

Image Clustering by Neural Network (SOM) using Contents of Color and Texture

Divakar Singh BUIT, BU Bhopal divakar.singh@buit.ac.in **R. C. Jain** SATI Vidisha dr.jain.rcd@gmail.com

Abstract - The rapid development in computer technology for multimedia databases, digital media results in increase in the usage of digital images. Vast amount of data can be hidden in the form of digitized image. Image mining is used to extract such kind of data and potential information from general collections of images. Image Clustering groups the images into classes of similar images without prior knowledge. CBIR has extensive potential applications. Visual content of still images are used by CBIR to search for similar images in large scale. Thus the search for the relevant information in the large space of image and video databases become more challenging and interesting too. This paper discuss the concept of image clustering by self-organizing map (SOM) using the contents color and texture as image features for improving user interaction with image retrieval systems. The visual content of an image is analyzed in terms of low-level features extracted from the image. For color feature extraction, HSV color model and texture Gabor filter is presented.

Keywords - ANN (Artificial neural network), SOM (Self-Organizing Map), color, color feature extraction, color histogram, HSV color model, Gabor Filter, texture.

I. INTRODUCTION

Data clustering is an interesting approach for finding similarities in data and putting similar data into groups. Clustering partitions a data set into several groups such that the similarity within a group is large than that among groups[1]. Whenever the large amount of data is presented then it will be helpful to summarise the huge data into group of data and or clusters of small group in order to understand it and analyze it efficiently. Making clusters of images that produces same behaviour is a difficult task for human being to do it manually with large amount of image data. Image clustering by SOM uses the contents of image such as color [3,4,8,9,10].

Image Mining basically falls in two steps the former is feature extraction and second part is grouping or clustering of image. For each image in a database, feature vector capturing certain essential properties of image is computed and stored in a feature database. Clustering algorithm is applied over this extracted feature to form the group [12, 25].

This paper is organised as follows: section1 introductory overview, section2 related work, section3 proposed work, section4 experimental results, section5 conclusion and references.

I-A. Data Clustering Overview

As mentioned earlier, data clustering is concerned with the portioning of a data set into several groups such that the similarity within a group is larger than that among groups. Clustering methods fall in two categories they are partional and hierarchical approaches. K-MEANS, SOM and Fuzzy C-Means come under partional clustering method that are widely used for satellite images. K-means clustering is mainly utilized for large data sets but it is sensitive to small cluster[2]. In offline mode, the system is presented with a training dataset, which is used to find the cluster canters by analyzing all the input vectors in the training set.

I-B. SOM

In this paper use of self organizing map is proposed for the training and clustering the images features. The unsupervised SOM training is used to train the system. Self-organizing maps (SOM) is a data visualization technique originated by Kohonen in 1990 which can reduce the dimensions of data. Humans cannot easily visualize high dimensional data. The SOM is trained through the use of iterative unsupervised learning to produce a feature map from the input data to the output space such that the topological properties of this network is preserved. In other words, similar n-dimensional data are grouped together and projected onto grids of the output layer. Each input is connected to all output neurons. Every neuron is associated with a weight vector with the same dimensionality as the input vectors. When input data (query issued by users) is given to the network, its Euclidean distance to all weight vectors can then be computed. The distance on the competition grids implies a similarity between the input data such that the pattern recognition can be completed [24].



I-C Color contents

Fig. 1. SOM

It is the most widely used technique. It does not depend on the image size or orientation. Color searches involves



comparing color histograms[3], based on the feature like color one can cluster the image into small groups easily. Color Feature Extraction

Sticker and Orengo [20]. Use three central moments of an image's color distribution in which Pij is the value of kth color component of ij-image pixel and P is the height of the image, and Q is the width of the image. They are mean, standard deviation and skewness. The formulae are as follows:

Means :

$$Ek = \frac{1}{PQ} \sum_{i=1}^{P} \sum_{j=1}^{Q} P \ ijK$$

Mean is the average value in the image.

