National Library of Medicine

Board of

Scientific

Counselors

Meeting

September 26 - 27, 2002

Biomedical Imaging Research & Development

Content-Based Image Retrieval (CBIR) of Biomedical Images

**A report to the
Board of Scientific Counselors**

September 26-27, 2002

Communications Engineering Branch

**Lister Hill National Center for Biomedical Communications
National Library of Medicine**

Contents

1 Introduction and background

> *We are, it appears, on the hinge of an important historical swing back towards what may be called the primacy of the image.* [1]

The ever-increasing volume of medical images, the economic impracticality of manually indexing these images, and the inadequacy of human language alone to describe image contents that are visually recognizable and medically significant, such as shape and geometry, color, texture of objects within images, all provide impetus for research and development toward practical *Content-Based Image Retrieval* (CBIR) systems that could become a standard offering of the medical library of the future.

Content-Based Image Retrieval refers to the retrieval of images that are indexed by descriptors (features) derived "directly" from the image pixels. CBIR also implies novel query methods for retrieval, such as finding images similar to an input example image, or to an input sketch. CBIR descriptors may include texture, boundaries of objects, geometric relationships among objects, grayscale or color histograms, as well as more abstract descriptors, such as Fourier or Wavelet transform coefficients, or transformed and reduced boundary coordinates. CBIR is different from conventional image retrieval systems in two ways: first, the *methods used to index* (a conventional system has a human indexer enter text that describes image contents), and secondly, in the *methods used to retrieve* (conventionally, retrieval of images is by relational database queries on the text used to index the image.)

It should be recognized at the outset that CBIR and, indeed, the entire field of work to exploit information from medical images is still at the research level, as articulated in the following quotation by Casimir Kulikowski:

> *"...it is still early days for medical imaging informatics. The structuring of medical information and knowledge...is only the beginning of a challenging scientific, technological, and empirical process that will have to draw on a multitude of disciplines. Besides the central computer vision, graphics and visualization, database, software engineering, artificial intelligence, distributed processing, networking, and interface research, basic work is needed in the underlying psychophysics, cognitive science, and mathematical/statistical modeling of the imaging modalities and their interpretation by synergistic human-machine methods. Considerable efforts will be needed to connect existing text-based semantic models with visual/graphical models of medical information and knowledge so that practitioners, researchers, educators, students, and patients can not only navigate but also understand the results of querying and interacting with our evolving multimedia databases."* [2]

The potential pay-off from medical imaging informatics, and from CBIR in particular, is highly significant. According to a 2001 technical review [3] of the field, CBIR is critical in digital libraries for patient care, clinical diagnosis and decision making in large-scale clinical trials, managing large-scale protein image databases, and in biomedical education.

In this introductory section we summarize the biomedical imaging work done at the Communications Engineering Branch (CEB). Sections on the objectives and significance of the ongoing and

proposed CBIR work follow. We then discuss CBIR in technical detail, including descriptions of the in-house and collaborative work done to date. A section describing the CBIR prototype system that we have created is next, followed by a section describing the proposed future CBIR work. The body of the report concludes with a summary recapitulating the presentation.

CBIR research is highly technical, mathematical, and requires the exploration of numerous alternative technical solutions. To maximize chances of success, alternatives may be pursued in parallel, with little a priori knowledge of the outcomes. For example, we have pursued several parallel strategies for x-ray segmentation, as will be shown. In addition, we have found it advantageous to conduct research activity in an order that may be logically out-of-sequence. For example, we have investigated the classification of vertebral shapes (by using manual segmentations of vertebrae) before solving the computer-assisted segmentation problem. For readability and to retain the flow of the discussion in this report, we place the most mathematical and detailed parts within the appendices.

1.1 Biomedical Imaging: CEB retrospective

The origins of this project lie in an effort a decade ago to preserve a collection of spine x-rays acquired as part of a periodic nationwide survey of public health conditions called the National Health and Nutrition Examination Survey (NHANES). The U.S. National Center for Health Statistics (NCHS) conducts this survey by setting up tractor trailer-based Medical Examination Centers appropriately equipped and staffed at selected locations throughout the U.S. Data is gathered on a sampling of the population at each site. The second such survey, NHANES II, yielded a broad spectrum of information on each of 25,286 participants, of which 20,322 were both interviewed and examined [4]. The data taken included medical examination data, demographic information, and blood chemistry analyses. In addition, a subset of the participants received a detailed examination that included radiographs of the cervical and lumbar spine. This resulted in a collection of approximately 17,000 films. The third survey, NHANES III, produced an estimated additional 10,000 films of hands, wrists and knees, though these have not been publicly released.

Biomedical imaging R&D within CEB has concentrated largely, though not exclusively, on the x-ray set collected by NHANES II. Since an original objective was to investigate the technologies for archiving and disseminating these x-rays, the 10,000 cervical radiographs and 7,000 lumbar spine radiographs were digitized at 146 dpi using a Lumisys laser scanner, and currently reside in 140 GB of magnetic storage on a CEB server, from which they are publicly accessible for FTP transfer. In addition, low resolution versions of the images may be displayed in response to queries to the WebMIRS system (described below). As mentioned, along with the x-ray images, the NHANES II survey also included information on demographics, health questionnaire responses and physician's examination results. Over 2000 fields of such information were collected on each surveyed person, providing a large body of textual information, most of which is publicly available through the WebMIRS system. It is this image set and the associated text that have been used in several CEB projects and that now inform the present project whose incentives include: (a) the classification of the images for biomedical researchers, in particular the osteoarthritis research community – a longstanding goal for us as well as our collaborators at NCHS and the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS); (b) the capability to retrieve images based on geometric characteristics of the vertebrae – of interest to the vertebral morphometry community; and

(c) the development of automated or computer-assisted classification and retrieval methods, highly desirable to offset the high cost of manual classification by medical experts. Two NIH workshops identified 25 biomedical features of interest in the NHANES II spine x-ray images, viz., anterior osteophytes, disk space narrowing, subluxation in the cervical spine and spondylolisthesis in the lumbar spine, posterior osteophytes, plate erosion or sclerosis, vacuum phenomenon, abnormalities, ankylosing spondylitis, apophyseal OA, congenital/developmental disease, DISH, evidence of surgery, fracture, infection, disc calcification, neuropathic spine, osteopenia, Paget's disease, rheumatoid arthritis, spondyloarthropathy, spondylosis deformans, anterior ligamentous calcification, congenital fusion, and tumor. However, only 3 features could be reliably and consistently detected, viz., anterior osteophytes, disc space narrowing, and subluxation for the cervical spine and spondylolisthesis for the lumbar spine. We have concentrated the automated classification and CBIR efforts on these.

The CEB projects using the NHANES II x-ray images as the principal image data have been favorably reviewed by the Board of Scientific Counselors and are summarized below.

DXPNET. The *Digitized X-ray Prototype workstations linked via InterNET* project, a collaborative effort among NLM, NIAMS, and NCHS, developed systems for the collection of controlled radiological interpretations of the NHANES II digitized x-ray images by multiple readers (domain experts) in the distributed environment of the Internet [5]. For this project CEB accomplished: the software development and system integration necessary to allocate images to multiple, geographically-distributed readers; transmit the images from CEB to the readers using specially-developed multisocket transmission methods [6] for enhanced efficiency; display the images on high-resolution (2Kx2.5K) Megascan monitors hosted by Sun workstations at reader sites; acquire image interpretations through an onscreen template (designed in collaboration with NIAMS domain experts); transmit the interpretations to a central database maintained by CEB; and allow for system monitoring and intervention for quality control.

DXPNET eventually led to the set of biomedical imaging research activities that are ongoing today, including WebMIRS (which evolved from an initial non-Web, Sun workstation-based, Medical Information Retrieval System), the FTP X-ray Archive, and the Digital Atlas of the Cervical and Lumbar Spine. One technical accomplishment related to the DXPNET project, in collaboration with NCHS domain experts, was to establish the level of digitization to be used for the 5,000 hand x-ray films collected by the NHANES III survey, conducted 1988-1994. A multiple-reader data collection was carried out at the CEB site, using high-resolution Megascan monitors, display and data collection software developed by CEB, and data consisting of 49 hand radiographs collected from the Pima Indian population. Each radiograph was digitized at three different resolutions: 2001x1634, 3000x2400, and 4900x3000, all with a pixel depth of 12 bits. Two radiologists independently graded each of the three sets of digital images for degree of bone erosions. A pre-existing grading of the original film by a different group of radiologists was used as a reference. Based on the outcome of the study, the NCHS elected to digitize the NHANES III x-ray films at the medium resolution (3000x2400), striking a balance between effective image quality and storage and transmission requirements [7].

WebMIRS. The *Web-based Medical Information Retrieval System* [1] is a research and development project to design and evaluate technology for providing wide-area network access to biomedical

---

[1]WebMIRS can be accessed at http://archive.nlm.nih.gov/proj/webmirs.php

User Category	WebMIRS	FTP X-ray Archive
Academia	63	65
Corporate	31	33
Hospital/Medical	17	27
Government	10	4
Military	2	2
Personal	2	13
Other/Unknown	10	10

Table 1: WebMIRS and FTP X-ray Archive: Users by Category

WebMIRS Usage Category	Count
General Research	44
Image interest	22
Epidemiology	16
Information technology	15
Casual	15
Classroom/education	6
Building similar system	6
Rheumatology	4
Other/Unknown	7

Table 2: How WebMIRS is used.

databases containing both text and images, and as an end-to-end research and analysis tool.

Specific achievements for WebMIRS include:

- Database access through the leading, standard Java-enabled World Wide Web browsers (Netscape/Internet Explorer), on PCs, Sun Solaris machines, and potentially any platform providing the same level of Java support;
- Optimum GUI design for relational database query capability;
- Retrieval of text as well as the associated x-ray images (if available);
- User capability to save results for subsequent review and documentation;
- User capability to export results to standard statistical tools (SAS, SUDAAN) for analysis;
- Support for access to multiple databases;
- Access to coarse-level segmentation data marked by medical expert for a subset of the images for use in image processing research.

A view of the WebMIRS query results screen, in the optional "Image View" mode, is shown in Figure 1(a). The textual data in the bottom window corresponds to the highlighted x-ray image of a particular survey subject. Two databases are supported by the WebMIRS system; these databases have been created from the data in NHANES II and III surveys. The WebMIRS NHANES II database contains demographic, anthropometric, adult health questionnaire and physical examination data for 20,322 NHANES II survey participants. The WebMIRS NHANES III database contains demographic, adult health questionnaire, youth health questionnaire, and laboratory data for all of the approximately 30,000 NHANES III survey participants. The 17,000 cervical and lumbar spine images collected during NHANES II are available for display through WebMIRS, in low-resolution form, in addition to this text survey data. For both databases, the statistical weights and survey design variables are available for proper analysis of the data.

Our design and development of WebMIRS has benefited from close collaboration with NCHS and NIAMS, and has also incorporated the advice and recommendations of statisticians expert in the nuances of the use and interpretation of health survey data, including Dr. Donna Brogan [8] of Emory University and Dr. Barry Graubard [9] of the National Cancer Institute. Currently, WebMIRS has approximately 135 registered users in the U.S. and in 28 foreign countries. Current WebMIRS users are predominantly in the academic world; a breakdown of users by category is

FTP X-ray Archive Usage Category	Count
Image processing/compression/display/database	42
Medical research/other medical interest	32
Medical education	21
Art/illustration	20
General research	11
Ph.D. thesis	4
Physiology/kinematics	2
Other/Unknown	22

Table 3: FTP X-ray Archive Usage by Category

given in Table 1. The ways in which WebMIRS is being used is given in Table 2. An example of a classroom use of WebMIRS is as a hands-on tool in a graduate class in public health statistics at Columbia University.

FTP x-ray image archive. All 17,000 NHANES II x-ray images have been made publicly available through an FTP archive that is publicized on the CEB Website. To view these images in full spatial and grayscale resolution, CEB has developed a Java image viewer that is publicly available from the same site. For users preferring to use standard and widely available image viewers, 550 of these images have been converted to the standard TIFF 8-bit form and made publicly available also, along with coarse segmentation data acquired by a medical expert. There are 154 registered users of the FTP archive. A breakdown of these users, by category, is given in Table 3. These images, and the segmentation data, have been accessed for use in a number of technical papers, and for four Ph.D. theses; one of these was recently published [10] and directly addresses the digital library problem of searching large image collections for images satisfying user criteria imposed on objects in the image. A breakdown of the ways in which these images are being used is given in Table 3.

Are the x-ray images useful for other than spine data? We see evidence that exploitation of the images for purposes other than spine information may be possible. Researchers in Spain [11], for example, have developed image processing algorithms for the automatic localization of landmarks within the skull and the extraction of geometric measurements derived from these landmarks. This process has application to the practice of orthodontics, and, when carried out manually, can take 10-15 minutes per image.

Digital Atlas of the Cervical and Lumbar Spine. 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 in collaboration with NIAMS and NCHS [12] to fill a perceived need for such reference data for osteoarthritis in the cervical and lumbar spine, especially since a standard reference [13] of photographs of these features is out of print and difficult to obtain. Important features of the Atlas include:

- 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;
- Display of single or multiple Atlas images simultaneously;
- Built-in image processing capability;
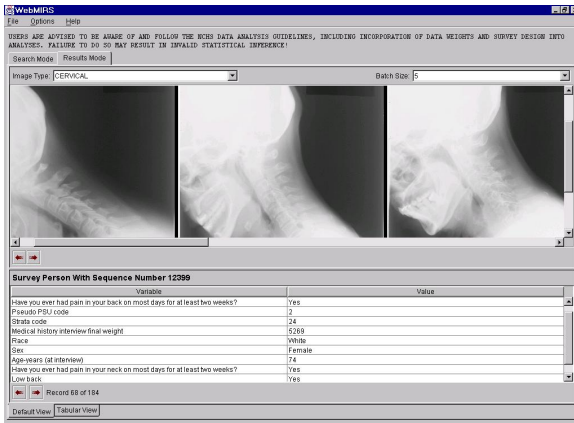
(a)
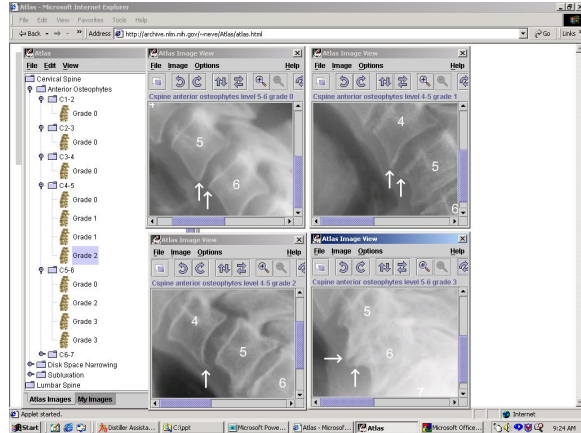
(b)
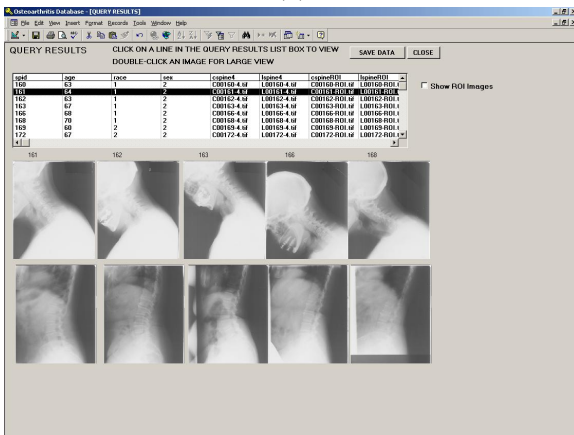
(c)

(d)

Figure 1: Related CEB Projects: (a) WebMIRS Query Results Screen, Image View (b) Atlas Images for CSPINE Anterior Osteophytes (c) Spine X-ray Multimedia Database Implemented with Microsoft Access (d) Visible Human Image Downloading Application Using Lossless Compression.

- Capability to add user-provided images to the Atlas, without code modifications.

An example of an Atlas display is given in Figure 1(b). 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.

Multimedia databases on commodity products. The migration of usable multimedia database products from the high-end of the workstation and software spectrum to the consumer or commodity end has been a topic of research interest in CEB. Toward this goal, a multimedia database product has been developed with Microsoft Access software that accesses the database from a CD-ROM drive. This product provides access to back pain-related survey data for 14,000 adults in the NHANES II health survey, along with digitized x-ray images for 550 of the older survey participants. A graphical user interface provides simple query capability by checkbox and mouse clicks, as well as the capability to export results in several formats, including ASCII, Excel, and HTML. A query results screen from this application is shown in Figure 1(c).

Other biomedical imaging R&D. Apart from the x-ray images, biomedical imaging R&D at CEB has involved the Visible Human data. Lossless compression techniques for exploiting slice-to-slice pixel similarities were the subject of one collaborative effort [14]. This work resulted in the development of a method for lossless compression of the Visible Human images by a process of removing the image background, and then applying Adaptive Arithmetic Coding to the remaining foreground image. The compression ratio achieved over the entire Visible Human 15 GB dataset was 9.2:1 a three-fold improvement over conventional lossless common techniques, a significant step for improving transmission efficiency over the Internet [15]. Figure 1(d) shows one screen of the application developed to allow downloading of Visible Human images using this lossless compression.

In addition, AnatQuest, a system analogous to MapQuest, makes VH images easily available to a large constituency. It is a web interface with a viewport into the anatomy so that a user may: navigate in 3 dimensions, along z- as well as x- and y-axes; zoom to navigate through high resolution images; display both raw image slices (sagittal, coronal, axial), as well as rendered organs (one by one, or streaming video, depending on user bandwidth available); show anatomic labels to identify structures; dissect rendered images to reveal internal structure. The AnatQuest viewer fetches desired images from a database server, and transfer mechanisms accommodate low bandwidth connections by transferring only portions sufficient to fill the display frame of the viewer. For high bandwidth connections, streaming video will give continuous motion display of organs. The system is based on Java Advanced Imaging (JAI: Java 3D) released for Web, which possesses image processing features, e.g., contrast enhancement, scaling, cropping.

2 Project significance

The significance of images as a key component of biomedical information, and therefore an area of concern to the NLM mission, was acknowledged at least as long ago as 1986:

> *In the area of nontextual signals, the field most in need of initiative by the NLM is that of the handling of images which are important to the biological community.* [16]

An NIH workshop sponsored by the National Cancer Institute was convened in 1993 on the specific topic of medical image databases and arrived at recommendations [17] for future research that included:

- Development of descriptive language to relate image features to disease
- Advanced non-textual query capability
- Image indexing using image features
- Support for spatial relations for queries to detect change by shape and size, but are robust enough to allow for deformations
- Tolerance due to levels of uncertainty in identifying objects with vague boundaries
- Cohesive unification of data from various sources
- Development of imaging tools for automatic segmentation, labeling and organization of normal and abnormal anatomy for retrieval, analysis and classification

Our R&D project in CBIR reflects many of these recommendations. We believe that the NHANES images are particularly helpful in CBIR research since they are associated with textual data that could serve as ground truth. Nevertheless, the tools and techniques we will use or develop will address problem domains extending beyond these x-ray images to CT scans, MRI, functional MRI, medical videos, and other image modalities relevant to biomedicine.

3 Objectives

Our overall goal is to make a significant contribution to the image- and multimedia-rich digital library of the future by advancing CBIR techniques applied to the NHANES x-ray images, and eventually to other biomedical images.

Specific objectives include:

- Conduct R&D in the steps needed for content-based indexing, viz., shape segmentation, feature extraction, feature vector organization and classification.
- Conduct R&D into techniques for image-based retrieval.
- Develop the algorithms needed to implement both indexing and retrieval.
- Design and develop a next-generation CBIR system (incorporating these algorithms) to serve as a platform for evaluating techniques that index and retrieve the NHANES images in an effective manner.
- In the long term, extend CBIR techniques developed for the NHANES images to other biomedical images.

