SIIM 2008 Abstract Submission Form
Descriptive Abstract
Email to: abstracts@siimweb.org

Abstract Title:
A Unifying Framework for Combining Content-Based Image Retrieval with Relational Database Queries for Biomedical Applications

Topic:
TRIP™

Preferred Presentation:
Paper Presentation (10-12 Mins)

Background:
Biomedical research scientists often conduct large-scale experiments involving multi-modality data such as neuronal recordings, CT scans and fMRI images. The non-textual nature and increasingly large amount of these data necessitates that they be retrieved not only via standard relational database queries over the meta data (such as patient information, reason for examination and date of exam), but also via measures of data similarity (such as image texture or 3D shape). Traditional relational database retrieval methods are designed for efficiency of retrieval and are extremely effective for exact retrieval on alphanumeric data. Content-based image retrieval (CBIR) methods were designed for retrieving images that are similar to a query image according to one or more domain-dependent distance measures [1]. CBIR systems cannot directly use the efficient indexing methods of relational database systems. Our goal is to develop a unified methodology for organization and retrieval of biomedical data from scientific experiments that makes use of both the similarity-based retrieval methodology of CBIR systems and the efficiency of relational database systems.

We have designed a probabilistic database framework that unifies the relational database and content-based retrieval methodologies to provide efficient and effective retrieval of multimedia data for biomedical applications. Our methodology can be summarized as follows. Queries are phrased in an extended SQL that allows both standard database constraints and similarity-based retrieval predicates. These extended-SQL queries are translated into two parts: 1) similarity-based searches for data that match the provided query data and return results in order of similarity in the form of probabilistic relations and 2) a standard SQL query that uses both the standard relations containing meta data and the probabilistic relations to answer the original query. The similarity searches uses image indexing techniques [2] to rapidly return images and signals similar to those provided in the query in order of similarity. Then the translated query combines the results from both the probabilistic relations and standard relations to efficiently answer the query.

Evaluation:
In order to evaluate our methodology, we have implemented distance measures for three different biomedical applications [3]. The first application comes from the University of Washington Eye Laboratory where scientists are studying cataracts of the eye using slit-lens images of mouse eyes and are interested in similarity searches based on features extracted from these images by computer vision techniques [4]. The second application comes from the Pediatric Imaging Research Laboratory of Children's Hospital and Regional Medical Center in Seattle whose researchers study craniofacial disorders in children [5]. They are interested in

shape-similarity measures on 3D skull data obtained from various modalities such as CT scans and stereo cameras. The third application involves the study of language sites in the brain by researchers in the departments of Surgery and Psychology at the University of Washington [6]. This application includes MRI data, fMRI data, and single-unit recording data in the form of time signals. Researchers are interested in similarity-based searches involving multiple different similarity measures.

Motivated by these and other biomedical applications, we have implemented a first-prototype retrieval system that includes all aspects of our design, but is currently limited to queries that invoke only a single similarity-based search and have limited alphanumeric constraint capabilities. Each application has an associated Postgress database that includes the alphanumeric data and contains file names for the signal and image data. Two databases have been constructed so far: one for the mouse eye experiments and one for the skull shape experiments. The mouse eye database contains a mouse relation with meta data on each mouse and an experiments relation containing meta data for each experiment and the file names of the associated slit-lens images. Similarly, the skull database contains a patient relation for meta data on the patients and an exam relation that contains meta data for each exam and file names of three planes from each CT scan. These planes have been used in our previous work [5] to analyze the shapes of the skulls. A graphical user interface allows the user to construct a query. The user first selects a database and is shown the schemas of the relations from that database, in order to choose the fields to be returned by the query. Next he or she selects a query image and is able to add constraints on the meta data. Finally he or she chooses one or more similarity measures, assigns weights to them, and clicks on the Search button. The results that satisfy the constraints are then displayed for the user in the graphical interface in order of similarity to the query image. To evaluate the prototype system, we have developed a questionnaire for users from the different applications to complete after trying out the graphical interface, which is available on the web. The questionnaire contains questions regarding ease of use, ability to construct meaningful queries in the particular domain, and acceptability of the computation time to answer the query. The questionnaire also asks users to suggest improvements for the next version of the system. It does not deal with accuracy of retrieved results, since they have been evaluated in our previous publications [3, 4, 5, 6].

Discussion:

The result of our research so far is a working system implemented within the probabilistic framework we have defined. Initial feedback from our questionnaires indicate that the system is easy to use and the retrieval times are reasonable, but the users are eager for additional capabilities. Our ongoing work is to extend the system to allow multiple similarity retrievals and multiple alphanumeric constraints whose results are combined in our probabilistic framework to produce the final answers to the query. Executed queries will produce results in both XML and HTML as well as displaying results on the graphical user interface

Conclusion:

We have designed a generic framework for combining standard relational database queries with similarity-based queries on any kind of multimedia data. We have shown this framework to be useful for queries from two different biomedical applications. We have implemented a first prototype system to test our framework and designed a questionnaire by which our users can evaluate it. With this beginning system in place, we can now extend the system to more complex queries.

