International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Content Based Image Retrieval Using Automatic Image Annotation

Ms. Kiran Gotmare[1], Sapna Khapre[2]

[1]Department of Computer Science and Engineering, G. H. Raisoni College of Engineering and Technology for Women
Nagpur, Maharashtra, India

[2]Professor, Department of Computer Science and Engineering, G. H. Raisoni College of Engineering and Technology for Women
Nagpur, Maharashtra, India

Abstract: ***Generally images are stored in the database by its name and search operation is performed by its name only. As the database contain large no. of images and other data, it is difficult to find the required image. The performance of existing method is inefficient and less accurate. Hence an approach is proposed in which image is identified by its contents. First the image is stored in the database and then while extracting the features it is considered as a single segment. Different features are extracted from the image. If there is only one content in the image then whole image is considered as single unit. Then each content belong to a particular image is tagged using a tag file. Precision and Recall values will be used while evaluating the performance of CBIR systems. User gives the query as input to the system which is actually the content in the image. To retrieve the image it tag each segment with other, and search image in database, if match found then required image will be given as output. As images take more time and space while transmission, this system sends tag file through network to reduce time and space. If images are similar then differentiate using another feature like color, texture, version or format.***

Keywords: Content-based image retrieval, feature selection, image annotation, object recognition, text-based image retrieval.

1. Introduction

As the database contain large no. of images and other data, it is difficult to find the required image. Exhaustive search is infeasible due to its extensive time requirement. The performance of existing method is inefficient and less accurate. Dealing with large databases makes the annotation problem more difficult and therefore an effective approach is needed to manage such databases. It is a process of assigning keywords to digital images depending on the content information. It is a mapping from the visual content information to the semantic context information. The advances of information and communication technologies allow the creation of image archives extensively. As a result, the size of images database archives is increasing rapidly.

Even though with the enhance of image search technology whether in the system itself in the stand alone system or search images via the Web, the search via image content is still potential field to enhance. In general, people do not spent time **labeling, organizing** or **tagging** their image collections.

- There are already **billions of digital images** stored on personal computers and in commercial databases.
- To find a required image for an ordinary user is a challenging task i.e. Image tagging & Retrieval.

The proposed scheme overcomes the drawbacks of existing scheme such as inefficiency and inaccuracy. It provides less search time and high retrieval accuracy. It stores the image in the database on the basis of its contents, and retrieves by contents only.

2. Related Work

2.1 Indexing and Storing of Image

Image is described in terms of high dimensional vectors. The vector can be of pixel value or texture. Normalization and quantization technique is used. An indexing scheme called MSIDX use value cardinality of descriptor to achieve high retrieval accuracy and low search time. It increases the probability of to similar image to lie in a close constant range. One of the descriptor used in [1] is CEDD incorporates color and texture information in histogram. Image is split into blocks according to texture types. Bins re formed by respective color and texture. Information is normalized to keep computational power low. It supports dynamic indexing and storing of new image content by real time insertion technique and efficient search time and retrieval accuracy. This method perform **Preprocessing** for indexing and storing of pre-existing image, **insertion** for real time indexing and storing of image **query processing** for searching similar image to a posed image query, **deletion** for deleting already indexed image.

2.2 Relation between Image Content and Tag

Tags are words or phrase which describes the content in images. Tags are extensively used on Web to describe image content. Two level data fusions between image content and tag is proposed in [2]. 1) Graph is built to fuse the visual feature-based image similarity graph with the image-tag bipartite graph. 2) A random walk model which utilizes a fusion parameter to balance the influences between the image contents and tags. User provides numerous tags to the web images, which gives meaning description to images. It helps

to those images having little or no textual context. Tag data help to improve various tasks in image retrieval.

Visual features and tags are two different but closely related aspects of images. Global features are extracted from images then infer image similarity graph and then form hybrid graph with image-tag bipartite graph. one part of the weighted edges are connecting different images and the weights represent the similarities between them, while the other part of the weighted edges are bonding the images and tags with the weights reflecting the co-occurrent frequencies, then random walk model use fusion parameter to balance importance between image content and tag [2].

Figure 1: Example image with its tags.

2.3 Feature Extraction

The global feature is efficient in computation and storage due to its compact representation. Types of global feature extraction are done in [2].

- **Grid Color moment:** It is use to extract color feature from image. An image is partitioned into a 3*3 grids. Three types of color moments are extracted from grid i.e. color mean, color variance and color skewness in each color channel R, G, and B.
- **Local binary pattern (LBP):** is defined as a gray-scale invariant texture measure, derived from a general definition of texture in a local neighborhood.
- **Edge:** An edge orientation histogram is extracted for each image. First each image is converted into a grayscale image and then employ edge detector to obtain the edge map for computing the edge orientation histogram.

2.4 Image Annotation

Image annotation is important to make huge unlabeled image indexable to existing text-based indexing and search method. In image annotation semantic tags or labels are assigned to digital image. Two key techniques are employed in [3]. One is to map the high-dimensional image visual features into hash codes and other is to implement it as a distributed system so that search and mining processes are provided as Web services. As no training data set is required [3] enables annotating with unlimited vocabulary and is highly scalable and robust to outliers. Since the surrounding text of Web image is noisy, data mining techniques can be adopted to remove noise and to figure out silent terms and phrases from search result to annotate the image.

