2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing

Image Retrieval in Multimedia Databases: A Survey

[1]Yihun Alemu, [2]Jong-bin Koh, [1]Muhammed Ikram, [2]Dong-Kyoo Kim
Department of Computer Engineering, Ajou University
South Korea, Suwon
[1]{yihuna,engr.ikram}@gmail.com, [2]{nitefly,dkkim}@ajou.ac.kr

Abstract
In the past decades the advancement in the area of database management systems shifts towards multimedia. Multimedia information is very expressive, self explanatory, narrative, etc. Now a day the development of digital media, advanced network infrastructure and the easily available consumer electronics makes the multimedia revolution to run in an alarming rate [4]
Inline with the advancement of database technology that incorporates multimedia data, an open question that always rose in the technology is how to retrieve/search images in the multimedia databases. There are a huge number of research works focusing on the searching mechanisms in image databases for efficient retrieval and tried to give supplementary suggestions on the overall systems.

The growing of digital medias (digital camera, digital video, digital TV, e-book, cell phones, etc.) gave rise to the revolution of very large multimedia databases, in which the need of efficient storage, organization and retrieval of multimedia contents came into question. Among the multimedia data, this survey paper focuses on the different methods (approaches) and their evaluation techniques used by many of recent research works on image retrieval system.
Many researchers develop and use lots of approaches towards image retrieval. This paper, in general, classified image retrieval into text based and content based, including the newly growing ontology based image retrieval system as one focus.
We address, in this paper, the challenges, techniques and evaluation methods used in image retrieval systems through a detailed look of most recent works

1. Introduction

Multimedia databases have emerged as the natural solution to manage the vast growing of multimedia information. Multimedia databases and the internet not only enhance each other, but have a dramatic impact on developers, network administrators, content providers and users [17]. By its very nature, multimedia databases need a large storage area than other conventional databases since they stores mainly of images and audio video. An image database contains a large collection of images with similar features that makes the querying mechanism problematic. And currently, many of the research works on image databases rely on low level features-either text or image features, that lead it with a number of limitations to get an exact query result.
The task of image retrieval is to find and retrieve the most similar figure for a given query. However, image retrieval is a complex process that inherits techniques from different fields like pattern matching, information retrieval and computer graphics.
Typically image in multimedia database is searched based on keywords, features and/or concepts. But the problem, specifically in feature based retrieval, lies on the vast number of attributes to express the image: size, color, shape, texture, location, position, domain, etc. Moreover, the complexity in the nature of two-dimensional image data gives rise to a host of problems that alphanumeric information systems were never designed to handle [5].
Given tremendous amount of image data, capabilities to support efficient and effective image retrieval have become increasingly important. There are two general approaches for image retrieval: the *text-based approaches* apply traditional text retrieval techniques to image annotations or descriptions where as the *content-based approaches* apply image processing techniques to extract image features and retrieve relevant images. One obvious approach is to describe the image contents verbally, typically using keywords. Once the verbal descriptions are obtained, text search techniques can be applied to retrieve images in the database allowing query-by-keyword, but this assumption is seldom met since manual labeling is too expensive. The other approach is to represent images with nonverbal descriptions which can be reliably computed from images. Such descriptions are image features based on color, shape, and texture. Conventional content-based image retrieval (CBIR) methods use these image features to define image similarity [21].

In general, both in text based and content based image retrieval systems, the bottleneck to the efficiency of the retrieval is the semantic gap between the high level image interpretations of the users and the low level image features stored in the database for indexing and querying. In other words, there is a difference between what image features can distinguish and what people perceives from the image since human perception is complex and seems to be dependent on context, purpose,

978-0-7695-3762-7/09 $26.00 © 2009 IEEE
DOI 10.1109/IIH-MSP.2009.159

emotion, psychological ground and many more individual cases.

2. Text Based Image Retrieval

With the development of the Internet, and the availability of image capturing devices, efficient image searching, browsing and retrieval tools are required by users from various domains, including remote sensing, fashion, crime prevention, publishing, medicine, architecture, etc. To address these problems, many general purpose image retrieval systems have been developed under two basic frame works: text based and content based [1].

It is back to 1970 that the text based image retrieval has got its attention. Early on, keyword-based search has become the leading paradigm for querying multimedia databases; while it has lots of limitations.

In text based image retrieval systems the first thing to do is providing textual descriptions/annotations for images, which is very tiresome, inefficient, and expensive and has a problem of undesirable mismatch due to annotation impreciseness. Text annotation is extremely tedious in large image collections. To provide text descriptions or annotations, two approaches can be applied. The first approach acquires descriptions/annotations *manually* by human annotators. The second approach is to *automatically* annotate images using machine-learning techniques that learn the correlation between image features and textual words from the examples of annotated images [3]. In both cases manual labeling is too expensive while automatic methods are not reliable. Automatic image classification yields limited access, in such a way that only few objects (like faces or cars) can be recognized reliably from general images [5], [21].

