-based medical article record system (see MARS); WebMIRS, web-based medical information retrieval system (see MIRS); WILL, workstation for interlibrary loan.

* Corresponding author. Tel.: +1 301 435 3208; fax: +1 301 402 0341.
E-mail addresses: rlong@mail.nih.gov (L.R. Long), santani@mail.nih.gov (S.K. Antani), gthoma@mail.nih.gov (G.R. Thoma).

0895-6111/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compmedimag.2004.09.015

the major part of the paper, biomedical imaging. We are working with two major biomedical datasets: the 17,000 spine X-rays and associated health survey data collected in the second national health and nutrition examination survey (NHANES II), and the 60,000 uterine cervix images and associated clinical data collected in the National Cancer Institute's Guanacaste Project. Each dataset has its own requirements, partly determined by the characteristics of the specific images, but also determined by the user communities for each dataset. The biomedical imaging work done for each dataset is presented in separate sections in this paper, followed by an enumeration of critical issues that continue to confront the technical worker in this field, and our conclusion.

2. Document imaging

Current research focuses on document image analysis and understanding (DIAU) techniques applied to the problem of extracting bibliographic data from scanned journal articles to populate MEDLINE®, the NLM's major database of 12 million citations to the biomedical journal literature [1]. Research activities in page segmentation [2,3], zone labeling [4], zone text reformatting [5] and lexical techniques, and the incorporation of a commercial OCR system, have resulted in a production system that automatically produces the titles, author names, institutional affiliations and abstracts from the TIFF pages of journal articles. Rule-based algorithms that drive each stage of the data extraction rely on geometric and contextual features derived from the OCR output (which in addition to ASCII text provides bounding boxes and text attributes) as well as layout analysis. While these features are identified and incorporated into the rules by manual effort, recent work has been done in generating features dynamically [6–8]. Investigations have also been done to substitute rule-based labeling with an artificial neural network approach [9,10].

Lexical analysis has focused on pattern matching techniques to reduce the uncertainty (low confidence) of OCR output, and the use of historic affiliation data to correct the errors in small text and italics incorrectly recognized by the OCR. Both techniques are designed to reduce the manual labor involved in text verification [11,12].

In addition to this core research into DIAU algorithm design, the implementation of a robust production system (MARS, for Medical Article Records System) has required such ancillary work as: the use of speech recognition to enable hands-free operation of scanners [13]; the extension of the capabilities of the commercial OCR system to include the recognition of Greek letters and biomedical symbols [14]; the design of time stamp instrumentation to monitor the operational efficiency of the modules in the system; the development of an isomorphic technique to display multiple instances of the same detected character for rapid text verification [15], among others.

Recent work has focused on the extraction of bibliographic data from online (Web-based) medical journals indexed in MEDLINE [16]. This is motivated by the rapidly increasingly number of journals that are Web-based, and in many cases without paper equivalents. Here too the identification of suitable heuristic rules for algorithms to segment pages, label zones and reformat zone text is crucial. An added complication is that articles can appear in either PDF or HTML formats. The system, called WebMARS, first downloads and classifies the articles [17], converts the PDF files to HTML, and then the daemons implementing the zoning, labeling and reformatting algorithms operate on these files to extract bibliographic data. In contrast to the crisp rules used in MARS (required by the presence of OCR errors), fuzzy rules are used for labeling in WebMARS [18]. As a first step in creating fuzzy rules for labeling, feature statistics (normalized and smoothed histograms) are used to create membership functions for each label. This step is then followed by heuristic rules that rely on text content (words characteristic of a bibliographic element) and syntax of the element (e.g. @ symbol in an email address.)

These current lines of research have historic precedents at the LHNCBC. Since the early 1980s, document imaging has been an active research area in this organization, motivated originally by the possibility of preserving books and journals at the NLM as digital images in place of microfilm, and the accompanying need to understand the feasibility of the existing technologies, and the equipment and labor costs of a production system [19,20]. Our approach was to build a series of prototypes of a system for document capture, storage and retrieval [21]. These prototypes, controlled by PDP-11 class computers initially and later IBM PC machines, were used as testbed systems for experiments that addressed engineering feasibility, image enhancement and compression techniques, conversion rates and costs. A significant finding was that digital imaging costs and production times were comparable to those for microfilming [22,23].

A challenge at the time was the capturing of the page images of books and bound journals, many of these fragile, rapidly and safely. We built a 'book scanner' equipped with split trays that held a document face-up captured by a CCD camera [24]. This system went through three generations, each improving on the scanning resolution and image processing functions. Alternative approaches to document storage were evaluated: by optical disks [25] and RAID systems [26].

In the 1990s, subsequent work in document imaging was motivated by yet another NLM mission objective, viz. the interlibrary loan service. Here the approach was to build and test two systems, SAIL and WILL, each implementing a different vision of delivering documents (articles from biomedical journals) from one library to another. Both systems accessed NLM's automated request routing service to download interlibrary loan requests from health science libraries nationwide. The systems then parsed the requests

to identify the requester, the citation to the article requested, and the form in which the article was to be sent (fax, email or postal mail). SAIL automatically retrieved previously scanned articles from optical disk storage, converted the TIFF images to the appropriate formats and filled the requests [27,28]. The WILL system was designed under the assumption that most requested articles would not be scanned and stored, and therefore printed out an informative slip for an operator to retrieve the paper journal from the library's stacks. The operator would then scan the article, the only manual step, after which the machine would process the TIFF images to the appropriate format for meeting the request [29,30]. Experiments with these systems, at different medical libraries, showed that the WILL approach was more useful in light of the observed request patterns, leading to commercial acquisitions based on this design.

While these systems addressed interlibrary loan services (which deliver documents from one library to another, that in turn provides its patrons with the documents requested), a convenient way to deliver documents directly to users over the Internet was needed. This inspired the development of DocView, a Windows client software that allowed users to receive TIFF documents, enhance them and organize them for later use [31,32]. DocView has over 16,000 users in 190 countries. Most users receive documents from their libraries or document supply houses. In response to feedback from DocView users, a companion Web-accessible file-conversion service, DocMorph, was developed to allow users to upload files to a server and receive them in a desired format, e.g. TIFF images to PDF [33]. DocMorph, with over 11,000 users by 2003, is designed to convert 50 file formats to PDF, text (through OCR), and speech. This last facility has been used to serve sight impaired patrons by creating audio from TIFF documents [34].

While DocMorph has been used widely, in response to user feedback a bulk file conversion system, MyMorph has been developed as a user client. The MyMorph client software allows the upload of thousands of files at a time to the DocMorph server, thereby substantially reducing user interaction with the system [35]. This suggests a role for MyMorph in the file migration stage of an electronic preservation system, an objective of increasing interest to the library and archival communities.

3. Biomedical imaging for spine X-rays

3.1. NHANES databases

The National Health and Nutrition Examination Surveys (NHANES), conducted by the National Center for Health Statistics (NCHS), part of the Centers for Disease Control and Prevention (CDC), are a family of national health surveys that have been conducted for more than 30 years. 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 [36]. The second and third surveys (NHANES II and NHANES III, respectively) are major biomedical datasets that we have incorporated into our work [37] and are both described below. NHANES II has been a major source of image data for our work and that survey is emphasized in this paper.

3.1.1. NHANES II

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 7000 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 [38]. This data was collected by NCHS mobile examination centers (MECs). The MECs, consisting of four trailer trucks connected with walkways and organized into medical stations, are shown in Fig. 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 the resulting 140 GB of data is currently stored on a Sun A5000 RAID system.

3.1.2. NHANES III

NHANES III was conducted 1988–1994 and included participants as young as 2 months; this survey had no upper age limit. The NHANES III database consists of approximately 30,000 records with more than 2000 fields/record, including demographic data such as age, height, weight, race, sex; anthropometric data; physician's examination data, laboratory data, and health questionnaire data [39]. Hand and knee X-ray films were collected by the NHANES III survey, but NCHS has not released this data for public use. 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 for a multiple-reader study to determine the best digitization 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 digitization levels (50, 100, and 150 μm, respectively). The results, published in the Journal of Digital Imaging [40], were used by NCHS in the subsequent digitization of these

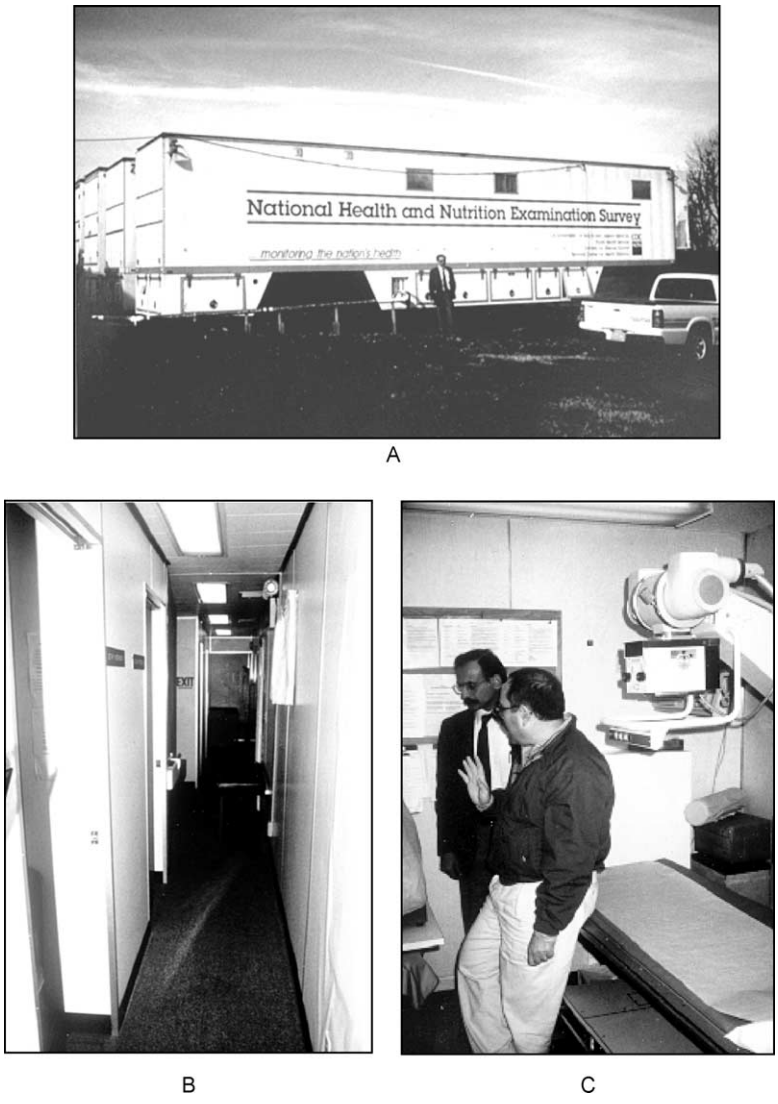


Fig. 1. An NCHS Mobile Examination Center. (A) Three trailers joined with walkways to make a single MEC; (B) view down the hallway of one trailer, showing some of the examination stations; (C) view inside the X-ray station.

images. Images from this digitization have been used in research into the automated assessment of hand/wrist radiographs for arthritis, by the use of neural networks [41].

3.1.3. Biomedical review

Two workshops [42] 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 subluxation, 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. The candidate and final features from the NIH workshops are given in Table 1.

3.1.4. 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 viewing the images at 1K×1K spatial resolution on PC monitors and verifying 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 done by a medical expert in radiology, who reviewed images and films

Table 1
Candidate biomedical features for the NHANES II X-rays

Anterior osteophytes[a]
Disc space narrowing[a]
Spondylolisthesis (L-spine)[a]
Subluxation (C-spine)[a]
Posterior osteophytes
Plate erosion/sclerosis
Vacuum phenomenon
Abnormalities (if any noted)
Ankylosing spondylititis
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)

[a] Indicates the final feature set of interest from the NIH workshops.

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 2051 image/film pairs for overall visual quality and scored 1625 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 [43]: anterior osteophytes: 94.2%; subluxation (or spondylolisthesis in lumbar spine): 94.1%; and disc space narrowing: 93.7%.

