VOL 2 (2018) NO 1

e-ISSN : 2549-9904
ISSN : 2549-9610

INTERNATIONAL JOURNAL ON INFORMATICS VISUALIZATION

A New Approach Towards Image Retrieval Using Texture Statistical Methods

R.Tamilkodi[#], G.Rosline Nesa Kumari[*], S.Maruthu Perumal[**]

[#] *CSE Department, Saveetha University, Assoc.Prof.GIET, Rajahmundry-533296, India*
** Saveetha University, Tandalem, Chennai- 600077, India*
[**] *CSE Department, NBKR Institutes of Science and Technology, AP, India*
Email : tamil_kodiin@yahoo.co.in, rosemaruthu@gmail.com

***Abstract*— Texture is a possession that represents the facade and arrangement of an image. Image textures are intricate ocular patterns serene of entities or regions with sub-patterns with the kind of brightness, color, outline, dimension, and etc.This article proposes a new method for texture characterization by using statistical methods (TCUSM). In this proposed method (TCUSM) the features are obtained from energy, entropy, contrast and homogeneity. In an image, each one pixel is enclosed by 8 nearest pixels. The confined in turn for a pixel can be extracted from a neighbourhood of 3x3 pixels, which represents the fewest absolute unit. We used four vector angles 0, 45, 90,135 to carry out the experimentation with the query image. A total of 16 texture values are calculated per unit. Compute the feature vectors for the query image by calculating texture unit and the resultant value is compared with the image database. The retrieval result shows that the performance using Canberra distance has achieved higher performance.**

***Keywords*— Texture, Color, Statistical, Energy, Entropy, Contrast, Homogeneity, Retrieval.**

I. INTRODUCTION

Texture analysis is the fundamental task used in the content based image recovery which specifies the properties of image content. It contains the significant information about the structural plan of surfaces and their bond to the adjoining surroundings. In general, texture can be defined as a usual replication of elements or patterns on a surface. The repetition frequencies results in textures can appear to be random and unstructured. A number of techniques have been used for measuring the texture features such as Gabor filter, fractals, wavelets, cooccurence matrix etc.

Texture has many diverse extents; there is no solo way of texture depiction that is ample for an assortment of textures. It refers to a class of statistical measures and models that exemplify the spatial variations surrounded by descriptions resources of extracting information. Texture is an aerial erect that defines local spatial association of spatially altering ethereal principles that is recurring in a section of outsized spatial position. Since an image is completely in the form of pixels, texture can be distinct as a unit consisting of equally associated pixels and collection of pixels. This cluster of pixels is called as texture primitives or texture essentials.

In the early 70's Haralick et al [8] anticipated cooccurence matrix representation of texture feature. This approach explored gray level spatial reliant of texture.Tamura et al [5] explored texture depiction from unlike angle and projected a computational estimation on six visual properties like coarsness,disparity,directionality,linelikeness,constancy and unevenness. The QBIC system and MARS system further improved Tamura's texture depiction [1]. In the early 90's the wavelet transform was introduced for texture representation.

In texture analysis, the most important task is to extract a texture feature which specifies textual characteristics of the original image [11-13]. The texture feature of an image is extracted by mean and variance of the wavelet sub bands. But wavelet [2-4] loses their universality in capturing the edge discontinuities in image which is important in texture representations. This proposed method (TCUSM) overcomes the weakness of conventional wavelets to obtain images with less computational complexity.

The main aim of the paper is to retrieve the images accurately with the combined texture features based on orientation, directionality and regularity. The characteristics features of texture are investigated in classification experiments with images taken from Brodatz album [15].The classification accuracy rates are compared with other methods and the results are found to be improved significantly.

This paper is organized as follows: section 2 explains the model behind the work proposed and the methods of detecting the characteristics of texture. Section 3 with proposed set of texture images and the experiments shown with Corel Dataset. Finally, the conclusion about proposed method has been highlighted.

II. Image Retrieval Using Proposed Method (TCUSM)

The goal of this work is to study, if the values of the features describing a texture are the same or very similar with the query image compared with the database image are retrieved. Before we are going to say about our work, we want to know the terms used in our methods. Let us see about those methods how it extracts the feature form the images. Figure 1 shows the block diagram of our proposed method TCUSM.This projected method works by supplying the query image which is fed for feature extraction through various statistical methods like contrast, homogeneity, energy and entropy. Once the value of texture unit if found a comparison is made with the image database by applying the same statistical methods to find the value of the texture spectrum.Simliarity measures are carried out to find the distance measure between the two textures unit obtained and if it shows equal then the corresponding image can be retrieved.

