Middle-East Journal of Scientific Research 24 (Techniques and Algorithms in Emerging Technologies): 309-317, 2016
ISSN 1990-9233;

DOI: 10.5829/idosi.mejsr.2016.24.TAET23543

Image Retrieval Based on Similarity Search Using the Combination of Fisher Criterion Based Genetic Algorithm

T. Sowmi and E. Saravana Kumar

P.G Scholar, Department of Computer Science and Engineering,
Adhiyamaan College of Engineering, Hosur, Tamilnadu, India

Abstract: In multimedia community, the retrieval of relevant images became the major research issue due to the wide availability of data. From the enormous review on hashing based techniques, the retrieval rate and the time complexity was high related to similarity search. So, a novel fisher criterion based genetic algorithm is incorporated. For time consumption in searching process, four feature extraction methods are expected to reduce the redundant features. As a consequence, fisher criterion based genetic algorithm is applied to optimize the feature subset by considering the fisher fitness value. The extensive experiments are conducted with four publically available datasets and compared with state-of-the-art hashing techniques.

Key words: Hashing · Fisher criterion · Genetic Algorithm · Dimensionality reduction · Feature Extraction

INTRODUCTION

Due to the rapid development of digital technologies, many photo sharing website such as Google+, Facebook, Twitter etc., are becoming the most frequently using website [1]. While comparing with earlier days, at present the usage of those website are increased as well as the storage of multimedia data as text, image, video and audio which has been shared and uploaded by the user are also increased. From the survey on multimedia data, Parfeni on 2011 [2] reported that the Flickr website was uploaded with 6 billion images. Later on Jeffries [3] reported that the uploaded images per day became to 3.5 million images on 2013. So retrieving the relevant images (i.e., similarity search) based on feature extraction method became the most challenging task in multimedia community as well as it became the current research issue in multimedia retrieval.

The similarity search [4] is defined as a general term which is used to search the similar information from large information repositories. This search is mainly used where the objects containing repository do not possess in natural order. Under similarity search, nearest neighbor search is considered as the subclass of it and it has the task of identifying the similar samples. If the dimension of the image is high, it leads to curse of dimensionality [1] and it affects the retrieval efficiency. But this search is applicable only for the low dimension of the image.

The dimensionality of the image can be reduced in two ways as feature extraction [5] and feature selection [6,7] process. Feature selection is considered as the process of selecting the optimal feature subset based on the objective function and feature extraction converts the high dimensional data onto a low dimensional space. Criterion for the feature reduction is based on problem settings. Moreover, to retrieve the relevant image for the query image is infeasible because the dimensionality of the image is high (i.e., curse of dimensionality). The high dimension of the image are encoded to low dimension for the efficient similarity search in large scale database. To address the above problem, many optimization techniques were used. Recently many hashing based optimization techniques were considered. They commonly used feature extraction methods to extract the features from the images and then used Locality Sensitive Hashing (LSH) to generate the binary bit [8,9]. But the recognition rate was low interms of considering the precision and recall rates [10].

From the survey on genetic algorithm [11], Fisher Criterion based Genetic Algorithm is considered to improve the retrieval efficiency which is a popular and good choice for retrieval process. Genetic algorithm is the

Corresponding Author: T. Sowmi, P.G Scholar, Department of Computer Science and Engineering,
Adhiyamaan College of Engineering, Hosur, Tamilnadu, India.

one which is mostly used in feature selection algorithm. In common, feature selection is considered as the process of selecting the best feature subset from the feature set. Normally, they used classifiers as SVM, K-nearest neighbor to generate the objective function. This procedure of evaluating the objective function is time consuming, because by using the classifiers, each and every subset will be retrained and then again classification is performed for the feature subset. And also the quality of the features subset (i.e., Objective function) may not produce the same result as another one [12]. So to consume the time for evaluating the objective function, fisher criterion optimization technique is used which is considered as one of the feature selection algorithm. It eliminates the redundant features where most of them considered that reducing the features tends to the best retrieval performance. This algorithm undergoes three technologies for feature selection as filter, wrapper and embedded method [13]. In filter, based on ranking algorithm the best feature subset is selected and in wrapper and embedded methods, the optimal feature subset is selected in means of considering the performance of the classifiers. But in combination of fisher criterion with genetic algorithm, it estimates the fitness score (i.e., Fitness function) for the feature vector. Initially, instead of selecting the individual features this criterion finds the optimal feature subset interms of consuming time and retrieval efficiency. Then based on the fisher fitness evaluation, the optimization process is processed which is described in section IV-B.