Another problem in the use of keywords is the complexity between words and concepts due to synonymy (different words denote the same concept) or homonymy (same word denotes different concepts) and many search criteria cannot be well described by a few keywords [12]. Current researches [12], [3], [7] conclude that key word based searching can not be successful alone. In finding a target image, the content of the image and the image features play an important role. It is also suggested that [12] since images might have no annotations, or incompletely annotated, a joint use existing text annotations and visual features can provide a better retrieval results.

In recent years indexing and retrieval approaches based on keywords and visual features together are getting attentions. A work by [27], which can be a good example of such approach, solved to some extent the problem of synonymy and homonymy using LSI (latent semantic indexing). LSI can also be applied to the joint visual and keyword-based feature vectors in order to find a hybrid reduced representation that links sets of keywords and images. Unfortunately, to identify meaningful relations between keywords, LSI needs high amounts of data. This requirement can only be met when a relatively large quantity of text—rather than just a few keywords—is associated to every image.

It is those important difficulties that boosted research activities in the field of CBIR in the early 1980s. CBIR came into being to solve those problems inherited by text based systems, and currently many of the researches in image retrievals concentrated on the advancement of CBIR systems in many directions from simple low level features to a combined visual and human interactive systems for better achievement. The next part of this survey paper reviewed in detail about such systems with respect to the current state of the art.

3. Content Based Image Retrieval (CBIR) systems

Many of today's image retrieval systems rely on CBIR with varied techniques ranging from single feature vector to combined visual and conceptual image content descriptions and ontology. This survey tried to give detailed view of most of the traditional and recent CBIR systems with respect to their methodological approach and performance analysis. In CBIR, images are indexed by their visual content, such as color, texture, and shapes. Moreover, the fundamental difference between content-based and text-based retrieval systems is that the human interaction is an indispensable part of the latter system. Humans tend to use high-level features (concepts), such as keywords, text descriptors, to interpret images and measure their similarity. Eventhough the field of CBIR has been extensively researched in recent years, none of the proposed approaches has achieved satisfactory performance due mainly of the semantic gap; expressing the discrepancy between the low level features that can be readily extracted from the images and the high level descriptions that are meaningful to the users [1], [12]

Image content can be described at various levels. It may regard perceptual features (also known as content-dependent metadata) like color, texture, shape, structure and spatial relationship, or semantic primitives (also known as content-descriptive metadata) such as the identification of real-world objects and the meaning of the images [19] , and image retrieval using low-level visual features is a challenging and important issue in content-based image retrieval. However, most of the CBIR systems focused on perceptual (low level) features.

3.1. Low level (basic) features

Multimedia contents, in general, can be represented as keyword based, feature based and/or concept based;

most of the CBIR systems based up on the basic low level features (color, texture, shape and spatial relationship). As most researchers agree, due to the difficulty of inferring semantic meaning from low level features, none of CBIR systems are satisfactory. Including yahoo and Google, many well known search engines currently are limited to textual keywords. Due to the increasing demand of better management and retrieval of multimedia data, the next generation multimedia databases are looking for better performance from CBIR systems. Despite the difficulty of extracting exact images features, the vital step in CBIR system is feature extraction. From the vast amount of research works based on low level features, color is the dominant one due to its robustness and independent to the size and orientation of images.

It is also said by [28] as, most systems use color and texture features, few systems use shape feature, and still less use layout features. The retrieval on color usually yields images with similar colors. Retrieval on texture does not always yield images that have clearly the same texture, unless the database contains many images with a dominant texture. Searching on shape gives often surprising results. Apparently the shape features used for matching are not the most effective ones.

Color is a visual feature which is immediately perceived when looking at an image. It is one of mostly used visual features in retrieval and it can also be used to find the location in an image and to differentiate a large number of objects [22]

Texture is an innate property of virtually all surfaces. It represents the regularity, smoothness/coarseness of the image. Texture gives a direction sense to the spatial arrangement of image intensities [22]. Color histogram, color coherence vector (CCV) and color moments (CMM) are the common approaches for color feature representation of an image where as filter banks and AM-FM models have been used as a common model to represent texture of an image. Those techniques, however, are all global methods in that they extract the visual metadata of the whole image (globally). Global methods have the advantage to have a compact representation and that the extracted visual metadata can be, under certain constraints, efficiently compared and indexed using a Spatial Access Method. The main disadvantage is that in these approaches, there is no information about the spatial distribution of colors inside the images. Thus, images with very different spatial layout may have similar representations [19]. In order to take spatial distribution into account, several regional methods have been developed [19], [14], [25], [21], [1]. The color based clustering (CBC) system is the one advocating regional systems. In regional methods, an image is segmented in a set of regions accordingly to a predefined visual property and each region is represented individually. [19] uses a single linkage clustering algorithm, which is a hierarchical type agglomerative algorithm together with their defined distance function for similarity matching.

In their experiment, the comparison has been done against the global (GCH, CMM and CCV) and partition based (Grid and CSH /color shape histogram/) color histogram approaches based on 20,000 Corel Corp Jpeg images, and the CBC system out performs all the above 5. However the CBC system didn't compare with other low level feature based systems to check its performance.