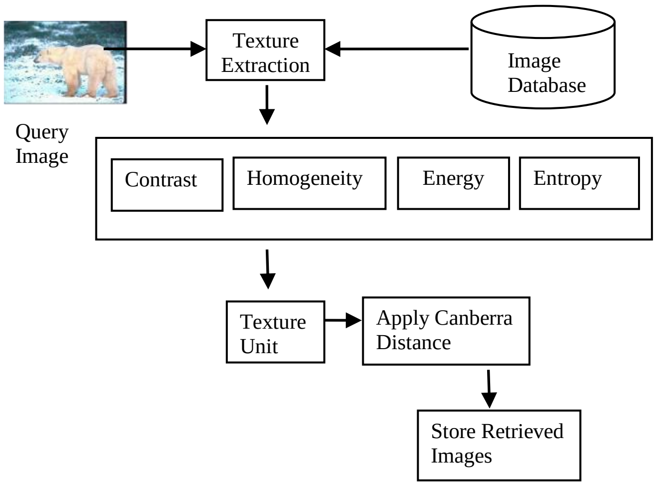


Fig.1.Block Diagram of proposed method

A. *Cooccurence Matrix*

The gray level cooccurence estimates image properties associated to information which considers the correlation along with pixels or group of pixels. Haralick [8] recommended the use of grey level cooccurence matrices which became the famous and widely used texture features. It is based on the joint probability distributions of pairs of pixels. It shows often the grey level occurs at a pixel located at a fixed geometric position relative to each other pixel, as a function of the grey level.

The following figure shows the 3x3 image and its gray level cooccurence matrices. The number of threshold levels is 4.Consider the image as shown in Fig.2. if we use the position operator "one pixel to the right and one pixel down then we get the gray level cooccurence matrix as shown. The 2 in the cooccurence matrix indicates that there are two occurrences of a pixel with gray level 1 immediately to the right of the pixel level as well as the down of the pixel level 1.

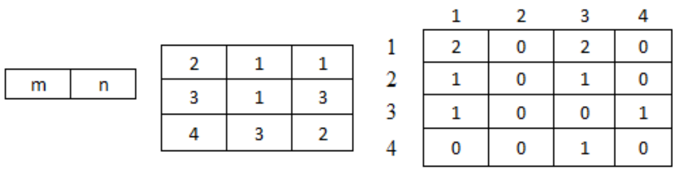


Fig.2. working out of Cooccurence matrix

The grey level co occurrence matrix calculates how frequently a pixel with gray level value m occurs flat, perpendicularly or crosswise to nearby pixels with the value n. The size of cooccurence matrix will be the number of threshold levels. When we consider neighboring pixels, the distance between the pair of pixels is 1. On the other hand, each altered relative position between the two pixels to be compared creates a different co-occurrence matrix.

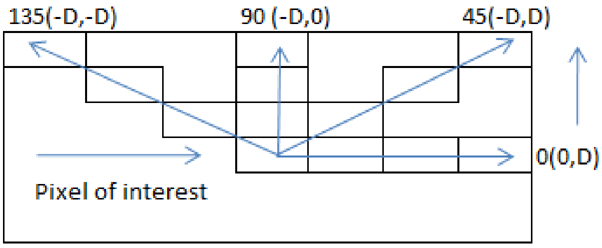


Fig.3.Directional analysis P_0, P_{45}, P_{90}, P_{135}

Statistical features of grey level were one of the most primitive methods worn to categorize textures.Haralick [8] recommended the use of grey level cooccurence matrices (GLCM) to take out second order statistics from an image.GLCMs have been used very effectively for texture categorization in evaluations.

Matrix element P(m,n) in a grey level cooccurence is 2.Order probability of varying grey level m to n when affecting distance d in the direction ɵ of the image, or equivalent (Δi, Δj).

