Text-based, Content-based, and Semantic-based Image Retrievals: A Survey

Mohammed Alkhawlani
Dept.of Information Systems, Faculty of Computers and Information, Mansoura University
Mansoura, Egypt

Mohammed Elmogy
Dept. of Information Technology,Faculty of Computers and Information, Mansoura University
Mansoura, Egypt
Email: melmogy [AT] mans.edu.eg

Hazem El Bakry
Dept. of Information Systems, Faculty of Computers and Information, Mansoura University
Mansoura, Egypt

Abstract—**Image retrieval from databases or from the Internet needs an efficient and effective technique due to the explosive growth of digital images. Image retrieval is considered as an area of extensive research, especially in content based image retrieval (CBIR). CBIR retrieves similar images from large image database based on image features, which has been a very active research area recently. The content, that can be derived from image such as color, texture, shape…etc., are called features. This paper will present a survey and discuss the current literature of different types of image retrieval (IR) systems. An overview of the important techniques in image retrieval will be discussed. Finally, some urgent challenges in IR, that have been raised recently, will be presented as well as possible directions for future research.**

Keywords: Content-based image retrieval, Text-based image retrieval, Semantic-based image retrieval,Automatic Image Annotation.

I. INTRODUCTION

In recent years, collections of digital images are created and increased rapidly. In many areas of academia, commerce, government, medicine, and Internet, a huge amount of information is out there. However, we cannot access or make use of this information unless it isorganized to allow efficientbrowsing, searching, and retrieval. One of the main problems is the difficulty of locating a desired image in a large and varied collection. While it isperfectly feasible toidentify a desired image from a small collection simply by browsing, more effective techniques are needed with collections containing thousands of items. Image retrieval attracts interest amongresearchers in the fields of image processing, multimedia, digital libraries, remote sensing, astronomy, databaseapplications, and other related areas[1].

Image retrieval has been a very active research area since 1970s, with the thrust of two major research communities: database management and computer vision[2]. Therefore,

image retrieval can be defined as the task of searching for images in an image database. As shown in Fig. 1, image retrieval techniques can be classified into three categories: text-based image retrieval (TBIR), content-based image retrieval (CBIR), and semantic-based image retrieval (SBIR). The next subsections will discuss these categories in more detail.
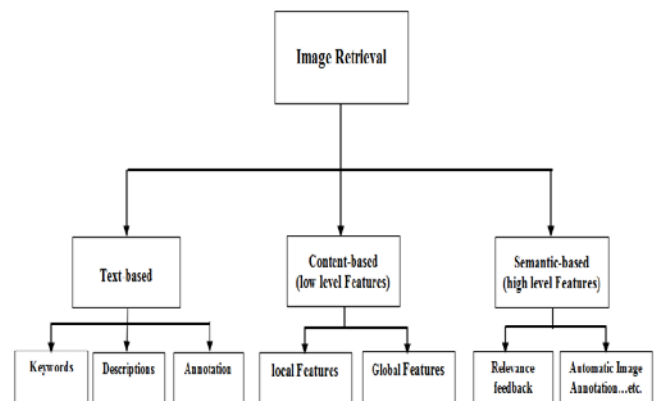

Figure 1. Image retrieval categories.

A. Text-Based Image Retrieval (TBIR)

TBIR can be traced back to the late 1970s. A very popular framework of TBIR was first annotated the images by text and then used text-based database management systems to perform image retrieval[2]. TBIR is used to manually annotate the image in the database with annotations, keywords, or descriptions.This process is used to describe both image contents and other metadata of the image such as: image file name, image and imageformat, image size,and image dimensions. Then, the user formulates textual or numeric queries to retrieve all images that are satisfying some of the

criteria based on these annotations,as shown in Fig. 2. However, there are some drawbacks in TBIR [3]. The first drawback is that the most descriptive annotations must usually be entered manually. Manually annotation for a large image database is impractical. The second drawback is that the most images are very rich in its content and has more details. The annotator may give different descriptions to images with similar visual contents. Also, textual annotations are language-dependent[4].
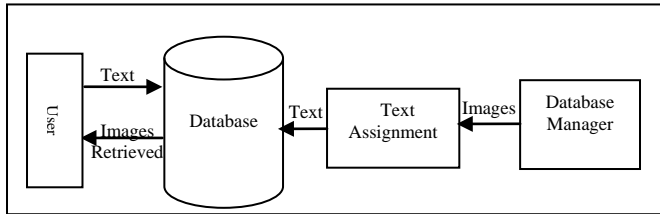

Figure 2. A typicalText-Based Image Retrieval system.

B. Content-Based Image Retrieval (CBIR)

CBIR is considered as an active and fast advancing research area. It is also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR)[5]. The term CBIR seems to have originated with the work of Kato [6] for the automatic retrieval of the images from a database based on the color and the shape. After that, the CBIR term has widely been used to describe the desired images retrieving process from a large collection of databasebased on image visual contents, normally called as features (color, shape, texture…etc.)[7]. In the early 1990s, as a result of the advances in the Internet and techniques of digital image production, a huge amount of digital images are produced in sciences, education, medicine, industry, and other fields available to the users that increased dramatically and make the drawbacks faced by TBIR became more and more tough. This needs formed the driving force behind the emergence of CBIR techniques [8]. The advances in CBIR researches mainly contributed by the computer vision community [2]. The used techniques and algorithms originate from many fields such as object recognition and signal processing [7]. However, in the last decade CBIR has received much attention which is motivated by the need to efficiently handle the rapidly growing amount of multimedia data[9]. It covers versatile areas, such as image segmentation, image feature extraction, representation, mapping of features to semantics [10].Research and development issues in CBIR cover a range of topics, most important are: understanding image users' needs and information-seeking behavior, identification of suitable ways of describing image content, extracting such features from raw images and matching query and stored images in a way that reflects human similarity.

