Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IMPACT FACTOR: 6.017

IJCSMC, Vol. 6, Issue. 7, July 2017, pg.128 – 137

A Study on Web Images Retrieval Using Content Based Image Retrieval Methods

John Bosco.P[1]*, S.K.V.Jayakumar[2]

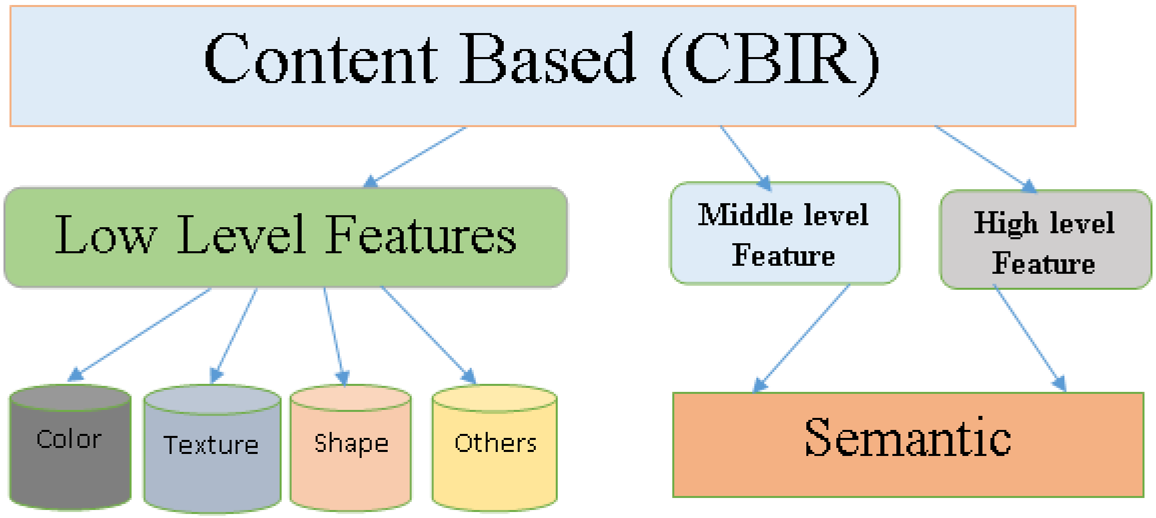
[1]Department of Computer Science, Pondicherry University, Puducherry-605016
[2]Department of Computer Science Pondicherry University Puducherry-605016
Corresponding Author: boscoboss84@gmail.com

Abstract: Content based Image retrieval has wide application in social media, Medical Image, Remote sensing etc. CBIR aims at retrieving the most similar images in a huge database images by taking into account image visual properties. Our extreme goal is ignore the semantic gap between query image and retrieved images. We discuss about the various kind of method in which color, Texture, Shape features of the images are used to improve the similar image retrieval results in the system. Number of techniques has been developed for CBIR methods. This paper proposed for inclusive survey of all the techniques, Such as color histograms (CH), Color moments, and Texture features are GLCM, Gabor Transforms and Tamura Features, with shape features are region moments like that. Also covers the various kinds of Distance Metric (DM), and clustering Techniques, Re-ranking with performance evaluation. In further of discus the various data sets used in CBIR system are also mentioned. Finally we have been discussed recently ten various standard publications papers.

Index Terms— CBIR, Web Images, Distance Metrics (DM).

Introduction.

Content-based image retrieval (CBIR) mostly widely used system in image processing techniques .An image retrieval system more important play role to search images in pool database. Content-based image retrieval is mainly there two approaches used by internet image scale search engines, Firstly is text-based image search engines that is many commercial internet scale image search engines using the text based image retrieval[1],[2].An Image is indexed by its visual content which visual semantics described by low and high level features. One is low level features are represented is color and texture, shape and layout. Another one is high-level feature represented is spatial and semantic and context information etc. Text-based image retrieval methods are traditional image retrieval system [3], [5].The keyword for the images was created by human operators. User types query keywords in the hope of finding a certain type of images the search engine returns a lot of images and unwanted images it is well known that text based image search suffers from the ambiguity of query word. Text-based query keywords for images in a large database they cannot describe the content of the images accurately the return results are inefficient, ambiguity images, expensive and may not capture every text query that describes the image. [6], [14].Secondly CBIR system help in finding efficient image. It is uses the semantic contents of an image such as texture, color and Shape to shape to represent and index the image. Generally, Image descriptor may include both visual and semantic, visual information may include color texture and shape [7], [29], and 31].Visual or object content is descriptor can be either global feature or local features. A global features uses the visual features the whole images and Local descriptor uses features of region or objects in to describe the image content[5].These features dividing three levels that are low middle and high[7,29] CBIR system relies on color ,texture, and shape which are small level image features. CBIR methods search for one specific image which search based on the esthetic value of the images. Many CBIR system has developed. This method techniques tools, and algorithm that are used originated from fields such as statistics and pattern recognition. We have to discuss some important techniques and methods. [28].