Query-by-Keyword (QBK) is conducted which filter out semantically dissimilar images. It reduces image space and time for successive retrieval. Encoding of image visual features into hash code which are binary bitwise and efficient distance measure is used in [3].

Figure 2: Image with a tag.

Generally color feature encode global information such as intensity, illumination, saturation whereas texture features discover local information such as gradient, interest point, etc. Both of them are very powerful and have been widely used in many tasks such object recognition, segmentation, and image retrieval. Annotation precision is defined as the number of images that assigned the keyword correctly, divided by the total number of images predicted to have the keyword. The annotation recall is defined as the number of images that assigned the keyword correctly, divided by the number of images that assigned the keyword in annotation.

2.5 Attribute Adaptation

The goal of [4] is to predict attributes for target image given a well defined label set attribute. It extracts multiple type of visual feature to represent each image in target set. It takes advantage of transfer learning as well as semi supervised learning to improve the performance of attribute prediction for new image. It removes the shared irrelevance and noise by using a robust loss function. To perform image attribute adaptation [4] minimize the mismatch of different distribution. It helps to get optimal common feature space for domain adaptation.

2.6 Projection

Locality preserving projection [5] considers within locality structure as well as between locality structures while processing. It compromise between global and local structure and use especially on high dimensional data set. It is based on adjacency graph that use neighborhood information of the data points for within-locality and dissimilarity graph that use interneighbor information for between-locality. Then both graphs are combined to get the projection. It is use in training data set and suitable for high dimensional and dynamic databases. [5] Preserves the local structure of data set and retain the between-locality information contained in the data set.

2.7 Detection of Duplicates

Copy retrieval scheme based on local features and can deals with very large databases in terms of quality and speed. Selection of features is based on distribution of the feature distortion. As it is based on local features approximation is strong and reduces the amount of data to explore. The original document is not identified by previously inserted mark but by the content-based extracted signature. First the local features are extracted called neighbors and then in post processing it count the number of local matches to detect the duplicates. Features are not extracted from each point but from unique key points[6].

3. Proposed Methodology

Image is first goes through processing stage where features are extracts from image. Each feature is considered as one content and tagging is provided to it. While the feature extraction is going on, whole image is considered as one unit so as to avoid segmentation. Because in segmentation sometimes the content of image get partitioned and that portioned piece of content will not get tagged with any keyword. After the interested feature is extracted by system the tagging is made i.e. particular name is given to the feature as per the content. The tag file describes the context information of the content. The tag file is store in the database and requires less space. When user sends a query to the system regarding the content, first the annotation is built on it. Annotation is compared with the inbuilt annotation file, if match found, then particular image is given as output. . If only one content is available or content having only one color then whole image is considered as one unit. As the feature extraction is made manually the accuracy is 100%. Precision and recall is employed to evaluate the performance. The main objective of this proposed scheme is to overcome the drawbacks of existing scheme such as inefficiency and inaccuracy.

4. Conclusion

By using the vector, this approach proposed scheme that can easily fragment the images by its contents. Unique tag is applied to each content so that no confusion is made while searching and retrieving image. Each fragment will be tagged with a tag file, to map visual content information to semantic context information. Using this mapping it can retrieve the accurate image. If same tag is available in more than one image then all the related images are given as output. With the images in output all the related tags and class of all tags is also shown so that user can choose their image of interest. If images are similar then differentiate using another feature of image like color, texture, version or format. Instead of sending image through network if tag file or annotation is send then it saves transmission time and memory. Search and Retrieval operation takes less time. Provide more accurate information.

References

[1] Eleftherios Tiakas, Dimitrios Rafailidis, Anastasios Dimou, and Petros Daras, “MSIDX: Multi-Sort Indexing for Efficient Content-Based Image Search and Retrieval” IEEE Transactions On Multimedia, Vol. 15, No. 6, October 2013
[2] Hao Ma, Jianke Zhu, Michael Rung-Tsong Lyu, and Irwin King, “Bridging the Semantic Gap Between Image Contents and Tags” IEEE Transactions On Multimedia, Vol. 12, No. 5, August 2010
[3] Xin-Jing Wang, Lei Zhang, Xirong Li, and Wei-Ying Ma, “Annotating Images by Mining Image Search Results” IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 30, No. 11, November 2008
[4] Yahong Han, Yi Yang, Zhigang Ma, Haoquan Shen, Nicu Sebe, and Xiaofang Zhou, “Image Attribute Adaptation” IEEE Transactions On Multimedia, Vol. 16, No. 4, June 2014
[5] X. He and P. Niyogi, “Locality preserving projections,” in Proc. NIPS, 2003
[6] Alexis Joly, Olivier Buisson, and Carl Frélicot “Content-Based Copy Retrieval Using Distortion-Based Probabilistic Similarity Search” IEEE Trans. On Multimedia, Vol. 9, No. 2, February 2007