Image Retrieval Systems

Ambrose G.Tuscano *

Department of Computer Science, University of Maryland, Baltimore County

May 2020

Abstract

With the recent information overload, focus is on retrieving relevant data from a collection of those resources. Information Retrieval systems are designed to compare the similarity between the queried data and the data in the database, wherein it selects one with the most similarity to the Queried Data. In this paper, we will focus on Image Retrieval techniques. This paper starts with discussing the working of text based image retrieval then the content-based retrieval. We highlight the various techniques of content based image retrieval such as retrieval by color, shape and the texture and the various algorithms that are used for content based image retrieval.

Keywords— Text Based Image Retrieval, Content Based Image Retrieval

*atuscan1@umbc.edu

Contents

1 Introduction **3**

2 Image Retrieval Methods **4**

3 Text Based Image Retrieval **4**

4 Content Based Image Retrieval **4**

5 Discussion on the strategies used **5**
5.1 Text Based Image Retrieval 5
5.1.1 Text-based Indexing System 5
5.1.2 Image Retrieval using Automatic Image Annotation: 6
5.1.3 Progressive Multi-Instance Learning 6
5.2 Content Based Image Recognition 6
5.2.1 Color Based Features 6
5.2.2 Texture Based Features 6
5.2.3 Shape Based Features 7
5.2.4 Techniques used in CBIR 7

6 Challenges and Discussion **7**

7 Conclusion **8**

1 Introduction

With the Digitalization of Data, huge image collections are rapidly being created and made available through the World Wide Web. The amount of information that can be retrieved from such images is massive and at the same time complicated to work on. To query this database of images, and to find the most similar images from these databases to desired images is a topic of research in Algorithms and Computer Vision. On going research aims at systems to auto index such images for improving the retrieval times and improving technologies in similarity matching. The query can be in any format be it text, similar image, sketch with the aim being to use the idea in the query to find images which match in the database. The System have been developed in many ways to cover various topics of interest, but the basic framework is the same.

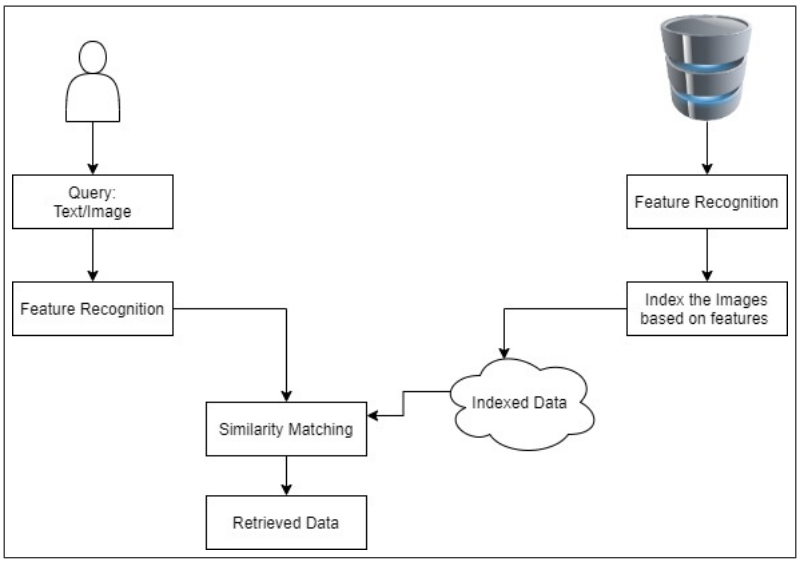

Figure 1: Basic Structure of a Image Retrieval System

Image Retrieval Systems are mainly classified as text-based and content-based image retrieval systems. A Legacy approach, Text-based system searches the images based on the Labels, annotations, keywords,websites and more textual data associated with the images. TBIR Systems need manual keywording of the image, so that the images can be indexed. This is time consuming and has human errors, thus making just this approach Unreliable.

On the other hand the Content-based system searches the images based on the visual features of the image.CBIR Systems rely on the visual contents of the query image to retrieve images from an image collection. They focus on retrieving images based on local features of images which include color, texture, shape , and global features like histograms of oriented gradients (HOG), Shape Matrices, Invariant Moments,,etc..

In the past, Global features were used for the tasks related to Retrieval, but such systems don't quite well capture the complexities of the image. And so, currently Local features are being favoured for the image recognition part. It is also seen that using a hybrid approach improves the feature detection significantly. For example, using color and shape to identify cars is more reliable than using color or shape alone. The context information is also helpful for recognizing objects. With that know, many image retrieval techniques are developed which include but are not limited to techniques which require indexing done before based on some approach, techniques where each database's image's local features are compared with query image features, and also techniques where the objects in the images are segmented before to map the semantic gap using Computer Vision can be observed.