Figture 1.CBIR Framework

1. Methods of feature extraction.

The low-level features generally used in CBIR are color, texture, shape and edge. Color, texture are important properties in content-based image retrieval system [1].The CBIR retrieval system commonly follows a similar image pattern. Image is represented by features that are a vector of global features or a set of local features [12, 15].

1.1. Color.

Color features methods is most widely used in QBIR which can be using lot of methods such as color Histogram, color moments, scalable color, dominant color, HSV and RGB Histogram.

The important feature extraction of images, color features are defined RGB and HSV histogram [4].The color features are extracted from color moments, color histogram, invariant color histogram [23, 13].

1.2. Color Moments.

Color moments defined are first, second and third central moments. Each image color channel is used in representing the color feature vector in the Hue, Saturation and Value in HSV color space [21].

1.3. Color Harmonization.

Color harmonization is defined as measure the quality of color distribution of an image. They are following formula [35].

$$harmonic(I_i) = F_{\min(I_i(m,\propto))}$$

$$F\big(I_i,(m,\propto)\big) = \sum_{pj \in I_i \| H(P_j) - E_{T_{m(\propto)}} \| \times s(p_j)}$$

Where H,S and V are Hue, Saturation, and value respectively., p_j denotes j-th pixel in $I_{i,}$ $_{Tm}$ (a) is the harmonic templates. And m is one of the seven Templates, and (a) is orientation for, m.

1.4. Hue, Saturation, and Value (HSV).

HSV and HSL both are common cylindrical coordinate representations of points in an RGB methods.Hue (H), Saturation(S), and Value (V) representations as for each pixel j in image $I_{i,}$ we computer their average for I_i

$$\mathrm{hue}(\mathrm{I_i}) = \sum_{j} \mathrm{h}(\mathrm{p_j})/_{|\mathrm{I_i}|}$$

Saturation (I_i), and Value (I_i) can be easily obtained by replacing H (p_j) with S ((p_j)) and V (p_j), respectively.

1.5. Color Histogram

Color histogram is two types that are (i) GCHs (Global Color Histograms) (ii) LCHs (Local Color Histogram) Color is RGB and HSV histograms are used to represented the color channel of an image(RGB), The concept of local Binary Patterns (LBP).and Local Ternary Patterns(LTP) with Directional Binary Code(DBC) has been utilized to extract features from individual color channel (R,G,B)[26,32],

1.6. RGB Histogram.

Colors are defined in three-dimensional color spaces. These could either be RGB (Red, Green, and Blue), HSV (Hue, Saturation, Value) or HSB (Hue, Saturation, and Brightness).

1.7. HSV Histogram.

HSV denotes the Hue, Saturation and Value. Which is represented by image histogram. It is a graphical representation of the number of pixel values in an image.

They have a lot of methods like is a color histogram, Global Color Histogram (GCHs) and Local Color Histograms (LCHs) that denotes the single and global pixel values.[21].

1.8. Local Binary Patterns (LBP).

The LBP is texture features methods. Texture feature in a local neighborhood. In general given 3x 3 pixel pattern, which is calculated from center and high pixel value with its neighborhoods as the following equations [11].

$$LBP_{N,R} = \sum_{i=0}^{N-1} 2^i \, X \, f_1(p_i - p_c)$$

Here *n* denotes the number of neighbors, R is a radius of the neighborhoods, p_c denotes grey value of Centre high pixels. P_i the gray values of its neighbors.

