International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 2 Issue 3, March - 2013

Content-Based Image Retrieval System in Medical Applications

[1]R.Senthil Kumar,
Asst.Professor, Dept of Computer Science,
Dharmapuram Adhinam Arts College,
Mayiladuthurai (Tamilnadu)

[2]Dr.M.Senthilmurugan,
Professor, Department of Computer Applications,
Bharathiyar College of Engineering & Technology,
Karaikal (U.T of Puducherry)

***Abstract*— Content-based visual information retrieval (CBVIR) or Content-based image retrieval (CBIR) has been one of the most vivid research areas in the field of computer vision over the last 10 years. In the medical field, especially digital images, are produced in ever increasing quantities and used for diagnostics and therapy. The Radiology Department of the university Hospital of Geneva alone produced more than 25,000 images a day in 2012, The cardiology is currently the second largest producer of digital images with digital imaging and communication in medical (DICOM), a standard for image communication has been set and patient information can be stored with the actual images. The main aim of this paper is to improve the overall CBIR system Performance using medical applications.**

***Keywords*—** Medical Imaging (MI), Content-Based Visual Information Retrieval (CBVIR), Content-Based Image Retrieval (CBIR), Digital Imaging and Communication in Medicine. (DICOM).

1. INTRODUCTION

Image retrieval has been an extremely active research area over the last 10 years, but first review articles on access methods in image database appeared already in the early 1980s. The growing number of digital image acquisition and storage systems in clinical routine rises demands for new access methods. Still, most picture archiving and communication systems (PACS) only use textual information to access a patient's image data, which has been mainly entered manually. Content-based image retrieval (CBIR) depends on automatically extracted content-descriptions for each images as well as their storage and comparison upon a query. The vast majority of users do not want to retrieve images simply on the basis of similarity of appearance. They need to be able to locate picture of a particular type of object, phenomenon or event [2].

Gudiva and Raghavan [1] have drawn a useful distinction between retrieval by primitive image feature (Such as color, texture or shape) and semantic feature (such as the type of object or even depicted by the image) Eakins [3] has taken this distinction further, identifying three distinct levels of image query each of which can befurther subdivided.

Level 1: Retrieval by primitive features. Such as color ,texture, shape. (eg. "find all pictures containing yellow or blue stars arranged in a ring")

Level 2: Retrieval by derived attribute or logical feature (eg. "find pictures of a passenger train crossing a bridge")

Level 3: Retrieval by abstract attributes (eg. "find pictures illustrating pageantry")

1.1. Content-based image retrieval systems

Although early systems existed already in the beginning of the 1980s [4] the majority would recall systems such as IBM's Query By Image Content (QBIC) as the start of content-based image retrieval [5,6]. The commercial QBIC system is definitely the most well-known system. Another commercial system for image [7] and video[8] retrieval is virage that has been well known commercial customers such as CNN. CBIR is an image search systems that search for image by image content (ie) Keyword-based Image/Video Retieval (eg) Google image search, YouTube.

1.1.1 Basic Components of CBIR

- Feature Extractor
  - Create the metadata
- Query Engine
  - Calculate Similarity
- User Interface.

1.1.2. Kinds of CBIR

- General: We try to match a query image to an arbitrary collection of images.
- Application specific: We try to match a query image to a collection of images of a specific type (eg. Finger prints, X-ray images of a specific organs.

1.1.3. Image Retrieval Phases

There are four phases of retrieves the images from the database based on the geometric properties of the input query images boundry.

(i) Image database generation phase
  - input images from scanner, thinning editings
(ii) Outline-based image retrieval phase
  - Identify the global shape similarity
(iii) Global hash-table generation phase

- Combines the individual hash-tables

(iv) Index-based image retrieval phase.

- Invariant features of the input query image to compute the indices.

1.1.4. Applications of CBIR

- Consumer Digital Photo Albums
  - Digital Cameras
  - Flicker
- Medical images, Crime Prevention
- Digital Museum
- Trademarks search
- MPEG-7 content descriptors

1.1.5 CBIR Working Principle

- Extract features from images.
- Let the user do query
  - Query by Sketch
  - Query by Keywords
  - Query by Examples
- Refine the result by relevance feedback
  - Give feedback to the previous result.

