



Content-based image retrieval system using ORB and SIFT features

Payal Chhabra¹ · Naresh Kumar Garg¹ · Munish Kumar²

Received: 4 October 2017 / Accepted: 9 August 2018
© The Natural Computing Applications Forum 2018

Abstract

Measures of components in digital images are expanded and to locate a specific image in the light of substance from a huge database is sometimes troublesome. In this paper, a content-based image retrieval (CBIR) system has been proposed to extract a feature vector from an image and to effectively retrieve content-based images. In this work, two types of image feature descriptor extraction methods, namely Oriented Fast and Rotated BRIEF (ORB) and scale-invariant feature transform (SIFT) are considered. ORB detector uses a fast key points and descriptor use a BRIEF descriptor. SIFT be used for analysis of images based on various orientation and scale. *K*-means clustering algorithm is used over both descriptors from which the mean of every cluster is obtained. Locality-preserving projection dimensionality reduction algorithm is used to reduce the dimensions of an image feature vector. At the time of retrieval, the image feature vectors are stored in the image database and matched with testing data feature vector for CBIR. The execution of the proposed work is assessed by utilizing a decision tree, random forest, and MLP classifiers. Two, public databases, namely Wang database and corel database, have been considered for the experimentation work. Combination of ORB and SIFT feature vectors are tested for images in Wang database and corel database which accomplishes a highest precision rate of 99.53% and 86.20% for coral database and Wang database, respectively.

Keywords CBIR · ORB · SIFT · *K*-means · LPP

1 Introduction

In the development of Internet and multimedia, expansive measure of images is produced and dispersed yet to share and store such a lot of information effectively is a noteworthy issue. The most effective method to retrieve desired information from a bulky database is as yet a major issue. So, content-based image retrieval (CBIR) techniques are used to retrieve the images from bulky database based on

the desired information. In CBIR, the stored database and query image features are extracted and compared with each other, from which most comparable outcomes are come back to the client. In the present days, many of the real-world image-based retrieval applications used CBIR strategies. For example, physician may use CBIR strategy to retrieve comparative patient issues from a database, so that a best decision can be taken by the physician in the treatment of patients. The CBIR computation complexity is high because of the huge measure of database. Numerous strategies have been proposed in the past decades for retrieval of images based on their contents from bulky database, but issues for extracting features and results of the proposed systems are not appropriately satisfactory. To dump these issues, we present another technique using SIFT and ORB feature extraction techniques. We utilized two component extraction techniques in our examination work in light of the fact that a solitary element extraction strategy does not demonstrate all elements in an image, so the combination of extraction techniques preferred over a solitary one. The SIFT descriptor utilizes a 128-element feature dimensions in one key point and ORB descriptor

✉ Munish Kumar
munishcse@gmail.com

Payal Chhabra
payal.chhabra09@gmail.com

Naresh Kumar Garg
naresh2834@rediffmail.com

¹ Department of Computer Science and Engineering, GZS Campus College of Engineering and Technology, Maharaja Ranjit Singh Punjab Technical University, Bathinda, Punjab, India

² Department of Computational Sciences, Maharaja Ranjit Singh Punjab Technical University, Bathinda, Punjab, India

uses 32 elements in one key point, which require a high memory space for storing features and high complexity, so to reduce the space and complexity problem our system uses a *K*-means clustering algorithm and locality-preserving projection (LPP) over both descriptors. This paper is divided into five sections. In Sect. 1, introduction of present work has been discussed. Related work is exhibited in Sect. 2. In Sect. 3, we have presented a proposed framework for content-based image retrieval system. Section 4 presents experimental results in view of the proposed framework. At long last, conclusion and future extension are presented in Sect. 5.

