A Review on various approaches used in Content Based Image Retrieval

Mr Sharad Yenkar Author[*1] , Dr A P Thakare Author[*2]

Department of Electronics & Telecommunication,

Sant Gadge Baba Amravati University, Amaravati,India

A Review on various approaches used in Content Based Image Retrieval

Abstract

One of the most exciting and dynamic areas of computer vision science was image recovery. Image databases are automatically scanned, downloaded, cataloged, and accessed by means of content-based image retrieval systems (CBIR). Important characteristics for content-based image recovery systems include color and texture qualities. For this reason, content-based image retrieval (CBIR) is a desirable source of precise and speedy retrieval. A multitude of methods have been developed recently to enhance CBIR performance. This essay addresses the advantages and disadvantages of CBIR as well as the reasons it is significant presently. In addition to applications, other approaches for feature extraction in CBIR are also covered.

Keywords: Image recovery, Image Processing, CBIR, Feature Extraction

I. Introduction

When it comes to image processing, database applications, and multimedia databases for image sorting, images are crucial. Stated differently, this knowledge area is commonly referred to as an image recovery if the returned information consists of an image set (Kaur and Sohi, 2016). This data is utilized as a space for image searches. Pictures are employed today for efficient operation in many different industries, including corporations, architecture, advertising, fashion, crime prevention, and historical study. The term "large collection of these images" refers to the image database.
One of the many database-dependent computer systems that are necessary for seeing and gathering photos is image retrieval. There are two methods for retrieving images: content-based image recall (CBIR) and text-based image retrieval (TBIR). The manual annotation process, which is challenging and expensive for large databases, is the disadvantage of TBIR. Scholars have demonstrated a greater interest in creating methods for image recovery that rely on autonomously generated features including color, form, and texture. This approach is commonly known as content-based image recovery, or CBIR (Kumar et al., 2015). The semantic distinction between low visual features and high image semantics is the main focus of CBIR research [Bansal et al., 2016]
Computer vision is used in huge databases to tackle picture retrieval problems, a process known as CBIR. "Content-based" refers to the search engine's assessment of the image's real content. In this context, "content" can refer to any {etails that can be taken out of the image itself, such as colors, shapes, textures, or other characteristics.

- Since color-based images are independent of image size or orientation, one of the most widely utilized techniques is to analyze them based solely on color.
- Shape: When we talk about shape, we're talking about the form of the area that we're looking for, not the shape of an image.
- Texture: Texture metrics scan images for observable patterns.

This method of image retrieval uses the form, color, and texture of a photo along with user-based query searches to find photos from huge databases. According to Ali and Sharma (2017), CBIR is an interface that bridges the semantic gaps between the low-level perception of a computer system and the high-level perception of a human brain. The computer system cannot perform complicated visual tasks at the same speed as the human brain. Visual image material in CBIR is represented as image features that are extracted by computationally expensive and widely dimensional feature extraction techniques, which seem to be domain-specific automatic feature extraction techniques. Thus, human intervention is not included (Wang et al., 2018).

II. Various Feature Extraction Techniques

The process of turning an image's raw pixel values into interpretable data that may be utilized with other methods, including machine learning, is known as feature extraction. When dealing with enormous image sizes, this method helps finish jobs like image matching quickly. There is a reduction in the object representation. This is the process for clearly identifying and extracting the most pertinent information from unprocessed data. The extraction feature in image processing creates derived values that are meant to be non-redundant and useful, starting with the initial set of measured data. The lowering of dimensionality has an impact on feature selection. An algorithm can be used to minimize the amount of data input when it becomes too big to handle.

In order to develop a classification of patterns, feature extraction is a crucial phase that seeks to extract the pertinent data that defines each class (Soora and Deshpande, 2017). To create feature vectors, pertinent features are taken out of objects or alphabets in this process. A discernible degree of the image's features are involved in feature extraction (Balan and Sunny, 2018). To extract the image's features that require testing, methods such as Average RGB, Color Moments, Co-occurrence, Local Color Histogram, Global Color Histogram, and Geometric Moments are employed. Conversely, feature matching entails matching the extracted characteristics to yield visually similar outcomes.

