

Object Based Image Retrieval from a Repository

Laxmidevi Noolvi, M V Sudhamani

Abstract—Today is a digital world. Due to the increase in imaging system, digital storage capacity and internetworking technology Content Based Retrieval of Images (CBIR) has become a vibrant research spot. The CBIR systems helps user to browse and retrieve similar kind of images from huge databases and World Wide Web. The Object based Image Retrieval (OBIR) Systems are the extension to the CBIR technique where it retrieves the similar images based on the object properties. So far massive amount of work has been done in this field of research. A plenty of the techniques and algorithms are published in the different papers. This paper provides brief survey on basic and recent approaches and techniques explained in different papers.

Keywords: CBIR, OBIR, Features, Semantic gap, Similarity Measures.

I. INTRODUCTION

The advancement and wide spread of technology has been lead to digital revolution. The substantial increase in the digital storage capacity and digital imaging creates mountain of image database. To handle and make use of stored images, efficient and effective tool is necessary, which is capable of retrieving images from large database has became crucial. The image retrieval systems started emerging in the early days of 1990's. Those systems were based on the text keyword search which is called as text or keyword based image retrieval systems (TBIR). In these traditional TBIR systems images were attached with some keywords which describe that image and similar images are retrieved by matching keywords of the image present in the database. Such system requires manual labelling and also categorizing large scale of images so that those are not feasible for larger collection of images.

In contrast to TBIR, content-based image retrieval (CBIR) method was gradually taking place later than. CBIR effectively addressed the issues with the text annotation. CBIR works based on the visual information which is embedded in the image itself, called as content present in an image. The image information may be texture, shape, color and spatial location of image. These are often called as low level features of image. CBIR is basically a technique wherein user submit query image to the system and it provides relevant images to the users. There are two types of CBIR systems, one is based on the high level features and another is on low level features.

However, CBIR systems focused on low level methods of different feature extraction, the accuracy of the results obtained subsequently increased but still image retrieval with these low level features has been restricted to very low efficiency. The high level concepts (semantics) in the users mind will not be portrayed by such features. These low level features are just statistical representations which are far away from the human perception. This mismatch between these features is termed as semantic gap. This has become open challenge in the current research work on CBIR.

Many researcher have been addressed this issue of semantic gap to some extent so far. These systems are based on the objects and their properties present in the image. Because an object relates more to human perception than any other attributes of an image. They are called as Object based image retrieval (OBIR) systems. On the other side many of the techniques have emerged with new techniques such as Relevance Feedback, Intermediate Description termed as Object Ontology method, use of machine learning classifiers [16]etc The paper organization is as follows. Basic architecture of CBIR system explained in section 2. Applications of CBIR system are given in section 3. Brief survey on basic and recent Feature extraction techniques used by different researcher is given in section 4. Different similarity/distance measures used for retrieval are summarised in the section 5. Section 6 gives concise performance evaluation measures. Finally paper end with concluding the work.

II. BASIC CBIR SYSTEM ARCHITECTURE

The architecture of the CBIR system showed in Figure 1. The system can be explained in two phases, in the offline phase, the repository images visual features are extracted then stored in the features database. In the online phase, the query image is submitted through graphical user interface (GUI) by the user and vector of features of a query image are extracted in the similar manner as explained in the offline phase. The query image features are compared with the features in the database using similarity measures. Finally, the visually similar images are retrieved.

Revised Manuscript Received on December 15, 2019.

Correspondence Author*

LaxmideviNoolvi, CSE, RNSIT, Bangalore, India. Email: laxmidevinoolvi@gmail.com

Dr. M V sudhamani, ISE, RNSIT, Bangalore, India. Email: mvsudha_raj@hotmail.com

Object Based Image Retrieval from a Repository

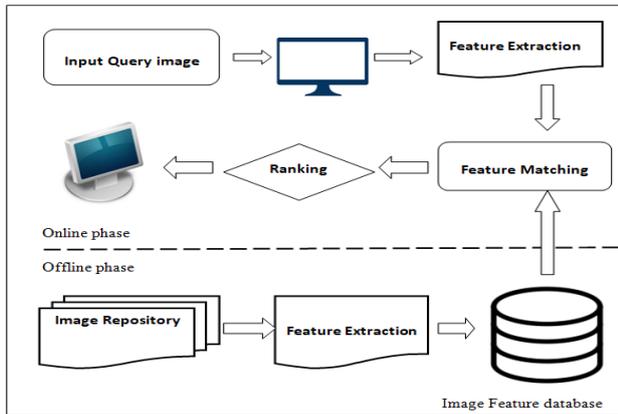


Figure 1: Basic Architecture of CBIR

III. APPLICATIONS OF CBIR

The CBIR is applied in wide range of real time applications are listed below:

