international journal of medical informatics xxx (2008) xxx–xxx

ELSEVIER

journal homepage: www.intl.elsevierhealth.com/journals/ijmi

SPIRS: A Web-based image retrieval system for large biomedical databases

William Hsu[a], *Sameer Antani*[b,*], *L. Rodney Long*[b], *Leif Neve*[b], *George R. Thoma*[b]

[a] *Medical Imaging Informatics Group, University of California, Los Angeles, USA*
[b] *National Library of Medicine, U.S. National Institutes of Health, Bethesda, MD, USA*

ARTICLE INFO

Article history:
Received 20 February 2008
Received in revised form
11 August 2008
Accepted 16 September 2008

Keywords:
Medical informatics applications
Information storage and retrieval
Content-based image retrieval
Visual access methods
Web-based systems

ABSTRACT

Purpose: With the increasing use of images in disease research, education, and clinical medicine, the need for methods that effectively archive, query, and retrieve these images by their content is underscored. This paper describes the implementation of a Web-based retrieval system called SPIRS (Spine Pathology & Image Retrieval System), which permits exploration of a large biomedical database of digitized spine X-ray images and data from a national health survey using a combination of visual and textual queries.

Methods: SPIRS is a generalizable framework that consists of four components: a client applet, a gateway, an indexing and retrieval system, and a database of images and associated text data. The prototype system is demonstrated using text and imaging data collected as part of the second U.S. National Health and Nutrition Examination Survey (NHANES II). Users search the image data by providing a sketch of the vertebral outline or selecting an example vertebral image and some relevant text parameters. Pertinent pathology on the image/sketch can be annotated and weighted to indicate importance.

Results: During the course of development, we explored different algorithms to perform functions such as segmentation, indexing, and retrieval. Each algorithm was tested individually and then implemented as part of SPIRS. To evaluate the overall system, we first tested the system's ability to return similar vertebral shapes from the database given a query shape. Initial evaluations using visual queries only (no text) have shown that the system achieves up to 68% accuracy in finding images in the database that exhibit similar abnormality type and severity. Relevance feedback mechanisms have been shown to increase accuracy by an additional 22% after three iterations. While we primarily demonstrate this system in the context of retrieving vertebral shape, our framework has also been adapted to search a collection of 100,000 uterine cervix images to study the progression of cervical cancer.

Conclusions: SPIRS is automated, easily accessible, and integratable with other complementary information retrieval systems. The system supports the ability for users to intuitively query large amounts of imaging data by providing visual examples and text keywords and has beneficial implications in the areas of research, education, and patient care.

Published by Elsevier Ireland Ltd

* *Corresponding author at*: CEB/NLM/NIH MSC 3824, Building 38A, 8600 Rockville Pike, Bethesda, MD 20894, USA. Tel.: +1 301 435 3218.
E-mail address: santani@mail.nih.gov (S. Antani).

1386-5056/$ – see front matter. Published by Elsevier Ireland Ltd
doi:10.1016/j.ijmedinf.2008.09.006

1. Introduction

Medical imaging has become an important tool in not only objectively documenting patient presentation and clinical findings, but also understanding and managing various diseases. Image data provides tangible visual evidence of disease manifestation. For example, in spine X-rays, shape deformations of the vertebral body (e.g., protrusions of the bone) are often indicative of osteophytes, which commonly appear in conditions such as osteoarthritis. As medical imaging becomes more widely used in patient care and generates larger amounts of data, tools that allow users to retrieve and manipulate such data effectively are becoming essential. Hospitals and medical centers have been adopting technology such as Picture Archiving and Communication Systems (PACS), Hospital Information Systems (HIS), and Radiological Information Systems (RIS) to assist in the digital collection, organization, and storage of patient data. The goal of these systems is to make patient data more accessible; in reality, the amount of data that is entered and stored in these systems have created a new challenge for effective information indexing and retrieval. Historically, PACS have limited users to query by certain keywords (e.g., unique patient identifier, fields in the image header). However, these keywords often may not capture valuable visual characteristics contained within the image, thereby limiting the power of posed clinical queries and reducing the overall usefulness of the data. As imaging data becomes more abundant, technologies are needed to extract and structure the content of medical images and enable more efficient image retrieval for patient care, education, and research.

Over the past two decades, content-based image retrieval (CBIR) systems have been researched to address the problem of indexing and retrieval of visual data in a variety of domains [1]. Rather than limiting queries to textual keywords, CBIR allows the user to provide a visual query consisting of a sketch, an image, a color characteristic, texture sample, or a specified region of interest (ROI), which is then used to find other similar images in a database. Here, similarity is defined as minimizing the difference between features of interest in the query image and indexed images in a database. The problem of retrieving information based on image content has been researched by various groups since the early 1990s resulting in the development of tools such as Query by Image Content (QBIC) [2], Virage [3], and Blobworld [4]. CBIR in medicine has been an active area of research [5], but only a small number of proposed systems have been demonstrated in the clinical environment. While large biomedical image collections exist, such as the National Cancer Imaging Archive (NCIA) or the Lung Imaging Database Consortium (LIDC) created under the aegis of the Cancer Imaging Program[1] at the U.S. National Cancer Institute (NCI), these efforts have concentrated on data collection and transmission but have left development of applications to the research community. Early efforts in medical CBIR include Comparison Algorithm for Navigating Digital Imaging Databases (CANDID) [6], which is based on a technique similar to the *n*-gram method for searching free-text documents: a signature is computed that represents the content of the image in an abstract sense using texture features. A preliminary evaluation showed that the technique used in CANDID correctly classified computed tomography (CT) lung images of patients with scleroderma and vasculitis. I2C [7] integrates tools for defining image analysis routines based on specific image classes; some of the algorithms are interactive, while others are automated. This system has been implemented in a mini-PACS system and deployed on the web. ASSERT [8] requires a physician to delineate the ROI containing the pathology on which the features are indexed. The system has been evaluated clinically on high-resolution CT lung images. More recently, Image Retrieval for Medical Applications (IRMA) [9] has been developing a retrieval system that supports semantic and formalized queries. For example, IRMA would be able to automatically determine the staging of a patient's therapy or retrieve a set of images with similar diagnostic findings. While the IRMA system is a global feature system, which means it extracts a single feature vector from the entire image, efforts are underway to develop a multilevel distributed CBIR system [10] that supports both global and local features. These developments are pioneering efforts in CBIR and have helped demonstrate its potential use in the healthcare environment. However, widespread adoption and application of CBIR in the clinical environment have yet to be achieved. This lack of utilization is attributed to several limitations: (i) the difficulty of integrating current implementations with existing healthcare systems [5]; (ii) the amount of human expert intervention required for image indexing (e.g., I2C and ASSERT); and (iii) the lack of evaluation using a large set of text and imaging data. CBIR has the potential of improving many areas in healthcare, including research (e.g., finding the prevalence of a pathological feature in a large survey collection), education (e.g., creation of teaching files), and clinical decision support (e.g., diagnosing a patient given a set of image findings).