3.2. Prototype systems, tools, and data repositories for NHANES

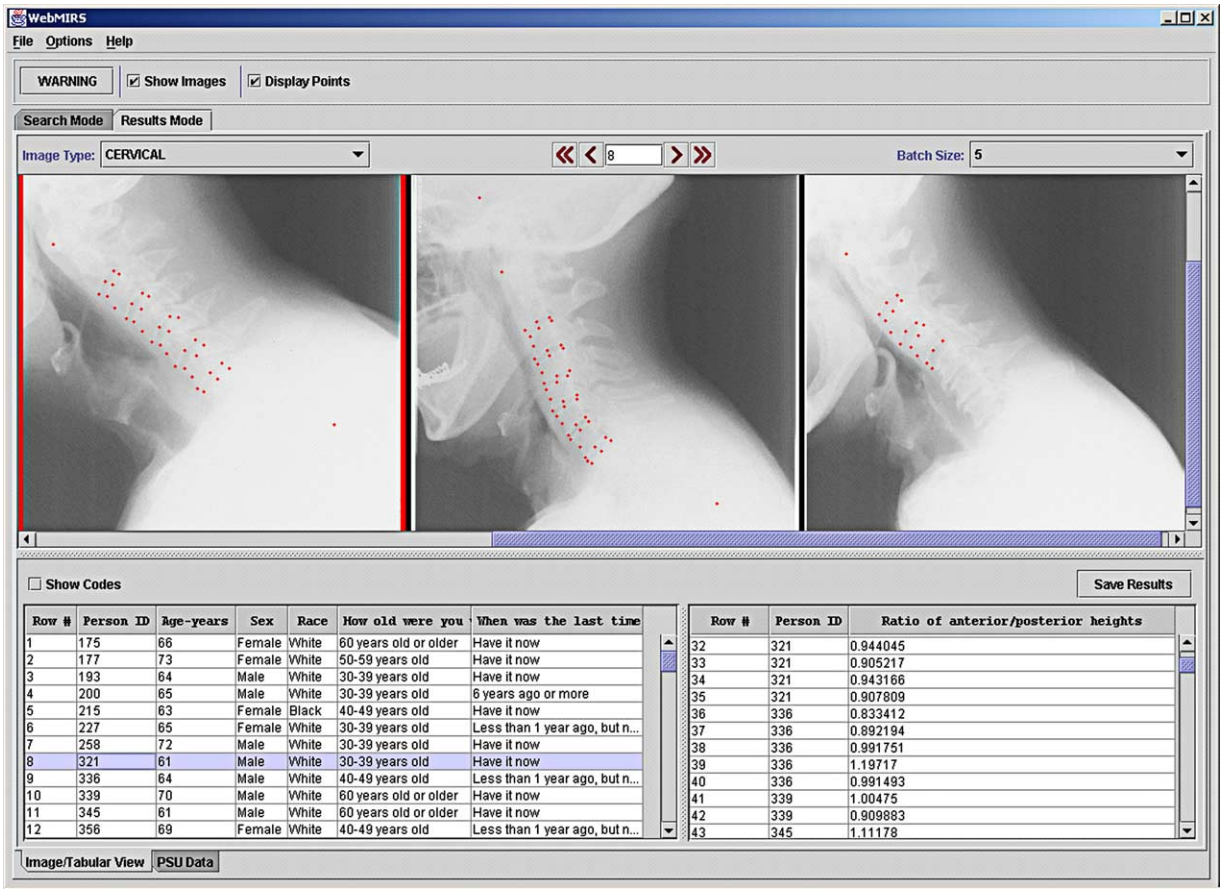
3.2.1. WebMIRS

Our initial work to create a multimedia database system was a Sun platform dependent application called the Medical Information Retrieval System [44] (MIRS). The first Web-based MIRS (WebMIRS) [45] 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 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 2 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 Fig. 2). 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 five hundred and fifty of the images; this data consists of nine-point boundary landmarks placed on the vertebrae in these images under supervision of a board-certified radiologist (see Fig. 3). This quantitative vertebral data may also be used for WebMIRS query and retrieval. Fig. 2 in fact gives an example of a WebMIRS query, 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. These nine-point radiologist landmarks ('morphometric marks') have become a basic reference set for first-order evaluation of segmentation algorithms for us and our collaborating researchers. The present WebMIRS system is version 1.0.10 and is deployed as a Java application using Sun's Java Web Start 1.4.2 technology. Most WebMIRS users operate on the PC platform.

3.2.2. Digital atlas of the cervical and lumbar spine

A digital atlas of the cervical and lumbar spine [46] was developed, using a subset of these images which were interpreted to consensus by a panel of three medical experts convened by the NIAMS. 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 to fill a perceived need for such reference data for osteoarthritis in the cervical and lumbar spine, especially since a standard reference [47] of photographs of these features is out of print and difficult to obtain. Important features of the Atlas



Row #	Person ID	Age-years	Sex	Race	How old were you	When was the last time
1	175	66	Female	White	60 years old or older	Have it now
2	177	73	Female	White	50-59 years old	Have it now
3	193	64	Male	White	30-39 years old	Have it now
4	200	65	Male	White	30-39 years old	6 years ago or more
5	215	63	Female	Black	40-49 years old	Have it now
6	227	65	Female	White	30-39 years old	Less than 1 year ago, but n...
7	258	72	Male	White	30-39 years old	Have it now
8	321	61	Male	White	30-39 years old	Have it now
9	336	64	Male	White	40-49 years old	Less than 1 year ago, but n...
10	339	70	Male	White	60 years old or older	Have it now
11	345	61	Male	White	60 years old or older	Have it now
12	356	69	Female	White	40-49 years old	Less than 1 year ago, but n...

Row #	Person ID	Ratio of anterior/posterior heights
32	321	0.944045
33	321	0.905217
34	321	0.943166
35	321	0.907809
36	336	0.833412
37	336	0.892194
38	336	0.991751
39	336	1.19717
40	336	0.991493
41	339	1.00475
42	339	0.909883
43	345	1.11178

Fig. 2. WebMIRS query results including display of nine-point morphometric points; ratio of anterior/posterior vertebra heights is part of the returned data in this query.

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 Fig. 4. 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.

3.2.3. FTP archive

All 17,000 NHANES II X-ray images have been made publicly available through an FTP archive that is publicized on the NLM Communications Engineering Branch 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 PhD theses, including two published recently [48,49], 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 nine-point radiologist marks.

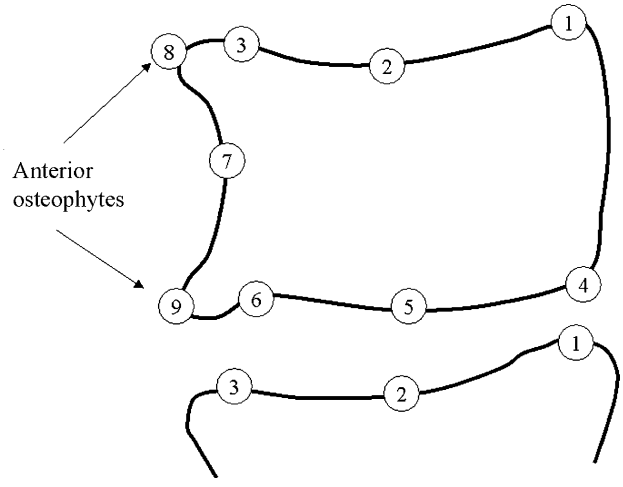


Fig. 3. Cervical spine vertebrae illustrating nine-point morphometry.

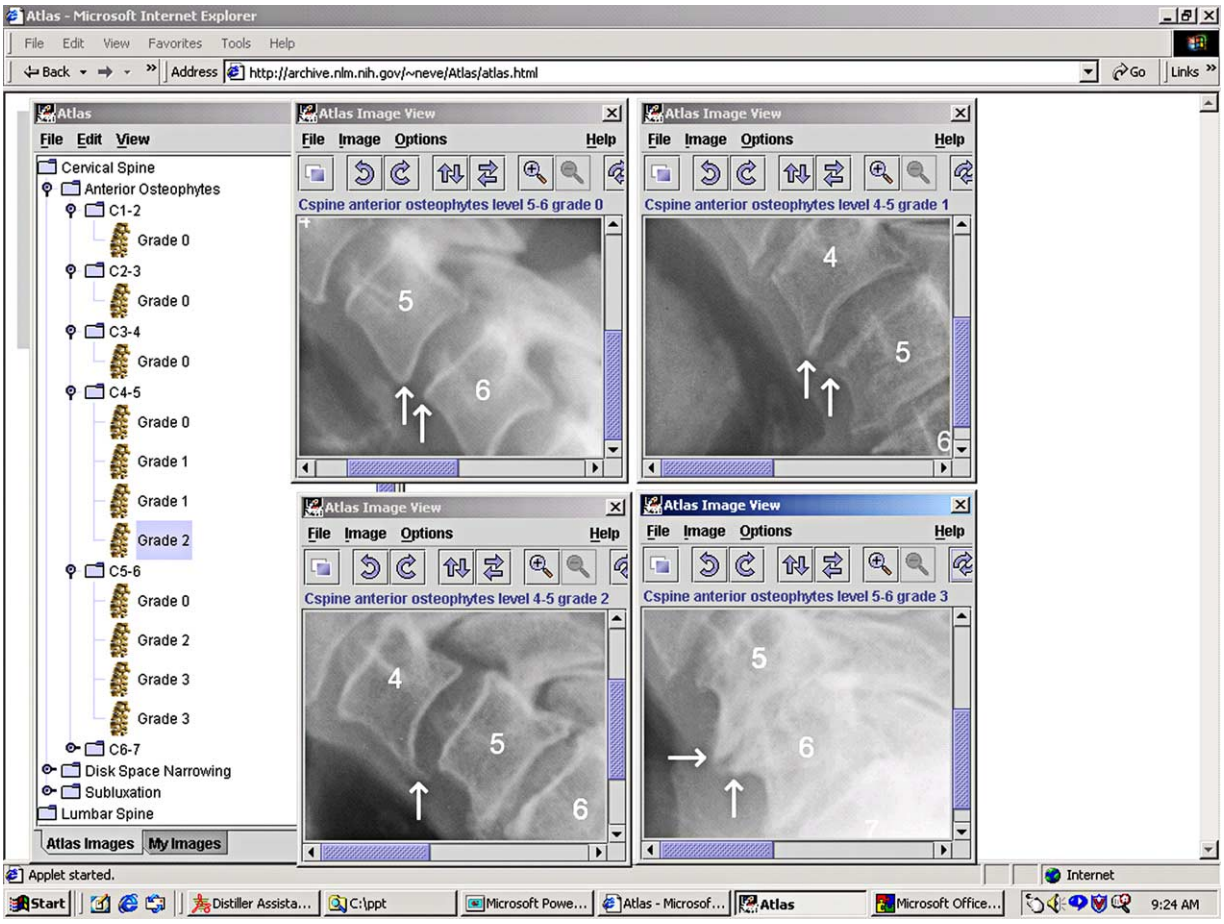


Fig. 4. Digital atlas of the spine.

3.3. Image processing R&D with NHANES

3.3.1. 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) [50,51], and incorporates both vector and scalar quantization.

3.3.2. Segmentation

A goal of our segmentation work has been to develop a suite of segmentation tools representing leading segmentation techniques for research and comparative evaluation. This work has included active contour segmentation (ACS), active shape modeling (ASM), and a 'Hierarchical Segmentation' approach that combines multiple approaches, including the Generalized Hough Transform (GHT) and active appearance modeling (AAM). Currently Live Wire segmentation is being implemented as well. Each method investigated represents 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 approaches, including deformable template methods, with some of those based on statistical models. For each segmentation method investigated, we summarize the concept, main features, level of interactivity, results obtained, and reference the responsible collaborating researchers. All of the segmentation implementations described were implemented in MATLAB.

3.3.2.1. Active contour segmentation (ACS).

Concept: our ACS algorithm is a classical snake model combined with an initial contour created from a priori information and a search constraint on the contour. The algorithm minimizes an objective function by seeking a contour with maximized gradients along normals to the contour, and minimized contour length ('maximum edge strength and maximum smoothness'). The objective function has heuristically-determined weights for these two factors.

Main features: segmentation is one vertebra at a time. The initial contour is a template created by averaging manually-segmented vertebral shapes. The algorithm

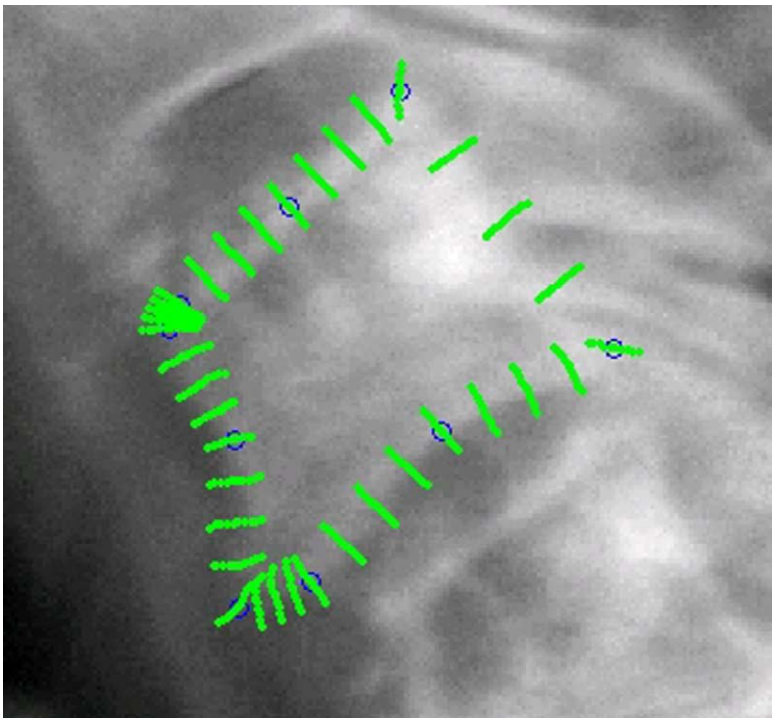


Fig. 5. Orthogonal curves replace normal line segments in defining the search grid used in Active Contour Segmentation. Points on the solution contour are constrained to lie on these non-intersecting curves.

constrains solution contours to lie on a grid between an 'inner contour', inside the template, and an 'outer contour', outside the template. A novel feature is that the grid line are 'orthogonal curves', calculated by numerically solving a boundary-value partial differential equation (see Fig. 5). This approach is given by Tagare in [52]. The original implementation used a grid created by simple normal line segments; however, in cases where the vertebra has a narrow protrusion, these normals can be self-intersecting, resulting in bad segmentations in many cases. The implementation of the core ACS algorithm is due to Tagare and the orthogonal curves implementation is due to X. Qian (see Fig. 6 for the ACS user interface).