- Searching Web (WWW)
- In the military tasks
- Face detection for Security purpose
- Prevention of crimes
- Geographical information and remote sensing systems
- Engineering and Architectural Designs
- Preventing fake registration of trademarks in Intellectual property rights protection
- In search of new Fashion and interior designs
- In the field of Advertising and Journalism
- Cultural heritage
- Home entertainment
- Automatic Medical diagnosis
- Education and training

IV. BRIEF SURVEY ON FEATURE EXTRACTION OF BASIC AND RECENT TECHNIQUES

Feature extraction plays very vital role in fast and efficient retrieval of visually similar images to the query image from the image repository. At the early stage of CBIR, it was rely on simple visual features of image. The features extracted must be invariant to changes and must be discriminative in nature as a consequence they act as significant for image retrieval. Image retrieval systems approaches can be broadly classifies as two categories; first, which uses basic primitive low level features and second, semantic CBIR which tries reduce semantic gap by the use of advanced methods.

A. Basic Feature extraction approaches

At the early stage of CBIR, it was bring into being on simple visual features of an image. Plenty of work on CBIR was done; initially it was based on shape, texture and color. This section presents the basic and recent work on CBIR.

Color features

Color is very important not only adding attractiveness to objects but also provide additional information. Color is used as outstanding feature in many CBIR systems. Its high dimensional value makes its discrimination superior than single gray values. Color features may be calculated based on either global (whole image) or local (Partitioned image). At the first foot step of CBIR used color moments, color histogram [22], Color Correlogram, Color Coherence Vector (CCV), and Invariant color features [8]. In [18] Color Histogram for K-Means (CHKM), histogram intersection, color co occurrence matrix, dominant color descriptor are explained, Zernike chromaticity distribution moments, color difference histogram that aggregates the visually identical color differences between two pixels under different surroundings, and also it presented the comparative study of color feature extraction methods.

In [5] effective color feature depends on its color space used (RGB, HSV, Lab, etc...). Summary of color space comparisons presented and entropy of color histogram bins used as visual features. In [21] explained the Color Fuzzy Correlogram (CFC) and Color Layer Based Texture Elements Histogram (CLBTEH). In CLBTEH color space will be quantized into 8 layers in Hue, 3 layers in saturation and value respectively then 16 texture elements are convoluted and constructed histogram of each texture element present in each color layer it will result in 224 (16 x 14) dimensional feature vector. In CFC it integrates the advantages of color Correlogram and color auto Correlogram. It divides the image I into the range of $[I_{min}, I_{max}]$ and takes central pixel and finds the fuzziness with the surrounding pixels with different distances based on ambiguity function. Then add all the fuzziness values of the central pixel and surrounding pixels that becomes the fuzzy correlation value of central pixel. Finally same color values of fuzzy correlation values are added up. These CFCs are values of all quantized sub images and calculated at different distances.

The Quantized Hue, Saturation and Value of an image with associated semantic color names for segmented regions are used in order to reduce gap between semantic and low (machine) level features [9, 11]. Color Strings Comparison explained in [17]. In [3], The input color images is color clustered using clustering algorithm then representative color are assigned to every cluster based on average value of colors in each cluster, and fraction of pixels in particular color cluster and total pixels in image taken as feature vector. In [27] converted RGB to YCbCr color space and then applied the canny edge operation on Y plane, later combined Cb and Cr components formed color edge image, later converted back to RGB image, on each individual RGB image channel histogram of edges are calculated.

Texture feature

The many of real-time images are described by their textures. Therefore, texture is essential feature in providing visual semantics for image retrieval.