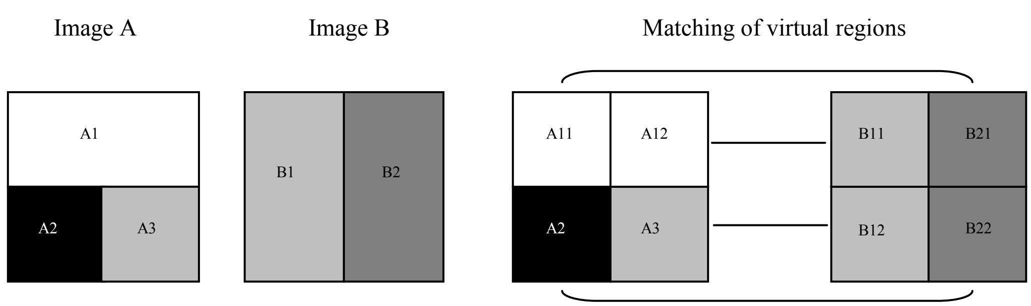


Fig. 1 decomposition of real regions into virtual regions for matching, CBC system (taken from [19])

To increase the effectiveness of color based image retrieval, multi resolution similarity technique has been proposed [14]. To improve the quality of color histogram search and to get a better query result, [14] proposes to take partitions/sub histograms instead of taking a histogram as a whole to represent an image; this will allow the user to search on different resolutions based on a hierarchical histogram representation.

The multidimensional histogram, actually, is a bit complex but it is a good alternative for the common histogram based image similarity search. In their system, an image is divided into regions with homogenous color distribution and a histogram is developed to each region of an image. Such statistical information of an image serves as abases for similarity measures. The system differs from other related work in that the similarity measures lies on the sub histogram families than the entire histogram and the effectiveness measure is independent of the size of the result set.

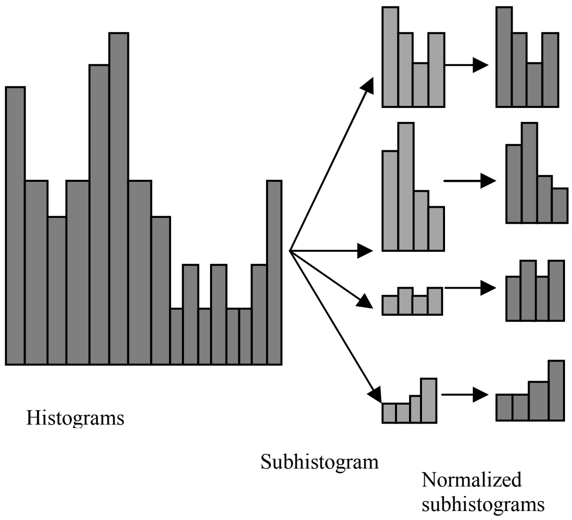


Fig.2 Histograms with subdivisions and normalizations (taken from [14])

The system uses sub histograms of images and a similarity measure for the sub histogram than the entire histogram as a whole and it goes on through a hierarchical similarity to find the target image. The system is evaluated based on a new effectiveness measure in addition to the common precision/recall measures. The effectiveness measure has four parameters; the number of relevant images in the database (R), the number of returned images (E), the number of returned relevant images (Rr), and the number of missed relevant images (Rm = R−Rr) , where the detail of the formula is there in the reference. The system is compared with several approaches (Adapt, WBIIS, Color histogram and Color layout) and it proves to be better than those approaches, however it needs still some level of improvement by incorporating multiple image features. Further more, it opens an issue for how much the level of detail goes for histogram subdivision to support user's queries, which is not answered there. And since subdivision of histograms have no limit; it may lead to an unmatched query result.

A further recent work to low level feature based CBIR incorporates users. Incorporating human intuition and emotion into retrieving images is vital for effective result [23], [20]. Such kind of human oriented systems facilitate the search in both explicit and implicit queries. In such method, [23] use IGA (interactive genetic algorithm) together with wavelet transformation. The IGA adopts the users' choice as fitness and a user can increase/decrease the effectiveness of a color indefinitely and interactively until he/she gets his/her desire. However the efficiency test [23] used is not good, they evaluate it with respect only to comparing with their own past work and use ten students for performance analysis. Despite their evaluation, it is new to use psychological test in their work for measuring the users' satisfaction

With same color feature again, attribute combination is also a new alternative approach for better image retrieval. As people cognition to color is considered to be multi layer and multi profile, attribute combination of different granularity, under local and global color feature based systems, can extract color features of individual child/segmented blocks of images more efficiently [25]. Image retrieval based on quotient space granularity theory includes two parts: describing images under different granularities to obtain attribute features and select a reasonable optimal guide line function to combine attribute functions in different levels [25]. It adopts global color feature based on histogram and local color feature based on DCT and SVD. It combines the color histogram approach (global) and local color feature extraction (DCT and SVD) with multi level quotient space granularity combination and get a better performance with respect to those independently applied methods.