Standard deviation:

It is the square root of variance of the distribution.

$$SDK = SQRT \left(\frac{1}{PQ} \sum_{j=1}^{P} \sum_{j=1}^{Q} (P \ ijK - EK) 2 \right)$$
Skewness:

$$Sk = \left(\frac{1}{PQ} \sum_{j=1}^{P} \sum_{j=1}^{Q} (P \ ijK - EK) 3) 1/3\right)$$

It is the measure of asymmetry in the distribution.

I-D Color model

I-D.1 HSV Color Model

HSV (hue, saturation, value) color models were developed to be more "intuitive" in manipulating with color and were designed to approximate the way humans perceive and interpret color. Hue defines the color itself. The values for the hue axis vary from 0 to 360 beginning and ending with red and running through green, blue and all intermediary colors. Saturation indicates the degree to which the hue differs from a neutral gray. The values run from 0, which means no color saturation, to 1, which is the fullest saturation of a given hue at a given illumination.

In the HSV color model maximum saturation of hue (S=1) is at value V=1 (full illumination).

The HSV color space is essentially a cylinder, but usually it is represented as a cone or hexagonal cone (hexcone) as shown in the Figure 2 "HSV Solid", because the hexcone defines the subset of the HSV space with valid RGB values. The value V is the vertical axis, and the vertex V=0 corresponds to black color. Similarly, a color solid, or 3D-representation, of the HLS model is a double hexcone (Figure 2 "HSV Solid") with lightness as the axis, and the vertex of the second hexcone corresponding to white [10].

Copyright © 2012 IJECCE, Al



Fig. 2 HSV Solid

In the HSV "hexcone" model, value is defined as the largest component of a color, our M represents chroma. This places all three primaries, and also all of the "secondary colors" - cyan, yellow, and magenta - into a plane with white, forming a hexagonal pyramid out of the RGB cube.

I-D.2 Texture

Texture is the regular repetition of an element or pattern on a surface. The notion of texture generally refers to the presence of a spatial pattern that has some properties of homogeneity [27].

Gabor filter, named after Dennis Gabor, is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. In the spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. The Gabor filters are self-similar: all filters can be generated from one mother wavelet by dilation and rotation.



Gabor filters have been shown to posses optimal localization properties in both spatial and frequency domain and thus are well suited for texture segmentation problems A Gabor filter can be viewed as a sinusoidal plane of particular frequency and orientation, modulated by a Gaussian envelope.



258



FIGURE (a) an image consisting of five different textured regions: cotton canvas (D77), straw matting (D55), raffia (D84), and herringbone weave (D17), and pressed calf leather [26]. (b) the goal of texture classification is to label each textured region with the proper category label: the identities of the five texture regions present in (a). (c) The goal of texture segmentation is to separate the regions in the image which have different textures and identify the boundaries between them. The texture categories themselves need not be recognized. In this example, the five texture categories in (a) are identified as separate textures by the use of generic category labels (represented by the different fill patterns).

A two dimensional Gabor function g(x,y) and its Fourier transform G(u,v) can be written as:

$$g(x, y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right) \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi j Wx\right]$$
(1),

$$G(u,v) = \exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right]\right\}$$
(2),

II. RELATED WORK

Artificial neural network have been used widely in today's world. In variety of disciplines, clustering technique is used as pattern recognition [15], speech analysis [16] and information retrieval [17, 18] etc, the self-organizing NN also known as Kohonen network. As discussed in[12], a number of visual features can be used for describing color, shape and motion characteristics of shots. The SOM clustering has been applied for each feature. After human evaluation of clustering results the following feature were selected: color [13] structure and region based shape for representing similarity by color and shape correspondingly. Now day's remote sensing images are used extensively for finding patterns of interest[1,24]. compared the effect of K-means, Fuzzy-means clustering etc based on color features [3] [25]. CBIR describes the process of retrieving desired images from a large collection on the basis of features such as color [4,8,9,10] and shape[5,6,7], BTC with local average threshold has been proposed. The most popular color space is RGB which stands for RED-GREEN-BLUE as discussed in[14], eigen approach is adopted using color as retrieval feature and this approach performs better than histogram matching.[22] proposed a universal model for the CBIR system by combining the color, texture and edge density features. The advantages of local and global features together have been utilized for better retrieval efficiency. The results are quite good for most of query images.