There is no precise definition for texture only it represents certain regularities in the intensity values. In [20, 22] explained Texture approaches are divided into four classes 1. Structural approach: It is arrangement of intensity values of pixels. The morphological operators, edge base properties are best representation of these. 2. Statistical approach: These are calculated statistics of the pixel values. Gray level co occurrence matrix features such as correlation, contrast, and entropy, energy, and homogeneity, first order histogram based features. 3. Model based approach: these are analytical models of two type fractals and stochastic model. 4. Transform based features: these are calculated features on transformed space of image such as Gabor transform, Fourier transform and wavelet transform.

The Diagonal texture structural descriptors (DTSD) that calculates sum of sub divided 4 x 4 sub-image diagonal values then based on pre defined equations find the specific number for each divided section, finally finds histogram of such values as features, detailed in [26]. The Tamura features include statistical features of six types; roughness, regularity, coarseness, contrast, directionality and line likeness. Wold's feature explains perceptual characteristics of texture features such as harmony, evanescent and in deterministic. Simultaneous auto regressive (SAR) model and multi resolution auto regressive (MRSAR) model explained in [8].

Shape features

Shape is very good perceptual view of object or region. The various works on CBIR systems used shape features for retrieve images. The shape of object is geometric property and well represented when it is used with effective segmentation algorithms. Where segmentation is separating objects from its background. Basically calculating shape features of object mainly divided into two types [2, 14, 19, 22]. The region based features and boundary or contour based features. In the former approach features are calculated statistics of shape boundary or contour pixels. In the region based features different statistics of whole region are calculated.

The IBM's QBIC system [1], which retrieves the images based on the user drawn sketches of shapes of desired object in an image. In [2], contour based features such as chain code which is represented as sequence of edge code calculated based on boundary traversal either clockwise or anti-clockwise to the direction from current to next pixel. Because of non unique chain code variant method called derivative of the chain code was developed as derivative chain code (DCC). The Fourier descriptor [8] which represent object shape with Fourier transform of contour pixels. UNL Fourier descriptor [2] is improved version of Fourier descriptor. Histogram of edge directions explained in [22].

The region based features such as circularity, area, major axis; eccentricity and minor axis are simple form of features. The moment invariants [2] are the general features for image retrieval, due to their invariance property with respect to rotation, translation. The turning angle is angle between the pixels on contour and counter clock wise tangent as arc

length function [8]. The Pseudo Zernike moments and Zernike moments are also used in work [22]. In [14] some of the work combined shape features to increase the retrieval rate.

B. Semantic CBIR/OBIR approaches

Even though massive amount of work has been carried out using visual features (shape, color and texture) and got appreciated results but when it comes to human thinking level using only these features cannot provide retrieved results up to the understanding of human beings [11]. After all we human are ultimate users of these systems. The systems must reduce crack between high semantic features and machine centric low features in line to connect to human perception. This is the current open challenge in this field of CBIR systems. Here we explain some of such works, which addressed this issue to an extent.

In [4], used support vector machine (SVM) with basic CBIR which will classify retrieved results based on relevant feedback. The retrieved images are classified as relevant and non-relevant examples and learned by SVM classifier. The learning algorithm will involuntarily update weights for positive relevant images. This avoids the users from manual assignment of preference weight for all positive relevant images and it also makes use of negative image information. The superior results are gained as compared to that of systems with positive feedbacks.

In [15], explained object based image retrieval, there are three contributions, first, a novel method which is modification of scale invariant feature transform (SIFT) descriptors called Root-SIFT is a square-root kernel (hellinger kernel) not a benchmark Euclidean distance that measures the SIFT descriptor similarity; second, query expansion: it is distinctive approach where non relevant image called negative images are taken into account and trained SVM classifier which provides weight vector for retrieved images. The inverted index used for retrieval; (iii) an image augmentation method is used for further improvement. In [16], the proposed technique, explains different approaches for reduction of the gap in two types of features. The low level features (texture, colour and shape) are computed and then applied Relevance Feedback and Adaptive Clustering on initially retrieved results and then obtained the final results and it helps to bridging the gap.