The remainder of this paper is described as follows: In section 2, the combination of fisher criterion based genetic algorithm is survived. In section 3, the extraction of features to remove the redundancy is discussed and in section 4, the working of fisher criterion and Genetic Algorithm are described. In section 5, the result for the optimization process as well as the comparison results for four datasets are shown. Finally, the conclusion of this paper is described in section 6.

Related Works: In this literature, fisher criterion and Genetic algorithm are survived related to image mining.

Fisher Criterion: Feiping Nie *et al*. [14] used trace ratio criterion algorithm for feature selection [14]. Laplacian score and fisher score are the two feature selection algorithm used which determines the score for estimating the fittest solution. The main idea of Laplacian score was to evaluate the features by considering the locality preserving power, because the two data points (i.e., feature values) may lag to the same portion of the image. Based on the laplacian score, the power of the locality was preserved. Then, the fisher score selection algorithm proceeded in calculating the mean and weighted distance for each class in the database. In combination of Laplacian and fisher scores, the objective function was predicted and they directly optimized the scores for the feature subset selection process.

The combination of Fisher's Criterion and Linear Discriminant Analysis [15] for Face recognition was proposed by Marryam Murtaza *et al*. [15]. By combining these two algorithm, they overcame the inadequacy of Linear Discriminant Analysis (LDA) and Maximum Margin Criterion (MMC) which was one of the form of conventional LDA. Generally, LDA is considered as a supervised batch classifier and functions as converting the high dimensional input data to the low dimensional data. Under the reasonable computational cost, they fought against the singularity of within class scatter matrix where the number of samples in the intra class is smaller than the dimensionality of the samples. LDA and MMC reduces the computational complexity in the feature free subspace by using the minimum Redundancy Maximum Relevance (mRMR) algorithm.

Quanquan Gu *et al*. [13] generalized the fisher score for feature selection [13]. Fisher score was determined as the supervised feature selection strategy. It selects the optimal features independently based upon their score which has been estimated by the fisher criterion. The filter based fisher criterion was usually derived as a binary selection of features which maximize the performance of the selection process. The fisher score was calculated in terms of considering the distance between local points. From the fisher score the top ranked n numbers were selected because the scores were determined independently it neglects the combination of feature. The selection procedure deals with heuristic algorithm which was suboptimal solution.

Zhi-Wei Hou *et al*. [16] proposed Kernelized Fuzzy Fisher Criterion based clustering Algorithm [16]. Based on the estimation of Euclidean distance, the clusters were formed. For each cluster, mean and the weighted distance were calculated and it acted as a threshold point.

Genetic Algorithm: A genetic algorithm based wrapper feature selection method for classification of hyper spectral images using support vector machine [17] was proposed by Li Zhuo [17]. They combined the optimization algorithm with the SVM classifier, which reduces the computational complexity for obtaining the optimal feature subset and also improved the Classification Accuracy Rate (CBR). The CBR can be predicted as the quality measure for the optimized feature subset. Using this hybrid algorithm, the feature subset as well as the kernel SVM parameters were determined at the same time.

Self-optimizing Image segmentation system based on genetic algorithm [18] was described by Bir Bhanu et al [18]. In common, the segmentation problem was considered as an optimization problem and it was a difficult task of understanding any automated image process. The segmentation process was adapted by incorporating the genetic algorithm with the self-optimizing technique. The hyperspace of segmentation parameter was efficiently searched by GA and found the approximate global maximum solution. In this process after getting the image it analyze and finds the characteristics of the image and passes the information along with the external variables to the genetic learning component. Normally the segmentation algorithm finds the global optimum solution instead of finding the local solution by Jing Kong (2009) which was considered as a practical and effective segmentation algorithm.