As illustrated in Fig. 3,a typical CBIR is divided into off-line feature extraction and on-line image retrieval. In off-line stage, the system automatically extracts features of each image in the database and stores them in a features database

(features of the images are extracted and represented with feature vectors). In on-line stage, user input an image query to the system. The features of the query image are extracted and represented. The similarity was measured between the feature vector of the query image and the feature vectors of the images in the database. Then, the retrieval process is performed by applying an indexing scheme to provide an efficient way of searching the image database. Finally, the system returns the images that are most similar to the query image[11].
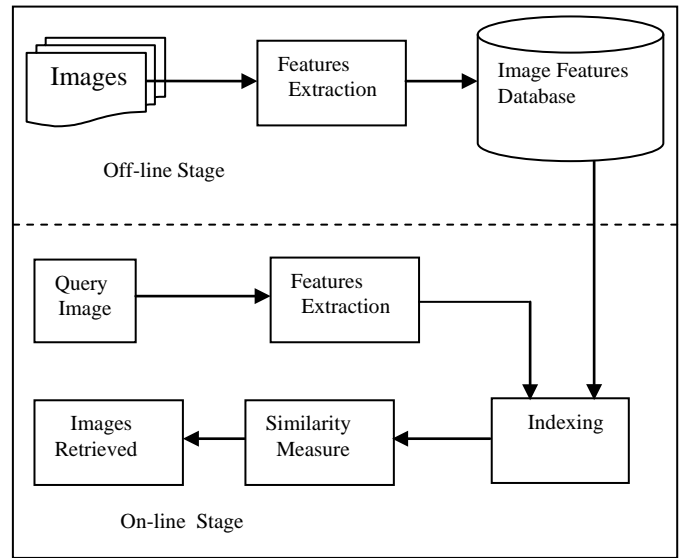

Figure 3. A typical Content-Based Image Retrieval system.

Another technique for image retrieval used to integrate text and image content to enhance the retrieval accuracy. Both the text and content-based techniques have their own characteristics, advantages, and disadvantages. By combining them, parts of their disadvantages can be overcome [12].

C. Semantic-Based Image Retrieval (SBIR)

In general, the problem of CBIR is the semantic gap between the high-level image and the low-level image. In other words, there is a difference between what image features can distinguish and what people perceives from the image. As shown in Fig. 4, SBIR can be made by extraction of low-level features of images to identify meaningful and interesting regions/objects based on the similar characteristics of the visual features. Then, the object/region features will go into semantic image extraction process to get the semantics description of images to be stored in database.Image retrieval can be queried based on the high-level concept.Query may be done based on a set of textual words that will go into semantic features translator to get the semantic features from the query. The semantic mapping process is used to find the best concept to describe the segmented or clustered region/objects based on the low features. This mapping will be done through supervised or unsupervised learning tools to associate the low-level features with object concept and will be annotated with

the textual word through image annotation process[1,13]. Semantic content obtained either by textual annotation or by complex inference procedures based on visual content[14].
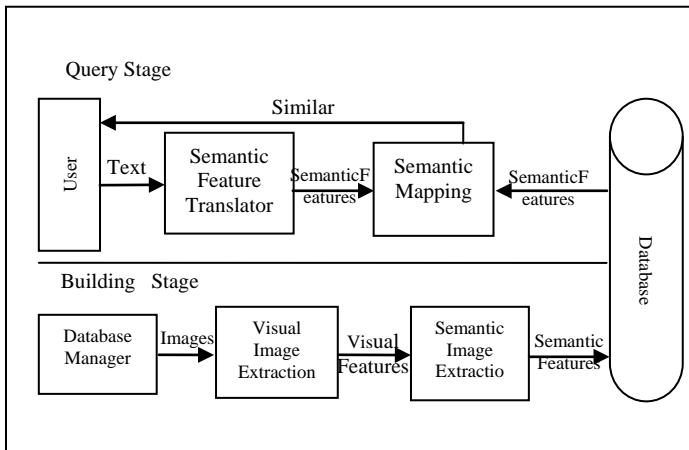

Figure 4. A typical Semantic-Based Image Retrievalsystem[1].

The rest of this paper is arranged as follows. SectionII gives an overview of the most common features used in image retrieval. SectionIII reviews current research in IR.Section IV lists the image retrieval challenges and introduces some of future research directions. Finally, the conclusion and the future work are presented in sectionV.

II. COMMON FEATURES FOR IMAGE RETRIEVAL

A feature is defined as capturing a certain visual property of an image[15]. In general, image features can be either global or local[16]. The global features describe the visual content of the entire image,whereas local features describe the regions orobjects (i.e. a small group of pixels) of the image content. The advantage of global extraction is its high speed for both extracting features and computing similarity. However, global features are often too rigid to represent an image. Specifically, they can be oversensitive to location and hence fail to identify important visual characteristics[15]. Local-feature approaches provide a slightly better retrieval effectiveness than global features[17]. They represent images with multiple points in a feature space in contrast to single point global feature representations. While local approaches provide more robust information, they are more expensive computationally due to the high dimensionality of their feature spaces and usually need nearest neighbors approximation to perform points matching[18,19]. Several important features that can be used in IR will be elucidated in the next subsections.

A. Color Features

The color has widely been used in IR systems, because of its easy and fast computation[20,21]. Color is also an intuitive feature and plays an important role in image matching. Most IR systems use color space, histogram, moments, color coherence vector, and dominant color descriptor represent color. The color histogram is one of the most commonly used color feature representation in image retrieval. The original idea to use histogram for retrieval comes from Swain and Ballard [21], who realized the power to identify an object using color is much larger than that of a gray scale[7,22].

Although the global color feature is simple to calculate and can provide reasonable discriminating power in image retrieval. It tends to give too many false positives when the image collection is large. Many research results suggested that using color layout is a better solution for image retrieval. To extend the global color feature to a local one, a natural approach is to divide the whole image into sub-blocks and extract color features from each of the sub-blocks. The advantage of this approach is its accuracy while the disadvantage is the general difficult problem of reliable image segmentation[23,24].