Interactivity: The user manually positions, scales, and orients the initial template.

Performance: No performance evaluation has been done on ACS at this time.

3.3.2.2. Active shape modeling (ASM).

Concept: The ASM formulation that we have followed is that described by Cootes and Taylor [53] of the University of Manchester.

Main features: 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.

Interactivity: The user manually positions, scales, and orients the initial template.

Performance: Our initial evaluation of ASM for NHANES II cspine images was done by Sari-Sarraf [54, 55] on a test set of 40 images. For each image a template with 80 points was manually created, using the nine-point radiologist marks described earlier as guides; the template modelled the vertebrae shapes from the bottom of C2 through C6. Using the leave-one-out approach, 40 ASM shape and grayscale models were created, by sequentially omitting one image (the 'test image'), and including the remaining 39; for a given test image, we refer to these 39 images as the 'model images' for that test image. The ASM model created by leaving out one image was then applied to test the performance of ASM on that image. In this fashion, ASM was evaluated on an image which it has not previously 'seen'. The algorithm was evaluated by taking the manually-created

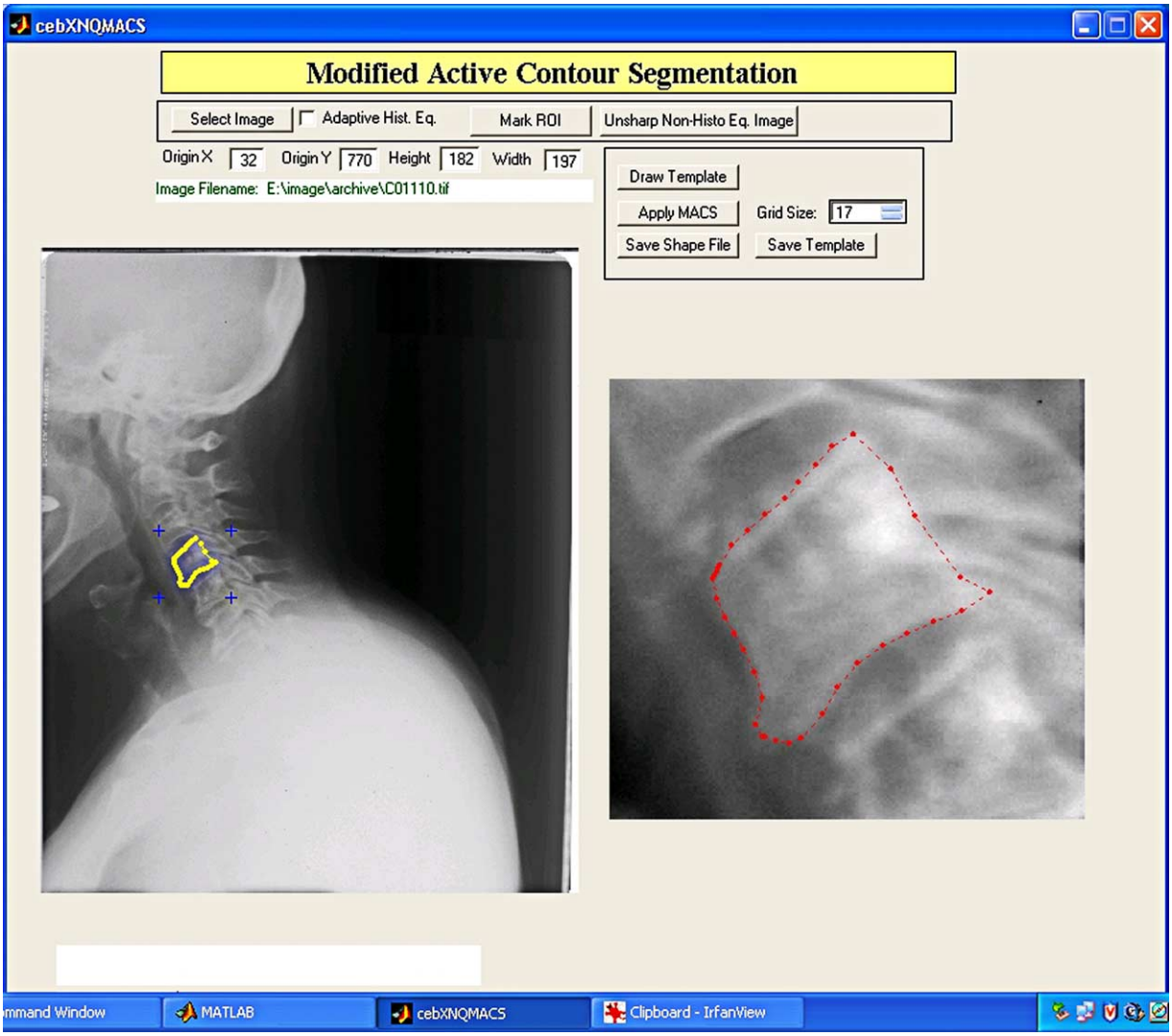


Fig. 6. Active Contour Segmentation user interface. Solution curve is shown on right panel in zoomed view.

template for an image as 'truth' for that image, and computing the mean (averaged over the 80 points in the template) point-to-point pixel error (MPPE) between truth and the ASM solution. The template used to initialize ASM for a test image was the ASM shape model created from the 39 model images for that test image. The template was manually placed by positioning the C4 template shape on top of the C4 vertebra in the test image. Results were as follows, in units of pixels: over the 40 test cases, the minimum MPPE was 12.8; average MPPE was 114.7; and maximum MPPE was 282.9. More extensive ASM evaluations are described in the section below on hierarchical segmentation.

3.3.2.3. Hierarchical segmentation. A significant aspect of our work has been the investigation of hybrid segmentation methods carried out in sequential coarse-to-fine steps. Two such hierarchical segmentation approaches have been developed. Both use the generalized Hough transform (GHT) for the first step: finding the approximate location of the target vertebrae; both have a core segmentation method for simultaneously segmenting multiple vertebrae; and both have a final step for refining the vertebral boundaries. The first method, developed by Zamora, uses Active Shape Models as its core method; the other, developed by Howe, relies on Active Appearance Models. We precede the discussion of the Zamora and Howe algorithms with a discussion of the GHT, which is common to the two approaches.

3.3.2.4. Generalized Hough transform (GHT).

Concept: The basic GHT formulation for detection of arbitrary shapes in images was given by Ballard [56], and GHT for the NHANES II images has been extensively investigated by Sari-Sarraf, Tezmol, and Gururajan [57,58]. The GHT is a template matching method. Matching is done up to position, orientation and scale, but the shape of the template does not deform.

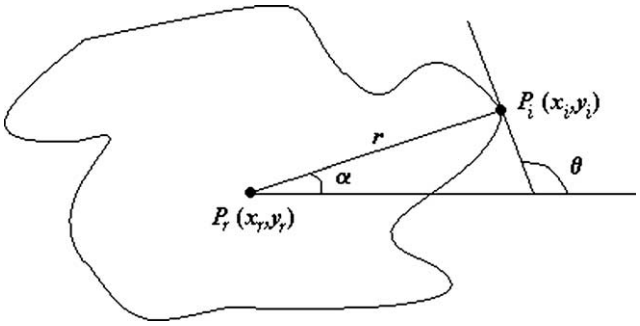


Fig. 7. Geometry for building *R*-table for generalized hough transform (GHT).

Main features: Geometry for the template to be matched in given in Fig. 7 [57].

In this figure $P_r(x_r, y_r)$ is a reference point which is the origin of an axis system fixed in the template, and $P_i(x_i, y_i)$ is an arbitrary point on the template boundary, which is specified by the (distance, angle) pair (r, α). The 'edge direction angle' θ is determined by the intersection of the tangent line through P_i and the horizontal axis. This geometry allows the construction of a discretized '*R*-table' that defines the template in terms of edge direction angles θ_i and the corresponding (r_i, α_i) pairs, as illustrated in Fig. 8.

If the template is rotated through an angle φ, relative to an image-fixed coordinate system, and scaled by a quantity s, then these relationships hold

$$x_i = x_r + sr_i \cos(\alpha_i + \varphi)$$

$$y_i = y_r + sr_i \sin(\alpha_i + \varphi)$$

between a boundary point and the reference point. We observe that in a given image, the gradient direction is computed as $\arctan(G_y/G_x)$, where (G_x, G_y) is the gradient vector at (x_i, y_i), and the edge direction differs by $\pi/2$. We may then apply the GHT algorithm: fix a scale factor s and orientation angle φ; then, at each candidate boundary point in the image, compute the gradient vector and the corresponding edge direction angle θ_i; use the R-table to find a corresponding (r_i, α_i) pair. Finally, use the above equations to solve for the corresponding (x_r, y_r). This procedure uniquely identifies a bin (x_r, y_r, s, φ) in 4-D Hough space corresponding to the candidate boundary point (x_i, y_i). A counter for the bin is incremented. For a fixed

θ_i	(r,α)
$\Delta\theta$	(r_1,α_1), (r_2,α_2)
$2\Delta\theta$	(r_3,α_3)
$3\Delta\theta$	(r_4,α_4), (r_5,α_5), (r_6,α_6)
…	…

Fig. 8. *R*-table for GHT.

(scale, orientation) pair (s, φ), this procedure is repeated for all candidate boundary points in the image. Then the procedure is repeated over the discretized range of possible scales and orientations. At the end of all processing, the bin with the greatest count identifies the position, scale, and orientation of the template in the image. The candidate boundary points in the image for which we compute Hough bins are edge points that we have detected in the image; hence the detection of edges in the images is crucial to the performance of the algorithm. Tezmol applied unsharp masking, filtering for noise removal, binarization, and finally Sobel edge detection to produce edge images, as shown in Fig. 9; Gururajan developed an alternative technique based on first producing a phase congruency map of the image, followed by filtering, binarization and application of the Sobel operators. It is this latter approach that was used to produce the results described below.

Interactivity: None, but see the discussion below about incorporating an interactive feature to disambiguate cases where the template is misplaced along the vertebral column.

Performance: In the most recent work [58], the GHT was tested on 273 cspine images and 262 lspine images. 'Truth' segmentations were manually done for 100 images of each type. For cspine, vertebrae segmented were bottom of C2 through C5; for lspine, the vertebrae were L1–L5. Each segmentation contained 100 points. For the cspine images a template was created by finding the mean of the 100 cspine segmentations; similarly, for the lspine images. For cspine, the scale range used was [0.8, 1.25] at increments of 0.05; orientation range in degrees was $[-25, 25]$ with a 5° increment. For lspine, the scale range used was [0.85, 1.05] at increments of 0.05; orientation range in degrees was $[-5, 5]$ with a 5° increment. To drive a measure of performance, a bounding box which closely fit the cspine template was created, and similarly, for the lspine template. When the template had been placed on an image by the GHT algorithm, the number of truth points lying within the template's bounding box were counted. For the images other than the 100 segmented of each type, the coarser radiologist points (see Section 3.2.1) were available and were used as truth. For the cspine images, the mean percentage and standard deviation percentage of points within the bounding box were (80, 23); median percentage of points within the bounding box was 82. Corresponding numbers for the lspine were mean/standard deviation percentages of (75, 27) and median percentage of 83. These numbers alone give little insight into the usefulness of the GHT for first approximation segmentation of the images. A visual inspection was done on a total of 473 cspine images processed by the GHT (including the above 273); in a number of cases where the GHT placed the template inaccurately, the misplacement was approximately correct, up to 'sliding along the vertebral column'. That is, instead of begin placed correctly on C2–C5, the template may have been placed on C3–C6, for example. Many of these misplacements were determined to correspond to

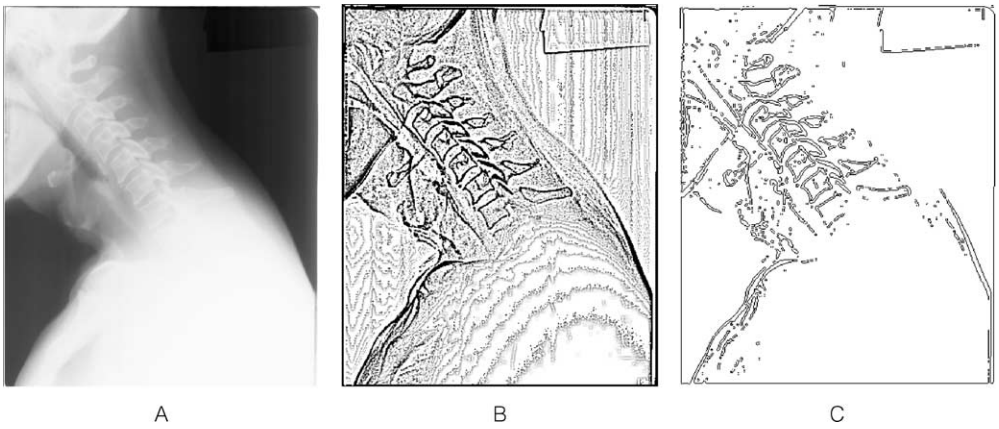


Fig. 9. Generalized Hough Transform pre-processing. (A) Original cervical spine image; (B) image after unsharp masking to enhance edges; (C) image after edge detection.

failure of the algorithm to properly distinguish the correct template bin in Hough space from among the top three bins. Hence, an interactive capability was proposed (and later incorporated) wherein the GHT would return the template placements corresponding to the top three bins, and allow the user to choose among these best three candidates. When the user was allowed to consider these top three candidates, more than 400 of the GHT results were judged acceptable of the 473 visually inspected. Similarly, for the lspine images, 472 were processed by GHT (including the above 262) and visually inspected,and more than 400 were judged acceptable, again, if the top three candidates are considered. For the GHT, computation time is an issue; in the above work, the images were reduced to one quarter of their original size horizontally and vertically (original size for cspine was 1462×1755; for lspine, 2048×2488). On a 2.1 GHz AMD microprocessor, processing time for a cspine image was 55 s; for an lspine image, 110 s. Fig. 10 shows an example cspine template used, a plot of the bins in Hough space, and a resulting template placement, with bounding box.