3.2. Combined features

Using a combination of visual features and concept oriented interactive manner yields a more effective search result in image databases compared to the traditional CBIR system using single visual feature and simple linear combined low level visual features. To address the challenge of semantic gap reduction for image retrieval, many of recent CBIR systems tried to make full use of image information to extract features.
Many methods have been used in CBIR system, including methods based on color [19], [5], [14], [23], [25], based on texture [2], based on shape [6], based on spatial relations [13] and so on. Most CBIR systems allow query formulation with user setting of relative importance of features (e.g., color, texture, shape, etc) to mimic the user's perception of similarity. Similarity matching is crucial in retrieval systems to understand new concepts with existing ones. Similarity matching function is multi level, until the most primitive distance measures or correlation metrics are used at the leaf level. They, [10], proposed a learning algorithm that adapts similarity matching function from ranking errors feedback by the user. User can reorder the retrieved images and the system uses both short term and long term (dubbed) learning in user specific mode. The

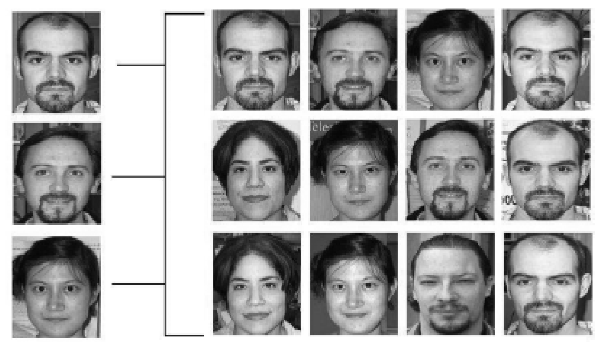
overall approach is good specifically for homogeneous data set; it was also tested based on facial features (chin, eyes, nose, etc.) and achieved a better result.

In CBIR), image has various inherent features which reflect its content such as color, texture, shape, spatial relationship features etc. How to organize and utilize these features effectively and improve the retrieval performance is a valuable research topic [24]. The techniques of organizing and searching images based on their content are still in its early stage. And research by content now is very demanded specially in the area of multimedia databases. In similarity measures using shape, texture, color, etc. weight has limited expressive power, for instance the size and color of an image may not be equally treated with shape of the same image. Such and related difficulties lead [11] to propose another alternatives-the object query language for describing the image in terms of regions and properties of regions.

In their scenario, a user can construct, with the help of the system, his/her query automatically based on a set of images selected and/or rejected by the user. In addition, to have a sufficient number of similar properties, the system composes a set of regions having a vast amount of properties like saturation, intensity, area, perimeter, adjacency, etc. using an efficient segmentation process. This work uses a genetic algorithm with abstract data types and inheritance classes to represent example and counter example images in different regions. The structure of the algorithm uses 9 colors, 5 sizes, 5 vertical and horizontal positions. It contributes a user friendly query based system on simple predicates and unary/binary operations, it also avoids the explicit weight measures and uses properties list by end users based on the given example and counter example images; and it is flexible for reusing the system for other cases by incorporating additional features through its object orientation nature [11].

Some of the limitations, they pointed out, to be adjusted later through extra works include; segmentation and semantic overhead to express images, no relevant feedback mechanism to facilitate subsequent queries and it misses the idea of aggregate functions to compare images from sets of regions in a many-to-many relationship. Even though it is efficient for geometric figures (mainly of regular polygons), it might be challenging for area, intensity, color, etc of irregular shaped bodies.

Similarly, [9] proposed multiple query examples to retrieve not only images similar to individual examples but also images which actually represent a combination of the content of query images. Using machine learning, for generating the most appropriate feature combination from multiple examples, can yield a better search result. The approach by [9] implemented only for facial images in which the database of images are assumed to be of same type, which is for limited domain. The usual precision/recall performance measure has been used based on the number of independent components and window size, and evaluate their system by varying these performance. To improve performance and to differentiate more informative and less informative regions in human facial images, the system incorporates learning.



Query images Retrieved images using regression analysis

Fig. 3. Results of facial image retrieval from a learned system where three query images are used (taken from [9])

Like the simple visual feature based traditional CBIR system, clustering/segmentation is one of the modern image retrieval approaches proposed by many researchers. An approach proposed by [22] uses a new clustering algorithm that converts image information into a ratio (Fisher Discriminant) that will depend on the inter cluster and intra cluster distances. The Fisher Discriminant is effective in classification of images and provides a logical group of the image pixels, not only considers the means of the classes but also the variance of each class in the classification process [22]. This work also employs a distance function to rank images in the database according to their similarity. The similarity is measured as the distance between the statistical distributions like means and variances that represent the cluster centers and the pixel attributes. A widely used clustering method is the *k-means* algorithm where the pixels in the image are initially assigned to a fixed number of classes. The main disadvantage of this approach (k-means) is that the process has to be iterated many times before the final classification yielded, and also the effect of each pixel on the objective function is not taken into consideration. In contrast, [22] used the *j-means* algorithm to overcome those problems in k-means. In the retrieval process, two features are extracted from each segment; color (based on histogram) and texture (being represented by calculating busyness factor t for each pixel). Moreover, the classification using J-means clustering is very effective in the sense that for each pixel assignment the means and the variances of all the classes are taken into account. After the first iteration, most of the pixels are assigned to proper classes. The spatial constraints are activated in the subsequent iterations. One major problem with this spatial term is at the edges, because at the edges the neighboring pixels need not belong to the same class, and hence they introduced an edge pixel detector so that the spatiality is not enforced on the edge pixels.