III. PROPOSED WORK

On the basis of above discussion, first part of evaluation calculate color moment for each of three color components. Each color components yields a feature vector of three elements as discussed in section 1.3 that is mean ,standard deviation and skewness. Thus total nine features vectors are calculated for one image. In the Second part calculate texture moment for that a total of 24 features (8 each sd, mean and energie) per image texture features are extracted. In the Third part Neural network based Clustering is applied over feature database and get output of the proposed algorithm.



Fig. 3 Feature extraction & Clustering

III-A. Color feature extraction algorithm

Proposed algorithm for color feature extraction using HSV color model

Step1: Input an RGB image.

Step2: Convert RGB to gray scale

An RGB color images is an M*N*3 array of color pixels and is an stack of three gray scale images is fed into the red, green, blue inputs of a color monitor, produce the color image on the screen.

Step 3 : The average values for the RGB components are calculated for the all images[12].

R average= summation of all the RED pixels in the image R (P) $\,$

No. Of pixels in the image P

G average= summation of all the GREEN pixels in the image G (P)

No. Of pixels in the image P

B average= summation of all the BLUE pixels in the image B (P)

No. Of pixels in the image P

Step 4: The HSV values of a pixel can be transformed from its RGB representation according to the following formula.

$$H = \arctan \{ 3(G - B) \} / \{ (R-B) + (R-B) \}$$

$$S = 1 - [min\{ R, G, B \} / V]$$

$$V = (R + G + B) / 3$$

Step 5: Represent color by a three dimensional vector X=(x1,x2,x3), where

 $X1 = S * V * \cos (H)$ $X2 = S * V * \sin (H)$ X3 = V

Copyright © 2012 IJECCE, All right reserved



Step 6: Compute mean, standard deviation and skewness of an image.

Step 7: Construct feature vector for color representation using above defined feature components.

Feature Vector Color = (Mean, Standard Deviation, Skewness).

III-B. Proposed algorithm for texture feature extraction using Gabor Function.

Input : 1000 input images(i.e. Each image is of size 384*256 pixels).

III-C. Unsupervised SOM clustering algorithm is as follows

Input : features of 1000 input images(i.e. Color feature 09X1000,Texture feature 24X1000,color+texture 33X1000)

Output : 10 Clusters in term of Hits(With grid size 5*2=10).

Step1: initialize the feature vector to key frame (Grid size 5*2) of shot as cluster c1 and add 1 to its count.

Step 2. If the key frame pool of a segmented shot is empty, then go to step6. Otherwise, take a feature vector of next frame fj of next shot.

Step 3. Compute the Euclidean distance between this input feature vector and the frames in the cluster, and find the minimum distance.

Step 4. If the minimum distance is higher than the automatic threshold value, include the feature vector of new frame into new cluster cj. Assign a counter to this new frame and set the counter value to one, then go back to step2.Otherwise,add a new frame into existing cluster and update the weights of the nearest node(winner) according to Eq(5). Increase the counter value of the winner frame by one. Decrease the learning rate $Wji(k+1) = Wji(k) + \eta(t)\nabla(xi(k)) + Wji(k)...$ (5)

Where,

Wji :- is the ith weight of the jth (winner) frame nearest to the input feature vector,

Xi:- is the ith element of the input vector,

Output : Texture feature 24X1000.

Step 1: Assume, g(x,y) be the mother Gabor wavelet. Step 2: Obtain self-similar filter dictionary by appropriate dilations and

rotations ofg(x,y) through the generating function:

gmn(x,y)=a-mG(x',y'), a > 1,m,n = integer

 $x'=a\text{-}m(x\cos +y\sin),$ and $y'=a\text{-}m(-x\sin +y\cos)\dots$ (3)where =n /K and K is the total number of orientations. The scale factor a-m

in (3) is meant to ensure that the energy is independent of m. [27].

Step 3: The Gabor wavelet transform dilates and rotates the two-dimensional Gabor function.

Step4: The image is then convoled with each of the obtained Gabor functions.

This gives a total of 24 features (8 for each angle sd, Eni mean and energy) per image texture features are extracted.