4 CBIR - Research status

CEB researchers have collaborated with the biomedical and academic communities in the image segmentation and feature classification, and also in the area of image compression. In this section we discuss our CBIR work and related technical approaches and issues using the general scheme given above. The techniques and work products described in this section led to the creation of a CBIR prototype (CBIR1) in 2001. Discussed in Section 4.4.5, CBIR1 used simple 9-point vertebra boundaries. Ongoing work to create and further develop CBIR2, a more advanced prototype is described in Section 5.

4.1 Definitions

CBIR comprises both indexing and retrieval, defined as follows:

- Indexing (Section 4.3) - the computer-assisted data reduction of images into mathematical features; Indexing may be subdivided into the steps of

  - Segmentation - the computer-assisted determination of the boundaries of the objects of interest;

  - Feature extraction - the reduction of the boundaries that result from the segmentation process into mathematical "feature vectors" that capture the shape properties of interest;

- Feature vector organization - the organization of the feature vectors in the database into a structure optimized for searching efficiency;
- Classification - the computer-assisted labeling of segmented objects into categories of interest.

- Retrieval (Section 4.4) - the user interaction to retrieve desired images from the database; this comprises

  - User query formulation - the method used to specify the query; the advanced methods of interest are *query by image example* and *query by sketch* applied via different search paradigms, such as *target-search*, *category search*, and *open ended browsing* (Section 4.3.3).;
  - User query feature extraction - the reduction of the user's exemplary image or sketch into a mathematical feature vector compatible with the feature vectors stored in the database;
  - Query search space strategy - the search strategy used to efficiently search the database of feature vectors for those near to the user query feature vector; usually, a search strategy is defined by the choice of organization of the feature vectors;
  - Similarity matching method - the method used to compare input and stored image features and to measure similarity between them.

4.2 Related Work

Research in content-based image retrieval today is a lively discipline, expanding in breadth. As happens during the maturation process of any discipline, after early successes in few applications, research is now concentrating on deeper problems, challenging the hard problems at the crossroads of the discipline from which it was born: computer vision, databases, and information retrieval. [18]

Research in Content-Based Image (and Video) Retrieval has widened in scope from retrieval of family photographs, and speciality images such as flower collections, to trademark identification, video analysis and is now at the footstep of biomedical images. A variety of approaches has been taken for achieving the goal of CBIR, and initial successes have opened many avenues for research. The popular approaches and features are well documented in [19, 20]. An important step in building a CBIR system for a set of images is identifying the image features that can be used to achieve the seemingly dichotomous task of efficiently discriminating between the images while determining similarity between like images at a "semantic" level. Thus, much research has been devoted to identifying meaningful image features that uniquely identify the image, developing methods for computing similarity among these features, and applying the feature similarity to image semantics. There has also been much interest in identifying suitable data organization approaches to increase search efficiency.

In work toward identifying suitable image features, typically color, texture, and shape have been used. However, spine x-rays for example, do not have color and do not exhibit any repeating textures. Object shape is thus the only useful feature. The object could be a portion of the spine

or just an individual vertebra. If the spine is viewed as being composed of individual vertebrae, then suitable image segmentation methods need to be identified or developed for extracting the vertebral boundary. Available techniques and adopted approaches are discussed in 4.3.1. Once the vertebra boundary has been extracted, the next step is to represent it in a form that has low dimensionality while capturing the features of interest in the shape. The methods that achieve this and that determine similarity between shapes are discussed in Appendix A.

There has been a lot of interest, recently, in developing biomedical image databases and indexing techniques that support content-based image retrieval [21, 22, 23, 24], but much work needs to be done for developing mature production-level CBIR systems. The various complexities in indexing biomedical images, and challenges in indexing of their "content" have been discussed in [17, 25]. Once suitable image features have been identified, the problem of storing these image features also is very challenging. Not only is the indexing strategy important for efficient retrieval, but it must also address the variety of queries that can be posed to the system. Several studies have been done in developing indexing structures to make the search effective and efficient [26, 27, 28, 29, 30, 31, 32, 33].

4.3 Indexing

The complexity and context-relatedness of medical image content should dismiss false hopes that image indexing can occur fully or that there exists some universal primitive. [25]

It has been observed [34] that the problem of image indexing has been approached by two different groups, using distinctive methodologies. One group, clustered largely around the more traditional, text-oriented library informatics world, has approached the problem as a task in efficiently adding text descriptors to images. An example of this approach is the Greater Cincinnati Memory Project [35] which undertook indexing a collection of thousands of historical photographs: an on-screen template was developed that allowed an indexer to manually enter descriptive text in a template. The second group, clustered largely around computer science work, has approached the problem through image processing. This second group has come to be largely identified with CBIR. We also note a third research direction, proposed by Goodrum [36], that seeks to combine image processing with text labeling of images. Specifically, the hypothesis is that, for certain image classes, images that are found to be similar in terms of well-chosen primitive characteristics detectable by image processing may be reasonably labeled with similar text labels. Some success has been reported for this approach [37] for labeling of photographs (e.g., "sunset") according to the clustering of certain color combinations. In the biomedical domain, similar attempts are at a very early stage; Ogiela [38] has published some beginning results in some efforts to automatically generate linguistic descriptions of lesions in the coronary artery and in the urinary tract, based on patterns of stenosis and dilatation that have been detected in these vessels with image processing.

While, as the quote at the beginning of this section indicates, fully-automated indexing of our images may be a distant goal, our intent is to research how far computerized methods can replace or support labor-intensive activities of the human indexer.

4.3.1 Segmentation

The segmentation problem has been recognized as one of the most urgent problems in digital imaging and remains as one of the grand challenges of the field [10]:

> *Medical images often contain anatomical structures that are not clearly captured by the image sensor or that are easily confused with similar adjacent structures... The permanent quest for segmentation tools that can handle such complex images has led to numerous methods solving the segmentation problem for specific organs, imaging sensors or medical problems. However, there is still little progress in the development of generic and robust methods.* [10]

As this statement emphasizes, robust segmentation tools applicable generally to biomedical images are non-existent. Instead, the most reliable and robust tools available are those that have been developed for specific classes of biomedical images. Since there is no a priori knowledge of which specific techniques would prove most fruitful for our images, we have conducted several efforts in parallel. In this section we discuss this work, done both in-house and collaboratively, and provide rationale, work summary, technical background as necessary to motivate the more important issues, and accomplishments. Our segmentation work consists of the following:

- Manual 9-point Segmentation: acquisition of coarse segmentations by medical experts as reference data for validating segmentation algorithms.
- Region Localization: computer-assisted high-level, pre-segmentation understanding of image contents.
- Active Shape Modeling (ASM): segmentation by deformable template created by statistical samples.
- Active Contour Segmentation: segmentation by mathematical energy model.
- Generalized Hough Transform (GHT): segmentation by template matching using Hough bin counting.
- Computer-assisted Dense Manual Segmentation: segmentation by manual point selection, edge detection, and spline curve fitting.

Each of these is described in the subsequent sections, though we concentrate in particular on Active Shape Modeling (ASM), which has come to be the focus of our collaborative segmentation work.

4.3.1.1   Manual 9-point segmentation.   At an early state of our segmentation work, and in parallel with segmentation algorithm development, we obtained segmentation data created by human medical expertise. The purpose of this task was to acquire reference data as a guideline for validating vertebrae segmentation algorithms. These data consisted of (x, y) coordinates for specific geometric locations on the vertebrae; a maximum of 9 points per vertebra were collected. Points 1-6 corresponded to the standard 6 points commonly collected in the field of vertebral morphometry, point 7 was the anterior midpoint of the vertebra; points 8 and 9 were also collected if the vertebra had anterior osteophytes, and corresponded, respectively, to the maximum protrusion of these osteophytes on the anterior top and bottom of the vertebral body. (An osteophyte is a "bony outgrowth or protuberance" [39].) The points collected are illustrated in Figure 2 (a).
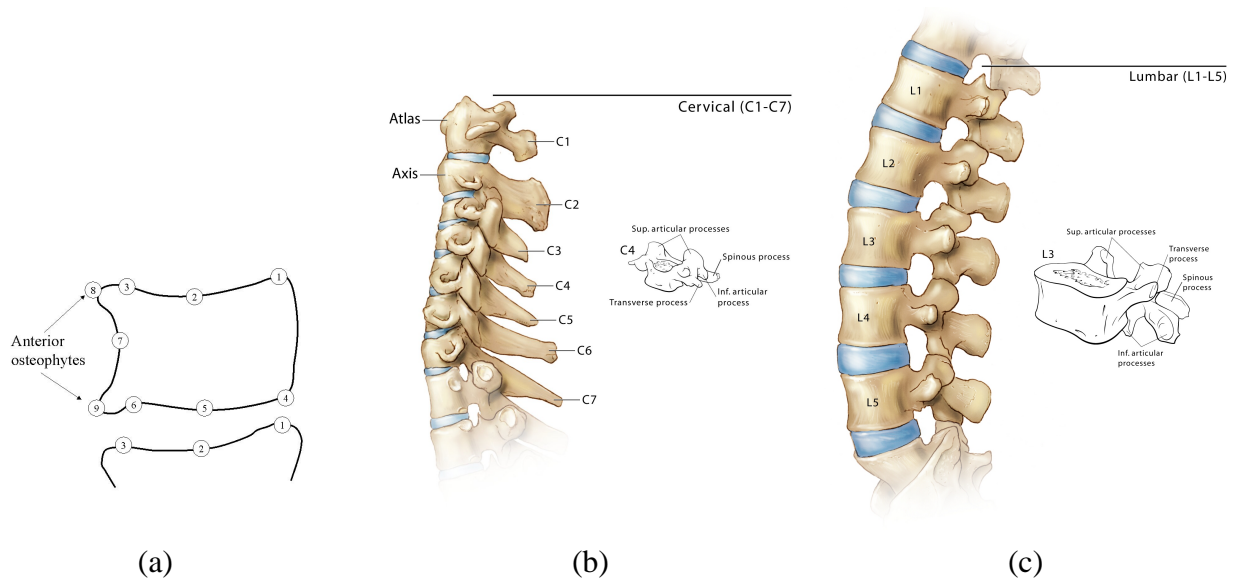
Figure 2: Spine figures: (a) Nine-point vertebral data collection (b) Illustrations of Cervical Spine and (c) Lumbar Spine Anatomy

For this task, we began with the requirements of collecting enough data to reasonably represent significant variance in vertebral shape due to individual body types, collecting data with a significant number of pathological or otherwise abnormal vertebrae, and maintaining biomedical integrity of the data. The trade-offs among these requirements resulted in a decision to collect data on 550 unique x-ray images, 275 cervical spine images and 275 lumbar spine images. The data were collected under supervision of a board-certified radiologist with expertise in interpreting spine x-rays; he was assisted by an engineer with an established history of successful research and development within the biomedical image processing community, who developed specialized software for the data acquisition. In order to maximize the probability of abnormal vertebrae appearing in these images, all of the images used were for people 60 years and more of age. For purposes of future analysis, half of the spine images were chosen from people answering "no" to the "neck pain" question on the NHANES II health questionnaire, and half from those answering "yes".

Only one radiologist was available as reader, so the data collected do not have the significance of a carefully-controlled, multiple expert reader study. However, in order to get some measure of the possible intra-observer variance in point placement, we interspersed 50 duplicate images with the 550 unique images in a blind fashion. The total number of images read, then, was 600.

The data points were collected for all of the vertebrae with sufficiently visible boundaries. Typically, for the cervical spine images, these included the vertebral boundaries from the bottom of C2 through C6, though in a few cases C7 and T1 were also visible. For the lumbar spine boundaries, the collected data typically spanned L1-L5, although in a few cases parts of the boundaries for the thoracic T12 and sacral S1 were collected. Additionally, a few special points were collected: for the cervical spine, these included a point marking the approximate center of gravity of the C1 vertebra. The nomenclature for the vertebrae and the overall spine anatomy are illustrated in Figures 2(b) and (c)[2].

---

[2]Artwork in Fig. 2(b) and (c) by David Blume of Lister Hill Center's APDB and coordinated by Joe Fitzgerald.

The collection of this data has proved to be useful for both internal and external research purposes. Uses to date include the following:

1. Both in-house and external researchers have used it as reference data for assessing the accuracy of coarse-level region segmentation algorithms. Both Long [40] and Zamora [41] have created algorithms to estimate the spine region location and/or orientation: Zamora placed the spine region within oriented rectangular bounding boxes, as described in Section 4.3.1.2.1; Long estimated curves that lie within and that bound the spine region (see Section 4.3.1.2.3). Both used the radiologist points as measures of accuracy of their algorithms. We note that the placement of boundary points on objects within digitized images is subject to inter-reader and intra-reader variances, even when done by subject matter experts. A carefully designed, multiple reader study is required to accurately estimate these variances. The goal of our work has not been to establish point placement and associated variances with highest accuracy, but rather to acquire data of sufficient accuracy to allow us to bootstrap to the next level in our segmentation work: automated or computer-assisted segmentation, rather than purely manual segmentation.
2. The collected data has also been used in-house to develop dimensional data for simple trapezoidal models of the vertebrae: these measurements include anterior, posterior, and medial heights, and spacing between adjacent vertebrae. These geometric values derived from the marked points have been made available as auxiliary data in the WebMIRS NHANES II database, as well as in the CEB CBIR1 Prototype. Note that the geometric measurements are pixel-valued, since we do not have physical measurement data for the x-ray images.

4.3.1.2 Region localization. "Region localization" refers to the estimation of boundaries within the image that enclose objects of interest at a coarse level of precision. Region localization is important for

- Initializing higher-precision algorithms for fine level object segmentation. Such segmentation algorithms as Active Shape Modeling, discussed later in detail, are local optimizers; i.e., they seek the best object segmentation within small neighborhood of their beginning points. Thus, good starting points (approximate position/orientation/scale) of the objects to be segmented are critical.
- Providing methods for assisting human experts in rapid image display and review (independent of its use in initializing a segmentation process). For example, with an algorithm that can automatically, rapidly, and with high probability identify the spine region with a marked rectangle, this region of interest can be automatically zoomed on the display as the reviewer moves from image to image, even though the location and orientation of the spine may vary appreciably in these images.

4.3.1.2.1 Spine region localization with oriented rectangles. One of our collaborators, Zamora, provides an initial estimate of cervical spine location and orientation automatically, with no a priori information [41]. Zamora's algorithm locates the spine area as an oriented rectangle, as shown in Figure 3.

This algorithm assumes that a line passing through the maximum amount of bone structure in the image will lie over a large part of the spine area. Given a line passing through the image, the
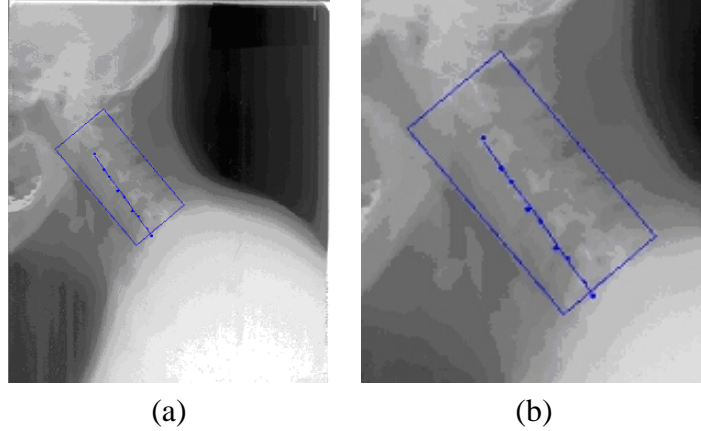
Figure 3: Automatic Estimation of Location of the Cervical Spine (a) Zamora algorithm results. (b) Enlarged region.

algorithm computes a grayscale integral for that line by summing the image grayscale values along the line. The algorithm repeats this process for a number of lines transversing the image at various rotation angles, and determines the maximum line integral value. This maximum line integral defines the line that is taken to estimate the spine orientation. Once the spine orientation is known, the image is oriented with the spine vertical in the image, and the top and bottom boundaries of the spine are estimated by detecting horizontal edges in the oriented image. The spine side boundaries are estimated from a priori estimates of spine width in the images.

4.3.1.2.2 Spine region localization with Generalized Hough Transform. We investigated this as an alternative region localization technique for the purpose of initializing the ASM algorithm. Initial results were so promising that we have continued to explore it as a segmentation method in its own right. The Generalized Hough Transform (GHT) is an algorithm inspired by the standard Hough Transform, which is commonly used to detect straight lines in images. The generalization of this algorithm allows the detection of arbitrary shapes in images. The GHT as applied to the spine x-rays operates on a template of the vertebrae and finds the best match in the image to the template while varying scale, position, and orientation. The matching is effectively an exhaustive search of the entire image, using the Hough "bin counting" method. Because the GHT is not a deformable template method, the resulting match cannot be expected to agree with the vertebrae in fine detail: the matching accuracy is limited by that obtainable by modifying the position, scale, and orientation of the input template. Nevertheless, preliminary results obtained by this technique have shown promising results up to the expected accuracy limits. Because the matching method is an exhaustive search, the method as currently implemented is very time consuming. Current collaborative research is exploring ways to increase the accuracy of the matching and to reduce the computation time to within practical limits.

4.3.1.2.3 Other spine region localization methods. Additional in-house work has focused on methods to fix an anatomy-based coordinate system in the images, to classify gross-level features of skull, shoulder and background, to automatically estimate spine location and orientation with approximating curves, and to semi-automatically identify and label individual vertebrae in the cervical spine. Figure 4 shows results of the spine estimation with approximating curves, for two different cervical spine images. In this Figure, four approximating curves have been automatically
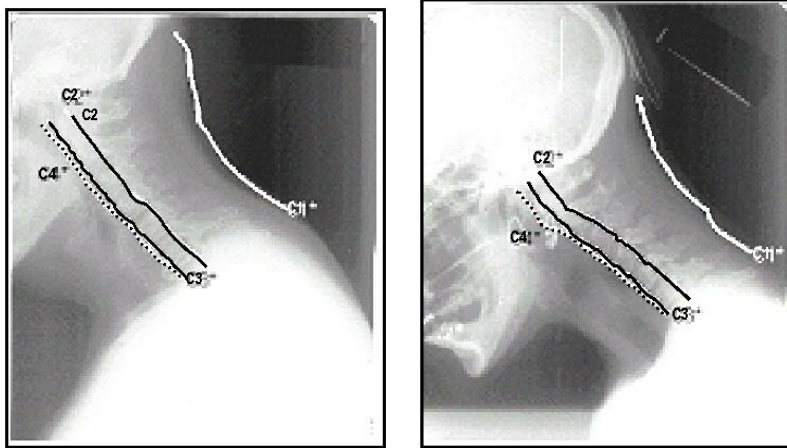
Figure 4: Estimating the Spine Location and Orientation with Approximating Curves: Automatic Process (applied to two different cervical spine images).

generated for each image: one curve (C_4) estimates the tissue boundary at the front of the spine, one curve (C_3) estimates the bone boundary at the front of the spine, one curve (C_2) estimates the curve of the spine which follows the greatest concentration of bone mass, and one curve (C_1) estimates the boundary between the neck area behind the spine and the background. (Note that the curve nomenclature, such as "C_4", should not be confused with the nomenclature for vertebrae in the cervical spine, such as "C4".)

4.3.1.3 Segmentation by Active Shape Modeling. A large part of our segmentation work has concentrated on Active Shape Modeling (ASM), although we have not ignored other approaches. The ASM formulation that we have followed is that described by Cootes and Taylor [42] of the University of Manchester. For CEB segmentation work, ASM has represented a significant advance beyond heuristic, edge detection methods which have yielded very little promise of success in segmenting irregular, noisy images, into the domain of model-based, statistical, deformable template methods. An important development for our work was the publication, described below, of the success of the ASM algorithm for a class of images similar to the spine x-rays.