2 Related work

Soltanshahi et al. [6] have used clustered scale-invariant feature transforms for CBIR. They considered SIFT feature extraction technique and use two approaches for SIFT descriptor. In the principal approach, they cluster a SIFT descriptor into 16 clusters, and in the second approach, key focuses are registered in view of direction. Alhassan and Alfaki [1] have introduced a color and texture fusion-based technique for CBIR. Peizhong et al. [5] have extracted texture and color-based feature elements. In texture features, local binary pattern (LBP) strategy is used, and in color features, Colour Information Feature (CIF) technique is used. LBP is used for rotation-invariant and multitermination imperatives, but alone LBP does not function admirably when similitude measure depends on color so the CIF method is used for color features. Kaur and Sohi [4] have exhibited a novel method for content-based image retrieval using HSV, Gabor wavelet, color moment, edge gradient, and wavelet transform feature extraction techniques. They have considered four similarity measure methods, i.e., Chi-square, Euclidean, Manhattan, and Chebyshev removal measure for arrangement reason. Fadaei et al. [2] have used dominant color descriptor (DCD), wavelet and curvelet features for giving the novel structure to content-based image retrieval system. DCD approach utilizes an HSV shading area to concentrate highlights. Wavelet texture-based feature utilizes global wavelet decomposition, and curvelet method was used for a multiscale transformation. This paper utilizes curvelet and wavelet features to remove the noise and translation problems. They have additionally explored particle swarm optimization (PSO) algorithm for fusion of all three features. Giveki et al. [3] have presented histograms of gradient (HOG), local binary pattern (LBP), local ternary pattern (LTP), local derivative pattern (LDP), and SIFT highlight extraction procedures, and lastly they registered outcomes and saw that the combination of SIFT and LDP give back a highest precision rate.

3 Proposed CBIR system

Proposed CBIR system consists of various phases, namely database collection, feature extraction, classifier prediction, and retrieved images. Block diagram of the proposed framework is depicted in Fig. 1. In database collection phase, we have considered two public datasets, namely Wang and corel image datasets. In feature extraction phase, we have considered SIFT and ORB feature extraction methodology for extracting image feature descriptor. After that, we have applied *K*-means clustering algorithm to generate *K* number of clusters using the descriptor array and LPP is also used to reduce dimensions of the feature vector.

3.1 Phases of proposed CBIR methodology

- Query Image input.

In this phase someone have to input a query image.

- Feature Extraction

The most important part of a CBIR system is to extract the meaningful features of an image. In present work, authors have considered two techniques for feature extraction namely, SIFT and ORB, for extracting descriptors of an image. After that, *K*-means clustering algorithm has been applied to generate *K* number of clusters using the descriptor vector and LPP is used to reduce dimensions of the feature vector.

- Database Image Feature

Store the SIFT and ORB feature vectors in training database.

- Testing Data Features

At this stage user enters a testing dataset for content-based image retrieval.