B. Texture Features

Texture is a property that represents the surface and structure of an image. Texture can be defined as a regular repetition of an element or pattern on a surface. Image textures are complex visual patterns composed of entities or regions with sub-patterns with the characteristics of brightness, color, shape, size…etc.[10].The commonly known texture descriptors are Wavelet Transform[25],Gabor-filter[26],and Tamura features[27].

C. Shape Features

Shape can generally be defined as the description of an object regardless of its position, orientation, and size. Therefore, shape features should be invariant to translation, rotation, and scale for an effective IR. In the direction of using shape as an image feature, it is necessary to determine object or region boundaries in the image and this is a challenge[28]. Compared with color and texture features, shape features are usually described after images have been segmented into regions or objects. Since robust and accurate image segmentation is difficult to achieve, the use of shape features for image retrieval has been limited to special applications where objects or regions are readily available. In general, the shape representations can be divided into two categories, boundary-based that uses only the outer boundary of the shape and region-based that uses the entire shape region[29]. The most successful representatives for these two categories are Fourier descriptorand moment invariants[2,30].

D. Spatial Location Features

Spatial location is also significant and is used for region segmentation. Spatial location is described as top/bottom, top left/right and back/front as per the position of an object in an image. For example, the sea and sky may have the same characteristics of texture and color but the spatial information is not similar. Sky typically represents the above portion whereas sea is at the below portion of an image. Hence, the spatial information of multiple objects in an image extracts significant information for retrieval of images. Most spatial information is presented in terms of2D strings[31]. The 2D

string spatial quad-tree used for spatial information representation[32].

E. Local Image Features

Local features are small square, sub-images extracted from the original image[33]. They can be consider to have two different types:

- The patches: They extracted from the images at salient points and dimensionality reduced using Principal Component Analysis (PCA) transformation [34].
- SIFT descriptors [35]: They extracted at Harris interest points [36].

To use local features for image retrieval, three different methods are available[33]:

- Direct transfer: The local features extracted from each database image and from the query image. Then, the nearest neighbors for each of the local features of the query searched and the database images containing most of these neighbors returned.
- Local feature image distortion model (LFIDM): The local features from the query image compared to the local features of each image of the database and the distances between them summed up. The images with the lowest total distances are returned.
- Histograms of local features: A reasonably large amount of local features from the database is clustered and then each database image represented by a histogram of indices of these clusters. These histograms are then compared using the Jeffrey divergence.

III. COMMON TRADITIONAL AND Current Research in Image Retrieval

There have been numerous previous efforts in image retrieval. This section gives a brief overview of some previous work targeting the TBIR, CBIR, SBIR and Automatic Image Annotation (AIA)

A. TBIR Systems

The text-based approach is a traditional simple keyword based search. The images are indexed according to the content, like the caption of the image; filename, title of the web page, and alternate tag...etc. and stored in the database. Processing a user query could involve a stop word removal and stemming. Some of the keyword based image retrieval approaches are bag of words. Image retrieval is then shifted to standard database management capability combined with information retrieval techniques. Some commercial image search engines, such as Google Image Search and Yahoo Image Search, are keyword-based image retrieval systems [37] .

Liet al.[38] proposed a TBIR approach that can effectively exploit loosely labeled Web images to learn robust SVM classifiers. First, they partitioned the relevant and irrelevant Web images into clusters, and then they treated each cluster as a "bag" and the images in each bag as "instances". To predict the labels of instances (images), they proposed a progressive scheme called PMIL-CPB to automatically and progressively select more confident positive bags, which leads to more robust classifiers. They conducted comprehensive experiments using the NUS-WIDE [39] data set and Google data set [40], and the results clearly demonstrate the effectiveness of this approach.

B. CBIR Systems

This section discusses CBIR from three different aspects: common features, local features, and hybrid techniques. Several methods are proposed for images retrieval based on similarity between common features, such as color, shape…etc. This is done by retrieving the stored images from a collection by comparing the stored features in the database with automatically extracted features of the processed image. The use of local features in image retrieval have been proposed to resist many problems, such as the illumination changes, rotation, viewpoint variations…etc. The interest points are defined as the salient image patches that contain rich local information about an image[41].The following subsections, discuss some studies that are concerned these three types.

1) Common Feature

Chen et al. [42] proposed a soft computing approach which is called unified feature matching (UFM). In this retrieval system, an image is represented by a set of segmented regions. Each region is characterized by a fuzzy feature (fuzzy set) reflecting color, texture, and shape properties. The resemblance between two images is defined as the overall similarity between two families of fuzzy features and it quantified by a similarity measure (UFM measure). This UFM measures has two major advantages. The first is that the UFM approach reduces the adverse effect of inaccurate segmentation. It makes the retrieval system more robust to image alterations. The second is that the UFM is better in extracting useful information under the same uncertain conditions.

Chang et al. [43]presented an image retrieval method based on region shape similarity. This method consists of several steps. At first, the images segmented into primitive regions. Then, they are combined to generate meaningful composite shapes, which used as semantic units of the images during the similarity assessment process. They employed three global shape features and a set of normalized Fourier descriptors to characterize each meaningful shape. All these features are invariant under similar transformations. Finally, they measured the similarity between two images by finding the most similar pair of shapes in the two images. There are two potential problems with this method in handling images that are more complex. The first problem is that it is quite hard for machines to determine meaningful regions. The second problem is that many features and similarity models are

proposed, but none of them has been proved to be identical to the human vision model.

Banerjee et al. [44]presented a robust technique for extracting edge map of an image which is followed by computation of global feature using gray level as well as shape information of the edge map. They used blurred image as input and used the concept of Top and Bottom of the intensity surface to extract possible candidates for the edge map. The similarity between the features vectors of two images are computed by Euclidean distance metric.