Attakitmongcol K and Srikaew A [19] proposed a new approach for optimization in watermaking by using genetic algorithm [19]. Discrete Multiwavelet transform was used to propose the spread spectrum image watermaking algorithm which improved the visual quality of watermaked images and robustness of the watermark. Khaled Loukhaoukha et al (2010) described Multi-objective Genetic Algorithm for Image watermarking based on singular value decomposition and Lifting Wavelet transform. In this they used Multiple Scaling Factors to achieve the highest robustness without losing watermark transparency. But determining the optimal values for Multiple Scaling Factors was quite difficult.

Texture-Based Identification and Characterization of Pneumonia Patterns in Lung [20] was proposed by Anup R. Aswar, Kunda P. Nagarikar *et al*. [20]. They surveyed that the Identification and characterization of diffuse parenchyma lung disease (DPLD) patterns was difficult. So an automated scheme for volumetric quantification of interstitial pneumonia (IP) patterns was implemented which was the subset of DPLD. This algorithm gave a deep understanding of feature selection technique. FCM considered images as separate points. Because the spatial dependence was not considered by fuzzy function.

Xiabi Liu *et al*. [12] proposed feature selection method and genetic optimization algorithm [12]. They selected the fittest features among the various extraction method interms of bag-of-words, wavelet transform method and histogram. They also used classifiers as SVM, KNN and Naive Bayes to label the images in the dataset. Based on that classification the retrieval process was processed. Among these classifiers SVM classifier performed best in classifying the labels among the classes. But KNN and Naive Bayes classifiers lagged in their performance.

Feature Extraction: The feature extraction methods is used to extract the low dimensional features from the high dimensional data [5]. When the input (i.e., query image) given to the process, it may be too large to process with redundant pixel values. So to transform into the reduced set of features (i.e., feature vector), the feature extraction method is used. Then the desired task can be performed by considering the feature vector instead of using the complete data's. In this implementation, four extraction methods are considered which are discussed as follows.

GIST [21], commonly used in web scale image search. It normally retrieves the values from the same landmark and is used for image completion which does not require any form of segmentation. The image is divided into 4-by-4 grid for which the intensity of the images are extracted. It focus on the shape of the image and the relationship among the outlines of the image.

Pixel [22], estimates a normalized value for the pixel value of the image. The pixel value of the image ranges from 0 to 255. By estimating the normalized value, it contains about 512 dimension as in feature vector.

Bag-Of-Words [1], similar patches are grouped into same cluster by using k–means clustering. Considering the pixel values, the range among them are determined and based on the range, centroid is calculated. The distance which are close to the centroid are grouped into the same cluster. Then the codewords are generated by rectifying the cluster formation.

Scale Invariant Feature Transform [23], which is shortly declared as SIFT. The main function of this descriptor is to detect the keypoints in the gray scale image and performs localization and filtering to remove the unstable points. Detection and Localization takes place by using Gaussian and laplacian function.

Methodology: In previous work, the combination of fisher criterion based genetic algorithm was considered for image classification [12]. Generally, while considering the

Genetic Algorithm, two main problems takes place. 1) How to evaluate the feature subset? 2) How to perform search? The feature subset is obtained by considering the fisher fitness score and where the search is performed by evaluating the distance among the feature subset. The estimation of fitness value and Euclidean distance calculation are discussed elaborately in section A and C.

Fisher Fitness Evaluation: The fisher fitness criterion is considered as a feature selection algorithm [12] interms of removing the redundant features for easy search. This algorithm is used as a threshold point for generating the binary bit among the feature values (i.e., Decimal values). It estimates the fitness function (i.e., Objective function) by calculating the mean and distance among all the classes. In general, for example if the dataset contains 100 images and separated as 10 classes (i.e., category) and each class contains of 10 images. The algorithm for fitness estimation is presented in table 1. The mean and weighted distance calculation for estimating the fitness function are described as follows,

Mean Calculation: The mean is generally calculated by considering the dimension of the image for the whole dataset as well as for each classes. The mean estimation for separate class is given as,

$$m_1 = \frac{\text{Sum of dimension for each image}}{\text{Number of images in the class}}$$

The mean estimation for the whole dataset,

$$m_2 = \frac{\text{Sum of category and dimension of each image}}{\text{Number of images in the dataset}}$$

Weighted Distance Calculation: Two forms of weighted distance are calculated. First, the average weighted distance for the whole dataset as well as for the corresponding mean is calculated by considering the dimension and mean for all images in the dataset. Second, the average weighted distance between the classes is calculated by the sum of mean estimation for whole dataset and for separate classes.