In [9], the paper is having three parts, segmenting image, color naming for regions and then assigning ranks to images. Initially, for all database images, each is partitioned into regions. After that, colour name is defined semantically for each region. In the method uniformly divide the Hue value into ten colour names, violet, blue, red, yellow, aqua, aquamarine, magenta, orange, green and purple. Saturation and intensity values are divided (inconsistently) into 4 parts respectively as adjectives that indicates saturation and luminance. In the retrieval process, it is done through by comparing the query images regions and database images regions and selected as candidate images which are having same color names.

The results are further re-ranked based on their similarity to the input query image.

In [7], the work is on retrieval of generic image from the large collections without commanding restrictions such as the accessibility of query-images or annotations for the images. A novel intermediate-level descriptors for color features are presented for each region attributes. The simple vocabulary called as object ontology is formed using the intermediate level descriptors. In this provided method, in the first part some image regions are given as results to the user for their query, then results are evaluated by the user, and relevant objects and not relevant objects are marked; then these are given as input to train or learned by a support vector machine (SVM); and later for re-ranking a constrained similarity measure (CSM) is utilized.

In [10], explains equivalence between the current object-based image retrieval and conventional keyword annotated retrieval and presents techniques of objects retrieval by visual phrases. The approach presents algorithms to build the visual phrases using the feature descriptors in the image and phrases indexing. Later image similarity is measured based on visual phrases.

In [23], presented experimental analyses based on feature descriptors such as SIFT and SURF. The experiments are conducted on database which is consists of multi-objects formed using the Columbia Object Image Library (COIL-100) dataset. The SIFT and SURF features are compared and matched among the query and extracted features of image collection. The similar images are retrieved by similarity score which is obtained by matching.

In [24], the novel approach for image retrieval with novel feature extraction, features matching measures and it provides facility for selecting sub-space. The approach proceeds with clustering by organising the images collection, and then selects the best feature, matching measure by the help of relevance feedback provided by user. The feature Selection and matching measure selection is visualized to the user through clustering.

In [28], contributed the novel work with fusion of spatial colour features information, shape features and then object recognition. The RGB image features are extracted by spatial color arrangements; these are fused with the shapes formed by intensity edge connecting and interest points of the gray image. The experiments are carried out on 10 highly recognised dataset benchmarks and they got the significant results as compares to the state-of-art interest point descriptors methods.

In [29], addressed the issue of considering the holistic features used in the convolution neural network activation. The compact holistic feature does not capture the representation of the small objects present in an image. They considered the local features for end-to-end similarity learning and applied the correlation layers to the features for local aggregation. In [30], proposed region based deep learning technique for instance retrieval, it first segments the interested regions (RoIs) from an image, and then extracts CNN features for each RoIs from the fully-connected final

layer of network. In the proposed method, target objects are identified even in the cluttered backgrounds by visual matching using RoIs-based features that best describe the patterns.

V. SIMILARITY/DISTANCE MEASURES

The similarity/ distance metrics are used to capture the relation among the database images and query image by computing the similarity or distance among query image features and database images features. The minimum distance ensures that two images are alike and those images are ranked with high score, accordingly further the images are organised based on their score. Many papers have used different techniques of similarity matching. The Histogram Intersection Distance is widely used in color based CBIR systems [12, 14]. The simple Euclidean distance which is sum of square root distance between the two feature vectors, it is explained in many [3, 5, 8, 9, 11, 12, 22, 24]. The Manhattan distance which is absolute distance explained in [22, 25]. Minkowski distance is also widely used in image retrieval [22]. Another distance measure called d1-distance defined as ratio of distance between pair of feature vector to the total sum of feature vector pair values and addition with some constant [26].