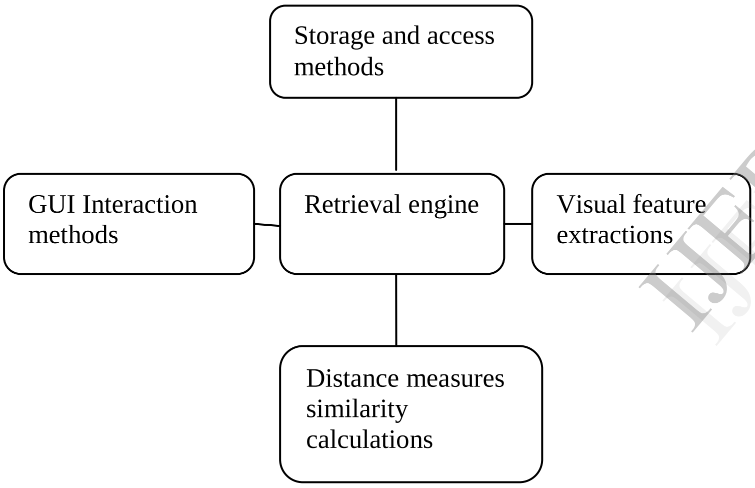

Fig.1 The principal components of all content-based Image retrieval systems.

Most of these systems have a very similar architecture for browsing and archiving/indexing images comprising tools for the extraction of visual features, for the storage and efficient retrieval of these features, for distance measurements or similarity calculations and a type of graphical user interface(GUI). This general system setup is shown in Fig.1.

1.2.1 Visual features

Visual features were classified in[9] into primitive features such as color or shape,logical features such as identity of objects shown and abstract features such as significance of scenes depicted. Even systems using segments and local features such as Blobword[10,11] are still far away from identifying objects reliably. No system offers interpretation of images or even medium level concepts as they can easily be captured with text. This loss of information from an image to a representation by feature is called the semantic gap[12].

Another gap is the sensory gap that describes the loss between the actual structure and the representation in a (digital) image.

1.2.1 Color

In Stock photography (large, varied databases for being used by artists, advertisers and journalists), color has been the most effective feature and almost all systems employ colors. Although most of the images are in the red, green, blue (RGB) color space, this space is only rarely used for indexing and querying as it does not correspond well to the human color perception. Other spaces such as hue, saturation, value (HSV) and LUV spaces are much better with respect to human perception and are more frequently used. This means that differences in the color space are similar to the differences between colors that human perceive. In specialized fields, namely in the medical domain, absolute color or grey level features are often of very limited expressive power unless exact reference points exist as it is the case for computed tomography images.

1.2.2 Texture

Texture measures have an even larger variety than color measures. Some of the most common measures for capturing the texture of images are wavelets and Gabor filters These texture measures try to capture the characteristics of the image or image parts with respect to changes in certain directions and the scale of the images. This is most useful for region or images with homogeneous texture.

1.2.3 Local and global features

Both color and texture features can be used on a global image level or on a local level on parts of the image. The easiest way to use regional feature is to use blocks of fixed size and location, so called partitioning of the image [7] for local feature extraction. When segmenting the image into areas with similar properties the locally extracted features contain more information about the image objects or underlying structures.

1.2.4 Segmentation and shape feature

Fully automated segmentation causes many problems and is often not easy to realize. In image retrieval, several system attempt to perform an automatic segmentation of the images in the collection for feature extraction. To have an effective segmentation of images using varied image databases the segmentation process has to be done based on the color and textual properties of the image regions. After segmentation the resulting segments can be described by shape features that commonly exists, including those with invariants with respect to shifts, rotations and scaling.

1.3 Storage and access methods

Common storage methods used are relational databases inverted files, self-made structure or simply to keep the entire index in the main memory which will inevitably cause problems, when using large databases. These methods often need to use dimension reduction techniques or pruning methods to allow for an efficient and quick access to the data.