Table 1: Fisher Criterion Algorithm

Algorithm 1: Fisher criterion Evaluation
Input:Feature vector, Number of category in the dataset, Number of images in each category
Step 01: Construct the mean for each category as well as for the whole dataset as (1) and (2).
Step 02: Construct the average weighted distance for the overall dataset and between the classes as in section IV-A-2
Step 03: Estimate the fitness value.Output: Fitness value

Fitness Function: The fitness function is calculated by dividing the two forms of weighted distance. Based on the fisher fitness value, the binary bit is estimated. The binary bit acts in the range of 0's and 1's. The decimal values are converted to the binary string for easy convergence. The fitness value act as a threshold point for the feature vector. The values higher than the fitness value is generated as 1 or else it is generated as 0.

Optimization Process: In Genetic Algorithm optimization process, it includes 5 main terminologies as population initialization, selection, crossover, mutation and termination [12,4]. The flow of these terminologies are described as follows and refer table 2 for GA.

Population Initialization: The population is initialized by considering the feature vector which has been estimated by the extraction method. The extraction methods are briefly described in section III. In GA, feature weight vector is considered as an individual. Each individual is encoded in binary string as 0's and 1's.

Selection: The selection operator is considered as a main operator in Genetic Algorithm. The individuals are selected by considering the fisher fitness evaluation which is discussed in section IV-A. The top most two maximum individual fittest is selected which act as a parent individual. The probability for selecting the parent individual is calculated by considering the fitness value for each and every image in the dataset. This probability estimation for each individual in % form declares the chance of each individual getting selected for a parent individual.

Crossover: The crossover operator is used to produce two individual set based on the selected parent individuals.

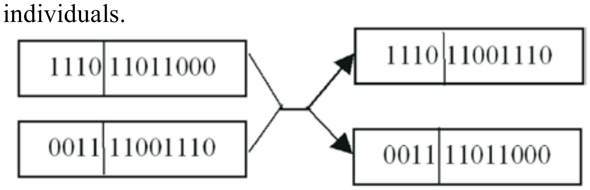


Fig 1. Crossover Example

In this operator, the mean for the parent individuals fittest is calculated and it acts as a threshold point for generating new individual set. The points larger than the threshold point are formed in one set and smaller values are formed in another set. The example is shown in figure 1. The mean estimation for parent individual is given as,

$$m = \frac{\text{Sum of the fitness value of parent individual}}{2}$$

Table 2: Genetic Algorithm

Algorithm 2: Genetic Algorithm

Input:
Feature vector, Fitness value
Step 01: Initialize the population by considering the feature vectors
Step 02: Estimate the fitness value for each image as in algorithm 1
Step 03: Select the top two fittest value as parent individual
Step 04: Construct the mean for two fittest value of parent individual as m in (3)
Step 05: if m >feature vector, set in s_1
Step 06: else
Step 07: set in s_2
Step 08: Construct the mean for s_1 and s_2 as m_{max} and m_{min} in (4)
Step 09: if $m_{max} > s_1$ and $m_{min} > s_2$, then the values are accepted; otherwise discarded
Step 10: until convergence criterion satisfied
Output: Reduced feature subset

Mutation: Generally, mutation operator changes one bit in the newly generated offspring (i.e., child chromosome) and uses the changed bit individual for the next generation population. But in image retrieval process, as it proceeds in the crossover operator, the mean is calculated separately for each individual set and it acts as a threshold point. The values larger than the threshold point are accepted and other values are discarded. Finally, one set is formed by considering the larger values. The example for mutation operator is shown in figure 2.

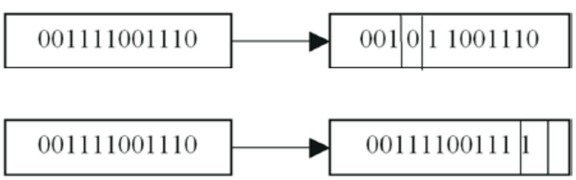


Fig. 2: Mutation Example

The mean estimation for two individual sets,

$$m_{max}, m_{min} = \frac{\text{Sum of dimension of the image}}{\text{Number of images in the set}}$$

Termination: In GA, the stopping criterion is defined in efficient manner. The algorithm is terminated when it reaches the maximum fittest value. The stopping criterion is given here is when the difference between two adjacent fitness value reaches the value 0.002, the algorithm gets terminated.

