International Journal of Trend in Scientific Research and Development (IJTSRD)
Volume 4 Issue 2, February 2020 Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

An Impact on Content Based Image Retrival: A Perspective View

Shivanshu Jaiswal[1], Dr. Avinash Sharma[2]

[1]PG Scholar, [2]Head of Department,
[1,2]Department of CSE, MIT, Bhopal, Madhya Pradesh, India

ABSTRACT

The explosive increase and ubiquitous accessibility of visual data on the Web have led to the prosperity of research activity in image search or retrieval. With the ignorance of visual content as a ranking clue, methods with text search techniques for visual retrieval may suffer inconsistency between the text words and visual content. Content-based image retrieval (CBIR), which makes use of the representation of visual content to identify relevant images, has attracted sustained attention in recent two decades. Such a problem is challenging due to the intention gap and the semantic gap problems. Numerous techniques have been developed for content-based image retrieval in the last decade. We conclude with several promising directions for future research.

Key words: *content-based image retrieval, visual representation, indexing, similarity measurement, spatial context, search re-ranking.*

How to cite this paper: Shivanshu Jaiswal | Dr. Avinash Sharma "An Impact on Content Based Image Retrival: A Perspective View" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-4 | Issue-2, February 2020, pp.210-212, URL: www.ijtsrd.com/papers/ijtsrd29969.pdf

IJTSRD29969

I. INTRODUCTION

Content based image retrieval (CBIR) has been an active research area since 1970. It applications has increased many fold with availability of low price disk storages and high speeds processors. Image databases containing millions of images are now cost effective to create and maintain. Image databases have significant uses in many fields including medicines, biometric security and satellite image processing. Accurate image retrieval is a key requirement for these domains. Researchers have developed several techniques for processing of images databases [1]. These include techniques for; sorting, searching, browsing and retrieval of images. Traditional image retrieval approach interprets image by text and then use textual information to retrieve images from textbased database management system. This method has several drawbacks; it uses keywords associated with images to retrieve visual information. It is very tedious and time consuming. It is hard to describe the contents of different types of images with textual representation. Keywords due to their subjective natures fail to bridge the semantic gap between the retrieval system and the user demands; consequently the accuracy of the retrieval system is questioned. The keyword for describing images becomes inadequate in large databases. It is not scalable.

Content Based Image Retrieval (CBIR) is a powerful tool. It uses the visual cues to search images databases and retrieve the required images. It uses several approaches and techniques for this purpose. The visual contents of images, such as color, texture, shape and region, are extensively explored for indexing and representation of the image contents. These low level features of an image are directly related to the contents of the image. These image contents could be extracted from image and could be used for measuring the similarity amid the queried image and images in the database using different statistical methods. In content-based retrieval systems different features of an image query are exploited to search for analogous images features in the database.

Various techniques based on texture features have been proposed in the literature. These include both statistical approaches and spectral approaches. Mostly these techniques are not able to capture accurate information. Color is most reliable feature which is easier to implement for retrieval of image. Color is easier to implement because it is robust to background compilation. It is free of image size and its orientation. The most common approach for color features extraction of images is histogram. Color histogram illustrates the color distribution in image and it entails low computational cost. Color is also insensitive to trivial deviations in the assembly of image. The main shortcoming of color histogram is that they cannot fully consider spatial information and they are not exclusive [2]. Different images having same color distribution yield almost similar histograms. Besides, in diverse lighting conditions analogous images having same point of view generate dissimilar histograms. Despite of using the information extracted from image, most of the CBIR systems yield imprecise outcomes. Because it is challenging to relate the low-level features with the high-level user semantics. This problem is known as semantic gap [10]. To over-come the problem of semantic gap, relevance feedback methods are used in [1]. Relevance feedback method provides a mechanism for CBIR system to allow the system to learn about the features best serve the user's interests. This method enable user to assess the images retrieved by the current query and assign them values which indicates their relevance.

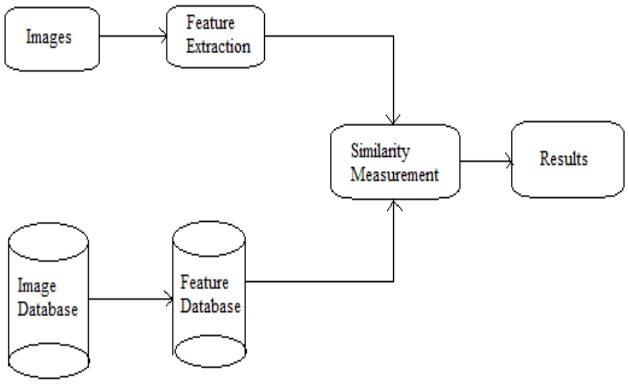

Figure1: CBIR System

II. LITERATURE STUDY

Shaoyan Sun et. al, [1] 2018, Image retrieval has achieved remarkable improvements with the rapid progress on visual representation and indexing techniques. Given a query image, search engines are expected to retrieve relevant results in which the top-ranked short list is of most value to users. However, it is challenging to measure the retrieval quality on-the-fly without direct user feedbacks. In this paper, we aim at evaluating the quality of retrieval results at the first glance (i.e., with the top-ranked images). For each retrieval result, we compute a correlation based feature matrix that comprises of contextual information from the retrieval list, and then feed it into a convolutional neural network regression model for retrieval quality evaluation.