From a X xY image with G grey levels and f(x,y) is the intensity, then

$$P(m,n|\ \Delta i, \Delta j)=ST(m,n|\ \Delta i, \Delta j),$$

$$\text{where}\quad S = \frac{1}{(x-\Delta i)(y-\Delta j)},$$

$$T(m,n|\ \Delta i, \Delta j) = \sum_{X=1}^{(y-\Delta j)} \sum_{Y=1}^{(x-\Delta i)} C$$

$$\text{and}\quad C=\begin{cases} 1 \text{ if } f(x,y) \text{ where } m = f(x+\Delta i, y+\Delta j)=n \\ 0 \text{ else} \end{cases}$$

Alternative notation depends on direction and distance (m,n | d, ɵ).

B.*Contrast*

Contrast is a local gray level disparity in the grey level cooccurence matrix. It can be contemplated as a linear dependency of grey levels of neighbouring pixels.

$$\text{Contrast} =\sum_{m,n} |m- n|^2\, p(m,n)^2$$

where m & n are the horizontal and vertical coordinates and p is the pixel value. If the neighbouring pixels are very similar in their grey level values then the dissimilarity in the image is very short. The grey level variations show the distinction of texture itself. It measures the amount of variation flanked by the maximum and smallest value of a position of adjoining pixels.

C. Homogeneity

It assumes big values for minor gray tone differences in couple elements. Homogeneity process the regularity of the non-zero entries in the cooccurence matrix. It weights values by the inverse of dissimilarity weight.

$$\text{Homogeneity} = \sum_{m,n} \frac{P(m,n)}{1+|m-n|}$$

It is more sensitive to the presence of near diagonal elements in the matrix. It has maximum value when all elements in the image are equal. In grey level cooccurence matrix contrast and homogeneity are strongly but inversely correlated in terms of equivalent distribution in the pixel pairs. It means homogeneity decreases if contrast increases while energy is kept constant. Therefore high homogeneity refers to textures that contain ideal repetitive structures while low homogeneity refers to big variation in both texture elements and their spatial arrangements. The range of homogeneity is equal to [0,1].if the image has little variation then homogeneity is high equal to 1 and if there is no variation then we can say that the homogeneity is low equal to 0.

D. Energy

It measures the uniformity of texture that is pixel pair repetitions. It detects disorders in textures. Energy reaches a maximum value that equal to one. The higher the energy value, the bigger the homogeneity of the texture.

$$\text{Energy} = \sum_{m,n} P(m,n)^2$$

The range of energy is [0,1] where energy is 1 for a constant image. All the features discussed are connected in certain manner. Contrast calculates the variation of grey level pairs but with a different weight. Homogeneity weights values by the inverse of contrast weight, which means lower the homogeneity, higher the contrast. Energy is actually local homogeneity and entropy is the opposite of energy.

E.Entropy

It is a measure of randomness, having its highest value when the elements of P are all equal. In the case of a checkerboard, the entropy would be low.

$$\text{Entropy} = \sum_{m,n} P(m,n) \log P(m,n)$$

III. Functioning Theory Of Projected Technique TCUSM

First a query image is fed in to the system for the conversion of RGB into grey level image. As from the characteristics of Texture we can able to find out all the statistical methods and the resultant is stored with texture unit. In this proposed method (TCUSM) the features are obtained from energy, entropy, contrast and homogeneity. In an image, each pixel is surrounded by 8 neighbouring pixels. The local information for a pixel can be extracted from a neighbourhood of 3x3 pixels, which represents the smallest complete unit. We used four vector angles 0, 45, 90,135 to carry out the experimentation with the query image. A total of 16 texture values are calculated per unit. Compute the feature vectors for the query image by calculating texture unit and the resultant value is compared with the image database.

Steps involved in proposed method TCUSM:

1. Read the Query Image
2. Convert into gray level
3. Apply Characteristics of texture
4. Find the value of Texture Unit
5. Repeat the steps 1 to 4 for database image.
6. Compare the Texture Unit value with the database image value.
7. If both values are similar retrieve the images.
8. Stop computation.

IV. Performance Evaluation

To evaluate the retrieval efficiency of the proposed system, we use the performance measure calculation by using the formulas for recall and precision. The one which computes the system to retrieve all the images which is related to the query image supplied where as precision computes the aptitude of the system to retrieve only the images that are exactly matches with the query image**.**

Using precision and recall to measure the accuracy of retrieval system is still the most prominent technique.

$$\text{Precision} = \frac{R_r}{T_r}$$

$$\text{Recall} = \frac{R_r}{T}$$

Where

R_r -Represents the number of appropriate images retrieved.