1.9. Local Ternary Patterns (LTB).

LTP methods are texture features methods, which is introduced by Tan and Triggs.Local ternary pattern (LTP) operator extends LBT to 3- value codes (HSV).Local pattern for each pixel (i,j) the total image is represented by Histogram [11,27].

$$H_{LP}(l) = \sum_{i=1}^{N1}\sum_{j=1}^{N2} f_2(LP(i,j),l); l \in [0, 2^p - 1)]$$

1.10. Invariant Color Histogram.

One of the important calculated the color histogram, in this histogram are computed variable kernel density calculated. This histogram invariant to highlights, shading, and reflections [25, 26].

$$\mathrm{f(x)} = \frac{1}{n}\sum_{i=0}^{n} \frac{1}{\sigma(xi)} \, \mathrm{k} \, \frac{(x-xi)}{\sigma(xi)}$$

Here ,kernel K is a function satisfying ,∫ K(x)dx =1.The kernel centered on xi, has its own scale parameter $\sigma(x_i)$,i=1,2,….,n, which is scale parameter is a function of the RGB- values and the color space transform.

2. Texture Feature Extraction.

Texture feature extraction is very computationally intensive for individual pixels.it is one way that can be used to help in segmentation and classification of image. Texture features give us information about the spatial arrangement of color and intensities in an image. We have to discuss about some texture methods. Such as GLCM, wavelet transformation, Tamura features soon on [6, 7, and 9].

2.1 GLCM.

The Gray-level co-occurrence matrix (GLCM) extracts from the texture features [18]. They has four co-occurrence matrices of four different orientations (horizontal 0^0, *vertical* 90^0 and two diagonals 45^0 and 135^0) are constructed. The GLCM is created from a grayscale images. They have to construct Five higher order features (i)

Energy, (ii) Contrast,(iii) Homogeneity, (iv) Correlation, (v) Entropy can be derived from the gray level-occurrence matrix(GLCM) of the horizontal orientation, to give a five dimensional feature vector. They are followings equations.

$$\text{Contrast} = \sum_i \sum_j (i-j)^2 \, p(i,j)$$

$$\text{Entropy} = \sum_i \sum_j (i-j) \log P(i,j)$$

$$\text{Energy} = \sum_i \sum_j \; p2\,(i.j)$$

$$\text{Correlation} = \sum i \sum j \frac{(i-\omega i)(j-\omega i)p(i,j)}{ai\;aj}$$

$$\text{Homogeneity} = \sum_i \sum_j P(i,j)\,/1+|i-j|$$

The GLCM is a statistical way to describe texture features, it is calculated for each segment. Here p is denote matrix we can define a matrix p(i,j) that counts number of a matrix. Example consider one image I of dimensions MxN with x=1,2,3,…M and y=1,2,3,…N and position P stands for one elements to the right if ,i=I(x,y) and j= I(x+1,y) then the GLCM is P(i,j)=P(i,j)+1.

2.2 Energy Filters and Edginess factor.

This method is based on the calculation for each pixel of the filtered image [30], The mean and variance of the four square neighborhoods in which each pixel is a corner, and the final value for that pixel the mean of the neighborhood with the lowest variance, which is supposed to be more homogeneous and, consequently, should contain only one type of texture (no borders).The edginess factor is a feature that represents the density of edges present in a neighborhood [25, 31].

2.3 Gabor filters.

Gabor filters are directly related to Gobar wavelets in which is A set of Gabor filters with different frequencies and orientations may be helpful for extracting useful features from an image[28].They are several regions based on scalable color and homogeneous texture information using mixture Gaussian algorithm[25].

2.4 Tamura Features.

Tamura features are six features Tamura features which is Coarseness, directionality, line likeness, contrast and regularity, roughness, Coarseness. Directionality and contrast features are very strong and important also efficient features [29].

2.5 Wavelet transform method.

Wavelet transform techniques are time frequency which can be either low or high pass filter.it is analysis signals which is a function of time and frequency measured. This technique provides a robust methodology for texture analysis in different scales. Wavelet transform many methods such as Continuous wavelet transform (CWT), Discrete wavelet transform (DWT), FWT (Fast Wavelet Transform).

3. Shape Extraction.

Shape features are important features techniques, they are many approaches using the shape feature extraction [8, 10]. Shape features must have some essential properties that is identity ability, translation, rotation, scale invariance, noise invariance soon on .some shape features moments such as the Hu Moments invariants and invariant Zernike moments .This way can improve the CBIR system results.

