A Comparative Analysis of Various Local Feature Descriptors in Content-Based Image Retrieval System

Lalit Kumar Tyagi¹, Rama Kant ² and Anish Gupta³

¹Department of Computer Science & Engineering, GL BAJAJ Institute of Technology and Management, Gr. Noida, India
²Department of Computer Science & Engineering, GL BAJAJ Group of Institutions, Mathura, India
³Department of Information Technology, ABES Engineering College, Ghaziabad, India
E-mail: gupta.anish1979@gmail.com

Abstract. Image acquisitions are increasing day by day due to progress in social networking and digital technologies. Nowadays, with the evolution of various image capturing devices, an enormous quantity of complex images is being produced. Content-primarily based image Retrieval (CBIR) is the answer to access images without difficulty wherein proper indexing and association are required. It makes CBIR a distinguished field in computer vision research. There are several uses of CBIR systems in day to day life for example medical, internet, scientific research and various other communication media. In the CBIR system, the user gives a query to obtain images from large datasets having a large number of images. In information transfer via electronic media using particular formats of data, images play an essential role. The extraction of information from communicated images is necessary with extra processing. In this paper, a comparative survey has been carried out on different content-based image retrieval implementation. These methods are determined by several authors for the feature extraction process of images and classification. This will help to plan the strategy for optimizing the CBIR system.

1. Introduction
Web browsing and viewing photos are being designed for people around the world because of the Internet of Things (IoT) [1], Cameras-embedded devices and the rapid growth of Internet technology. Many new image database-based implementations are yet to be investigated due to fast universal access to the internet and digital images. The aim of image search is to efficiently recover relevant visual documents from a huge-scale visual transcription to a linguistic or visual query. Recently, the multimedia and computer vision communities have shown their interest in image search white its exploration since 1990 [2]. In standard image search engines, surrounding web image meta data content, such as titles and tags, is used for indexing visual multimedia. Content-based image retrieval (CBIR) is given preference due to incompatibility of textual data with visual content. The text and image recovery systems based on content make CBIR a very active research field. The text-based image recovery system can be traced back to the late seventies, which many database management systems are using to retrieve images with manually annotated by text descriptors. These systems are obsoleted now-a-days, due to the subjectivity of human experience in annotation process[3]. In the early eighties, to
address the shortcomings of text-based retrieval systems, CBIR systems were introduced. Visual content \[4, 5\] is used to search images from massive image collections based on user requests for the query object and typically extracts any of the active colour, texture, shape, etc. attributes. For content-based object retrieval \[6, 7, 8\], various approaches have been developed to obtain image characteristics. Image processing with many types of objects is a critical and daunting issue in today’s technological world.

Studies also suggested methods of extraction of object features using features of local frequency. Several approaches are proposed to develop feature based on the interaction between the middle and adjacent pixels, based on local neighbourhood image pixels.\[9, 10, 11, 12\]. The nearest neighboring pixels play an important role in the local feature extraction. Using local pixel intensity values, the primary objective is to provide an effective local feature predictor that collects information based on the reciprocal relationship of neighboring pixels and integrates features based on the closest adjacent pixels to the centre pixel. The image characteristics are mainly divided into three major groups in the CBIR system: colour, form, and form \[13, 14\]. Hypothetically, in order to allow a clearer differentiation in the evaluation process of resemblance, certain characteristics need to be united. The overall CBIR device architecture is seen under Figure 1. With the consumer, the overall CBIR system starts. Next, it incorporates the input image as a submission into the process. Then, for the recovery of the relevant image, scan the image and all the images that are available in the archive in the same manner. Second, there may be a few pre-processing approaches to images, which also rely on a single purpose for improved recovery. An approach to image enhancement prior to computational processing is image preprocessing. If an algorithm’s input information is very broad and is considered to be redundant, it can be transformed into a reduced set of main features called the Function Vector seen in Figure 2. This is a method called Extraction of Functions. All pertinent information from the input data is stored in the derived functions, meaning that all the necessary tasks can be done using a simplified collection of features instead of the total function. Some of the basic features that \[15\] has retrieved are colour, shape, local descriptors, and type. When the query image joins the system as a whole, its feature space is compared to the \[16\] database function. The most applicable images are returned according to the measure of the gap. Some techniques that provide satisfaction measures called relevance feedback can be used to assess the degree of relevance of the extracted objects. We can clearly understand the overall operation of the CBIR method from Figure 1 through feedback on relevance. By upgrading the query and comparative calculations, this will increase the efficacy of the overall extraction process. In order to minimise user engagement and ignore multiple iteration filtering, computerised feedback and machine self-training are used.

