A Study on the Content-Based Image Retrieval System for Medical Applications

Kien-Ping Chung, Lance Chun Che Fung and Hong Xie
School of Information Technology
Murdoch University
South Street, Murdoch
W.A. 6150, AUSTRALIA
{k.chung, l.fung, h.xie}@murdoch.edu.au

Abstract

The storage of images in databases is gaining popularity due to the rapid development of cheap and easy image capturing equipment. The principle of Content Based Image Retrieval (CBIR) system normally depends on specific features such as colour, texture, shape or semantic meaning of the image. On the other hand, content-based medical images retrieval (CBMIR) system presents different challenges. This paper discusses the comparison on the key features of a Content-Based Image Retrieval (CBIR) system and its potential adaptation as a content-based medical images retrieval (CBMIR) system. This will lead to subsequent studies and investigation in the prospective research areas.

1. Introduction

Over the last decade, there is an increasing trend of storing non text-based data in databases. Image in particular, has been gaining popularity as an alternative and sometime more viable option for information storage. While this presents a wealth of information, however, it also causes a great problem in retrieving appropriate and relevant information during searching. This has resulted in a growing interest and much active research on the subject of the extraction of relevant information from non text-based databases. Over the past few years, researchers have achieved certain degree of success in these fields. This is shown as evidence in the increasing number of commercially available search engines for images.

Content-based image retrieval (CBIR) system is a type of system which retrieves images based on their features such as colour, texture, shape or semantic meanings of the image. Upon reviewing the CBIR systems that have been reported, it was found that such systems can be grouped into two main categories, namely, *generic* and *domain-specific* systems. The generic retrieval systems like Yahoo Image Surfer and Lycos image library contain all types of different images. On the other hand, domain-specific retrieval systems only contain images that are closely related to a specific application area. A content-based medical image retrieval (CBMIR) system is a typical example of a domain-specific retrieval system.

During the past two decades, the development of new modalities such as Computed Tomography (CT), Magnetic Resonance Imagining (MRI), and Picture Archiving and Communication Systems (PACS) have resulted in an explosive growth in the number of images stored in databases. Until recently, textual index entries are mandatory to retrieve medical images from a hospital image archive system. However, the development of CBIR techniques has not only created new possible ways of retrieving images, but also opened out opportunities for other related applications.

It is however simplistic to consider that one can directly apply a generic CBIR system to a medical image database. In this study, it is considered that it will be of interest to readers the characteristics and the development trend of CBMIR systems from the perspective of one whose background is in the discipline of CBIR systems. In this paper, the applications of such technology are discussed. The development trend of CBMIR and when appropriate, comparisons between the two systems are also drawn. Finally, this paper concludes by summarizing the differences between the two systems.

2. Medical Image Database

First of all, three characteristics of a medical image database are identified. Each of these characteristics of the system presents a different challenge to the research community. The following sub-sections provide a detailed discussion on the characteristics of the CBMIR systems.

2.1 Heterogeneity

Medical image is only a general phrase many used to describe images which have captured information about the human body. It is actually a broad field that consists of image classes such as photography (e.g., endoscopy, histology, dermatology), radiographic (e.g., x-rays), and tomography (e.g., CT, MRI, ultrasound). It imposes unique, image-dependent restrictions on the nature of features available for abstraction. Each of the image classes possesses its unique characters in terms of size, shape, colors and texture of the region of interest. Thus, the visual appearance of the same organ or part of the human body will be different under different modalities. Furthermore, it is also possible that the interest in the same image may be different by systems or users for different application. It is therefore not difficult to deduce that different approaches will be required for different modalities and systems for different applications. These approaches may include the design of user interface, indexing structure, feature extraction and query processing units.

2.2 Imprecision

Imprecision has been a problem for the CBIR systems. Likewise, CBMIR systems suffer the same problem. Tagare et al. [1] have identified three components in imprecision, i.e., semantic imprecision, feature imprecision, and signal imprecision. These imprecision are caused by the ambiguity of human language, disagreement in the observation of an image and quality of information captured in the image.