1) Color-based CBIR Features

The visual characteristic of an object that is produced by light that is emitted, transmitted, or reflected is called color. From a mathematical perspective, the color signal is an extension of the scalar signals of vector signals. An image with a histogram can be used to extract color properties. The disadvantage of the color histogram is that two of the same color may be identical to one another. A popular technique for representing color content is the color histogram (Latif et al., 2019).

A easy way to generate a histogram is to read each pixel in a picture only once and increase the corresponding bin of the histogram. This will count the number of pixels of each type. According to Putri et al. (2017), color histograms are largely invariant to translation, scale changes, modest off-axis rotations, translation of the image axis, and partial occlusion.

2) Spatial Features

An object's gray level, amplitude, and spatial distribution define its spatial properties. One of the object's most fundamental and significant characteristics is its amplitude. In X-ray images, amplitude reflects the absorption characteristic of body masses and enables bone differentiation from tissue.

3) Transform Features

Generally speaking, an image's transformation yields information about the data frequency domain. Zone filtering is used to obtain the properties of the picture alteration. Another name for the feature mask is an aperture or split feature mask. High-frequency components are commonly employed for edge and boundary detection. Angular slits can be utilized for orientation detection. Additionally, extraction from the Transform function is required when the input data originates within the Transform coordinate.

4) Shape Features

Shape characteristics, which are either based on border shape information or boundary plus inside content, are the main basis for shape representation. Different kinds of form features are created for object recognition, and they are assessed according to how well they enable the retrieval of comparable shapes from the database (Liu et al., 2020). The physical makeup and profile of the thing are referred to as its form. It is common practice to classify and match shapes, identify artifacts, or compute classes using spatial attributes. Other characteristics that are employed in the formal object extraction process include moment, perimeter, area, and orientation.

5) *Texture-based CBIR*

Texture is a repeated information pattern or structural structure with frequent intervals. Surface applies, in a general sense, to surface characteristics and appearance of an object by its elementary component size, form, density, arrangement, and proportion. As texture feature extraction a basic stage for collecting these features via a texture analysis process. Texture feature extraction is a key function in different image processing applications, such as remote sensing, medical imaging, and content-based image recovery, given the significance of texture information (Chaugule and Mali, 2014).

With texture segmentation, an image can be divided into multiple unbundled sections according to texture qualities, all of which are homogeneous with respect to other texture features. For the purpose of applying surface or scene rendering texture mapping, texture synthesis is a popular approach that turns normally small textures into large ones (Hoang, 2019). Using texture information, the shape from texture reconstructs the three-dimensional surface geometry. Texture extraction is an inevitable step in all these approaches.

6) Statistical-based Feature Extraction

Statistical techniques use the non-deterministic characteristics that control the relationships between an image's gray levels to indirectly quantify texture. By computing local features at each place in the image and generating a collection of statistics from the distributions of the local features, statistical techniques are utilized to examine the spatial distribution of gray values (Tahir and Fahiem, 2014).

The statistical methods can be classified into first order (one pixel), second order (pair of pixels) and higher order (three or more pixels) statistics. The first order statistics estimate properties (e.g. average and variance) of individual pixel values by waiving the spatial interaction between image pixels. The second order and higher order statistics estimate properties of two or more pixel values occurring at specific locations relative to each other. The most popular second order statistical features for texture analysis are derived 93 from the co-occurrence matrix.

These traits are easily distinguishable, in contrast to structural aspects. In contrast to other features, statistical features remain unaffected by distortions or noise. Many methods exist for extracting statistical characteristics, such as n-tuples, crossings and distances, zoning, and histogram estimations (Vithali and Kumbharana, 2015).

7) Structural-based Feature Extraction

Texture is represented by well-defined primitives and a hierarchy of those primitives' spatial arrangements in structural techniques. The basic definition is required for the texture description. A strong symbolic description of the image is one of the benefits of the structural technique based feature extraction; nonetheless, this property is more helpful for image creation than analysis jobs. Because micro- and macro-texture can vary greatly, this approach is not suitable for natural textures (Vithlani and Kumbharana, 2015).

