

Chapter 37

Survey and Analysis of Content-Based Image Retrieval Systems



Biswajit Jena, Gopal Krishna Nayak, and Sanjay Saxena

Introduction

The advent of the Internet facilitated the exchange and querying of information. Over the years, the methodology adopted by users to query data has witnessed considerable changes primarily because of the querying and data retrieval mechanisms on user-end getting easier, interactive, and friendly. The earlier years of the Internet era saw a greater amount of text-based data being generated, queried, and transferred but as the feature of multimedia got incorporated with the textual Web pages and applications; the shift has been toward image-based retrieval schemes. Earlier, this was brought forward by text-based image retrieval systems [1], where the visuals were annotated manually by textual phrases or words. When such a system was used to query a particular image from a large database of textually annotated images, it often would suffer from imprecision in the search results. This was chiefly because different humans might perceive an image differently. Also, the process of annotating an image was time-consuming and required a lot of human effort. To combat this issue, the CBIR was put forward in the early 1980s [2]. Since then, it has become a lively research area backed by a variety of different individual fields like pattern recognition, machine learning, computer vision, and databases to name a few.

The fundamental goal of any CBIR system is feature extraction [1, 2]. An object in such a system is described in terms of its low-level features like texture, color, and shape.

Often, human beings tend to identify an object from its color; thus, we can closely link color with the visual perception of an object in the human mind. To study this feature, a lot of techniques are applied to color perception and color spaces. Color

B. Jena (✉) · G. K. Nayak · S. Saxena
International Institute of Information Technology, Bhubaneswar 751003, India
e-mail: biswajit310@gmail.com

histograms are one of them. In this, the focus is on the distribution of color in an image regardless of the spatial location where that particular color might be found in an image. They are typically employed on three-dimensional color spaces like HSV and RGB and are extremely useful because of their flexibility, low computational complexity, and compact representation. Another important low-level feature is the texture which is responsible for defining the spatial positioning of colors or various intensities in an image. Gray-level co-occurrence matrix (GLCM), Gabor filter, wavelet transform, and curvelet transform are some ways of texture representation [6–9]. The shape-based features are extremely helpful yet a challenging problem faced by CBIR systems. They help us to index the objects. An object's shape helps it to describe it more meaningfully yielding an efficient search result. But such features have a complex implementation in CBIR systems that are desired to be 2D or 3D and must show invariance to properties like translation, rotation, and scaling. Generally, shape descriptors are of two types, i.e., contour-based (representing the boundary) and region-based (representing the entire region) [3, 14].

Keeping these low-level features into consideration, a lot of CBIR systems like QBIC, Netra, Photobook, Virage, FIRE, LIRE, etc., have been developed. The ultimate goal is to minimize the interval between these low-level numerical and statistical features and high-level perception of them by the human mind. The ability with which a CBIR system is able to reduce this gap portrays its efficiency and usability. Thus, a lot of work is ongoing so as to develop a semantically and statistically synchronized CBIR system.

The organization of our paper is such that in Sect. 37.2, we discuss the basic design of a CBIR system. In Sect. 37.3, we analyze the various CBIR systems available in the literature, and in Sect. 37.4, we conclude the paper.

A Typical CBIR System

In this section, we discuss the design of a typical CBIR system. All CBIR systems begin with users generating query images. The query image may undergo some kind of preprocessing which is essential for enhancing the quality of the query image and suppress the unwanted features like noise or distortion present in the image. If the data is voluminous and redundant, it becomes difficult for algorithms to handle it. So, a feature vector is created comprising all the important data. This phenomenon is known as feature extraction [4]. The extracted features are a set of reduced relevant features that hold all the necessary information on which all tasks are performed. Now, then there is a comparison between the extracted feature vector and feature databases. The degree of similarity between the feature vector of the query image and the corresponding images present in the feature databases defines the degree of precision with which an image is retrieved by a CBIR system.