The goal of our work is to develop a retrieval system that implements recent developments in feature representation, efficient indexing, and similarity matching; supports whole and partial shape matching, which enables a wide range of meaningful queries to be posed; and utilizes a distributed framework, which is customizable to the needs and constraints of healthcare environments. The result, Spine Pathology & Image Retrieval System (SPIRS),[2] provides a Web-based interface for performing image retrieval on a database of digitized spine X-ray images using the morphological shape of the vertebral body. A query editor enables users to pose queries by sketching a unique shape or selecting or modifying an existing shape from the database. Additional text fields enable users to supplement visual queries with other relevant data (e.g., anthropometric data, quantitative imaging parameters, patient demographics). These hybrid text-image queries may be annotated with pertinent pathologies by selecting and weighting local features to indicate importance. Results appear in a customizable window that displays the top matching results and related patient data. SPIRS addresses current

[1] http://imaging.cancer.gov/.

[2] http://archive.nlm.nih.gov/spirs.

Please cite this article in press as: W. Hsu, et al., SPIRS: A Web-based image retrieval system for large biomedical databases, Int. J. Med. Inform. (2008), doi:10.1016/j.ijmedinf.2008.09.006

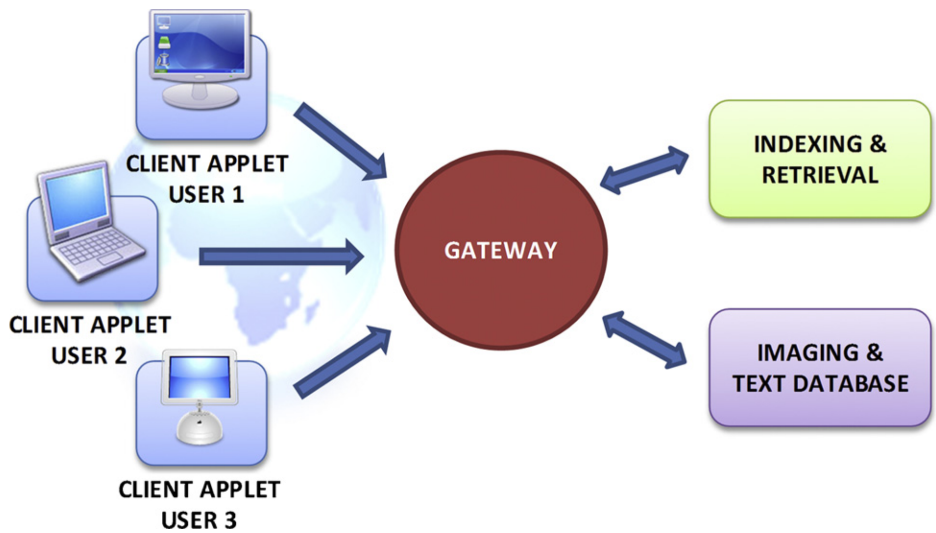


Fig. 1 – The distributed architecture of SPIRS. User clients communicate through the Internet to a gateway, which connects each client to the biomedical database and indexing and retrieval algorithms. The gateway acts as a mediator, authenticating users and ensuring that the data is sent to the appropriate user.

Table 1 – A subset of the text data collected as part of the NHANES II study.

Data groups	Example fields
NHANES II text data fields	
Demographical information	Age, gender, race, ethnicity, income, education, marital status
Anthropometric data	Height, weight, bitrochanteric breadth, elbow breadth, upper arm girth, triceps skinfold, subscapular skinfold, handedness
Medical history	Health status, use of medications, hospitalization, tuberculosis, eating problems, chronic conditions, anemia, diets, smoking
Health history supplement	Joint pain data including back disk and neck questions

limitations of CBIR implementations by: (i) utilizing open standards to communicate among components, which can be extended to support data encryption to meet privacy regulations; (ii) implementing algorithms that automate shape extraction and representation; (iii) utilizing feature indexing methods for efficient retrieval; (iv) combining text and image feature queries for hybrid queries; and (v) performing evaluation of a variety of whole and partial shape retrieval algorithms using a large biomedical dataset. Much work has been done in the past on visual querying paradigms, as reviewed in [11]. SPIRS utilizes the query-by-example paradigm, in which users provide an example shape or image to the system as a method of finding similar images. In addition, it combines visual and text queries to provide users with greater flexibility in retrieving relevant results. SPIRS is a working proof-of-concept that demonstrates the capability of accommodating large amounts of text and imaging data that is expected in the modern research and healthcare environments.

2. Methods

The system's distributed architecture, shown in Fig. 1, consists of four components: (i) the client applet, which provides a front-end for users to pose queries and interact with results; (ii) a gateway that acts as an mediator between client and server-side components; (iii) the indexing and retrieval server, which performs the feature representation and similarity matching; and (iv) the databases containing images and associated text data. The following sections describe the process by which imaging features are extracted and indexed and how the user interacts with SPIRS to perform queries and interact with the results.

2.1. *Biomedical database*

At the U.S. National Library of Medicine (NLM), the focus of CBIR research has been to develop systems capable of performing a range of queries on large medical multimedia databases comprising various biomedical images and patient health data. One such database contains digitized spine X-ray images and associated metadata from a nationwide survey, the National Health and Nutrition Examination Survey (NHANES),[3] conducted regularly by the National Center for Health Statistics (NCHS) at the Centers for Disease Control and Prevention (CDC) in the United States. The goals of NHANES include estimating prevalence of selected diseases, monitoring disease trends, and studying the relationship between nutrition and health. The Lister Hill National Center for Biomedical Communications, an intramural research and development division of the NLM, maintains data from

[3] NHANES: http://www.cdc.gov/nchs/nhanes.htm.

Please cite this article in press as: W. Hsu, et al., SPIRS: A Web-based image retrieval system for large biomedical databases, Int. J. Med. Inform. (2008), doi:10.1016/j.ijmedinf.2008.09.006

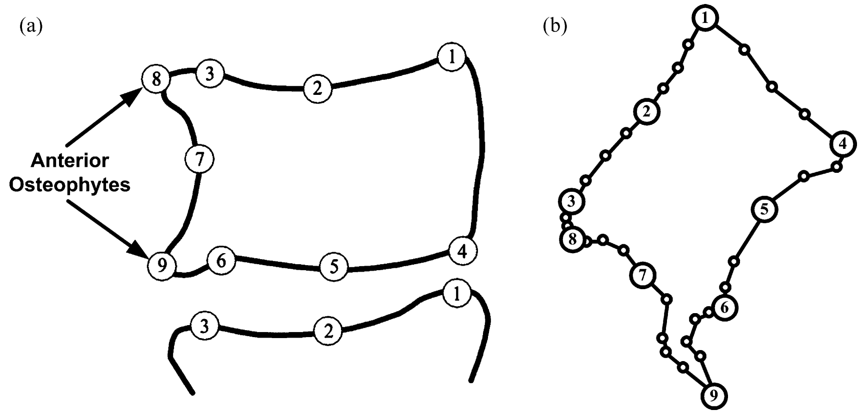
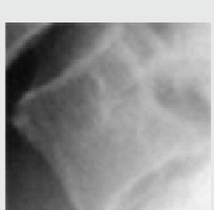
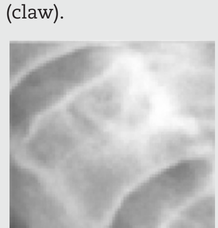


Fig. 2 – Illustration showing the (a) 9-point model and the (b) 36-point model for vertebral shape description. On both models, points 8 and 9, if not coincident with points 3 and 6, respectively, indicate the existence of osteophytes.