The structural qualities serve as the foundation for both geometrical and topological features. Various global and local factors can be reflected by the geometrical and topological features that exhibit strong tolerance to distortions and style changes. Such a representation can also encode some information about the composition of the item or provide some information about its structure. According to Mohamed et al. (2015), the four categories can be divided into various topological and geometric representations:

a) Trees and Graphs

b) Coding

c) Topological Structure Extraction and Counting

d) Geometric feature estimation and approximation

8) Euler Number based Feature Extraction

The number of items in the area less the number of holes in the objects is the image's Euler number. The Euler numbers for the numbers "0," "4", "6," and "9" are 0 and 1, "2, "3,," "5,," and "7" are 1. One inherent feature of objects that can be used to characterize the shape topology is the Euler number. Two equations based on pixel geometry and connection qualities are proposed by the authors in (Azula et al., 2014). These equations can be used to efficiently determine the Euler number of a binary digital image with either thick or thin boundaries. The method used by the authors computes this feature, but it also retrieves the underlying topological information that the image's shape pixels supply. A theoretical proof is also provided for the accuracy of calculating the Euler number with the revised equations (Pal et al., 2018).

An integer known as the Euler number is used to classify data based on the quantity of holes in it. A character with no holes is assigned a positive number; a character with one hole is assigned a zero; and a character with two holes is assigned a negative number.

The system developed by Matsuoka et al. (2018) takes a handwritten text image as input, processes it through various stages, and outputs a text based on the features extracted from each character using Diagonal Feature Extraction and Euler Number classification using the Modified One-Pixel Width Character Segmentation Algorithm. To evaluate the system, a total of 100 handwritten text images are employed. Character recognition rate was 88.7838%, while word recognition rate was 50.4348% for the system.

9) Utilizing Geometrical Features for Feature Extraction

In (Mistry et al., 2017), the local characteristics (landmarks) of a set of emotion expressions (anger, happiness, sadness, and surprise) are extracted using a geometric-based features extraction operation from images of the BOSPHORUS database as a training stage. The classification operation is then completed by applying the threshold method (Euclidean distance) between the distances of the neutral image and the expression image. During the testing phase, the trained system is used to feature extraction and classification for stereoscopic 3D video movies. This technique is applied to 40 recorded 3D video clips, 10 videos for each of the four basic emotional expressions; the discrimination ratio is 85%.

(Raveena et al., 2017) discusses feature extraction and classification for Malayalam OCR (Neha et al., 2015) utilizing geometrical properties. The amount of features that are extracted has no bearing on the handwriting recognition system's accuracy. Only the features that we were able to extract are used. Language-focused characteristics known as geometrical features are able to categorize every character in a language in a unique way. This system makes use of a number of geometrical aspects. This system uses multiclass SVM and four step preprocessing.

III. Related Work

(Choudhary et al., 2014) presented an integrated content-based method for retrieving images that extracts texture and color information. while extracting color features from color images, the color moment (CM) is utilized, and while extracting texture features from grayscale images, the local binary pattern (LBP) is employed. The image's texture and color features are then integrated to create a single vector feature. Euclidian distance ultimately compares the feature vectors of the database photos with the query images to achieve similarity matching. Face recognition is the primary use of LBP.

(Zhu, 2014) proposed the Plane Semantic Ball (PSB), a new Adaptive Index Structure GPU designed to maximize parallel hardware acceleration while also minimising the function of the retrieval process. Semantics is integrated into the creation of representative pivots in PSB, and several balls are chosen to cover more specific reference features. PSB divides the online retrieval of CBIR into discrete parts that are efficiently executed on the GPU. The suggested method can attain high speed with minimal information loss, according to comparative studies conducted with the GPU-based brute force method. Furthermore, using two common picture datasets, Corel 10 K and GIST 1 M, PSB is contrasted with the cutting-edge random ball cover (RBC) method.