T_r -Represents the entire images retrieved

T-Represents the total appropriate images retrieved.

Precision and recall results are tabulated in table 1.

Table 1
Precision And Recall Values Using Proposed Method (TCUSM)

Category of Images	**Average Precision**	**Average Recall**
Autumn	0.898	0.255
Dog	0.956	0.301
Bear	0.841	0.269
Ship	0.921	0.248

Flowers	0.937	0.355
Train	0.895	0.278
Castle	0.861	0.299
Cloud	0.901	0.298
Cat	0.954	0.247
Horses	0.869	0.248
Average (%)	0.9033	0.2798

Graph representing the performance analysis of texture feature using TCUSM

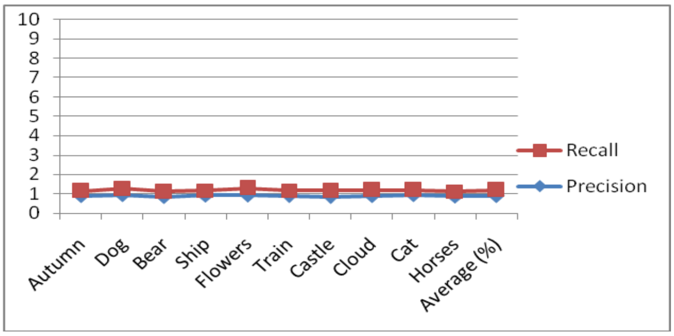


Fig.4 Graph represents the average precision and recall

V. Similarity Measure

A query image is any one of the images from image database. This query image is processed to compute the feature vector. Distance metrics are calculated between the query image and every image in the database. This process is repeated till all the images in the database have been compared with the query image. After completing the distance algorithm an array of distances is obtained and which is then sorted.

The similarity functions seek calculates the content difference between two images based on their features. One of the images is given as search parameter and another is stored in the database and had their features previously extracted. There are different ways to show the results for users, the most common is use the ranking method and present thumbnails according to the similarity degree in relation to a query. The judgement of how similar a database image is to a query is dependent on which image distance measure or measures are used to judge similarity. Calculate Canberra Distance to get the similarity distance between images. In the present work we have carried out Canberra similarity distance measure described as below.

A. Canberra Distance

It is the sum of absolute values of the differences between ranks divided by their sum, thus it is a weighted version of the distance. It examines the sum of series of a fraction difference between the coordinates of a pair. Each term of fraction difference has value lies between 0 & 1. The Canberra Distance itself is not between 0 & 1. If one of the coordinator is 0, the term become unity regardless the other value, thus the distance not be affected. Note that if both coordinate are 0's, we need to be defined as $\frac{0}{0} = 0$. This distance is very sensitive to a small change when both coordinate near to zero.

$$D_{mn} = \sum_{k=1}^{i} \frac{|x_{mk}-x_{nk}|}{|x_{mk}|+|x_{nk}|}$$

For finding the calculation input the values of query image compare it with Database image and store the resultant. From the resultant values we can able to know the matching relevant images.

VI. Experimental Results

We followed an approach to carry out an assessment and modifications to the retrieval of images based on the features experienced by means of a cautiously particular division of the Corel image dataset and the vector space likeness compute. We selected random sample of images from the Corel collection to carry out the distance vector calculation for the retrieval process. We have categorized the images in various forms like Autumn, Bear, dog, horse, castle, cloud, train, ship, cat, Flowers, Ship etc.We reorganised the Corel sample dataset as many images with similar concepts were not in the same group and some images with different images were in the same group in the original database, each group includes more than 100 images and the image group are category-homogeneous. These concept groups were used in the evaluation of the results of our technique.

For each feature we evaluated performance in the configuration described to improve performances were devised and evaluated. The general themes considered were how best to represent an entire image, how to accommodate differing sizes and scale of images and how to cope with the regional qualities of textures. These evaluations were run on the Corel data. The best performing features from the initial evaluation were then tested on the image data set. Tests were run with each texture feature combined with a high performing colour feature.