This Paper will cover the existing systems and techniques followed by them. And a short analysis of these techniques. Finally, ending the paper with a conclusion.

2 Image Retrieval Methods

Problem with Image Retrieval is that such systems are complicated to implement. On what basis would one classify the images, how would one index the image so as to find it while checking for similarity, what algorithms should be implemented to find what the content of the image is, how to improvise over the semantic gap in the high level features and low level features, are few of the things that are addressed in any Image retrieval system.

Though currently implemented in many ways, the Image Retrieval systems are developed based mainly on the ideas used in Text Based Image Retrieval System and Content Based Image Retrieval Systems. The other methods include:

- Sketch Based Image Retrieval
- Semantic Based Image Retrieval
- Annotation based image retrieval

We will discuss about Text Based Image Retrieval System and Content Based Image Retrieval Systems and analyse few of the techniques used to implement them.

3 Text Based Image Retrieval

The text based image retrieval utilizes the method of capturing the metadata of the image location, such as labels, keywords, captioning or descriptions to the images and works considering them as feature vectors to calculate the similarity score and thus finding relevant images. Three Ways to make a database for querying images are:

- We can Manually Assign Keywords to each image
- We can use use text already associated with the images
- Though complicated another way is to analyse the image content to automatically assign keywords using Computer Vision techniques to facilitate querying later.

The input query here is also in the form of a textual data too to facilitate the similarity checking. Thus we generate a description vector of the images in the database using the annotation. Next the vector, generated by the user query is then matched with the vectors stored in the database.
This method is time consuming and may lead to a lot of human errors, which may affect the recall and accuracy of the system. Another drawback is about the interpretation for images by different people leading to a loss in the value of objects in the image, and thus wrong annotation. Moreover, the system may fail to recognize input like similar words, homonyms, spelling differences and at times abbreviations of the queried data too.

Here traditional text information retrieval methods like bag of words, tfidf to calculate the relevance and similarity matching using methods like cosine similarity, Smooth Inverse Frequency, Jaccard matching, Latent Dirichlet allocation are being used. [1]

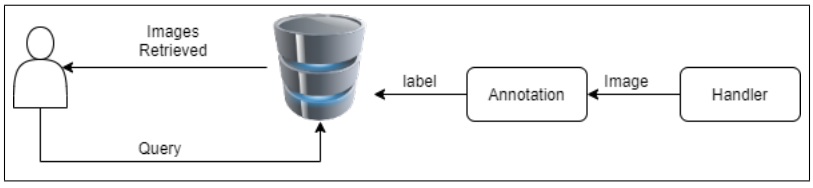

Figure 2: Basic Structure of a Text Based Image Retrieval System

4 Content Based Image Retrieval

In seeking to overcome limitations of the Text Based Image Retrieval, a content-based retrieval (CBR) system was proposed in the early 1990's [2] They are the images retrieval systems which retrieve images

based on image visual contents, normally called as features. This features basically are related to color, shape and texture, while other high level features like image objects(using detection algorithm) can also be considered as features. The visual features cannot completely characterize semantic content for image descriptions, but they are easier to integrate into mathematical formulations.

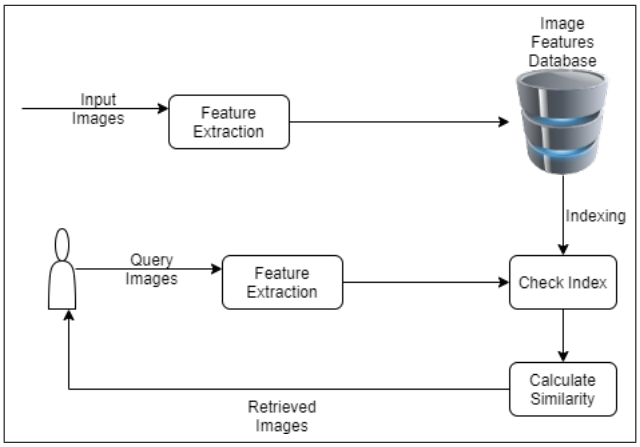

Figure 3: Basic Structure of a Content Based Image Retrieval System

The Content Based Image Retrieval System workflow is first get the image features, index those features and perform similarity computations of the stored images with the users query using certain algorithms to provide answers to the user's query. Lastly, the user feedback is used for checking the relevancy of retrieved images, the type of the user-interaction,etc.