As demonstrated by (Kaur et al., 2016), CBIR combines an image's texture, color, and borders instead of keywords or other pertinent information. This work offered an organized review of the literature on a range of imaging modalities, covering accessible approaches and fundamental ideas along with research gaps. This research studies and applies retrieval strategies based on features including HSV, Color Moment, HSV and Color Moment, Gabor Wavelet and Wavelet Transform, and Edge Gradient. Sensitivity, Specificity, Retrieval Score, Error Rate, and Accuracy are only a few of the characteristics used in the quality assessment process for the image retrieval techniques that have been researched and the proposed technique. The performance assessment's experimental results show that the recommended approach works better than alternative methods.

(Meena et al., 2016) presented a method for recovering images linked to a query image from a vast collection of unique photographs. In order to extract the many features included in the image, it uses a method based on image segmentation. The aforementioned functions are equivalent, and the picture data is arranged in decreasing order of similarity in vectors known as the database image's feature vectors. Cloud computing handles the same. The CBIR framework is a program developed for the Windows Azure platform.

It is necessary to process a lot of images concurrently in order to identify them based on how similar they are to a user's query image. The method is implemented in several instances on Windows Azure virtual machines hosted in Microsoft data centers. The CBIR removes human dependency at the extraction step by representing visual picture material as automatically extracted image features that don't require personal involvement.

A retrieval model based on fusion was presented by Alhassan et al. (2017) as a means of combining texture and color characteristics from images using various fusion techniques. The results indicate that the CombMEAN fusion approach has the best and highest precision value and has outperformed both the individual color and texture retrieval model in both the top 10 and the top 20 images after the implementation of the proposed Wang image dataset retrieval model, which is widely used in CBIR.

CBIR relies on the image extraction feature, which is a visual characteristic that is automatically extracted—that is, without the need for human intervention. For the extraction feature, (Ali and Sharma, 2017) employed the SIFT feature extraction technique, which essentially provides us with an important spot in the image. In order to lower the complexity, cost, energy, and time consumption, we employ the BFOA (Bacteria foraging optimization algorithm) optimization technique. The SIFT image feature algorithm yields a set of unvalued imaging features. Then, in order to verify the similarity, a deep neural network is built, and the texting and validation stages are executed in accordance with the training, resulting in improved performance over earlier methods.

Using a range of distance measures, (Mistry et al., 2017) offer an efficient hybrid feature-based CBIR system. When using frequency domain functions like moments and spatial domain features like color auto programming and color moments in the HSV histogram, the Gabor wavelet transform is applied. Furthermore, an effective CBIR system is developed using color and edge directivity descriptor features to improve the accuracy of binarized statistical image features.

(Seth and Jindal, 2017) offer a retrieval strategy based on the image's combination of color, texture, and edge information. Sensitivity, Specificity, Retrieval Score, Error Rate, and Accuracy are only a few of the characteristics used in the quality assessment process for the image retrieval techniques that have been researched and the proposed technique. (Stubendek and Karacs, 2018) presents a novel method for enhancing the performance of automatic target recognition (ATR) by adaptively modifying the Gabor filter. The Gabor filter is a two-layer network structure in which the weights between the input and output layers form a linear classifier and the input layer represents an adaptive nonlinear extraction feature.

To facilitate image search, content-based image retrieval, or CBIR, takes features out of pictures. After reviewing some of the fundamental CBIR techniques for deriving color, texture, and shape features, (Chatterjee et al., 2019) demonstrates how image features can be derived from the compressed domain—specifically, from JPEG images—without requiring full picture decompression.

(Unar et al., 2019) proposed a critical method for CBIR that combines textual and visual features to produce identical images. The query image is first divided into textual and non-textual categories by the approach. A picture of a query is considered textual when it contains any text, and the text is recognized as a Text Word Bag. If the query image is classified as non-textual, the visual highlighting elements are retrieved and molded as a bag of visual words. Next, related images from the top are obtained based on the fused function vector, which is the result of fusing visual and textual information. Three recovery methods are supported: image query, keywords, and both.