Deselaers et al. [36] discussed a large variety of features for image retrieval and compare them quantitatively on four different tasks: stock photo retrieval, personal photo collection retrieval, building retrieval, and medical image retrieval. For their experiments, five different, publicly available image databases are used and the retrieval performance of the features analyzed in detail. This allows comparison the findings features from this work with other features that were not covered or that will be present in future. The main question addressed in their work, which features are suitable for which task in image retrieval, thoroughly investigated. Finally, they would like to find better image descriptors and methods to combine these appropriately.

Ahmed et al. [45]presented an approach for graph matching that resembles the human thinking process. The image represented by Fuzzy Attributed Relational Graph (FRAG) that describes each object in the image by all its attribute and spatial relation. They proposed a color feature representation based on Fuzzy concepts. The proposed model is applied to real images evaluated by different users with different perspectives and gives satisfactory results. They found that there is still need to enhance this proposed system by modifying the fuzzy membership functions to improve the image feature representation.

Chakravarti et al.[46]conducted a project that implement and test a simple color histogram based search and retrieve algorithm for images.. The study found the technique to be effective by analysis using the RankPower measurement[47]. The strengths of this algorithm is that it is relatively easy to implement from a coding standpoint. In addition, their system allows images retrieval that have been transformed in terms of size as well as translated through rotations and flips. The main weak point, the implementation of color histograms does not necessarily allow the relevant images as seen by the algorithm to be the same as the relevant images as seen by a human. The results mixed and not consistently accurate.

Yusof [48]presented a dominant color descriptor (DCD) technique for medical image retrieval. The medical image system collected and stored images in medical database. The purpose of DCD technique is to retrieve medical image and display similar images using queried image. A DCD specifies a small number of dominant color values and the statistical. It used Euclidean distance formatching the similarity. Simple application has been developed and tested using DCD. He would like to improve capability of DCD in term of time for retrieving process from large medical image database.

2) Local Features

Jégou et al.[49]showed that exploiting distances between local descriptors significantly improves the accuracy of image search. Firstly, they introduced a distance criterion that provides additional information about correct matches. Secondly, they exploited the distances between SIFT descriptors[35] and the reciprocal neighbors to further refine the similarity measure between descriptors. This method remains too costly to be applied even on a few thousands images. Finally, they asked a question that is how to better approximate the proposed reciprocal nearest neighbor method.

Chen et al.[50]presented local image descriptor using vector quantization SIFT (VQ-SIFT)technique for more effective and efficient image retrieval. Instead of SIFT's weighted orientation histograms, they applied vector quantization (VQ) histogram as an alternate representation for SIFT features. They mentioned that the VQ-based local descriptors are more robust to rotation, projective transforms, and illuminations. Experimental results showed that SIFT features using VQ-based local descriptors can achieve better image retrieval accuracy than the conventional algorithm while the computational costs significantly reduced.

Putraet al.[51]proposed a feature extraction method for image retrieval system based on local self-similarity partitioned iterated function system (PIFS). The feature of image represented by using distance and angle of a pair of range and domain block position from fractal code. The proposed system uses variant contrast and spatial dimensions fractal feature for comparison. The testing of system's performance used 1000 images that have been stored in a database with 10 different categories with different number of image for each category. Images classified as homogeneous and heterogeneous image based on the complexity of the background. System accuracy is determined by using homogeneous image type thatgives better accuracy than heterogeneous image type. They would like to improved performance of this system especially for heterogeneous images.

3) Hybrid Techniques

Luo et al.[52]proposed a hybrid image retrieval system for the Web. The system performs image retrieval in two steps. In the first step, a text-based image meta-search engine retrieves images from the Web using the text information on the host page to provide an initial image set.The second step is used to overcome the disadvantage of low-precision rate of the text-based approach. They used the CBIR approach to re-filter the search results. They used the recall and precision measures for retrieval performance evaluation. There need improved performance using more elaborate visual features to describe the image content.

Dinakaran et al.[12] proposed an interactive image retrieval system, integrating text and image content to enhance the retrieval accuracy. In addition, they presented a refining search algorithm to optimize user searching time on a retrieved image and improve the system quality. Query

refinement comprises query expansion and query re-weighting. Query expansion allows user to expand the query for desired image search. To expand the query, user has to find the other relevant terms that are again time consuming and cause distraction in the search.

C. SBIR Systems

In order to improve the retrieval accuracy of CBIR, the researches focused to reduce the 'semantic gap' between the low and high features. Many recent worksare proposed to cover different aspects in this area, including using ontology to define high-level concepts.

Wu and Hsu[53]proposed a framework that employs ontology and MPEG-7 descriptors to deal with problems arising from representation and semantic retrieval of images. The framework allows for the construction of incrementally multiple ontologies, and shares ontology information rather than building a single ontology for a specific domain not only between the image seekers but also between different domains. The similarity between query ontology and domain ontology for matching relevant images estimated by using Naïve Bayesian inference. This framework consists of three major processes: RDF translation, indexing user query, and matching process. In addition, it provided a relevant feedback mechanism.

Wang et al.[54]proposed a novel image re-ranking framework. The framework has offline and online parts. At the offline stage, the reference classes (which represent different concepts) related to query keywords automatically discovered and training images automatically collected in several steps. At the online stage, the search engine according to the query keyword retrieves a pool of images. Since all the images in the pool are associated with the query keyword according to the word-image index file, they all have pre-computed semantic signatures in the same semantic space specified by the query keyword.They created three data sets to evaluate the performance of this approach in different scenarios. The proposed framework can be improved along several directions: Finding the keyword expansions used to define reference classes can incorporate other metadata and log data besides the textual and visual features.