VI. RESULTS AND DISCUSSIONS

The massive work on content based image retrieval systems used precision or accuracy and recall [22] as measure of performance. Most of the works are compared with these measures. Precision (P) is fraction of Total number of Similar images Retrieved (TSR) to the Total Images Retrieved (TIR) by the system. The equation is given in (1).

$$P = \frac{TSR}{TIR} \quad (1)$$

Recall (R) is ratio of Total number of Similar images Retrieved (TSR) and Total Images present in the Database (TID). The equations is given in (2).

$$R = \frac{TSR}{TID} \quad (2)$$

The F-Measure or F-Score (F) is combined measure which denotes importance of precision over recall and it is also represents the harmonic mean of precision and recall [23]. The equations as given below,

$$F = 2 \times \frac{P \times R}{P + R} \quad (3)$$

In [4], to enhance the performance of these systems a concept called Relevance Feedback has been incorporated, where if the results of retrieved images are not satisfied by the user perception then images are further re-ranked and classified using the standard classifier. In [10], to improve the speed of the retrieval of images the indexing technique with inverted indexed retrieval is used. The images are indexed using the visual phrases.

VII. CONCLUSION

This paper deals with the brief explanation of various types of content based image retrieval systems and list of applications. It also explained the different basic and semantic features extraction methods of CBIR and OBIR systems. It provides the complete outline of the techniques which are present in this field of work. The different similarity/distance measures used for indexing and matching are dealt in this paper and finally the performance evaluation measures which are more crucial for the end users are outlined. This work can be taken it as summary of techniques which are available for the any CBIR/OBIR; it can help the beginners to start with their research work in this field.

REFERENCES

1. M. Flickner et al., "Query by Image and video content: The QBIC System," IEEE, 1995.
2. Mehtre, Babu M., Mohan S. Kankanhalli, and Wing Foon Lee. "Shape measures for content based image retrieval: a comparison," *Information processing & management* 33, no. 3, 1997.
3. Mehtre, Babu M., Mohan S. Kankanhalli, and Wing Foon Lee. "Content-based image retrieval using a composite color-shape approach," *Information Processing & Management* 34, no. 1, 1998.
4. Hong, Pengyu, Qi Tian, and Thomas S. Huang. "Incorporate Support Vector Machines to Content-Based Image Retrieval with Relevant Feedback," *In International Conference on Image Processing*, pp. 750-753. 2000.
5. Zachary, John, S. Sitharama Iyengar, and Jacob Barhen. "Content based image retrieval and information theory: A general approach," *Journal of the American Society for Information Science and Technology* 52, no. 10, 2001.
6. Müller, Henning, Wolfgang Müller, David McG Squire, Stéphane Marchand-Maillet, and Thierry Pun. "Performance evaluation in content-based image retrieval: overview and proposals," *Pattern recognition letters* 22, no. 5, 2001.
7. Mezaris, Vasileios, Ioannis Kompatsiaris, and Michael G. Strintzis. "An ontology approach to object-based image retrieval," *In Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, vol. 2, pp. II-511. IEEE, 2003.
8. Long, Fuhui, Hongjiang Zhang, and David Dagan Feng. "Fundamentals of content-based image retrieval," *Multimedia Information Retrieval and Management*, Springer, Berlin, Heidelberg, 2003.
9. Liu, Ying, Dengsheng Zhang, Guojun Lu, and Wei-Ying Ma. "Region-based image retrieval with high-level semantic color names," *In Multimedia Modelling Conference, 2005. MMM 2005. Proceedings of the 11th International*, pp. 180-187. IEEE, 2005.
10. Zheng, Qing-Fang, Wei-Qiang Wang, and Wen Gao. "Effective and efficient object-based image retrieval using visual phrases," *In Proceedings of the 14th ACM international conference on Multimedia*, pp. 77-80. ACM, 2006.
11. Liu, Ying, Dengsheng Zhang, Guojun Lu, and Wei-Ying Ma. "A survey of content-based image retrieval with high-level semantics," *Pattern recognition* 40, no. 1, 2007.
12. Yue, Jun, et al. "Content-based image retrieval using color and texture fused features," *Mathematical and Computer Modelling* 54.3-4, 2011.
13. Singha, Manimala, and K. Hemachandran. "Content based image retrieval using color and texture," *Signal & Image Processing* 3.1, 2012.
14. Abuhaiba, Ibrahim SI, and Ruba AA Salamah. "Efficient Global and Region Content Based Image Retrieval," *International Journal of Image, Graphics & Signal Processing* 4.5, 2012.
15. Arandjelovic, Relja, and Andrew Zisserman. "Three things everyone should know to improve object retrieval," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2911-2918. IEEE, 2012.
16. Dar, Showkat Ahmad, and Zahid Gulzar Khaki. "Content Based Image Retrieval," *IOSR Journal of Computer Engineering (IOSR-JCE)*, Volume 12, Issue 2, 2013.
17. Jenni, Kommineni, Satria Mandala, and MohdShahrizalSunar. "Content Based Image Retrieval using colour strings comparison," *Procedia Computer Science* 50, 2015
18. Srivastava, Divya, Rajesh Wadhvani, and Manasi Gyanchandani. "A Review: Color Feature Extraction Methods for Content Based Image Retrieval," *International Journal of Computational Engineering & Management* 18, no. 3, 2015.
19. Patel, Mitisha Narottambhai, and Purvi Tandel. "A Survey on Feature Extraction Techniques for Shape based Object