Fig.5.Sample query images from Corel Database

Query Image: Bear

Fig.6.Retrived images after applying proposed method (TCUSM)

The retrieval results for bear is illustrated in Fig.5.The distance metrics used is Canberra distance. when an image is addressed; a searching process will retrieve the closest image s in the database. The retrieved images are displayed according to their distance with the image query. The image that indicating perfect match will be retrieved first. Table 1 depicts the retrieval results for the sample images taken from the Corel dataset. The average precision is 8.172 and the average recall 2.821.

Proposed method we have used four statistical measures namely energy,entropy,homogeneity and contrast which shows proposed method is efficient than DWT-GLCM Methods. Canberra distance measures gives better retrieval results both in the proposed methods. Average precision is calculated and the graph is drawn between Average precision and recall as shown in Fig.4.The method of image retrieval is illustrated in the block diagram Fig.1.In this study the image query and image database that are represented in the RGB color space is converted to gray. Firstly all images are stored to the database. An image in the database is to be considered as a match to a query if the distance of feature vector image in the databases and the distance between feature vector of a query image is equal to zero. The retrieval result is based on the similarity computation.

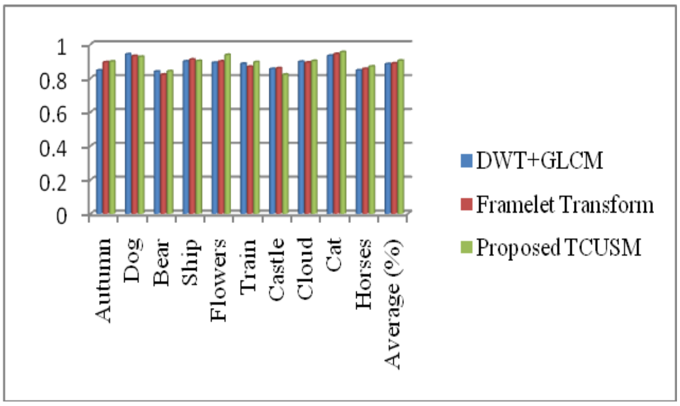
A. Assessment with other methods

The comparison of the proposed methods is carried with S.Sulochanan et.al [14], Metty Mustikasari et.al [10]. GLCM using Image sub blocks and DWT retrieval methods are compared with the proposed statistical method TCUSM.

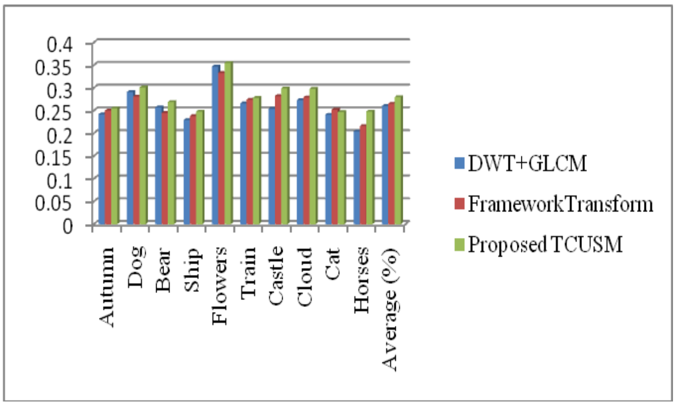
TABLE 2.
COMPARISON OF PROPOSED TCUSM WITH OTHER METHOD

Category of Images	DWT+GLCM		Framelet Transform		Proposed TCUSM	
	Precision	Recall	Precision	Recall	Precision	Recall
Autumn	0.845	0.242	0.894	0.25	0.898	0.255
Dog	0.941	0.291	0.931	0.281	0.956	0.301
Bear	0.839	0.257	0.821	0.245	0.841	0.269
Ship	0.899	0.229	0.91	0.238	0.921	0.248
Flowers	0.891	0.347	0.899	0.333	0.937	0.355
Train	0.885	0.266	0.867	0.274	0.895	0.278
Castle	0.854	0.254	0.859	0.282	0.861	0.299
Cloud	0.897	0.273	0.892	0.279	0.901	0.298
Cat	0.932	0.241	0.943	0.252	0.954	0.247
Horses	0.846	0.204	0.855	0.216	0.869	0.248
Average (%)	0.8829	0.2604	0.8871	0.265	0.9033	0.2798

Assessment chart with Precision



Assessment chart with Recall



VII. CONCLUSION

The search for the relevant information in the large database has become more challenging. In this paper a new proposed method TCUSM was framed for retrieval of images from a large database which gives better results and higher precision when compared to other methods. In this paper we have used four different features are obtained from an image based on energy, entropy, contrast and homogeneity. The proposed method TCUSM generally reduces the computational time and at the same time increases the user relations and also the exactness is better level as the images are retrieved on the basis of pixel in turn.