the second survey, NHANES II, which was collected between 1976 and 1980 and featured over 20,000 participants. Each participant's record includes over 2000 textual data points such as health questionnaire answers, anthropometric information, and results from a physical exam (Table 1). This textual data is stored in a MySQL relational database. Supplementing the textual data is a collection of 17,000 cervical and lumbar spine X-ray images that were taken from patients aged 25–74. These images were originally on film and subsequently digitized using a 146 dpi scanner resulting in 140 Gigabytes (GB) of data [12]. The collection is considered valuable to radiologists, bone morphometrists and researchers interested in osteoarthritis and musculoskeletal diseases, and medical educators. A panel of domain experts reviewed a sample of the data and identified 23 key biomedical features exhibited in the X-ray images. Of these, anterior osteophytes (AO), subluxation in the cervical spine and spondylolisthesis in the lumbar spine, and disc space narrowing (DSN) were determined to be frequently occurring and reliably detectable using vertebral shape characteristics. However, consistently identifying images that exhibit particular types and severities of these pathologies is extremely tedious for such a large collection. SPIRS and related tools [13], developed to analyze the images and populate the database, automate many tasks and enable retrieval of the images using sketches of the vertebral boundary with the desired feature. While the intent of the NHANES dataset is not clinical decision support, the combination of health survey and imaging data is useful in retrieving examples of abnormal vertebral images for research and education. For instance, SPIRS can be used to determine what types of features (e.g., protrusion on the anterior edge of the cervical vertebra) are consistently associated with a particular symptom (e.g., neck discomfort).

This paper focuses on identifying images in the NHANES II dataset that contains AOs using SPIRS. AOs typically appear on the two anterior "corners" in the sagittal view as bony protrusions, shown in Fig. 2. Depending on their appearance, AOs may be classified into one of two groups and assigned a grade of severity. The Macnab classification assigns an osteophyte as either a claw, which occurs when a spur of triangular shape rises from the vertebral rim and curves toward the adjacent disk, or a traction, which protrudes horizontally and does not extend across the inter-vertebral disc space. The different grades of severity are: slight, moderate, and severe, depend-

Table 2 – List of criteria used to classify the degree of severity for an AO.

Severity	Slight	Moderate	Severe
AO severity grading criteria			
Features	No narrowing or a $<15^\circ$ angle by the AO from the expected normal anterior face of the vertebra or protrusion's length begin <1/5 of the vertebra width (traction) or height (claw).	Mild narrowing or a [15–45°] angle by the AO from the expected normal anterior face of the vertebra or protrusion's length being (1/5–1/3) of the vertebra width (traction) or height (claw).	Sharp/sever narrowing or a $\geq 45^\circ$ angle by the AO from the expected normal anterior face of the vertebra or protrusion's length being >1/3 of the vertebra width (traction) or height (claw).
Example image			

Please cite this article in press as: W. Hsu, et al., SPIRS: A Web-based image retrieval system for large biomedical databases, Int. J. Med. Inform. (2008), doi:10.1016/j.ijmedinf.2008.09.006

ing on the degree of narrowing as depicted in Table 2 [14–16]. When the user provides SPIRS with a query shape, it is used to retrieve other similar shapes in the NHANES database along with associated text data and Macnab classification.

The SPIRS framework is also being used to query an image database of cervicographic images (cervigrams) created by the National Cancer Institute (NCI) and the National Library of Medicine (NLM) for the study of uterine cervix cancer [17]. The database contains approximately 100,000 cervigrams taken as part of the *Guanacaste* and *ALTS* projects that study the natural history of human papillomavirus infection and cervical neoplasia. In addition to cervigrams, correlated clinical, cytologic, and molecular information are also available. Unlike NHANES II, the intent of this dataset is to be used in clinical practice as a method for identifying whether patients have precursors to cervical cancer given the appearance of their cervigram.

2.2. *Segmentation*

In our continuing research, we have explored and implemented various algorithms to extract, index, and retrieve shapes of vertebral bodies from X-ray images. Segmentation is the process of characterizing features (e.g., shape, color, texture) that can be represented in a unique form that is useful for CBIR. In spine X-ray images, the vertebral shape is an image characteristic of high significance. We have considered the combination of manual and semi-automated algorithms including active contour segmentation, LiveWire, active shape models (ASMs), and hierarchical segmentation [12] to obtain the boundaries of the vertebral body. Here, we examine two types of hierarchical segmentation algorithms: *Zamora hierarchical segmentation* [18], which is based on an ASM but captures the shape of bone spurs by using a combination of Generalized Hough Transform (GHT), and a deformable model (DM), and *Howe hierarchical segmentation* [19], which combines the GHT with two active appearance model (AAM) steps applied initially to the entire set of vertebra and then to each individual vertebra. The performance of these algorithms is presented in Section 4.

Automated segmentation methods are necessary to make CBIR practical because the time required for manual segmentation would be prohibitive for large datasets such as NHANES II. However, their outputs need to be comparable to those done manually by an expert. To this end, an evaluation tool has been developed called PathVa (Pathology Validation) [20], which is used to collect ground truth data. Content experts log into the imaging database, review images, mark the pathology data, and validate or create boundary segmentations. The tool has also been useful in finding other pathologies inside the NHANES II dataset, such as spinal stenosis, which was not originally studied from the dataset.

While the aforementioned algorithms are specific to extracting shape boundaries, different image characterization methods may be used to extract other features such as color and texture. For example, in uterine cervix images, color tone and surface texture are two major features used by physicians to differentiate tissue types and to identify the developing stage of cervical neoplasia. Therefore, in addition to defining the boundaries of different regions, these images are characterized using a combination of color moments descriptor, Discrete Wavelet Transform, and lesion size.

2.3. *Indexing and retrieval*

Once segmented, the vertebral shape boundaries are then treated as closed polygons and represented in a variety of forms, such as Polygon Approximation, Fourier Descriptors, or geometric shape properties [21]. Traditionally, these representations are stored as a feature vector and compared individually with the query. However, for large image archives, this linear comparison approach would be impractical. SPIRS therefore uses a coordinate tree to efficiently index shapes and optimize retrieval time. Embedded with the indexing process are appropriate distance measures, which identify matching shapes [22]. Several similarity measures such as Procrustes distance, Fourier descriptors, shape features, invariant moments, polygon approximation for tangent space matching, and token evaluation in multi-scale space have been implemented for matching complete vertebral shapes [23]. However, these algorithms have performed with only 56% accuracy in retrieving pathologically relevant vertebrae images [24]. To improve retrieval results, a partial