References:

[1] H. Muller, N. Michous, D. Bandon, and A.Geissbuhler.A review of content based image retrieval systems in medical applications – clinical benefits and future directions. International Journal of Medical Informatics, 73:1–23, 2004.

[2] A. P. Berman and L. G. Shapiro. A flexible image database system for content-based retrieval. Comput.Vis. Image Underst., 75(1-2):175–195, 1999.

[3] L. G. Shapiro, I. Atmosukarto, H. Cho, H. J. Lin, S. Ruiz-Correa, and J. Yuen. Similarity-Based Retrieval for Biomedical Applications. Case-Based Reasoning on Signals and Images. P. Perner (Ed.), Springer, 2007.

[4] J. Yuen. Feature-Based Classification of Mouse Eye Images. Undergraduate thesis, University of Washington, 2006.

[5] H. J. Lin, S. Ruiz-Correa, R.W. Sze, M. L. Cunningham, A.V. Hing, M.L.Speltz, and L. G. Shapiro. Efficient symbolic signatures for classifying craniosynostosis skull deformities. In Proceedings of the Workshop on Computer Vision for Biomedical Image Applications, pages 302–313, 2005.

[6] S. Rolfe, An Independent Component Analysis Tool for Exploring Functional Connections in the Brain, MS Thesis, Electrical Engineering Department, University of Washington, June, 2007.

Keywords:

content-based image retrieval

probabilistic database

similarity-based search

distance measures

medical imaging

biomedical research tools

Enter Keyword 7

Roger A. Bauman, MD, Student Paper Award

To be eligible for the student paper award, the author must be a medical student, resident, graduate or postdoctoral student, fellow, or in a full-time training program. The applicant must be a student at the time of submission, the first author on the paper and the presenter of the paper at the SIIM 2008 meeting to be eligible.

Are you a student submitting your paper for the Bauman Student Paper Award? ☐ Yes ☒ No

SIIM New Investigator Travel Awards

Travel awards are open to individuals with an area of study and interest in imaging informatics. At the time of submission, applicants must be in a full-time training program or must have completed their training program within two years of the SIIM 2008 Annual Meeting. This includes full-time undergraduate and graduate students, medical students, residents, and post-doctoral fellows. Applicants must be first author on the paper/abstract and presenter of the paper/poster/demonstration at the SIIM 2008 Annual Meeting to be eligible.

Are you submitting for a SIIM New Investigator Travel Award? ☐ Yes ☒ No

---

Primary Author:

First Name: Indriyati

Middle Name/Initial: Enter Name

Last Name: Atmosukarto

Degree: M.S.

Title: Graduate student

Department: Computer Science and Engineering

Organization: University of Washington

Address: 185 Stevens Way

City: Seattle

State/Province: WA

Zip Code: 98105

Phone: 206-685-6765

Fax: Enter Fax

Email: indria@cs.washington.edu

SIIM Member? ☐ Yes ☒ No

Secondary Authors:

Author 2:	**Author 3:**	**Author 4:**
First: Ravensara	First: Joshua	First: Linda
Middle: S.	Middle: D.	Middle: G.
Last: Travillian	Last: Franklin	Last: Shapiro
Degree: Ph.D	Degree: M.S.	Degree: Ph.D

Title: Medical Informaticist Organization: Children's Hospital and Regional Medical Center	Title: Researcher Organization: University of Washington	Title: Enter Title Organization: Enter Organization
Author 5: First: James Middle: F. Last: Brinkley Degree: M.D., Ph.D. Title: Professor Organization: University of Washington	**Author 6:** First: Dan Middle: Last: Suciu Degree: Ph.D. Title: Professor Organization: University of Washington	**Author 7:** First: John Middle: I. Last: Clark Degree: Ph.D. Title: Professor Organization: University of Washington
Author 8: First: Michael Middle: L. Last: Cunningham Degree: M.D., Ph.D Title: Professor Organization: University of Washington	**Author 9:** First: Enter Name Middle: Enter Name Last: Enter Name Degree: Enter Degree Title: Enter Title Organization: Enter Organization	**Author 10:** First: Enter Name Middle: Enter Name Last: Enter Name Degree: Enter Degree Title: Enter Title Organization: Enter Organization
Author 11: First: Enter Name Middle: Enter Name Last: Enter Name Degree: Enter Degree Title: Enter Title Organization: Enter Organization	**Author 12:** First: Enter Name Middle: Enter Name Last: Enter Name Degree: Enter Degree Title: Enter Title Organization: Enter Organization	

Presenter (if different than Primary Author):

Presenter's First Name: Enter Name

Middle Name/Initial: Enter Name

Last Name: Enter Name

Degree: Enter Degree

Title: Enter Title
Department: Enter Dept
Organization: Enter Organization
Address: Enter Address
City: Enter City
State/Province: Enter State/Prov
Zip Code: Enter Zip
Phone: Enter Phone
Fax: Enter Fax
Email: Enter Email