4.3.1.4 A motivating study. Researchers associated with Cootes carried out successful segmentation on images of the lower thoracic and upper lumbar spine (T7-L4) acquired by dual x-ray absorptiometry (DXA), reported by Smyth [43] in 1999 in the journal Radiology [3]. Smyth used a single shape template to model the 10 T7-L4 vertebrae in a collection of 78 DXA images acquired from females aged 44-80. In order to run the ASM algorithm, the user was required to manually initialize the ASM search by anchoring the template to the image with three manually-placed points, after which the algorithm would deform and move the template to seek to the vertebrae.

To get a measure of the accuracy of the ASM segmentation, Smyth manually marked 73 landmark points on each vertebra in every image in consultation with a radiologist, and compared the converged ASM vertebra boundaries with the manually-marked boundaries. Further, to estimate how these errors would compare with reproducibility errors occurring when acquiring boundary markings manually with multiple human readers, Smyth had four readers independently mark six

---

[3]Thanks to Dr. Carl Jaffe of Yale for providing us with this significant article.

boundary points on each vertebra in each of the images, and computed variances in point placement across these readers. Significant results that he reported, as a function of vertebrae T7-L4, were (1) ASM successfully converged for 94% of the L3 vertebra with an increasing rate of success up to 99.2% convergence for T7; (2) the maximum error observed between the manually marked boundaries and the ASM boundaries was 1.61 pixels, for the L4 vertebra; (3) for all but two of the vertebra, the ASM/manually-marked error observed was less than the reproducibility error of the four human readers (exceptions were L1, where the ASM/manually-marked error was 1.17, and the four-reader error was 1.12; and L4, where the respective errors were 1.61 and 1.47).

Smyth provided the most complete published results of successful semi-automated segmentation we had seen for spine images acquired with a modality similar (although likely of better contrast quality) to our x-ray images. (Additionally, Gardner [44] reported that spine x-ray segmentations had been obtained at a useful level by using deformable models in an interactive system for digitized lumbar spine images.) Using the ASM Toolkit, an add-on to the MATLAB mathematical software system, that is developed by Visual Automation, Ltd., and which incorporates the ASM algorithm, we conducted research on the effectiveness of ASM for segmenting our x-ray images. The results, described below, were sufficiently promising to justify additional work to carry out a comprehensive research effort for the evaluation and potential use of ASM in segmenting the x-ray images.

Because of the significance that ASM has assumed in our work, we present a technical overview of the algorithm theory and operation here.

4.3.1.5 The ASM algorithm. The ASM algorithm, as described in [42], consists of three ingredients: (1) *a template*, which is the "average shape" of an object that we want to search for in an image; a template is composed of a discrete number of "landmark points" that define the shape; (2) *a shape model*, that encapsulates information that constrains how the template is allowed to deform; and (3) *a grayscale model*, that encapsulates information about the expected image pixel values (grayscale values) at each of the landmark points. The ASM template is created by collecting landmark points for multiple instances of the object of interest. For the cervical spine C2 vertebra, for example, this would mean collecting landmark (boundary) points for C2. Each collected set of landmark points defines one "sample shape vector". Writing each sample shape vector as a column vector and organizing the collection of such vectors as a matrix yields the "sample shape matrix". The shape model is computed by linear algebra and statistical operations on the total set of sample shape vectors (computing mean and covariance of the sample shape matrix, and eigenvector/eigenvalue pairs of the covariance matrix), and provides a method of describing in precise mathematical terms:

- the manner in which the sample shapes vary from the mean shape; the various departures from the mean are defined by the eigenvectors (or "modes of variation") of the shape matrix covariance; for example, deformation of the lower anterior corner of a vertebrae, caused by the presence of an osteophyte there, would be expected to be represented as a "mode of variation" in the shape model, provided that our sample shapes include both normal vertebrae and those with lower anterior osteophytes;
- the significance of a mode of variation from the mean shape; each mode of variation is an eigenvector of the shape matrix covariance, and has an associated eigenvalue; the magnitude
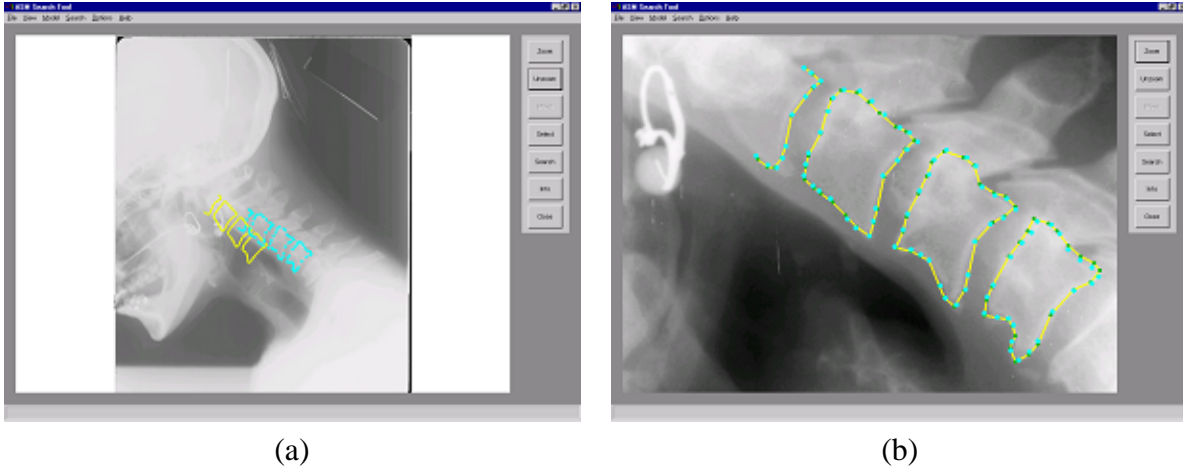
16

Figure 5: In-house work with ASM: (a)Placing the template for ASM segmentation (Truth Segmentation is shown in background) (b) The converged segmentation.

> of this eigenvalue, relative to the sum of all eigenvalues for all modes of variation, is a measure of the proportion of total variance from the mean shape represented by this mode of variation; this allows us to find those modes of variation that explain (for example) 95% of the total variance from the mean shape; keeping only these modes of variation in the shape model allows an economy of calculation when the algorithm is executed.

4.3.1.6 ASM algorithm operation. To operate ASM, the template is placed within an image (either randomly, or intelligently, if a priori information is available), with the goal of converging to a target object by rotating, translating, scaling, and deforming. The template initially becomes the "current shape". ASM is an iterative process where, at each step: (1) grayscale values from the image are sampled at lines normal to each landmark point on the current shape; these sampled values are compared to the expected grayscale values at each landmark point, and the landmark point is then replaced with the sample point that most closely corresponds to the expected value (using Mahalonobis distance as a measure of closeness); (2) the shape model is used to constrain the shape produced by the grayscale model to lie within reasonable distance of the shapes represented by the shape model; and (3) an updated position, orientation, and scale ("pose") is estimated.

When the step-to-step differences between the current shape and the previous shape become sufficiently small, the algorithm has converged. Figures 5(a) and (b), both results of in-house work, illustrate the process of placing an ASM template on a target image and show the resulting, converged segmentation.

4.3.1.7 ASM research to date *Initial ASM evaluation on the x-ray images.* Collaborative ASM work that has been carried out at the Computer Vision Laboratory at Texas Tech University has included an experimental evaluation of ASM on a set of cervical spine images, analysis of algorithm failure modes and development of strategies to address these modes, and development of software to efficiently collect landmark points on the boundaries of vertebrae in x-ray images, with computer assistance. The results of this phase were presented in a technical report [45] to CEB, along with the associated data, including manually-derived vertebra landmarks used for template building and converged vertebra boundaries produced by ASM. The ASM implementation available in
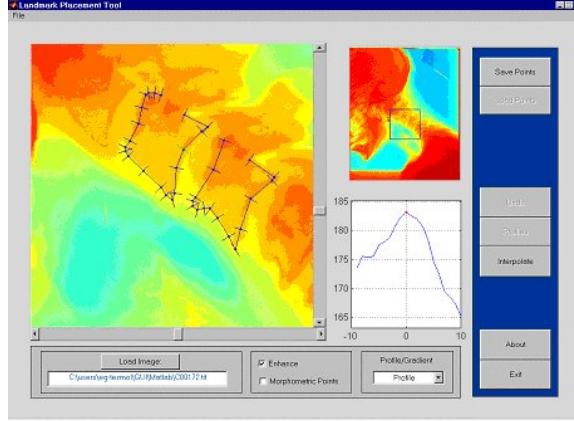
17

Figure 6: Landmark Tool

the MATLAB ASM Toolkit was used for the work. The test image set consisted of 40 NHANES II cervical spine images; in each image vertebrae C2-C7 was manually segmented using approximately 15 points per vertebra. (The ASM Toolkit software imposes a constraint on the number of total landmark points that may be used in a template, hence the somewhat sparse segmentation.)

Development of the Landmark Tool software. From preliminary in-house experience with the algorithm [46], it was recognized that proper collection procedures for the landmark data would be critical to the success of the algorithm, and this issue was addressed early in the work, by development of the Landmark Tool. The Landmark Tool allows the placement of landmark points more accurately on mathematical edges of the object, rather than relying completely on the human observer's subjective judgment of edge location. The Landmark Tool is shown in Figure 6. The tool allows the user to mark a candidate landmark point, then to get a display of information about the pixel values in the neighborhood of that point. (Specifically, the user may display plots of the pixel grayscale values, or slopes of these values, along short line segments perpendicular to the estimated boundary of the vertebra.) These plots allow the user to intelligently reposition landmark points more closely to the mathematical edge points and reduce the "observer/algorithm data gap" in creating the ASM template.

Development of techniques for ASM initialization. A second problem addressed in this work was the anticipated requirement for an automated method for initializing the ASM algorithm with an estimate of the location and orientation of the C2-C7 vertebrae. ASM is a local optimization algorithm: it searches for the object location within relatively small neighborhoods of the current shape. (Even though the algorithm incorporates a multiscale approach, searching first a coarse version of the image, then as it converges on an object in that level, searching a finer and finer level, it still shows failure to converge when initialized too far from the actual object location and orientation.) The method by Zamora, discussed in Section 4.3.1.2, was developed to address this problem, as is the Generalized Hough Transform method, also discussed in that section.

Extended ASM R&D. The extended ASM work currently under way includes extending the capabilities of "classic" ASM as formulated by Cootes to potentially yield better performance for the x-ray images, as measured by (1) capability to initialize with small or no human intervention; (2) capability to reduce or avoid cases of convergence to spurious solutions due to "vertebrae shift"[4];

---

[4]"Vertebrae shift" occurs in those cases where we seek to segment cspine vertebrae C2-C6, for example, but ASM

and (3) and capability to segment with sufficient accuracy to be able to classify the vertebrae for biomedical conditions of interest, for example, presence of anterior osteophytes. Toward this goal, a hybrid algorithm is being studied that (1) incorporates the Generalized Hough Transform for initialization, (2) contains a modified form of ASM to incorporate the use of edge information, and (3) includes a "deformable model" (DM) algorithm based on the work of Tagare (see the following section) to do fine-level segmentation on high-interest parts of the vertebrae (in particular, the corners).

4.3.1.8 Active Contour Segmentation. Also known as *snakes*, ACS is another segmentation technique that we researched in collaboration with Dr. Hemant Tagare of Yale. We pursued this area with the prospect of obtaining a vertebrae segmentation tool within a relatively short time, since ACS is a reasonably simple implementation (unlike the ASM algorithm, for example).

Beginning with a prior shape to be matched to image data, and an a priori position and orientation of the shape, the active contour method deforms the shape by minimizing an *objective function* that is constructed to be a function of the total shape curve. For our work the objective function was expressed as the sum of two terms: the first term is the sum of slopes (gradients) of image grayscale values along normals to the shape curve; the second term is the weighted length of the shape curve. Appropriate signs are given to these two terms so that minimizing the objective function achieves a balance between maximizing the sum of slopes along the shape curve and minimizing the length of the shape curve itself. The rationale for this formulation is as follows. Vertebra boundary points are expected to satisfy the following mathematical property: the image grayscale values at a boundary point will have the steepest slope in the direction perpendicular to the boundary, within a local neighborhood of the boundary point. This accounts for the first term in the objective function. However, if this is allowed to be the only term in the objective function, the algorithm tends to maximize the sum of slopes by deforming the shape curve into a zig-zag curve, where the sum of slopes has been maximized by enlarging the length of the curve and producing an unrealistic vertebral shape. The second term, then, is added as a constraint on the length of the solution shape curve. A weighting term allows an empirically- determined balance between the two terms to be set: at present, the selection of the value for the weight is based on informal observation of the results across a range of trial values.

The process begins by the user selecting a region of interest (ROI) that contains the object (vertebra) to be segmented. Within this ROI, the user places a template (the a priori or beginning shape) near the vertebra; the user may rotate and scale the template to achieve the best approximate match of the template shape to the vertebral shape. After template placement, the user starts the active contour algorithm. The algorithm creates a grid of points in a band of the image that is centered about the template, and searches for the shape curve that yields the global minimum (within the band) of the objective function. Rather than doing an exhaustive computation of the objective function for every possible shape curve within the band, the algorithm achieves the same goal by incorporating dynamic programming. Dynamic programming is an optimization method well-known in the engineering/computing community [47, 48] and is appropriate for optimization problems such as this one, where the objective function may be evaluated point by point along a candidate solution shape curve, and the evaluation at any point depends only on the evaluation at

gives us a segmentation of C3-C7, instead.

the previous point.

The ACS algorithm has been incorporated into the segmentation tool used in our prototype CBIR system.

4.3.1.9   Computer-assisted dense manual segmentation.   For the feature classification work described in Section 4.3.4, vertebral boundaries were generated by our collaborator Dr. Joe Stanley of the University of Missouri by a combination of manual and computer processing. In this approach, first, the original image is processed by using the Kirsch edge detection operator. Next, a human operator manually marks vertebral boundary points using the Kirsch edge areas as guides; then, spline curves are fit through the manually marked points and, finally, these spline curves are resampled at equal intervals of arc length to obtain the final boundary points. The importance of this method is that it is a practical way to obtain usable segmentations for a sufficient number of vertebrae to support the feature classification work.

4.3.2   Feature extraction

The features extracted at this stage define and describe the image content. Image content to the user is often at the semantic level. For example, the user may be seeking images with cases of "vertebral fracture", "disc compression", "Anterior Osteophytes of grade 3", etc. Mapping such a high level query to suitable image characteristics is a challenging task. However, a middle ground is achievable by identifying appropriate image features and representing them in a form to enable responses to similar queries by defining areas in the segmented structures and assigning meaning to the measurements on those. It is thus important to identify the image features to be extracted to enable responses to queries relevant to the class of images in question. The cervical and lumbar spine x-rays images are grayscale images with little or no useful texture information in them. This eliminates two significant (robust, and popular) indexing features, viz. color and texture.

Shape based image matching techniques have been used in limited ways and are still not mature enough to provide a high degree of automated semantic labeling. Additionally, there has been little study on forming shape-based indices for any large set of images. Most shape databases have been applied to trademark databases where global shape features and possibly color and texture, are used.

What is shape representation? Boundary data is extracted as (x, y) coordinates in the image space and needs to be represented in a form suitable for archiving, indexing, and similarity matching. A shape representation method converts a dense 2D representation of a boundary, i.e., the (x, y) coordinates of boundary points, into a form that has the following properties [49, 50, 51]:

- *Invariance*: two boundaries that have the same shape should have the same representation.
- *Uniqueness*: two boundaries with different boundaries should have different representations.
- *Stability*: a small change in the boundary should cause a small change in its representation.
- *Efficiency*: the representation should be computationally efficient to compute and store.
- *Ease of implementation*: the representation should be the least complex of competing methods to implement.
- *Computation of shape properties*: the representation should support computation of shape properties such as symmetry, area, perimeter, etc.

20

- *Meaningful representation*: The representation should retain properties of the shape that are meaningful to the application.

These requirements may be extended to include matching of partial boundaries or specific local regions in the boundary. In addition to the shape representation properties identified above, other properties more relevant to retrieval are described in [51].

- *Geometric Invariance* (also known as Similarity Invariance): the representation should be invariant to rotation, translation, and scaling.
- *Compact representation*: essential shape information should be maintained in a significantly reduced manner.
- *Fast matching speed*: efficient shape feature computation and fast retrieval in a large image database.
- *High quality image retrieval*: retrieved images should be relevant to the inquirer.

In light of these properties, we investigated suitable shape representation and similarity methods in the literature and evaluated them for applicability to the vertebral shapes. The goals of this study were:

- To evaluate several candidate shape description algorithms for efficient indexing and retrieval of vertebra boundary shapes.
- To determine the effect of number of boundary points on shape description. In particular, to determine the critical number of boundary points needed for describing the global vertebral shape while retaining important characteristics such as Anterior Osteophytes.
- To determine if a vertebral boundary shape needs to be divided into segments (possibly overlapping) for queries on localized boundary segments of interest.

Several approaches to shape content-based image retrieval are seen in the literature, though most have been applied to non-medical images. A brief survey on these methods applicable to our images are presented in Appendix A. The process of identifying relevant and applicable shape methods, criteria for selecting the methods for evaluation, and the results are presented below.

4.3.2.1   Criteria for representing vertebral shapes.   We observe that most shape representation methods published in the literature use the global shape characteristics for indexing; i.e. the final shape representation is controlled by the distribution of all the boundary points in the image space. Such an approach may not be suitable for biomedical shapes, such as vertebral shapes, for the following reasons:

- **Similarity in Anatomical Shapes:** While there may be great disparity in shape boundaries between objects from different anatomy classes, there is high intra-class similarity. In our case, we generally observe high similarity in the outlines of the sagittal view of vertebrae from different regions on the spine. Shape methods that depend on global shape characteristics are likely to produce very similar representations for these shapes, making subtle discrimination between the shapes difficult. Desirable methods would need to include *local* shape features to discriminate between various vertebrae across the images. Techniques could include varying the density of representative points at various regions on the object boundary based on the insignificance and/or discriminative power of those regions.

21

- **Boundary Representation:** In biomedical images, minor differences in outlines of the anatomy could be significant. Shape representation methods tend to reduce the number of points in a shape outline in order to avoid excessive dimensionality. So, it is possible that these differences may not be accurately retained in the indexed shape representation. Effective methods may need to represent the boundary as separate segments while maintaining relationships between these segments.
- **Localized Queries:** Very often the queries may be made on local portions of the vertebral shapes, such as anterior osteophyte, or curvature in the top and bottom of the vertebrae. The practical use of this system is likely to concentrate on specific features on the vertebra(e) and it is unclear if conventional shape indexing methods are able to support such localized queries. The approach suggested above may be helpful in querying local portions of the contour.
- **Edge interpretation:** Far and near edges appear very often in biomedical images. How can such complex shapes be represented? Should bifurcation be treated as a separate image structure? Should there be a relationship graph between the edges? While technical solutions can be developed to address this issue, its biomedical significance needs investigation.

4.3.2.2 Evaluation of shape representation and similarity methods. We have evaluated select shape representation methods from the literature to determine if they are suitable for application to biomedical images. Most shape-based CBIR methods, to date, have been applied to trade-mark databases, fish images, silhouettes of tools, etc. For example, Jain and Vailaya [52] have evaluated shape retrieval methods applied to trademark image databases.

The methods published in the literature can be grouped based on their characteristics. For example, methods that employ intrinsic shape characteristics such as centroid, eccentricity, elongation, invariant moments, etc., methods that use polygonal approximations of object contours, methods that use deformable templates, and those based on frequency operations such as Fourier Transform or Wavelets.