The three types of imprecision described by Tagare et al. [1] are very much similar to the problems experience by the general CBIR systems. One would argue that the semantic imprecision is similar to the problems of polysemy and synonymy occurring in terms of query processing. Feature imprecision also occurs in general CBIR system when users have different interpretation on the image, or the inability by the user to describe the semantic content of the image. This is a common problem in the CBIR system as the scope of the collected images are broader than images collected for medical field. Lastly, signal imprecision is quite common among images captured in the outdoor environment.

2.3 Dynamism of Indexing Structure

As described in the previous section, the human interpretation of a medical image may vary from person to person. The interpretation of an image from the same person may also change as the person gains more experience. Thus, the area of interest for the same image may change as the interpretation of the image changes. For systems that index images by their semantic contents or the visual features of the area of interest, such changes may result in a need to modify the indexing structure in order to adapt to the user's knowledge. However, traditional indexing structures reported so far are static. The process of re-organizing the indexing structure is mostly manually driven. Hence, a significant overhead is included. Ideally, the indexing structure for medical images should be dynamic while keeping the overhead for re-organizing the indexing structure to minimal. Preferably, very little manual interaction should be required.

2.4 Remark

All of the issues described above are not unique to CBMIR systems. As a matter of fact, these issues are also faced by CBIR systems. To a degree, the magnitude of some of these issues faced by CBIR is larger that what those in CBMIR systems. The reason is being that the scope of the medical images is bounded within the medical domain. Hence, certain assumptions can be made prior to analysing the images.

3. Applications

Medical imagery is an exciting field for researchers of CBIR. It not only contains vast amount of image resources that the researchers can work on, it also provides practical applications that research theories can be applied to. Due to these reasons, there has been a steady growth in developing medical applications with the use of CBIR techniques. The CBMIR systems are grouped into two categories mostly according to their input data format and to a certain extent, the domain scope of their applications.

Traditionally, there are two standard approaches in querying the system, namely, *query by keywords* or *image examples*. In query by example, the diagnostic system is one of the applications where researchers have been focusing on. As the name implies, the output from these systems is the diagnostic result derives from the system's input image. Until now, the systems reported have only been designed to support specific medical tasks such as retrieval of tumor shapes in mammograms [2], identification of lung disease from computed tomography [3], differentiation of Mantle Cell Lymphoma (MCL) from Chronic Lymphocytic Leukemia (CLL) or Follicular Center Cell Lyphoma (FCC) using pathology images [4], and retrieving of spine in the x-ray database [5]. All these systems are designed to query by image. The region of interest for the input image is partially or automatically selected by the system. Manual interrelation is required when the image resolution is low, or with an inability of applying image models to capture the visual features in interest. It is worth noting that some authors are quite cautious in using the term "diagnostic" for their proposed systems. Instead, these authors prefer to call them "decision support" systems.

In addition to the diagnostic systems, Liu et al. [6] have developed a teaching assistant system for the identification of lung diseases from tomographic images. This system allows the professor to select images with similar texture but may be not belonging to the same disease. The objective is to teach interns to learn how to distinguish various disease images with similar texture.

Currently, the tradition picture and archiving image system (PACS) is used for searching medical images in many hospital or clinical systems. The images stored in the PACS system are normally organized according to their semantic content, or by patient's details, or, other related information. Systems that allow the users to retrieve images via the patient's details and other related information are normally based on the patients' history.

In summary, systems that browse or search images via the semantic content are mostly used for education or decision support purposes. The image grouping and organization is mostly manually driven, and these images are mostly indexed by a simple phrase or term.

4. Overall Framework

Figure 1 is a possible framework for CBIR systems. The communication between the modules is mostly bi-directional, implying this framework allows bi-direction interaction between users and the systems. This framework is an extension from Rui et al. [7]. It is felt that the query processing unit should be related to feature extraction unit, and the user should have a role in defining the image label.