3.1 Hu's moments set.

Hu moments discussed the translation and rotation with scaling.it is consist of nonlinear centralized moment's expression. We assume the image functions as f(x,y) and the standard two- dimension moments as.

$$m(u,v) = \iint_{-\infty}^{\infty} f(x,y)xy^{uv}dxdy$$

Here u, v denotes are positive integers that are the order of the moment.
The central moment is defined as,

$$cm = \iint_{-\infty}^{\infty} f(x,y)(x-\bar{x})^{u}(y-\bar{y})^{v}dxdy$$

Where $\bar{x} = \frac{M_{10}}{M_{00}}$ and $\bar{y} = \frac{M_{10}}{M_{00}}$ this description possesses the ability of data retention and shift invariance.

3.2 Invariant Moments.

This moment important and most frequency used in shape features descriptors. Zernike moments can be used as an effective descriptor of global features. The geometric moment function M_{pq} of order (x+y) is defined by the equation.

$$M_{pq} = \sum_x \sum_y x^p y^q x(x,y) \quad pq = 0,1,2....$$

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,\bar{y})$$

Where (x_c,y_c) is the center of the object.to obtain the ability of scale invariant ,the central moment should be normalized. **Where** $\varphi_{p\ q}= \varphi_{pq} / \varphi^r_{00}$, γ=p+q+r /2**.** Denote the components of the centroid moments? The following equations is canalized moments..

$\varphi_1 = \eta_{20} + \eta_{02}$

$\varphi_{2=}$ $(\eta_{20} + \eta_{02})^2 + 4\eta_{11}{}^2$

$\varphi_{3=}$ $(\eta_{30} - 3\eta_{12}) + (3\eta_{21} - \eta_{03})^2$

$\varphi_{4=}$ $(\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$

$\varphi_{5=}$ $(\eta_{30} + 3\eta_{12}) + (\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] +$
$(3\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + (\eta_{30} + \eta_{12})^2$

$\varphi_{6=}$ $(\eta_{20} - \eta_{02})^2 [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11} (\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$

$\varphi_{7=}$ $(3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2] - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} + 3\eta_{12}))(\eta_{30} + \eta_{12}) [(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$.

3.3 Histogram of Edge Direction.

Edge directions are quantized into a number of bins. They have 73 bins are there in edge histogram which is 73 bins are denoted the edges shapes in an images, first 72 bins are denotes in edge direction and last because 73 bins is used in non-edges pixels [1, 22].

4. Clustering Techniques.

Clustering is a collection of data objects that are similar to one another within the similarity clustering and dissimilarity. Clustering is the classification of objects into different groups or more precisely the partitioning of a data set into subsets. Many methods are there.1.K-means clustering 2.support vector machine (SVM) 3.Fuzzy Clustering 4.Hierarchical clustering and 5.Rank SIFT

4.1. K-means Clustering.

K-means clustering methods in which K-initial points are chosen to represent initial cluster centers. It is employs a greedy iterative approach to find a clustering that minimizes image visual gap. As such it can converge to a local optimal instead of a globally optimal clustering. They have many clustering techniques clustering algorithms adapted to Color, Texture, and Shape feature extraction.

4.2. Support Vector Machine method.

This is (SVM) re-ranking or clustering mechanism methods. SVM used in relevance feedback image retrieval methods which classification of positive and negative approach [25].relevance feedback mechanism is effective retrieval system that is computed the low and high level semantic features. Its main aiming is improving classification mechanism. There three steps 1.It maps the similarity between query images and database images.2.It is calculated the distance.3.It is form optimizations. Another one is SVM-based RF (Relevance feedback) methods with using probabilistic and weighted feature kernel function.

4.3. Rank-SIFT**.**

Rank-SIFT algorithm is the revised version of SIFT (Scale Invariant Feature Transform) algorithm which uses ranking techniques to improve the retrieval performance of the SIFT algorithm. K-Means can be applied on SIFT [13], SIFT is widely using the feature extraction method which is extraction from huge database images [5]. The Mutual KNN (k-nearest neighbor graph) Graph is defining neighborhood for each data paint which is using improving performance dataset [8].

4.5. Fuzzy Clustering.

Fuzzy clustering methods are generation of higher relevance clustering which belongs to more than one cluster. We have recently proposed a fast and exact HCM (Hard c-means) variant that is called Weighted Sort-means (WSM) that utilizes data reduction and increased nearest neighbor pixel values searched.