D. AIA Systems

Automatic image annotation is the process by whicha computer system automatically assigns metadata in the form of captioning or keywords to a digital image[55].Image annotation refers to labelling of images with a set of predefined keywords based on contents of image. This can be helpful for filling the huge gap between low-level features obtained from the image and high-level semantics derived from image. The basic concept of image annotation is to automatically learn about semantic concepts from large number of sample images and use these concepts to label new images [56].

Shaoet al.[57]proposed an image annotation system shown in Fig. 5. First, images are segmented to regions; this stage is only required for visual descriptors that are based on object shapes. Next, MPEG-7 visual descriptors [58] are extracted from the image. Finally, support vector machines [59]are used to classify the visual descriptors into different image categories such as landscape, cityscape.
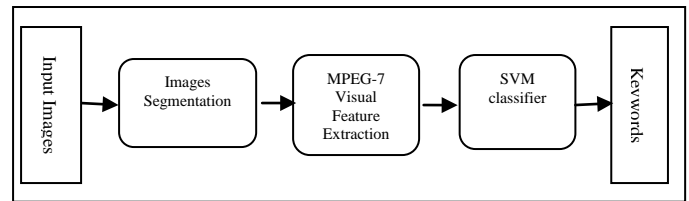

Figure 5. Image Annotation System[57].

AIA is a challenging task due to various imaging conditions, complex and hard to describe objects, a highly textured background and occlusions[60]. However, it enables users to retrieve images by text queries and often provides semantically better results than CBIR. In recent years, it is observed that image annotation has attracted more and more research interests[61]. Many Different strategies including co-occurrence model [62], machine translation model [63], latent space models [64,65],classification approaches [55,66] and relevance language models [67] have been proposed in the literature and each strategy tries to improve previous one [61].

The state-of-the-art techniques of automatic image annotation can be categorized into three categories[68,69]: generative models, discriminative models, and nearest-neighbor-based methods. The generative models treats image and text as equivalent data. It attempts to discover the correlation between visual features and textual words on an unsupervised basis by estimating the joint distribution of features and words[70]. Its contain topic models and mixture models, and can be constructed by estimating the probability distribution over image features and high-level semantic concepts [71]. Probabilistic Latent Semantic Analysis (PLSA)[72] , Hierarchical Dirichlet process model [73], andMarkovian Semantic Indexing (MSI)[74]are typical generative models. The discriminative models treat each annotated keyword as an independent class, and creates different classifiers for different keyword a separate classifier[71] (e.g., support vector machine classifier (SVMs) [59],Gaussian Mixture Model [75]). This model computes similarity at the visual level and annotates a new image by propagating the corresponding words [70]. The third type isnearest-neighbor-based methods. These approaches assume that semantic-relevant images share similar visual features and treat annotation as a retrieval problem. For each keyword, a seed image will receive relevance votes from its visual neighbors that are labeled with this keyword by users and the votes can be weighted according to their visual similarities. In this situation, it is extremely important to choose a criteria to define which images are neighbors. Recently, the nearest-neighbor-based methods have captured more attentions from researchers due to its good annotation performance and simplicity in principle [71].

Yang et al.[76] presented a hybrid generative-discriminative classifier to simultaneously address the extreme data-ambiguity. They abstracted AIA as a generic learning task called many-class multi-label multi-instance classification (M3C), which aims at learning decision rules from data that are ambiguous in both input and output and involve massive classes. They suggested AIA system that builds the multi-instance multi-label corpora by segmenting each image into instances (i.e. regions) and using words in the caption as labels. Each region is represented using the standard bag-of-discrete-feature framework [77].They test the method on two data sets that are from popular online annotation engines on two real-world benchmarks. One limitation of this model is that, by assuming instance exchangeability it does not account for context correlations. They plan to empirically compare their model with state-of-art annotation algorithms (e.g., MIR-FLICKR set [78]) to provide a big picture of existing annotation methods.

Xu and Mu [71] used a modified keywords transfer mechanism base on image-keyword unidirectional graph to derive a great annotator. They assigned a few given keywords to test image if most of its neighbor images share the given keywords. They first build an image-keyword unidirectional graph to capture the relationships between images and keywords (the correlations between an image and different keywords are not necessarily the same), then they represented the images with six low level features which are used to find the nearest-neighbors of test image. They used a modified keywords transfer mechanism that employs visual similarity and image-keyword unidirectional graph to annotate new image.They found the used method achieved better annotation performance than two of the most well-known methods which are the Annotation by Image-Concept Distribution Model (AICDM) [79] and Label Transfer mechanism(LT) [80]on the PascalVOC07 database[81].

IV. IR CHALLENGES AND RESEARCH DIRECTIONS

Bridging the semantic gap for image retrieval still considered a big challenge. Even though there are a lot of efforts and works on image retrieval research, but it is not enough to provide satisfactory performance. However, there are still some spaces, which need to be improved besides the challenges that is associated with mapping low level to high-level concepts. Also overcome of the semantic gap in the broad domain database is complex because the images in broad domains can be described using various concepts.There are needto see better supportfor the image retrievalbased semantic concept with a focus on the retrieval by abstract attributes, involving a significant amount of high-levelreasoning about the meaning and purpose of the objects. In addition, the extracted semantic features should be applied for any kind of image collection. Moreover, there is need to effective ways retrieve of similar images that are conform to human perception and without human interference.

In image annotation some issues poses a challenge: images have to be automatically annotated with meaningful labels , Developing of an improved image segmentation algorithms toidentify objects that facilitate annotation process andannotation to different Languages needs to handle with Large volumes of databases and also many things are hard to be expressed.

