Content based Image Retrieval System using Local Feature Extraction Techniques

Abhishek Madduri
Engineering Management, Duke University
Durham, North Carolina, USA

ABSTRACT
This article presents a Content Based Image Retrieval (CBIR) system based on SURF-ORB-AdaBoost methodology. Content-based image retrieval is a process that applies computer vision approaches for seeking and overseeing extensive image collections more efficiently. With the development of expansive digital image collections activated by fast advances in electronic capacity limit and processing power, there is a developing requirement for devices and computer systems to help productive browsing, searching, and retrieval for image collections. Hence, the aim of this article is to build up a content-based image retrieval system. In this article, the authors presented a combination of SURF and ORB features for image retrieval from a large image collection. Initially, SURF (Speeded Up Robust Features) and ORB (Oriented Fast Rotated and BRIEF) features are extracted from the given query image. Subsequently, K-Means clustering algorithm is used to analyze the data and the LPP dimensionality reduction method is used to reduce the space complexity of the system and increase the performance of the system. After that, classifier is applied to extract the relevant image. A precision rate of 97.8% has been reported using the proposed CBIR system for the Wang image dataset.

Keywords
CBIR; SURF; ORB; K-Means; LPP; BayesNet; Random Forest

1. INTRODUCTION
Content Based Image Retrieval (CBIR) is used to find the specific kind of image from a large image collection database. Searching, indexing, and browsing images among the expansive image databases are exceptionally disputable assignments. But, in case of CBIR systems the task of searching for desired images which are located at distributed locations is so easy using feature extraction methods. Image features are extracted in two domains either transformed domain and spatial domain. Transformed domain uses transformed data to extract features and spatial domain uses pixel values for extracting image features. Both spatial and transformed domains extract either global or local features. In the literature, various methods using global and local features have been presented for CBIR [1]. The image descriptor is one of the most important factors which affect the performance of a CBIR system in terms of accuracy and speed. The image descriptor extracts the features from an image based upon the contents in an image. Nowadays, CBIR systems are used for many applications of actual world image-based retrieval. For example, doctors may use CBIR process to retrieve comparative patient issues from a database, so that a better decision can be taken by the physician in the treatment of patients. The CBIR computation complexity is high because of the large database. A good number of techniques have been presented for retrieval of images based on their contents from a large database, but the major problem is features extracted from images and experimental outcomes are not capable for CBIR system. In this proposed system, the authors have considered two local image feature extraction methods, namely, SURF and ORB. These both feature extraction methods are used to extract image descriptor vector. After that, K-means clustering algorithm is used to analyze the data and the LPP dimensionality reduction method is used to reduce the space complexity of the system and increase the performance of the system.

2. RELATED WORK
A lot of feature extraction methods already exist in the literature for retrieval of content-based images. CBIR works on very different guidelines as compared to text-based image retrieval. In content-based image retrieval process, the properties of the query image are extracted and compared with the properties of the images stored in the large dataset. The most common features utilized are numerical measures of texture, shape, and color; thus, essentially all current CBIR frameworks, regardless of whether business or trial, works in this way. Some methods are designed to retrieve images with respect to global image features and while others retrieving images with respect to local image features. When the analysis is based upon frequency, global features are used for lower frequency and local features are used for higher frequency. Batur et al. [2] have used SIFT (Scale Invariant Feature Transform) feature extraction method for extracting image features and Euclidean distance method is used for measuring similarity distance between query image features and database image features. The main disadvantage of this method is that the author is not using any dimensionality reduction method, because the length of the SIFT feature vector is 128 dimensions which consume more memory space and take more time to retrieve images. Srivastava and Khare [3] have used a combination of two methods, namely, Wavelet and LBP have been used to find texture features at three orientations: horizontal, vertical, and diagonal. LBP lacks direction information so that is why the combination of LBP (Local Binary Patterns), and wavelet provides direction to extract shape from texture features. Similarity measure is done using a Euclidean distance method. Cedillo-Hernandez et al. [4] have used SURF feature extraction method to extract the contents of each video in the database. In this method, the computation cost of SURF depends upon the size of the image, which is 2 times faster than CIF. The feature extraction process is done in two stages if the content is video. Firstly, extracting the down sample luminance frame and in the second stage SURF descriptor is extracted. The feature extraction process did in one stage if the content is an image, the SURF descriptor is directly extracted from the query image. The similarity matching process is done using a Euclidean distance method. Kaur and Sohi [5] have used three feature extraction methods, namely, texture, color, and edges.