5. Re Ranking Methods.

Re ranking system is improving the text based image retrieval methods.Re-ranking based on the visual The most search engines is using the text-based image retrieval as textual information is sometimes retrieved is noise images and irrelevant images are retrieved. This problem is reduced using re ranking images represents to their visual information's.re ranking system is improving similarity images which considerations of an images similarity and diversity images that is effective and efficient system. Such as 1.semantic re ranking 2.tag based re-ranking 3.visual re-ranking 4.Bayesian re-ranking5.contextual re ranking 6. RL Sim re-ranking models so on. Visual re-ranking is combined textual and visual cues. Bayesian re-ranking is drives the best re-ranking images by maximizing visual information while minimizing irrelevant images reduced.re ranking is ordering the initial ranked list based on visual pattern that is called contextual based video and image search re-ranking.

6. Distance Measure Techniques.

Distance measure are similarity metrics, they are used for comparing the similarity of two images [3]. There are different kinds of similarity measure techniques like (i).Euclidean distance,(ii).Histogram intersection (iii).Murkowski distance (iv).Quadratic distance (v).Manhattan distance (vi).Mahalanobis(vi).Sum of absolute differences (vii).Sum of squared absolute differences (viii) city block differences (ix).Canberra distance (x).Maximum value differences and Minkowski distance [31].

6.1. Euclidean Distance.

This distance metric is mostly used for measurement in CBIR system [1],let q be query image and t be the target database image and let $P_q(Zq)$ and $P_t(z_t)$ be their respective probability densities [25,26].

$$D_{Euc}(q_i, t_i) = \sum_{i=1}^{n} (q_i - t_i)^2$$

6.2. Histogram Intersection.

In this metric for comparing histogram [14] .which is calculate the common part of the two histogram, the histogram intersection of two histogram H and H' is measures the following equation.

$$\mathbf{d_n(H, \bar{H}) = \sum_{m=1}^{m} min(H_m, \overline{H_m})}$$

6.3. City Block Distance.

This metric is called the Manhattan distance. The city block calculates the robustness to outliers this distance metric is computed by the sum of absolute between two feature vectors of images.

$$\Delta \boldsymbol{d} = \sum_{i=1}^{n} |\boldsymbol{Q_i} - \boldsymbol{D_i}|$$

6.4. Bhattacharya Distance.

It is measures the similarity of two discrete or continuous probability distributions. This metrics between the query image q and the target images t in the database. They are following equation. [26].

$$D_{Bhat}(q,t) = \frac{1}{8}(\mu_q - \mu_t)^T \left[\frac{\Sigma_q - \Sigma_t}{2}\right]^{-1} (\mu_q - \mu_t) + \frac{1}{2}\ln\frac{|(\Sigma_q - \Sigma_t)/2|}{\sqrt{|\Sigma_q||\Sigma_t|}}$$

Here $\boldsymbol{\mu_q}$,μ_t are the denotes the mean vectors, $\sum_{\mathbf{q}}$ and $\sum_{\mathbf{t}}$ are the covariance's of the image pixel values.

6.5. Sum of absolute difference (SAD).

This metric is very straightforward distance metric and extensively used for combining the distance between the images in CBIR to get the similarity.

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

Here n defined the number of features, i=1, 2, 3…n. Qi is query image and Di is target images.

6.7 Mahalanobis Distances.

This metric also called quadratic distance. This metrics between query image and database images which are measured the similarity metrics.

$$D_{Maha}(q,t)=(\mu_q-\mu_t)^T\sum^{-1}(\mu_q-\mu_t).$$

7. Dataset.

CBIR system is using a lot of image dataset which can allow the comparison of image retrieval systems under different kind of images.1.Corel Dataset 2. WANG Database 3.UW Database 4.IRMA-10000 Database 5.ZUBUD Database 6.UCID Database 7. Caltech Database 8. Holidays Dataset. [10, 22, 15, 20, 26].The Corel image database contains a very huge of images of various contents like That Beach, Roses, and Horses. Several Database contains natural images and it has 101 different objects categories e.g. The Olivia Dataset, Flickr 18 and IAPR TC-12 Benchmark also contains natural image datasets. Corel datasets consist of 1000 images with under 10 different categories [5].UKBench has a total of 10,200 images. WANG Database subset of 1000images of the Corel photo database. Also ZuBuD dataset [15, 8, 4].

8. Performance Measures.

The CBIR systems are using several performance metrics. This is a help to find retrieval measured. Generally, we have using two metrics one is precision another one is Recall measurement. We have to discuss some important metrics formula. It mostly using CBIR system the following the equation [13, 18, 22].