4.3.2.2.1 Shape representation methods selected for evaluation. We implemented one method in each category: global shape properties, polygon approximation, scale space filtering, and Fourier descriptors. Global shape properties such as size, perimeter, convex perimeter, elongation, roughness, and compactness, etc., can be used for measuring similarity and are covered in [53]. Invariant moments have also been used for discriminating shapes [54]. Multi-stage modification using invariant moments has shown very good results [52]. Other methods using higher order moments include generalized complex moments [55] and Zernike moments [56]. Multi-scale shape representation has been used to smooth and simplify the contours [57, 58, 50]. Here, the curvature function was then used to analyze the smoothed curve in order to determine the critical points on the curve. Another polygon curve representation method is done in the tangent space, also called "turn function". This method uses curve evolution to remove small variations and less significant features and then represents the curve in tangent space [59, 60, 61]. Shapes or contour points have also been described in the frequency domain [62, 63]. With a proper representation of the contour points of a polygon, Fourier descriptors have been used to measure shape similarity that is insensitive to translation, rotation, and scaling [64]. The similarity function has to be insensitive to starting point of the polygon curve as this may affect results for methods such as polygon representation where results depend on the order in which points are processed. In our research,

	Comparison to	
Method	to self	to altered
Invariant Moments	100%	100%
Turn Function	100%	89%
Token Description	100%	90%
Fourier Descriptor	100%	100%

Table 4: Results of evaluation from shape representation and similarity testing on 40 shapes.

a method in each of these categories was implemented. A shape similarity measure was designed for each method so that the input shape can be compared against the shapes in the database and the similarity scores can be obtained to evaluate the performance of each method. The selected methods and similarity measures are described in greater detail in Appendix A.

4.3.2.3 Evaluation strategy and results 20 sets of boundary points were used for testing. The same 20 sets of data were rotated by 20 degrees, rescaled by 1.2, and starting point shifted by 40 boundary points to create another 20 sets of shape data. This amounted to a total test data of 40 sets. Testing was done in two parts. First, the shape was compared against itself and then compared against the altered counterpart of itself to test the robustness of geometrical invariance. Secondly, the retrieval resulting was compared between similar shapes and shapes that are very distinct. Table 4 shows the statistics result from an initial evaluation on all four methods.

The turn function method has 44 scores out of 400 that are lower than the score when compared against its altered counterpart. Token description missed 40 of them and has 90% accuracy. The invariant moments and Fourier descriptors did not miss any.

The second test was done by dividing shapes into two groups. One group has shapes that are highly similar and the other one has shapes that are completely different. Only the invariant moments and Fourier Descriptor methods were compared in the second test because the other methods already failed the first test. The results from this test revealed that the invariant moments worked better with shapes that are significantly different whereas the Fourier descriptors worked better with the shapes with subtle variations. In conclusion, we note that for vertebral shapes:

- Global shape properties and invariant moments work effectively distinguish shapes with significant differences.
- Curve evolution works better than scale space filtering in smoothing contour noise.
- Curvature function analysis is a good tool for analyzing polygon curves.
- Matching two turn functions and matching tokens are computationally intensive tasks.
- Fourier Descriptor method provides better geometrical invariant properties for matching and calculating the similarity measure.
- While an overall global match is achieved, category based and local feature based matching is challenging without significant modifications to these methods.

4.3.3 Feature vector organization

This section presents the candidate methods found in the literature that may be applicable to the CBIR system being developed. We are currently evaluating their suitability for implementation with our images and extracted feature vectors.

Here the objective is to organize the features extracted from the images so that the similarity computation can be efficiently performed for retrieval of relevant images. That is, the structure should allow pruning of a large image database with minimal computation into a small set of candidate image features on which detailed comparisons could be made.The types of structures seen in the literature are based on two criteria; (a) type of image features, and (b) types of queries supported. The type of image features would determine the comparison strategy and the organization of each vector so as to minimize computation.

In order to determine the feature vector structure suitable for queries supported by the CBIR system, it is necessary to understand the various types of queries that can be posed. Unlike traditional databases, where the match is usually exact even if it falls within a specified range, multimedia databases usually have to deal with approximation and fuzziness in their matches. Content-based searches can be classified into three types, the *target search* where a specific image is sought, the *category search* where one or more images from a category are sought, and *open-ended browsing* in which the user seeks an image by specifying visually important properties [65]. Additionally, the queries tend to have semantic connotations making organization of data indices very challenging. The efficiency of the database system is achieved by minimizing the search space within which the similarity function can be executed.

We identify the following approaches in organizing such data [66]:

- point access methods (PAMs) such as k-D trees and grid files,
- spatial access methods (SAMs) such as N-trees, linear quadtrees, z-ordering, R-trees,
- vector space clustering such as generalized k-nearest neighbor (k-NN), agglomerative cluster trees, etc.

PAMs. Each record, a k dimensional feature vector, is viewed as a point in a k-dimensional space. The PAMs access the points by their address. Two methods are proposed for such access, the grid file and the k-D tree. A grid file can be viewed as a generalization of extendible hashing in multiple dimensions. A grid is imposed on the address space where the grid size adapts to the data density in a particular area. The access is thus reduced to an addressing problem. A k-D tree, divides the address space in disjoint regions, through "cuts" on alternating dimensions/attributes [66]. It is structurally a binary tree with a record on every node, however a different discriminator is used at each level. The k-D tree is reported to easily handle exact-match queries, range queries and nearest-neighbor queries.

SAMs. Linear quadtrees tries to impose a linear address space to multidimensional data. The process takes its approach from quadtree structures for 2D images and extends it to higher dimensions. The z-ordering and other approaches are techniques for efficiently visiting the cells in this structure. R-Trees [26] can be thought of as a B-tree extended for multidimensional objects. In this structure, a spatial object is represented by its minimum bounding rectangle (MBR). Nonleaf nodes contain pointers and the MBR for all the nodes below it. Each leaf node contains the object identifier and its MBR. It is deemed to be one of the most successful spatial access methods for target searches, range queries, nearest neighbor type queries, insertions, splits, spatial joins, etc. Its variations, the R^*-trees and the Hilbert R-trees achieve improved performance and space utilization for more complex algorithms.

Clustering. This technique follows from pattern classification literature. The k-dimensional feature vector is considered as a point in space and all points "near" it are clustered together. The clusters then form super-clusters and agglomerative clustering techniques are used to retrieve the relevant feature vectors. A variant of this approach has been used in [23]. Although clustering techniques are efficient, the high dimensionality of the feature vector may cause loss in search efficiency for normally distributed data. This problem has been referred to the *curse of dimensionality*. A variety of clustering techniques can be found in [67].

In general, similarity search structures need to be designed so that performance and accuracy can be easily traded. The problem with tree structures when applied to high dimensional data is that the search engine has to search a significant portion of the data before it can arrive at a decision. Kurniawati et. al [68] propose an improved similarity search tree structure, called SS^+-Tree, for searches in a high-dimensional feature space. They employ the widely used k-means clustering algorithm. A clustering of the nodes at a particular level in the search tree is done and the centroid of this cluster is used to determine if a new level should be created. Pun and Squire [69] use Correspondence Analysis and the χ^2 metric to determine the independence between the features and form a suitable index structure to use this independence effectively. White and Jain [70] propose a variant of the k-D tree called the VAM k-D tree. This is so named because its split orientation is based on its *variance* and the split position is *approximately* the *median*. The authors also describe a method to construct an R-tree based on the VAM $k - d$ tree called the VAMSplit R-tree.

We are currently studying these methods for applicability to the feature vectors formed from our images. At first glance, any of these methods is likely to provide a solution our problem. However, efficiency in accessing and updating the database is likely to determine, after an empirical evaluation, the method of choice for our data.

4.3.4   Feature classification

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

We have conducted research in feature classification in collaboration with Dr. Joe Stanley of the University of Missouri. The biomedical features investigated are among those features specifically identified as of interest by two workshops convened by the NIAMS. In work done to date, the following features have been investigated: anterior osteophytes of the cervical spine; anterior osteophytes of the lumbar spine; disc space narrowing of the cervical spine; disc space narrowing of the lumbar spine;

Anterior osteophytes: The approach to the work was as follows: obtain "truth" classifications for spine vertebrae from medical experts; obtain segmentations for spine vertebrae; classify the vertebrae using automated techniques; compare the automated results against the expert "truth".

For all of the required segmentations, the method used was that described in Section 4.3.1.9. All classifications of the vertebrae were carried out using artificial neural network technology.

For the anterior osteophytes of the cervical spine, a total of 704 vertebrae were used, with 352
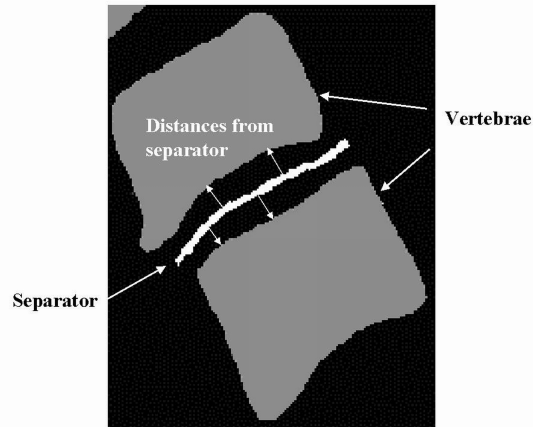
Figure 7: Vertebrae Separator Created for Measuring Disc Spacing

"truth" classified as normal and 352, as abnormal, by a board-certified radiologist. For a vertebra shape, 32 features were derived for use in classifying the shape as normal/abnormal for anterior osteophytes. These features included radius of curvature and gradient measures along the shape, and a mathematical morphology feature that measures how much the shape protrudes from its average local neighborhood. The vertebrae data was divided into training, validation and test sets, following standard practice in development of neural network classifiers: the weights in the network were iteratively modified to make the network classifications approach the "truth" classifications for the training data; this training phase was terminated when the classification on the validation data reached a minimum; then the network was used to classify the test data, and the network performance was taken as the rate of correct classification achieved on this data. On the test data, an overall agreement score of 85% was achieved, as compared to the given "truth".

Similarly, for the anterior osteophytes of the lumbar spine, a total of 782 vertebrae were used, with 391 "truth" classified as normal and 391, as abnormal, by the same board-certified radiologist. The same feature set and test procedure was used for the lumbar spine as for the cervical spine. An overall agreement score of 71% was achieved, as compared to the given "truth" for the lumbar spine. It was noted that the poorer performance of the network for the lumbar spine was perhaps due in part to the lower contrast of these images and the resulting ambiguities in segmentation.

Disc space narrowing: In order to investigate classification techniques for disc space narrowing, a reference set of "truth" readings was obtained from collaborating radiologists at Phelps County Medical Center in Rolla, Missouri. Using the CEB-developed Digital Atlas of the Cervical and Lumbar Spine as a standard, 50 cervical spine images were interpreted for disc space narrowing. Specifically, levels C3/4, C4/5, C5/6, and C6/7 were interpreted in each of these images. Two radiologists carried out the interpretation independently. Similarly, 50 lumbar images were interpreted for narrowing, at L3/4,L4/5,and L5/S1, by one radiologist.

An algorithm for assessing disc spacing was developed. The algorithm operates on an image region containing two adjacent, segmented vertebrae (and the space between them). The algorithm computes a "vertebrae separator", a curve with points lying equidistant between the adjacent vertebrae boundaries; for each point on this separator, the "distance to a vertebra" is taken to be the Euclidian distance to the closest neighboring point on one of the vertebra. Figure 7 illustrates the separator created by this algorithm.

Using statistics of these pointwise distances, the "disc space distance" may be characterized. The mean, standard deviation, minimum, and maximum of the pointwise distances have been used in the work to date. One result that has been obtained is the performance of an experimental classifier using the minimum distance measurement described above. In a trial data set of 159 vertebrae, a correct classification rate of 86% was achieved. The classification work has resulted in published and pending technical articles [71, 72]. The capability of computer systems to assist in the classification of image information into biomedical feature categories is expected to increase in importance as algorithm intelligence, computing power, and volume of image data increase. This research is currently in progress.

4.4 Retrieval

This section describes concepts and work accomplished for the retrieval side of CBIR systems. We introduce the concepts of retrieval, then illustrate how each of these concepts was implemented in a prototype CBIR system.

4.4.1 User query formulation

User queries to a CBIR system are usually formulated as *query by example* or *query by sketch*. The general system that we aim for supports both these type of queries, as well as standard text queries, and hybrid queries using both text and image/sketch inputs. We focus, though, on the novel query types of image example and sketch. For these, the user interface must support (1) selection of an example image and capability for the user to localize the object of interest within that image; this may be accomplished by user interaction, for example, to draw a rectangle around the object of interest, and then to invoke a (possibly interactive) segmentation capability; the result is shape boundary for the example object; (2) capability for the user to employ drawing tools to create the boundary shape for the object of interest. We note that an example image indicated by the user may be chosen from the image database itself, in which case, shape information is already available without the need for further segmentation and feature extraction.

4.4.2 User query feature extraction

Once the boundary shape has been determined, by either segmentation of an example image, or by sketch, features are derived from the shape in exactly the same manner as deriving features from the shapes that populate the database. The result is that, just as for the shapes in the database, the input shape is described by an N-dimensional "feature vector", where N is the number of individual features.

4.4.3 Similarity Matching Method

The measure of similarity of two shapes is the mathematical distance between the feature vectors for the two shapes. Distance may be defined as ordinary Euclidean distance or the L_2-norm as shown in Equation 1, where D is the distance between shapes S_a and S_b, and a_i and b_i are feature vector elements. The feature vectors are of length n.

$$\|d\|_2 = | S_a - S_b |^2 = \sqrt{\sum_{i=1}^{n} | a_i - b_i |^2} \tag{1}$$

The distance may also be a mathematical metric more suited to the particular type of feature vectors being used. The most common type of query is expected to have this basic formulation: given an example shape S_0, find all shapes S in the database within a distance D of S_0; i.e., find all S, such that $d(S_0, S) < D$, where d is the distance measure being used. Other metrics used include the distance measures such as the Procrustes distance, Bookstein test, Kendall test [73], Hausdorff distance, χ^2-metric, Kolomogorov-Smirnov distance, the Mahalanobis distance [19], etc. The clustering methods commonly used are k-means clustering, fuzzy k-means, Minimum Spanning Tree (MST), etc. Details on these can be found in a pattern classification text such as [67]. A review of statistical pattern recognition and clustering and classification techniques is covered in [74].

Hence, CBIR queries are inherently of a different "species" than relational DBMS queries, and biomedical information systems of the future have to contend with this new query domain.

4.4.4 Query search space strategy

In order to efficiently carry out the spatial domain searches required by CBIR queries, careful attention to data organization is required as the size of the database increases. Considerable research has been carried out in the computer science community to develop data organizations, usually in the form of tree structures, that reduce the expected amount of searching required to satisfy a particular spatial query. Among the proposed (and implemented) structures are R-trees, (and variants R^+, R*) KD-trees (and variants KDB, SKD), BD-trees, and others. These approaches usually allow an input N-dimensional vector to be compared to the high level nodes in the tree; as long as a node is beyond a threshold distance from the input vector, none of the points subordinate to that node need to be searched. The idea is to allow efficient homing in on the region of interest in query space.

4.4.5 The CBIR1 prototype: CBIR on 9-point vertebral shapes

As an initial R&D testbed we implemented a small CBIR system (CBIR1) with initial, coarse-level shape retrieval capability. The main characteristics of this system were as follows: there were 118 images in the database; each image had 4-5 vertebral shapes that were defined by the radiologist 9-point boundary marks; the text data in the system consisted of a subset of the NHANES II health survey data for the subjects of these images. We discuss here each of the retrieval factors introduced above.

4.4.5.1 User query formulation. Both query by image example and query by shape were supported. The example images were restricted to those already in the database. No segmentation capability was provided: when query by image example was used, the associated 9 point shape from the database was taken as the query shape. When query by sketch was done, the user was allowed to define a query shape by positioning 9 points with a mouse.

Figure 8: 9-point System: (a) Main Query Screen (b) Main Results Screen (c) Query by Sketch Results.

4.4.5.2 User query feature extraction The "features" stored in the database for the 9-point shapes were simply the 9 points themselves. Hence, no special feature extraction capability was provided.

4.4.5.3 Similarity matching method. The similarity matching distance used was the sum of squared point-to-point Euclidean distance between two shapes, after the shapes had been translated, rotated, and scaled to produce the minimum such distance, frequently called "Procrustes distance". In other words, the two shapes were put into the closest match by translation, rotation, and scaling, then the distance was computed as described above.

4.4.5.4 Query search space strategy. No special spatial data organization was created. For the testbed, each query resulted in an exhaustive, linear search of the entire database. Image, text, and hybrid queries were supported by the system.

Figure 8(a) shows the main query screen of the system on which the options for specifying text as well as the initial options for specifying image queries were made. Figure 8(b) shows the main results screen; in the case of image query or hybrid query, the returned image results are given in the order of similarity to the input image. Figure 8(c) shows an example of the results obtained through query by sketch with this system.

5 Content-based image retrieval prototype (CBIR2)

In this section we describe the design architecture of our current prototype system, CBIR2. Our goal is creation of an advanced query capability that supports query by image example and/or image sketch for accessing the indexed images while enabling combined retrieval of associated information in the text database. Envisioning a combined biomedical image and text information retrieval system is a natural progression for us. As described in Section 1.1, we have already developed and deployed the WebMIRS system which allows retrieval of text and images by queries to the NHANES text data, but does not use image features for retrieval. Additionally, we developed

a prototype CBIR1 system for a small subset of NHANES II x-ray images as an initial test of data retrieval by any combination of health survey text data, vertebral dimensional data, and vertebral shape (discussed in Section 4.4.5).

5.1 System overview

The prototype CBIR2 system is composed of two separate systems, viz., the indexing system and the retrieval system. The indexing system includes methods for automated image segmentation, image feature extraction, image feature organization, and data organization structure for the text data associated with the images. The retrieval system provides the interface and the methods for retrieving the image and text data through combined queries. It also includes methods for determining similarity between features extracted from the query visual and those stored in the database. These systems are explained in greater detail in Sections 5.3 and 5.4.

The system is designed to enable content-based retrieval of objects of interest that have been segmented from the images. Additionally, it is designed to support combined text and image queries on the NHANES dataset. The queries could be by image example, which could eventually include user submissions, or by user-sketch in which the user sketches the object, which could eventually include queries to portions of the object of interest. In order to achieve this, it is necessary to index the image and text data. The text data can be indexed in a traditional relational database management system, such as MySQL, Oracle, or Informix, and queried using SQL. An example of this is the WebMIRS system.

Image features are organized to enable similarity based retrieval. In our system, it is necessary to segment the vertebrae from the digitized cervical and lumbar x-ray images. Presence of features such as anterior osteophytes, disc space narrowing, subluxation, spondylolisthesis are determined through computation of appropriate measurements. The next step is to index the vertebra subimages using the boundary data. In order to determine similarity between vertebrae, it is beneficial if the number of points defining the boundary are reduced to a few essential landmark points that uniquely identify the shape and significant pathology. Content-based retrieval is performed by defining a similarity metric and organizing the vertebra boundary information and associated features for efficient retrieval.

5.2 System data

The image quality in the spine x-ray images is fairly poor with ambiguous vertebral boundaries, making a reliable segmentation a challenging task. Current implementation contains human assisted boundary segmentation using the Active Contour Segmentation Tool, shown in Figure 9.