 $\eta(t)$:- is the learning rate (0 < $\eta(t)$ < 1),and

k:- is the iteration number,

 $\eta(t)$ becomes smaller as time t increases.

Step 5: go to step2 until all feature vectors of shots are exhausted.

Step 6: Compute the mean weight of each cluster and stop processing.

IV. EXPERIMENTS

Proposed scheme has been performed with 10 classes of images which can be downloading from the website wang.ist.psu.edu/iwang/test1.tar. In this proposed scheme the database has 10 classes of images, 100 images in each class. Each image is of size 384*256 pixels. The system is developed in Matlab. We define unichrome feature as values that are extracted from a single color layer of hue, saturation, value. The first part of evaluation computes color moments for each of the three color components. Each color component yields a feature vector of three elements i.e. mean, standard deviation and skewness.

Fi Color = (Ei, SDi, Si) for ith color component.

The Second part of evaluation computes texture moments for each of image. Each image component yields a feature vector of three elements i.e. standard deviation, mean, and energy.

Ti Texture = (SDi, mni, Eni i) for ith texture component . IV-A Confusion Matrix for Color Contents

Class of images										
1	11	2	37	28	17	1	1	0	2	'African human'
3	2	2	5	18	0	1	30	39	0	'Beach'
5	1	2	23	37	1	1	11	19	0	'Historical place'
0	5	66	9	2	3	10	5	0	0	'Bus'
0	0	0	0	0	0	0	0	0	100	'Dinosaur'
10	0	0	2	62	3	0	17	4	2	'Elephant'
3	31	13	18	0	3	32	0	0	0	'Roses'
73	7	0	0	6	11	0	1	2	0	'Horse'
0	0	11	10	4	1	3	46	25	0	'snow'
5	23	2	3	11	43	8	1	3	1	'Breakfast'

IV-B Confusion Matrix for texture contents: Cluster->from 1 to 10



Cla	ss of	imag	es							
0	4	19	18	21	5	21	2	7	3	'African human'
1	0	6	6	40	1	6	5	25	10	'Beach'
0	2	18	18	11	11	16	12	7	5	'Historical place
13	1	6	1	2	28	18	25	6	0	'Bus'
0	0	7	2	1	0	3	17	68	2	'Dinosaur'
0	0	19	21	22	0	4	8	24	2	'Elephant'
0	0	1	7	18	0	0	0	0	74	'Roses'
2	0	2	22	55	0	0	3	1	15	'Horse'
0	1	12	16	32	2	4	8	16	9	'snow'
0	9	27	14	3	12	17	9	8	1	'Breakfast'

V-C Confusion Matrix for color and texture contents: Cluster->from 1 to 10

Clas	ss of i	imag	es							
2	23	11	32	28	1	0	1	0	2	'African human'
4	0	4	4	17	1	1	30	39	0	'Beach'
5	0	5	21	37	1	1	12	18	0	'Historical place'
2	1	16	6	2	10	57	5	1	0	'Bus'
0	0	0	0	0	0	0	0	0	100	'Dinosaur'
8	2	2	1	64	0	0	16	4	3	'Elephant'
15	11	23	15	0	32	4	0	0	0	'Roses'
77	4	1	1	7	0	0	1	9	0	'Horse'
0	0	8	3	6	3	11	47	22	0	'snow'
14	52	6	3	11	8	1	1	3	1	'Breakfast'

F-measure	Recall	Precision	Class of images
34.5945	32	37.6471	'African human'
39.7959	39	40.6250	'Beach'
26.5682	36	21.0526	'Historical place'
64.7398	56.00	76.7123	'Bus'
97.0873	100	94.3396	'Dinosaur'
47.2324	64	37.4269	'Elephant'
41.0256	32	57.1429	'Roses'
66.6666	75	60.0000	'Horse'
44.1314	47	57.1429	'snow'
53.0612	52	54.1667	'Breakfast'



Fig 4.1 Som for Simple hits for 10 cluster(Color Conents)



Fig 4.2 Som for Simple hits for 10 cluster(texture contents)



Fig. 4.3 Som for Simple hits for 10 cluster (color and texture contents)

Table 1.1 shows the result in term of F-measure, Recall and Precision(Color contents)

IV-D1 Recall consists of the proportion of target images that have been retrieved among all the relevant images in the database.