On the other hand, the object recognition and detection is one of the most challenging problems in image retrieval. The addressing of this problem is urgent.In general, the object should be recognized regardless of the illumination changes, changes of size (scale), rotation, background clutter, viewpoint change, and occlusion. However, this cannot be obtained in most CBIR because it does nothave adequate ability to capture important properties of object.In support of better object recognition, the interest point's detectors were introduced to represent the local features of images in image retrieval systems. Many algorithms have been developed for the purpose of detecting and extracting the interest points like Scale-Invariant Feature Transform (SIFT)[35],Speeded up robust features (SURF)[82],and Oriented FAST and Rotated BRIEF (ORB) [83].

However, there are many issues of intense research that still unsolved or solved with limited success for example: The optimization of feature processing time and query response time are important for a huge image database. The selection of algorithms, parameters etc. is very important because specific one is not necessary suitable for all application e.g. the segmentation algorithm used for natural images may not be suitable for medical images. An appropriate index structure that allows efficient searching of large image database is still a problem under research [84].

V. CONCLUSION

This paper attempted to provide an overview of the most common techniques of different types of image retrieval systems. Most systems used low-level features, few systems used semantic feature. Global featuresfail to identify important visual characteristics of images but it's very efficient in computation and storage due to its compact representation. From another perspective, local features that can be extracted from images handle partial image matching or searching for images that contain the same object or same scene with different viewpoints, different scale, changes in illumination, occlusions,,…etc. Therefore,local features can identify important visual characteristics of images but it ismore expensive computationally. The semantic features that based on keywords or annotationsmaybe very subjective and time consuming. Whereas,the semantic features that based on visual contentis complex becauseof the inference procedures. Automatic image annotation is good approach toreduce the semantic gap, butit still achallenging task due tothe different conditions of imaging, occlusionsand the complexity, and difficulty to describe objects. In future, there is a need to work more and more with available techniques to deal with the semantic gap to enhance image retrieval.As a survey, it is very difficult to include each and every aspect of all works.

However, this paper focused to give overviewof the most common traditional andmodern typesof IR approaches.

REFERENCES

[1] H.W. Hui, D. Mohamad, and N.A. Ismail, "Approaches, challenges and future direction of image retrieval," Journal of Computing, vol. 2, 2010, pp. 193–199.

[2] 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, 1999, pp. 39–62.

[3] S.-K. Chang and A. Hsu, "Image information systems: where do we go from here?," IEEE transactions on Knowledge and Data Engineering, vol. 4, 1992, pp. 431–442.

[4] H. Tamura and N. Yokoya, "Image database systems: A survey," Pattern recognition, vol. 17, 1984, pp. 29–43.

[5] G. Mailaivasan Parthiban Karthikram, "Tag based image retrieval (TBIR) using automatic image annotation," International Journal of Research in Engineering and Technology, vol. 03, 2014.

[6] T. Kato, "Database architecture for content-based image retrieval," SPIE/IS&T 1992 Symposium on Electronic Imaging: Science and Technology, International Society for Optics and Photonics, 1992, pp. 112–123.

[7] M. Singha and K. Hemachandran, "Content based image retrieval using color and texture," Signal & Image Processing: An International Journal (SIPIJ), vol. 3, 2012, pp. 39–57.

[8] S. Mitra and T. Acharya, Data mining: multimedia, soft computing, and bioinformatics, John Wiley & Sons, 2005.

[9] N. Chauhan and M. Goyani, "Enhanced Multistage Content Based Image Retrieval," IJCSMC, vol. 2, Issue. 5, 2013, pp. 175–179.

[10] N. Chauhan and M. Goyani, "Enhanced Multistage Content Based Image Retrieval," IJCSMC, vol. 2, Issue. 5, 2013, pp. 175–179.

[11] H. Zhang H. Dagan F. Long and D. Fen, Fundamentals of content based image retrieval," in D. Feng, W. Siu, H. Zhang (Eds.): "Multimedia Information Retrieval and Management. Technological Fundamentals and Applications,, Springer-Verlag, .

[12] B. Dinakaran, J. Annapurna, and C.A. Kumar, "Interactive image retrieval using text and image content," Cybern Inf Tech, vol. 10, 2010, pp. 20–30.

[13] H.H. Wang, D. Mohamad, and N. Ismail, "Image Retrieval: Techniques, Challenge, and Trend," International conference on Machine Vision, Image processing and Pattern Analysis, Bangkok, Citeseer, 2009.

[14] N. Shanmugapriya and R. Nallusamy, "Anew content based image retrieval system using GMM and relevance feedback," Journal of Computer Science 10 (2): 330-340, 2013.

[15] R. Datta, D. Joshi, J. Li, and J.Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," ACM Computing Surveys (CSUR), vol. 40, 2008, p. 5.

[16] A. Halawani, A. Teynor, L. Setia, G. Brunner, and H. Burkhardt, "Fundamentals and Applications of Image Retrieval: An Overview.," Datenbank-Spektrum, vol. 18, 2006, pp. 14–23.

[17] M. Aly, P. Welinder, M. Munich, and P. Perona, "Automatic discovery of image families: Global vs. local features," Image Processing (ICIP), 2009 16th IEEE International Conference on, IEEE, 2009, pp. 777–780.

[18] A. Arampatzis, K. Zagoris, and S.A. Chatzichristofis, "Dynamic two-stage image retrieval from large multimedia databases," Information Processing & Management, vol. 49, 2013, pp. 274–285.

[19] A. Popescu, P.-A. Moëllic, I. Kanellos, and R. Landais, "Lightweight web image reranking," Proceedings of the 17th ACM international conference on Multimedia, ACM, 2009, pp. 657–660.

[20] Y.-K. Chan and C.-Y. Chen, "Image retrieval system based on color-complexity and color-spatial features," Journal of Systems and Software, vol. 71, 2004, pp. 65–70.

[21] J.M. Fuertes, M. Lucena, N. Perez de la Blanca, and J. Chamorro-Mart□nez, "A scheme of colour image retrieval from databases," Pattern Recognition Letters, vol. 22, 2001, pp. 323–337.