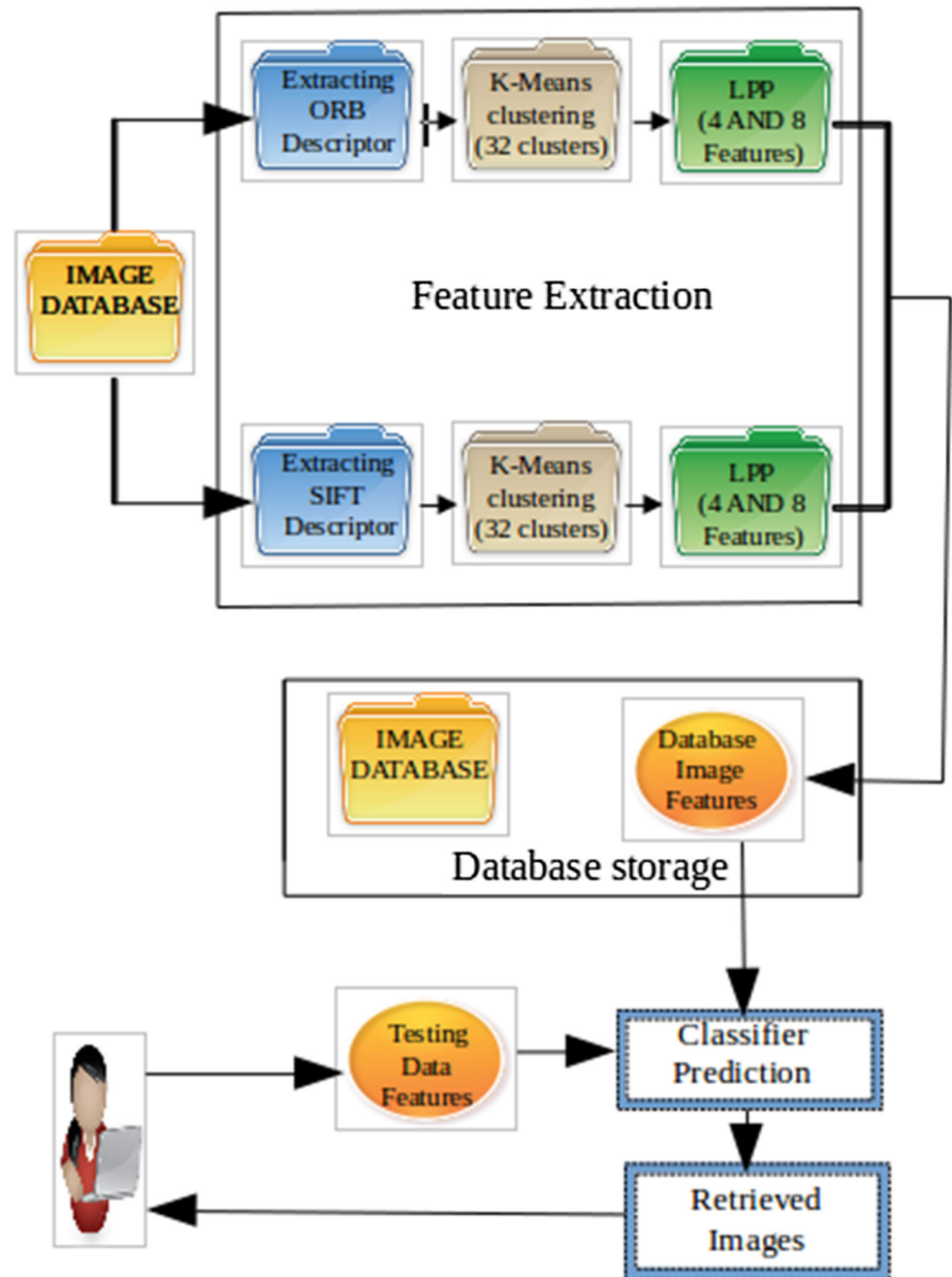
- Classifier Prediction

Classification is used to decide the probability of a one set of information belongs to another set. Classification uses a training dataset, in which class labels are already assigned for all objects. In present work, authors have considered three classification techniques, namely decision tree, MLP, and random forest.

- Retrieved Images

The most similar predicted results are returned to the user. Similar images are retrieved according to testing data features.

Fig. 1 Block diagram of proposed CBIR system



3.2 Algorithm

Input Images in Wang and corel image dataset and query image.

Step 1 Extracting a feature descriptor for each image from a dataset using ORB and SIFT feature extracting methods.

Step 2 K-mean clustering algorithm is used on feature descriptor of every image. Generate 32 clusters for every descriptor array from which the mean of every cluster is computed and finally we get a 32-dimensional feature vector for every descriptor.

Step 3 LPP dimensionality reduction algorithm is used to reduce the feature vectors into 4 and 8 components.

Step 4 Integrate the both feature vectors ORB and SIFT and stored the combined feature vector into database storage.

- Step 5** Train the proposed system using a combination of ORB and SIFT feature vectors.
- Step 6** Input the query image and extract the ORB and SIFT features of a questioned image.
- Step 7** The trained classifier used to predict the similarity between testing data features and the whole database feature vector, classifier predict the class of testing data features.
- Step 8** Retrieve the most similar images in class as output.