Shaoyan Sun et. al, [2] 2017, Similarity measurement is an essential component in image retrieval systems. While previous work is focused on generic distance estimation, this paper investigates the problem of similarity estimation within a local neighborhood defined in the original feature space. Specifically, our method is characterized in two aspects, i.e., "local" and "residual". First of all, we focus on a subset of the top-ranked relevant images to a query, with which anchors are discovered by methods such as averaging or clustering. The anchors are then subtracted from the neighborhood features, resulting in residual representations.

Wengang Zhou et. al, [3] 2017, In content-based image retrieval, SIFT feature and the feature from deep convolution neural network (CNN) have demonstrated promising performance. To fully explore both visual features in a unified framework for effective and efficient retrieval, we propose a collaborative index embedding method to implicitly integrate the index matrices of them. We formulate the index embedding as an optimization problem from the perspective of neighborhood sharing and solve it with an alternating index update scheme.

Ziqiong Liu et. al, [4] 2017, Recently, feature fusion has demonstrated its effectiveness in image search. However, bad features and inappropriate parameters usually bring about false positive images, i.e., outliers, leading to inferior performance. Therefore, a major challenge of fusion scheme is how to be robust to outliers. Towards this goal, this paper proposes a rank-level framework for robust feature fusion. First, we define Rank Distance to measure the relevance of images at rank level. Based on it, Bayes similarity is introduced to evaluate retrieval quality of individual features, through which true matches tend to obtain higher weight than outliers. Then, we construct the directed Image Graph to encode the relationship of images. Each image is connected to its K nearest neighbors with an edge, and the edge is weighted by Bayes similarity.

Wengang Zhou et. al, [5] 2017, The explosive increase and ubiquitous accessibility of visual data on the Web have led to the prosperity of research activity in image search or retrieval. With the ignorance of visual content as a ranking clue, methods with text search techniques for visual retrieval may suffer inconsistency between the text words and visual content. Content-based image retrieval (CBIR), which makes use of the representation of visual content to identify relevant images, has attracted sustained attention in recent two decades. Such a problem is challenging due to the intention gap and the semantic gap problems.

III. IMAGE CONTENT DESCRIPTOR

An image content Descriptor can be local or global. It can be specific as well as general. Global uses features of thewhole image and local divides image into parts first. A simple method of partition is to use a division i-e cut image into regions having equal shape and size. They may not be meaningful and significant regions but it is a process to represent global features of any image. Partition the image into similar and homogeneous areas is an improved method with the use of some standard such as the Region Segmentation algorithms. Another complex method is to obtain semantically meaningful objects by Object Segmentation. The image content is further classified into two broad categories as visual content and semantic content.

A. Visual Content

The visual content is further classified into two main classes

1. General Visual Content

When the features or content of the query image are visible and are generally perceived then they fall into this category. Included common visual contents are the features like shape, texture, color, structure, spatial relationship etc.

2. Domain Specific Visual Content

When the query is based on such content that requires some domain knowledge then those query image content are domain specific visual content like Human Face Detection needs some prior information about the human facial characteristics. These characteristics are not general; these are specific i-e only related to human facial features.

B. Semantic Content

The semantic contentsare either described by textual explanation or by using the complex interpretationmeans which are based on some visual-content. 2.2. Image Retrieval Gaps: The differences between images stored in database and the query image in retrieval are called gaps. The degree of difference will be how far the two images are i-e the gap between the two images. The gaps are divided into two categories: semantic and sensory.

1. Semantic Gaps

The mismatching of information of visual query data and the stored image information in the database is obtained. This selected gape to match the image on the similarities basis is called semantic gap. User entered some queries for which optical likeness does not match completely with human perception. By which a semantic gap between CBIR system and the user is obtained. Semantic retrieval has some limitations. A difficulty present in it is that most of the images have more than one semantic interpretation. Because images used for training have usually short description in form of a caption, therefore, some features might never be recognized. This helps to decrease the amount of images instances used for training and weakens the system's capability to be trained for the concepts that are rare and which have a high variable visual appearance. Semantic retrieval system has a limited vocabulary so it mostly generalizes everything other than the semantic space i.e. for which is not trained.

2. Sensory Gaps

These are the gaps between the real object and the information in form of computational description obtained from capturing that object in an image form. It is the shortcoming of the image capturing device.