A Gabor filter bank with smooth parameters, orientation, and Gaussian envelope frequencies determine the parameters of a Gabor filter-based feature extractor. The literature suggested a variety of parameters, and filter banks produced by these parameter settings typically perform admirably. However, filter banks constructed in this manner cannot be the best when it comes to pattern classification. A novel method for designing Gabor bank filters is suggested by (Liu et al., 2020) by incorporating feature selection into the design phase. The arrangement of the stream banks selects two streams. First off, the choice of filter lowers the computational complexity of texture extraction by forming a compact Gabor filter bank. Second, the Gabor filter bank was created with a higher sample-to-feature ratio in mind, in order to achieve low-dimensional representation.

IV. Conclusion

Pictures are the most noticeable component of all media files. An efficient and dependable approach for monitoring large databases is image retrieval. With the content-based image recovery approach, an image query can be addressed by the user to obtain images from the database in a manner akin to the image query. Although CBIR's current operational scope is relatively limited, it has great potential in a number of areas, such as image search for media, home entertainment, criminal detection, missing vehicle detection, and consumer products—finding any particular product by its color, shape, size, or features, for example.

The development of efficient and dependable recovery methods has been aided by the sharp increase in image storage capacities.As image compression, digital image processing, and image extraction technologies progress, CBIR's research continues to increase at a steady pace. Additionally, the advancement of CBIR technology has a significant impact on the creation of faster, less expensive memories as well as strong processing power. There are numerous possible CBIR applications with this architecture. Further research and development is being done on the efficiency and design of content-based picture recovery systems.

References

1. Alhassan, A. K., & Alfaki, A. A. (2017). Color and texture fusion-based method for content-based Image Retrieval. *2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*.

2. Ali, A., & Sharma, S. (2017). Content based image retrieval using feature extraction with machine learning. *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)*.

3. Azuela, H. S., J., Santiago, R., López, A., Peña Ayala, A., Cuevas Jimenez, E. V., & Rubio Espino, E. (2014). Alternative formulations to compute the binary shape Euler number. *IET Computer Vision*, 8(3), 171–181.

4. Balan, S., & Sunny, L. Z. (2018). Survey on Feature Extraction Techniques in Image Processing. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 6(3), 217-222.

5. Bansal A., Agarwal R., & Sharma R. K. (2016). Statistical feature extraction based iris recognition system. *Indian Academy of Sciences*, 41(5), 507-518.

6. Chatterjee, S., Dey, D., Munshi, S., & Gorai, S. (2019). Extraction of features from cross correlation in space and frequency domains for classification of skin lesions. *Biomedical Signal Processing and Control*, 53, 101581.

7. Chaugule, A., & Mali, S. N. (2014). Evaluation of Texture and Shape Features for Classification of Four Paddy Varieties. *Journal of Engineering*, 2014, 1–8.

8. Choudhary, R., Raina, N., Chaudhary, N., Chauhan, R., & Goudar, R. H. (2014). An integrated approach to Content Based Image Retrieval. *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*.

9. He, T., Liu, Y., Yu, Y., Zhao, Q., & Hu, Z. (2019). Application of Deep Convolutional Neural Network on Feature Extraction and Detection of Wood Defects. *Measurement*, 107357.

10. Hoang, N.-D. (2019). Automatic detection of asphalt pavement raveling using image texture based feature extraction and stochastic gradient descent logistic regression. *Automation in Construction*, 105, 102843.

11. Kaur, M., & Sohi, N. (2016). A novel technique for content based image retrieval using color, texture and edge features, *2016 International Conference on Communication and Electronics Systems (ICCES)*.

12. Kumar, K., Li, J.-P., & Zain-Ul-Abidin. (2015). Complementary feature extraction approach in CBIR. *2015 12th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP).*

13. Latif, A., Rasheed, A., Sajid, U., Ahmed, J., Ali, N., Ratyal, N. I., Khalil, T. (2019). Content-Based Image Retrieval and Feature Extraction: A Comprehensive Review. *Mathematical Problems in Engineering*, 2019, 1–21.