A number of retrieval approaches are used in different CBIR systems. Query by example, iterative search, semantic retrieval, and relevance feedback [5] are some of them. In query by example, an exemplary image is fed into the CBIR system with

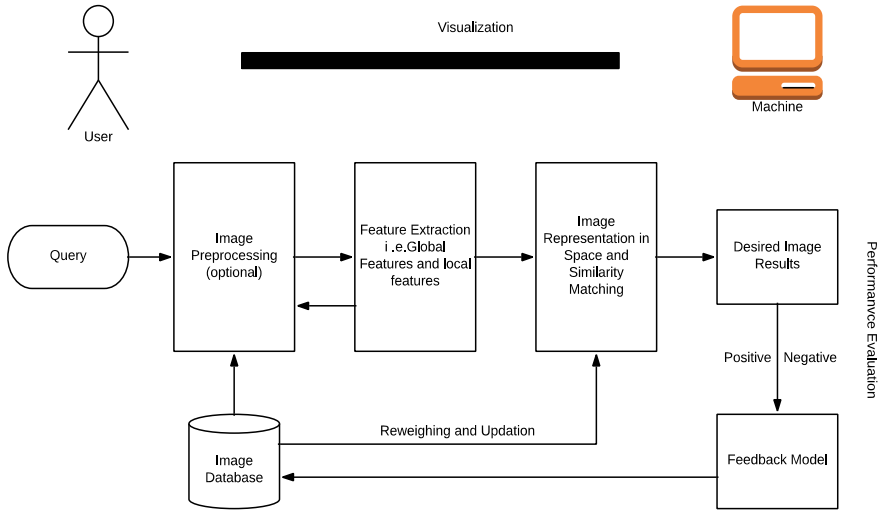


Fig. 37.1 Visualization of CBIR at the user and machine level

which it tries to match the query image. In iterative search, several machine learning methodologies are employed to retrieve the relevant data. In semantic retrieval, the user queries for images by words and phrases. This is the most difficult means of implementing a CBIR as the queries made by users are open to interpretation and it is difficult for a machine to semantically determine the correct meaning which the human user might be querying. Relevance feedback involves human intervention. The images retrieved are classified into relevant, irrelevant, and neutral by users. Then, this updated information is used to make a new search. Thus, a progressive search with relevant user feedback refines the overall search. Figures 37.1 and 37.2 show a typical CBIR system’s functionality.

Analysis of Different CBIR System

1. Analysis-1 [6].

The important techniques used here are basically for feature extraction. Here, we can use both spatial and frequency domain techniques of feature extraction. The spatial domain techniques include color moments, color auto-correlogram, and HSV histogram features. Similarly, the various frequency domain feature extractor techniques are Gabor wavelet transform [15], SWT [16], and BSIF [17].

The precision was obtained using Chebyshev, Cosine, L1, and L2 distance metrics for different classes of images. The experimentation takes place considering four special feature descriptors.

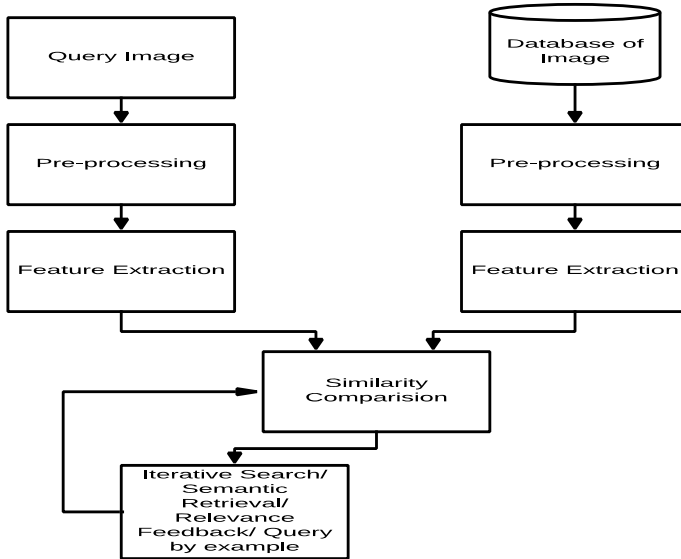


Fig. 37.2 A typical CBIR system's framework

The various distance metrics that are used in this study are Minkowski, City block, Mahalanobis, and Euclidean. These give an average accuracy near to 65 percent with respect to all the classes of images with respect to spatial, frequency domain features as well as CEDD [18], BSIF fusion features, and hybrid features.

2. Analysis-2 [7].

The theme of this article is to do the post-CBIR system, in which the large database is first partitioned into various clusters based on various image features like color, size, texture, etc., for efficient image retrieval. For clustering purposes, the ACPSO clustering algorithm [19] is chosen over the commonly used clustering algorithms like PSO [20], K-means [21], ACO [22], etc.

After that, the highly efficient ACPSO techniques are employed for image retrieval from the clustered large database of images.