As follows, the document is structured. The literature survey is listed in section II. The different CBIR methods used in literature are described in Section III and several CBIR methods are compared. The conclusion in which we explore the potential nature and utility of this survey is Section IV.

2. Literature Review
In the CBIR literature, several research and surveys have shown that the very challenging problems created by particular approaches are recognised and analysed. The Local Binary Pattern (LBP) is an important image descriptor used in numerous applications for computer vision such as facial recognition, object identification, target detection, image retrieval, etc. Due to its performance, various LBP variants have been developed to resolve its limitations. These methods generate the pattern from the predefined set of image pixels that limit the amount of information they capture. Srivastava et. al \[17\], introduced the CBIR technique of multi-scale Local Binary Pattern (LBP). The Local Binary Pattern is determined on various
The final function vector is formed by the Gray Level Co-occurrence Matrix (GLCM). The multi-scale LBP method overcomes the shortcomings of the single-scale LBP and acts as a more precise function descriptor. It efficiently catches certain textures’ large-scale dominant characteristics that are not accomplished by single-scale LBP. The semantic disparity between low-level image characteristics and high-level human comprehension, which reduces the accuracy of CBIR, is a significant problem.

A new Local Neighborhood Discrepancy Pattern (LNDP) image recovery approach for local features has been implemented by Verma et al. [18]. Centered on its neighbouring pixel relationship, the classical local binary pattern (LBP) transforms each image pixel into a binary pattern. But the descriptor of the LNDP function differs from the local binary pattern, since it transforms the reciprocal association of all neighbouring pixels into a binary pattern. All LBP and LNDP complement each other as they use local pixel intensity to obtain different information. Both LBP and LNDP features were merged in this work to obtain most of the information that can be obtained using variations in local frequency. A colour image retrieval system was introduced by Zhou et al. [19] using an appropriate combination of two colour and local directional pattern (LDP) histogram types, respectively. In addition, in order to increase retrieval performance, considering the importance of various feature elements, an effective standardisation of features and a new distance metric were proposed for extraction of features.
A spatial pattern descriptor known as Local Derivative Radial Pattern (LDRP) was implemented by Fadaei et. al [20] for content-based image retrieval representation. The square or circle variance of gray-level pixels is the origin of all previous local trends. Since the intensity relationship of pixels will constitute several of the individual textures along a line, these methods do not have the optimum capacity to reflect texture detail. LDRP is based on the gray-level variation of the pixels in a line and the weighted combinations. Multi-level coding, however, is used rather than binary coding in multiple directions. Yalavarthi et. al [21] implemented it using improved Gabor wavelet transformation to maximise retrieval effectiveness. In order to construct non-orthogonal functions, Gabor wavelet transformation (GWT) is generally based on combining plane wave and Gabor function features. The challenge property of images with GWT is categorised, multiple scaling and orientation with various filters, to reduce extra information.

Kumar et. al [22] implemented a contourlet tetra pattern that is a revised local pattern version that uses contourlet directions to extract the image tetra pattern. The difference between both the local tetra pattern (LTrP) and the contourlet pattern approach is that LTrP uses first-order spatial derivatives to extract directions, while this method uses contourlet transformation to find directions. Contourlet transformation is used to find directions based on the assumption that it helps to efficiently represent the images in multiple directional bands with more precise directional information than in spatial derivatives. In other work Kumar et. al [19] introduced a new methodology for selecting the image pixels used to extract the pattern based on the image characteristics. This method uses characteristics of the image line / curve to extract the local pattern that the author calls it as local curve pattern. The characteristics of line and curve are considered since they are the dominant components of an image and are used to efficiently represent the image.

Sharif et. al [23] proposed an approach based on combining visual terms and BRISK descriptors with a visual-bag-of-word approach based on scale-invariant function transformation (SIFT) that greatly improved the performance of CBIR. The suggested solution uses features of SIFT-BRISK to depict the visual quality of images in a more realistic context, due to the creation of two dictionaries. On the basis of speed-up robust features (SURF) and directed gradient histograms (HOG) feature descriptors, successful visual fusion methodology has been applied. HOG is used to extract global features, while in Mehmood et. al [24], SURF is used to extract local features. For large-scale image retrieval, global features are used, while local features work best with close visual presence on those systems that accept semantic queries. SURF is invariant in size and rotation and performs well with low lighting compared to the HOG descriptor. For applications based on scene-recognition-or activity-recognition, HOG, on the other hand, performs higher. The graphic mixture of SURF and HOG attribute descriptors is carried out in this article. A low level and shape-based image retrieval system has been developed by Jin et. al [25]. In this approach, the low-level image features are used to determine the salient area of the image as it has simple characteristics of computation and is easy to implement. And to quantify similarities between salient regions, the shape characteristics of the salient region are used as the shape function has very good invariability in translation, rotation and scaling, which is very suitable to define the boundary details of the zone.