1.4 Important techniques

There is a large number of important techniques to improve the performance of retrieval systems. One of the most prominent technique is a relevance feedback, that is well known from text retrieval. This technique has proven to be important for image retrieval as well because often unexpected or unwanted images shown up in the result of a similarity query. Other technique from the artificial intelligence community are also used for image retrieval such as long-term learning from user behaviour based on data mining in usage log files using the well known market basket analysis.

2. Types of medical imaging

In modern medicine, medical imaging has undergone major advancements. Today this ability to achieve information about the human body has many useful clinical applications. over the years,different sorts of medical imaging have been developed each with their own advantages and disadvantages.

X-ray based methods of medical imaging include conventional X-ray, Computed tomography(CT) and mammography. To enhance the X-ray image, contrast agents can be used for example for angiography examination.

Molecular imaging is used in nuclear medicine and uses a variety of methods to visualize biological process taking place in the cells of organisms. Small amount of radioactive markers called radio pharmaceutical are used for molecular imaging. other types of medical imaging are magnetic resonance imaging(MRI) and ultrasound imaging. unlike conventional X-ray, CT and molecular imaging, MRI and ultrasound operate without ionizing radiation. MRI uses strong magnetic fields, which produce no known irreversible biological effects in humans. Diagnostic ultrasound systems use high-frequency sound waves to produce images of soft tissue and internal body organs.

2.1 Challenges in CBIR system

Four main challenges are :

- computational complexity of indexing and retrieval process,which mainly represents the elapsed time for indexing and retrieval purposes.
- memory and disk space requirements.
- semantic retrieval performance,which shows the accuracy of the system and
- usability.
-

2.2 The need for content-based medical image retrieval

It is important to explain the needs and to discuss possible technical and methodological improvements and the resulting clinical benefits. The goal of medical information systems have often been defined to deliver the needed information at the right time, the right place to the right person. In order to improve the quality and efficiencyof care processes. Such a goal will most likely need more than a query by patient name,series ID or study ID for images. For the clinical decision-making process it can be beneficial or even important to find other images of the same modality, the same anatomic region of the same disease. DICOM headers have proven to contain a fairly high rate of errors. for example for the field anatomical region, error rates of 16% have been reported. This can hinder the correct retrieval of all wanted images.

Clinical decision support techniques such as case-based reasoning or evidence-based medicine can even produce a stronger need to retrieve images that can be valuable for supporting certain diagnoses. It could even be imagined to have image-based reasoning (IBR) as a new discipline for diagnostic aid.

Besides diagnostics ,teaching and research especially are expected to improve through the use of visual access methods as visually interesting images can be chosen and can actually be found in the existing large repositories.

2.3 Use in various medical departments

There are also several other departments where CBIRs have been implemented. A categorization of images from various departments has been described. A classification of dermatologic images, cytological specimens have already been described. Pathology images have often been proposed for content-based access as the color and texture properties can relatively easy be identified. within the radiology departments , mammographics are one of the most frequent application areas with respect to classification and content-based search. The negative psychological effects of removing tissue for false positive patients have been described of one of the principal goals to be reduced. Ultrasound images of the breast are used. Another active area is the classification of high resolution computed tomography (HRCT) scans the lung as done by the Assert project. Magnetic resonance images (MRI) of the brain are used to demonstrate the image search algorithm. CT brain scans the search for medical tumours by their shape properties(after segmentation) have been described. Functional photon emission tomography (PET) images and spine X-ray are used for retrieval. Table 1. Various image types and the systems that are using these images.

Images used	Names of the system
HRCTs of the lung	ASSERT
Functional PET	FICBOS
Spine X-rays	CBIR2
CTS of the head	MIMS
Mammographics	APKS

3. Techniques used in medical image retrieval

The various techniques that are currently – used or that have been proposed for the use in medical image retrieval applications.

3.1 Query formulation

Most systems in CBIR use the query by example (s) QBE paradigm which needs an appropriate starting image for querying. This problem of a sometimes missing starting image is known as the page zero problems. If text is attached to the images then the text can be used as a starting point. The starting point does not need to be further defined but the images of the case can be used directly. The use of sketches has already been proposed in generic image retrieval. and it has also been proposed for the use in medical applications. Consisting the difficulty in exact drawing and the need for some artistic skills and time, this method will only be applicable for a very small subset of queries, such a tumor shapes or spine X-rays where outline are possible directly in the images. For general image retrieval sketches are too time-consuming and the retrieval results often not exact enough.