3.3.2.5. Zamora hierarchical segmentation.

Concept: This hybrid segmentation method, developed by Zamora and Sari-Sarraf, is described in detail in [49]. The ingredients are (1) the GHT for initial placement of a rigid template on the vertebrae; (2) ASM for deforming the template non-rigidly, beginning at the position, orientation, and scale output by the GHT; and (3) a final 'deformable model' (DM) step to adjust the ASM output at the vertebral corners, where the ASM segmentations tend to miss irregular bone deformities and protrusions, which have proved elusive to model by statistical shape sampling, as is required by ASM. (Finding enough examples to represent the corner variability has not been successful to date.)

Main features: Zamora made two modifications to the use of the base GHT described above. The first was to allow interactive access to the top three bins returned by the GHT when the template created as the mean vertebral shape from a training set of manually segmented vertebrae was used. The second modification was to input not only

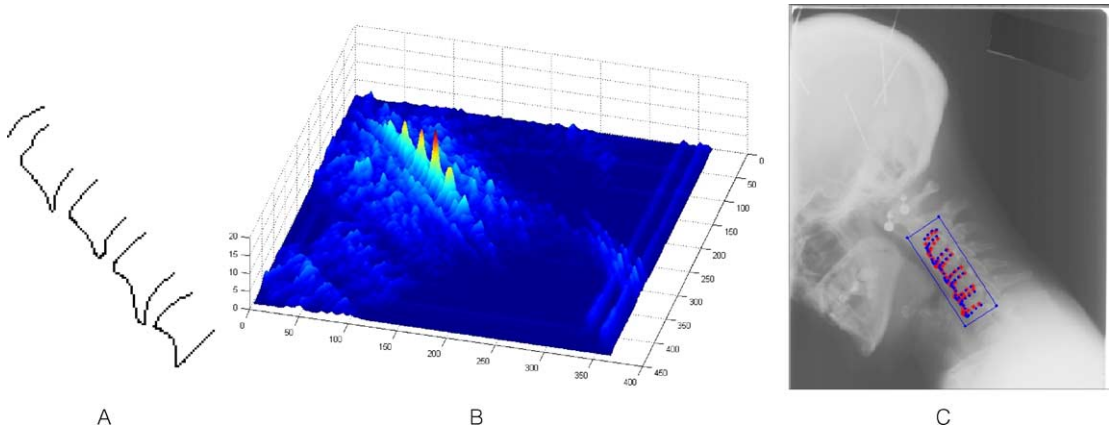


Fig. 10. Generalized Hough Transform processing and result. (A) Example cervical spine shape template; (B) Hough bin values plotted in x–y space. Axes are x and y position coordinates; the plot is for one fixed value of orientation and one fixed value of scale; highest peak is GHT estimate for positioning template in the image; (C) example GHTfinal result, shown by overlaying GHT-positioned/oriented/scaled template onto original image. A computed bounding box for the result is also shown.

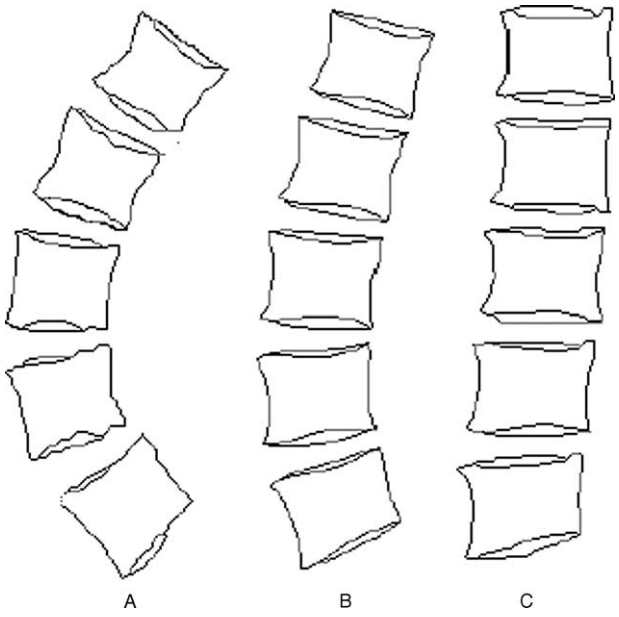


Fig. 11. Lumbar spine shape templates computed by principal components analysis (PCA), used for Generalized Hough Transform step in Zamora Hierarchical Segmentation. These templates represent the mean and extremes of curvature found in the sample set of images. (A) maximum curvature template; (B) mean shape template; (C) minimum curvature template.

the mean shape as a search template, but also two additional shapes that represent extremes of bending as determined by a principal components analysis of the training data (i.e. the image data used to create the vertebral templates) (see Fig. 11).

Hence, for each GHT run, five solutions were output for the operator to choose among (the three top bin solutions when the mean shape template was used; and the top bin solution when each of the two extreme-shape templates were used). For the ASM algorithm, Zamora applied a pre-processing step to compensate for a phenomenon he had previously identified in [59]: for a number of the images, the grayscale values measured along normals to the vertebrae boundaries carry insufficient information for the ASM algorithm to localize the boundary. (ASM depends on such grayscale information to deform the template points to the vertebral edge locations in the image.) Zamora's pre-processing created multiple edge images by unsharp masking, averaged these edge images, and then blended the resulting mean edge image with the original grayscale image. The final deformable model (DM) step takes the two anterior corners of the segmentation produced by ASM as input, with each corner consisting of five points from the segmentation that are centered around a point tagged as a 'vertebra corner point' in the original template input to the GHT. The DM is formulated as a standard 'snake' that deforms to minimize the sum of external and internal energy terms; hence this final step allows the corners to deform outside of the shape constraints of the ASM model and, except for the smoothness constraint imposed by the internal energy term, is constrained only by the image data itself, not by a shape a priori, as in ASM.

Interactivity: The operator chooses among five alternative template placements by the GHT. Performance: The algorithm was tested on 100 cspine images and 100 lspine images. For the cspine images templates were created by manual segmentation that contained 80 points and spanned the bottom of C2 through C6; for the lspine images, templates were likewise created that contained 200 points and spanned L1-L5. These templates were used as the ground truth for evaluating performance, and the average pixel distance between the manually-created template for an image and the shape output by the algorithm was used as a performance measure. Performance was computed at each of the segmentation stages, so that the improvement due to each stage could be assessed. Results were as follows, for the cspine images, percentages of test images with errors of 20 pixels or less: after GHT alone, 65%; after GHT + ASM, 75%; and after GHT + ASM + DM, 75% (negligible improvement with DM). For the lspine images, for errors of 50 pixels or less: after GHT, 40%; GHT + ASM, 47%; and GHT + ASM + DM, 49%.

3.3.2.6. Howe hierarchical segmentation.

Concept: This hybrid segmentation method, developed by Howe and Sari-Sarraf, is described in detail in Ref. [60]. Like the Zamora algorithm, it is a three-step process: (1) the GHT for initial rigid template placement; (2) active appearance models (AAM) for deforming the template non-rigidly, beginning from the template output by GHT; and (3) a final application of AAM to each individual vertebra to allow the algorithm to adjust the segmentation for each vertebrae without being influenced by image data associated with the other vertebrae.

Main features: Like Zamora, Howe allowed interactive access to the top three bins returned by the GHT when the template created as the mean vertebral shape from a training set of manually segmented vertebrae was used. Unlike Zamora, no additional templates from principal components analysis were used. In AAM, the shape and texture ('appearance') of the entire object to be segmented is used. The AAM approach implemented generally follows that given by Cootes [53], although the AAM models that were created included pixels in the exterior neighborhoods of the vertebrae, as well as the vertebrae interiors; in addition, pixels near the borders of the vertebrae were weighted to be more significant in the model fitting, than pixels relatively far from the borders. Fig. 12 gives an example of the lumbar spine model.

Interactivity: The operator chooses among three alternative template placements by the GHT. Performance: Cspine and lspine AAM models were built, following the procedure described below, again using 100 cspine and 100 lspine images, respectively, that were manually

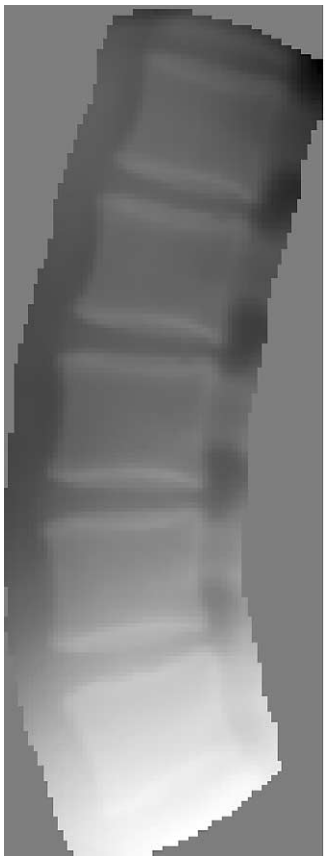


Fig. 12. Example active appearance model for lumbar spine. This model was computed as the mean shape and texture of the lumbar spine in 100 sample images. Shapes were manually segmented. In this work, texture was taken as graylevel pixel values.

segmented as for the Zamora algorithm testing. For the GHT step, the cspine template was the mean shape computed from these images, and similarly, for the lspine template. Testing was done in leave-one-out manner for each of the 100 cspine images, where for each image in turn (the test image), an AAM model was built using the other 99 images; then this model was applied to the test image. The same process was applied to the lspine images. Results were as follow. For cspine images, percentages of images having errors of 10 pixels or less: after GHT alone, 10%; after GHT + multi-vertebrae AAM, 60%; and after GHT + multi-vertebrae AAM + single-vertebrae AAM, 65%. For the lspine images, for errors of 25 pixels or less, the following success rates were achieved: after GHT alone, 21%; GHT + multi-vertebrae AAM, 67%; and GHT + multi-vertebrae AAM + single-vertebra AAM, 68%.

3.3.3. Automated classification of biomedical features

Concept: Automated classification of NHANES II images for biomedical features by shape has been investigated by Stanley [61,62], using artificial neural networks and clustering techniques.

Main features: Cherukuri and Stanley [61] have investigated the application of artificial neural networks to the discrimination of lumbar spine vertebrae for the presence of anterior osteophytes. Four geometric features were derived and tested individually. The most successful was obtained by first finding the exclusive-OR of a vertebral area and the area within the convex hull of the vertebra; and, for this region, finding the area of the largest connected component on the anterior side of the vertebra. Charmathy and Stanley [62] have applied *K*-means and self-organizing map clustering techniques to the task of scoring adjacent vertebrae for disc space narrowing on a 0–3 scale, representing normality (0) to maximum abnormality (3). Features were derived by first computing an separator curve equidistant between adjacent vertebrae and then measuring various quantities relative to this separator, such as minimum Euclidean distance from the separator to a vertebra, and normalizing the quantities to obtain features that are intended to be invariant with respect to vertebral size.