The tool allows the user to place an initial template on the vertebra and apply the ACS algorithm. The user may enhance the image using histogram equalization before applying the method. The tool allows the user to create a template by marking points around the vertebra. The template can then be saved for future use. The position and size of the template can be controlled by rotation, translation and scaling prior to invoking the ACS algorithm. After the segmentation, the user can accept and save or discard the segmentation results. If the segmentation is accepted, the tool estimates the location of the next vertebra and places the template on it and the process is repeated.
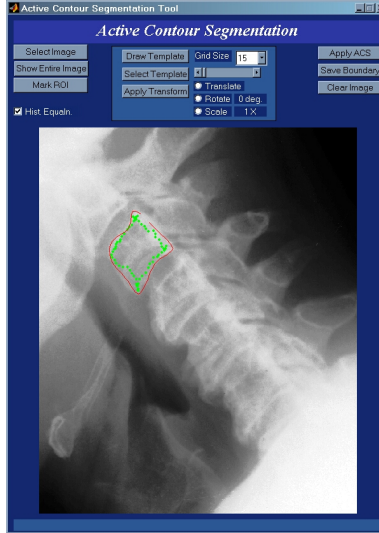
Figure 9: Active Contour Segmentation Tool: Main Window

```
<cebimg>
<head>
<Type>cervical spine</Type>
<Mode>x-ray</Mode>
<View>lateral</View>
<Dbsource>NHANES II</Dbsource>
<Imgid>C00215</Imgid>
<Coordsys>image</Coordsys>
<Origin>(0,0)</Origin>
<Filecreator>DMK</Filecreator>
<Creatorlevel>engineer</Creatorlevel>
<Createdate>19-Jul-2002 15:55:33</Createdate>
</head>
<obj>
<Id>7314166634788</Id>
<Anatomyid>vertebra</Anatomyid>
<Anatomyid2>C1</Anatomyid2>
<Anatregion>face</Anatregion>
<Geomtype>region</Geomtype>
<Segmentor>DMK</Segmentor>
<Segsoftware>acsegment</Segsoftware>
<Segdate>19-Jul-2002 15:55:25</Segdate>
<Bound>
<Pts>(356,1102),(361,1105),...</Pts>
<OrientPts>(-20,-48),...</OrientPts>

<BoundBox>(349,1186),(349,1041),(469,1041),(469,1186)</BoundBox>
<Theta>72</Theta>
<OrientBox>(446,1025),(340,1060),(387,1200),(492,1165)</OrientBox>
<Origin>(408,1106)</Origin>
</Bound>
<Manual>
<Pts>(352,1113),(355,1124),...</Pts>
<OrientPts>(3,-60),(14,-58),...</OrientPts>
<BoundBox>(352,1190),(352,1025),(482,1025),(482,1190)</BoundBox>
<Theta>88</Theta>
<OrientBox>(479,1022),(349,1027),(355,1193),(486,1188)</OrientBox>
<Origin>(412,1108)</Origin>
</Manual>
<Template>
<Pts>(344,1106),...</Pts>
<OrientPts>(-19,-64),...</OrientPts>
<BoundBox>(342,1191),(342,1030),(470,1030),(470,1191)</BoundBox>
<Theta>75</Theta>
<OrientBox>(449,1016),(328,1049),(372,1208),(492,1174)</OrientBox>
<Origin>(410,1107)</Origin>
</Template>
</obj>
...
</cebimg>
```

Figure 10: XML style **.cbr** file for storing segmented vertebra information

If at any point the segmentation is not acceptable, the user can perform manual segmentation on the vertebra.

We have incorporated the manual segmentation function in our prototype to allow other CBIR tasks to proceed while we develop more reliable automated segmentation techniques. The tool generates segmentation output in a XML style file and is given a **.cbr** file extension. A sample file snippet for a segmented cervical spine x-ray is shown in Figure 10. As the figure shows, we store for each object, the template, and automated, and manual segmentation results. This way, entries can be modified as needed following future developments.

The **.cbr** file records the information about an image in a **cebimg** structure. The information about the image, database source, view (e.g., lateral, sagittal, AP), and the human segmentor are retained in the header structure called **head**. Additionally, the coordinate system origin is also specified. This is used by the objects within the image as a reference point. The image identifier is the same as that in the text database, allowing cross indexing across the image and text databases. The segmented objects are stored within the **cebimg** structure as **obj** structures. These contain
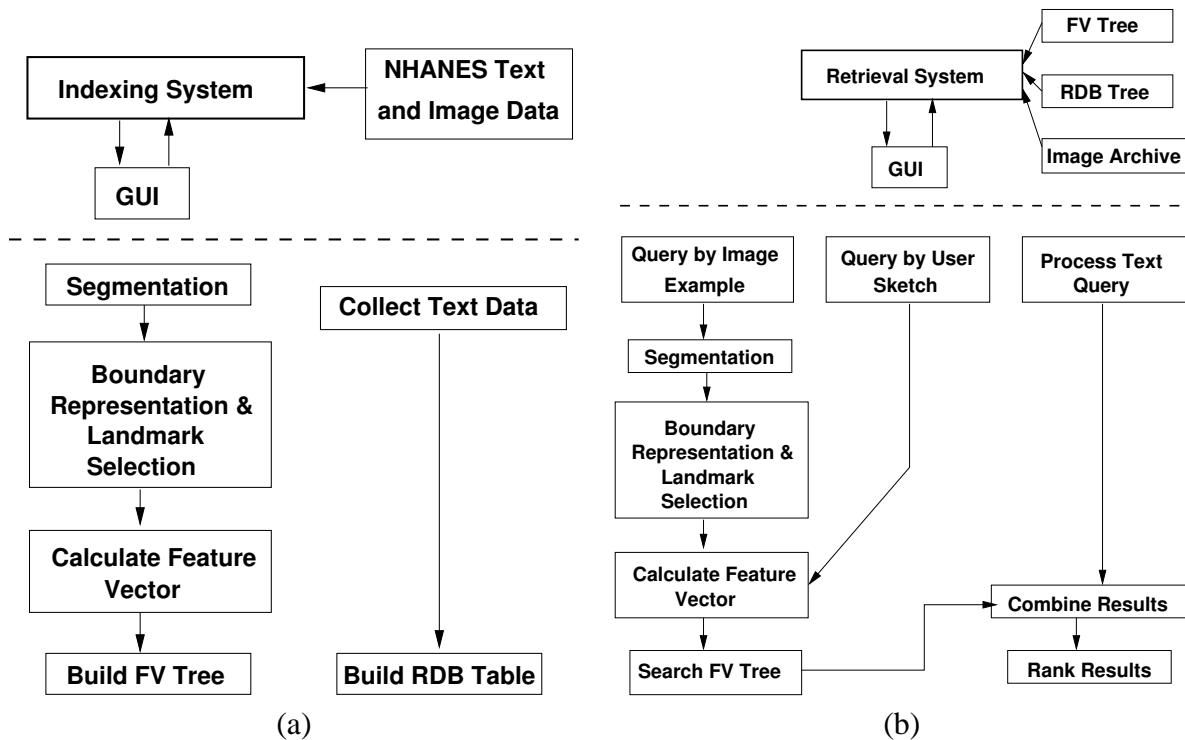
Figure 11: Block Diagram of the CBIR system: (a) Indexing System and (b) Retrieval System

a unique object identifier, anatomy identifiers, region of the anatomy, the segmented boundary points, the bounding box, the oriented bounding box, etc. In addition, the structure also allows the storage of the manually modified boundary points, if any, and the template used for segmentation. The unique object identifier and the variety of object boundary data extracted allow many versions of the segmented object to be retained. This has been designed so that a variety of database schema can be maintained. For example, the current best segmentation could be exposed for user CBIR searches, while research could proceed with segmentations available via other schemas as we improve the state-of-the-art in CBIR.

In addition to the data in the `.cbr` file, landmark points that have been marked on the vertebra boundary by a board-certified radiologist are also available for 550 cervical and lumbar x-ray images.

5.3 Indexing system

The Indexing system consists of methods to:

- extract image features
- transform the image features into a form suitable for similarity matching, and
- organize image and text data.

The architecture of the indexing system is shown in Figure 11(a). The indexing process is currently semi-automated and done via a graphical interface. This interface allows indexing of two types of data. The text data is organized as fields in a relational database table from which data can be
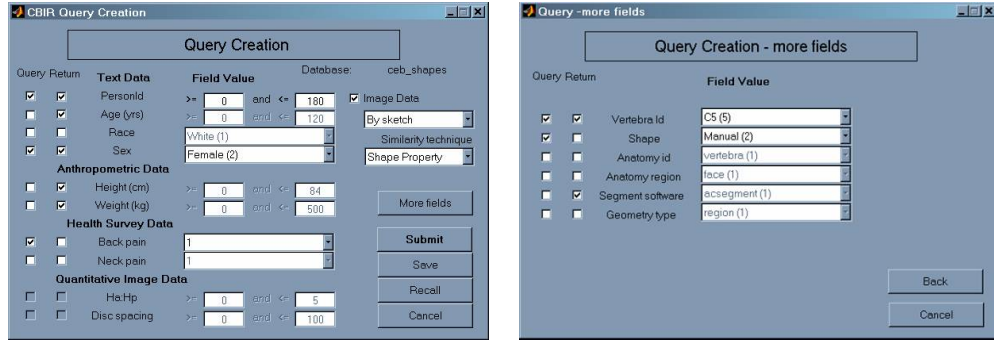
(a)   (b)
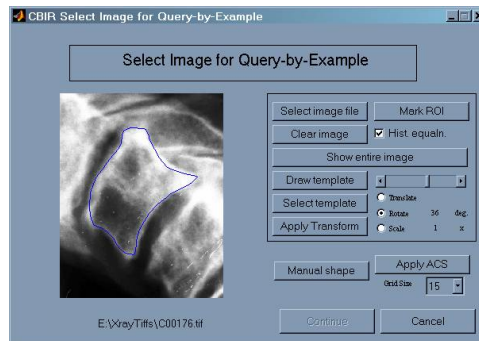
(c)

Figure 12: Query Dialogs. (a) Main Screen (b) Options Dialog (c) Query-by-Example Dialog.

retrieved using a relational database manager such as MySQL. The indexing of the image data on the other hand is a more involved process.

The first step in indexing the images is the segmentation of the objects of interest, the vertebrae. In our CBIR prototype, the shapes are segmented using the ACS tool described above. If the ACS segmentation were deemed unsatisfactory, then the boundary was marked manually. Both the automatically segmented boundary and manual boundary are stored in the database. The system is designed to replace this boundary information with the output of more reliable segmentation algorithms as they are developed.

Following the acquisition of the boundary outline of the vertebra shape, The dense boundary points stored in the **.cbr** file are reduced to a small set of meaningful representative points by a shape representation algorithm. This coarse boundary and a binary image representation of the vertebra are used to find meaningful shape features that are invariant to translation, rotation, scaling and starting-point shift. Issues related to this are discussed in Section 4.3.2.2. A feature vector is then created from various computed features and organized into a data structure for efficient retrieval. While research is proceeding toward determining an effective feature vector organization strategy, we are currently using a flat structure and linear search for retrieval. Having an inefficient, but working system enables us to improve on various modules as the research evolves. We are currently working on developing a suitable structure for efficient organization of image features and retrieval of the image data.

5.4 Retrieval system

The Retrieval system provides an interface to, and includes the set of methods for, content-based image retrieval and retrieval of text data. The system uses the following data to support user queries:

- the Feature-Vector indexing structure,
- the image archive, and
- the Relational Database Tree for text data.

The basic types of queries are to the text data, image data and combined queries to both. The retrieval of the text data is supported through Open-Database Connectivity (ODBC) protocol to retrieve results using the MySQL DBMS. The queries to the image data can be specified in using an example image to retrieve images that are visually similar or by drawing a sketch of the indexed feature, in this case the vertebra boundary. The system presents the user with a GUI for creating queries and supports text, image example, and image sketch queries, and queries that combine text and image example or image sketch. Figure 12(a) and (b) show the initial screen and the options screen for generating the basic query. The retrieval paths for image-example based queries and sketch-based queries are the same except for the feature extraction phase necessary for the former. The same feature extraction phase as in the indexing process is applied to the example image. The
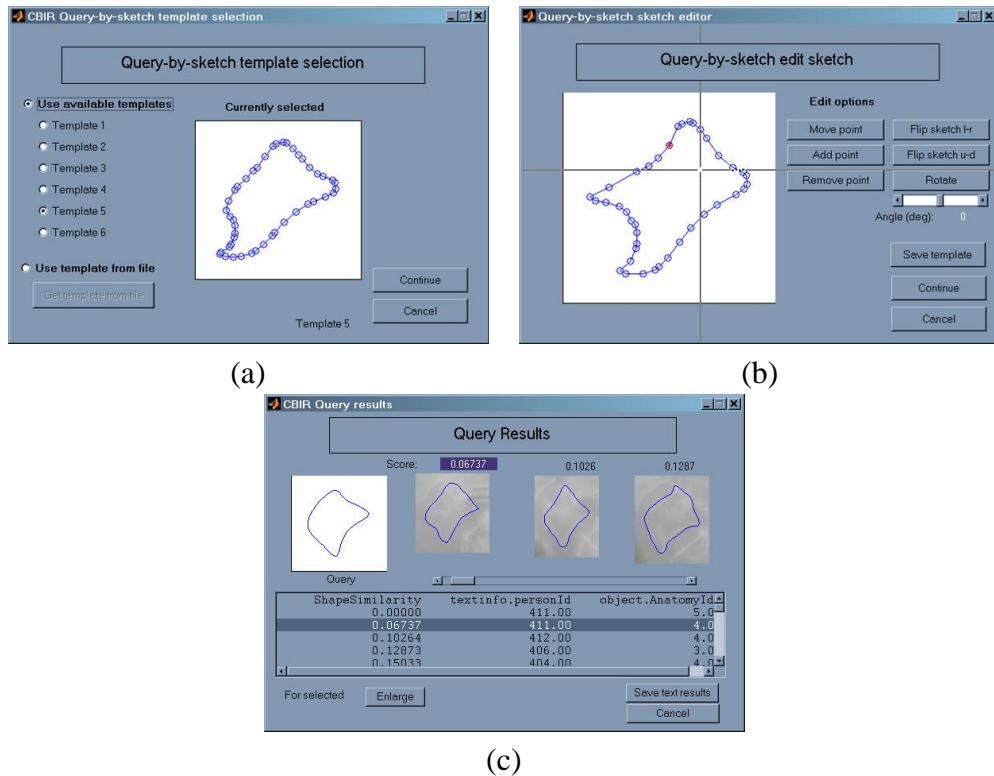

(a)   (b)

(c)

Figure 13: (a) Query-by-Sketch (Template) Dialog (c) Query-by-Sketch (User Sketch) (c) Results

image features in the query are then matched by a shape similarity algorithm to determine the similarity distance between the query and the database shape. The greater the distance between

34

two feature vectors the greater is the dissimilarity. The system allows users to specify an image for an image-example based query as shown in Figure 12(c). For an sketch-based query, the users may choose to either use one of the provided templates, or use their own template and modify it or draw an outline from scratch. The dialog boxes for these are shown in Figure 13(a) and (b).

At present, the retrieval is based on the global vertebra shape, but we intend to enhance the shape matching to allow searches on local regions in the vertebra boundary, e.g., to search for an anterior osteophyte. This and related objectives are itemized in Section 6.1.2. The results from the query are presented as shown in Figure 13(c).

5.5   Status

The CBIR2 prototype system described above enables content-based access to vertebra boundaries extracted from the scanned x-ray images and associated text data from the NHANES II survey. This system is a significant improvement over the earlier prototype that used the 9 points marked by a radiologist. This system extracts, with minimal human intervention, the entire vertebra boundary and stores it in a **.cbr** file structure. The image data can be retrieved by shape using one of four shape representation and similarity methods described earlier. The indexing and retrieval process is done via a GUI which greatly simplifies the process.

Despite these advantages, much work needs to be done. The system therefore is designed to be a CBIR research testbed for evaluating the enhancements. The modular design of the prototype allows us to replace modules with improved ones, thereby allowing testing of newly developed methods on live data.

Several areas are under investigation in the current prototype. Shape segmentation still requires some human intervention. We intend to develop a more automated segmentation technique. In addition, the issue of validation of the segmented boundary is yet to be addressed. The shape representation methods too are limited in their abilities. While they allow matching of the entire vertebra shape, it is not possible to pose queries on local features of importance such as anterior osteophytes. Finally, we are still in the process of identifying a suitable indexing structure for efficient storage and retrieval of the feature vectors.

Our proposed goals that address and improve upon the current shortcomings in CBIR are described in the next section.

6   Proposed work

6.1   Shape-based retrieval

6.1.1   Shape segmentation

- Development of reliable shape segmentation algorithms (possibly) trained to extract desired features from vertebral x-ray images. ASM achieves this to an extent. Collaboration with researchers at Texas Tech University aims to do this with Generalized Hough Transform (GHT), ASM, and Deformable models.
- Improve active contour methods to segment shapes. Near term improvements could include efforts in easing/minimizing human interaction.

6.1.2   Shape representation and similarity

Shape representation and similarity techniques implemented in the system allow matching of global shapes. However, they need to be made more sensitive subtle shapes corresponding to the pathology.

- Polygon approximation algorithms and Fourier Descriptors show potential as good methods for matching shapes. However, several enhancements are necessary for them to allow matching of subshapes. Fourier Descriptors work very well for matching overall shapes, however, they do not support matching of local regions. The polygonal approximation algorithm identifies and retains points that are significant to the shape. However, the method does not necessarily retain points that correspond to significant pathologies. These modifications would help support queries on specific pathology.
- Additionally an open-contour similarity algorithm is necessary for matching portions of shape.
- The similarity matching algorithms support rotation, translation, and scaling invariance. However, in both the polygon approximation and Fourier Descriptor algorithms rotation invariance is achieved by searching through the bend-angle or turn-angle plots. These techniques need to be made more efficient, possibly through use of landmark features.

  Traditionally shape representation methods are disjoint from the image data, in that they operate purely on the geometrical boundary information. However, image features could be used as landmarks in order to achieve these goals.

6.2   Indexing and feature organization for CBIR

Hurdles in extracting image features and organizing them for efficient search remain to be addressed:

- Continued refinement of accurate, labor-efficient segmentation tools for x-ray image shape extraction, toward the goal of allowing this process to be done at the skill level of a trained technician, thus overcoming the economic barrier to obtaining this biomedically important image information
- Research toward the representation of vertebrae shape boundaries in compact feature vectors that capture all of the biomedically relevant shape information
- Development and validation of efficient storage mechanisms for organizing and retrieving these feature vectors. (This work is currently underway by one of our collaborators, who is studying methods for optimizing the structure of spatial data trees for specific datasets [23].)

6.3   CBIR evaluation

A critical issue for CBIR systems is the evaluation of their performance. Current approaches require the creation of benchmark queries and "expected query results" created by the human observer, using largely subjective criteria. These results become the "ground truth" for evaluating the performance of the CBIR system against the benchmark query. For example, in the work

of Brodley [75] using CBIR on high resolution computed tomography (HRCT) images of the lung, a radiological lung specialist was shown the top four images returned as most similar to an example query; the specialist scored the returned images on a scale from 1-5 according to the level of agreement of the radiologist with the CBIR similarity estimation. Several basic measures are computed relative to this "ground truth", including:

- Recall - (returned relevant results)/(all relevant results)
- Precision - (returned relevant results)/(all returned results)
- Fallout - (returned irrelevant results)/(all irrelevant results)

From these fundamental measures, calculated over a range of benchmark queries, plots may be derived that characterize the system performance. For example, "retrieval effectiveness" is calculated by expressing precision as a function of recall. A desirable system characteristic is to have high precision values for all values of recall [76].

6.4 Evaluation with biomedical domain experts

Close coordination of our work with the biomedical community is a critical factor for synchronizing our common goals and producing results that are, in the final analysis, to be judged by their worth to users whose domain of interest is biomedical, rather than computer science or engineering. Toward this end, we have worked with rheumatologists and radiologists from NIH and medical schools to create WebMIRS and the Digital Atlas, to study digital resolution levels to use for hand x-ray film, to collect coarse segmentation data, and collect multi-reader interpretations of disc space narrowing and subluxation. We will continue this direction by:

- Demonstrating the CBIR2 system at national rheumatological conferences, such as the American College of Rheumatology and the Radiological Society of North America meetings and initiating dialog with potential collaborators.
- Soliciting biomedical collaboration by direct invitation to researchers who have published in the field of vertebral morphometry or have otherwise expressed interest in the area of x-ray image retrieval from large collections by image characteristics

Significant biomedical validation is required for:

- Accuracy and usefulness of vertebral shapes that are produced by the system segmentation and which are represented by feature vectors in the database
- Accuracy of the data returned by shape similarity queries and usefulness of the query types supported
- Accuracy and usefulness of the biomedical classifications produced by the indexing system

6.5 Web-aware biomedical CBIR

It is a goal of this project to investigate the potential usability of CBIR techniques on widely-accessible Web-based systems such as the WebMIRS system. The results of this work will determine the practicality of building CBIR capability into the present WebMIRS system or of architecting a new Web-based system on the foundation of the CBIR work.