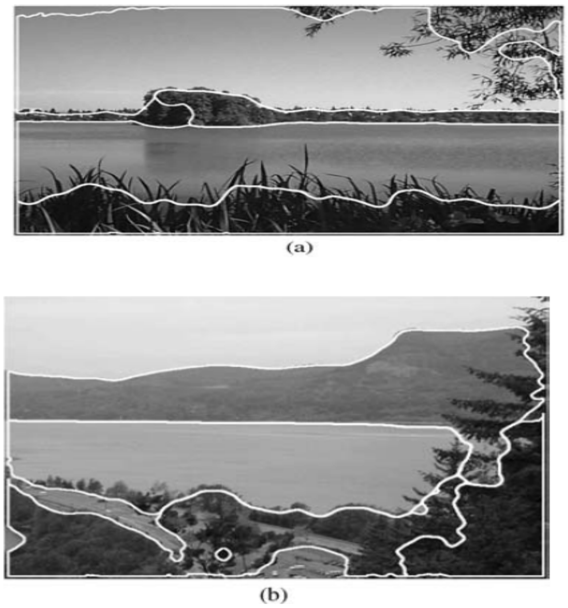


Fig. 4 the clustering of color images (taken from [22])

This work, [22], used four different kinds of evaluation techniques; precision, recall, false alarm rate and miss rate and it is compared with three image retrieval approaches in which two of them were clustering-based image retrieval using J-means clustering and k-means clustering respectively and the other one is image retrieval using global features. It is observed that the new j-means algorithm is better in retrieval. Even though including spatial terms has a significant impact in retrieval process, it would have been more expressive if it uses additional features (than color and texture) with an advanced user interactive algorithms. In addition, automatic and self organized computation of the number of segments can lead the system for better retrieval.

It is obvious that, as the combined features increase, let alone the time and space complexity, the retrieval effectiveness also get improved. A system developed by [15] is based on the combination of four image features; color feature (HSV color histogram), texture feature (co-occurrence matrix), shape feature (moment invariant based-on threshold optimization), and spatial relationship feature (based-on the Markov chains). As said by [15], Color feature is often broadly used to describe the images which are difficult to be segmented and needn't to consider space information. Texture feature is more efficient to describe images with complicated textural content such as lawn, sandlot, etc. Shape feature can describe the shape information of image object well, especially when the image with clear edges and object. Spatial relationship can describe the inter-relationship of objects in an image.
Segmentation based on image feature vectors is different from the classical segmentation which is based on spatial area. With this idea, [15] used ISODATA clustering method for image segmentation.

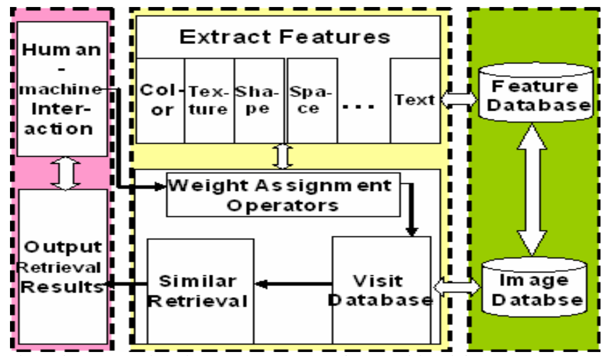


Fig. 5 image retrieval system structure for combined features

An interesting approach in their experimental method is composing different images from different domains (landscape, animal, cars, constructions, etc), and the system is compared with single low level feature based retrieval systems.
A recent work on CBIR [26] incorporates granularity theory to the combined low level visual features. They propose color information of image (based on histogram) and texture feature (based on gray level co-occurrence matrix/GLCM/) as both methods are considered to be quick and effective classical algorithms. The work contributes much in the application of granularity theory in image retrieval, but it also lacks on comparing it to other combined feature approaches. Here in [26], a better result obviously obtained compared to methods adopting only single retrieval algorithms. The absence of automated/intelligent behavior can be taken as one limitation to be included for further work and their optimal guideline function is not fully described too, as stated there.
To make retrieval more intelligent and interactive, a CBIR system should also incorporate other related complementary disciplines such as the knowledge based system, computer vision, data mining, neural networks and so on. A more recent approach proposed by [24] focuses on intelligent concept oriented search. As they said, traditional image retrieval based on visual-based matching is not effective in multimedia applications. Consequently, the modeling of high-level human sense

for image retrieval has been a challenging issue over the past few years. Most of contemporary CBIR systems explored to look for the primary features to present an image. However, the related efforts are not satisfactory enough due to the gap between low-level features and high-level concepts, and hence problems still exist in traditional classification-based image retrieval. For example, different users making different query-terms perhaps want to look for the same kind of images.

