Global Journal of Computer Science and Technology
Graphics & Vision
Volume 12 Issue 11 Version 1.0 Year 2012
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

# Performance of Retrieval Information System for Medical Images

By Hazem (Moh'd Said) Hatamleh, Ahmed A.M Sharadqeh, As'ad Mahmoud Alnaser, Omar Alheyasat & Ashraf Abdel-Karim Abu-Ein

*Al-Balqa' Applied University, Jordan*

*Abstract* - Problem statement : In this article the principles of building knowledge and retrieval information system will be applied to medical images in some studied hospital. Since there is a huge number of medical images this system will organizes and manages the operation of retrieving and displaying such images to the persons who need such images in short time and in high quality services.

Approach : From relevance assessments we can compute measures of retrieval performance such as: Recall (R), discrimination DC, and Precision. Results: both recall and precision of the system are linearly depend on relevant items correctly retrieved.

Conclusion : Number of retrieved images from huge total number of medical images in some hospital determine the systems' recall, discrimination, and precision of the retrieval information system.

*Keywords :* *Knowledge systems, retrieval information systems, medical images.*

*GJCST-F Classification:* *J.3*

PERFORMANCE OF RETRIEVAL INFORMATION SYSTEM FOR MEDICAL IMAGES

*Strictly as per the compliance and regulations of:*

# Performance of Retrieval Information System for Medical Images

Hazem (Moh'd Said) Hatamleh α, Ahmed A.M Sharadqeh σ, As'ad Mahmoud Alnaser ρ, Omar Alheyasat Ѡ & Ashraf Abdel-Karim Abu-Ein ¥

*Abstract -* Problem statement : In this article the principles of building knowledge and retrieval information system will be applied to medical images in some studied hospital. Since there is a huge number of medical images this system will organizes and manages the operation of retrieving and displaying such images to the persons who need such images in short time and in high quality services.

Approach : From relevance assessments we can compute measures of retrieval performance such as: Recall (R), discrimination DC, and Precision. Results: both recall and precision of the system are linearly depend on relevant items correctly retrieved.

Conclusion : Number of retrieved images from huge total number of medical images in some hospital determine the systems' recall, discrimination, and precision of the retrieval information system.

*Keywords: Knowledge systems, retrieval information systems, medical images.*

## I. Introduction

With the availability of digital medical image acquisition devices and the rapid growth of computing power, effective retrieval of digital medical images in a large database is a challenging research point. The health and medical related professions use and store the patients' visual information in the form of ultrasound, X-rays or other types of scanned images for the purpose of diagnosis and monitoring. The availability and optimal use of these medical images with respect to medical diagnosis and allied purposes is a function of how these images are stored and retrieved, [1, 2].

Region based signature can be acquired by image segmentation. Reliable segmentation is also critical to get the image shape description. However, content-based medical image retrieval that confront many image types, some of them even have not a clear object, so some strategies for dealing with this problem is to reduce dependence on accurate image segmentation for a practical image retrieval system. The increasing reliance of modern medicine on diagnostic techniques such as radiology, histopathology, and computerized tomography has resulted in an explosion in the number and importance of medical images now stored by most hospitals. While the prime requirement for medical imaging systems is to be able to display images relating to a named patient, there is increasing interest in the use of CBIR techniques to aid diagnosis by identifying similar past cases. Most development work in the PACS (Picture Archiving and Communication Systems) area is still directed towards providing basic functionality (ensuring that medical images can be successfully digitized, stored and transmitted over local area networks without loss of quality) and usability (providing user-centered interfaces and integrating image storage and retrieval with wider aspects of patient record management). However, experimental content-based retrieval systems are beginning to have some impact. Due to the problem of achieving the accurate image segmentation for medical image retrieval; this thesis will deal with strategies to reduce the dependence on accurate image segmentation for a practical image retrieval system. It is desirable for the relevance feedback based on the user participation in image retrieval system to solve the problem of the semantic gap and the image retrieval with low level visual features. Through the user's feedback, the corresponding high-level semantic will be obtained based on machine learning theory. SWETA et al. 2006, Medical images are a critical component of the healthcare system with great impact on the society's welfare. Traditionally medical images were stored on film but the advances in modern imaging modalities made it possible to store them electronically. Thus, this paper proposed a novel framework for classifying various strategies for storing, retrieving and processing digital medical images. In addition to a detailed discussion, the assessment of the classification framework includes a potential usage scenario of the framework. For researchers, this study identified an important strategies and points out future research directions while, for practitioners, the proposed framework might help medical users develop a lucid understanding of the different approaches and their advantages and disadvantages. Cosmin S, 2011, the article presented a software system that implemented a multimedia database management server. The software has a modularized architecture controlled by a relational database system. An element of originality is that along