[22] M.J. Swain and D.H. Ballard, "Color indexing," International journal of computer vision, vol. 7, 1991, pp. 11–32.

[23] T.S. Chua, K.-L. Tan, and B.C. Ooi, "Fast signature-based color-spatial image retrieval," Multimedia Computing and Systems' 97. Proceedings., IEEE International Conference on, IEEE, 1997, pp. 362–369.

[24] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz, "Efficient and effective querying by image content," Journal of intelligent information systems, vol. 3, 1994, pp. 231–262.

[25] J.R. Smith and S.-F. Chang, "Transform features for texture classification and discrimination in large image databases," Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference, IEEE, 1994, pp. 407–411.

[26] P. Wu, B. Manjunath, S. Newsam, and H. Shin, "A texture descriptor for browsing and similarity retrieval," Signal Processing: Image Communication, vol. 16, 2000, pp. 33–43.

[27] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," Systems, Man and Cybernetics, IEEE Transactions on, vol. 8, 1978, pp. 460–473.

[28] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," Signal Processing: Image Communication, vol. 17, 2002, pp. 825–848.

[29] D. Zhang and G. Lu, "Review of shape representation and description techniques," Pattern recognition, vol. 37, 2004, pp. 1–19.

[30] Y. Rui, A.C. She, and T.S. Huang, "Modified Fourier descriptors for shape representation-a practical approach," Proc of First International Workshop on Image Databases and Multi Media Search, 1996, pp. 22–23.

[31] S.-K. Chang, Q.-Y. Shi, and C.-W. Yan, "Iconic indexing by 2-D strings," Pattern Analysis and Machine Intelligence, IEEE Transactions on, 1987, pp. 413–428.

[32] H. Samet, "The quadtree and related hierarchical data structures," ACM Computing Surveys (CSUR), vol. 16, 1984, pp. 187–260.

[33] T. Deselaers, D. Keysers, and H. Ney, "Features for image retrieval: A quantitative comparison," Pattern Recognition, Springer, 2004, pp. 228–236.

[34] T. Deselaers, D. Keysers, and H. Ney, "Discriminative training for object recognition using image patches," Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, IEEE, 2005, pp. 157–162.

[35] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," International journal of computer vision, vol. 60, 2004, pp. 91–110.

[36] T. Deselaers, D. Keysers, and H. Ney, "Features for image retrieval: an experimental comparison," Information Retrieval, vol. 11, 2008, pp. 77–107.

[37] J. van Gemert, "Retrieving images as text," 2003.

[38] W. Li, L. Duan, D. Xu, and I.W.-H. Tsang, "Text-based image retrieval using progressive multi-instance learning," Computer Vision (ICCV), 2011 IEEE International Conference on, IEEE, 2011, pp. 2049–2055.

[39] T.-S. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y. Zheng, "NUS-WIDE: a real-world web image database from National University of Singapore," Proceedings of the ACM international conference on image and video retrieval, ACM, 2009, p. 48.

[40] R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman, "Learning object categories from Google's image search," Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, IEEE, 2005, pp. 1816–1823.

[41] K.VELMURUGAN, "A Survey of Content-Based Image Retrieval Systems using Scale-Invariant Feature Transform (SIFT)," International Journal of Advanced Research in Computer Science and Software Engineeringe, vol. 4, Issue 1, 2014.

[42] Y. Chen and J.Z. Wang, "Looking beyond region boundaries: Region-based image retrieval using fuzzy feature matching," Matching", Proc. Multimedia Content-Based Indexing and Retrieval Workshop, INRIA, Citeseer, 2001.

[43] C. Chang, W. Liu, and H. Zhang, "Image retrieval based on region shape similarity," Photonics West 2001-Electronic Imaging, International Society for Optics and Photonics, 2001, pp. 31–38.

[44] M. Banerjee and M.K. Kundu, "Edge based features for content based image retrieval," Pattern Recognition, vol. 36, 2003, pp. 2649–2661.

[45] H.A. Ahmed, N. El Gayar, and H. Onsi, "A New Approach in Content-Based Image Retrieval Using Fuzzy Logic," INFOS, 2008, p. 8.

[46] R. Chakravarti and X. Meng, "A Study of Color Histogram Based Image Retrieval.," ITNG, 2009, pp. 1323–1328.

[47] X. Meng, "A comparative study of performance measures for information retrieval systems," Information Technology: New Generations, 2006. ITNG 2006. Third International Conference on, IEEE, 2006, pp. 578–579.

[48] M.K. Yusof, "Effectiveness of Dominant Color Descriptor Technique in Medical Image Retrieval Application," World Academy of Science, Engineering and Technology, International Science Index, 2010.

[49] H. Jégou, M. Douze, C. Schmid, and others, "Exploiting descriptor distances for precise image search," INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE (inria), vol. 2, 2011.

[50] et al Qiu Chen, "Local Image Descriptor using VQ-SIFT for Image Retrieval," World Academy of Science, Engineering and Technology Vol:59.

[51] K.G.D. Putra, A. Purnawan, and M. Artana, "A New Technique Based on PIFS Code for Image Retrieval System," Applied Mathematical Sciences, vol. 7, 2013, pp. 5957–5967.

[52] B. Luo, X. Wang, and X. Tang, "World Wide Web based image search engine using text and image content features," Electronic Imaging 2003, International Society for Optics and Photonics, 2003, pp. 123–130.

[53] W.H.H. Roung Shiunn Wu, "A Semantic Image Retrieval Framework Based on Ontology and Naive Bayesian Inference," International Journal of Multimedia Technology, Vol. 2, Issue 2, Pages 36-43.

[54] X. Wang, S. Qiu, K. Liu, and X. Tang, "Web image re-ranking using query-specific semantic signatures," IEEE, 2013.

[55] C. Chang E et al., "Content-Based Soft Annotation for Multimodal Image Retrieval Using Bayes Point Machines," CirSysVideo, vol. 13(1), 2003, pp. 26–38.