6.6 Other images and modalities

It is a goal of this project to extend the CBIR techniques developed to at least one other large collection of images, of a different anatomical type and a different image modality. Two such image classes that are under investigation are (1) digitized color images of the cervix which are of high interest to the NCI; and (2) digitized color endoscopy images. Both of these image classes have been proposed as collections of interest by biomedical researchers who have requested technical assistance from CEB.

6.7 Milestones

Year 1: Complete candidate algorithms for all stages of CBIR applied to the NHANES II x-rays, incorporate these algorithms in an advanced CBIR2 prototype, and demonstrate this prototype at national forums.

Year 2: Using the CBIR2 prototype as a testbed, conduct biomedical validation studies on accuracy and usefulness of vertebral shape data for reliable indexing and retrieval; complete the indexing of NHANES II x-rays; select a second image collection for CBIR research.

Year 3: Enhance CBIR2 prototype with algorithms suitable for indexing and retrieval of the second image collection; conduct biomedical validation studies; complete indexing the second image collection.

7 Summary

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

Acknowledgements

University of Missouri-Rolla, and Daniel Abodeely, M.D., and Greg Cizek, M.D., of the Phelps County Medical Center, Rolla, Missouri. All physicians who have collaborated in this work are board-certified in radiology.

References

1. Jorgensen C. Image access: Bridging multiple needs and multiple perspectives; introduction and overview. Journal of the American Society for Information Science and Technology. 2001; 52(11):906–910.

2. Kulikowski CA. Medical imaging informatics: Challenges of definition and integration (editorial). Journal of the American Medical Informatics Association. 1997 May/Jun;4(3):252–253.

3. Wang JZ. Wavelets and imaging informatics: A review of the literature. Journal of Biomedical Informatics. 2001;34:129–141.

4. U. S. Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics, Hyattsville, MD. Public Use Data Tape Documentation, Health History Supplement, Ages 12-74 Years,Tape Number 5305, National Health and Nutrition Examination Survey, 1976-1980. 1983 August.

5. Long LR, Ostchega Y, Goh GH, Thoma GR. Distributed data collection for a set of radiological x-ray interpretations. Proceedings of SPIE Storage and Retrieval for Image and Video Databases V. 1997; 3022:228–237.

6. Long LR, Berman LE, Neve L, Roy G, Thoma GR. An application-level technique for faster transmission of large images on the internet. Proceedings of SPIE Multimedia Computing and Networking. 1995;2417:116–129.

7. Ostchega Y, Long LR, Goh GH, Hirsch R, Ma LD, Scott WWJ, Johnson W, Thoma GR. Establishing the level of digitization for wrist and hand radiographs for the third national health and nutrition examination survey. Journal of Digital Imaging. 1998;11(3):116–120.

8. Brogan D. Software for sample survey data, misuse of standard packages. Encyclopedia of Biostatistics, P Armitage, T Colton, eds., vol. 5. John Wiley & Sons, New York. 1998; 4167–4174.

9. Korn EL, Graubard BI. Analysis of Health Surveys. Wiley-Interscience. 1999.

10. Ghebreab S. Strings and Necklaces: On Learning and Browsing Medical Image Segmentations. Intelligent Sensory Information Systems Group, University of Amsterdam, Amsterdam, Netherlands. 2002.

11. Grau V, Alcaniz M, Juan MC, Monserrat C, Knoll C. Automatic localization of cephalometric landmarks. Journal of Biomedical Informatics. 2001;34(3):146–156.

12. Long LR, Pillemer S, Goh GH, Berman LE, Neve L, Thoma GR, Premkuman A, Ostchega Y, Lawrence R, Altman RD, Lane ND, Scott Jr WW. A digital atlas for spinal x-rays. Proceedings of SPIE Medical Imaging: PACS Design and Evaluation. 1997;3035:586–594.

13. The Epidemiology of Chronic Rheumatism, vol. II of Atlas of Standard Radiographs of Arthritis. F. A. Davis Company, Philadelphia, PA. 1963.

14. Meadows S, Thoma GR, Long LR, Mitra S. Entropy encoding of difference images from adjacent visible human digital color photographic slices for lossless compression. Proceedings of SPIE Medical Imaging: Image Processing. 1997;3031:749–755.

15. Zamora G, Wilson M, Mitra S, Thoma G. An innovative web-based system for high lossless compression and fast, interactive transmission of visible human color images. The Third Visible Human Conference Proceedings. 2000; CDROM.

16. National Library of Medicine Long Range Plan: Obtaining factual information from databases. Tech. rep., National Library of Medicine. 1986 June.

17. Zink S, Jaffe CC. Medical imaging databases: A National Institutes of Health workshop. Investigative Radiology. 1993;28(4):366–372.

18. Smeulders AWM, Worring M, Santini S, Gupta A, Jain R. Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000 December; 22(12):1349–1380.

19. Antani S, Kasturi R, Jain R. A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video. Pattern Recognition. 2002;35(4):945–965.

20. Rui Y, Huang TS, Change SF. Image retrieval: Current techniques, promising directions, and open issues. Journal of Visual Communications and Image Representation. 1999;10(1):39–62.

21. Bredno J, Kohnen M, Dahmen J, Vogelsang F, Wein BB, Lehmann TM. Synergetic impact obtained by a distributed development platform for Image Retrieval in Medical Applications (IRMA). Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Media Databases. 2000;3972:321–331.

22. Lehmann TM, Wein BB, Dahmen J, Bredno J, Vogelsang F, Kohnen M. Content-based image retrieval in medical applications: A novel multistep approach. Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Media Databases. 2000;3972:312–320.

23. Qian X, Tagare HD. Optimally adapted indexing trees for medical image databases. IEEE International Symposium on Biomedical Imaging. 2002; .

24. Robinson GP, Tagare HD, Duncan JS, Jaffe CC. Medical image collection indexing: Shape-based retrieval using kd-trees. Computerized Medical Imaging and Graphics. 1996;20(4):209–217.

25. Tagare HD, Jaffe CC, Duncan J. Medical image databases: A content-based approach. Journal of the American Medical Informatics Association (JAMIA). 1997;4(3):184–198.

26. Guttman A. R-trees: A dynamic index structure for spatial searching. ACM SIGMOD Record - Proceedings of Annual Meeting. 1984;14(2):47–57.

27. Sellis T, Roussopoulos N, Faloutsos C. The R^+-tree: A dynamic index for multidimensional objects. Proc. 13th VLDB Conference. 1987; 507–518.

28. Beckmann N, Kriegel HP, Schneider R, Seeger B. The R^*-tree, an efficient and robust access method for points and rectangles. Proceedings of ACM SIGMOD International Conference on Management of Data. 1990; 322–331.

29. De Jonge W, Scheuermann P, Schuf A. S^+-trees: An efficient structure for the representation of large pictures. CVGIP: Image Understanding. 1994 May;59(3):265–280.

30. Lin KI, Jagadish HV, Faloutsos C. The TV-tree: An index structure for high dimensional data. The VLDB Journal. 1994;3(4):517–542.

31. Korn F, Sidiropoulos N, Faloutsos C, Siegel E, Protopapas Z. Fast nearest neighbor search in medical image databases. Proc. 22nd VLDB Conference. 1996; 215–226.

32. Ankerst M, Kriegel HP, Seidl T. A multistep approach for shape similarity search in image databases. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1998 November/December; 10(6):996–1004.

33. Chen JY, Bouman CA, Dalton JC. Hierarchical browsing and search of large image databases. IEEE Transactions on Image Processing. 2000;9(3):442–455. Unread.

34. Chu H. Research in image indexing and retrieval as reflected in the literature. Journal of the American Society for Information Science and Technology. 2001;52(12):1011–1018.

35. Tull L. Cooperative digital imaging projects: the Greater Cincinnati memory project. The Electronic Library. 2002;20(1):43–48.

36. Goodrum AA, Rorvig ME, Jeong KT, Suresh C. An open source agenda for research linking text and image content features. Journal of the American Society for Information Science and Technology. 2001;52(11):948–953.

37. Stan D, Sethi IK. Mapping low-level image features to semantic concepts. Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Media Databases. 2001;4315:172–179.

38. Ogiela M, Tadeusiewicz R. Semantic-oriented syntactic algorithms for content recognition and understanding of images in medical databases. Proceedings of IEEE International Conference on Multimedia and Expo (ICME2001). 2001; .

39. Spraycar M, (Ed.) Stedman's Medical Dictionary. Williams & Wilkins, Baltimore, MD, 26th ed. 1995.

40. Long LR, Thoma GR. Localization and feature indexing in spine x-rays. Journal of Digital Imaging. 2001;10(4):939–956.

41. Zamora G, Sari-sarraf H, Mitra S, Long LR. Estimation of orientation and position of cervical vertebrae for segmentation with active shape models. Proceedings of the IS&T/SPIE Medical Imaging 2001: Image Processing. 2001;4322:378–387.

42. Cootes TF, Taylor CJ. Statistical models of appearence for computer vision. Tech. rep., University of Machester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, Manchester, M12 9PT, U.K. 2001 February.

43. Smyth PP, Taylor CJ, Adams JE. Vertebral shape: Automatic measurement with active shape models. Radiology. 1999;211(2):571–578.

44. Gardner JC, Heyano SL, Yaffe LG, von Ingersleben G, Chesnut CHI. A semi-automated computerized system for fracture assessment of spinal x-ray films. Proceedings of SPIE Medical Imaging: Image Processing. 1996;2710:996–1008.

45. 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.

46. Long LR, Thoma GR. Segmentation and image navigation in spine x-rays. Proceedings of SPIE Medical Imaging: Image Processing. 2000;3979:169–179.

47. Geiger D, Gupta A, Costa LA, Vlontzos J. Dynamic programming for detecting, tracking, and matching deformable contours. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1995 March;17(3):294–302.

48. Raphael C. Coarse-to-fine dynamic programming. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2001 December;23(12):1379–1390.

49. Mokhtarian F, Mackworth AK. Scale-based description and recognition of planar curves and two-dimensional shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1986;8(1):34–43.

50. Mokhtarian F, Mackworth AK. A theory of multiscale, curvature-based shape representation for planar curves. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1992;14(8):789–805.

51. Hoffman ME, Wong EK. Content-based image retrieval by scale-space object boundary shape representation. IS&T/SPIE Conference on Storage and Retrieval for Media Databases. 2000;3972:86–97.

52. Jain AK, Vailaya A. Shape-based retrieval: A case study with trademark image databases. Pattern Recognition. 1998;31(9):1369–1390.

53. Ang YH, Li Z, Ong SH. Image retrieval based on multidimensional feature properties. Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases III. 1995;2420:47–57.

54. Hu MK. Visual pattern recognition by moment invariants. IRE Transactions on Information Theory. 1962;8:179–187.

55. Kim YS, Kim WY. Content-based trademark retrieval system by using visually salient feature. Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 1997; 307–312.

56. Ip HHS, Shen D, Cheung KKT. Affine invariant retrieval of binary patterns using generalized complex moments. Second International Conference on Visual Information Systems (VISUAL'97). 1997; 301–308.

57. Del Bimbo A, Pala P. Shape indexing by multi-scale representation. Image and Vision Computing. 1999;17(3-4):245–261.

58. Bengtsson A, Eklundh JO. Shape representation by multiscale contour approximation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1991;13(1):85–93.

59. Latecki LJ, Lakämper R. Application of planar shape comparison to object retrieval in image databases. Pattern Recognition. 2002;35(1):15–29.

60. Latecki LJ, Lakämper R. Shape description and search for similar objects in image databases. State-of-the-Art in Content-Based Image and Video Retrieval, RC Veltkamp, H Burkhardt, HP Kriegel, eds., vol. 22 of Computational Imaging and Vision. Kluwer Academic Publishers. 2001; 69–96.

61. Arkin EM, Chew LP, Huttenlocher DP, Kedem K, Mitchell JSB. An efficient computable metric for comparing polygon shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1991; 13(3):209–216.

62. Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and Machine Vision. Brooks/Cole Publishing Company, second edition ed. 1999.

63. Gonzalez R, Woods R. Digital Image Processing. Prentice Hall, second ed. 2002.

64. Zahn C, Roskie R. Fourier descriptors for plane closed curves. IEEE Computer. 1972;C-21(3):269–281.

65. Meilhac C, Nastar C. Relevance feedback and category search in image databases. IEEE International Conference on Multimedia Computing Systems, vol. 1. 1999; 512–517.

66. Zaniolo C, Ceri S, Faloutsos C, Snodgrass RT, Subrahmanian VS, Zicari R. Advanced Database Systems. Morgan Kaufmann Publishers, Inc. 1997.

67. Duda RO, Hart PE, Stork DG. Pattern Classification. John Wiley & Sons, Inc., New York, NY, second ed. 2000.

68. Kurniawati R, Jin JS, Sheperd JA. The SS^+-tree: An Improved Index Structure for Similarity Searches in a High-Dimensional Feature Space. Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases V. 1997;3022:110–120.

69. Pun T, Squire D. Statistical structuring of pictorial databases for content-based image retrieval systems. Pattern Recognition Letters. 1996;17(12):1299–1310.

70. White D, Jain R. Similarity indexing: Algorithms and performance. Proceedings of IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases IV Vol. SPIE 2670. 1996; 62–73.

71. Stanley RJ, Cherukuri M, Long LR, Thoma GR. The application of shape features to cervical spine vertebra image analysis. 2002. Submitted to IEEE Transactions on Medical Imaging.

72. Stanley RJ, Long LR. A radius of curvature-based approach to cervical spine vertebra image analysis. 38th Annual Rocky Mountain Bioengineering Symposium, vol. 37. 2001; 385–390.

73. Dryden I, Mardia K. Statistical Shape Analysis. John Wiley & Sons, Inc. 1998.

74. Jain AK, Duin RPW, Mao J. Statistical pattern recognition: A review. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2000 January;22(1):4–37.

75. Brodley C, Kak A, Shyu C, Dy J, Broderick L, Aisen AM. Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision. Proceedings of The Sixteenth National Conference on Artificial Intelligence. 1999; 760–767.

76. Smith JR. Quantitative assessment of image retrieval effectiveness. Journal of the American Society for Information Science and Technology. 2001;52(11):969–979.

77. Del Bimbo A, Pala P. Effective image retrieval using deformable templates. Proc. International Conference on Pattern Recognition. 1996; 120–124.

78. Scarloff S. Deformable prototypes for encoding shape categories in image databases. Pattern Recognition. 1997;30(4):627–641.

79. Pala P, Santini S. Image Retrieval by Shape and Texture. Pattern Recognition. 1999;32(3):517–527.

80. Adoram M, Lew MS. IRUS: Image Retrieval Using Shape. IEEE International Conference on Multimedia Computing Systems, vol. 2. 1999; 597–602.

81. Günsel B, Tekalp AM. Shape similarity matching for query-by-example. Pattern Recognition. 1998; 31(7):931–944.

82. Lei Z, Tasdizen T, Cooper D. Object signature curve and invariant shape patches for geometric indexing into pictorial databases. Proceedings of IS&T/SPIE Conference on Multimedia Storage and Archiving Systems. 1997;3229:232–243.

83. Alferez R, Wang YF. Image Indexing and Retrieval Using Image-Derived, Geometrically and Illumination Invariant Features. IEEE International Conference on Multimedia Computing Systems, vol. 1. 1999; 177–182.

84. Petrakis EGM, Milios E. Efficient retrieval by shape content. IEEE International Conference on Multimedia Computing Systems, vol. 2. 1999; 616–621.

85. Mokhtarian F, Abbasi S. Retrieval of Similar Shapes under Affine Transform. Third International Conference on Visual Information and Information Systems (VISUAL'99), Appears in Lecture Notes in Computer Science No. 1614. 1999; 566–574.

86. Mokhtarian F, Abbasi S. Shape similarity retrieval under affine transforms. Pattern Recognition. 2002;35(1):31–41.

87. Sharvit D, Chan J, Tek H, Kimia BB. Symmetry-based indexing of image databases. Journal of Visual Communications and Image Representation. 1998;9(4):366–380.

88. Rui Y, Huang TS, Mehrotra S, Ortega M. Automatic matching tool selection using relevance feedback in MARS. Second International Conference on Visual Information Systems (VISUAL'97). 1997; 109–116.

89. Kliot M, Rivlin E. Invariant-based shape retrieval in pictorial databases. Computer Vision and Image Understanding. 1998;71(2):182–197.

90. Derrode S, Daoudi M, Ghorbel F. Invariant content-based image retrieval using a complete set of fourier-mellin descriptors. IEEE International Conference on Multimedia Computing Systems, vol. 2. 1999; 877–881.

Questions for the Board

1. Does the Board recommend an extension of our CBIR research to include biomedical images from different domains (e.g., specific clinical specialities) or of different data types (e.g., 3D imagery or video) important to biomedical research and practice?

2. We formulate user queries as query by example image, query by sketch, standard text query, and a hybrid of text and image/sketch. Does the Board see any of these as exceptionally important to future multimedia biomedical databases, or all of these as equally important?

3. Of the biomedical features in spine x-rays considered important by the NIH workshop, only 3 could be reliably detected in the NHANES films. Does the Board recommend that we pursue image processing techniques to enhance the digital images to attempt to extract these other features? Alternatively, can the Board suggest other sets of spine x-ray images in which these features are present?

Appendix A   Shape representation and similarity algorithms

In Section A.1 we present a brief survey of the literature on shape representation and similarity methods. Following this, in Section A.2, we present details on the methods used in the evaluation and CBIR2 prototype system. The shape used to explain these algorithms is a vertebra boundary segmented from a cervical spine x-ray image from the NHANES II dataset.

A.1   Shape representation and similarity

Different approaches are taken for matching shapes in CBIR systems. Some researchers have used them as a matching tool in image example type queries. Others have projected its use for user sketch type queries. The argument for the latter is that in a user sketch the human perception of image similarity is inherent and the image matching subsystem does not need to develop models of human measures of similarity. One approach adopts the use of deformable image templates to match user sketches to the database images [77, 78, 79]. Since the user sketch may not be an exact match of the shape in the database, the method elastically deforms the user template to match the image contours. An image for which the template has to undergo minimal deformation, or loses minimum energy, is considered the best match. A low match means that the template is lying in areas where the image gradient is 0. By maximizing the matching function and minimizing the elastic deformation energy, a match can be found. The distance D between two images (or image regions) is given by

$$\mathcal{D}(T, I, u) = \int \int_S (Tu_1(x_1, x_2), u_2(x_1, x_2)) - I(x_1, x_2))^2 \mathrm{d}(x_1, x_2). \tag{A-1}$$

In Equation A-1, x_1 and x_2 are coordinates of some point on the grid on surface S, $u = (u_1, u_2)$ defines the deforming function causing the template and the target image to match, resulting in new coordinates given by $u_1(x_1, x_2)$ and $u_2(x_1, x_2)$ for template T and image I. In the following, $\mathcal{J}(f)$ is the total amount of bending of the surface defined by $(x_1, x_2, f(x_1, x_2))$ is measured as in Equation A-2.

$$\mathcal{J}(f) = \int \int_S ((f_{x_1 x_1})^2 + 2(f_{x_1 x_2})^2 + (f_{x_2 x_2})^2) \mathrm{d}(x_1, x_2). \tag{A-2}$$

The deformation energy, given by $\mathcal{F}$ in Equation A-3, is thus $\mathcal{J}(u_1) + \mathcal{J}(u_2)$. This is the balance between the amount of warp and the energy associated with it, where μ controls the stiffness of the template.

$$\mathcal{F} = \mu(\mathcal{J}(u_1) + \mathcal{J}(u_2)) + \mathcal{D}(T, I, u). \tag{A-3}$$

Adoram and Lew [80] use Gradient Vector Flow (GVF) based active contours (snakes) to retrieve objects by shape. They note that deformable templates are highly dependent on their initialization and are unable to handle concavities. The authors present results by combining GVF snakes with invariant moments. Günsel and Tekalp [81] define a shape similarity based directly on the elements of the mismatch matrix derived from the eigenshape-decomposition. A proximity matrix is formed using the eigenshape representation objects. The distance between the eigenvectors of the query and target object proximity matrices forms the mismatch matrix. The elements of the mismatch matrix indicate the matched feature points. These are then used to determine the similarity between the shapes.