*Author α σ ρ Ѡ ¥ : Al-Balqa' Applied University, Jordan.*
*E-mail α : hazim-hh@bau.edu.jo*
*E-mail σ : sharadqh_78@yahoo.com*
*E-mail ρ : asad1_99@yahoo.com*
*E-mail Ѡ : omarheyasat@hotmail.com*
*E-mail ¥ : ashraf.abuain@fet.edu.jo*

with the classical functionality of such a system, it includes specialized modules for extracting texture and color characteristics from images and for executing content based retrieval queries. This system is tested from the speed point of view both for images processing speed and retrieval speed. Mohammad O, et al. 2001, in this paper, they introduced intermediate features. These were low level "semantic features" and "high level image" features. That is, in one hand, they can be arranged to produce high level concept and in another hand, they can be learned from a small annotated database. These features can then be used in an image retrieval system. They reported experiments where intermediate features are textures. These are learned from a small annotated database. The resulting indexing procedure is then demonstrated to be superior to a standard color histogram indexing. Johan M, et al. 2007, their research project was addressing the problem of content-based medical image retrieval in large databases. They were exploiting grids to tackle the computational requirement of this problem. They developed strategies to optimize the load distribution over the very large scale EGEE grid infrastructure, taking into account its properties and load. They have explored several strategies to identify relevant images. Texture features extracted using Gabor filters proved to be an efficient and relevant mean of indexing medical databases. The texture features could be correlated to image modality, tissues, and subtle changes such as myocardium tissues variation during the cardiac cycle. Henry J.et al. 1998, the emergence of Multimedia Electronic Medical Record Systems (MEMRS), architectures that integrate medical images with text-based clinical data, will further hasten this trend. The development of these systems, storing a large and diverse set of medical images, suggests that in the future MEMRS will become important digital libraries supporting patient care, research and education. The representation and retrieval of clinical images within these systems is problematic as conventional database architectures and information retrieval models have, until recently, focused largely on text-based data. Medical imaging data differs in many ways from text-based medical data but perhaps the most important difference is that the information contained within imaging data is fundamentally knowledge-based. New representational and retrieval models for clinical images will be required to address this issue. Within the Image Engine multimedia medical record system project at the University of Pittsburgh they were evolving an approach to representation and retrieval of medical images which combines semantic indexing using the UMLS Metathesuarus, image content-based representation and knowledge-based image analysis. Johan Montagnat et al. 2004, in this paper they studied the impact of executing a medical image database query application on the grid. For lowering the total computation time, the image database is partitioned in subsets to be processed on different grid nodes. A theoretical model of the application computation cost and estimates of the grid execution overhead are used to efficiently partition the database. They showed results demonstrating that smart partitioning of the database can lead to significant improvements in terms of total computation time. Weidong C. et al. 2001, various picture archiving and communications systems (PACS) have been developed to deal with this growing volume of data generated by different systems, providing digital image acquisition, archiving, retrieval, processing, and distribution and communication and display functions.1 In a PACS system, archived medical images can be quickly retrieved electronically. PACS workstations can take advantage of image analysis and processing software to manipulate and enhance image data. However, current PACS systems are expensive and complex. They require a large number of review and display workstations, and each workstation usually requires dedicated licensed software and considerable maintenance support. The high cost of dedicated PACS workstations prevents their deployment at all locations where they would be useful, for example, elsewhere in the same institution and at remote sites such as other institutions or the home. Here, they presented a prototype Web-based medical image data access and manipulation system.

## II. Information Retrieval Systems

Information retrieval systems are everywhere: Web search engines, library catalogs, store catalogs, cookbook indexes, and so on. *Information retrieval* (IR), also called *information storage and retrieval* (ISR or ISAR) or *information organization and retrieval*, is the art and science of retrieving from a collection of items a subset that serves the user's purpose; for example:

- Web pages useful in preparing for a trip to Europe;
- Magazine articles for an assignment or good reading for that trip to Europe;
- Educational materials for a learning objective;
- Digital cameras for taking family photos;
- Recipes that use ingredients on hand;
- Facts needed for deciding on a company merger.

The main trick is to retrieve what is useful while leaving behind what is not. [3, 4].

## III. Utility, Relevance, and IR System Performance

Utility and relevance underlie all IR operations. A document's utility depends on three things, topical relevance, pertinence, and novelty. A document is *topically relevant* for a topic, question, or task if it contains information that either directly answers the question or can be used, possibly in combination with other information, to derive an answer or perform the task. It is *pertinent* with respect to a user with a given purpose if, in addition, it gives just the information needed; is compatible with the user's background and cognitive style so he can apply the information gained; and is authoritative. It is *novel* if it adds to the user's knowledge.