Here, [24] adopts the application of data mining together with query decomposition techniques for image searching. Their intelligent concept oriented search (ICOS) uses the image-concept query decomposition (ICQD) technique to decompose a query image into several concepts to touch the user's mind.

Concept oriented search get the desired high level image through the help of ICQD. It includes effective annotation, association, mining and visual ranking for conceptual objects and has intelligent search method to enhance high level concept image retrieval.

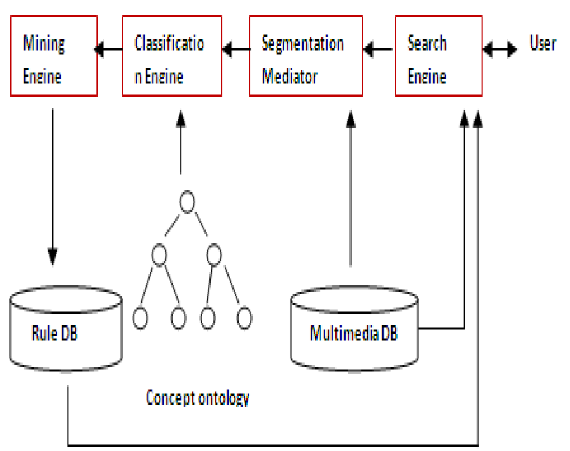


Fig. 6 the frame work of ICOS (from [24])

The system has two major phases; training phase (to generate association rule) and concept retrieval phase (in which the images matching the concept queries from the users). It also uses an integrated classification, mining and search engines.

The system has been evaluated based on precision, F-measures (for image-concept query decomposition under different classification thresholds) and performances for different sizes of experimental data. One of its strong side, unlike most CBIR systems, is evaluating the efficiency with a database of images of non fixed size. It looks better than the traditional CBIR systems using sequential comparison strategy, despite the overheads on concept ontology.

Above are the detailed reviews on major recent works of CBIR systems. Researchers in the area are still doing more on image retrieval by expanding the scope to the advanced retrieval strategy incorporating high level semantic and ontology scenarios.

Semantic based image retrieval: it is among the recent works of image retrieval systems that focuses mainly on the techniques which can reduce the semantic gap of CBIR systems [1], [12]. An improved support vector machine (SVM)-based active relevance feedback frame work together with a hybrid visual and conceptual content representation and retrieval is one approach proposed by [12]. The approach employs the global color, texture and shape where the shape content is described by Hough transform. For the semantic resource it uses WordNet as it is better in many aspects to that of Cyc and ConceptNet. However computing a conceptual feature vector for every image is the major overhead of the method. From the experimental evaluation of [12], employing both visual and concept-based feature vectors visibly improves the quality of the results compared to using visual features alone, and projecting the keywords on all the key concepts gives better performance than projecting only on their key super-concepts. Further more, [12] believed also that the joint use of the conceptual content representation together with their new relevance feedback framework contribute to a significant improvement of the retrieval results.

A similar approach has also been proposed by [1], using fuzzy based system to decrease the semantic gap in CBIR systems. In the system, a fuzzy modeling approach has been developed to model the expert human behavior in the image retrieval task, which makes the system more powerful and expert based.

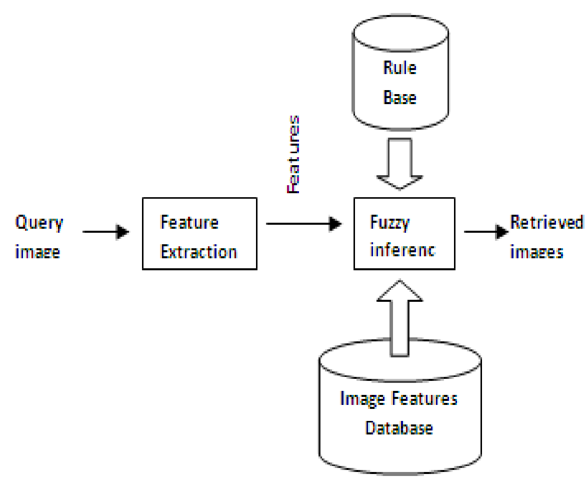


Fig. 7 structure of the proposed fuzzy system for image retrieval (as given by [1])

The evaluation shows that, the proposed fuzzy system improves the precision–recall performance of the CBIR systems.

Ontology based image retrievals: it is beyond the two major image retrieval paradigms, text based and content based. In practical applications both have limitations. Ontology-based image retrieval has the potential to fully describe the semantic content of an image, allowing the similarity between images and retrieval query to be computed accurately [8], [7]. Semantic technologies like ontology and the XML markup language provide tools for a promising new approach to image retrieval based on implementing semantic understanding of image content. Ontology-based image retrieval has two components: semantic image annotation (focuses mainly on the description of image content) and semantic image retrieval (to allow searching and retrieval based on image content).
Ontology refers to a formal representation of a set of concepts within a domain and the relationships between those domains, and semantic annotation is to describe the semantic content in images and retrieval queries. Through semantic annotation, both images and retrieval queries can be formalized as XML files [8]. An ontology based image retrieval system basically uses these two concepts: ontology and semantic annotation. Besides this, semantic annotation of an image or retrieval query still needs the intervention of a human being which is an open research area making it automated and interactive similar to some of the combined visual feature based CBIR systems do.

