International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Re-Ranking Images Retrieval: Combined Multiple Features Method

John Bosco P, S.K.V.Jayakumar

Abstract: ***Content-Based Image Retrieval (CBIR) grown rapidly in multimedia field, image retrieval, pattern recognition, etc. CBIR provides an effective way of image search and retrieval from the pool image databases. Learning effective relevance measures plays a critical role in improving the performance of image retrieval systems. In this paper present a Combined multiple features method which is two key parameters (i) Feature extraction, (ii) Similarity metrics for content-based image retrieval method. Feature extraction and similarity metrics important role in Content-Based Image Retrieval. We define hybrid feature extraction and similarity method for finding the most similar images retrieved. Combined features extraction using the various image features. These papers explain some important distance metrics such as Euclidean distance and City block distance. The experiments are performed using the various kinds of databases such as WANG Database, Corel Dataset. The experimental result shows that the proposed method is proved more effective than existing methods.***

Keywords: CBIR, Multiple features, Feature Extraction, WANG Dataset.

I. INTRODUCTION

Content-Based Image Retrieval (CBIR) is a widely used system in Digital image processing techniques. Digital image retrieval system plays a key role in searching images in pool database. It is mainly used by internet image scale search engines through two approaches. Firstly, it is text-based image search engine which happens to be a marketable internet scale image search engines that uses the texts based image retrieval [1-5]. Most of the search engine works on Text based Approaches while traditional search method functions on using manual annotation of images. Generally Text based search engine is not a much more efficient system. Text based image retrieval having some disadvantages which are follows 1.Text Query web image searching do not describe the images of the content.2.It do not Capture user intention of images accurately.3.User may not have more sufficient knowledge about the content of the user query image hence the ambiguity of the query key words may lead to the retrieval of irrelevant Images.4.Time consuming as the user also have to spend more time to find the exact image he desired. Second one is CBIR Approaches systems are widely used in multimedia search engines. In (CBIR) Content or image based image retrieval overcomes the keyword based image retrieval or concept based image indexing. Content based retrieval system uses several departments such as medical, engineering and army field besides various commercial system methods [6-7]. An Image is indexed by its visual Content where visual semantics are described by low and high level features. Low level features represent color and texture, shape and layout while high-level feature represents spatial and semantic and context information etc. These features can be divided into three types that are high, middle and low level features [10], CBIR system relies on color ,texture, and shape which are small level image features, The number of existing CBIR search engines has developed more re-ranking methods.

Manuscript published on 30 September 2019
* Correspondence Author
JOHN BOSCO P*, Department of Computer Science and Engineering, Pondicherry University, Puducherry, India. Email: boscoboss84@gmail.com
S K V JAYAKUMAR, Department of Computer Science and Engineering, Pondicherry University, Puducherry, India. Email: skvjey@gmail.com

II. PROPOSED METHOD

CBIR aims at avoiding the use of the textual descriptions. CBIR Approaches system widely used in multimedia search engines. The major CBIR system follows a low-level feature like texture, shape, and color [5,6]. The CBIR systems have more challenges such as semantic gap, ambiguity results, more feedback and limited performance level [9, 10].Hybrid approaches computes the various feature extraction and various similarity metrics content of images attained from Web image retrieval. Hybrid techniques compute the color features, textual features and edge detection to extract and compute the texture feature of images represented in horizontal vertical and diagonal direction methods. Such kind of methods exploits the use of the image visual information to refine the preliminary text-based queries [12]. The re-ranking for image search results can accomplish considerable performance improvements.

Color Features

Algorithm for Color Features

Input: Conversion of RGB images to HSV color space.
Output: Find the color moments (Mean & standard deviation)
Begin
Step1: Read the input image.
Step2: Convert from RGB to HSV components features.

$$\mu_{H,S,V} = \frac{1}{N}\sum_{j=1}^{N} p_{ij} \qquad (1)$$

$$\sigma_{H,S,V} = \left(\frac{1}{N} \sum_{j=1}^{N} \left(p_{ij} - \mu_i \right)^2 \right)^{\frac{1}{2}} \quad (2)$$

Step3: Extract the mean and standard deviation for H, S and V using the Corel Dataset.