A different approach to CBIR based on shape has been through use of implicit polynomials for effective representation of geometric shape structures [82]. Implicit polynomials are robust, stable and exhibit invariant properties. The method is based on fixing a polynomial to a curve patch. A vector consisting of the parameters of this curve is used to match the image to the query. A typical database would contain the boundary curve vectors at various resolutions to make the matching robust. Alferez and Wang [83] present a method to index shapes which is invariant to affine transformations, rigid-body motion, perspective transforms and change in illumination. They use a parameterized spline and wavelets to describe the objects. Petrakis and Milios [84] use a dynamic programming based approach for matching shapes at various levels of shape resolution. Mokhtarian and Abbasi [85, 86] apply the Curvature Scale Space based matching for retrieval of shapes under affine transform. Mokhtarian and Mackworth [49, 50] have earlier presented studies on representation of 2D planar curves using the scale-space theory. This has been further explored by Hoffman and Wong [51]. Sharvit et al [87] propose the use of shock structures to describe shapes. The describe the symmetry-based representation as one which retains the advantages of the local, edge-based correlation approaches, as well as of global deformable models. It is termed as an intermediate representation. Two benefits of this approach have been outlined; the computation of similarity between shapes and the hierarchical symmetries captured in a graph structure. Rui et al [88] propose the use of multiple matching methods to make the retrieval robust. They define the requirements of the parameter as invariance and compact form of representation. The authors define a Modified Fourier Descriptor (MFD) which is an interpolated form of the low frequency coefficients of the Fourier Descriptor normalized to unit arc-length. They also calculate the orientation of the major axis. The matching of the images is then done using the Euclidean distance, MFD matching, Chamfer distance and Hausdorff distance. Although these matching tools have been used in this system, they can also be used to match shapes which have been specified using other appropriate descriptors. Jain and Vailaya present a study of shape based retrieval methods with respect to trademark image databases [52]. An invariant-based shape retrieval approach has been presented by Kliot and Rivlin [89]. Semi-local multi-valued invariant signatures are used to describe the images. Such representation when used with containment trees, a data structure introduced by the authors, allows for matching shapes which have undergone a change in viewpoint, or are under partial occlusion. It also allows retrieval by sketch. The invariant shape re-parametrization is done by applying various transforms (translation, rotation, scale) to the curve signature. Translation, rotation and scale invariance, which is imperative for shape based retrieval, can also be achieved through the use of Fourier-Mellin Descriptors. Derrode et al [90] base their system on these and describe them to be stable under small shape distortions and numerical approximations. The Analytical Fourier-Mellin Transform (AFMT) is used to extract the Fourier-Mellin Descriptors. An alternate form of representing shape has been to transform the nodes into a statistical shape space and compare the points [73].

There are several techniques for shape representation and similarity that have been developed over the years. However, almost all of these techniques had been developed to address some particular application of shape representation. The methods can be classified into those that generate a shape representation based on all the boundary points and those that take the local curvature into consideration. None of these methods take the local context of the boundary shape into consideration when determining the representation.
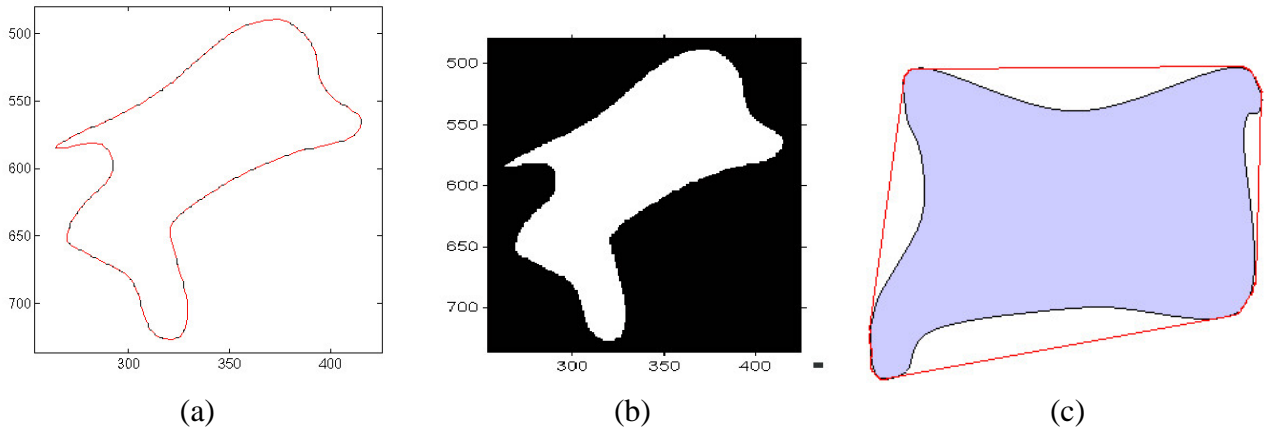
(a)　　　　　　　　　　　(b)　　　　　　　　　　　(c)

Figure A-1: Global Shape Properties: (a) Original contour (b) Binary image (c) Convex perimeter of a shape

A.2　Algorithms selected for evaluation

I **Global Shape Properties and Invariant Moments.** Global shape properties and moments are features intrinsic to a shape. The properties include area, center of mass, perimeter, axis of rotation, etc. In order to compute these, the shape contour must be converted to a binary image for processing which will give the same weight for each pixel inside the shape contour. Figure A-1(a) and (b) show the original shape contour and its binary image counterpart.

A **Global Shape Properties** Several global shape properties were calculated for selection. They include:

- *Center of gravity*: The average of x and y coordinates of all pixels inside the shape contour. This is calculated as the first order moment along each axis.
- *Area*: The total number of pixels inside the object contour.
- *Perimeter*: The total length of the contour in number of pixels.
- *Convex perimeter*: An approximation of the perimeter of the convex hull of an object. An example (red line) is shown in Figure A-1(c).
- *Major axis length and angle*: The major axis of the result of least-squared error fit of an ellipse.
- *Minor axis length and angle*: The minor axis of the result of least-squared error fit of an ellipse.
- *Compactness*: This is defined as $perimeter^2/4\pi.Area$
- *Roughness*: Defined as $roughness = perimeter/convexperimeter$. A smooth convex object, such as a perfect circle, will have the roughness of 1.0
- *Elongation*: The ratio of major axis length to minor axis length.

Besides the center of gravity and area which do not meet the invariant requirements for shape-based retrieval, these shape properties can be used to create a multi-dimensional feature space for calculating the distance between two shapes (in the feature space). The distance between two points in the feature space is a measure of shape similarity. The farther the two points are, the less similarity between the two shapes.

B **Invariant Moments** For a 2D continuous function $f(x, y)$, the moment of order $(p + q)$ is defined as:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy \tag{A-4}$$

The central moments are calculated by shifting the origin to the center of the image given by $(\bar{x}, \bar{y})$ and defined as:

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \tag{A-5}$$

In the discrete domain, the integral is changed to a summation resulting in:

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \tag{A-6}$$

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \tag{A-7}$$

The first and second order of the central moments can be derived from Equation A-7 and expressed as:

$$
\begin{array}{ll}
\mu_{00} = m_{00} & \mu_{20} = m_{20} - \bar{x} m_{10} \\
\mu_{10} = 0 & \mu_{02} = m_{02} - \bar{y} m_{01} \\
\mu_{01} = 0 & \mu_{11} = m_{11} - \bar{y} m_{10}
\end{array}
$$

The normalized central moments would then be:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^r}, \text{where } r = \frac{p + q}{2} + 1 \tag{A-8}$$

The relevant Hu [54] invariant moments then are given by

$$
\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2
\end{aligned}
$$

II **Scale Space Filtering.** Scale space filtering reformats the shape boundary points to represent the shape at different levels of detail. It is said to follow human perception of shapes [57]. It provides more detail at scale higher level and progressively reduces the detail level until the shape becomes an oval shape. While capable of shape matching, a problem with this method is that the shape shrinks as it progresses from high detail level toward low detail level making comparison scale sensitive.

1.) **Curve Smoothing.** A closed planar curve can be parameterized according to its length and be expressed as $c(t) = \{x(t), y(t)\}$, where x and y represent the coordinates for each boundary point and t is the normalized length from a selected starting point. To smooth a
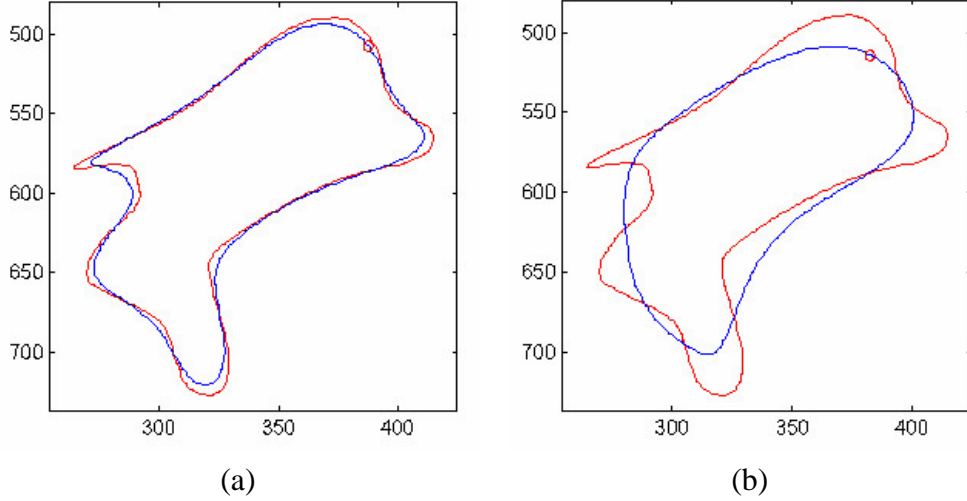
Figure A-2: Curve Smoothing (a) $\sigma = 4.0$ (b) $\sigma = 1.5$

curve, functions $x(t)$ and $y(t)$ can be convolved with a one-dimensional Gaussian kernel at different levels of abstraction as expressed below:

$$x(t, \sigma) = \int_{-\infty}^{t} x(s)g(t - s, \sigma)ds$$

$$y(t, \sigma) = \int_{-\infty}^{t} y(s)g(t - s, \sigma)ds$$

$g(s, \sigma)$ is a Gaussian function with selectable variance s for different levels of detail. Figure A-2 shows the results of scale space filtering at different levels of detail. The shape will eventually become an ellipse if we continue reducing the value of σ.

2.) **Shape Representation and Token Descriptions.** The curvature function of $c(t, \sigma)$ at $\{x(t, \sigma), y(t, \sigma)\}$ can be expressed as:

$$\Gamma(t, \sigma) = \frac{X'(t, \sigma)Y''(t, \sigma) - X''(t, \sigma)Y'(t, \sigma)}{((X'(t, \sigma)^2 + Y'(t, \sigma)^2)^{3/2}} \tag{A-9}$$

where X', Y' and X'', Y'' are the first and second derivatives of $y(t, \sigma)$ and $y(t, \sigma)$, the parameterized boundary coordinate functions.

The critical points on the polygon can be determined by searching for the minima on the curvature function. Once the critical points are determined, the polygon can be broken down to small segments (shape tokens). A polygon smoothed by a Gaussian function with $\sigma = 4.0$ is shown in Figure A-3(a). Gaussian function window size was increased at a step of two pixels until five tokens were left. Figure A-3(b) shows the curvature function when the specified number of tokens was obtained. The dots in red on the plot indicate the minima which correspond to the selected points. Figure A-3(c) shows a small segment of the polygon defined by two critical points. The visual aspect of each token (segment) is represented by three properties. They are symmetry (s), length (l), and orientation (o). As shown in the figure, the symmetry is calculated as $s = s_1 - s_2$ and the orientation is the angle from horizontal of the line connecting the center of gravity of the token and
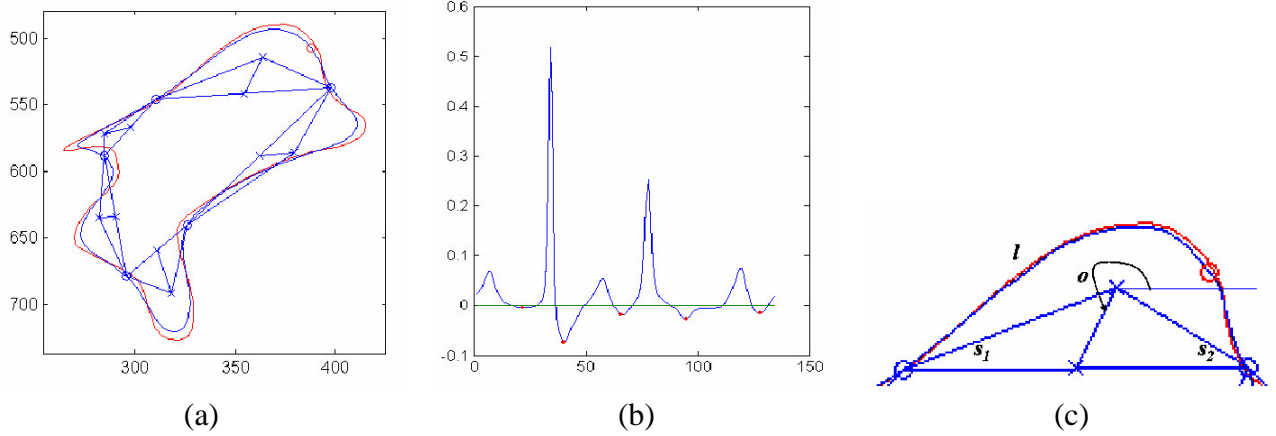
Figure A-3: Curvature Function analysis (a) at $\sigma = 4.0$ (b) plot of function (c) Token detail.

the mid point of the line connecting two adjacent critical points. The length of token l was calculated and normalized by the total length of the polygon. The token orientation cannot be used for similarity measurement because it does not meet rotation invariance requirement.

3.) **Similarity Metric.** The L_2-norm is used to compute the distance between two shapes. However, a problem with this method is that it is not inherently rotation invariant. The effect of rotation on a shape is that the curvature function shifts to the right or left depending on the direction of rotation. The values at the extreme ends of the plot wrap around and appear at the other end. Rotation invariance can be achieved if a correspondence can be determined between the tokens. This can be done, for example, by selecting a reference token (such as the longest token) and shifting the other so that they match up.

III **Polygon Approximation.** Polygon approximation or curve evolution is a process that eliminates insignificant shape features and reduce the number of data points. The resultant representation is one that uniquely describes the shape. The approximated curve was then converted to tangent space for similarity measurement.

1.) **Curve Evolution.** Curve evolution is used to reduce the influence of noise and to simplify the shapes by removing irrelevant and keeping relevant shape features. This is achieved by iteratively comparing the relevance measure of all vertices on the polygon. Higher relevance value means that the vertex has larger contribution to the shape of the curve. For each iteration, the vertex that has the lowest relevance measure is removed and a new segment is established by connecting the two adjacent vertices. The relevance measure is calculated as

$$K(s_1, s_2) = \frac{\beta(s_1, s_2)l(s_1)l(s_2)}{l(s_1) + l(s_2)} \tag{A-10}$$

where β is the turn angle and l is the normalized length for shapes s_1 and s_2. The relevance measure is in direct proportion to the turn angle and the length of the curve segment. Figure A-4(a) shows the result of 100 iterations of the curve evolution algorithm on the original curve shown in Figure A-1(a).

2.) **Tangent Space.** The turn angle was calculated by referencing to the horizontal line as shown in Figure A-4(a). The smoothed curve was then represented by the turn function
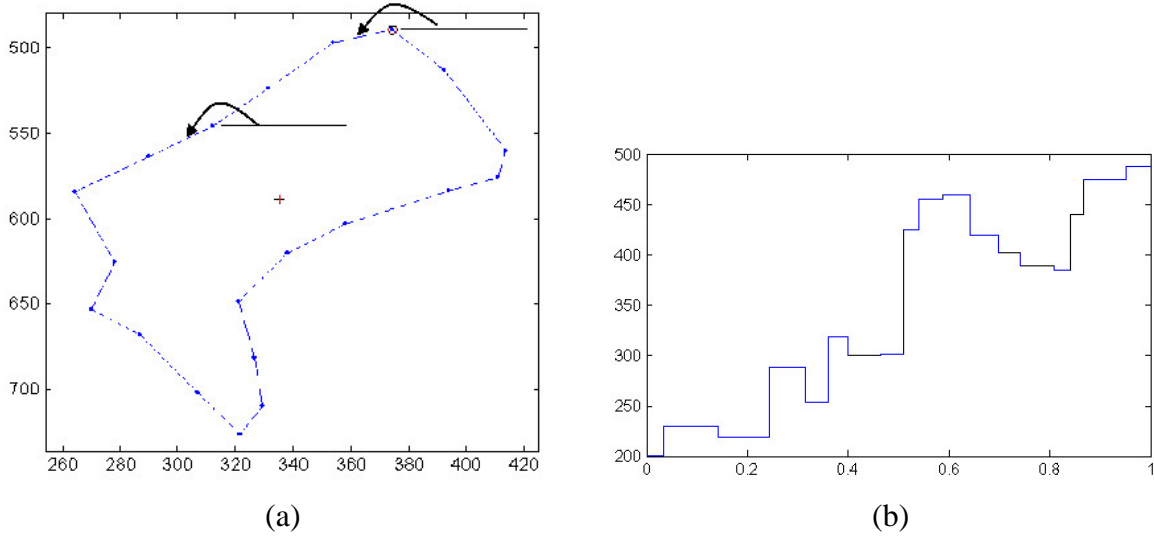
A-6

Figure A-4: Curve evolution (a) result and turn function and (b) plot of turn function.

shown in Figure A-4(b) which is the function of turn angle verses the normalized length. The length was normalized to 1.0 and shown in the X axis. The Y axis represents the turn angle.

Representing shape in tangent space meets the invariant requirements for shape-based retrieval. It is translation invariant because the turn angles and length do not contain information about the shape location. It is scaling invariant because it uses normalized length. For rotation and starting point shift, the turn function remains the same shape expect shifting vertically when there is a rotation and moving horizontally when there is a shift in starting point. Figure A-5(a) illustrates the rotation and starting point shift invariants. The dotted blue lines represent the turn function of the original curve. It shifted to the left (red) when the starting point is shifted by 20 and it moved down (green) when the original shape was rotated by 20 degrees.