Interactivity: None. Performance: For the lumbar spine anterior osteophyte discrimination, 572 vertebrae were used for training, with half known to be normal and half abnormal; 108 different vertebrae were used for testing, half normal and half abnormal. Twenty individual test and training sets where generated by randomly selecting 572/108 vertebrae from among the available vertebral data. The neural network used had a $4\times4\times1$ architecture (four input nodes for the four features; one hidden layer; one output node), with sigmoid function at input and hidden layer nodes, and a linear transfer function at the output. The vertebrae were manually segmented, using the nine-point radiologist marks as a guide. Results were as follows: the mean correct classification of normal vertebrae in the test set was 88.6%; the mean correct classification of abnormal vertebrae was 90.5%. For the work in classifying disc space narrowing by grade, 294 adjacent vertebral pairs ('interfaces') of the cervical spine were used. Twenty runs were made; for each run 80% of the data was randomly assigned to the training set, and the remaining 20% to the test set. The mean percentages of correct classifications over the 20 runs were, for grades 0–3, respectively, 90.4, 85.2, 93.8, and 82.1%.

3.3.4. Content-based image retrieval

Content-based image retrieval (CBIR) is the focus and principal goal of our work in biomedical imaging. CBIR methods could provide access to specific visual biomedical features in large image archives. For biomedical image databases the notions of image content, content representation, and content similarity have high relevance. For example, in the biomedical context, it is to be expected that many important queries may be intended to retrieve pathological cases which have only subtle differences from the normal. Hence, it is important that, at the time of indexing of biomedical images for later retrieval by image

contents, the extracted image features are carefully preserved by the methods selected to represent them (which usually are oriented toward data reduction and hence risk sacrificing details); and it is equally important that, at retrieval time, the similarity measures used are sufficiently fine-grained, so that subtle similarities between a query image and a database image are recognized.

We have implemented several modular prototype CBIR systems for retrieval for a subset of the NHANES II spine X-rays. We initially used the nine-point model for vertebral shape representation [63], and Procrustes distance [64] to measure shape similarity. This system evolved to accept more complex vertebral segmentations and alternative shape representations, viz. Fourier descriptors, Polygon Approximation, Curvature Scale Space, invariant moments and other shape features [65]. The system supported retrieval based on shape similarity to a sketch or example vertebral image, as well as by using text as part of the retrieval process. (NHANES II health survey data was included in the database, as well as vertebral shapes.) The shapes are segmented using the active contour segmentation (ACS) method described above, with operator interactivity to adjust the ACS output when required. The system is composed of two subsystems, one for indexing (segmentation, principally), and one for retrieval. The indexing system included the ACS segmentation algorithm and feature extraction, using one of the alternative shape representations. The retrieval system provided the interface and the methods for image and text retrieval. For this it included a query-by-sketch and a query-by-image example capability and methods for computing similarity between the user query and shapes stored in the database. In addition, text retrieval via SQL and methods to combine the text and image queries were also included.

In order to determine the quality of retrieval, several evaluation studies were conducted [66]. A summary of the results from these experiments are presented in Table 2. The table also gives the average displacement in the results ranking from the position expected in the truth ranking.

3.4. Partial shape matching

Whole shape matching refers to the matching of shapes with respect to their global shape characteristics. In query-by-sketch and query-by-example paradigms of shape-based CBIR, whole shape matching methods may be ineffective or difficult to use for the retrieval of shapes that have localized shape features of interest, such as protrusions at one vertebral corner. To address this drawback, we have investigated partial shape matching (PSM) methods that allow the user to sketch or identify only the local region of interest on the vertebral boundary. Figs. 13 and 14 illustrate the user interface for whole shape matching query results, and for partial shape matching query creation, respectively.

Our PSM method uses shapes with a fixed number of boundary points, and Procrustes distance for measuring similarity [67]. A performance evaluation has been conducted. For this a data set was generated from a total of 206 spine NHANES II X-ray images (106 cervical and 100 lumbar images) from subjects who were 60 years of age or older. The age criterion was used because of higher prevalence of degenerative joint disease and hence, expected vertebral shape deformity, in this population. Vertebrae from the selected 206 images were segmented using the active contour segmentation (ACS) algorithm previously described. The process yielded 896 segmented vertebrae, composed of 407 cervical (C3–C7) and 489 lumbar (L1–L5) vertebrae. Each vertebra was represented by 36 points.

Two classification schemes for anterior osteophytes were established by a medical expert to evaluate the accuracy of the PSM algorithm. One is Macnab's classification, established by Macnab and his coworkers in 1956 on radiological and pathological bases [68,69]. The second classification is a severity grading system which was defined by a collaborating medical expert, based on judgment of reasonable criteria for assigning severity levels to the Macnab classes. There are two Macnab classes of osteophytes: claw and traction. Claw osteophyte ('spur') arises from the vertebral rim and curves toward the adjacent disk. It is often triangular in shape and curved at the tips. Traction spur protrudes horizontally, is moderately thick, does not curve at the tips, and never extends across the intervertebral disk space. Vertebrae that do not exhibit anterior osteophytes (AO) that fall under these categories are labeled Normal. There are three grades of severity in system: slight, moderate, and severe. A grade of Slight is

Table 2
Results of preliminary vertebral shape retrievals, using query-by-sketch, and whole shape matching

Experiment	Evaluated against	Summarized results
Fourier descriptor on shape data reduced using polygon approx.	Similarity list using Procrustes distance on nine-point shape data	On the average 54.7% of retrieved shapes matched the top 25 ranked shapes in similarity list. Average displacement = 23.4
Fourier descriptor on automatic nine-point data from modified polygon approx.	Similarity list using Procrustes distance on nine-point shape data	On the average 70% of retrieved shapes matched those in the similarity list. Average displacement = 9.0
Fourier descriptor on shape data reduced using polygon approx.	Similarity list by medical professional on different grades of pathology	Average relevance = 5.7/10. Average accuracy = 3.8/10 using staggered scoring (Scores averaged over 48 cervical and lumbar queries).

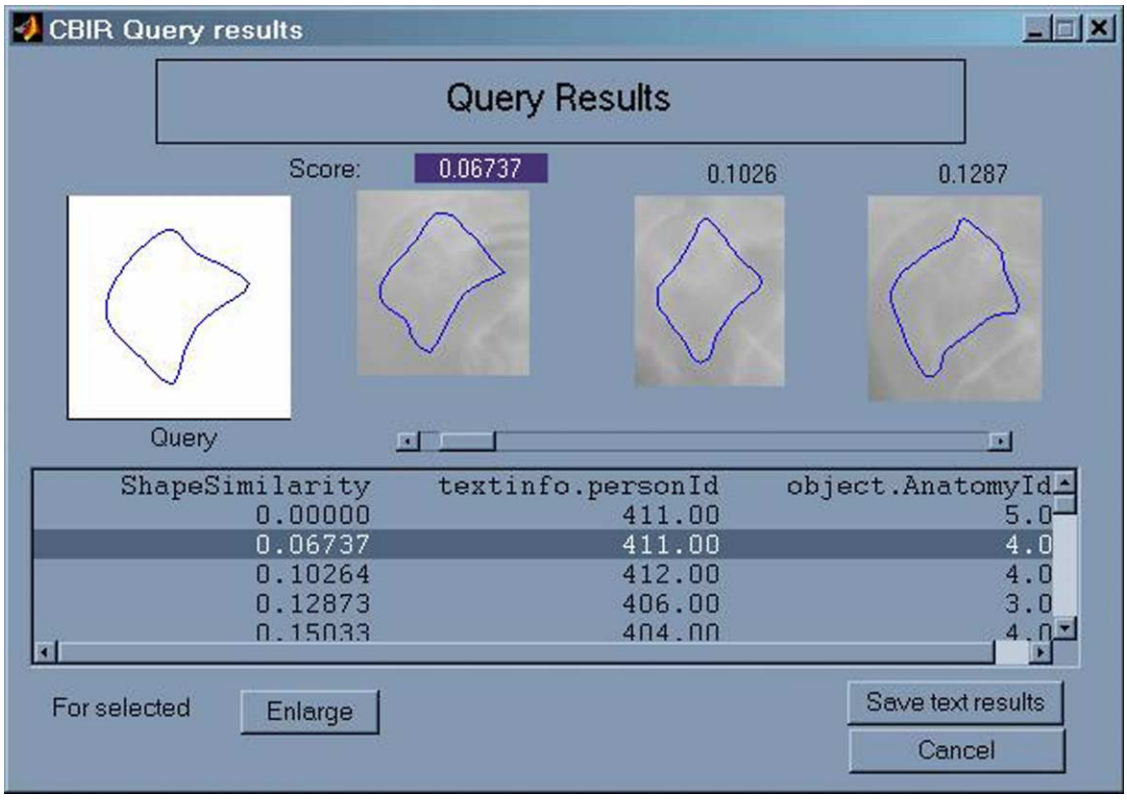


Fig. 13. Prototype CBIR system, query results, whole shape matching.

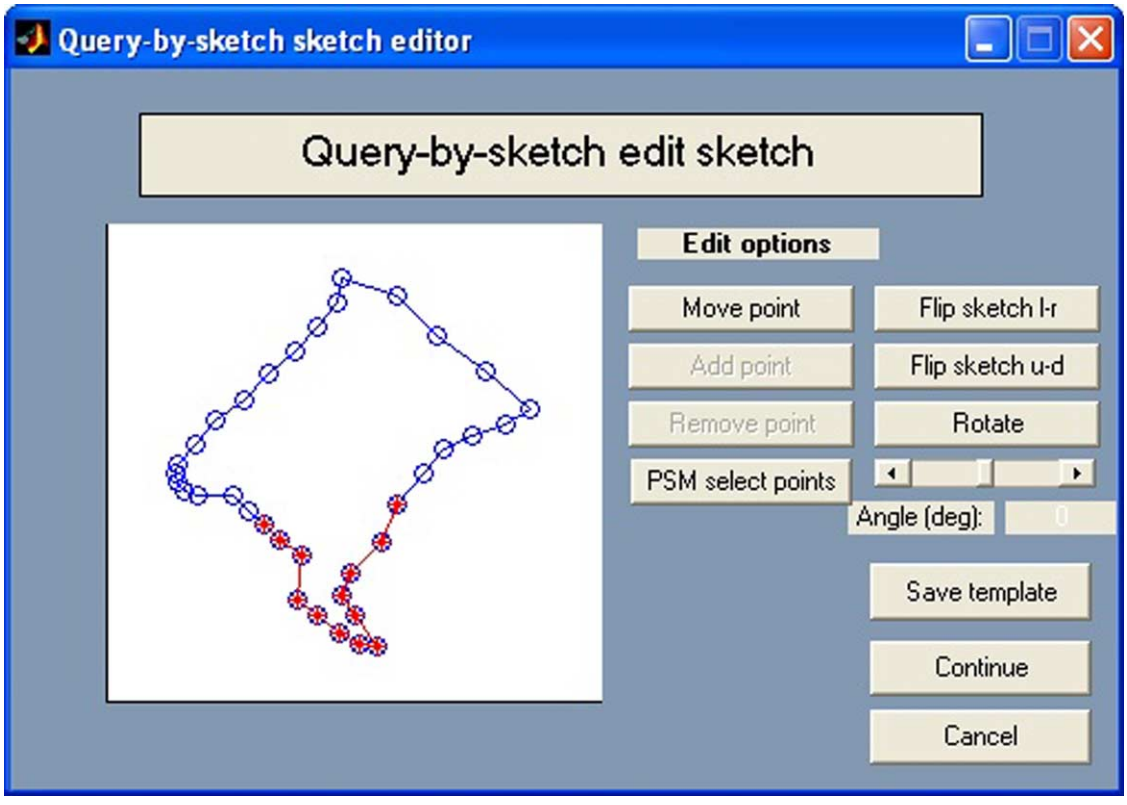


Fig. 14. Example of query-by-sketch and identification of region of interest for partial shape matching.

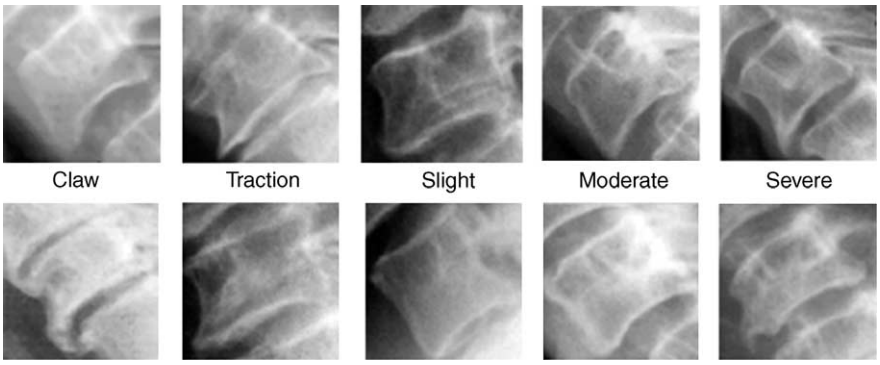


Fig. 15. Claw and traction spurs, and examples of severity of each.

assigned when the vertebral body is approximately square. It may have a slight protuberance, where the tip of osteophyte is round and no narrowing is observed at the base of the protuberance. Moderate grade is characterized by evident protuberance from the ideal horizontal or vertical edge of the vertebra. The bounding edges of the AO form an angle of at least 45° and the osteophyte has a relatively wider base than severe grade. Severe grade is characterized by presence of hook, the angle is less than 45° and has a narrow base, or protrudes far (about 1/3 of the length of the horizontal border) from the normal (ideal 90°) vertebral corner, as shown in Fig. 15 below.