The resulted accuracy for K-means is 0.91, ACO is 0.88, PSO is 0.96, and ACPSO 0.98 is noticed.

3. Analysis-3 [8].

Patterns from images were extracted using the techniques like local binary patterns, local mesh patterns, local texton XOR pattern, and local ternary co-occurrence pattern comparing their histograms and then constructing feature vectors concatenating all the histogram and comparing the query image with the database image.

To measure the performance of the system, various performance parameters used are average retrieval rate or precision, recall, and precision.

The results show that on different quantized levels, Corel-1k, Corel-k, and Corel-10k vary and gradually quantized level of 4 and 8 clearly outperforms the other level (all in multiples of 4) at each of the databases.

4. Analysis-4 [9].

Here, GLCM [23] is used as well as color histogram so that the color and texture features can be useful in the classification of cow type. The various features for GLCM are energy, contrast, homogeneity, correlation, and entropy. The angle for each is 0, 45, 90, 135° with an average of 1.

The system is trained with 100 images and tested with 20 images of five different classes. Based on these test images, various performance metrics such as recall, precision, and accuracy are measured using a confusion matrix.

The results obtained from the system show that being a robust system for image retrieval with an accuracy of 95%, precision, and recall of 100%.

5. Analysis-5 [10].

The techniques that are used for feature extraction are GLCM and discrete wavelet transform (DWT) and its combination for both texture-based and color-based features.

The database used for retrieval purposes is the WANG image database. It has 1000 color images. The average RA for GLCM texture features was 0.33, but with DWT and GLCM, the average RA increased to 0.43, wherein in both cases texture feature were taken into consideration. So, it is proved that the combined effort always produces a good result. Again, while considering both color and texture feature average RA stands with 0.77.

6. Analysis-6 [11].

The techniques that are used for feature extraction are color, texture, intersecting cortical model (ICM), and K-means clustering method.

The performance parameter considered here is precession. It is a very effective feature extraction performance measure.

The various distance metrics parameters used in this approach are Euclidean, City block, and Canberra. Out of the distance metric, the Canberra distance generates very less distance retrieved images. On the other hand, the K-means method gives the best results in terms of similarity measurements.

7. Analysis-7 [12].

The CBIR techniques are using the image transform for feature extraction such as contourlet, ridgelet, and shearlet transforms, and for also classification of the methods such as Naïve Bayes, K-nearest neighbor (k-NN), and multi-class support vector machine (multi-SVM).

The metrics are commonly used to measure the quality of the retrieval process that is false positive, true positive, false negative, and true negative. These are used in a confusion matrix which is normally a table which holds the values of true positives, false positives, true negatives, and false negatives which can be used to represent the

set of test data for describing the performance of a classification model or classifier. Classifiers are compared using sensitivity, specificity, accuracy, error rate, Jacquard coefficient, F-measure.

The multi-SVM and Naïve Bayes classifiers fetch better results. When the sensitivity increases, the accuracy rate also increases and resulting in the multi-SVM classifier outperforming with 90.76% accuracy for ridgelet transform. The sensitivity, specificity, and accuracy of the multi-SVM are much higher for contourlet transform when compared to the others. The multi-SVM classifier gives better results when used with shearlet transform. The sensitivity, specificity, and accuracy of the multi-class SVM are much higher than others. The error rate produced by the multi-SVM classifier is also very much low and only 2.8%. The shearlet extracts more features and thus outperforms the other two classifiers because of its ability to handle and extract more features from the images than the ridgelet and contourlet.

8. Analysis-8 [13].

In the era of computation and the Internet, creating huge multimedia databases to retrieve these multimedia data in an efficient way is always a challenge. The CBIR system is a monumental achievement in this direction. The multimedia data is retrieved based on various features such as color, size, and texture. But in this study, an efficient index-based technique is developed for retrieval purposes.

For achieving state-of-the-art results in terms of accuracy and efficiency an attempt is made by this indexing CBIR system.

For the experimentation of the index-based CBIR system, medical images are indexed on MATLAB coding. Then, the query image and corresponding results in images are compared for the retrieval process.

Conclusion

The CBIRs allow us to search and query images efficiently from a large image database. From the papers we surveyed, we can ascertain that the performance of a CBIR system depends on the accuracy and the retrieval time in which it is able to produce the results. A lot of different schemes are employed to increase precision though most of them still do not hold effective in case of large databases primarily due to the increase in the search time. Thus, the careful study of various systems will help us in developing a robust and efficient CBIR system.