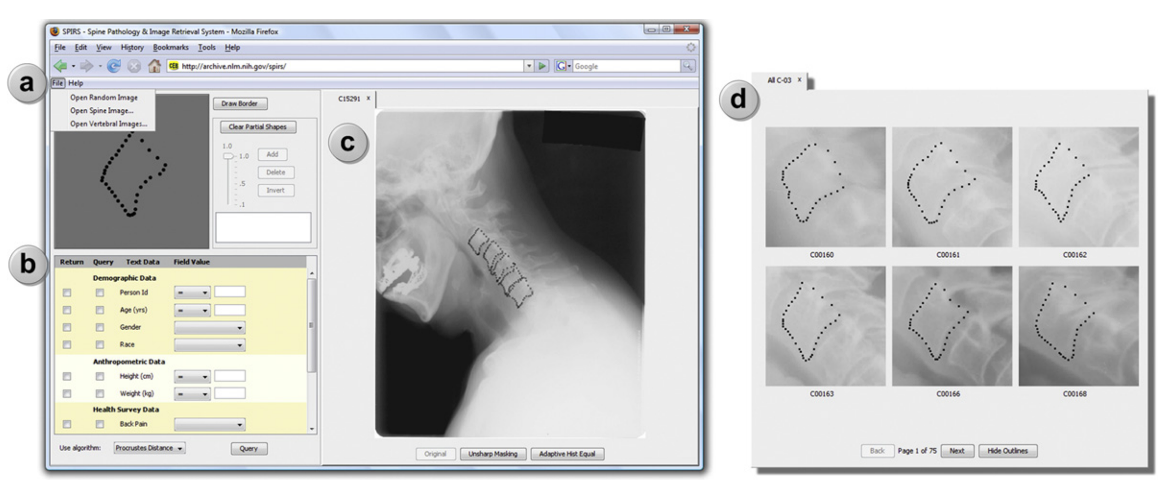


Fig. 3 – A screen capture of the SPIRS client applet (left) and a cropped screen capture of the overall results view tab shown in (d). The default view consists of: (a) the query menu; (b) query editor; and (c) overall view.

Please cite this article in press as: W. Hsu, et al., SPIRS: A Web-based image retrieval system for large biomedical databases, Int. J. Med. Inform. (2008), doi:10.1016/j.ijmedinf.2008.09.006

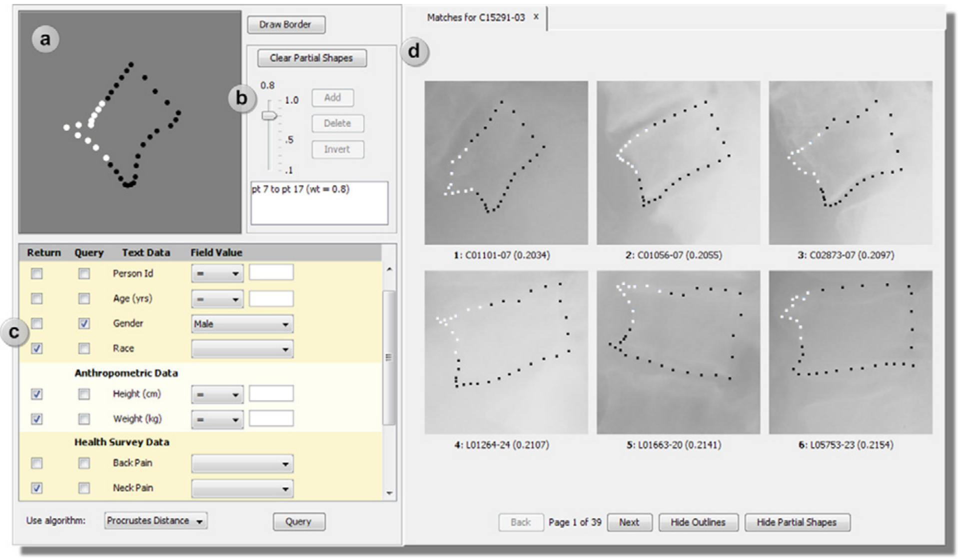


Fig. 4 – Screen capture of SPIRS performing a partial shape query. The user selects and modifies a shape in the (a) visual query component and selects a subset of points in the shape to emphasize in the query using the (b) partial shape query tool. The user can also specify any textual keywords and values to be returned using (c) text query component. A series of shapes that are similar to that of the query shape, with the partial shape emphasized in a lighter color (made white in this image for clarity), are returned and shown in (d) the cropped view. .

shape matching algorithm has been explored with Procrustes distance and Fourier descriptors; this algorithm is described further in Section 3.2.

2.4. *Client applet*

The client, shown in Fig. 3, is a Java applet that runs in a Web-browser and provides the user with the necessary functionality to query for particular shapes of interest and interact with the results. The applet has been developed with the following considerations: (i) it is a thin client that provides the look-and-feel of a standalone application but uses minimal computing resources; (ii) it does not require installation of files onto the local drive; (iii) is easily modifiable without affecting core-CBIR functionality; and (iv) it is platform independent. The client interface is comprised of the following components, described below.

- *Query Menu.* A potential difficulty in using a retrieval system based on a query-by-example paradigm is that users may need assistance in finding an appropriate starting image or shape as the basis of a query. SPIRS addresses this problem, known as the *page zero problem* [21], by providing a Query Menu (Fig. 3a), which allows users to find the desired vertebral shape by either retrieving a random image, specific patient image, or a particular type of cervical or lumbar vertebra (e.g., third cervical vertebra, C3). A query vertebral shape can be selected from the returned X-ray images in the Overall View or from the selection of vertebral image crops in the Cropped View shown in Fig. 3d.
- *Overall View.* The Overall View (Fig. 3c) displays the entire patient X-ray and allows the user to examine and select any of the segmented vertebral shapes. In addition to the unprocessed version of the scanned image, two forms of enhanced images are also available for improved visualization of subtle detail, viz., one processed with unsharp masking, which enhances edges, and the other an adaptive histogram-equalized image, which improves contrast. The user may use the mouse to hover over and view the vertebral outlines as image overlays. A mouse click on the desired shape selects and transfers it to the query editor.
- *Cropped View.* The Cropped View (Fig. 3d) allows a user to inspect multiple vertebral shapes at once. Each displayed vertebra is cropped from the original patient X-ray using the boundary information associated with its shape. The crops are then normalized so that each vertebra is facing in the same direction, which simplifies the comparison between various vertebrae.
- *Query Editor.* The Query Editor (Fig. 3b) enables the user to pose queries using a query vertebral shape and/or using textual information pertaining to health history, anthropometric data, quantitative imaging parameters, and demographic data. The visual query component of the editor is a canvas where the user may choose to draw an entirely new shape or edit points that have been imported from an existing shape found using the Query Menu. The Query Editor also supports multiple partial shape query specification. This feature provides the functionality to select parts of the vertebral shape enabling the algorithms to focus on these boundary intervals, which may exhibit significant pathology. Partial shape querying is shown in Fig. 4 as the highlighted interval along the vertebral boundary. The textual query component allows users to specify expressions and keywords that filter the results based on survey data captured as part of NHANES II. Presently, potential text queries include the patient's gender, race, height,

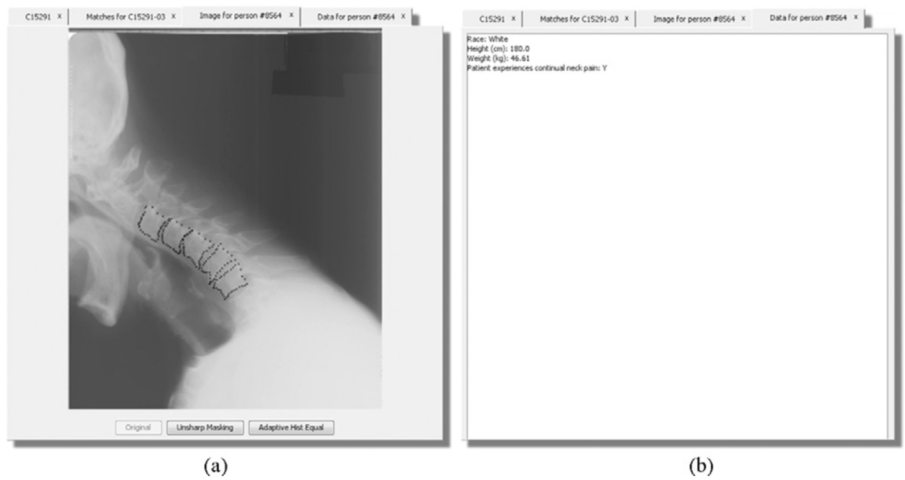


Fig. 5 – Upon selecting a particular shape of interest from the returned shapes, the user is presented with (a) the original patient image from which that shape was segmented and any associated textual data that the user may have selected (b).

weight, and quantitative imaging parameters such as spacing between discs; additional terms from the survey will be added in the future. The system supports relational operators (e.g., greater than, less than), which provides flexibility in querying a range of values (e.g., return all matching results for patients who are older than 55 years old). Prior to executing a query, the user selects a retrieval algorithm from a list of implemented algorithms with trade offs between performance and accuracy.