3.3 Techniques

ORB and SIFT methodologies are considered for feature extractions. *K*-means applied for clustering and LPP is used for reducing the dimensionality of the feature vector.

Feature extraction methods	ORB and SIFT
Number of cluster used in <i>K</i> -means	32
LPP reduced dimensions	4 and 8

3.4 Feature extraction techniques

Feature extraction is the most imperative part of in the proposed work. In this phase, we have to extract meaningful feature elements for image retrieval from bulky database. In present paper, we have considered two feature extraction techniques, namely, ORB and SIFT.

3.4.1 ORB (Oriented FAST Rotated and BRIEF)

ORB is developed in “openCV labs” which uses Features from Accelerated Segment Test (FAST) key point detector and binary BRIEF descriptor. ORB is used to extract the fewer but the best features from an image. The cost of computation is also less when contrasted with SIFT and SURF, but the magnitude is faster than SURF. Few samples of original database images are depicted in Fig. 2. In

Fig. 3, we have shown features extracted using ORB key point.

ORB firstly apply the FAST key point detector, which detects the large number of key points and then ORB uses a Harris corner detector to find good features from those key points. Extracted good features generate better results and less sensitive to noise. Using the patch moment, in ORB we can find the centroid of image, as following:

$$m_{pq} = \sum_{x,y} x^p y^q (x, y)$$

Using the intensity centroid of image patch, we can find out an orientation of corners:

$$C\left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right)$$

From the center of patch to centroid the angle is given by:

$$\text{atan2}\left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right) = \text{atan2}(m_{01}, m_{10})$$

ORB uses *OFAST* for the quick key focuses on direction and *RBRIEF* for the BRIEF descriptor with orientation (rotation) angle. The sample used in BRIEF descriptor is transformed using the orientation. Each sample point takes 5×5 windows of pixels and points are taken from 31×31 sample space. One window acquires 32-element ORB descriptor vector.

3.4.2 SIFT (scale-invariant feature transform)

The SIFT detector extracts number of attributes from an image in such way which is reliable with changes in the lighting impacts and perspectives alongside other imagining viewpoints. The SIFT descriptor will distinguish nearby elements of an image. In Fig. 4, we have shown features extracted using SIFT.

The major stages in SIFT are:

- Detector
 - Find Scale Space Extreme Detection
 - Key point Localization and filtering
- Descriptor



Fig. 2 Few samples of original database

Fig. 3 Features extracted using ORB key points



Fig. 4 Features extracted using SIFT key points



Orientation Assignment
Feature Description

- Similarity

Feature Matching

3.5 K-means

In *K*-means clustering algorithm the Euclidean distance method or max min measure is used to calculate a distance between centroid of cluster and object. We have a set of n descriptor array and *K*-means clustering algorithm is used to cluster n descriptor array into K number of clusters and minimize the intra cluster variance as described in following equation.

$$arg_s = \min \sum_{i=1}^k \sum_{x_j \in S_i} X_j - \mu_i$$

In present work, we are using $K = 32$, clusters $S_i = 1, 2 \dots K$ and μ_i is center of all points X_j in S_i .

- Step 1* From a set of n descriptor array randomly select 32 instances.
- Step 2* Assign every instance to its closest center point.
- Step 3* Update a cluster center.
- Step 4* After assigning all instances and finally 32 clusters are formed each cluster with different no of instances according to closest center points.
- Step 5* Calculate a mean of each cluster and as a final step we have a 32 means for 32 clusters.