'Ground truth' for the test was created by having the medical expert label each vertebrae for anterior osteophytes, using the classifications: normal, claw spur, or traction spur, and, if either claw or traction spur was present, by labeling the spurs as Slight, Moderate, or Severe.

The medical expert selected a total of 28 vertebral shapes from the data set as queries using combinations of the classifications and locations of the osteophytes. For example, queries were created for Superior and Inferior Claw type with Slight, Moderate and Severe pathologies; and similarly for Superior and Inferior Traction type. In addition, queries were developed for Superior and Inferior Normal vertebrae. The evaluation studied the top 10 vertebral shapes returned by the PSM algorithm and compared their pathology classification to the query shape classification. Initial results show a near 100% retrieval precision for normal vertebra and 65% retrieval precision for claw and traction types averaged over all grades of severity. Sample results from a PSM query are shown in Fig. 16.

4. Biomedical imaging for uterine cervix images

4.1. The Guanacaste database

The Guanacaste Project [70] is an intensive, population-based cohort study of human papillomavirus (HPV) infection and cervical neoplasia among 10,000 women in Guanacaste, Costa Rica, where the rates of cervical cancer are perennially high. State-of-the-art visual, microscopic, and molecular screening tests are being used to examine the origins of cervical precancer/cancer and to explore which factors make a geographic region 'high risk'. The Guanacaste study has completed its field phase after 7 years of follow-up, and now has changed into a variety of subprojects based on collected specimens, visual images, and outcomes. The National Cancer Insititute (NCI) is examining several potentially important etiologic co-factors, such as chronic inflammation and endogenous hormone levels, which may contribute to cervical cancer risk. Most ambitiously, over 20,000 DNA and 20,000 plasma specimens are being tested for HPV DNA and antibodies, respectively, to determine how type-specific HPV DNA types (there are over 40 types of cervical HPV) and antibodies influence outcome. NCI and NLM are collaborating to develop methods to permit exploration of visual aspects of HPV and cervical neoplasia. In etiologic studies NCI will relate the numbers of infecting viral types with numbers and positions of lesions. NCI will be able to follow the topographic progression and regression of lesions. For screening research NCI will be able to use 60,000 digitized uterine cervix images from the Guanacaste Project to optimize and standardize visual screening of the cervix. NLM has the role of developing tools and technologies used in these studies.

4.2. Tools and prototype systems for the uterine cervix images

4.2.1. Boundary marking

The boundary marking tool (BMT) [71] allows the collection of a detailed set of data relevant to the evaluation of uterine cervix images, including adequacy of the image for visual evaluation, visual diagnosis, presence of the cervix, and the marking of a set of '3/6/9/12' orientation landmarks, as well as regions corresponding to aceto-whitening, invasive cancer, squamous metaplasia, Nabothian cysts, the os, the squamocolumnar junction

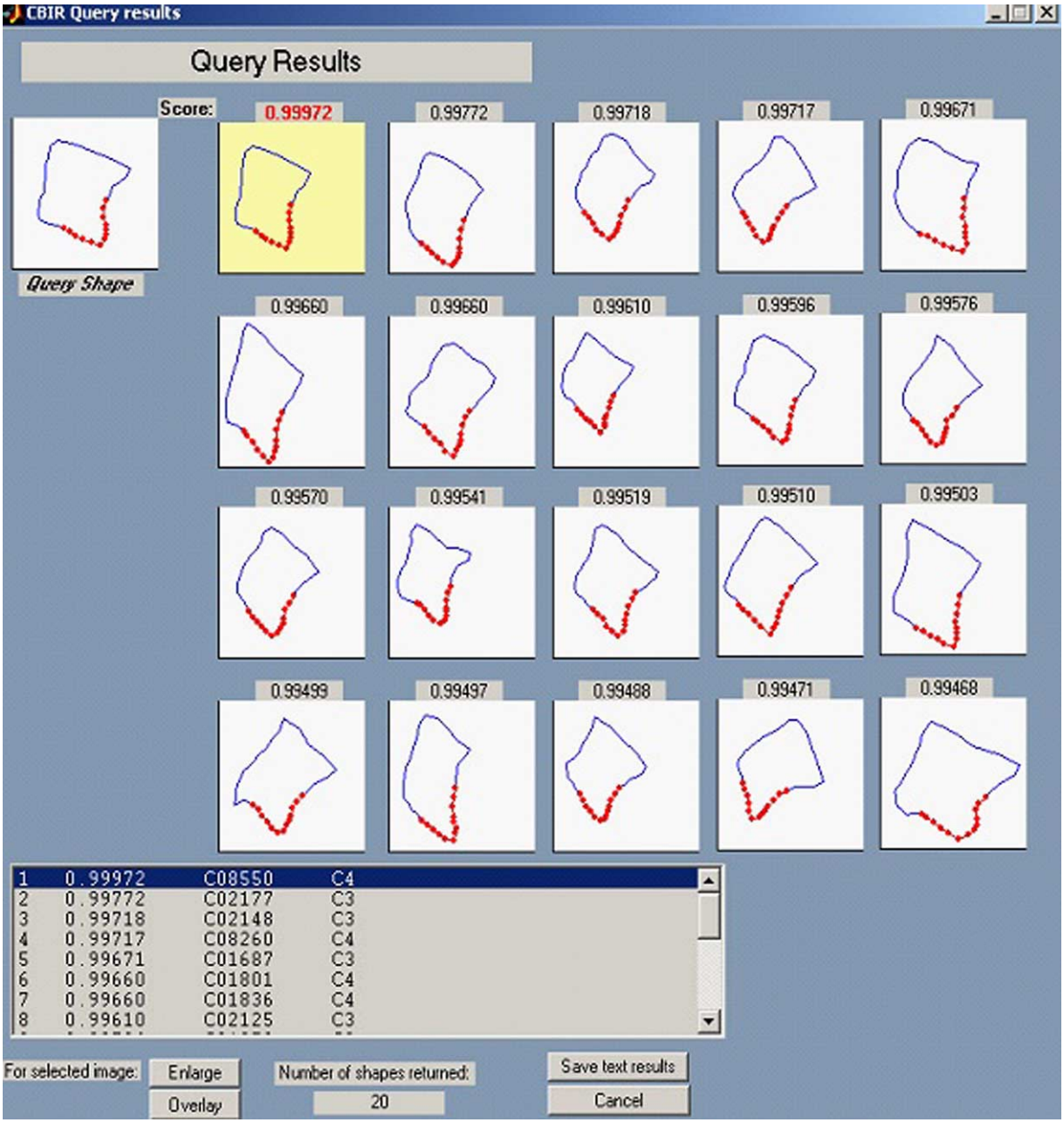


Fig. 16. Results of query-by-sketch, Partial Shape Matching. In the upper left, the user sketch for the query can be seen. The part of the shape (lower corner) for this particular query has been marked by the user. The rectangular display of images to the right contains the query results, ordered by shape similarity (as calculated by the algorithm), to the user's query. The text display at the bottom identifies the images by database ID and anatomical name (e.g. 'C4' means 'cervical spine vertebra 4').

(SCJ), and polyps. The BMT user interface is illustrated in Fig. 17.

Where appropriate, the expert is provided with a pull-down menu of choices. For example, for 'visual diagnosis', the pull-down menu includes the choices of normal, cervicitis, metaplasia, condyloma, low-grade lesion, high-grade lesion, and invasive cancer. When the expert draws boundaries on the image the boundaries are shown in color-coded form. Capability is provided to indicate that a boundary is partially obscured and to record an assumed boundary path for the obscured area. For particular regions, the expert may display a detailed view that allows sub-classification of the region contents. For example, for acetowhite regions, the presence or absence of visual 'tiling' (mosaicism) is recorded, along with a classification of the mosaicism as coarse or fine. In addition, the margins and color of the region are classified using a standard four-level Reid index. All BMT outputs are saved as records in a central MySQL database that resides on a server at NLM; these outputs include the spatial boundary data, which is recorded as a set of (x,y) pixel coordinates in a standard image coordinate system. The BMT is a cross-platform Java

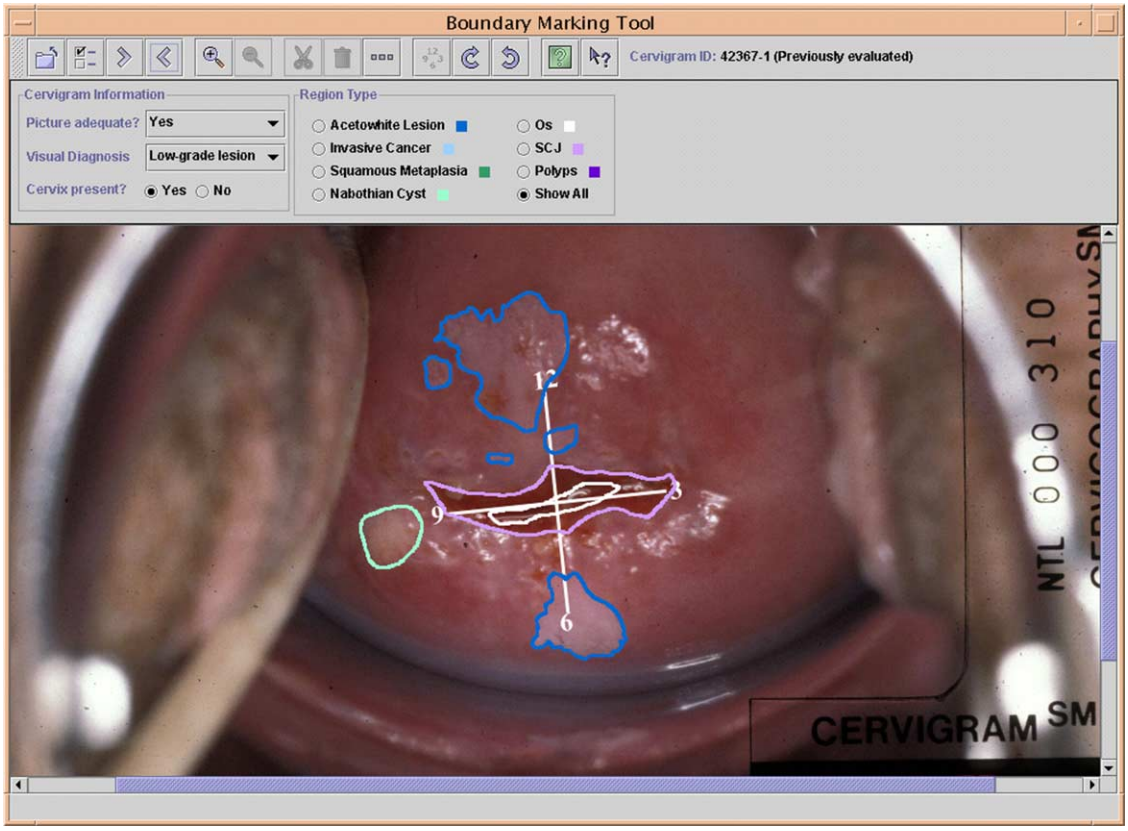


Fig. 17. Boundary Marking Tool for uterine cervix images; regions are color-coded according to region type (acetowhite lesion, invasive cancer, squamous metaplasia, Nabothian cyst). The 3/6/9/12 anatomical 'landmarks' are manually placed by the user to provide the standard orientation for the cervix that is used by researchers and clinicians in this field; their placement is independent of the regions marked in the image.

application that runs on the user's desktop and loads digitized cervigrams from local storage (e.g. DVD) while saving results to the central MySQL database by means of the MySQL connector/j package which implements Sun's Java Database Connectivity (JDBC) API.

The BMT has been already used in a study developed to explore HPV tropisms for glandular or squamous epithelium. Pictures from 1016 patients were reviewed to record the junction between those epitheliums by drawing a boundary around it. The objective of this study is to determine whether a bigger area of glandular epithelium exposed on the ectocervix correlates with infections for specific types of HPV. Currently the evaluation of the pictures has been finished and results are being analyzed.

4.2.2. Multimedia database tool

The multimedia database tool (MDT) [72] has all the functionality of the current WebMIRS system: capability to query a database of text and related images through a Graphical User Interface over the Web, capability to show query results with a multiple image display and concurrent display of associated text data, and capability to export query results to text files for statistical analysis or other purposes. In addition, the MDT adds capability for distributed data collection from remote users; one application of this new capability is to allow users at many geographically dispersed sites to view images and record interpretations for those images into the central database. To support this new capability, a system of user privileges is also being incorporated into the MDT that will restrict database components that are user-writeable and also restrict write privileges to those users authorized by the database administrator. The MDT system has also been designed to support a broad class of text/image databases. The MDT user interface is illustrated in Fig. 18, for a database of uterine cervix images.