14. Liu, Z., Lai, Z., Ou, W., Zhang, K., & Zheng, R. (2020). Structured optimal graph based sparse feature extraction for semi-supervised learning. *Signal Processing*, 107456.

15. Matsuoka, Y. R., R. Sandoval, G. A., Q. Say, L. P., Y. Teng, J. S., & Acula, D. D. (2018). Enhanced Intelligent Character Recognition (ICR) Approach Using Diagonal Feature Extraction and Euler Number as Classifier with Modified One-Pixel Width Character Segmentation Algorithm. *2018 International Conference on Platform Technology and Service (PlatCon)*.

16. Meena, M., Singh, A. R., & Bharadi, V. A. (2016). Architecture for Software as a Service (SaaS) Model of CBIR on Hybrid Cloud of Microsoft Azure. *Procedia Computer Science*, 79, 569–578.

17. Mistry, Y., Ingole, D. T., & Ingole, M. D. (2017). Content based image retrieval using hybrid features and various distance metric.

Journal of Electrical Systems and Information Technology, 5(3), 874-888.

18. Mohamed A., Nasein D., Hassan H., & Haron H. (2015). A Review on Feature Extraction and Feature Selection for Handwritten Character Recognition. *(IJACSA) International Journal of Advanced Computer Science and Applications*, 6(2).

19. Neha J. Pithadia1, & Nimavat, V. D. (2015). A Review on Feature Extraction Techniques for Optical Character Recognition. *International Journal of Innovative Research in Computer and Communication Engineering*, 3(2), 1263-1268.

20. Pal, M., Bhattacharyya, S., & Sarkar, T. (2018). Euler number based feature extraction technique for gender discrimination from offline Hindi signature using SVM & BPNN classifier. *2018 Emerging Trends in Electronic Devices and Computational Techniques (EDCT)*.

21. Patel M. N., & Tandel P. (2016). A Survey on Feature Extraction Techniques for Shape based Object Recognition," International Journal of Computer Applications, 137(6), 16-22.

22. Putri, R. D., Prabawa, H. W., & Wihardi, Y. (2017). Color and texture features extraction on content-based image retrieval. *2017 3rd International Conference on Science in Information Technology (ICSITech)*.

23. Raveena, P. V., James, A., & Saravanan, C. (2017). Extended zone based handwritten Malayalam character recognition using structural features. 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT).

24. Schroder, J., Goetze, S., & Anemuller, J. (2015). Spectro-Temporal Gabor Filterbank Features for Acoustic Event Detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23 (2), 2198–2208.

25. Seth N., & Jindal S. (2017). A Review on content based image retrieval, *International Journal of Computers and Technology* 15(14).

26. Soora, N. R. & Deshpande, P. S. (2017). Review of Feature Extraction Techniques for Character Recognition. *IETE Journal of Research*, 68(3), 280-295.

27. Stubendek, A., & Karacs, K. (2018). Shape Recognition Based on Projected Edges and Global Statistical Features. *Mathematical Problems in Engineering*, 2018, 1–18.

28. Tahir M., & Fahiem M. A. (2014). A Statistical-Textural-Features Based Approach for Classification of Solid Drugs Using Surface Microscopic Images. *Computational and Mathematical Methods in Medicine*, 2014, 1-12.

29. Unar, S., Wang, X., Wang, C., & Wang, Y. (2019). A decisive content based image retrieval approach for feature fusion in visual and textual images. *Knowledge-Based Systems*.

30. Vithlani P., & Kumbharana C. K. (2015). Structural and Statistical Feature Extraction Methods for Character and Digit Recognition. *International Journal of Computer Applications*, 120(4), 43-47.

31. Wang, H., Li, Z., Li, Y., Gupta, B. B., & Choi, C. (2018). Visual saliency guided complex image retrieval. *Pattern Recognition Letters*.

32. Zhu, L. (2014). Accelerating Content-Based Image Retrieval via GPU-Adaptive Index Structure. *The Scientific World Journal*, 2014, 1–11.