- *Results View.* The Results View offers two displays. The Cropped View (Fig. 3d) shows returned matching vertebrae, while the Overall View (Fig. 3c) appears when a user selects one of the resulting vertebral shapes. In the Overall View, the entire X-ray of the patient is shown with the matching vertebra highlighted. If the user selects certain text fields from the text database to be returned as part of the query, a separate tabbed window appears, presenting the requested information. The information displayed may be customized by selecting the appropriate checkboxes in the text query component of the Query Editor. A unique feature of this system is the ability to use resulting shapes as queries. This feature can be considered as a form of iterative relevance feedback by enabling the user to specify a result that best matches the search criteria and use it as the basis for a new query. Relevance feedback performs analysis on user interaction with the system to automatically give more weight to certain results that may be of interest to the user. More advanced relevance feedback methods (*e.g.,* linear weight-updating approach [22], addition of a short-term memory model to cache images with positive feedback [23]) have been explored as standalone MATLAB programs and are under development for use with the Web-based system.

2.5. *Gateway*

SPIRS utilizes standard Web communication interfaces to enable interaction with its components. The gateway is a Java servlet (e.g., Apache Tomcat) that acts as a mediator between clients, which could be applets or remote applications, and server-side components. It manages multiple simultaneous connections (users) as a separate session and queues requests to the core-CBIR engine. The servlet automatically generates the appropriate statements in structured query language (SQL) query syntax for interacting with a MySQL database. It also is responsible for packaging and transmitting responses back to the requesting client. Other types of databases may be supported by customizing the gateway configuration file to support the different database schema and query syntax. In addition, the core architecture of the gateway makes interaction with other geographically distributed CBIR systems possible: recent work has combined the IRMA[4] system, which is a system located in Germany for CBIR of medical images using overall (global) image intensity features with the SPIRS shape matching engine [10,25]. While it appears as if the IRMA system provides shape query support for subtle localized pathology, in actuality, its interface is using the SPIRS server located at the NIH in the USA for computing shape-similarity of spine X-ray images. Queries executed within IRMA are routed to SPIRS via the gateway. Information is transmitted between systems using XML (extensible markup language). The entire transaction is divided into three primary events—*querystatus, query,* and *queryresult*. Each element in the XML file is designed for a particular event. The <*querystatus*> element is used to determine whether a desired service is available and to obtain a list of currently available services. The <*query*> element contains information about shape queries such as the vertebra contour, partial shape indices and weights, maximum number of results desired, and range of similarity result scores. SPIRS responds with the <*queryresult*> element populated with matching image-vertebra tuples and similarity scores.

3. Querying

Although much CBIR research has focused on overcoming technical challenges such as developing accurate and efficient

[4] IRMA Web site: http://www.irma-project.org/index_en.php.

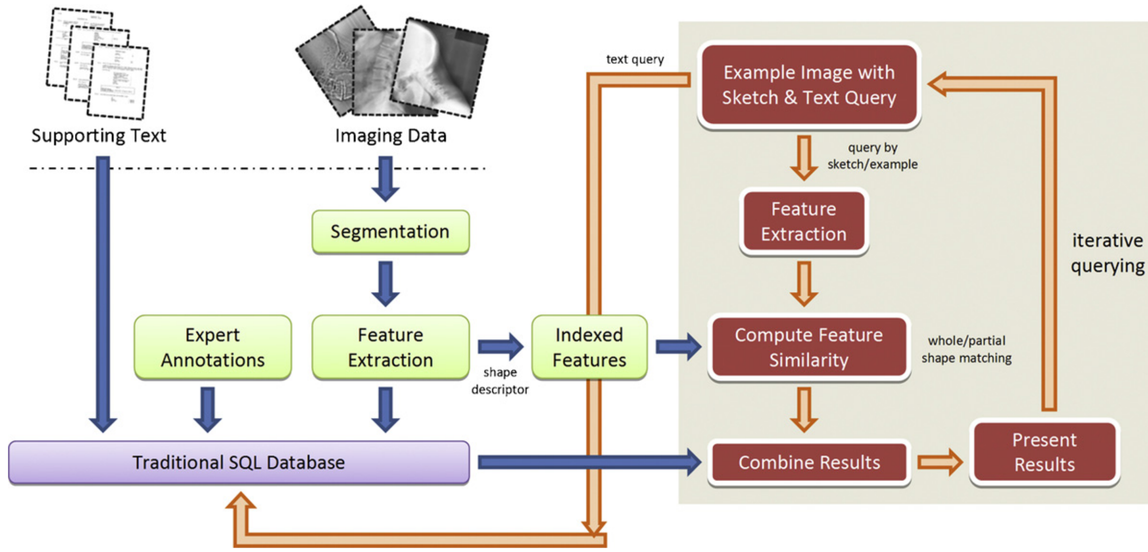


Fig. 6 – The query handling process (shaded region) in the context of the overall retrieval system.

retrieval algorithms, in order for widespread CBIR adoption and use, systems need to provide users with the ability to exploit the capabilities of these algorithms in managing large biomedical multimedia databases. With this goal, we believe that SPIRS allows users to pose meaningful queries that have both clinical and research applicability. The following sections describe how users pose queries and interact with the results.

3.1. *Query handling*

Figs. 4 and 5 illustrate how users interact with the system to pose a query. In this particular example, a researcher is interested in finding any association as to whether male participants in the survey data who exhibit a particular protrusion on the anterior of their cervical vertebra complain of neck pain. The user would first utilize the Query Menu to find the initial shape of interest, whether by using a random image, selecting a shape from a specific participant's image, or specifying a specific vertebra type and number (e.g., C3). Once a query shape has been selected, the user may modify the shape to accentuate a particular feature (e.g., move the points to create the desired protrusion) in the Query Editor (Fig. 4a) and demarcate a subset of the shape to emphasize in the query using the partial shape functions (Fig. 4b). The user may also input additional textual and numerical data to further constrain the returned results such as limiting the results to only male patients (Fig. 4c). The user also selects the desired parameters to be returned. Once the query is executed, the results are displayed as cropped images depicting the shape and sorted by similarity where smaller values indicate that the shape is more similar to the query shape (Fig. 4d). Selecting a particular shape returns the entire patient image (Fig. 5a) from which the cropped image was extracted. If the user selected to have patient data returned with the query, a separate tabbed pane (Fig. 5b) provides that information (e.g., whether the patient experiences continual neck pain). The user may iteratively pose queries by selecting a returned shape and using that shape as the basis of a new query.