Table.1 .***Comparison of the CBIR methods used in recent papers.***

S no	Paper	Feature Extraction Methods	Distance Measure Methods.	Clustering Techniques	Advantage	Dis Advantage
1.	Improving content – based image retrieval with compact global and local multi features[et al Ahmad Alzu bi, Sep2016]	HSV color histogram and color moments with chromaticity moments and GLCM and wavelet moments with LBP, GIST, SIFT LBP, HOG.	Euclidean Cosine, correlation, Manhattan	Not used.	This system using lot of techniques used. Combined Local and Global features method, HSV, SIFT, GLCM, SURF, LBP, HOG.	It takes more matching times. Which is not using re ranking techniques.
2.	SIMIR:New Mean SIFT color multi clustering image retrieval.[et al Hadjer Lacheheb Feb 2016]	HSV and SIFT two feature methods are used.	LIRE and FIRE two image retrieval methods are used.	K-means and Multi clustering.	We have using multi clustering Techniques and multi search approaches techniques are used.	The system not using more measurement techniques.
3.	Hard versus fuzzy c-means clustering for color quantization [Quan Wen et al 2011].	They have two types of clustering techniques are used. Such asK-means and fuzzy means. .(Median cut, WAN ,Oct,Wu)	Mean absolute error (MAE) and Mean squared error (MSE).	Five data set they are Hats, Motocross, flowers and sill, parrots.	This System New method used, HCM better than FCM.Because very fast.	In this system don’t using huge dataset. Not using Texture and shape Features Extraction.

4.	Content based image retrieval system using clustered scale invariant feature transform.[et al Gholam Ali montazer 2015].	SIFT Algorithm K-means Algorithm	Euclidean distance Measure.	K-means clustering Techniques.	We introduce comparative between new method and existing method. Result efficiency Is 92.96	It is compared only autocorrelations and wavelet transform.
5.	Similarity measure for image re sizing Using SIFT user.[et al Shungang 2012].	The research approach using combining Seam Carving Scaling algorithm. SIFT feature method also used.	Euclidean distance Measure.	Not Used.	This approach is using effective image resizing algorithm used.	We could not get original features.
6.	Regional –based image retrieval using semantic clustering matrix and adaptive learning.[et al Felci Rajam 2012].	In this system proposed for finding low and high level features using semantic cluster matrix (SCM).	Euclidean distance, Bhattacharya distance, Mehalanobi Distance.	SVM(Support vector Machine) Semantic Clustering.	We have using effective system proposed. They have three methods as such semantic cluster.	Does not using re-ranking techniques.
7.	Fusion at Features Level in CBIR system Using Genetic Algorithm.[et al Chandrasekhar 2013]	This system denotes for images features. Which can form combined feature vector. Which is genetic algorithm. The methods are Edge Histogram, GCLM,GA	New measurement is cumulative match score, CMS.	KNN Classifier (K-Nearest Neighbor Algorithm).	Genetic Algorithm based on feature extraction. We can get original results.	This method is Old techniques. Also Shape Feature method is not used.
8.	Remotely sensed image retrieval based on region-level semantic mining.[et al Tingting Liu 2012]	This paper using GLCM, and mean feature methods are used .It is find high and low level features.	They are used on NDVI, NDBI Euclidean distance.	Image Smoothing.	The research proposed for high level significance.as such Mean, variance.ASM.	Does not using Shape feature extraction.
9.	A new SVM based relevance feedback image retrieval using probabilistic feature and weight kernel function [et al Xiang-yang Wang 2016].	PCA-Principle component analysis. AGMM, Global color Histogram (GCH), Pyramidal wavelet transform (PWT), YCrCb color space.	New method Kullback leibler divergence metrics. Also relevance feedback.	They have using SVM based RF clustering techniques.	This methods is using new features weights (PCA).lot of database set are used.	Not clearly mentioned re ranking methods.
10	Tag-Based Image search By social Re-Ranking[et al Dan Lu 2016]	Mean, standard deviation, skewness. HSV color space.HWVP-Hierarchical wavelet packet descriptor.	VR-View based re ranking. VUR, RR Relevance feedback .CRR, DRR, SR-Social re ranking.	Not used.	Very effective, to enhance the results and time save method.	It does not use more dataset.