Step4: The six dimensional color feature vector $\mu_H, \mu_S, \mu_V, \sigma_H, \sigma_S, \sigma_V$ is extracted.

End

The important features such as the extraction of images, color features are defined in the RGB and HSV histogram [4]. The color features are extracted from the color moments, color histogram, invariant color histogram [20].

Gray Level Co-occurrence Matrixes (GLCM)

GLCM set of features are based on second order statics. GLCM associated texture features calculation are image analysis techniques such as Energy, Entropy, Contrast, Homogeneity, Correlation etc. Gray Level Co-occurrence matrix characterizes spatial relationship of pixels and it's also known as Gray Level spatial dependence matrix which defines relationship between reference samples and the neighbor samples used for the GLCM, Gray Level Co-Occurrence matrix direction are Horizontal(0^0),Vertical(90^0),Diagonally up(130^0) and Diagonally Down(145^0) [8,9,10]. The following are the GLCM feature extraction techniques equations.

$$Contrast = \sum_i \sum_j (i-j)^2 p(i,j) \quad (3)$$

$$Entropy = \sum_i \sum_j (i-j) \log p(i,j) \quad (4)$$

$$Energy = \sum_i \sum_j p^2(i,j) \quad (5)$$

$$Correlation = \sum_i \sum_j \frac{(i-\omega_i)(j-\omega_i)\, p(i,j)}{a_i a_j} \quad (6)$$

$$Homogeneity = \sum_i \sum_j \frac{P(i,j)}{1+|i-j|} \quad (7)$$

Here p denotes matrix that can define a matrix p (i, j) value in pixels.

Edge Detection

Edge Detection finding for a neighborhood with strong sign of change. Edge detection simply finds the regions in an image where we have a sharp change in intensity change in color high values. [13].

Roberts Edge model

Roberts's edge detection is simple method which implements the first order partial derivative by using the Roberts cross gradient operator [14]. Roberts's operator contains of a pair of 2×2 convolution mask. The Roberts mask given below.

$$G_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} G_y = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (8)$$

$$|G| = \sqrt{G_x^2 + G_y^2} \quad (9)$$

$$|G| = |Gx||Gy| \quad (10)$$

$$\theta = \arctan(Gx + Gy) - 3\pi/4 \quad (11)$$

Sobel Edge Model

It is find the gradient magnitude on an image. It is uses a pair of 3×3 convolution masks. One calculating gradient in x direction, anther one is measure the gradient in y direction [15, 17]. The partial derivatives calculated as below.

$$G_x = (p_2 + 2p_3 + p_4) - (p_0 + 2p_7 + p_6) \quad (12)$$

$$G_y = (p_6 + 2p_5 + p_4) - (p_0 + 2p_1 + p_2) \quad (13)$$

The Sobel mask in matrix form are given as

$$G_x = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} G_y = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad (14)$$

Sobel Operator mask can calculated as Magnitude of the edge $|G| = \sqrt{G_x^2 + G_y^2}$,

Approximate strength $|G| = |Gx||Gy|$ the direction of edge $\theta = \arctan(Gx + Gy) - 3\pi/4$

Distance measure are similarity metrics, they are used for comparing the similarity of two images [21]. There are many similarity measure techniques such as (i) Euclidean Distance (ii) City Block Distance [22].

Euclidean Distance

This distance metric is mostly utilized for measurement in CBIR system [1], let q be query image and t be the target database image and let $P_q(q_i)$ and $P_t(t_i)$ be their respective probability densities [5-6].

$$D_{Euc}(q_i, t_i) = \sum_{i=1}^{n} \left(q_{i,} - t_i \right)^2 \quad (15)$$

City Block Distance

This metric is known as the Manhattan distance. The city block calculates the robustness to outliers where this distance metric is calculated by the sum of absolute between two feature vectors of images [10,11].

$$\Delta d = \sum_{i=1}^{n} |Q_i - D_i| \quad (16)$$

Algorithm for proposed Method

Input: Select Query Image from Dataset which is denoted by **User**$_{\text{IMAGE}}$. Dataset (DS) contains Number of images which is representing by **DS**$_{\text{NOIMAGES}}$.

Output: Retrieve the Number of images lR_1, lR_2, lR_3, lR_4, lR_5.... lR_n Є **$DS_{NOIMAGES}$** as most similar images for User Query Image.