The PACS systems discussed in the previous section are rather simple as a content-based retrieval system. At the minimum, a CBIR should consist of components as depicted in Figure 1. Clearly, the PACS systems lack the system module for feature selection, and quite possibly the images are indexed by a very simple one-dimensional structure grouped by the label given to the images.

Diagnostic systems reported so far are only designed for very specific application. It may not be possible for such systems to be transferred to other medical applications. The reason is obvious, different diagnostic system uses different visual features for identifying different medical cases. Thus, the feature extraction approach for each system will be different. Often, such systems are also static, implying that a significant overhead is required for a visual feature to be added, deleted or modified. Furthermore, the indexing structure applied for these systems are often not targeted for a large database and definitely not for image browsing.

Tagare et al [1] have identified several essential features for CBMIR systems: (a) non-textual indexing, (b) customized scheme, (c) dynamic modules, (d) similarity modules, (e) comparison modules, (f) iconic queries, (g) descriptive language, (h) multi-modality registration, (i) image manipulation. Researchers have generally viewed these necessities as the guideline for building a complete CBMIR system.

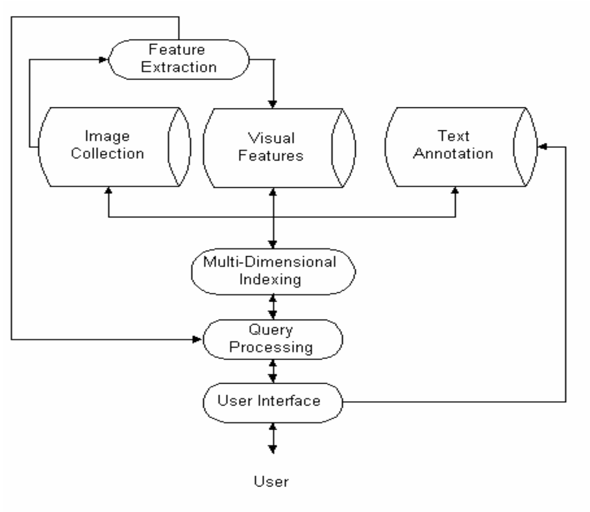

Figure 1. A possible framework for CBIR system.

Lehmann et al. [8] have proposed a medical image browsing and searching framework called Image Retrieval in Medical Application (IRMA). This framework has the potential of answering every system requirements as listed by Tagare et al [1]. IRMA is a multi-layer framework that provides separate layers which include: identification of image categories, extraction of image content and local features, indexing images based on their semantic content and image retrieval on the semantic level. The major difference between this framework and the more traditional CBIR framework [9] is that this framework uses prior knowledge of different medical modalities to determine the content of the images. This is however not possible for a general CBIR system.

5. User Retrieving Techniques

To date, much of the research effort for CBMIR systems has been spent on identifying features which uniquely identifying image, and the indexing structures for images in the system. Not much has been spent on the user interface of the system. It is true that most of the issues surrounding CBIR system in designing user interface are to do with limiting the domain scope of the input query, and to a degree, this is less of an issue for a narrow domain retrieval system such as CBMIR. However, most of the CBMIR systems reported are stilled targeted for trained personal, indicating that they are room for improvement.

The user interface for CBMIR system is based on the strategy adopted by the traditional CBIR system. The standard approach is query the system by image or keyword. As with any search engines, keywords have limited usefulness, in that it is difficult to assign keywords consistently and exhaustively to each image. In query by example, most of the systems provide users with editor for manipulating the input image, and some may even allow the user to query by sketch. However, the issue facing most of these systems is that it is relatively difficult for the users to specify the appropriate combination of visual features for a particular search. Systems proposed by Liu et al. [6], Comaniciu et al. [4] and Shyu et al. [3] all require user to select the region of interest (ROI) area in an image for the system to construct the proper query. The users of this type of systems are usually the experts in the related fields. Automatic image extraction function however is relatively rare.