8.1NDCG (Normalized Discounted Cumulative Gain).

We have similar the Normalized Discounted Cumulatively gain which is used in image retrieval [1, 3, 4, 5]. When there are more than two relevance levels. Given a ranked list, The NDCG at the depth *l* is defined as.

$$NDCG@l = Z_l \sum_{i=1}^{l} \frac{2^{r^i} - 1}{\log(1+i)}$$

Where r^i ,is relevance score of the i^{th} image, and Z_l is a normalized constant to guarantee that a perfect ranking ‘s NDCG@*l* is equal to 1.

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

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

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

Conclusion.

Our aim is to provide overview survey of research details on CBIR system which various kind of properties that is uses Color, Texture, Shape and clustering models, also We have been discussed combined Distance Metrics (DM) and clustering with Re-ranking methods. Our main techniques focus on reducing the visual information’s between Query image and database images. Our aiming results demonstrate that the proposed method based on color, Texture, Shape features of image sub-blocks has better retrieval performance. As further studies, Distance metrics (DM) and re-ranking also clustering methods. Additionally we have discussed some data set and evaluation metrics.

References.

1. Daniel Carlos,Guimaraes Pedronette,Ricardo da s.Torres.**Combining Re-Ranking and rank aggregation methods for image retrieval**.***Multimedia Tools***.2016 Apr,75,pp.9121-9144.
2. Xiaoou Tang,Ke Liu,Xiaoou Tang.**Query Specific Visual Semantic Spaces for Web Image Re-Ranking**.***IEEE Transaction***.*2014, pp.857-864*.
3. Hadjer Lachehed Saliha Aouat.**SIMIR:New mean SIFTcolor multi-clustering image retrieval**. ***Mulmedia Tools Appl***. 2016 Feb,pp:11042-015-3167-3.
4. Daniel Carlos,Guimaraes Pedronette,Ricardo da s.Torres.**A Correlation grapg approach for unsupervised manifold learning in image retrieval tasks**.***Neurocomputing.***2016,pp:66-79.
5. Gholam Ali Montazer,Davar Giveki.**Content based image retrieval system using clusterd scale invariant feature transforms**.***Optik*** .2015,pp:1695-1699.
6. LakkovidouC,N.Anangostopolos,A.Kapoutsis,Y.Boutalis,M.Lux and S.A.Chatzichristofis.**Localized global descriptors for content-based image retrieval.*EURASIP Journal on Advances in signal processing***.2015,Vol.80.
7. Dan Lu,xiaoxiao Liu,Xueming Qian .**Tag-Based Image Search by social Re-Ranking**.***IEEE Trans Syst on multimedia***.Aug 2106, Vol.18,No .8 .
8. Santhosh Kumar vipparthi,Shyam Krishna Nagar.**Color Directional Local Quinary Patterns for Content Based Indexing and Retrieval.*Human-centric Computing and information Sciences*** .2014,pp:4:6.
9. Songhe Feng, Hong Bao,Congyan Lang,De Xu.**Combining visual attention model with multi instance learning for tag ranking**.***Neurocomputing.***2011,Vol.74,pp:3619-3627.
10. Shungang Hua,Guopeng Chen,Honglei and Qiuxin Jiang.**Similarity measure for image resizing using SIFT feature.*EURASIP Journal in image and video processing***.2012,vol.6.
11. Eugene Santos Jr.Qi Gu.**Automatic content based image retrieval using semantic analysis.*J Intell Inf Syst***.2014,Vol,43,pp:247-269.
12. Xiang-Yang wang,lin-Linliang,Wei-Yi-LiDong-ming li,Hong-Ying yang.**A new SVM –based relevance feedback image retrieval using probabilistic feature and weighted kernel function**.***J.Vis Communication image***.2016,Vol.38,pp: 256-275.