## IV. Material and Method

Analogously, a soccer player is topically relevant for a team if her abilities and playing style fit the team strategy, pertinent if she is compatible with the coach, and novel if the team is missing a player in her position. Utility might be measured in monetary terms: "How much is is it worth to the user to have found this document?" "How much is this player worth to us?" "How much did we save by finding this software?" In the literature, the term "relevance" is used imprecisely; it can mean utility or topical relevance or pertinence. Many IR systems focus on finding topically relevant documents, leaving further selection to the user. Relevance is a matter of degree; some documents are highly relevant and indispensable for the user's tasks; others contribute just a little bit and could be missed without much harm (see ranked retrieval in the section on *Matching*). From relevance assessments measures of retrieval performance can be computed such as:

Recall (R),

$$R = \frac{RICR}{ARI} \quad (1)$$

Also discrimination DC,

$$DC = \frac{IICR}{AII} \quad (2)$$

Precision can be written as:

$$P = \frac{RIR}{AIR} \quad (3)$$

Where: RICR: relevant items correctly retrieved, ARI: all relevant items in the collection, IICR: irrelevant items correctly rejected, AII: all irrelevant items in the collection, RIR: relevant items retrieved, AIR: all items retrieved.

Evaluation studies commonly use recall and precision or a combination; whether these are the best measures is debatable. With low precision, the user must look at several irrelevant documents for every relevant document found. More sophisticated measures consider the gain from a relevant document and the expense incurred by having to examine an irrelevant document. For ranked retrieval, performance measures are more complex. All of these measures are based on assessing each document on its own, rather than considering the usefulness of the retrieved set as a whole; for example, many relevant documents that merely duplicate the same information just waste the user's time, so retrieving fewer relevant documents would be better.

## V. How Information Retrieval Systems Work

IR is a component of an information system. An information system must make sure that everybody it is meant to serve has the information needed to accomplish tasks, solve problems, and make decisions, no matter where that information is available. To this end, an information system must (1) actively find out what users need, (2) acquire documents (or computer programs, or products, or data items, and so on), resulting in a collection, and (3) match documents with needs.

## VI. Results

In this work the total number of medical images in some collection is 1000 image, and the relevant items correctly retrieved are in 9 trials are respectively (100, 200, 300, 400, 500, 600, 700, 800, and, 900) then Recall- RICR relationship can be represented as in figure (1).

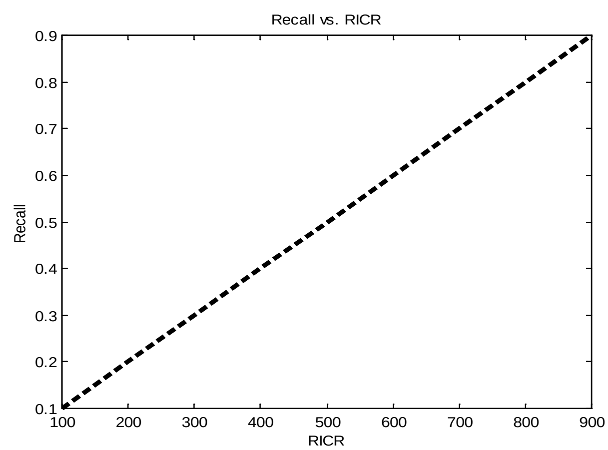


*Figure 1 :* Recall vs. RICR

Also discrimination DC varies linearly with rejected items for the same package of 1000 medical images. The precision of the system as in [3] can be reaching to 80%.

## VII. Discussion

It is obvious that the recall and precision of the image system is linearly depends on RICR, also the proportionality between such values is fully with about 80% .

## VIII. Conclusions

Number of retrieved images from huge total number of medical images in some hospital determine the systems' recall, discrimination, and precision of the retrieval information system.

## References Références Referencias

1. [C. Weidong 2001] F. Dagan, F. Roger: Web-Based Digital Medical Images, IEEE Computer Graphics and Applications.
2. [M. Johan2007] Tristan GLATARD, Diane LINGRAND: Texture based Medical Image Indexing and Retrieval on Grids, MEDICAL IMAGING TECHNOLOGY Vol.25 No.5 November 2007
3. [M. Johan2004] B. Vincent, M. Isabelle: Medical image databases content-based queries partitioning on a grid, MEDIGRID, French ministry for Reseach ACI-GRID project. http://www.creatis.insa-lyon.fr/MEDIGRID/.
4. [O. Mohamad2001] Bruno Jedynak Mohamed Daoudi, 2001, Image Indexing & Retrieval Using Intermediate Features, MM'O/, Sept. 30-Oct. 5, 2001, Ottawa, Canada.
5. [Cosmin Spahiu, 2011]: TESTING THE PERFORMANCES OF A MULTIMEDIA DATABASE SYSTEM" International Journal of Computer Science and Applications, Vol. 8, No. 1, pp. 123-135.
6. [SWETA 2006] D. ALINA: Strategies for Working with Digital Medical Images Proceedings of the 39th Hawaii International Conference on System Sciences – 2006.