Another advanced method proposed by [7] is the ontology with multi modality based image retrieval system. Researchers argue that such ontological searches for images are very efficient especially for images with complicated content and ambiguous semantics. Single modality refers to either the text or image features of a domain specific image, where as multimodality refers both. The goal is creating machine-process able queries for such kind of diverse domain specific images, in their example, using both text and image features.

Currently the work is applied to animal domain, categorizing the system into three aspects: first, domain knowledge of the animals including scientific name, distribution, habitat, etc, second, animal images and third the association between domain knowledge and image features for ontology construction. The ontology again has three main components: animal ontology, textual description ontology and visual description ontology. By further classifying each component hierarchically, the low level features of images will be converted to a set of terms and incorporate them into the knowledge base.

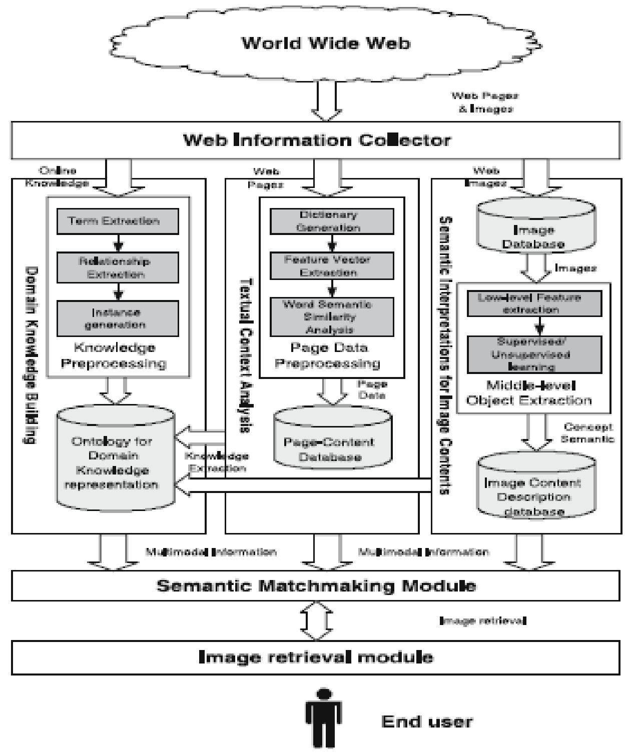


Fig. 8 work flow of the system proposed by [7]

The system uses the widely known description logic for representing knowledge in terms of classes, and relationships between classes use a match making algorithm to bind user specified queries with the knowledge base. The system is compared with Google (key word based approach) and other ontology based approaches. Its important privilege lies on its flexibility, easy to add more domains with out changing the entire architecture. One major bottle neck in such system is the extra burden in creating the ontology, knowledge base construction and the accuracy of image feature classification.

4. Conclusion

A wide variety of researches have been made on image retrieval in multimedia databases. Each work has its own technique, contribution and limitations.
Inline with the advancement of science and technology, images play an important role in transmitting information. Hence, together with the mechanisms for storing and processing in multimedia databases, retrieval of those stored images is always a question in the area.
As a survey paper, we might not include each and every aspect of individual works, however attempt has been made to deal with a detailed review of the most common traditional and modern image retrieval systems

from early text based systems to content based retrieval and ontology based schemes. We review those works mainly based on the methods/approaches they used to come up to an efficient retrieval system together with the limitations/challenges and the evaluation mechanisms used. And we tried to give a constructive idea for further work in each approach we reviewed.

We observed that, most of the single visual feature base CBIR systems use color and the combined system use a combination of almost all visual feature types. Further more, many of the works in image retrieval have been done for online system, and hence further work is needed to adapt those techniques in enterprise levels. On the other hand, most of the methods didn't include the self learning scheme. A user's query sometimes may give no result due to new parameters used beyond the knowledge base collections; in such cases it is better that the system stores the new knowledge from the user and use for other subsequent queries.

We believe that our work is not a complete, but a part in multimedia data; since multimedia data includes both audio and video (for instance MPEG-7 is the known technology that covers a wide range of applications including DVD, CD and HDTV). Hence, a more comprehensive review containing all multimedia data types is needed for more contribution in the area.