1. Collect the features Mean and Standard Deviation from the Dataset **$Images_{NOIMAGES}$.**
2. Calculate the Mean and Standard Deviation for the U_{IMAGE} and **$DS_{NOIMAGES}$** using Egs (1,2)
3. Calculate the Texture feature for the **$DS_{NOIMAGES}$** using equations
4. N=0, Count
5. For i=1 to N do
   - a_1. if the mean and standard deviation of DS $_{IMAGES}$ lies between User image and Dataset
     Images then select the specific image and store in New Database **(New_{DB}).**
   - a_2. Count=Count++…
   - a_3. New Dataset=No of Images.
   - a_4. End if.

   End.
6. Calculate GLCM texture features for $User_{IMAGE}$ and **New Dataset** $_{NOIMAGES}$.
7. Calculate Edge feature for **User** $_{IMAGE}$ and **$Newdataset_{NOIMAGES}$**
8. For each image in New Database **($Newdataset_{DB}$)** do

S_1. Find the similarity metric between **$User_{IMAGE}$** and **New_{DB}** Images.

S_2.Combine the calculated similarity metric of the Color features, GLCM texture features and Edge feature of the images Є **$NewDataset_{DB}$**.

End

9. Store the feature values in ascending order
10. Return the R Most similarity Images lR_1, lR_2, lR_3, lR_4, lR_5.... lR_n is retrieved.

III. EXPERIMENTAL EVALUATION

This section is very important to establish the experimental in various databases. We conducted experiments for demonstrating the effectiveness of our combined edge detection methods (CED).The experiment contained three important categories [24].

We have proved the method has efficient and accurate also proved Precision, Recall and Accuracy. Fig shows the Efficient results. Fig shows the sample Corel Data set images. We conducted experiments using retrieved re-ranking images from the popular Corel datasets which contains 1000 images including 10 categories in each category 100 images i.e. African faces, Roses, Bus, and Car etc. We gathered the top 20 re-ranking images from the result which is very similar to the query image and evaluated the effectiveness in comparing the compute the precision, recall, F- measure and error rate [22,23].

Fig 1.Retrieved result as query image as Mountains

We proposed method is measured in terms of Precision, Recall and Error rate as shown below. We conducted popular experiments using re-ranking retrieved images from Huge Dataset such Corel Dataset.

$$\text{precision} = \frac{\text{Number of relevant images retrieval}}{\text{Total number of images}}$$

$$\text{Recall} = \frac{\text{Number of relevant images retrival}}{\text{Number of relevant images in collection}}$$

$$\text{F} - \text{measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

Table I. The Average Values of Precision Recall

Images Names	Precision	Recall
African faces	84	42
Beach	70	35
Monuments	66	33
Buses	92	46
Dinosaurs	98	49
Elephants	86	43
Flowers	94	47
Horses	94	47
Mountains	72	36
Food	84	42

Table II. The Average Values of Precision Recall and F-Score

Dataset	Average precision	Average recall	F-Score
Corel Dataset	84	42	56.00
Wang Dataset	86	40	54.60