To solve these problems, in addition to the two query approaches using image or keyword, some systems also provide feedback strategies and ranking for refining the search query. The feedback strategies invite interactive inputs from the user to refine the query for subsequent retrieval. In addition to the relevant feedback, the ranking approach can also be incorporated to the user interface to provide a better capture on the relevancy of the result. In comparison to the CBIR systems, these approaches are simpler as the domain scope of CBMIR systems are better defined. Hence, the iteration cycles for the interaction between user and machine are less in CBMIR systems.

A popular approach for system such as MedPix is to use directory-like structure to guide user to construct a more accurate query. The advantage of this approach is the context of the query keywords are bounded by the directory entries established by the search engine site. In this way, the user is able to obtain more accurate search results.

6. Query Processing

Query processing, in any content-based retrieval systems, is a module between the user interface and the indexing structure. It acts as a module to bridge the semantic gap between the user's input and the actual query applied to the database. In short, it converts the user input into a feature vector to be applied for searching through the index tree. Thus, the approach applies to this component is tightly coupled with the design of the user interface and the image indexing structure employed by the system. Hence, the issues described in the previous section, that is, problems with keyword/s and image example, is mostly handled by this module.

6.1 Query by Keyword

One of the biggest challenges facing researchers in query with keywords is the ability to accurately represent the user input by the system-constructed query. One of the major reasons for the low accuracy of the search result is caused by miss-representation and the system's miss-interpretation of the user query. To a large degree, this is caused by the expressive nature of human language. Polysemy (word with multiple meanings), synonymy (different words with same meaning) and context sensitivity of a word or phrase are the primary reasons for the miss-interpretation of user inputs. In a narrow domain, these problems can be partly dealt with by applying techniques such as word dictionary, word stemming or thesaurus to reduce the ambiguity caused by the keywords. However, there is no CBMIR system to our knowledge that allows the user to construct more complex query such as "retrieve 5 images that looks 30% like the input image" or "red cat in the house".

6.2 Query by Image

In recent years, with the advancement of image processing techniques, query by image example has emerged as a preferred option for constructing searches in CBMIR systems. The reason being that query by image example can avoid the ambiguity issue surrounding with keyword query. Some systems also provide options for a user to specify the relative importance of each feature in the image, or functional features to let the user to manipulate the input image. All these extra options are designed for constructing queries that have a better representation of users' intention.

In query by image example, the query is constructed by extracting the relevant features from the input image and a search vector that uses these features. Weights can also be assigned to fine tune the importance of each element in the feature vector. Depending on the application, the weights of the feature vector can be explicitly assigned by users, or assigned by system through a system defined rule or relevance feedback from the user.

6.3 Relevance Feedback

The use of relevance feedback together with ranking is a means for the system to iteratively fine tune the feature vector through feedbacks from the user. It has not been found that any of the CBMIR systems are employing this technique. This may due to the fact that most of the CBMIR systems reported only apply to a very narrow scope domain. Nevertheless, if researchers are interested in integrating different medical modalities together, there is no doubt that such approach should have its place in CBMIR systems. Interested reader should refer to reference [10] for a comprehensive review on this technique.

7. Feature Extraction

Feature extraction is a core feature of any CBIR systems. This module is either directly or indirectly related to all the different components in a CBIR system. In-fact, the selection of the indexing structure and design of the query processing unit is directly affected by this module. In the following paragraphs, the attention focuses on the discussion of visual features. A discussion on semantic feature is left to the subsequent section.

In the general CBIR systems, the standard approach for features extraction is to extract features such as colour, texture and shape. In the more advanced systems, the geometry information of the object/s in the image can also be extracted and analysed. The indexing structure is then constructed base on the combinations of these features information. These features in a general domain system are extracted and analysed by generic image-processing approaches. At the minimum, these image-processing approaches have to be robust and insensitive to the image size, viewing orientation and in some cases, object occlusion. For each of these visual features, there are various representations that represent the human perception of that feature from different perspectives. What features and representations to use will depend on the application, and sometimes, the selection process can be an art form in itself.