Image Euclidean Distance: Euclidean distance is considered as estimating the distance measure between two images (i.e., pixel difference between two images). To calculate the distance measure, the dimension of the image must be same. If the image is in single channel, the absolute difference is computed from each channel of the other image. If both images are in multichannel, the values in each channel are compared separately [30]. The distance measure is calculated as,

$$d^2(x,y) = \sum_{k=1}^{N} (x^k - y^k)^2$$

Where, d represents the distance measure, x and y represents the images and k represents the dimension of the image.

Experiments: The experiments are conducted to verify the performance of proposed method with state-of-the-art hashing techniques. The implementation is carried out in Matlab with four publically available datasets.

Datasets: The experiment is conducted with four different publically available datasets as MIR Flickr [25], CIFAR-10 [26], NUS-wide [27] and SIFT-1M [23]. Each dataset contains about different collection of images in the resolution 256×256.

MIRFlickr, contains about 25,000 images with high clarity which are collected from the flickr website. The website contains about both image tags and image contents for research purposes.

CIFAR-10, the images in the dataset are represented in independent labels using the wordnet lexical database. The website contains about 60,000 images with 10 classes and 6000 images in each class.

NUS-wide, National University of Singapore created a dataset for web media search containing about 2,70,000 images with 1000 frequent tags. This dataset was considered as a large scale web image database.

SIFT-1M, contains about 1million feature vector which are extracted from the large set of images. The dimension of the image is 128D, each point representing the terms localization and orientation.

Performance Evaluation: The performance of the retrieval rate is evaluated by true positive, true negative, false positive and false negative rates which is considered interms of precision and recall. If a relevant image is retrieved for the given query image then it is considered as true positive otherwise it is considered as false negative and the meaning of true negative and false positive is as same [12].

Experimental Results: The experimental results are classified into three forms as feature extraction, dimensionality reduction and comparison with various hashing techniques. The results are shown as follows.

Results of Feature Extraction: In retrieval process, the features are extracted from the images which converts the high dimensional data to low dimensional data to remove the redundant features. Four different dataset is assigned with four different extraction method and the result is shown in table 3.

MIR-Flickr dataset is assigned with GIST feature, which estimates the normalized value based on the intensity and pixel value of the image. First, the gray scale image is divided into 4-by-4 grid and then the intensity for the image is calculated. The meaning of intensity indicates, in gray scale image, the color of the image will be in black and white where 0 is assigned for black color and 1 is assigned for white color. CIFAR-10 dataset is assigned with Pixel feature, it generally creates a normalized value by considering the pixel value of the image.

Table 3: Dimensions of the image

Datasets	MIRFlickr	CIFAR-10	NUS-wide	SIFT-1M
FeatureDimension of the image	GIST256	Pixel512	BOW100	SIFT100

NUS-wide dataset is assigned with BOW feature, which estimates the normalized value based on clustering. For the clustered values, the bag-of-words are generated where each point are specified in 100D.

SIFT-1M dataset is assigned with SIFT feature, based on localization and orientation the normalized value is estimated. First, the keypoints is detected by using Gaussian function and then the unstable points are removed by laplasian function. For the stable points, the orientation of the image is calculated which is estimated in 100D. The extracted feature values for each image is represented by 1×D (i.e., feature vector). The dimension of the image varies by considering the extraction methods.

Results of Dimensionality Reduction: For the dimensionality reduction, the Genetic algorithm is used. First, the fitness function is calculated by considering the mean and average weighted distance for each class as well as for the corresponding class. The value acts as a threshold point for binary bit generation. The result of fitness value for four datasets is specified in table 4.