In future, we can do formal verification to prove the correctness of methods implemented for CBIR as it’s distributed infrastructure for communication media [26][27]. We can also implement machine learning techniques for data analysis [28][29]. We also discussed CBIR strategies used by numerous writers and the application of them in this section. The benefits and drawbacks of the approaches are discussed in the next section.

3. Definition and Contrast of Methods for CBIR

This section provides the description of several CBIR methods in literature in Table 1 and Table 2. The comparison among CBIR methods is given in Table 3.
| S.NO | Technique Name | Authors | Description | Year |
|------|----------------|---------|-------------|------|
| 1    | CBIR with fusion of LDP and Color Histogram \[19\] | Ju-xiang Zhou, Xiaodong Liu, Tian-wei Xu, Jian-hou Gan, Wanquan Liu | A method of colour image recovery by using the fusion of two forms of histograms independently obtained from colour and LDP. | 2016 |
| 2    | CBIR using LDRP \[20\] | Sadegh Fadaei, Rassoul Amirfattahi, Mohammad Reza Ahmadzadeh | A local pattern descriptor is implemented for texture feature representation, which is referred to as LDRP. The gray-level pixels difference along a line is created together with their weighted combinations to calculate LDRP. In comparison, multi directed multi-level coding is implemented. | 2017 |
| 3    | CBIR based on shape similarity calculation \[25\] | Cong Jin, Shan-Wu Ke | To describe the image, image salient regions and shape representation are extracted which are used together with color features and texture information to perform image retrieval. | 2017 |
| 4    | CBIR using LCTP \[22\] | T.G.Subhash Kumar, V. Nagarajan | Contourlet transform is utilized to obtain the directions subject to the point that it aids to describe the images efficiently into various directional bands which will have more precise directional information than in the spatial derivatives. | 2017 |
### Table 2. (Continued....) Description of several CBIR methods

| S.NO | Technique Name | Authors | Description | Year |
|------|----------------|---------|-------------|------|
| 5    | CBIR using multiscale LBP [17] | Prashant Srivastava, Ashish Khare | Local binary patterns of separate sequences of eight neighbourhood pixels are measured on several scales instead of analysing related neighbourhood pixels. The final feature vector is generated by GLCM. | 2017 |
| 6    | CBIR using LBP and LNDP [18] | Manisha Verma, Balasubramanian Raman | LBP and LNDP features are fused for image extraction to retrieve salient information that is collected using variations in texture strength. Both the adjacent pixels’ reciprocal relationship is converted into a binary pattern. | 2017 |
| 7    | Scene analysis and search using local features and SVM for CBIR [23] | Uzma Sharif, Zahid Mehmood, Toqeer Mehmood, Muhammad Arshad Javid, Amjad Rehman, Tanzila Saba | A fusion technique focused on binary stable invariant scalable keypoints (BRISK) graphic terms and scale-invariant function transformation (SIFT). | 2018 |
| 8    | CBIR based on Visual Words Fusion Versus Features Fusion of Local and Global Features [24] | Zahid Mehmood, Fakhar Abbas et.al. | The fusion process of visual terms to fuse SURF (for local features) and HOG (for global features) characteristics. | 2018 |
| Scheme | Advantages | Disadvantages |
|--------|------------|---------------|
| CBIR using fusion of Color Histogram and LDP [19] | Efficient feature normalisation and a novel distance metric are implemented for the feature extraction to boost the efficiency of the retrieval. | The suggested work integrates the LDP function that extracts path information using a single colour channel. |
| CBIR using LDRP [20] | To extract significant data from the image, these patterns are good descriptors and suitable for representing texture features. | Calculation overhead in multi-level coding of micropatterns derivatives |
| CBIR based on shape similarity calculation [25] | Shape feature extraction technique is robust to scaling, translation and rotation. | Salient regions with similar contour affect the identification process which leads to failure of this system. |
| CBIR using LCTP [22] | Contourlet transform detects image discontinuities more efficiently and also gives more precise directional information as compared to the spatial derivatives. | This work is limited to greyscale images. |
| CBIR using multiscale LBP [17] | Single scale LBP was unable to capture large scale dominant features so this method eradicates this problem by capturing larger regions in some textures. | This method does not consider directional details for constructing feature vector. |
| CBIR using LBP and LNDP [18] | This took advantage of the mutual relationship between neighboring pixels and the relationship between center and neighboring pixels. | Computational overhead which leads to high Feature Extraction time and image retrieval time. |
| Scene analysis and search using local features and SVM for CBIR [23] | It performs better at poorly placed key points along the edges of an object inside an image. | It lacks spatial associations between the patches. |
4. Conclusions