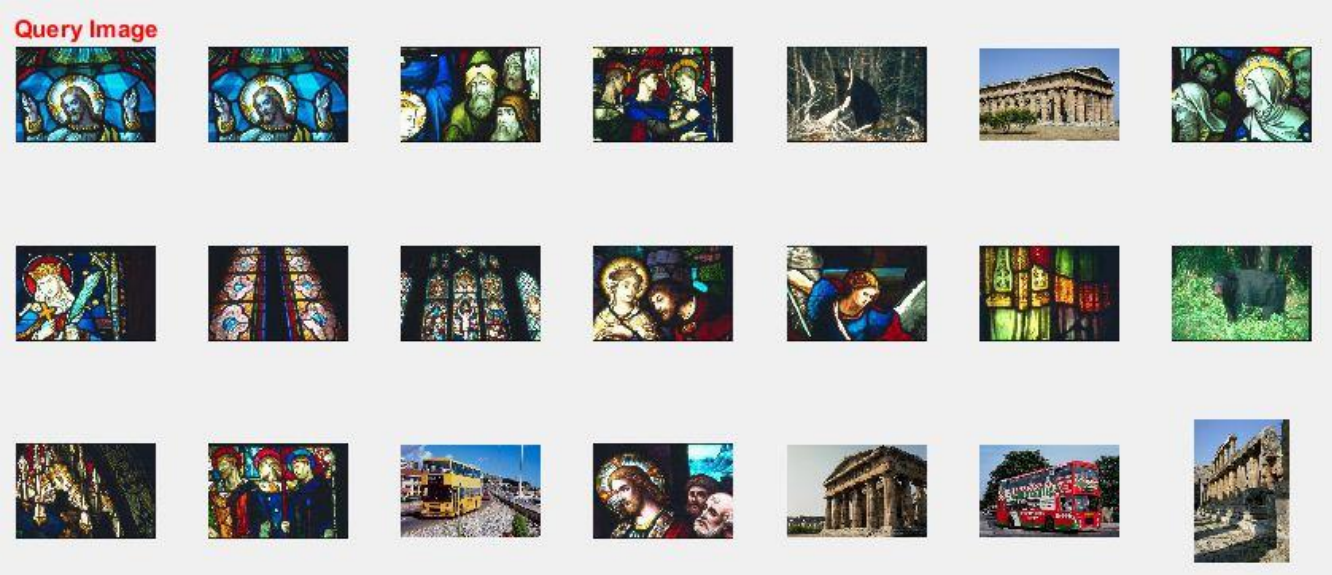

Fig 2.Retrieved result as query image as Jesus

To evaluate CBIR system we need a performance evaluation measure based on similarity image retrieved which help to finding for proposed system efficiency measurement.

We compute precision and error rate and recall also Accuracy. The user can be retrieved the top 20 images results are showed. This proposed work showing the top most 20 images are retrieved, First image as query image.