Table 4: Fitness value estimation for four datasets

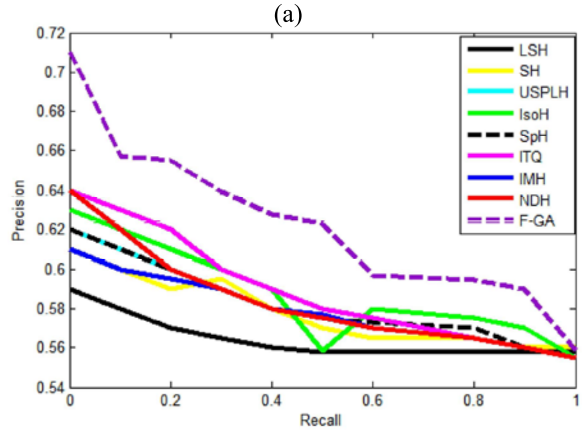
Dataset	MIRFlickr	CIFAR-10	NUS-wide	SIFT-1M
Fitness value	2.61333e-10	1.8453e-10	7.956e-10	1.0861e-08

The optimization process is initialized with 5 steps. First, the population is initialized by means of considering the feature vectors. Then the parent individual is selected by considering the fitness value. The top two fittest individuals are selected as a parent individual. Based on the parent individual, two new offspring are generated which comes under the crossover operator process.

The mean is calculated for two parent individuals which act as a threshold point. The mean value is compared with the feature vector. The values higher than the threshold point are stored in s_1 variable and smaller values are stored in s_2 variable where, two new individual set are formed.

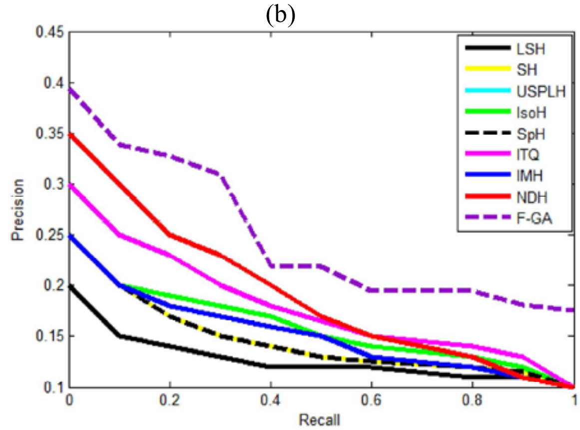
MIRFlickr

(a)



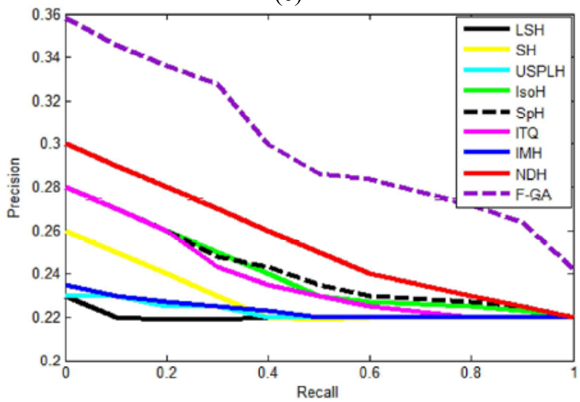
CIFAR-10

(b)



NUS-wide

(c)



SIFT-1M

(d)

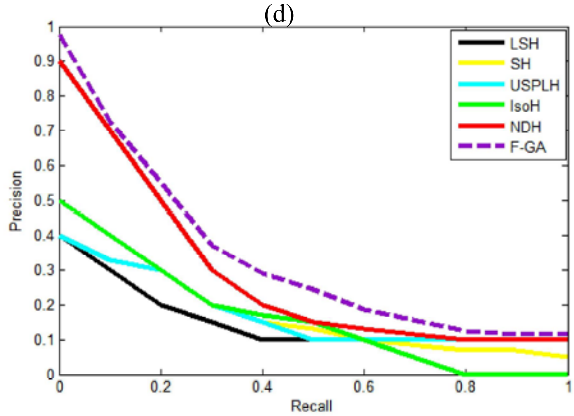


Fig. 3: Impact on the retrieval rate interms of precision and recall by considering the true positive and false negative rates. (a)-(d) indicates precision and recall of the 100 returned images for four datasets respectively.

Table 5: Dimensionality reduction

Datasets	MIRFlickr	CIFAR-10	NUS-wide	SIFT-1M
Number of selected features	128	256	50	50

In mutation operator, a mean is calculated for two sets individually where, this operator goes as same as crossover operator. The mean m_{max} and m_{min} acts as a threshold point and compared with s_1 and s_2 sets. The values larger than the threshold point are considered and remaining values are discarded.