References

1. Abolfazl L., M. Shahram Moin, Kambiz Badie: Semantic–Based Image Retrieval: A Fuzzy Modeling Approach. IEEE (2008)
2. B. Vermaa, S. Kulkarni. A fuzzy-neural approach for interpretation and fusion of colour and texture features for CBIR systems. Applied Soft Computing, 2004, 5(1): 119-130.
3. Chen Zhang, Joyce Y. Chai, Rong Jin: User Term Feedback in Interactive Text-based Image Retrieval. SIGIR'05, August 15–19, 2005, Salvador, Brazil.
4. Chrisa T. and Stavros C.: An MPEG-7 query language and a user preference model that allow semantic retrieval and filtering of multimedia content. Multimedia systems 2007, vol. 15 no 2(87 p)
5. Daniel L. Swets and John J. Weng: Efficient Content-Based Image Retrieval using Automatic Feature. IEEE 1995
6. George Gagaudakis, Paul L.Rosin. Incorporating shape into histograms for CBIR. Pattern Recognition, 2002, 35(1): 81-91.
7. Huan W· Song L · Liang-Tien Chia: Image retrieval with multi-modality ontology. Multimedia Systems (2008) 13:379–390
8. Hyvönen, E. Saarela, S., Viljanen, K.: Ontology based image retrieval. In: Proceedings of WWW2003 (2003)
9. Jayanta B., Koustav B., Santanu C.: Multiple Exemplar-Based Facial Image Retrieval Using Independent Component Analysis. IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 15, NO. 12, DECEMBER 2006
10. Joo-Hwee Lim, Jian Kang Wu, Sumeet Singh, Desai Narasimhalu: Learning Similarity Matching in Multimedia Content-Based Retrieval. IEEE Transactions on knowledge and data engineering, vol. 13, No 5, 2001
11. José Martinez and Stève Marchand: Towards Intelligent Retrieval in Image Databases. 1998 International Workshop on Multimedia Database Management Systems
12. Marin F., Nozha B., Michel C.: Semantic interactive image retrieval combining visual and conceptual content description. Multimedia Systems (2008) 13:309–322
13. Mario A. Nascimento, Veena Sridhar, Xiaobo Li. Effective and efficient region-based image retrieval. Journal of Visual Languages & Computing, 2003, 14(2): 151-179.
14. Martin H., Alexander H., Daniel K., Markus W.: Multi resolution similarity search in image databases. Multimedia Systems 10: 28–40 (2004)
15. Pengyu Liu, K. J., Zhuozheng W., Zhuoyi Lv: A New and Effective Image Retrieval Method Based on Combined Features. Fourth International Conference on Image and Graphics IEEE 2007.
16. Pentland, A., Picard, R.W., Sclaroff: Content-based manipulation of image database. International Journal of Computer Vision, Fall (1995)
17. R. B. Johnson: INTERNET MULTIMEDIA DATABASES. Printed and published by the IEE, Savoy Place, London WC2R OBL, UK. (1998)
18. Ren XuePing, Wan Jian, Xu XiangHua: An Approach ofImage Retrieval based on Bayesian and AAM. IEEE International Conference on Systems, Man, and Cybernetics October 8-11, 2006, Taipei, Taiwan
19. Renato 0. S., Mario A., Alexandre X.: An Adaptive and Efficient Clustering-Based Approach for Content-Based Image Retrieval in Image Databases. IEEE 2001
20. S. M. Ghazanfar, S. K. Hasnain: A Framework for Interactive Content-Based Image Retrieval. IEEE INMIS 2005
21. Shingo Uchihashi, and Takeo Kanade: CONTENT-FREE IMAGE RETRIEVAL BASED ON RELATIONS EXPLOITED FROM USER FEEDBACKS. IEEE 2005.
22. Sridhar R., Avula, Jinshan T., Scott T., Acton: An object-based image retrieval system for digital libraries. Multimedia Systems (2006) 11(3): 260–270
23. Sung-Bae Cho and Joo-Young Lee: A Human-Oriented Image Retrieval System Using Interactive Genetic Algorithm. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS PART A: SYSTEMS AND HUMANS, VOL. 32, NO. 3, MAY 2002.
24. Vincent. S., Ja-Hwung Su, Hao-Hua Ku, Bo-Wen Wang: Intelligent Concept-Oriented and Content-Based Image Retrieval by Using Data Mining and Query Decomposition Techniques. ICME 2008.
25. Xiangli Xu, Libiao Z., Zhezhou Yu, Chunguang Z.: Image Retrieval Using Multi-Granularity Color Features. IEEE 2008.
26. Xiangli Xu, Libiao Zhang, Xiangdong Liu, Zhezhou Yu, Chunguang Z.: Image Retrieval Using Multi-Granularity Features of Color and Texture. Fifth International Conference on Fuzzy Systems and Knowledge Discovery IEEE 2008.
27. Zhao, R., Grosky, W.I.: Narrowing the semantic gap—improved text based web document retrieval using visual features. IEEE Trans. Multimedia 4(2), 189–200 (2002)
28. Remco C. Veltkamp, Mirela Tanase: Content-Based Image Retrieval Systems: A Survey. Technical Report UU-CS-2000-34, October 2000
29. Atsuo Yoshitaka, Tadao Ichikawa: A Survey on Content-Based Retrieval for Multimedia Databases. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 11, NO. 1, JANUARY/FEBRUARY 1999