Firstly, HSV histogram is extracted from the image. Secondly, color moment is extracted from three planes red, green, and blue. Thirdly, Gabor wavelet method is used to calculate energy and amplitude. Fourthly, the wavelet moment method is used to calculate the mean coefficient and standard variation coefficient. At last, Sobel edge detection method is used to calculate the edge strength and gradient magnitude. This method shows 84.4% precision rate using Wang database. Yasmin et al. [6] have used image decomposition and EI (Edge and Inner) pixel classification methods. After that, clustering method is used to group the results based upon color. The decomposition process is done to control the orientation of the edges. In this method combination of color and shape features shows better results. Vinay et al. [7] have used ORB feature extraction method for extracting features from face images. After that, PCA (Principal Component Analysis) dimensionality reduction method is used to reduce the extracted dimensions by ORB method. Feature matching between the query image and the database image process is done by using FLANN (Fast Library for Approximate Nearest Neighbors).

3. DATASET

In this work, the authors have considered public dataset, namely, Wang image dataset consisting of 1000 images. This dataset consisting of 10-classes and each class contains 100 images; size of each image is 384×256, including African people, food, buildings, elephant, beach, horse, flower, mountain, bus, and dinosaurs [8]. A few samples of this dataset are shown below in Figure 1. Summary of dataset considered in this work is depicted in Table 1.

Table 1. Description of Wang image database

Data Set	Categories	Number of Images in Each Category	Total Number of Samples	Size of Each Image
Wang	10	100	1000	384×256

Figure 1. A few samples of Wang dataset

4. PROPOSED WORK

In the proposed work, the system is designed to have various components. The proposed CBIR system consists of various components, namely, query image input, feature extraction, testing data features, and similarity measure and retrieved images.

- *Query Image input*

In the proposed CBIR system, the authors have taken a public

database, namely, Wang image database and input the query image.

- *Feature Extraction*

Extracting of imperative features from an image is an important phase of a CBIR system. In this article, SURF-KMEAN-LPP and ORB-KMEAN-LPP methodologies are used for training the system. Initially, SURF and ORB methods have been used for extracting descriptor vector. After that, K-means clustering algorithm is applied to generate K=64 number of clusters using each descriptor vector from which mean of each cluster is generated. Now, after applying K-Means clustering algorithm, 64-dimensional feature vector have been extracted for both SURF and ORB feature extraction techniques. At last, LPP is used to reduce 64-dimensional feature vector into 16-dimensional feature vector.

- *Database Image Feature*

At this stage, the authors have combined both descriptor vectors SURF-KMEAN-LPP and ORB-KMEAN-LPP. Then, they stored both feature vectors in their own database.

- *Testing Data Features*

In this phase, the user will input the features of a testing dataset for retrieval of an image based on the contents.

- *Similarity Measure*

Classification is a decision-making phase, which is used to decide the probability of the testing data set related to training dataset. In classification phase, trained dataset contains all samples and class labels are assigned to each sample of training data. In this article, the authors considered three classifiers: BayesNet, Random Forest, and AdaBoost (Adaptive Boosting). Relationship between features of training dataset and testing dataset is predicted in this phase.

- *Retrieved Images*

Similar images are retrieved according to testing dataset features and returned to the user.

5. FEATURE EXTRACTION TECHNIQUES

Feature extraction is an important phase of proposed work, which is used to extract the imperative properties of the images for classification. Extracted features are used to represent the images for searching, browsing, and indexing of an image database. Feature extraction is not a big problem for image retrieval. But the significant phase of CBIR framework is to extract important information about the image. In this proposed work, authors have considered two feature extraction methods (SURF and ORB) with a combination of K-Means of LPP, i.e., SURF-KMEAN-LPP and ORB-KMEAN-LPP.

5.1 SURF (Speeded Up Robust Features)

SURF is a local feature extraction method. For extracting image feature key points, it utilized local invariant fast key-point detector and for extracting image feature descriptor, it utilized distinctive descriptor [4]. Processing of SURF method is fast and robust as compared with the SIFT feature extraction method. Firstly, it extracts the feature key-point of an image. After extracting key-point, it assigns orientation to the key-points, the orientation is assigned in circular motion with respect to the interested key-points. Then, the squared area is adjusted according to chosen orientation. At last,

feature descriptor is extracted using Haar wavelet responses. The length of the descriptor vector for one key-point in SURF is 64-dimensional feature vector.

5.2 ORB (oriented FAST rotated BRIEF)

ORB is also a local feature extraction method. For extraction, image feature key-points, it utilized a key-point detector, namely, FAST (Features from Accelerated Segment Test) and used BRIEF (Binary Robust Independent Elementary) for extracting image feature descriptor [7]. In ORB, the Harris corner detector method is used to find out best interested points from those which are detected by FAST key-point detector. For adding a direction to key-points, Orientation is applied to corners and the rotation is applied to binary test patterns using the direction of the patch. The length of the descriptor vector for one key-point in ORB is 32-dimensional feature vector.