This process continuous until it reaches the maximum fittest value among the adjacent generation. After termination, the dimensions of the image are reduced where the result is specified in table 5.

Comparison with State-of-the-Art Techniques: To prove the performance of fisher criterion based genetic algorithm, the retrieval rate is compared with state-of-the-art techniques. The retrieval rate is considered interms of true positive and false negative rates. The comparison of various optimization technique are listed as follows.

LSH [9], locality sensitive hashing for binary bit generation using random projections.

SH [28], Spectral Hashing learns the hashing function by means of considering the neighbors in input space.

USPLH [29], unsupervised sequential projection learning for hashing learns the hashing function in sequential manner.

IsoH [30], Isotropic hashing function learns the projection functions by considering the projected data in means of isotropic variances.

SpH [31], hypersphere based hashing function projects the coherent data points into a binary code.

ITQ [32], iterative quantization proceeds with the iteration process to achieve the initial projection matrix.

IMH [33], reflects the binary code based on geodesic distance.

NDH [1], Neighborhood Discriminant Hashing estimates the objective function by updating the transformation matrix.

While comparing with other optimization technique, the Fisher criterion based genetic algorithm provides better result in retrieval rate. The true positive rate is improved by 50% better than the existing hashing technique and the result is shown in figure 3.

Computational Complexity: In NDH method, the computational time was high for learning the compact binary codes [1]. So, for optimization process it consumed more time. In order to reduce the computational time for optimizing the feature subset, the binary bit generation and the calculation of objective function are determined at the same time using the novel proposed method.

Table 6: Time measured in seconds for overall process

Datasets	MIRFlickr	CIFAR-10	NUS-wide	SIFT-1M
NDHProposed method	9040	6530	12090	220100

The computation time calculation depends on the number of dimension and images considered for the experiments. The results for the comparison of proposed method with NDH method is shown in table 6.

CONCLUSION

This paper has proposed the combination of fisher criterion based genetic algorithm for image retrieval process. In hashing based techniques, if the images related to the query image are presented in other class, the retrieval rate becomes low and whereas, the time complexity was also high while learning the hashing function. Meanwhile, the novel method maximized the retrieval rate of about 50% better than state-of-the-art techniques and minimized the time complexity as much as possible

REFERENCES

1. Jinhui Tang and Zechao Li, 2015. Neighborhood Discriminant Hashing for Large-Scale Image Retrieval," Image Processing, vol. 24, no. 9, sep 2015.
2. Parfeni, L., 2011. Flickr boasts 6 billion photo uploads," Softpedia, Tech. Rep., Aug. 2011.
3. Jeffries, A., 2013. The man behind Flickr on making the service 'awesome again,"' The Verge, Tech. Rep., Mar. 2013.
4. Rajan, C. and N. Shanthi, 2015. Genetic based Optimization for multicast Routing algorithm for Manet' Sadhana - Academy Proceedings in Engineering Science, 40(7): 2341-2352.
5. Image retrieval- "Feature Extraction", url: https://feature extraction.com
6. Mohd Saberi Mohamad, Safaai Deris, Safie Mat Yatim and Muhammad Razib Othman, 2004. Feature Selection Method using Genetic Algorithm for the Classifcation of Small and High Dimension Data, First International Symposium on Information and Communications Technologies. October 7-8, 2004. Putrajaya, Malaysia.
7. Image mining- "Feature selection", url: https://feature selection.com
8. Kulis, B. and K. Grauman, 2009. Kernelized locality-sensitive hashing for scalable image search, in Proc. IEEE 12th Int. Conf. Comput. Vis., Sep./Oct. 2009, pp: 2130-2137.
9. Indyk, P. and R. Motwani, 1998. Approximate nearest neighbors: Towards removing the curse of dimensionality, in Proc. ACM Symp. Theory Comput., pp: 604-613.
10. Sowmi, T. and E. Saravana Kumar, 2015. A Survey on Similarity Search for Large Scale Database, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 4, Issue 9, September 2015.
11. Sowmi, T. and E. SaravanaKumar, 2016. A survey on Fisher Criterion based Genetic Algorithm for Feature Selection Method," South Asian Journal of Engineering and Teachnology, 2(12): 1-8.
12. Xiabi Liu, Ling Ma and Li Song, 2015. Recognizing Common CT Imaging Signs of Lung Diseases Through a New Feature Selection Method Based on Fisher Criterion and Genetic Optimization, IEEE Journal of Biomedical And Health Informatics, Vol. 19, No. 2, March 2015.
13. Quanquan Gu, Zhenhui Li and Jiawei Han, 2010. Generalized Fisher Score for Feature Selection,
14. Feiping Nie, Shiming Xiang, Yangqing Jia, Changshui Zhang and Shuicheng Yan, 2008. Trace Ratio Criterion for Feature Selection, Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence.
15. Marryam Murtaza, Muhammad Sharif, Mudassar Raza and Jamal Hussain Shah, 2014. Face Recognition Using Adaptive Margin Fisher's Criterion and Linear Discriminant Analysis (AMFC-LDA), The International Arab Journal of Information Technology, Vol. 11, No. 2, March 2014.
16. Zhi-Wei Hou, Liu-Yang Wang and Quan-Yin Zhu, 2010. Kernelized Fuzzy Fisher Criterion based clustering Algorithm," Distributed Computing and Applications to Business Engineering and science., pp: 87-91.
17. Li Zhuo and Jing Zheng, 2008. A Genetic Algorithm based Wrapper Feature Selection Method for Classification of Hyperspectral Images Using Support Vector Machine," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing 2008.
18. Bir Bhanu, Sungkee Lee and John Ming, 2009. Self-optimizing Image segmentation system using a genetic algorithm,
19. Attakitmongcol, K. and A. Srikaew, 2014. A new approach for optimization in watermaking by using genetic algorithm," IEEE Transaction on Signal Processsing," DOI: 10.1109, Sep 2014.