In comparison to the general domain retrieval system, the features selection process for CBMIR systems tends to be straightforward. Domain specific applications such as diagnostic systems [2, 3, 5] can apply their domain knowledge to assist the selection of important features required in identifying the disease, tumor or condition that the specialist is interested in.

8. Indexing

In order to make any CBIR systems truly scalable for large size image collection, the images are required to be indexed in a systematic manner. In a traditional database system, the data is indexed by a search key or combination of keys that uniquely identify an individual record. Often, a simple one dimensional data structure is adequate enough for indexing the data in such systems. However, images are more complex. Attempts to reflect this complexity usually results in images being represented by a set of values or attributes, commonly known as the feature vector. When represented in this manner, each value in the set becomes a point in an n-dimensional space, implying a multi-dimensional structure is required.

So far, the research efforts for indexing structures applied to CBMIR systems have been mostly revolved around two issues, and they are:

1. What data to be indexed?
2. How is the data organized?

These two issues are rather common in database and data structure communities. However, with the complexity of images and the high dimensionality of the visual features, the answers to these two questions may not be as trivial as it is for the traditional text database systems.

8.1 Indexing Value

In the previous section, we have discussed the possible visual features that can be used in indexing the image databases. However, visual features are only one of the possible features that can be used for indexing images. Depending on the application, image index structure can also be grouped by keywords, which is a great tool for capturing the semantic content of the images. In some cases, the image database may be better represented with the combination of semantic and visual features.

Cha and Chung [11] have proposed an approach that allows an image to be indexed by three separate indexing trees, and these index trees are: visual features, semantic and keyword indexing. The visual features are the combination of shape, color and texture. The semantic features of the system consist of a set of predefined attributes, and the keyword features are texts that are entered by users. The values of these attributes are stored in metadata format. These separate indexing structures provide the user with the flexibility to construct very complex query.

8.2 Indexing Structure

Indexing structure has been an interest for researchers for many years. This is mostly because it is essential to have a fast and efficient indexing structure for the database system to be scalable. As for CBMIR systems, many researchers have added two addition requirements to the system's indexing structure. The indexing structure has to be multi-dimensional and dynamic.

Multi-dimensional index is a structure that is often used in indexing large and complex data. These data include audios, videos, images and etc. Indexing tree is the most common used indexing structure for image database, and there are different types of indexing trees designed to accommodate different query requirements. Reader can refer to Reference [1] for a comprehensive review on the difference tree-based indexes available for image data.

One of the issues in applying indexing tree is the dimensionality of the index. The performance of the multi-dimensional indexing structure such as popular R-tree and R*-tree degenerates drastically with an increase in the dimensionality of the underlying feature space, this is mostly because the trees' fan-out decreases in inversely proportional to the dimensionality. To solve this problem, one promising approach is to first perform dimension reduction and then to use appropriate multidimensional indexing techniques for searching and retrieving images.

Even though the dimension of the feature vectors in most of the image retrieval systems is very high, not all the features posses the discriminatory power for being able to uniquely identify the images. There are various approaches in identifying the importance of an attribute in the feature space. Karhunen-Loeve transform (KLT), also known as principal component analysis (PCA), can be applied to identify the importance of the features in the principle component space. Different from KLT, Park et al. [12] used Quasi-Gabor Filter to reduce the dimensionality of the texture features.

Clustering is another powerful tool in performing dimension reduction. This approach cluster similar features together to perform recognition or grouping. This type of clustering is called row-wise clustering. Similarly, Zhang and Zhong [13] used self-organization map (SOM) neural networks as the tool for constructing the tree indexing structure. This approach provides the advantage of unsupervised learning ability, dynamic clustering nature, and the potential of supporting arbitrary similarity learning.