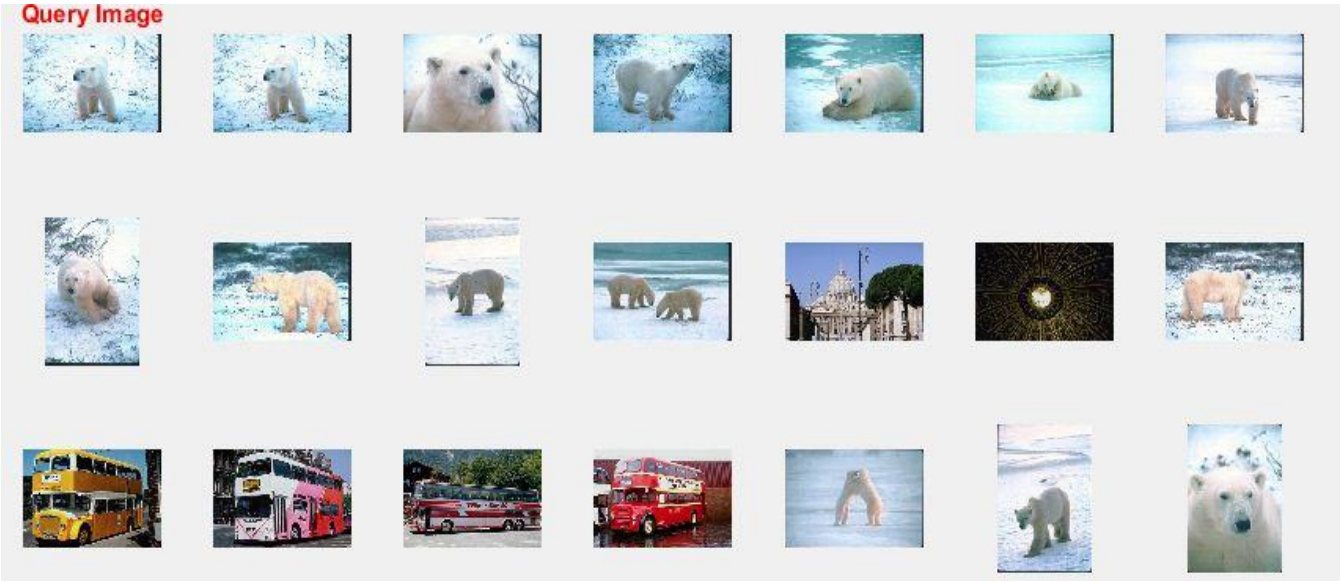

Fig 3.Retrieved result as query image as Lion

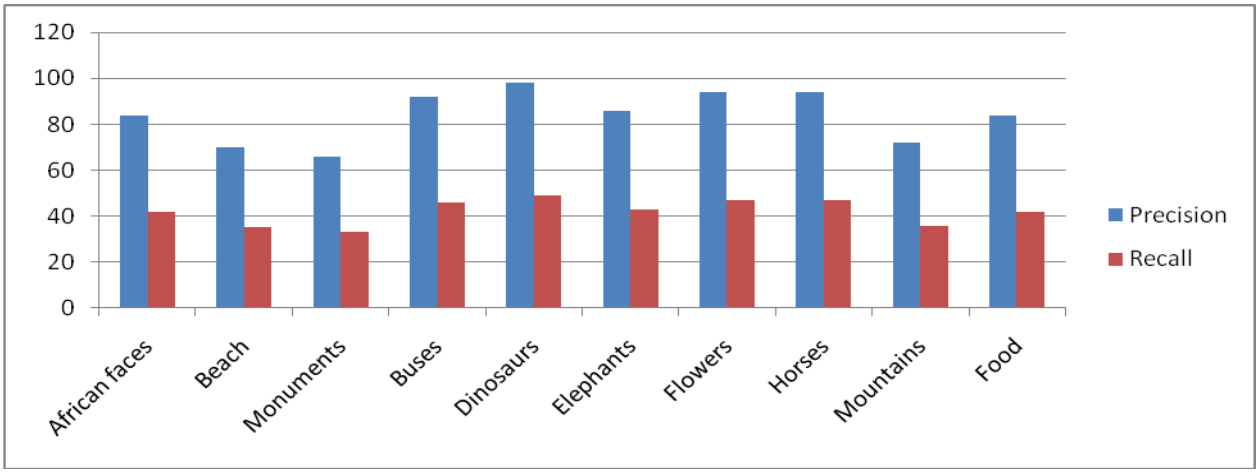

Fig.4 Precision Recall Using Euclidean Distance metrics

Fig 5.Retrieved result as query image as Bus

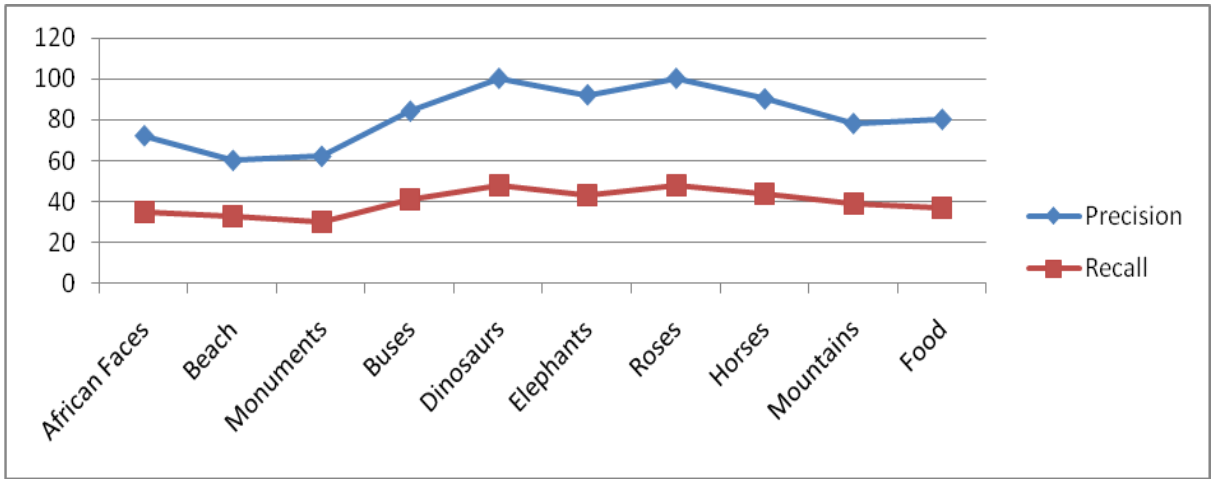

Fig.6 Precision Recall Using Euclidean Distance metrics

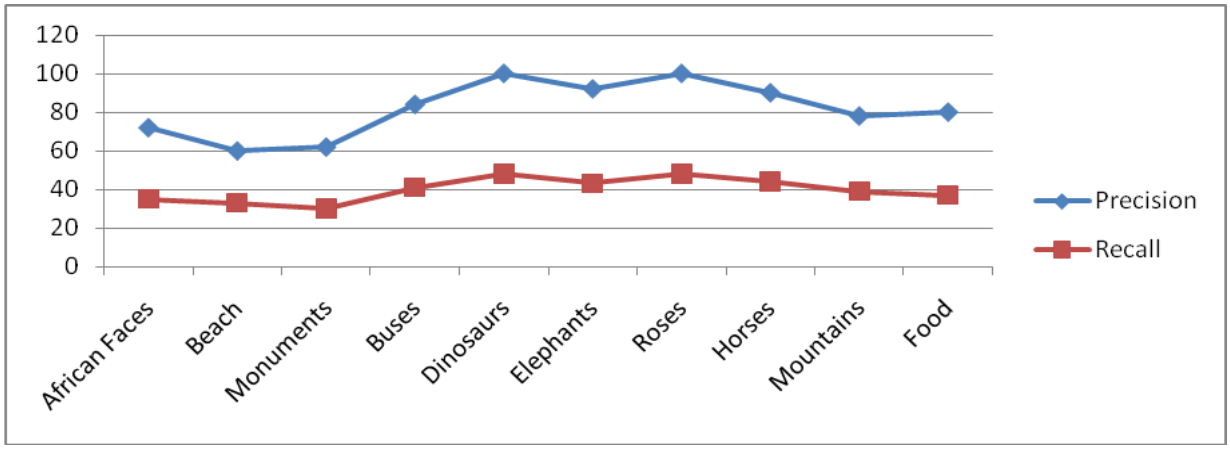

Fig.7 Precision Recall Using City Block Distance metrics

IV. CONCLUSION

In this paper, we discussed the comprehensive study on various type of feature extraction method and various distance metrics. In image processing recent research work of the content based image retrieval system and web image retrieval system. We have been discussed various combined feature extraction methods and distance metrics. In the feature extraction method combination of color and texture is giving better accuracy. In CBIR, We identified after the feature extraction some method applied classification. In Hybrid methods, we found semantic signature is giving better results. So we conclude that in both CBIR and web image retrieval system, performance depends on choosing the combined model of various feature extraction methods and most similar images retrieved.

REFERENCES

1. Guang-Hai Liu, Jing-Yu Yang, (2013) Content –based Image retrieval Using color difference Histogram, *Pattern Recognition.*Vol.46.pp.188-198.
2. Celia, I.Felci Rajam, "An Effective Content Based Image Retrieval Framework Using Machine Learning Techniques,"ICDE1second international conference Data Engineering and *Management* 2010, LNCS 6411, pp 162-169,2012.
3. M.E.EIAlami (2011) A Novel image retrieval model based on the most relevant features,"*Knowledge Based Systems*.Vol.24, pp. 23-32.