20. Anup, R. Aswar, Kunda P. Nagarikar and P.D. Pawar, 2015. Texture-Based Identification of Interstitial Pneumonia Patterns in Lung Multidetector CT, IETE 46th Mid Term Symposium “Impact of Technology on Skill Development” MTS-2015.
21. Matthijs Douze and Herve Jegou, 2009. Evaluation of GIST descriptors for web-scale image search, CIVR 09, July 8-10, 2009.
22. Karen Glocer, Damian Eads and James Theiler, 2005. Online Feature Selection for Pixel Classification, Appearing in Proceedings of the 22nd International Conference on Machine Learning, Bonn, Germany,
23. Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis., 60(2):. 91-110.
24. Liwei Wang, Yan Zhang and Jufu Feng, 2005. On the Euclidean Distance of Images, IEEE computer society, pp: 1334-1339.
25. Huiskes, M.J. and M.S. Lew, 2008. The MIR Flickr retrieval evaluation,” in Proc. ACM Int. Conf. Multimedia Inf. Retr., pp: 39-43.
26. Torralba, A., R. Fergus and W.T. Freeman, 2008. 80 million tiny images: A large data set for nonparametric object and scene recognition, IEEE Trans. Pattern Anal. Mach. Intell., 30(11): 1958-1970
27. Chua, T.S., J. Tang, R. Hong, H. Li, Z. Luo and Y. Zheng, 2009. NUS-WIDE: A real-world Web image database from National University of Singapore, in Proc. ACM Int. Conf. Image Video Retr., 2009.
28. Weiss, Y., A. Torralba and R. Fergus, 2008. Spectral hashing, in Proc. Adv. Neural Inf. Process. Syst., pp: 1753-1760.
29. Wang, J., S. Kumar and S.F. Chang, 2012. Semi-supervised hashing for largescale search, IEEE Trans. Pattern Anal. Mach. Intell., 34(12): 2393-2406
30. Kong, W. and W.J. Li, 2012. Isotropic hashing, in Proc. Adv. Neural Inf. Process. Syst., pp: 1646-1654.
31. Heo, J.P., Y. Lee, J. He, S.F. Chang and S.E. Yoon, 2012. Spherical hashing, in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., pp: 2957-2964.
32. Gong, Y., S. Lazebnik, A. Gordo and F. Perronnin, 2013. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval, IEEE Trans. Pattern Anal. Mach. Intell., 35(12): 2916-2929.
33. Shen, F., C. Shen, Q. Shi, A. van den Hengel and Z. Tang, 2013. Inductive hashing on manifolds, in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., pp: 1562-1569.