9. Conclusion and Remark

This paper briefly discussed the major components in the CBMIR systems, and a comparison is drawn with the generic CBIR systems wherever possible. Although the issues faced by both systems are common, the design approaches for the systems are quite dissimilar. Also, the approaches in designing the system components are vastly different. The strategies adopted in designing components such as user interface, query processing and the feature selection unit for the two systems are also vastly different. This is mostly because the scope of CBMIR systems is bounded by the systems' knowledge domain. Hence, certain assumptions can be made in CBMIR systems. It is recognised that CBMIR systems are still in an early stage of development. This is evidence in the lack of systems reported from the literatures. This should not be a surprise as the systems reported are only for diagnostic, decision support or teaching purpose. The application for these systems is very specific. The approach applied in these systems is also not generic enough to be transferred to other applications. Hence, there are ample rooms for future work in expanding the functionalities of both systems.

References

[1] H. D. Tagare, C. C. Jaffe, and J. Duncan, "Medical Image Databases: A Content-based Retrieval Approach," *Journal of the American Medical Informatics Association*, vol. 4(3), pp. 184-198, 1997.

[2] P. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, and Z. Prtopaps, "Fast and Effective Retrieval of Medical Tumor Shapes," *IEEE Transactions on Knowledge and Data Engineering*, vol. 10, pp. 889-904, 1998.

[3] C.-R. Shyu, A. Kak, C. E. Brodley, and L. S. Broderick, "Testing for Human Perceptual Categories in a Physician-in-the-loop CBIR System for Medical Imagery," presented at IEEE Workshop on Content-Based Access of Image and Video Libraries, Fort Collins, Colorado, 1999.

[4] D. Comaniciu, P. Meer, D. Foran, and A. Medl, "Bimodal System for Interactive Indexing and Retrieval of Pathology Images," presented at 4th IEEE Workshop on applications of Computer Vision, Princeton, New Jersey, 1998.

[5] L. R. Long, S. K. Antani, and G. R. Thoma, "A Prototype Content-based Image Retrieval system for Spine X-rays," presented at 16th IEEE Symposium on Computer-Based Medical Systems, New York, New York, 2003.

[6] C. T. Liu, P. L. Tai, A. Y.-J. Chen, C.-H. Peng, and J. S. Wang, "A Content-based Medical Teaching File Assistant for CT Lung Image Retrieval," presented at The 7th IEEE International Conference on Electronics, Circuits and Systems, Lebanon, 2000.

[7] Y. Rui, T. S. Huang, and S.-F. Chang, "Image Retrieval: Current Techniques, Promising Directions, and Open Issues," *Journal of Visual Communication and Image Representation*, vol. 10, pp. 39-62, 1999.

[8] T. M. Lehmann, B. Wein, J. Dahmen, J. Bredno, F. Vogelsang, and M. Kohnen, "Content-Based Image Retrieval in Medical Applications: A Novel Multi-Step Approach," presented at Proceedings of SPIE, 2000.

[9] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jai, "Content-Based Image Retrieval at the End of the Early Years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22(12), pp. 1349-1380, 2000.

[10] X. S. Zhou and T. S. Huang, "Relevance Feedback in Image Retrieval: A Comprehensive Review," *ACM Multimedia Systems*, vol. 8 No. 6, pp. 536 - 544, 2003.

[11] G.-H. Cha and C.-W. Chung, "An Indexing and Retrieval Mechanism for Complex Similarity Queries in Image Databases," *Journal of Visual Communication and Image Representation*, vol. 10, pp. 268-290, 1999.

[12] M. Park, J. S. Jin, and L. S. Wilson, "Fast Content-Based Image Retrieval Using Quasi-Gabor Filter and Reduction of Image Feature Dimension," presented at Fifth IEEE Southwest Symposium on Image Analysis and Interpretation, Santa Fe, New Mexico, 2002.

[13] H. J. Zhang and D. Zhong, "A Scheme for Visual Feature based Image Indexing," presented at Proceeding of SPIE Storage and Retrieval for Image and Video Databases, 1995.