3.2. *Partial shape matching*

A problem with many current CBIR approaches for medical images is that they operate on the image or region of interest (ROI) as a whole. Whole shape matching methods have been demonstrated to be more effective in identifying shapes with gross similarity because these algorithms operate over the entire shape. They are ineffective, however, in matching subtle shape characteristics, which may have critical pathology. It has been observed that pathologies of interest on a vertebral outline are often localized along a short interval on the boundary. For example, osteophytes are only expressed along the “corners” on the vertebral boundary as seen in the sagittal view. Partial shape matching (PSM) has been implemented in SPIRS allowing users to sketch or identify only the local interval of interest on the vertebral boundary. For instance, consider a user interested in identifying patients in the NHANES II data with a claw AO. In the query editor, the user would highlight the subset of points that comprise the spur, have the capability of altering its appearance by dragging the shape points, weight these points greater than the other points in the shape, and execute the PSM query. PSM has been demonstrated as an effective method to classify the severity of a pathology depicted in the query shape [23].

3.3. *Query execution*

Fig. 6 illustrates how the system handles a query and how query execution integrates with the overall CBIR system. When the user formulates a hybrid text-image query, the query is separated into its visual and textual components. Relevant features and annotations are extracted from the visual component and matched against the indexed features using specific similarity computation methods. Parameters from the textual component are executed against the database containing survey and patient data. The results of these individual searches are then combined before the results are presented to the user. The set of returned results consists of the intersection of the set of spine images with matching vertebrae shape profiles and the set of matching responses to text query components. For example, a user who is researching the correlation between certain shape features and the degenerative spine disease spondylosis may be interested in the intersection of results between a visual query depicting a traction spur and survey participants who have been diagnosed with spondylosis.

4. Results

To ensure the accuracy and utility of SPIRS in the research environment, we have evaluated the system in two parts: first, each component is evaluated individually, and then the components are evaluated as a complete framework. While many of the results for the individual components have been reported in previous publications, we summarize their results here. We then present results of the overall system in retrieving similar vertebral shapes and describe preliminary results showing how the SPIRS framework can be used to search uterine cervix images.

4.1. *Validation of segmentations*

Accurate segmentations are critical to the system's ability to capture subtle features that may be used for retrieval. In this work, two algorithms were explored: Zamora and Howe. The Zamora algorithm was tested using a random subset of 100 cervical and 100 spine images from the NHANES dataset. Success is defined as whether the converged solution is near (20 pixels for cervical spine, 50 pixels for lumbar spine) the actual boundary. While this measure is subjective and relies on the visual judgment of a domain expert, it is still useful for gauging the general abilities of the algorithm. It achieved 75% success for cervical spine and up to 49% for lumbar spine. For the Howe algorithm, the AAM was trained using a leave-one-out strategy (train on 99 images, test on 1 image) and achieved 65% accuracy for cervical and 67% for lumbar vertebra.

4.2. *Indexing performance*

The performance of the indexing scheme using kD-tree was compared with a metric tree. A sample of 2812 shapes was divided into groups of random sizes (434, 902, 1654, and 2812). Each set was indexed in the original shape space using a metric tree and optimal embedding into a kD-tree. After indexing, every shape was used as queries and the k-nearest neighbor vertebral images were retrieved using the Euclidean shape distance. Retrieval using a kD-tree consistently outperformed that of the metric tree, but both approaches were sub-linear in complexity [26].

4.3. *Overall evaluation*

To evaluate the overall performance of the system, three board certified radiologists were asked to (i) review and label a subset of 200 vertebrae for anterior osteophytes using the classifications and degree of severity discussed previously; (ii) generate 24 queries using exemplar cases from this set; and (iii) on (unranked) set of vertebrae that may be considered "similar" on the particular expressed pathology on each query vertebral outline. The evaluation studied the top 10 vertebral shapes returned by SPIRS and compared their pathology classification to the query shape classification. The retrieval algorithms achieved 68% relevance (precision and recall) when querying for specific osteophyte type (claw or traction) or severity (slight, moderate, severe) using PSM based on Procrustes distance method. The shortcomings in the performance have been linked: (i) to erroneous determination of the query semantics and (ii) to limitations of the shape matching algorithm. A corner-guided PSM using dynamic programming in initial experiments has shown a precision of 100% for the top match that gradually drops down to 85% as larger number of matches are considered [27]. In addition, SPIRS allows users to select a matching result as the basis of a new query. Traditionally, user interactivity has helped in minimizing similar problems with text retrieval, and user feedback has often been analyzed and employed to improve retrieval relevance. By allowing users the option to iteratively query and refine results, SPIRS implements a basic form of relevance feedback, which will be enhanced with a novel advanced weighted hierarchical feedback method using short-term memory [28] and other approaches as they are migrated from laboratory prototype routines. To quantify the improvements of using relevance feedback [28], eight queries were selected to have a user perform relevance feedback on a subset of 2000 shapes from the NHANES II dataset. The retrieval accuracy (number of relevant retrieved images over the total number of retrieved images) was calculated for all eight selected queries before and after each feedback refinement for the first three feedback iterations. An additional 22% performance improvement was observed through use of the linear updating approach after three iterations.

4.4. *Cervigram retrieval*

Work is also ongoing to evaluate the feasibility of using the SPIRS framework with the uterine cervix images. A preliminary study involves retrieving images in the database that contain the same region type as the query image. The SPIRS implementation was modified to search for similarity based on color, texture, and lesion size rather than shape. The user interface was also customized to incorporate additional tools for working with cervigrams. A set of 120 images with 422 expert-marked regions (e.g., acetowhite, blood, mucus) was presented to the system, and the fraction of the top five returned images that contained the same region type was computed. The system was able to differentiate between three types of tissue (acetowhite, blood, columnar epithelium) with up to 64% accuracy [29]. While additional improvements are needed to address the algorithm's ability to distinguish different tissues that have similar appearances, the evaluation demonstrates the feasibility of using the SPIRS framework in this domain.

5. Discussion

Tools for image management and CBIR systems that not only minimize the semantic gap [30] but also are sensitive to depicted pathology are increasingly necessary to interact with the growing volume of biomedical imaging data. In CBIR systems, images are characterized either using their content, which is captured through measurements of low level features such as color, texture, shape (edges), or by their location relative to some identifiable entity on the image. These image descriptors are quantified using a variety of particular techniques that attempt to ensure some basic properties of similarity:

- *Identity*: A database entity, if used as a query, should be found among the set of highly relevant matches.
- *Separability*: Features should be able to distinguish one object from another.

The descriptors can be applied at different scales that can be broadly categorized as *global* or *local* descriptors. Global descriptors are those that attempt to capture the visual content characteristic of the entire image on the assumption that *sufficient* separation can be achieved between different entities in the image database on those descriptors alone. Local descriptors, on the other hand, can be applied at various other image "levels" or "regions of interest" and attempt to not only separate images on the computed characteristic, but also support localized queries depending on the level of detail desired. Historically, global features have been used to characterize images in systems with a heterogeneous collection of images such as IRMA. However, when images result from a study that focuses on a particular aspect it is likely that the resulting database will be homogeneous. For example, the NHANES II collection has two kinds of images (cervical and lumbar spine) that are very similar within each group and require localized descriptors at various levels of detail to distinguish amongst them. Another example is ASSERT, which showed that the use of local features significantly improved the retrieval performance of HRCT images with pathologies similar to the query image [31]. This distinction between global and local descriptors is also applied to the vertebral outlines in the case of spine images and the AO pathology. Whole shape feature extraction algorithms are a type of global descriptors at this image scale and partial shape feature extraction algorithms are a type of local descriptors attempting to capture subtle pathology expressed along the vertebral shape.