An important R&D aspect of the MDT is the investigation into how capabilities originally designed for a particular database application may be generalized to serve many databases. Of particular interest is the investigation of how tools for the retrieval and display of spatial data may be incorporated into the system. For example, the functionality of the BMT (described above) is being incorporated into the MDT to allow the query, retrieval, and display of these regions superimposed on the uterine cervix images. The system is flexible enough that other image highlighting, such as WebMIRS vertebral segmentation, could also be displayed using the same mechanism.

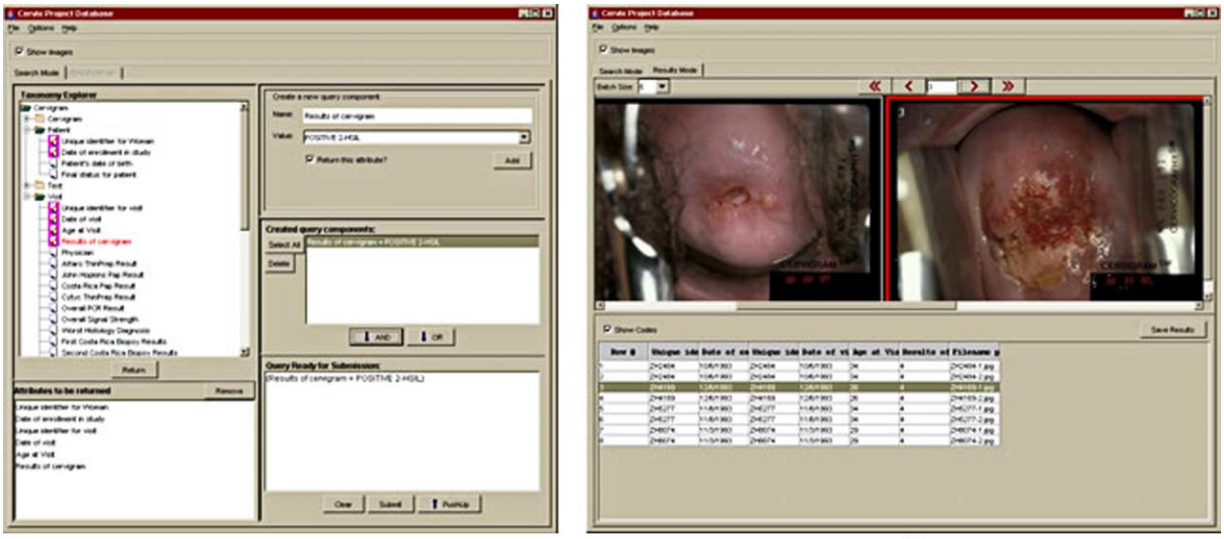


Fig. 18. Mutlimedia Database Tool. (A) Query-building screen with uterine cervix database tree; (B) results screen; records matching the query are returned. Text fields request by the user have their values returned in the table at the bottom of the screen. Corresponding images for each record are shown at the top of the screen in a horizontally-scrolled display.

4.3. Image processing R&D for the uterine cervix images

4.3.1. Compression

Compression for the uterine cervix images is being developed by Mitra [73] by methods similar to those for the NHANES II spine X-ray images, with incorporation of special processing to maintain color fidelity. Codebooks of various sizes and control parameters have been generated, and the HMVQ algorithm has been modified to allow accurate color quality control by allocating bit-rate independently by color channel. Alternative color transformations have been investigated. An enhanced LBG algorithm is currently under investigation to further optimize codebook optimization. Preliminary investigations on the effects of optimized color image compression on quality and bit rate control, as well as segmentation of acetowhite regions of the cervix have been done [74,75].

4.3.2. Segmentation

Preliminary work has begun into techniques for the indexing of the uterine cervix images by color and texture and the segmentation of significant regions, in particular, regions corresponding to the squamous epithelmium, the columnar epithielium, and acetowhitened regions. Greenspan [76] has done preliminary work that classified small patches chosen from these three regions in a test set of 30 images. Gaussian models were created for each region, and the classification was by color alone. Color features from 12 different color spaces were used, and classification errors ranging from 9 to 16% were observed. The general techniques being explored include the CBIR techniques of the Blobworld [77] system; these are being further investigated by Gordon [78]. Yang [75] has also reported initial segmentation investigation for these images, and King [79] has reported the development of a segmentation-based technique to allow these images to be registered with the same orientation with respect to anatomical landmarks.

5. Critical issues

In spite of the progress we have made many critical challenges remain. We enumerate major challenges below:

- Validation of segmentation, classification, and shape matching algorithms
- The integration of competing algorithms, developed by multiple collaborating research teams, into coherent research and operational systems
- Performance in terms of execution time for compute-intensive algorithms such as: the GHT; decompression for fast display of images using customized compression algorithms; and searching of N-dimensional space for similarity matching between a query object and objects in a database of feature vectors for images indexed by content
- Image color and spatial resolution quality preservation for Web-based biomedical image database systems
- Database system architectures that allow incorporation of new multimedia datasets without new programming, and with minimal labor effort of database administrators

Of these issues, algorithm validation is a recognized critical issue within the entire image processing community. Haralick [80] has provided an abstract mathematical

framework applicable to a broad class of problems, including curve fitting, local feature extraction and registration. For biomedical image segmentation in particular, validation approaches have been proposed by Yoo [81], Chalana [82], Udupa [83], and others. The work that we have done in segmentation and classification has incorporated biomedical expert (board-certified) judgment, but has been limited by the amount of data available from experts and the number of expert observers. For shape classification, we have incorporated judgment of medically trained observers, though not at the board-certified level. Increasing the rigor of our validation by building 'truth' sets from multiple biomedical experts and incorporating statistical validation methods as proposed in the references above remains an important goal.

6. Summary and conclusion

In the field of document imaging the work that we have done in document analysis and understanding, including the incorporation of OCR techniques and rule-based algorithms, are contributing to a labor reduction in the acquisition of titles, author names, institutional affiliations, and abstracts from scanned journal articles to populate NLM's MEDLINE system. In addition the DocView program has over 16,000 users who use this software to enhance and organize TIFF documents. DocMorph, with over 11,000 users, provides capability to convert over 50 file formats to PDF, to text (through OCR), and to speech.

In the field of biomedical imaging, the creation of operational image informatics systems that allow users to easily navigate, browse, and retrieve meaningful image contents from complex biomedical image data sets remains an unachieved goal, but a goal that we and our colleagues in the field of content-based image retrieval are approaching. We have created a multimedia database of 17,000 spine X-rays and national health survey data and the WebMIRS software that allows access to it, added initial image content 'truth' data in the form of nine-point morphometry data, and established an FTP archive to allow access to the spine images at full spatial resolution. We have created engineering-level 'truth sets' for first-order validation of our segmentation algorithms and, with model-based algorithms, have achieved over 65% success in our segmentations, using reasonably tight criteria. For biomedical feature classification of lumbar spine anterior osteophytes, we have achieved correct classification of normal vertebra and of abnormal vertebra in approximately 90% of the test cases, where the neural network classifications were judged against the expert classifications of a single expert medical observer. In the area of database retrieval of vertebrae by shape, we have created a prototype CBIR system that in tests to date has been able to return relevant results in the 50–70% range, among the top N results returned, as judged by medically trained observers. We are now expanding our work into the uterine cervix image domain, and have created software for collecting region-of-interest data from these images, which will be made available over the Web with a next generation WebMIRS system, the Multimedia Database Tool. Work is also under way in segmenting, indexing, and eventually retrieving these images by color and texture.

In the field of biomedical document processing, researchers continue to work toward long-range goals that include full-text searching of vast document stores, more streamlined Web delivery of biomedical documents, document understanding by natural language processing, linkage of heterogeneous databases, and automated recognition of machine or hand-printed multilingual text. In the field of biomedical image processing, researchers continue to pursue effective methods to deliver images and associated metadata over the Web, to efficiently segment and index biomedical images by visual or computable image content, and to effectively search and compare images by visual queries. In both these fields, the results that we have achieved to date are modest in comparison with these long term goals. Yet, in the sense of supporting user communities with document imaging and multimedia database R&D products, in demonstrating the use of OCR for labor reduction in document indexing, and in establishing benchmark results in techniques for X-ray image segmentation, shape classification by neural networks, and content-based image retrieval by shape, we believe that the work described here may be considered to have a measurable impact on progress in incorporating the methods of informatics in more integrated and useful ways within the biomedical sciences.

References

[1] Thoma GR. Automating data entry for an online biomedical database: a document image analysis application. Proceedings of fifth international conference on document analysis and recognition (ICDAR'99). Bangalore, India; September 1999. p. 370–3.

[2] Le DX, Thoma GR, Wechsler H. Classification of binary document images into textual or non-textual data blocks using neural network models. Machine Vision Appl 1995;8:289–304.

[3] Hauser SE, le DX, Thoma GR. Automated zone correction in bitmapped document images. Proceedings of SPIE: document recognition and retrieval VII, San Jose, CA, SPIE vol. 3967; January 2000. p. 248–58.

[4] Kim J, le DX, Thoma GR. Automated labeling in document images. Proceeding of SPIE. Document recognition and retrieval VIII, San Jose, CA, vol. 4307; January 2001. p. 111–22.

[5] Ford GM, Hauser SE, Thoma GR. Automated reformatting of OCR text from biomedical journal articles. Proceedings of 1999 symposium on document image understanding technology. College Park, MD: Institute for Advanced Computer Studies; April 1999, p. 321–5.

[6] Mao S, Kim J, le DX, Thoma GR. Generating robust features for style-independent labeling of bibliographic fields in medical journal articles. Proceedings of seventh world multiconference on systemics, cybernetics and informatics, Orlando, FL, vol. III; July 2003. p. 53–6.