REFERENCES

[1]. Peter Howarth and Stefan Rijger,"Evaluation of texture features for content based image retrieval" Springer 2004.

[2]. Bino Sebastian V.A.Unnikrishnan and Kannan Balakrishnana,Grey level co occurrence matrices: generalisation and some new features.IJCSEIT,Vol.2,No.2,April 2012.

[3]. B.Vijayalakshmin,V.Subbaih Bharathi,"A novel approach to texture classification using statistical feature"IJACT 2013.

[4]. G.N.Srinivasan,Shoba G.,"Statistican Texture Analysis", Proceedings of world academy of science ,engineering & technology vol.36 December 2008 ISSN 2070-3740.

[5]. S.Sulochana,R.Vidya,"Texture based Image retrieval using Framelet transform-gray level Cooccurence matrix(GLCM)", IJARAI Vol.2,No.2,2013.

[6]. K.Arthi,J.Vijayaraghavan,"Content based image retrieval algorithm using colour models",international journal of advanced research in computer and communication engineering vol.2,issue 3,March 2013.

[7]. Amanbir Sandhu,Aarti Kochhar,"Content based image retrieval algorithm using Texture, Color and shape for image analysis",IJCT,Vol.3,No.1,Aug 2012.
[8]. Haralick, R.M.,Watson,L.A facet model for image data .Comput.Vision Image process 15,113-129,1981.
[9]. S.Nandagopalan,B.S.Adiga,N.Deepak,"A universal Model for Content based image retrieval"IJECE,Vol.4,2009.
[10]. Metty Mustikasari,Sarifuddin Madenda,"Texture based image retrieval using GLCM and image sub block",IJARCSSE-International journal of advanced research in computer science and software engineering.Vol.5,Issue 3,March-2015.
[11]. Gerald Schaefer ,"Content based image retrieval",Computer vision: Mar 2013.
[12]. Michael Unser: Splines – A Perfect Fit for Signal and Image Processing, IEEE Signal Processing Magazine, November 1999, pp. 22-38.
[13]. A.Vadivel, S. Sural, and A. K. Majumdar, "An integrated color and intensity co- occurrence matrix," Pattern Recognit. Lett., vol.28, no. 8,pp. 974–983, Jun. 2008.
[14]. George Wolberg: Digital Image Warping, IEEE Computer Society Press, Los Alamitos, California, 1990.
[15]. Piyush Kothyari, Shriprakash Dwivedi, "Content based image retrieval using statistical feature and shape extraction", International Journal of innovative research in computer and communication Engineering, Vol.4, Issue 6, June 2016.
[16]. Abdolreza Rashno , Saeed Sadri," Content-based image retrieval with color and texture features in neutrosophic domain", 978-1-5090-6454-IEEE.2017.
[17]. M.Kamarasan,"Content-based Color Image Retrieval Based on Statistical Methods using Multiresolution Features", thesis awarded in Annamalai University, June, 2014.
[18]. S.Sulochana,R.Vidhya,"Texture based Image retrieval using frame let transform-Gray level cooccurence Matrix(GLCM)",IJARAI (International journal of advanced research in artificial intelligence,Vol.2,No.2,2013.
[19]. D. Tao, X. Tang, X. Li, and Y. Rui, "Kernel Direct Biased Discriminant Analysis: A New Content-based Image Retrieval Relevance Feedback Algorithm," IEEE Transactions on Multimedia (TMM), vol. 8, no. 4, pp. 716-727, August 2006.
[20]. J. Z. Wang, J. Li, and G. Wiederhold, "Simplicity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 23, no. 9, pp. 947-963, September 2001.
[21]. Giang Truong Ngo, Tao Quoc Ngo, Dung Duc Nguyen," Image Retrieval with Relevance Feedback using SVM Active Learning, ISSN: 2088-8708 Vol 6, No 6 December 2016
[22]. Baopeng Zhang, Hangzai Luo, Jianping Fan ,"Statistical modelling for automatic image indexing and retrieval" Neurocomputing,Volume 207, 26 September 2016.