Traditionally, computer-aided diagnosis (CAD) systems classify abnormalities in an image and assist in determining their degree of severity. This aspect is beginning to appear in CBIR systems as segmented (extracted) features are used for pathology classification on the images in the database as well as the query image. Unlike traditional text databases, however, an image database is often a collection of images stored on a file server. Relevant fields in text traditional text databases hold pointers to these images. This configuration, commonly found in PACS, RIS, and HIS, is challenging for use with CBIR applications. One reason is that the image "indices" are stored as long strings of numbers called feature vectors (segmented features and feature measurements). These "indices" are entities on which further computation is necessary to determine the degree of similarity between a query image and a candidate on the database. Several approaches have been proposed for indexing multidimensional vectors and include coordinate trees and metric spaces [30]. Our approach has been to use coordinate trees in SPIRS while we continue to evaluate other approaches.

In spite of the acknowledged importance of CBIR, several shortcomings in current approaches have prevented their widespread acceptance into medical research, practice, and education. We believe that a biomedical CBIR system should be easily accessible, extensible, and capable of supporting a rich set of segmentation, validation, indexing, query, retrieval, and visualization methods developed using open software and standards. With the SPIRS framework, we have implemented a system that incorporates these qualities to demonstrate how such system would perform on a large biomedical dataset such as NHANES II and the uterine cervix image database. In order to support unique requirements of different biomedical images it is necessary to either: (i) provide the capability to combine this system with others (possibly geographically distant) that have complementary features; or (ii) make the framework modular such that other databases and retrieval algorithms may be integrated. The SPIRS framework attempts to address these requirements through its client-server design, customizability of individual components (e.g., gateway, segmentation algorithms), use of a XML-based service protocol, and implementation in Java. The applet makes use of the Java Image I/O programming interface. While images in the NHANES II database are stored as JPEG (Joint Photographic Experts Group) format, SPIRS is capable of handling other common medical imaging formats such as DICOM (Digital Imaging and Communications in Medicine) provided the appropriate file reader packages. While most components in the framework require few changes to support different domains, the user interface needs the most customization to best suit a particular context. For example, in spine retrieval, the interface should be geared towards drawing and viewing vertebral shape boundaries. However, for cervical images, the interface should provide tools for identifying various cervix regions and weighing which features (color, texture, size) affect the results. Despite the advances that SPIRS has made, CBIR systems have yet to make inroads into routine clinical, biomedical research, or educational use. To better understand the underlying reasons, [32] presents a gaps-based framework to evaluate CBIR systems based on various characteristics broadly grouped into areas such as content, feature, performance, and usability. [33] details the results of evaluating SPIRS using this framework.

This work describes the basic implementation of the SPIRS framework and how it is being used in applications such as vertebral shape and how it can be extended to support uterine cervix image retrieval. Plans for SPIRS include a more exhaustive performance evaluation on a substantially larger dataset marked by multiple experts. In addition, a usability study will be conducted for ease of use, quality of expected outcome, value of SPIRS capability of correlating text and image data. As evidenced by the segmentation results, fully automated methods are still insufficient to accurately capture vertebral shapes (particularly lumbar spine) and osteophytes. In addition, overcoming the subjective evaluations such as those done to gauge the success of segmentations is an ongoing research topic [34]. While we believe the development of PathVa provides the evaluating domain expert with the tools to make the best determination of whether the segmentation is accurate, more objective quantitative measures are needed. Furthermore, using human experts judge segmentations or classify vertebral shapes introduces an element of potential bias. Currently, we address this issue by having a consensus opinion from at least two expert radiologists examining the vertebral shapes.

6. Conclusions

SPIRS is a Web-based distributed content-based image retrieval framework that supports hybrid visual and text queries, and may be applied to various applications in medicine. It implements novel shape representation and similarity matching embedded with an index tree that allows efficient retrieval. SPIRS aims to capture query semantics through support of advanced mechanisms like multiple partial shape matching and iterative querying that provides simple yet effective relevance feedback to the system. SPIRS is built using open standards and is simultaneously developed as a service, which enables its integration with other complementary information retrieval systems. Initial evaluation of the system has shown that the combination of partial shape matching and relevance feedback significantly improves the system's ability to retrieve similar results. Since its original release in August, 2006, SPIRS has undergone several revisions that have added support for additional retrieval algorithms (i.e., Fourier descriptors) and an improved, cleaner user interface for posing visual and text queries. Future goals for the project include: (i) additional data collection, assimilation, and validation; (ii) system feature enhancements; (iii) improved retrieval quality by learning from user feedback; and (iv) improved user interaction and visualization. Extracting and validating vertebral shapes is an ongoing process that will (at regular intervals) add to the 7200 shapes currently indexed. Planned feature enhancements include integration of a generalized shape segmentation toolbox, which is a currently a standalone application, incorporation of additional similarity algorithms, support for query of multiple vertebral shapes and disc space narrowing pathology, relevance feedback, and improved visualization algorithms.

Acknowledgements

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and Lister Hill National Center for Biomedical Communications (LHNCBC). The authors thank Dr. Alex Bui at the UCLA Medical Imaging Informatics Group for his valuable feedback.

Summary points

What is known on the topic

- Majority of current retrieval systems in clinical use rely on text keywords such as Digital Imaging and Communications in Medicine (DICOM) header information to perform retrieval.
- Content-based image retrieval has been widely researched in a variety of domains and provides an intuitive and expressive method for querying visual data using features such as color, shape, and texture.
- Current CBIR systems are not easily integrated into the healthcare environment; have not been widely evaluated using a large dataset; and lack the ability to perform relevance feedback to refine retrieval results.

What this study adds to our knowledge

- SPIRS implements a lightweight, Web-based CBIR system that is extensible and can be used with a variety of image types such as vertebral shapes and uterine cervix images.
- The distributed architecture and conformance to open standards allows the SPIRS framework to share functional resources with geographically separate CBIR systems such as IRMA.
- SPIRS enables allows users to pose hybrid visual and text queries to a large biomedical database for research, education, and patient care.

Contributions: *William Hsu* contributed to the initial design and implementation of the Web-based interface and gateway. He is the primary author of the MEDINFO 2007 submission. *Sameer Antani* Ph.D. is the lead researcher on the Content-Based Image Retrieval project at the National Library of Medicine, NIH. He has been directly involved with the reported research with several components developed by him, or under his direction or mentorship. *L. Rodney Long* M.A. jointly participates with Dr. Antani on several aspects of the project which include but are not limited to image segmentation, image transmission, and performance evaluation. *Leif Neve* is a developer on the project and directly contributes to the development, maintenance, and enhancements in SPIRS. *George R. Thoma* Ph.D. is the Chief of the Communications Engineering Branch and jointly participates on the project with Mr. Long and Dr. Antani. He contributes to the reasearch and provides guidance on biomedical imaging projects at the National Library of Medicine, NIH.