5.3 K-MEANS Clustering

Clustering in image processing is the method used for gathering picture descriptors into commonly restrictive clustering with various semantics [6]. Clustering technique is utilized for assignment of the collection and arrangement of objects such that items in a similar gathering (called a group) are more as compared to each other than to those in different gatherings (groups) [9]. K-MEANS clustering is a part of unsupervised learning in which the authors have clustered n-dimensional descriptor vector into a K number of clusters and minimize the intra group difference.

5.4 LPP (Locality Preserving Projection) dimensionality reduction

LPP method is used to reduce the space and increase the performance of the system. LPP concentrates on the neighborhood connection between the information. Reduction means discarding unimportant parts and reduces information while losing little as far as information [10]. LPP works in three steps: firstly, it generates the adjacency graph. Secondly, using similarity matrix weight is assigned to the graph and finally Eigenvalues and Eigenvector is computed.

6. EXPERIMENTAL RESULTS

Public dataset namely, the Wang image dataset is used for experimental work in this article. This work implements the content-based image retrieval techniques using the OpenCV computer vision library and the python programming language as it is a basic and effective programming language that gives functionalities to control images in conjunction with OpenCV. Proposed system uses a fusion of two local feature extraction methods, namely, SURF-KMEAN-LPP and ORB-KMEAN-LPP. For SURF-KMEAN-LPP, Initially, authors have extracted SURF descriptor vector for each image. Then, K-Means algorithm is used for clustering of a descriptor vector into 64 number of clusters and generate mean value of each cluster. After that, the 64-dimensional vector is reduced into 16-dimensional vector using the LPP dimensionality reduction method. Similarly, again, the authors have extracted features using ORB-KMEAN-LPP. After generating 16 dimensional vectors for both methods, they apply hybridization technique on both methods. On the final stage, they have 32-dimensional feature vectors. The effectiveness of the proposed system is evaluated based on precision, RMSE (Root Mean Square Error), AUC and Execution time parameter. For measuring the execution of the proposed work, authors have considered three classifiers, namely, BayesNet, Random Forest and AdaBoost. Tables 2-5 shows the classifier wise precision rate, RMSE, AUC (Area Under Curve) and execution time, respectively, of SURF-KMEAN-LPP and ORB-KMEAN-LPP Features. Proposed work also compares with the state-of-the-artwork as presented in Table 6.

Table 2. Classifier wise precision rate of SURF-KMEAN-LPP and ORB-KMEAN-LPP features

Feature Extraction Method	Length of Feature Vector	BayesNet	Random Forest	AdaBoost (Random Forest)
SURF-KMEAN-LPP	16	76.6%	94.3%	94.4%
ORB-KMEAN-LPP	16	85.7%	98.1%	98.3%
SURF-KMEAN-LPP + ORB-KMEAN-LPP	32	97.4%	98.7%	97.8%

Table 3. Classifier wise RMSE of SURF-KMEAN-LPP and ORB-KMEAN-LPP features

Feature Extraction Method	Length of Feature Vector	BayesNet	Random Forest	AdaBoost (Random Forest)
SURF-KMEAN-LPP	16	18.46	13.79	13.76
ORB-KMEAN-LPP	16	14.07	8.92	8.88
SURF-KMEAN-LPP + ORB-KMEAN-LPP	32	6.42	12.76	12.06

Table 4. Classifier wise AUC of SURF-KMEAN-LPP and ORB-KMEAN-LPP features

Feature Extraction Method	Length of Feature Vector	BayesNet	Random Forest	AdaBoost (Random Forest)
SURF-KMEAN-LPP	16	97.4	99.8	99.7
ORB-KMEAN-LPP	16	99.0	99.9	99.9
SURF-KMEAN-LPP + ORB-KMEAN-LPP	32	99.9	99.9	99.9

Table 5. Classifier wise execution time (in ms) of SURF-KMEAN-LPP and ORB-KMEAN-LPP features

Feature Extraction Method	Length of Feature Vector	BayesNet	Random Forest	AdaBoost (Random Forest)
SURF-KMEAN-LPP	16	0.12	1.22	0.80
ORB-KMEAN-LPP	16	0.05	0.57	0.81
SURF-KMEAN-LPP + ORB-KMEAN-LPP	32	0.06	0.86	1.32