3.) **Similarity Measurement** The distance (dissimilarity) between two turn functions Θ_A and Θ_B can be measured as

$$
\begin{aligned}
\delta_2(A, B) &= \|\Theta_A - \Theta_B\|_2 \\
&= \sqrt{\left(\int_0^1 |\Theta_A - \Theta_B|^2 \, ds\right)} \\
&= \sqrt{\min_{\theta \epsilon R, t \epsilon [0,1]} \left(\int_0^1 |\Theta_A(s+t) - \Theta_B(s) + \theta|^2 \, ds\right)} \quad \text{(A-11)}
\end{aligned}
$$

To measure the distance, the two turn functions must be aligned first. In most cases, the turn functions are not identical because of difference in shape. The alignment can only be achieved through minimizing the distance while shifting one turn function (query or database). In other words, the distance between two turn functions is obtained by performing a two-dimensional search to find the minimum distance. Another approach is to reduce the search to one dimension by calculating the best value of θ [61]. The best
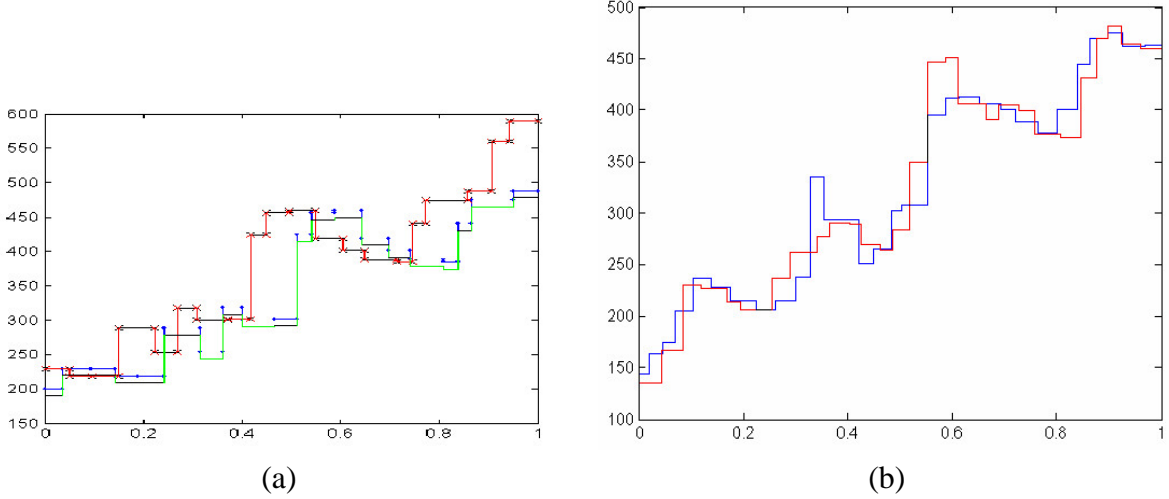
| (a) | (b) |

Figure A-5: Curve Evolution: (a) Rotation and starting point shift invariance (b) Curve evolution search result.

value of θ is a function of length shift t in the X axis to minimize

$$
\begin{aligned}
h(t, \theta) &= \int_0^1 \mid \Theta_A(s+t) - \Theta_B(s) + \theta \mid^2 ds & \text{(A-12)} \\
\text{when } \theta'(t) &= \int_0^1 (\Theta_A(s+t) - \Theta_B(s) + \theta) ds \\
&= \alpha - 2\pi t, \\
\text{where } \alpha &= \int_0^1 \Theta_B(s) ds - \Theta_A(s) ds
\end{aligned}
$$

Figure A-5(b) shows the one dimensional searching result. For each searching step in X (length) direction, the best value of θ was calculated according to Equation A-10. The distance δ was calculated and recorded. After shifting the turn function through the searching range, the minimum δ is the distance between the two turn functions.

IV **Fourier Descriptors.** The position of a point on a closed contour is a periodic function. Thus, the Fourier series may be used to approximate the contour. The resolution of the contour is approximation is determined by the number of terms in the Fourier series. Since simple operations such as scaling and translation are related to simple operations of the boundary's Fourier descriptors, they are attractive for use with boundary matching [64]. Rotation however requires the bend angle function to be computed.

Bend Angle. The bend angle verses normalized length function was calculated so that the shape representation meets the invariance requirements. The bend angle was calculated such that the clockwise turn gives a negative angle whereas a counter clockwise turn gives a positive angle as shown in Figure A-6. This method represents a closed polygon curve C (m vertices) as $\Theta(t)$, i.e., the bend angle as a function of length t. The parameter t is the normalized accumulated length. Because it does not contain orientation information, this representation meets the rotation invariance requirement. Normalized length makes it independent on the polygon size. The only issue left is the starting point shift invariance requirement which is taken care of by the shift invariance property of the power spectrum.
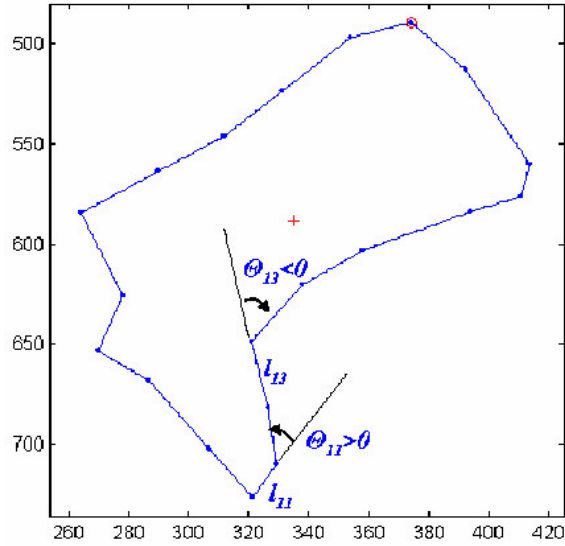
Figure A-6: Bend angle vs. normalized length.

The Fourier expansion of $\Theta(t)$ is expressed as

$$\Theta(t) = \mu_0 + \sum_{n=1}^{\infty}(a_n \cos nt + b_n \sin nt), \qquad \text{(A-13)}$$

where a_n and b_n are coefficients for each frequency component. The power spectrum of the bend angle function is invariant to the shift in length (t in this case). Because of this property, Fourier descriptors on a bend angle function meet all invariant requirements for shape-based retrieval. The similarity between shapes is the normalized difference between the Fourier descriptors of the shapes. The lower the difference, greater is the similarity.

Curriculum Vitae

George R. Thoma
Chief, Communications Engineering Branch, LHNCBC, NLM

Education and Training:

Swarthmore College, Swarthmore, PA   B.S.   1965   Electrical Engineering
University of Pennsylvania   M.S.   1967   Electrical Engineering
University of Pennsylvania   Ph.D.   1971   Electrical Engineering

Research and Professional Experience:

1984 - Present   Chief, Communications Engineering Branch, LHNCBC, NLM

- Directs and conducts R&D in projects involving image processing and communications engineering. Project descriptions and results appear in the published literature (see Publications), and some are synopsized in the branch's Web home page *http://archive.nlm.nih.gov*. Projects include the following:

  - MARS (Medical Article Records System): a system to automate the entry of bibliographic records from journal articles. This system based on document image analysis and understanding techniques is now serving one-third of the library's total production needs.

  - DXPNET, WebMIRS: an Internet-accessible multimedia database of digitized x-rays and associated text; content-based image indexing; online digital x-ray atlas.

  - Image compression techniques applied to the highly data intensive 24 b/p color cryosection Visible Human image collection.

  - AnatLine: end-user access to a database of Visible Human cryosection and rendered images.

  - DocView: research in document imaging for end-user Internet access to electronic documents.

1974 - 1984   Senior Electronics Engineer, Communications Engineering Branch, LHNCBC, NLM

- Conducted R&D in communications systems and signal processing techniques applied to electronic document storage , retrieval and display; biomedical image data compression, text recognition and image enhancement, satellite communications, echo suppression, videodisk premastering, motion-adaptive video compression.

1973 - 1974   Systems Engineer, General Electric Co., Space Division, Valley Forge, PA.
1971 - 1973   Systems Engineer, AII Systems, Moorestown NJ.
1968 - 1971   Ford Foundation Fellow, University Research Fellow and Postdoctoral Research Associate, University of Pennsylvania, Philadelphia.

Honors and Professional Activities:

- General Chair, 14th IEEE Symposium on Computer-Based Medical Systems, 2001

- Fellow of the SPIE (elected 1999)
- Awards: NIH Merit Award 1998, NLM Regents Award 1998, Federal Computer Week's Federal 100 Award in 1995, NLM Staff Recognition Awards (1997-2002); On-the-spot Award (1998); Special Act / Service Award (1997); Special Act Group Award (1992); Merit Awards (1984-97).
- Appointed to Internet2 Applications Strategy Council (2002).
- Appointed to Maryland Governor's Task Force on High Speed Networks (1998 - present).
- Keynote speaker at SPIE Electronic Imaging, San Jose CA, February 1996; and at 8th IEEE Symposium on Computer-based Medical Systems, Lubbock TX, June 1995.
- Appointed to Blue Ribbon Panel at National Institute for Standards and Technology for document file format standards (1994-95).
- Invited referee for: IEEE Journal on Special Areas in Communications; IEEE Computer Journal; IEEE Parallel and Distributed Technology; Information Processing and Management Journal; Journal of Clinical Engineering; SPIE and IEEE conferences.
- Association of Image Information Management Certificate of Service (1995)
- American College of Physicians Certificate of Accomplishment (1993)
- NASA Certificate of Appreciation (1993)

Publications (most recent only):

- Thoma GR, Ford G, Le DX, Li Z. Text verification in an automated system for the extraction of bibliographic data. Proc. 5th International Workshop on Document Analysis Systems, Springer-Verlag: Berlin, August 2002, 423-32.
- Thoma GR, Ford G. Automated data entry system: performance issues. Proc. SPIE: Document Recognition and Retrieval IX, Vol. 4670, January 2002, 181-90.
- Krainak DM, Long LR, Thoma GR. A method of content-based retrieval for a spinal x-ray image database. Proceedings of SPIE Medical Imaging 2002: PACS and Integrated Medical Systems. Vol. 4685, San Diego, CA, February 23-28, 2002.
- Long LR, Krainak DM, Thoma GR. Identifying image structures for content-based retrieval of digitized spine x-rays. Proceedings of SPIE Medical Imaging 2002: Image Processing. Vol. 4684, San Diego, CA, February 23-28, 2002.
- Walker FL, Thoma GR. A SOAP-enabled system for an online library service. Proc. Info-Today 2002, Medford NJ: Information Today, May 2002, 320-9.
- Thoma GR. Automating the production of bibliographic records for MEDLINE. Internal R&D report, CEB, LHNCBC, NLM; September 2001; 92 pp.
- Tran LQ, Moon CW, Le, DX, Thoma GR. Web page downloading and classification. Proc. 14th IEEE Symposium on Computer-Based Medical Systems. Los Alamitos, CA: IEEE Computer Society. July 2001, 321-6.
- Le DX, Tran LQ, Chow J, Kim J, Hauser SE, Moon CW, Thoma GR. Automated medical citation records creation for Web-based online journals. Proc. 14th IEEE Symposium on Computer-Based Medical Systems, Los Alamitos CA: IEEE Computer Society. July 2001, 315-20.
- Long LR, Thoma GR. Computer-assisted retrieval of biomedical image features from spine x-rays: progress and prospects. Proceeding of the 14th IEEE Symposium on Computer-Based Medical Systems (CBMS 2001), July 26-27, 2001, Bethesda, MD, pp. 46-50.
- Long LR, Thoma GR. Landmarking and feature localization in spine x-rays. Journal of Electronic Imaging 10(4), October 2001, 939-56.

Curriculum Vitae

L. Rodney Long
Electronics Engineer, Communications Engineering Branch, LHNCBC, NLM

Education and Training:

University of Texas, Austin TX	B.A.	1971	Mathematics
University of Texas, Austin TX	M.A.	1976	Mathematics
University of Maryland, College Park MD	M.A.	1986	Applied Mathematics

Research and Professional Experience:

1990-present,   Electronics Engineer,
                Communications Engineering Branch, LHNCBC, NLM

- Coordinates and conducts R&D in projects related to multimedia databases, image processing, and the dissemination of biomedical information over wide-area networks. Projects include:

  - Content Based Image Retrieval (CBIR): research into and development of prototype systems to allow efficient, computer-assisted indexing of the content of medical images, and the retrieval of images by similarity to image examples, sketches, and/or related text descriptors.

  - Biomedical Image Processing: research into biomedical image segmentation, region identification, and classification in support of CBIR.

  - WebMIRS: prototype Web-accessible multimedia database system.

  - Digital Atlas of the Cervical and Lumbar Spine: prototype digital atlas created with images graded to consensus by medical experts selected by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS).

1976-1990,   International Business Machines, Inc.

- Programmer for Space Shuttle and Global Positioning System projects, principally design and implementation of orbit modeling programs. Engineer/Scientist for analysis of radar applications.

Honors and Professional Activities:

- General Co-Chair of IEEE Symposium on Computer-Based Medical Systems, 2001
- NLM Special Achievement Award, 2000.

Publications (most recent only):

- Krainak DM, Long LR, Thoma GR. A method of content-based retrieval of a spinal x-ray image database. Proc. SPIE Medical Imaging 2002: PACS and Integrated Medical Systems. Vol. 4685, San Diego, CA, February 23-28, 2002, 108-116.

- Long LR, Krainak DM, Thoma GR. Identifying image structures for content-based retrieval of digitized spine x-rays. Proc. SPIE Medical Imaging 2002: Image Processing. Vol. 4684, San Diego, CA, February 23-28, 2002, 1204-1214.
- Long LR, Thoma GR. Landmarking and feature localization in spine x-rays. Journal of Electronic Imaging 10(4), October 2001, 939-956.
- Long LR, Thoma GR. Computer-assisted retrieval of biomedical image features from spine x-rays: progress and prospects. Proc. of 14th IEEE Symposium on Computer-Based Medical Systems (CBMS 2001), July 26-27, 2001, Bethesda, MD, 46-50.
- Long LR, Thoma GR. Identification and classification of spine vertebrae by automated methods. Proc. SPIE Medical Imaging 2001: Image Processing. Vol. 4322, San Diego, CA, February 19-22, 2001, 1478-1489.
- Long LR, Thoma GR. Feature indexing in a database of digitized x-rays. Proc. SPIE Storage/ Retrieval for Media Databases 2001. Vol. 4315, San Jose, CA, January 24-26, 2001, 393-403.
- Long LR, Thoma GR. Use of shape models to search digitized spine x-rays. Proc. IEEE Computer-Based Medical Systems 2000, Houston, TX, June 23-24, 2000, 255-260.
- Long LR, Thoma GR. Web Access to National Health Survey Text/Image Databases. Chap. 13 in Medical Image Databases, ed. by Wong, STC. Kluwer Academic Press, 1998, 325-353.
- 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. Proc. SPIE Medical Imaging 1997. Vol. 3035, Newport Beach, CA, February 22-28, 1997, 586-594.
- Ostchega Y, Long LR, Hirsch R, Ma LD, Scott, Jr. WW, Thoma GR. Establishing the level of digitization for wrist and hand radiographs for the third National Health and Nutrition Examination Survey. Journal of Digital Imaging, vol. 11, no. 3, August 1998, 116-120.
- Stanley RJ, Long R. A radius of curvature-based approach to cervical spine vertebra image analysis. Proc. 38th Ann. Rocky Mountain Bioengineering Symposium 37, April 2001, 385-391.
- Tezmol A, Sari-Sarraf H, Mitra S, Long R. Customized Hough transform for robust segmentation of the cervical vertebrae from x-ray images, Southwest Symposium on Image Analysis and Interpretation (SSIAI 2002), April 7-9, 2002, Santa Fe, New Mexico, 224-228.
- Thoma GR, Long LR. Compressing and transmitting Visible Human images. IEEE Multi-Media; New York: IEEE Computer Society. Vol. 4, No. 2; April-June 1997; pp. 36-45.
- Zamora G, Sari-Sarraf H, Mitra S, Long R. Analysis of the feasibility of using Active Shape Models for segmentation of gray scale images. Proc. SPIE Medical Imaging 2002: Image Processing. Vol. 4684, San Diego, CA, February 23-28, 2002, 1370-1381.
- Zamora G, Sari-Sarraf H, Mitra S, Long R. Estimation of orientation and position of cervical vertebrae for segmentation with active shape models. Proc. SPIE Medical Imaging 2001: Image Processing. Vol. 4322, San Diego, CA, February 19-22, 2001, 378-387.

Curriculum Vitae

Sameer K. Antani
Staff Scientist, Communications Engineering Branch, LHNCBC, NLM

Education and Training:

University of Pune, Pune, India	B.E.	1994	Computer Engineering
The Pennsylvania State University	M.Eng.	1998	Computer Science and Engineering
The Pennsylvania State University	Ph.D.	2001	Computer Science and Engineering

Research and Professional Experience:

4/2002 - Present	Staff Scientist, Communications Engineering Branch, LHNCBC, NLM
12/2001 - 4/2002	Post-Graduate Research Fellow, ORISE/NLM-Medical Informatics Training Program, LHNCBC, NLM

- Conducts and coordinates R&D in topics that include, but are not limited to, content-based image (and video) retrieval, multimedia databases, image analysis, and computer vision.

8/1995 - 5/2001	Researcher, Systems Administrator (Computer Vision Laboratory), and Instructor, Department of Computer Science and Engineering, The Pennsylvania State University
5/1998 - 8/1998	Research Intern, Philips Research
1/1995 - 5/1995	Project Officer, Advanced Systems and Computer Training Ltd., India
7/1994 - 12/1994	Instructor, Pune Institute of Computer Technology (PICT), India. (Affiliated to University of Pune)

Honors

- Elected President, PSU Engineering Graduate Student Council, (1997-1998) and university-wide student leadership positions. (1996 - 2000)
- Student Representative, Strategic Planning Committee, Department of Computer Science and Engineering, PSU, 1998, 1996.
- Awards: Rotary International Award for Computer Excellence, 1994; National Merit Scholarship, Government of India, 1987
- Elected as Chair, IEEE Student Chapter, Pune - (1993 - 1994).
- Reviewer for several prestigious conferences such as ACM Multimedia, IAPR Intl. Conf. on Document Analysis and Recognition (ICDAR), IAPR Intl. Conf. on Document Analysis Systems (DAS), IEEE Intl. Conf. on Computer Vision & Pattern Recognition (CVPR), IEEE Intl. Conf. on Image Proc. (ICIP) and journals such as ACM Multimedia Systems Journal, SPIE Journal of Electronic Imaging, Intl. Journal of Document Analysis and Recognition, and IEEE Trans. on Systems, Man, and Cybernetics.

Publications:

- Antani S, Kasturi R, Jain R. A Survey on the use of Pattern Recognition Methods for Abstraction, Indexing and Retrieval of Images and Video. Pattern Recognition Journal, 2002 Apr;35(4):945-965.

- Antani S, Crandall D, Narasimhamurthy, A, Mariano VY, Kasturi R. Evaluation of Methods for Detection and Localization of Text and Video. Document Analysis & Systems (Post Workshop Book) - In Print.
- Crandall D, Antani S, Kasturi R. Robust Extraction of Stylized Text from Digital Video. Accepted for publication in Intl. Journal of Document Analysis and Recognition (IJDAR).
- Antani S, Long LR, Thoma GR, Lee DJ. Evaluation of Shape Indexing Methods for Content-Based Retrieval of X-Ray Images. Submitted to IS&T/SPIE Symposium on Electronic Imaging Science and Technology - Conf. on Storage and Retrieval for Media Databases 2003.
- Antani S, Long LR, Thoma GR. A Biomedical Information System for Combined Content-Based Retrieval of Spine X-ray Images and Associated Text Information. Submitted to the Indian Conf. on Computer Vision, Graphics, & Image Processing - 2002.
- Crandall D, Antani S, Kasturi R, Robust Detection of Stylized Text Events in Digital Video. Intl. Conf. on Document Analysis and Recognition, 2001 (ICDAR'2001).
- Antani S, Crandall D, Kasturi R, Gargi U. Evaluation of Methods for Extraction of Text from Video. Intl. Wksp. on Document Analysis Systems, 2000 (DAS'2000).
- Antani S, Crandall D, Kasturi R. Robust Extraction of Text in Video. Intl. Conf. on Pattern Recognition, 2000 (ICPR'2000).
- Gandhi T, Kasturi R, Antani S. Application of Planar Motion Segmentation for Scene Text Extraction. Intl. Conf. on Pattern Recognition, 2000 (ICPR'2000).
- Antani S, Gargi U, Kasturi R, et al. A System for Automatic Text Detection in Video. 5th Intl. Conf. on Document Analysis and Recognition, 1999 (ICDAR'99).
- Antani S, Agnihotri L. Gujarati Character Recognition. 5th Int'l Conf. on Document Analysis and Recognition, 1999 (ICDAR'99).
- Gargi U, Antani S, Kasturi R. Indexing Text Events in Digital Video Databases. 14th Intl. Conf. on Pattern Recognition, 1998 (ICPR'98).
- Gargi U, Antani S, Kasturi R. Performance Characterization and Comparison of Video Indexing Algorithms. Intl. Conf. on Computer Vision & Pattern Recognition, 1998 (CVPR'98).