13. Bichen zheng,sangwon yoon ,sarah s.lam.**Breast cancer diagnosis based on feature extraction using hybrid of K-means and support vector machine algorithms.*Expert System with Applications.***2014,Vol.41.pp:1476-1482.
14. L.Koteswara Rao,D.Venkata Rao.**Local quantized extrema patterns for content-based natural and texture image retrieval.*Hum.cent.comput.inf.sci***.2015,pp:5:26.
15. Quan Wen,M Emre Celebi.**Hard versus fuzzy c-means clustering for color quantization.*EURASIP Journal on Advances in Signal Processing*** .2011,118.
16. Konstantin Markov and Tomoko Matsui.**High level feature extraction for the self-taught learning algorithm.*EURASIP Journal on audio,speech,and music processing*** *.2013,Vol.6.*
17. Ting Yao,chong-Wah Ngo and To Mei.**Circular Reranking for Visual Search.*IEEE Trans.Syst.*** *Aprl 2013,Vol.22,No.4 , pp.1644-1655.*
18. Chandrashekar Patil G,Mahesh Kolte T and Devendra Chaudhari S.**Fusion Features Level inCBIR System Using Genetic Algorithm**.***SEMCCO*** *,2013,Part II,LNCS* 8298 pp. 156-167.
19. Shaowei Liu,Peng Cui, Huanbo Luan,Wenwu Zhu,Shiqiang Tang and Qi Tian.**Social Visual Image Ranking for Web Image Search.***2013,Part I,LNCS 7732 pp*. 239-249.
20. Seetharaman K, Sathiyamoorthy S.**A unified learning framwork for content based medical image retrieval using a statistical model.*Computer and information Science***.*2016,Vol.* 28,pp:110-124.
21. Fazal Malik,Baharum Baharudin.**Analysis of distance metrics in content-based image retrieval using statistical quantized histogramtexture features in the DCTdomain.*Computer and information Science***,*2013,Vol.2*5,pp:207-218.
22. Gunhan park,Yunju Baek,Heung-Kyu Lee.**Re-ranking Algorithm using post-retrieval clustering for content-based image retrieval.*Information Processing and Management***.*2005,Vol.41,pp:177-194.*
23. Theo Gevers and Arnold W.M Smeulders .**PicTOSeek:Combining Color and Shape Invariant Features for Image Retrieval.*IEEE Trans.,Syst.*** *Jan 2000,vol.9,No.1.pp*.102-119.
24. Le An,Changjian Zou,Liyan Zhang,Bradley Denney.**Scalable attribute-driven face image retrieval.*Nuerocomputing***.*2016,vol.172,pp: 215-224.*
25. Ahmad Alzu bi,Abbes Amira ,Naeem Ramzan,Tareg Jaber.**Improving content –based image retrieval wiith compact glopal and local multi –features**.***Int J Multimed Info Retr***.2016.
26. Celia and I.Felci Rajam.**An Effective Content Based Image Retrieval Framework Using Machine Learning Techiniques.*ICDEM***.*2010,LNCS* 6411,pp.162-169,2012.
27. Swati V. Sakhare & Vrushali Nasre G.**Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture**.*International Journal of Computer Science & Informatics, 2011,Vol-I, Issue-II, pp* 57-61.
28. Chandan Singh,Neerja Mittal and Ekta Walia.**Complementary Features sets For optimal face recognition.*EURASIP Journal on image and video Processing*** **.***2014,vol.35.*
29. Bulent Cavusoglu .**Multiscale texture retrieval based on low-dimentional and rotation-invariant features of curvelt transform.*EURASIP Journal on image and video processing***. *2014:22.*
30. Tania Di Mascio , Luigi Laura, and Valeria Mirabella:**The Interface of VISTO, a New Vector Image Search Tool .*Human Interface,part I,HCII*** *.2007,LNCS 4557,pp*:417-426.
31. Xin Qi,Daihu Wang,Ivan Rodero,Javier Diaz-Montes,Rebekah H Gensure,Fuyoung Xing,Lin Yang .**Content –based histopathology image retrieval using CometCloud.***BMCBioinformatics* 2014,vol.15,pp:287.
32. Tingting Liu,Liangpei,Pingxiang Li,Hui Lin.**Remotely sensed image retrieval based on region-level semantic mining.*EURASIP Journal on image and video processing.*** 2012,vol.4.
33. Tasnim Sami,Nabeel Mohammed,Sifat momen.**Learning"initial feature weights"for CBIRusing query augmentation.*Int J Multimed Info Retr***,2016,vol.5,pp:125-132.
34. Jie Guo,Bin Song,Fang Tian,Hao Qin.**Texture classification with cross-covarience matrices in compressive measurement domain.***SIViP 2016,vol 10,pp:1377-1384.*
35. Shao-Hang Kao,Wei-yen Day,and Pu-Jen Cheng.**An Aesthetic – Based Approach to Re-Ranking Web Images.*AIRS*** *LNCS 2010,pp.*610-623.