Table 6. Comparison of proposed work with state-of-the-artwork

Author	Feature Extraction Method	DataSet	Precision
Fadaei *et al.* [11]	Dominant colour descriptor (DCD), Wavelet and multi-scale transformation curvelet	Wang image dataset	76.5%
Namdeo and Jadhav [12]	Curvelet transformation	Wang image dataset	80.0%
Guoand Prasetyo [13]	Bit Pattern Features (BPF) and Color Co-occurrence Feature (CCF)	Wang image dataset	77.9%
Singha and Hemachandran [14]	Wavelet and Color histogram	Wang image dataset	76.2%
Srivastava and Khare [15]	SIFT and Geometric	Wang image dataset	39.8%
Shrivastava and Tyagi [16]	Color histogram, Gabor texture feature, Fourier Descriptor Based Shape Feature	Wang image dataset	76.9%
Youssef [17]	Dominant color Extraction and GLCM	Wang image dataset	81.1%
Proposed Work	SURF-ORB-AdaBoost	Wang image dataset	97.8%

7. CONCLUSION

In this article, the authors have proposed an efficient CBIR system which retrieves the relevant images from a large database for a given query image. Initially, a CBIR system extracts SURF and ORB features from the query image as well as from the database images. Subsequently, K-Means clustering algorithm is used to analyze the data and the LPP dimensionality reduction method is used to reduce the space complexity of the system and increase the performance of the system. After that, classifier is applied to extract the relevant image. The proposed CBIR system was evaluated using public Wang image dataset. The experimental results show that the proposed CBIR system retrieves 97.8% precision rate.

8. REFERENCES

[1] Celik C, and Bilge HS 2017 Content based image retrieval with sparse representation and local feature descriptor: A comparative study. *Pattern Recognition*, 68(2):1-13.

[2] Batur A, Tursun G, Mamut M, Yadikar N, and Ubul K 2017 Uyghur Printed Document Image Retrieval Based on SIFT Features. *International Congress of Information and Communication Technology* (ICICT), 107:737-742.

[3] Srivastava P, and Khare A 2017 Integration of Wavelet

Transform, Local Binary Patterns and Moments for Content-Based Image Retrieval. *Journal of Visual Communication and Image Representation*, 42:78-103.

[4] Cedillo-Hernandez M, Cedillo-Hernandez A, Nakano-Miyatake M, and Perez-Meana H 2014 Content Based Video Retrieval System for Mexican Culture Heritage based on Object Matching and Local-Global Descriptors. *Proceedings of the International Conference on Mechatronics, Electronics and Automotive Engineering*, 38-43.

[5] Atoum, Issa & Ayyagari, Maruthi Rohit. (2019). Effective Semantic Text Similarity Metric Using Normalized Root Mean Scaled Square Error. *Journal of Theoretical and Applied Information Technology.* 97. 3436-3447.

[6] Yasmin M, Sharif M, Irum I, and Mohsin S 2017 An Efficient Content Based Image Retrieval using EI Classification and Color Features. *Journal of Applied Research and Technology*, 877-885.

[7] Vinay A, Kumar CA, Shenoy GR, Murthy NKB, and Natarajan S 2015 *ORB-PCA* Based Feature Extraction Technique for Face Recognition. *Procedia Computer Science*, 58:614-621.

[8] http://wang.ist.psu.edu/docs/related/

[9] Ayyagari, M. R. (2019). Integrating Association Rules with Decision Trees in Object-Relational Databases. *arXiv preprint arXiv:1904.09654*.

[10] Zhuo L, Cheng B, and Zhang J 2014 A Comparative Study of Dimensionality Reduction Methods for Large-scale Image Retrieval, *Neurocomputing*, 141:202-210.

[11] Fadaei S, Amirfattahi R and Ahmadzadeh MR 2017 A New Content-Based Image Retrieval System Based on Optimized Integration of DCD, Wavelet and Curvelet Features. *IET Image Processing*, 11(2):89-98.

[12] Namdeo HD, and Jadhav PD 2015 Content Based Image Retrieval Using Color and Texture. *International Journal of Electronics, Communication & Instrumentation Engineering*, 5(3):81-88.

[13] Guo JM, and Prasetyo H 2015 Content-Based Image Retrieval Using Features Extracted from Halftoning-Based Block Truncation Coding. *IEEE Transactions on Image Processing*, 24(3):1010-1024.

[14] Ayyagari, Maruthi Rohit. (2019). Efficient Driving Forces to CMMI Development using Dynamic Capabilities. *International Journal of Computer Applications.* 178. 24-29. 10.5120/ijca2019919024.

[15] Srivastava P, and Khare A 2016 Content-Based Image Retrieval using Scale Invariant Feature Transform and Moments. *Proceedings of IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics Engineering* (*UPCON),* 162-166.

[16] Ayyagari, M. R. (2019). Cache Contention on Multicore Systems: An Ontology-based Approach. *arXiv preprint arXiv:1906.00834*.

[17] Youssef SM 2012 ICTEDCT-CBIR: Integrating Curvelet Transform with Enhanced Dominant Colors Extraction and Texture Analysis for Efficient Content Based Image Retrieval. *Computer and Electrical Engineering*, 38:1358-137.