REFERENCES

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

[2] C.W. Niblack, R. Barber, W. Equitz, M.D. Flickner, E.H. Glasman, D. Petkovic, et al., QBIC project: querying images by content, using color, texture, and shape, in: Proceedings of SPIE, 1993, vol. 1908, 1993, pp. 173–187.
[3] J.R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, et al., Virage image search engine: an open framework for image management, in: Proceedings of SPIE, vol. 2670, 1996, pp. 76–87.
[4] C. Carson, M. Thomas, S. Belongie, J.M. Hellerstein, J. Malik, Blobworld: A system for region-based image indexing and retrieval, in: Third International Conference on Visual Information Systems, 1999, pp. 509–516.
[5] H. Müller, N. Michoux, D. Bandon, A. Geissbuhler, A review of content-based image retrieval systems in medical applications—clinical benefits and future directions, International Journal of Medical Informatics 73 (1) (2004) 1–23.
[6] P.M. Kelly, T.M. Cannon, Experience with CANDID: comparison algorithm for navigating digital image databases, in: Proceedings of SPIE, vol. 2368, 1995, p. 64.
[7] C.E. Chronaki, X. Zabulis, S.C. Orphanoudakis, I2Cnet medical image annotation service, Medical Informatics 22 (4) (1997) 337–347.
[8] C.R. Shyu, C.E. Brodley, A.C. Kak, A. Kosaka, A.M. Aisen, L.S. Broderick, ASSERT: A physician-in-the-loop content-based retrieval system for HRCT image databases, Computer Vision and Image Understanding 75 (1) (1999) 111–132.
[9] M.O. Güld, C. Thies, B. Fischer, T.M. Lehmann, A generic concept for the implementation of medical image retrieval systems, International Journal of Medical Informatics 76 (2–3) (2007) 252–259.
[10] S.K. Antani, T.M. Deserno, L.R. Long, G.R. Thoma, Geographically distributed complementary content-based image retrieval systems for biomedical image informatics, in: Proceedings of 12th World Congress on Medical Informatics (MEDINFO), vol. 12 (Pt. 1), 2007, pp. 493–497.
[11] T. Catarci, M.F. Costabile, S. Levialdi, C. Batini, Visual query systems for databases: a survey, Journal of Visual Languages and Computing 8 (2) (1997) 215–260.
[12] L.R. Long, S.K. Antani, G.R. Thoma, Image informatics at a national research center, Computerized Medical Imaging and Graphics 29 (2–3) (2005) 171–193.
[13] S. Antani, L.R. Long, G.R. Thoma, Content-based image retrieval for large biomedical image archives, in: Proceedings of 11th World Congress on Medical Informatics (MEDINFO), 2004, pp. 7–11.
[14] M.H. Heggeness, B.J. Doherty, Morphologic study of lumbar vertebral osteophytes. Grand rounds, Southern Medical Journal 91 (2) (1998) 187–189.
[15] D. Pate, J. Goobar, D. Resnick, P. Haghighi, D.J. Sartoris, M.N. Pathria, Traction osteophytes of the lumbar spine: radiographic-pathologic correlation, Radiology 166 (3) (1988) 843–846.
[16] I.A.N. Macnab, The traction spur: an indicator of segmental instability, The Journal of Bone and Joint Surgery 53 (4) (1971) 663–670.
[17] Z. Xue, L.R. Long, S. Antani, J. Jeronimo, G.R. Thoma, A web-accessible content-based cervicographic image retrieval system, in: Proceedings of SPIE, vol. 6919, 2008, pp. 691907 (1-9).
[18] G. Zamora, H. Sari-Sarraf, L.R. Long, Hierarchical segmentation of vertebrae from X-ray images, Proceedings of SPIE, vol. 5032 (2003) 631–642.
[19] B. Howe, A. Gururajan, H. Sari-Sarraf, L.R. Long, Hierarchical segmentation of cervical and lumbar vertebrae using a customized generalized Hough transform and extensions to active appearance models, in: Proceedings of the 6th IEEE Southwest Symposium on Image Analysis and Interpretation, 2004, pp. 182–186.
[20] S. Antani, J. Cheng, J. Long, L.R. Long, G.R. Thoma, Medical validation and CBIR of spine X-ray images over the Internet, in: Proceedings of SPIE, vol. 6061, 2006, pp. 60610J (1-9).
[21] S.K. Antani, M. Natarajan, J.L. Long, L.R. Long, G.R. Thoma, Developing a comprehensive system for content-based image retrieval of image and text from a national survey, SPIE Medical Imaging 5748 (2005) 152–161.
[22] R.J. Stanley, S. Antani, R. Long, G. Thoma, K. Gupta, M. Das, Size-invariant descriptors for detecting regions of abnormal growth in cervical vertebrae, Computerized Medical Imaging and Graphics 32 (1) (2008) 44–52.
[23] I.L. Dryden, K.V. Mardia, Statistical Shape Analysis, John Wiley & Sons Ltd., West Sussex, England, 1998.
[24] S. Antani, D.J. Lee, L.R. Long, G.R. Thoma, Evaluation of shape similarity measurement methods for spine X-ray images, Journal of Visual Communication and Image Representation 15 (3) (2004) 285–302.
[25] S.K. Antani, T.M. Deserno, L.R. Long, M.O. Guld, L. Neve, G.R. Thoma, Interfacing global and local CBIR systems for medical image retrieval, in: Proceedings of the Workshop on Medical Imaging Research (Bildverarbeitung fur die Medizin), 2007, pp. 166–171.
[26] X. Qian, H.D. Tagare, Optimal embedding for shape indexing in medical image databases, in: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention, vol. 8 (pt 2), 2005, pp. 377–384.
[27] X. Xu, D. Lee, S. Antani, L.R. Long, A spine X-ray image retrieval system using partial shape matching, IEEE Transactions on Information Technology in Biomedicine 12 (1) (2008) 100–108.
[28] X. Xu, D.J. Lee, S. Antani, L.R. Long, Relevance feedback for spine X-ray retrieval, in: Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems, 2005, pp. 197–202.
[29] Z. Xue, S.K. Antani, L.R. Long, J. Jeronimo, G.R. Thoma, Investigating CBIR techniques for cervicographic images, in: Proceedings of the Annual Symposium of American Medical Information Association (AMIA), 2007, pp. 826–830.
[30] H. Samet, Foundations of Multidimensional and Metric Data Structures, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.
[31] C. Shyu, C. Brodley, A. Kak, A. Kosaka, A. Aisen, L. Broderick, Local versus global features for content-based image retrieval, in: Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries, 1998, pp. 30–34.
[32] T. Deserno, S. Antani, R. Long, Ontology of gaps in content-based image retrieval, Journal of Digital Imaging (2008).
[33] L.R. Long, S. Antani, T.M. Deserno, G.R. Thoma. Content-based image retrieval in medicine: Retrospective assessment, state of the art, and future directions. International Journal of Healthcare Information Systems and Informatics (2009) 4(1) 1–16.
[34] Y.J. Zhang, A survey on evaluation methods for image segmentation, Pattern Recognition 29 (8) (1996) 1335–1346.

Please cite this article in press as: W. Hsu, et al., SPIRS: A Web-based image retrieval system for large biomedical databases, Int. J. Med. Inform. (2008), doi:10.1016/j.ijmedinf.2008.09.006