- Recognition," *International Journal of Computer Applications* (0975 – 8887) Volume 137 – No.6, March 2016.
20. S. Sonawane and V. S. Patil, "Image Retrieval Based on Texture-a Survey," *Proceedings of 55th IRF International Conference*, 22nd May, 2016.
21. Yang, Fu-ping, and Mei-li Hao. "Effective Image Retrieval Using Texture Elements and Color Fuzzy Correlogram," *Information* 8, no. 1, 2017.
22. S Rucha Patil, Dr. J Avinash, Agrawal, "Content-based Image Retrieval Systems: A Survey," *Advances in Computational Sciences and Technology*, ISSN 0973-6107 Volume 10, Number 9, pp. 2773-2788, 2017.
23. H. Kavitha, M. V. Sudhamani, Experimental analysis of SIFT and SURF features for multi-object image retrieval," *Int. J. Computational Vision and Robotics*, Vol. 7, No. 3, 2017.
24. Mosbah, Mawloud. "An Efficient and Effective Image Retrieval System on the basis of Feature, Matching Measure and sub-space Selection," *Journal of Information and Organizational Sciences* 42, no. 2, 2018.
25. Nazir, Atif, Rehan Ashraf, Talha Hamdani, and Nouman Ali. "Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor," *In 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1-6. IEEE, 2018.
26. Song, Wei, Yubing Zhang, Fei Liu, Zhilei Chai, Feng Ding, Xuezhong Qian, and Soon Cheol Park. "Taking advantage of multi-regions-based diagonal texture structure descriptor for image retrieval," *Expert Systems with Applications*, 2018.
27. Rehan Ashraf, Mudassar Ahmed, Sohail Jabbar, Shehzad Khalid, Awais Ahmad, Sadia Din, Gwangil Jeon. "Content Based Image Retrieval by Using Color Descriptor and Discrete Wavelet Transform," *Image & Signal processing*, 2018.
28. Khawaja Tehseen Ahmed, Shahida, Muhammad Amjad Iqbal, "Content Based Image Retrieval using Image Features Information Fusion," *Information Fusion*, 2019.
29. Chen, Zhenfang, et al. "Learning Local Similarity with Spatial Relations for Object Retrieval," *ACM MM*. Vol. 3. 2019.
30. Chen, Jingcheng, et al. "Instance Retrieval Using Region of Interest Based CNN Features," *Journal of New Media JNM*, 2019.

AUTHORS PROFILE



Learning.

Mrs. Laxmidevi Noolvi, she is currently pursuing her Ph.D degree in the department of Computer science and Engineering, RNSIT, Bangalore. of Industry experience. Her areas of research interests include Web Mining, Semantic Web, Artificial Intelligence and Machine



Dr. M V Sudhamani, currently working as Dean-R&D, Professor and HoD, Dept. of ISE, RNSIT. She is having Teaching, Research and Industrial experience of 25 years. She has specialization in Image Processing, Content-based Image Retrieval, Advanced Algorithms and Databases. Guided and guiding candidates for Ph. D degree. She has carried out two research projects from VTU and AICTE. She has served as member of Board of Examiners (BOE) and Board of Study (BOS) member in VTU and other autonomous institutions across India. She has organized two international conferences ICDECS 2011 and 2015, and one more in December 2019