5 Discussion on the strategies used

Though the models are same the stratergies to implement them change. In this section we discuss about a few of the approaches used to tackle such models.

5.1 Text Based Image Retrieval

The features for a system like this are the text data from the annoted image database, their weight components, etc.

5.1.1 Text-based Indexing System

This is similar indexing to traditional method used for Text based systems. It is done on a phased process. This strategy is implemented at [3]

- **Sentences Identification:**
  Involves process similar to tokenization and counts occurrences of the tokenized sentences.
- **Semantic Analysis:**
  This phase deals with deciphering the meaning of the words and later the sentences from its order of appearance.
- **Stop-word Removal and Synonyms Identification:**
  Words with more frequency which dont have any special meaning are removed.
- **Stemming:**
  Now, stemming algorithm is applied to identify and map all the words that have the same root, to the stem. This reduces the issues where similar words weren't recalled.
  Indexing:
  Done to increasing the processing speed and efficient retrieval of information.

- **Classification:**
  To perform the classification of sentences and to understand the similarity report classifiers like Naive Bayes are used.

5.1.2 Image Retrieval using Automatic Image Annotation:

This technique uses an approach where the features of a set of annotated images are extracted using the Speed-ed Up Robust Features (SURF) method and then annotated images are trained using the Support Vector Machine (SVM) classifier and based on the feedback they get annotated again. [4]. So, we do not need to extract the whole features from the images when a new image comes, rather we extract features from the objects of interest and matches those features against the different groups of images for the feature matching and effective retrieval based on object selection.

5.1.3 Progressive Multi-Instance Learning

Li et al. [5] 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 data set and Google data set, and the results were pretty accurate.

5.2 Content Based Image Recognition

CBIR is a move up on the text based techniques, here the image database is queried using a image and the images are retrieved based on their intrinsic features rather than the texts assigned to them. This helps in improving the accuracy and also minimizing human interference.

CBIR uses the features of the image like the color, shape, texture, etc. [Sharmin et al. 2002] for implementation. Content based system index the images automatically by using different techniques for their visual contents.

5.2.1 Color Based Features

Several methods for retrieving images on the basis of colour similarity have been described in the research, but most are variations on the same basic idea. Each image added to the collection is analysed to compute a colour histogram which shows the proportion of pixels of each colour within the image. The colour histogram for each image is then stored in the database. At retrieval time, the users Queried images histogram is checked for similarity with those images whose colour histograms match those of the query most closely. Such features can be extracted using techniques like Color Histogram, Color Correlogram, Color Auto-Correlogram, Color Coherence vector, and Dominant Color Descriptors. Xue et al [6] states separate color images and color histogram moment of extraction, and then two methods of extracting color feature vector weighted to achieve similar distance, similar to the last distance based on the size of the return search results, based on the realization of the characteristics of the color image Retrieval system and through rotation transitions.

Chang at [7] proposed color distributions, the mean value and the standard deviation, to represent the global characteristics of the image. Moreover, the image bitmap was used to represent the local characteristics of the image for increasing the accuracy of the retrieval system.

5.2.2 Texture Based Features

The ability to match on texture similarity can often be useful in distinguishing between areas of images with similar colour (such as sky and sea, or leaves and grass). Texture refers to visual patterns of homogeneity and do not result significantly in a single color.Textures at a certain scale are not textures at a coarser scale. Differently from color, texture is a property' associated with some pixel neighbourhood, not with a single pixel. Texture can be mapped by using some mathematical transformation over them and this can be useful for similarity checking. It is possible to calculate measures of image texture such as the degree of contrast, coarseness, directionality and regularity [Tamura et al, 1978], or periodicity, directionality and randomness

[Liu and Picard,1996]. Zhang et.al. at [8] proposed a system where Texture features are found by calculating the mean and variation of the Gabor filtered image. Rotation normalization is realized by a circular shift of the feature elements so that all images have the same dominant direction. The image indexing and retrieval are conducted on textured images and natural images. Sandhu et al [9] highlighter GLCM and histogram over shape and the texture properties as methods for texture based retrieval.

5.2.3 Shape Based Features

A shape is the form of an object or its external boundary, outline,or external surface, as opposed to other properties like color, texture or composition. Whereby image edges can be detected by using algorithms like Fourier Descriptors, Canny Algorithm , SIFT Descriptors, Moment In-variants and Eccentric and Axis Oriented methods. Also Grab cut and similar image segmentation techniques are useful. Mehtre at [10]discussed effectiveness of several shape measures for content based similarity retrieval of images. Jagadish [11] proposed to construct an index structure on the data such that given a template shape, matching shapes can be retrieved in time that is less than linear in the size of the database, that would be by means of an indexed lookup.