[7] Mao S, Kim J, Thoma GR. A dynamic feature generation system for automated metadata extraction in preservation of digital materials. Proceedings of international workshop on document image analysis for libraries (DIAL2004). Palo Alto, CA; January 2004. p. 225–32.
[8] Mao S, Kim J, Thoma GR. Style-independent document labeling: design and performance evaluation. Proceedings of SPIE—document recognition and retrieval. San Jose, CA; January 2004. p. 14–22.
[9] Le DX, Thoma GR. Automated document labeling using integrated image and neural processing. Proceedings of world multiconference on systemics, cybernetics and informatics, Orlando, FL, vol. 6; July 31–August 4, 1999. p. 105–8.
[10] Le DX, Thoma GR, Wechsler H. Document image analysis using integrated image and neural processing. Proceedings of third ICDAR'95. Montreal, Canada, August 14–16, vol. I; 1995. p. 327–30.
[11] Ford G, Hauser SE, le DX, Thoma GR. Pattern matching techniques for correcting low confidence OCR words in a known context. Proceedings of SPIE, document recognition and retrieval VIII, vol. 4307; January 2001. p. 241–9.
[12] Hauser SE, Browne AC, Thoma GR, McCray AT. Lexicon assistance reduces manual verification of OCR output. Proceedings of 11th IEEE symposium on computer-based medical systems. Los Alamitos, CA: IEEE Computer Society; 1998 p. 90–5.
[13] Hauser SE, Sabir TF, Thoma GR. Speech recognition for program control and data entry in a production environment. Proceedings of SPIE: intelligent systems in design and manufacturing II. vol. 3833 1999 p. 24–34.
[14] Le DX, Straughan SR, Thoma GR. Greek alphabet recognition technique for biomedical documents. In: Callaos N et al, editor. Proceedings of 6th world multiconference on systemics, cybernetics and informatics, vol. III, 2002. p. 86–91.
[15] Thoma GR, Ford G, Le DX, Li Z. Text verification in an automated system for the extraction of bibliographic data. Proceedings of fifth international workshop on document analysis systems. Berlin: Springer; 2002 p. 423–32.
[16] Le DX, Tran LQ, Chow J, Kim J, Hauser SE, Moon CW, Thoma GR. Automated medical records citation records creation for Web-based online journals. Proceedings of 14th IEEE symposium on computer-based medical systems. Los Alamitos, CA: IEEE Computer Society; 2001 p. 315–20.
[17] Tran LQ, Moon CW, Le DX, Thoma GR. Web page downloading and classification. Proceedings of the 14th IEEE symposium on computer-based medical systems. Los Alamitos, CA: IEEE Computer Society; 2001 p. 321–6.
[18] Kim J, Le DX, Thoma GR. Automated labeling algorithms for biomedical document images. Proceedings of seventh world multiconference on systemics, cybernetics and informatics, Orlando, FL. vol. V 2003 p. 352–57.
[19] Thoma GR, Hauser SE, Walker FL, Guy L. Electronic imaging techniques in the preservation of the biomedical literature, Proceedings of electronic imaging '88 West, Anaheim, CA, March 28–31; 1988. p. 906–13.
[20] Thoma GR, Walker FL. Archiving the biomedical literature by electronic imaging methods, Proceedings of ASIS, vol. 25; October 1988. p. 132–36.
[21] Suthasinekul S, Walker F, Cookson J, Rashidian M, Thoma G. Prototype system for electronic document image storage and retrieval. J Imaging Technol 1985;II(5):56–62.
[22] Hauser SE, Thoma GR, Gass SI. Factors affecting the conversion rate of bound volumes to electronic form. Proceedings of electronic imaging '89 east conference, Boston, MA; October 1989. p. 1041–6.
[23] Thoma GR. Electronic imaging for document preservation: system performance. In: Roth JP, editor. Chapter in case studies of optical storage applications. Westport, CT: Meckler Publishing; 1990. p. 65–72.
[24] Thoma GR, Guy L. Considerations in document image capture. In: Smith EV, Keenan S, editors. Information, communications, and technology transfer. North-Holland: Elsevier Science; 1987. p. 213–20.
[25] Thoma GR, Cookson JP, Walker FL, Hauser SE. Interfacing optical disks to a document image storage and retrieval system. J Imaging Technol 1986;12(5):288–92.
[26] Hauser SE, Berman LE, Thoma GR. Performance of RAID as a storage system for Internet image delivery. Proceedings of SPIE: multimedia storage and archiving systems, Boston, MA, November 18–19, vol. 2916; 1996. p. 14–22.
[27] Hauser SE, Hsu W, Thoma GR. Request routing with a back error propagation network. Applications of artificial neural networks IV, SPIE proceedings. vol. 1965 1993 p. 689–95.
[28] Hauser SE, Cookson TJ, Thoma GR. Using back error propagation networks for automatic document image classification. Applications of artificial neural networks IV, SPIE proceedings. vol. 1965 1993 p. 142–50.
[29] Thoma GR, Hauser SE, Le DX, Muller D, Walker FL. WILL: advances in the management of interlibrary loan. Internal technical report. vol. 1. Bethesda, MD: National Library of Medicine, Lister Hill National Center for Biomedical Communications, Communications Engineering Branch; 1995. 26 p.
[30] Thoma GR, Hauser SE, Le DX, Muller D, Walker FL. WILL: design of a standalone WILL unit. Internal technical report. vol. 2. Bethesda, MD: National Library of Medicine, Lister Hill National Center for Biomedical Communications, Communications Engineering Branch; 1995. 27 p.
[31] Walker F, Thoma GR. DocView: providing access to printed literature through the Internet. Proceedings of 10th integrated online library systems meeting, May 3–4. New York: Learned Information, Inc; 1995 p. 165–73.
[32] Thoma GR, Walker FL. Access to document images over the Internet. In: Rada R, Ghaoui C, editors. Medical multimedia. Oxford, UK: Intellect; 1995. p. 179–93.
[33] Walker FL, Thoma GR. Web-based document image processing. Procedings of SPIE: Internet Imaging. San Jose, CA; January 2000. p. 268–77.
[34] Walker FL, Thoma GR. Read it to me!. Proceedings of 21st national online meeting. Medford, NJ: Information Today, Inc; 2000 p. 473–83.
[35] Walker FL, Thoma GR. A SOAP-enabled system for an online library service. Proceedings of infotoday 2002. Medford, NJ: Information Today; 2002 p. 320–9.
[36] National Center for Health Statistics Web site: http://www.cdc.gov/nchs/about/major/nhanes/goals.htm
[37] Long LR, Antani S, Lee DJ, Krainak, Thoma GR. Biomedical information from a national collection of spine X-rays: film to content-based retrieval. Proceedings of SPIE medical imaging 2003: PACS and integrated medical information systems, 5033, San Diego, CA, February 15–20; 2003. p. 70–84.
[38] 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. p. 4.
[39] Plan and operation of the Third National Health and Nutrition Examination Survey 1988–94. National Center for Health Statistics. Vital Health Stat 1994;3(32).
[40] 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. J Digital Imaging 1998;11(3):116–20.
[41] Duryea J, Zaim S, Wolfe F. Neural network based automated algorithm to identify joint locations on hand/wrist radiographs for arthritis assessment. Med Phys 2002;29(3):403–11.

[42] Long LR. Workshops for the use of digitized NHANES II images. Technical summary for Communications Engineering Branch of the National Library of Medicine.
[43] Long LR. Quality control for the digitization of the NHANES II images. Technical summary for Communications Engineering Branch of the National Library of Medicine. International conference on document analysis and recognition ICDAR 95. Montreal, Canada, August 14–16, vol. I; 1995. p. 327–30.
[44] 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. p. 362–72.
[45] 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. p. 392–403.
[46] 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. p. 586–94.
[47] The epidemiology of chronic rheumatism. Atlas of standard radiographs of arthritis, vol. II. Philadelphia, PA: FA Davis Company; 1963.
[48] Ghebreab S. Strings and necklaces: on learning and browsing medical image segmentations. Amsterdam, The Netherlands: Intelligent Sensory Information Systems Group, University of Amsterdam; 2002.
[49] Zamora G. Dissertation for the degree of Doctor of Philosophy. Department of Electrical Engineering, Texas Tech University, Lubbock, TX; December 2002.
[50] 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, San Diego, CA, February 24–28; 2002. p. 371–82.
[51] Yang S, Mitra S. Content-based vector coder for information retrieval. BISC-FLINT international workshop, UC Berkley, Berkeley, CA, August 14-18; 2001.
[52] Tagare H. Deformable 2-D template matching using orthogonal curves. IEEE Transact Med Imaging 1997;16(1):108–17.
[53] Cootes TF, Taylor CJ. Statistical models of appearance for computer vision. Technical report, University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, Manchester, M12 9PT, UK; 2001 February.
[54] Sari-Sarraf H, Mitra S, Zamora G, Tezmol A. Customized active shape models for segmentation of cervical and lumbar spine vertebrae (contract technical report); 2000 December.
[55] 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. p. 378–87.
[56] Ballard DH. Generalizing the Hough transform to detect arbitrary shapes. Pattern Recogn 1981;13(2):111–22.
[57] Tezmol A. Customized Hough transform for robust segmentation of cervical vertebrae from X-ray images. Master's Thesis in Electrical Engineering, Texas Tech University; 2001.
[58] Gururajan A. Coarse segmentation of cervical and lumbar vertebrae using a customized version of the Generalized Hough Transform. Master's Thesis in Electrical Engineering, Texas Tech University.
[59] Zamora G, Sari-Sarraf H, Mitra S, Long R. Analysis of the feasibility of using active shape models for segmentation of gray scale images. Proceedings of SPIE medical imaging 2002: image processing. 4684, San Diego, CA, February 24–28; 2002. p. 1370–1381.
[60] Howe B. Segmentation of cervical and lumbar vertebrae in X-ray images using active appearance models and extensions. Thesis in Electrical Engineering, Texas Tech University; 2004.
[61] Cherukuri M, Stanley RJ, Long R, Antani S, Thoma G. Anterior osteophyte discrimination in lumbar vertebrae using size-invariant features. Comput Med Imaging Graphics 2004;28:99–108.
[62] Charmathy P, Stanley RJ, Cizek G, Long R, Antani S, Thoma G. Image analysis techniques for characterizing disc space narrowing in cervical vertebrae interfaces. Comput Med Imaging Graphics 2004; 28:39–50.
[63] Krainak DM, Long LR, Thoma GR. A method of content-based retrieval of a spinal X-ray image database. Proceedings of SPIE medical imaging 2002: PACS and integrated medical systems, San Diego, CA, February 23–28, vol. 4685; 2002. p. 108–16.
[64] Dryden IL, Mardia KV. Statistical shape analysis. New York: Wiley; 1998.
[65] Antani S, Long LR, Thoma GR. A biomedical information system for combined content-based retrieval of spine X-ray images and associated text information. Third Indian conference on computer vision, graphics, and image processing (ICVGIP '02), Ahmedabad, India; December 2002. p. 242–47.
[66] Antani S, Long LR, Thoma GR, Lee DJ. Anatomical shape representation in spine X-ray images. Proceedings of the third IASTED international conference on visualization, imaging and image processing (VIIP 2003), Benalmadena, Spain, vol. 1; September 2003. p. 510–5.
[67] Antani S, Xu X, Long LR, Thoma GR. Partial Shape Matching for CBIR of spine X-ray images. Proceedings IS&T/SPIE electronic imaging–storage and retrieval methods and applications for multimedia 2004. San Jose, CA. SPIE vol. 5307; January 2004. p. 1–8.
[68] Heggeness MH, Doherty BJ. Morphologic study of lumbar vertebral osteophytes. South Med J 1998;91(2):187–9.
[69] Pate D, Goobar J, Resnick D, Haghighi P, Sartoris DJ, Pathria MN. Traction osteophytes of the lumbar spine: radiographic-pathologic correlation. Radiology 1988;166(3):843–6.
[70] Herrero R, Schiffman MH, Bratti C, et al. Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. Pan Am J Public Health 1997;1(5):362–75.
[71] Jeronimo J, Schiffman M, Long LR, Neve L, Antani S. A tool for collection of region-based data from uterine cervix images for correlation of visual and clinical variables related to cervical neoplasia. Proceedings of IEEE computer based medical systems 2004, Bethesda, MD, June 24–25; 2004.
[72] Long LR, Bopf M, Coleman T, Antani S, Jeronimo J, Schiffman M, Thoma GR. An architecture for streamlining the implementation of biomedical text/image databases on the Web. Proceedings of IEEE computer based medical systems 2004, Bethesda, MD, June 24–25; 2004.
[73] Mitra S, Nutter B, Guo J. A Wavelet-based multi-spectral codec for digitized cervigram images. Technical progress report for the communications engineering branch of the National Library of Medicine, March 16; 2004.
[74] Guo J, Shrivastava P, Kepley K, Yang S, Mitra S, Nutter B. Bit-rate allocation control and quality improvement for color channels in HMVQ image compression. Proceedings of IEEE computer based medical systems 2004, Bethesda, MD, June 24–25; 2004.
[75] Yang S, Guo J, King P, Sriraja Y, Mitra S, Nutter B, Ferris D, Schiffman M, Jeronimo J, Long R. A multi-spectral digital cervigram™ analyzer in the wavelet domain for early detection of cervical cancer. Proceedings of SPIE image processing symposium, San Diego, California, February 15–20, vol. 5370; 2004.
[76] Greenspan H. Image content indexing of uterine cervix images. Technical progress report for the communications engineering branch of the National Library of Medicine, January 15; 2004.
[77] Carson C, Belongie S, Greenspan H, Malik J. Blobworld: image segmentation using expectation–maximization and its application to image querying. IEEE Transact Pattern Anal Machine Intell 2002; 24(8):1026–38.

[78] Gordon S, Zimmerman G, Greenspan H. Image segmentation of uterine cervix images for indexing in PACS. Proceedings of IEEE computer based medical systems 2004, Bethesda, MD, June 24–25; 2004.
[79] King PS, Mitra S, Nutter. An automated, segmentation-based, rigid registration system for cervigram images utilizing simple clustering and active contour techniques. Proceedings of IEEE computer based medical systems 2004, Bethesda, MD, June 24–25; 2004.
[80] Haralick RM. Validating image processing algorithms. Proceedings of SPIE medical imaging, February 12–18, vol. 3979; 2000. p. 2–16.
[81] Yoo TS, Ackerman MJ, Vannier M. Toward a common validation methodology for segmentation and registration algorithms. Medical image computing and computer-assisted intervention, lecture notes in computer science. vol. 1935. Berlin: Springer; 2000 p .422–31.
[82] Chalana V, Kim Y. A methodology for evaluation of boundary detection algorithms on medical images. Proceedings of SPIE medical imaging, vol. 2710; 1996.
[83] Udupa JK, LeBlanc VR et al. A methodology for evaluating image segmentation algorithms. Proceedings of SPIE medical imaging, vol. 4684, February 24–28; 2002. p. 266–277.

L. Rodney Long is an electronics engineer for the Communications Engineering Branch in the Lister Hill National Center for Biomedical Communications, a research and development division of the US National Library of Medicine. He has held his current position since 1990. Prior to his current job, he worked for 14 years in industry as a software developer and as a systems engineer. His research interests are in image processing and scientific/biomedical databases, with emphasis on the content-based retrieval of medical images. He has an MA in applied mathematics from the University of Maryland.

Sameer K. Antani is a scientist with the Communications Engineering Branch in the Lister Hill National Center for Biomedical Communications at the National Library of Medicine. He earned his BE (Computer) degree from the University of Pune, India, in 1994; and ME and PhD degrees in Computer Science and Engineering from The Pennsylvania State University in 1998 and 2001, respectively. His current research interests are in content analysis of visual information, multimedia databases, and medical image databases. His current projects include (complete and partial) shape representation and similarity algorithms for spine X-ray image retrieval, color- and texture-based segmentations of acetowhitened images of the uterine cervix, and image indexing algorithms.

George R. Thoma is Chief of the Communications Engineering Branch of the Lister Hill National Center for Biomedical Communications at the National Library of Medicine. In this capacity, he directs R&D programs in document image analysis and understanding, biomedical image processing, image compression, automated document image delivery, digital X-ray archiving, animated virtual books, and high speed image transmission. He earned a BS from Swarthmore College, and the MS and PhD from the University of Pennsylvania, all in electrical engineering. Dr Thoma is a Fellow of the SPIE, the International Society for Optical Engineering.