3.2 Text

Many systems propose to use text from the patient record or studies to search by content. The use of text for queries is undeniable efficient. The combination of textual with visual features or content and context of the images does have the most potential to lead to good results. Besides the free text that is frequently used for retrieval, medical patient records also contain very valuable structured information. Such as age, sex and profession of the patient. This information is just as important as free text to put the images into content.

4. Application fields in medicine and clinical benefits

This section gives a more ordered view on what in medicine image retrieval can be used for and what the effects can be if proper applications are developed.

Three image domains can instantly be stated for the use of content-based access methods: teaching, research and diagnostics.

First benefit will most likely be the domain of teaching. Here lectures can use large image repositories to search for interesting cases to present to the students. These cases can be chosen not only based on diagnosis or anatomical region but also visually similar cases with different diagnoses can be presented which can augment the educational quality. Research can also benefit from visual retrieval methods. Researchers have more options for the choice of cases to include into research and studies by allowing text-based and visual access. Finally, diagnostics will be the hardest but most important applications for image retrieval. To be used as a diagnostic aid, the algorithms need to prove their performance and they need to be accepted by the clinicians as a useful tool.

There are two principal ideas for supporting the clinical decision-making process. The first one is to supply the medical doctor with cases that offer a similar visual appearance. Another idea is the creation of databases containing normal cases and compare the distance of a new case with the existing cases doing thus dissimilarity retrieval as opposed to similarity retrieval.

5. Conclusion

Content-based visual information retrieval definitely has a large potential in the medical domain. Content-based methods can be used on a large variety of images and in a wide area of applications. It is clear that new tools and methods are needed to manage the increasing amount of visual information that is produced in medical institutions. To improve the overall CBIR system performance in different hardware platforms, having different technical capabilities and conditions, CBIR systems often analyse image content via the low-level features (color, texture and shape). Improving CBIR applications involves improving the retrieval performance, efficiency, adaptability and scalability of the overall system. Finally, the content-based access methods have an enormous potential when used in the correct ways.

REFERENCES

[1] V.N. Gudivada.V.V.Raghavan. Content-based image retrieval system. IEEE Comput. 28(9) (1995) 18-22.

[2] L.Armitage,P.G.B Enser,Analysis of user need in image archieves. J.Inform.Sci 23(4)(1997) 287-299.

[3] J.P.Eakins,Techniques for image retrieval Library and Information Brifings 85, British Library and south bank university,London 1998.

[4] N.S.Chang,K-S Fu,Query-by-pictorial example,IEEE Trans. Software Eng. SE 6 (6). (1980) 519-524.

[5] M.Flickner,H.Sawhnery,W.Niblack,J,Ashley,Q.Hunang,B.Dom,M.Gorkani,J.Hafner,D.Lee,D.Petkovic,D.Steele,P.Yanker,Query by image and video content: the QBIC System,IEEE comput.28(9)(1995) 23-32.

[6] W.Niblack,R,Barber,W.Equitz,M.D.Flickner,E.H.Glassman,QBIC project Querying images by content. vol.1908,1993,pp.173-187

[7] J.R.Bach,C.Futter A.Gupta A.Hampapur The virage image search engine.1996,pp.76-87

[8] A.Hampapur,A.Gupta,B,Horowitz,C,F,Shu,C,Fuller virage video engine.1997,pp.352-360

[9] J.P.Eakins,M.E.Graham,Content-based image retrieval,Tech Rep. JTAP-039,JISC Technology Application Program.New castle upon Tyne,2000

[10] C.Carson,M.Thomas,S.Belingie,J.M.Heller Stein a system for region-based image indexing & retrieval. The Netherlands,1999,pp.509-516

[11] S.Belongie,C,Carson,H,Greenspan,J,Malik color-and texture based image segmentation using Em and its application to content-based image retrieval. Bombay,India,1998,pp.675-682

[12] A.W.M.Smeulders,M.Worring,S.Santini,A.Gupta,R.Jain content-based image retrieval at the end of the early years. Machine Intel. 22(12) (2000) 1349-1380