IV. PROBLEM IDENTIFICATION

This work is inspired by the strong performance of convolution neural net- works (CNN) in image classification tasks, and the qualitative evidence of their feasibility for image retrieval provided. A subsequent report demonstrated that features emerging within the top layers of large deep CNNs can be reused for classification tasks dissimilar from the original classification task. Convolution networks have also been used to produce descriptors suitable for retrieval within the Siamese architectures. In the domain of \shallow" architectures, there is a line of works on applying the responses of discriminatively trained multiclass classification as descriptors within retrieval applications. Thus, uses the output of classifiers trained to predict membership of Flickr groups as image descriptors. Likewise, very compact descriptors based on the output of binary classifiers trained for a large number of classes (classiness) were proposed. Several work such as used the outputs of discriminatively trained classifiers to describe human faces, obtaining high-performing face descriptors. The current state-of-the-art holistic image descriptors are obtained by the aggregation of local gradient-based descriptors. Fisher Vectors is the best known descriptor of this kind, however its performance has been recently superseded by the triangulation embedding suggested. The dimensionality reduction of Fisher vectors is considered, and it is suggested to use Image-Net to discover discriminative low-dimensional subspace. The best performing variant of such dimensionality reduction is based on adding a hidden unit layer and a classifier output layer on top of Fisher vectors. After training on a subset of Image-Net, the low-dimensional activations of the hidden layer are used as descriptors for image retrieval. The architecture of retrieval therefore is in many respects similar to those we investigate here, as it is deep (although not as multi-layered as in our case), and is trained on image-net classes. Still, the representations derived are based on hand-crafted features (SIFT and local color histograms) as opposed to neural codes derived from CNNs that are learned from the bottom up.

There is also a large body of work on dimensionality reduction and metric learning. In the last part of the paper we used a variant of the discriminative dimensionality reduction similar to others. Independently and in parallel with our work, the use of neural codes for image retrieval (among other applications) has been investigated. Their findings are largely consistent with our; however there is a substantial difference from this work in the way the neural codes are extracted from images. Specifically, extract a large number of neural codes from each image by applying a CNN in a jumping window manner. In contrast to that, we focus on holistic descriptors where the whole image is mapped to a single vector, thus resulting in substantially more compact and faster-to-compute descriptors, and we also investigate the performance of compressed holistic descriptors. Furthermore, we investigate in details how retraining of a CNN on different datasets impact the retrieval performance of the corresponding neural codes. Another concurrent work investigated how similar retraining can be used to adapt the Image-Net derived networks to smaller classification datasets.

The problem identification in existing work is as follows:

1. Image Retrieval Quality becomes low due to high error rate.
2. The linear consistency of the prediction with ground truth labels becomes low, hence more consistent images may not retrieved properly.
3. Retrieval prediction becomes low then, insufficient images retrieved.

REFERENCES

[1] Shaoyan Sun, Wengang Zhou, Qi Tian, Fellow, Ming Yang and Houqiang Li, "Assessing Image Retrieval Quality at the First Glance", IEEE Transactions on Image Processing, 2018.

[2] Shaoyan Sun, Ying Li, Wengang Zhou, Qi Tian, Houqiang Li, "Local residual similarity for image re-ranking", Elsevier Journal of Information Sciences 417, 2017.

[3] Wengang Zhou, Houqiang Li, Jian Sun and Qi Tian, "Collaborative Index Embedding for Image Retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 2, No.5, December 2017.

[4] Ziqiong Liu, Shengjin Wang, Liang Zheng and Qi Tian, "Robust Image Graph: Rank-Level Feature Fusion for Image Search", IEEE Transactions on Image Processing, 2017.

[5] Wengang Zhou, Houqiang Li and Qi Tian Fellow, "Recent Advance in Content-based Image Retrieval: A Literature Survey", IEEE Transaction on Image Processing, 2017.

[6] Filip Radenovic Giorgos Tolias Ond_rej Chum, "CNN Image Retrieval Learns from BoW: Unsupervised Fine-Tuning with Hard Examples", IEEE Transaction on Image Processing, 2017.

[7] Giorgos Tolias and Herv´e J´egou, "Particular Object Retrieval with Integral Max-Pooling of CNN Activations", International CLR 2016.

[8] Wengang Zhou, Ming Yang, Xiaoyu Wang, Houqiang Li, Yuanqing Lin and Qi Tian, "Scalable Feature Matching by Dual Cascaded Scalar Quantization for Image Retrieval", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 2, No.4, January 2015.

[9] Shaoyan Sun, Wengang Zhou, Qi Tian, Houqiang Li, "Scalable Object Retrieval with Compact Image Representation from Generic Object Regions", ACM Transaction on Image Processing, 2015.

[10] Mattis Paulin, Matthijs Douze, Zaid Harchaoui, Julien Mairal, Florent Perronnin, Cordelia Schmid , "Local Convolutional Features with Unsupervised Training for Image Retrieval", IEEE International Conference on Computer Vision, 2015.