3.6 LPP (locality-preserving projection)

Locality-preserving projection (LPP) is utilized to diminish the higher-dimensional space into lower-dimensional space for information storage. LPP takes a shot at neighborhood

data about the informational index and in LPP, any information which have a place with a similar class. LPP means keeping the desired information and removing the undesired data. PCA keeps less information about the data than LPP. LPP for dimensionality diminishment works in three stages as following:

- 1 Adjacency graph for data is generated using space $x_i \in R^d$ and nodes for undirected graph $i = 1 \dots N$.
- 2 Weights are selected for graphs which are represented using similarity matrix $[p_{ij}]$.

$$p_{ij} = \begin{cases} e^{\| -x_i - x_j^2 / t \|}, & \text{if } x_i \in kNN(x_j) \text{ or } x_j \in kNN(x_i) \\ 0, & \text{otherwise} \end{cases}$$

- 3 Eigenvector and eigen value equation are computed:

$$XLX^T \sigma = \lambda XD X^T \sigma$$

where $\langle X = x_1 | \dots | x_n \rangle$, D is a matrix with sum of P . Using row or column diagonal terms, $d_{ii} = \sum_j p_{ij}$, L is Laplacian matrix $L = D - P$, σ projection matrix.

4 Dataset and experimental results

In this section, we have presented brief information about datasets and experimental results performed using the proposed framework for CBIR.

4.1 Dataset

In this proposed system, we used two image datasets one is Wang database and another one is corel database. Wang database comprises of 1000 images for 10 classes, including African people, food, buildings, elephant, beach,



Fig. 5 10-Class Wang dataset

horse, flower, mountain, bus and dinosaurs. In Fig. 5, we have shown a few samples of Wang database. Corel database comprises of 6410 images for 40 classes including car, dog, door, eagle, elephant, flower, indoor, mountain, train, woman, cat and many more. Few specimens of corel database appeared in Fig. 6. Brief information about these databases is presented in Table 1.

4.2 Experimental results

All images of dataset are tested using ORB and SIFT features. Each image will produce different array length due to different sizes of images, so we have applied K -means clustering algorithm, where $K = 32$ to uniform the dimensions of the feature set. After that, LPP