4. Hadjer Lachehed Saliha Aouat (2016) SIMIR: New: mean SIFT color multi-clustering image retrieval,"*Mulmedia Tools and Applications*, Vol.76, pp. 6333-6354.
5. Chaoran Cui, Peiguang Lin, Xiushan Nie, Yilong yin, Qingfeng Zhu(2017). Hybrid Textual-Visual Relevance learning for content-based image retrieval, *Journal of visual Communication and Image Representation*, Vol.40, pp.367-374.
6. Ceyhun Celik, Hasan Sakir Bilge (2017). Content based image retrieval with sparse representations and Local feature descriptors: A comparative study,"*Pattern Recognition*.Vol.68, pp.1-13.
7. Ahmad Alzu' bi, Abbes Amira,Naeem Ramzan(2015). Semantc content based image retrieval : A Comrehensive study, Journal of Visual Communication and Image Representation, Vol.32,pp. 20-54.
8. Mutasem K.Alsmadi, "Query Sensitive measure for content- based image retrieval using mete-Heuristic algorithm,"Journal of King Saud University-Computer and Information Sciences, Vol.29,pp.1-9,2017.
9. Xiang pan, Yongqiang Ye, Jinmei Cheng, Danwei, Bin Jiang, "Composite Derivative and edge detection," *Signal Image and video processing*, Vol. 8,pp. 523-531,2014.
10. Nishant Shrivastava, Vipin tyagi(2015). An efficient technique for retrieval of color images in large databases, *Computers and Electrical Engineering*, Vol.46, pp. 314-327.
11. R. Krishnamoorthy, S. Sathiya Devi(2013). Image retrieval using edge based shape similarity with multiresolutionenhanced orthogonal polynomials model,"*Digital Signal Process*.Vol. 23,pp. 555–568.
12. Song Renjie,zhang Ziqi,Liu Haiyang(2017). Edge Connection based Canny Edge Detection Algorithm," *Pattern Recognition and image Analysis*.Vol.27,pp.740-747.