[56] D. Zhang, M.M. Islam, and G. Lu, "A review on automatic image annotation techniques," Pattern Recognition, vol. 45, 2012, pp. 346–362.

[57] W. Shao, G. Naghdy, and S.L. Phung, "Automatic image annotation for semantic image retrieval," Advances in Visual Information Systems, Springer, 2007, pp. 369–378.

[58] B.S. Manjunath, P. Salembier, and T. Sikora, Introduction to MPEG-7: multimedia content description interface, John Wiley & Sons, 2002.

[59] X. Qi and Y. Han, "Incorporating multiple SVMs for automatic image annotation," Pattern Recognition, vol. 40, 2007, pp. 728–741.

[60] Y. Lei, W. Wong, W. Liu, and M. Bennamoun, "An HMM-SVM-based automatic image annotation approach," Computer Vision–ACCV 2010, Springer, 2011, pp. 115–126.

[61] H. Chaudhari and D. Patil, "A Survey on Automatic Annotation and Annotation Based Image Retrieval," International Journal of Computer Science and Information Technologies, vol. . 5 (2, 2014, pp. 1368–1371.

[62] H.T. Y Mori and R. Oka, "Image-to-word transformation based on dividing and vector quantizing images with words," In MISRM99 First Intl. Workshop on Multimedia Intelligent Storage and Retrieval Management, 1999.

[63] D.. P. Barnard K. de Freitas J.F.G. Forsyth Duygulu, "Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary," In: ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part IV, London, UK, Springer-Verlag, pp. 97–112.

[64] D. F. Gatica-Perez Monay, "On image auto-annotation with latent space models," MULTIMEDIA '03: Proceedings of the eleventh ACM international conference on Multimedia,ACM Press, vol. 275–278, 2003.

[65] N.E. Maillot J.-H. Lim T.-T. Pham and J.-P. Chevallet, "Latent Semantic Fusion Model for Image Retrieval and Annotation," Proc. 16th ACM Conf. Information and Knowledge Management(CIKM), 2007.

[66] G.C. Cusano C and R. Schettini, "Image Annotation Using SVM," Proceedings of the Internet Imaging IV, vol. vol. 5304.

[67] Jiwoon Victor Lavrenko Jeon and R. Manmatha, "Automatic image annotation and retrieval using cross-media relevance models," ACM SIGIR Conf. on Research and Development in Information Retrieval, 2003, pp. 119–126.

[68] J.-H. Su, C.-L. Chou, C.-Y. Lin, and V.S. Tseng, "Effective semantic annotation by image-to-concept distribution model," Multimedia, IEEE Transactions on, vol. 13, 2011, pp. 530–538.

[69] S. Zhang, J. Huang, H. Li, and D.N. Metaxas, "Automatic image annotation and retrieval using group sparsity," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, vol. 42, 2012, pp. 838–849.

[70] Z. Li, Z. Tang, W. Zhao, and Z. Li, "Combining Generative/Discriminative Learning for Automatic Image Annotation and Retrieval," International Journal of Intelligence Science, vol. 2, 2012, p. 55.

[71] G.-Q. Xu and Z.-C. Mu, "Automatic Image Annotation Using Modified Keywords Transfer Mechanism Base on Image-Keyword Graph," International Journal of Computer Science Issues (IJCSI), vol. 10, 2013.

[72] F. Monay and D. Gatica-Perez, "Plsa-based image auto-annotation: constraining the latent space," Proceedings of the 12th annual ACM international conference on Multimedia, ACM, 2004, pp. 348–351.

[73] O. Yakhnenko and V. Honavar, "Annotating images and image objects using a hierarchical dirichlet process model," Proceedings of the 9th International Workshop on Multimedia Data Mining: held in conjunction with the ACM SIGKDD 2008, ACM, 2008, pp. 1–7.

[74] Klimis S. Ntalianis Dionyssios D. Sourlas Konstantinos A. Raftopoulos and S.D. Kollias, "Mining User Queries with Markov Chains," Application to Online Image Retrieval IEEE Transactions On Knowledge And Data Engineering, vol. 25, No. 2, 2013.

[75] N.V. A.B. Chan P.J. Moreno G. Carneiro, "Supervised Learning of Semantic Classes for Image Annotation and Retrieval," Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. Vol.29,No.3, 2007, pp. 394–410.

[76] S.H. Yang, J. Bian, and H. Zha, "Hybrid generative/discriminative learning for automatic image annotation," arXiv preprint arXiv:1203.3530, 2012.

[77] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," Computer Vision–ECCV 2006, Springer, 2006, pp. 490–503.

[78] J. Verbeek, M. Guillaumin, T. Mensink, and C. Schmid, "Image annotation with tagprop on the MIRFLICKR set," Proceedings of the international conference on Multimedia information retrieval, ACM, 2010, pp. 537–546.

[79] J.-H. Su, C.-L. Chou, C.-Y. Lin, and V.S. Tseng, "Effective image semantic annotation by discovering visual-concept associations from image-concept distribution model," Multimedia and Expo (ICME), 2010 IEEE International Conference on, IEEE, 2010, pp. 42–47.

[80] A. Makadia, V. Pavlovic, and S. Kumar, "Baselines for image annotation," International Journal of Computer Vision, vol. 90, 2010, pp. 88–105.

[81] M. Everingham, L. Van Gool, C.K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," International journal of computer vision, vol. 88, 2010, pp. 303–338.

[82] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," Computer Vision–ECCV 2006, Springer, 2006, pp. 404–417.

[83] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," Computer Vision (ICCV), 2011 IEEE International Conference on, IEEE, 2011, pp. 2564–2571.

[84] G. ThakoreDarshak "Evaluation enhancement development and implementation of content based image retrieval algorithms".PhDThesis.Maharaja Sayajirao University .2013.