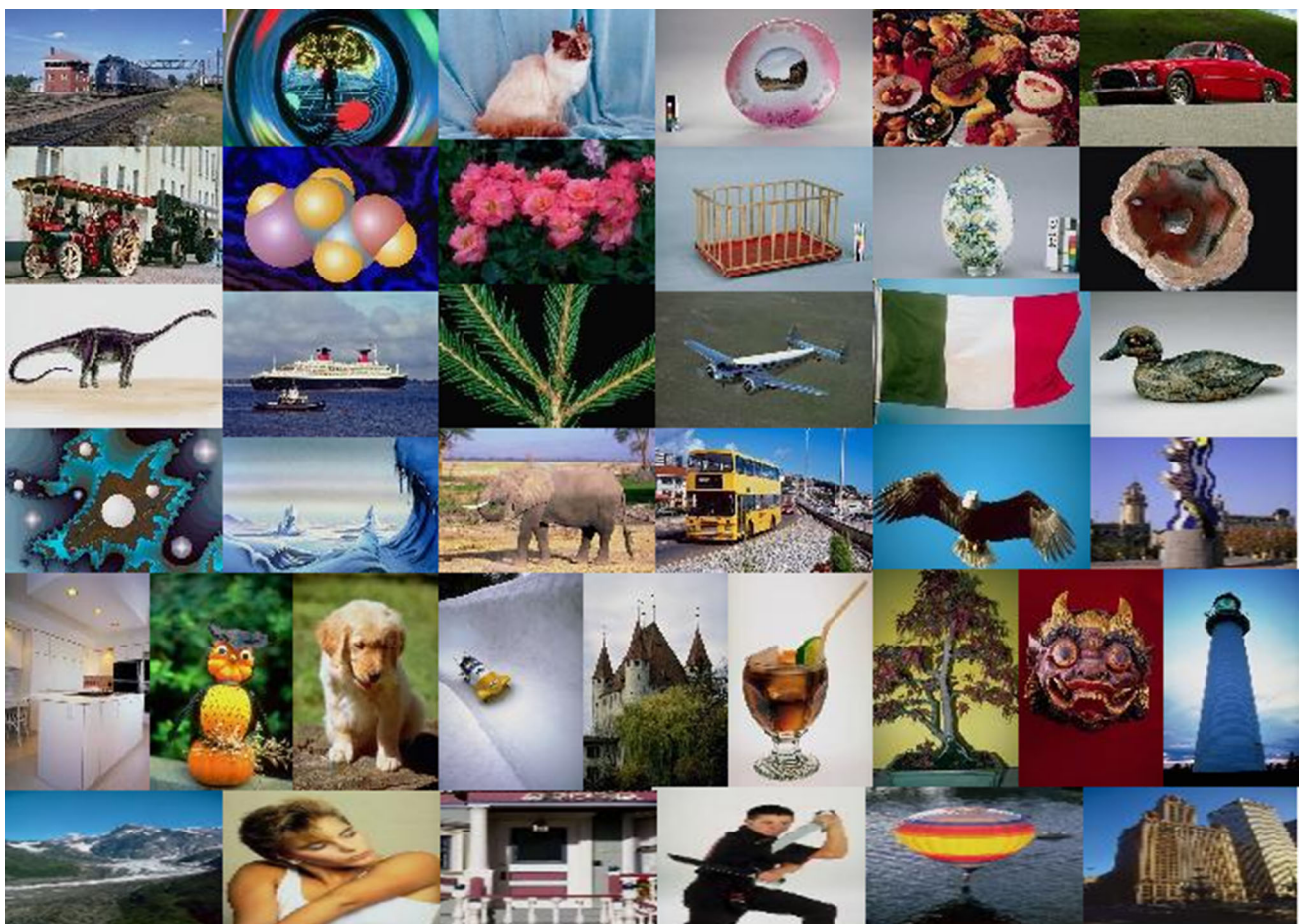


Fig. 6 Corel dataset

Table 1 Brief information about datasets

Database	Size of image	Number of classes	Total number of images
Wang database	384×256	10	1000
Corel image database	120×80 or 80×120	40	6410

dimensionality reduction method is used to reduce the dimensions into 4- and 8-dimensional feature vectors. We have performed this complete process for ORB and SIFT features individually and finally, combination of ORB and SIFT features was used for improving the performance of the proposed system. In the classification phase, we have used considered three different classifiers, namely decision tree, random forest, and multilayer perceptrons (MLPs). The performance of the proposed system is measured based on three parameters: precision rate, RMSE, and training time.

4.2.1 Precision

Performance of the proposed framework for CBIR is measured using precision. Precision is computed as number of relevant retrieved images divided by the total number of images retrieved (relevant and non-relevant). In Tables 2, 3, we have presented classifier wise precision rates achieved using the proposed framework with ORB and SIFT features.

4.2.2 RMSE

Root-mean-square error (RMSE) is used to measure the error between training and testing data by comparing the expected value with the observed value. RMSE for Wang dataset and corel dataset is presented in Tables 4, 5, respectively.

4.2.3 Training time

Time taken for retrieving the data from bulky dataset by using a decision tree, random forest and MLP classifiers is also considered. In Tables 6 and 7, we have presented time taken for Wang dataset and for corel dataset, respectively.

5 Conclusion

In this paper, a new technique, namely ORB, has been proposed for content-based image retrieval. ORB and SIFT features are considered for effectively retrieval of content-based images from bulky dataset. Size of SIFT and ORB descriptor requires a high memory space for storing features and high complexity, therefore, to reduce the space

Table 2 Precision rate for Wang dataset

Features 10-class dataset	No of feature vector dimensions	Decision tree (%)	Random forest (%)	MLP (%)
ORB	4	57.70	59.30	52.00
	8	56.40	67.00	32.50
SIFT	4	57.80	59.20	50.30
	8	54.50	61.70	63.20
ORB + SIFT	4 + 4	85.50	86.20	75.50
	8 + 8	84.20	84.40	34.33