Retrieval of content-based images over large image datasets is common among areas of communication. Important aspects of CBIR are reliability and output time. A comparative survey was conducted in this paper on certain techniques of extraction of features. There were few disadvantages to some listed strategies, such as loss of significant feature information, prolonged processing time, elevated computational complexity. Large data sets are found to always take longer to finish the search operation. That makes the system as a whole less reliable. Therefore, to achieve better results, an accurate, appropriate extraction of the image function is required. The cumulative manipulation of the database will be reduced as we reduce the problem associated with massive databases, and the result of an image retrieval process will be higher. Our research will then establish a new, improved strategy in the future to further refine the method of content-based image retrieval.

References

[1] Sethi P and Sarangi S 2017 Journal of Electrical and Computer Engineering 2017 1–25
[2] Zhou W, Li H and Tian Q 2017 arXiv preprint arXiv:1706.06646
[3] Sethi I K, Coman I L and Stan D 2001 Mining association rules between low-level image features and high-level concepts Data Mining and Knowledge Discovery: Theory, Tools, and Technology III vol 4384 (International Society for Optics and Photonics) pp 279–290
[4] Surya S and Sasikala G 2011 Indian J. Comput. Sci. Eng 2 691–696
[5] Singhai N and Shandilya S K 2010 International Journal of Computer Applications 4 22–26
[6] Liu Y, Zhang D, Lu G and Ma W Y 2007 Pattern recognition 40 262–282
[7] Müller H, Michoux N, Bandon D and Geissbuhler A 2004 International journal of medical informatics 73 1–23
[8] Smeulders A W, Worrning M, Santini S, Gupta A and Jain R 2000 IEEE Transactions on Pattern Analysis & Machine Intelligence 1349–1380
[9] Ojala T, Pietikäinen M and Harwood D 1996 Pattern recognition 29 51–59
[10] Ojala T, Pietikäinen M and Mäenpää T 2002 IEEE Transactions on Pattern Analysis & Machine Intelligence 971–987
[11] Tan X and Triggs W 2010 IEEE transactions on image processing 19 1635–1650
[12] Zhang B, Gao Y, Zhao S and Liu J 2009 IEEE transactions on image processing 19 533–544
[13] Smeulders A W, Worrning M, Santini S, Gupta A and Jain R 2000 IEEE Transactions on Pattern Analysis & Machine Intelligence 1349–1380
[14] Niblack C W, Barber R, Equitz W, Flickner M, Glassman H E, Petkovic D, Yanker P, Faloutsos C and TuBBin G 1993 QbiC project: querying images by content, using color, texture, and shape Storage and retrieval for image and video databases vol 1908 (International Society for Optics and Photonics) pp 173–187
[15] Yue J, Li Z, Liu L and Fu Z 2011 Mathematical and Computer Modelling 54 1121–1127
[16] Jenni K, Mandala S and Sunar M S 2015 Procedia Computer Science 50 374–379
[17] Srivastava P and Khare A 2018 Multimedia Tools and Applications 77 12377–12403
[18] Verma and Raman B 2018 Multimedia Tools and Applications 77 11843–11866
[19] Zhou J X, Liu X d, Xu T w, Gan J h and Liu W q 2018 International Journal of Machine Learning and Cybernetics 9 677–689
[20] Fadæs I, Amirfattahi R and Ahmadzadeh M R 2017 Signal Processing 137 274–286
[21] Yalavarthi A, Veeraswamy K and Sheela K A 2017 Content based image retrieval using enhanced gabor wavelet transform 2017 International Conference on Computer, Communications and Electronics (Comptelx) (IEEE) pp 339–343
[22] Kumar T S and Nagarajan V 2018 Signal, Image and Video Processing 12 591–598
[23] Sharif U, Mehmood Z, Mahmood T, Javid M A, Rehman A and Saba T 2019 Artificial Intelligence Review 52 901–925
[24] Mehmood Z, Abbas F, Mahmood T, Javid M A, Rehman A and Nawaz T 2018 Arabian Journal for Science and Engineering 43 7265–7284
[25] Jin C and Ke S W 2017 3D Research 8 23
[26] Gaur M and Kant R 2014 A survey on process algebraic stochastic modelling of large distributed systems for its performance analysis 2014 3rd International Conference on Eco-friendly Computing and Communication Systems pp 206–211
[27] Kant R and Gaur M International Journal of Computer Applications 975 8887
[28] Gupta A, Gupta M and Chaturvedi P 2019 Investing data with machine learning using python
[29] Gupta A, Lohani B and Kushwaha P 2014 International Journal of Advanced and Innovative Research Volume 3 201