5.2.4 Techniques used in CBIR

We discuss a few of the techniques used to retrieve a image in a CBIR.

- **Relevance Feedback**
  This systems are where the user refines the search results by marking images in the results as "relevant", "not relevant", or "neutral" to the search query, and at times adding additional information then repeating the search with the new information.
- **Gabor Filter**
  It is widely used for texture analysis. A two dimensional Gabor function g(x, y) consists of a sinusoidal plane wave of some frequency and orientation(Carrier), and two dimensional translated. Gaussian Envelope is used to modulate it. Gabor texture features are used as image content descriptors and efficiently are used to retrieve images [12]
- **SVM**
  A classifier can be learned from training data of relevance images and irrelevance images marked by users. Then the model can be used to find more relevance images in the whole database. Compared with other learning algorithms, the SVM approach is considered a good candidate because of its high generalization performance without the need to add apriori knowledge, even when the dimension of the input space is very high. Lei at [13] notes about how the SVM can be trained so the system can retrieve more images relevant to the query in the database efficiently.
  Douglas et.al at [14] proposes an iterative method for CBIR. The algorithm receives a database previously divided into N classes and apply DCT for feature extraction. Then it constructs N SVM machines and performs a selection of the candidate classes for query classification.
- **Neural Networks**
  A CNN doesn't need complex work like pre done feature extraction to work. Having a proper labelled data, we can train the system to learn the data features using complex layer structure. Though it is much of a research topic, such a method can be implemented.

  Research at [15] presents a model which is a combination of CNN topped with SVM to get relevant image.

6 Challenges and Discussion

Image Retrieval in itself is a tough task to work on. Further, the semantic gap arises due to difference between understanding low level and high level features of images and their interpretation. TBIR can help fill the semantic gap using better annotation, but that could be time consuming and also would see human errors. At the same time, CBIR systems with the current research don't seem to handle such features easily. Hence a hybrid Model is being researched on. Such a model works based a combination of both TBIR and CBIR techniques. Whereby, one queries a image data set by using text+Image query and all relevant images

to the text are retrieved, they are then compared with the Image and using CBIR techniques the best of them are chosen. This can be better implemented by using relevance feedback from user and improving search in real time using the refined query. This can handle recall and precision better.

7 Conclusion

In this paper, various techniques of image retrieval based on text and content of the image were analyzed. Each technique tries to solve the existing challenges faced by image retrieval system and is favourable in their individual uses. Various factors are responsible in affecting the performance of the system like image preprocessing, indexing algorithms, Annotation, etc. The factors that positively affect the system are researched more about and are combined to improve the system's performance.

Appendix A: Research conference and papers

References

[1] Syed Ali Jafar Zaidi, Attaullah Buriro,et.al., Implementation and Comparison of Text-Based Image Retrieval Schemes, (2019),

[2] Peter Enser, et.al, Image and Video Retrieval(2004)

[3] Avinash N Bhute and B.B. Meshram, Text Based Approach For Indexing And Retrieval Of Image And Video: A Review, (2014)

[4] Anurag Bhargava, et.al ., An Object Based Image Retrieval Framework Based on Automatic Image Annotation

[5] Wen Li et.al., Text-Based Image Retrieval using Progressive Multi-Instance Learning(2011)

[6] Bai Xue et.al., Research of Image Retrieval Based on Color, (2009)

[7] Tzu Chuen Lu, Color image retrieval technique based on color features and image bitmap, (2006)

[8] Dengsheng Zhang, et.al., Content-based Image Retrieval Using Gabor Texture Features

[9] Amanbir Sandhu and aarthi kocchar, Content Based Image Retrieval using Texture, Color and Shape for Image Analysis(2012)

[10] Babu M.Mehtre, et.al., Shape measures for content based image retrieval: A comparison(1997)

[11] H. V. Jagadish, A retrieval technique for similar shapes,(1991)

[12] Tudor Barbu, Content-Based Image Retrieval Using Gabor Filtering(2009)

[13] Lee Zhang, Support vector machine learning for image retrieval(2001)

[14] Douglas Natan Meireles Cardoso, Iterative Technique for Content-Based Image Retrieval using Multiple SVM Ensembles

[15] Mohamed Ouhda, Content-Based Image Retrieval Using Convolutional Neural Networks (2018)