Table 3 Precision rate for corel dataset

Features 40-class dataset	No of feature vector dimensions	Decision tree (%)	Random forest (%)	MLP (%)
ORB	4	72.80	77.32	68.43
	8	71.71	77.53	66.56
SIFT	4	93.39	97.81	75.45
	8	93.49	99.22	77.11
ORB + SIFT	4 + 4	94.74	99.27	92.77
	8 + 8	94.74	99.53	93.18

Table 4 RMSE for Wang dataset

Features 10-class dataset	No of Feature vector dimensions	Decision tree (%)	Random forest (%)	MLP (%)
ORB	4	24.74	21.83	23.92
	8	26.95	21.63	23.18
SIFT	4	25.47	22.30	25.44
	8	27.29	21.63	22.37
ORB + SIFT	4 + 4	16.51	14.18	17.24
	8 + 8	17.28	15.67	30.76

Table 5 RMSE for corel dataset

Features 40-class dataset	No of Feature vector dimensions	Decision tree (%)	Random forest (%)	MLP (%)
ORB	4	10.32	8.82	10.16
	8	10.89	8.78	10.59
SIFT	4	5.52	3.28	10.04
	8	5.55	2.74	9.69
ORB + SIFT	4 + 4	4.92	2.77	5.37
	8 + 8	5.00	3.34	5.46

Table 6 Time taken (in s) for Wang dataset

Features 10-class dataset	No of feature vector dimensions	Decision tree	Random forest	MLP
ORB	4	0.14	0.55	1.94
	8	0.05	0.32	2.36
SIFT	4	0.06	0.41	1.81
	8	0.05	0.50	2.4
ORB + SIFT	4 + 4	0.03	0.38	2.5
	8 + 8	0.05	0.50	3.77

Table 7 Time taken (in s) for corel dataset

Features 10-class dataset	No of feature vector dimensions	Decision tree	Random forest	MLP
ORB	4	0.19	4.74	77.92
	8	0.30	6.54	92.64
SIFT	4	0.13	3.42	79.86
	8	0.20	4.98	92.14
ORB + SIFT	4 + 4	0.17	4.28	91.98
	8 + 8	0.27	6.36	123.13

and complexity problem our system uses a K -means clustering algorithm and LPP over both descriptors. K -means reduce the descriptor into 32 clusters and LPP reduce into 4 and 8 components. Using 4- and 8-dimensional feature vector, we measure the precision, RMSE and time taken by the proposed CBIR system. Maximum precision rate of 86.20% and 99.53% has been accomplished for Wang dataset and corel dataset, respectively. We have also concluded that the proposed CBIR system performs better than already existing CBIR systems.

Compliance with ethical standards

Conflict of interest Authors have no conflict of interest in this work.

References

1. Alhassan AK, Alfaki AA (2017) Color and texture fusion-based method for content-based image retrieval. In: Proceedings of the international conference on communication, control, computing and electronics engineering (ICCCCEE), Khartoum, Sudan
2. Fadaei S, Amirfattahi R, Ahmadzadeh MR (2017) A new content-based image retrieval system based on optimized integration of

- DCD, wavelet and curvelet features. *IET Image Process* 11(2):89–98
3. Giveki D, Soltanshahi MA, Montazer GA (2017) A new image feature descriptor for content based image retrieval using scale invariant feature transform and local derivative pattern. *Opt Int J Light Electron Opt* 131:242–254
4. Kaur M, Sohi N (2016) A novel technique for content based image retrieval using color, texture and edge feature. In: *Proceedings of the international conference on communication and electronics systems (ICCES)*
5. Peizhong L, Guo J, Chamnongthai K, Prasetyo H (2017) Fusion of color histogram and LBP-based features for texture image retrieval and classification. *Inf Sci* 390:95–111
6. Soltanshahi MA, Montazer GA, Giveki D (2015) Content based image retrieval system using clustered scale invariant feature transforms. *Opt Int J Light Electron Opt* 126(18):1695–1699