References

1. Long F, Zhang H, Feng DD (2003) Fundamentals of content-based image retrieval. In: Multimedia information retrieval and management. Springer, Berlin, Heidelberg, 1–26
2. Gudivada VN, Raghavan VV (1995) Content based image retrieval systems. *Computer* 28(9):18–22

3. Kim W-Y, Kim Y-S (2000) A region-based shape descriptor using Zernike moments. *Sig Process Image Commun* 16(1-2):95–102
4. Kumar G, Pradeep KB (2014) A detailed review of feature extraction in image processing systems. In: 2014 Fourth international conference on advanced computing & communication technologies. IEEE
5. Zhang H-J, Zhong S, Zhu X (2006) Relevance maximizing, iteration minimizing, relevance-feedback, content-based image retrieval (CBIR). U.S. Patent No. 7,113,944. 26 Sep. 2006
6. Mistry Y, Ingole DT, Ingole MD (2018). Content based image retrieval using hybrid features and various distance metric. *J Electri Syst Inf Technol* 5(3):874–888
7. Meshram SP, Thakare AD, Gudadhe S (2016) Hybrid swarm intelligence method for post clustering content based image retrieval. *Procedia Comput Sci* 79:509–515
8. Bala A, Kaur T (2016) Local texton XOR patterns: a new feature descriptor for content-based image retrieval. *Eng Sci Technol Int J* 19(1):101–112
9. Sutojo T et al (2017) CBIR for classification of cow types using GLCM and color features extraction. In: 2017 2nd international conferences on information technology, information systems and electrical engineering (ICITISEE). IEEE
10. Atlam HF, Attiya G, El-Fishawy N (2017) Integration of color and texture features in CBIR system. *Int J Comput Appl* 164(3):23–29
11. Bhadoria S et al (2012) Comparison of Color, Texture and ICM Features in CBIR System. In: *Advanced materials research*, vol. 403. Trans Tech Publications Ltd
12. David HBF, Balasubramanian R, Pandian AA (2018) CBIR using multi-resolution transform for brain tumour detection and stages identification. *Int J Biomed Biolog Eng* 10(11):543–553
13. Rahmani MKI, Ansari MA, Goel AK (2015) An efficient indexing algorithm for CBIR. In: 2015 IEEE international conference on computational intelligence & communication technology. IEEE, pp 73–77
14. Jena B, Nayak GK, Saxena S (2019) Maximum payload for digital image steganography obtained by mixed edge detection mechanism. In: 2019 international conference on information technology (ICIT). IEEE, pp 206–210
15. Arivazhagan S, Ganesan L, Bama S (2006) Fault segmentation in fabric images using Gabor wavelet transform. *Mach Vis Appl* 16(6):356
16. Nason GP, Silverman BW (1995) *The stationary wavelet transform and some statistical applications*. Wavelets and statistics. Springer, New York, NY, pp 281–299
17. Kannala J, Rahtu E (2012). Bsf: Binarized statistical image features. In: *Proceedings of the 21st international conference on pattern recognition (ICPR2012)*. IEEE, pp 1363–1366
18. Chatzichristofis SA, Yiannis SB (2008) CEDD: color and edge directivity descriptor: a compact descriptor for image indexing and retrieval. In: *International conference on computer vision systems*. Springer, Berlin, Heidelberg
19. Ouadfel S, Batouche M, Ahmed-Taleb A (2012) ACPISO: a novel swarm automatic clustering algorithm based image segmentation. In: *Multidisciplinary computational intelligence techniques: applications in business, engineering, and medicine*. IGI Global, pp 226–238
20. Cui X, Thomas EP, Paul P (2005) Document clustering using particle swarm optimization. In: *Proceedings 2005 IEEE swarm intelligence symposium, 2005*. SIS 2005. IEEE
21. Kanungo T et al (2002) An efficient k-means clustering algorithm: analysis and implementation. In: *IEEE Trans Pattern Anal Mach Intell* 24(7):881–892
22. Kao Y, Cheng K (2006) An ACO-based clustering algorithm. *International workshop on ant colony optimization and swarm intelligence*. Springer, Berlin, Heidelberg
23. Sebastian V, Unnikrishnan A, Balakrishnan K (2012) Gray level co-occurrence matrices: generalisation and some new features. *arXiv preprint [arXiv:1205.4831](https://arxiv.org/abs/1205.4831)*