13. P. Srivastava, A. Khare,(2017). Integration of wavelet transform: Local Binary Patterns and moments for content-based image retrieval," *Journal of Visual Communication and Image Representation*, Vol.42,pp.78-103.
14. D. Vikram Mutneja(2015). Methods of Image Edge Detection: A Review,"*Journal of Electrical and Electranical system* ,Vol 4,pp.1-5.
15. A. Sengur, Y. Guo(2011). Color texture image segmentation based on neutrosophic set and wavelet transformation," *Computer Vision and image understanding*,Vol.115,pp.1134–1144.
16. Chaoran Cui,Peiguang Lin,Xiuushan Nie,Yilong Yin, Qingfeng Zhu (2017). Hybrid textual-visual relevance learning for content-based image retrieval, *Journal visual Communication .Image* ,Vol.48, pp.367-374.
17. Guang deng,Jean-Charles pinoli(1998). Differentiation-Based Edge Detection Using the Logarithmic Image Processing Model,*Journal of mathematical Imaging Visio*Vol.180,pp.161–180.
18. Yogita mistry,D.T.Ingole,M.D.Inglore(2017). Content Based Image Retrieval using hybrid features and various distance metric, *Journalof Electrical systems and information technology*, Vol.143, pp.1-15.
19. Guang-Hai Liu,Jing-Yu Yang(2013). Content –based Image retrieval using color differences histogram,"*Pattern Recognition*, Vol 46,pp. 188-198.
20. L.K.Pavithra,T.Sree Sharmila(2017). An efficient framework for image retrieval using color,texture and edge features, *Computers and Electrical Engineering*,Vol.131,pp.1-14.
21. J.Annrose,C.Seldev Christopher(2018). An Efficient image retrieval system with structured query basedfeature selection and filtering initial level relevant images using range Query, *Optik*.Vol.157,pp.1053-1064.
22. I.Felci Rajam, S ValliS(2011). RBIR:Semantic Region Based Image Retrieval by extracting the domonant Region and Semantic Learning,"*Journal of computer science*. Vol 7,pp.400-408.
23. Kehua Guo, Ruifang Zhang, Zhurong Zhou, Yayuan Tang, Li Kuang (2014). Combined Retrieval A Convenient and precise approach for Internet Image retrieval. *Information Sciences*, Vol.358.pp.151-163.

AUTHOR PROFILE

Mr. JOHN BOSCO.*P* pursed B.E Computer Science and Engineering from Anna university of Technology, Tiruchirappalli, India in year 2012 and M.E in Software Engineering from Anna University, College of Engineering Guindy campus, Chennai, Tamil Nadu, and India in year2014. He is currently pursuing Ph.D.in Pondicherry University, Puducherry, India In Department of Computer Science and Engineering undergoing research in the field of Web images Reranking, and content based Image retrieval (CBIR).

Dr. S.K.V.JAYAKUMAR received the PhD degree in Computer Science and Engineering, specialized in Services Computing from Pondicherry Central University. He is a graduate of Electrical Engineering and Masters in Computer Science and Engineering. He is an Associate Professor of Computer Science in School of Engineering at Pondicherry Central University. He has been in Teaching and Research for 22 years. His main research area is Data Science and Services Computing and has number of research publications as output to his research in the said specializations. He has published 29 research papers in International Journals and National and International Conferences. He got Principal Investigator for a UGC funded research project and established a state of the art Cloud Computing Laboratory at Pondicherry University He has good experience of handling funded projects in his specializations and organizing National and International Conferences.