Recall(R) = Number of Relevant Images Retrieved Total Number of Relevant Images

Precision consists of the proportion of relevant images that are retrieved.

IV-D 2 Precision (P) = Number of Relevant Images Retrieved

Total Retrieved Images

IV-D 3 F-measure is One measure of performance that takes into account both recall and precision.

$$F = \frac{2PR}{P+R} = \frac{2}{\frac{1}{R} + \frac{1}{P}}$$



F-measure(F) =



Fig. 5.1for precision and recall results(Color contents)



Table 1.2 shows the result in term of F-measure, Recall and Precision (texture contents)



F-measure	Recall	Precision	Class of Images
13.77049	21	10.2439	'African human'
26.22951	40	19.5122	'Beach'
16.58986	18	15.38462	'Historical place'
35.22013	28	47.45763	'Bus'
51.9084	68	41.97531	'Dinosaur'
18.32061	24	14.81481	'Elephant'
41.0256	74	61.15702	'Roses'
36.06557	55	26.82927	'Horse'
20.98361	32	15.60976	'snow'
24.88479	27	23.07692	'Breakfast'



Fig. 5.2 for precision and recall results (texture contents)

F-measure	Recall	Precision	Class of Images
34.5945	32	37.6471	'African human'
39.7959	39	40.6250	'Beach'
26.5682	36	21.0526	'Historical place'
64.7398	56.00	76.7123	'Bus'
97.0873	100	94.3396	'Dinosaur'
47.2324	64	37.4269	'Elephant'
41.0256	32	57.1429	'Roses'
66.6666	75	60.0000	'Horse'
44.1314	47	57.1429	'snow'
53.0612	52	54.1667	'Breakfast'

Table 1.3 shows the result in term of F-measure, Recall and Precision (color and texture contents)





Fig. 5.3 for precision and recall results (color andtexture contents)

As described in section 3.3, SOM clustering algorithm is applied with epoch=200 over the extracted feature and 10 clusters (i.e. grid size=5*2=10) are Table 1.1 shows the performance of clusters using SOM clustering algorithm. Result shows that Class 3 having lowest precision value and class 5th having highest value. similarly in case of recall class 1 having recall value and 5th class having highest recall value and Table 1.2 and 1.3 shows that Class 1 having lowest precision value and class 5th having highest value. similarly in case of recall value and Table 1.2 and 1.3 shows that Class 1 having lowest precision value and class 5th having highest value. similarly in case of recall 3rd having lowest recall value and 5th class having highest recall value.

V. CONCLUSION

Based on above discussion the content of an image can be expressed in terms of different features such as Texture. Here we have proposed a frame work of unsupervised clustering of image based on the Texture feature over the extracted dataset. The result of Table 1.1, 1.2 and 1.3shows that the SOM clustering algorithm produces better results that are very much acceptable. Class 5th having highest precision and recall value it means color content is highly effective in this case and also Class 5th and 7th having highest precision and recall value it means Texture content is highly effective in this Class.

REFERENCES

- Priti Maheshwary, Namita Srivastav "Comparison of clustering algorithm on region based remote sensing image similarity matching" international conference on content based image retrival ICCBIR-2008.
- [2] S.Vasuki, L.Ganesan and R.Shymala Sundari "Performance comparison of Hybrid Methods combining Clustering and region merging for segmentation of multi spectral images" international conference on content based image retrival ICCBIR-2008.
- [3] Dr. H.B.Kekre, Tasneem Mirza "Content Based image retrieval using BTC with Local average the sholding" international conference on content based image retrival ICCBIR-2008.
- W. Niblavk et al," quering images by color, texture and shape." Proc. SPIE, vol. 1908, pp. 173-187, 1993.
- [5] Comaniciu D., Meer P., For a,D.," Shape-based image indexing and retrieval for diagnostic pathology," IEEE Proceedings of 14th International Conference on pattern recognition, Volume1, Issue,16-20 Aug 1998.
- [6] Kunttu, I., Lepisto, L., Rauhamaa, J., Visa, A. "Multiscale fourier descriptor for shape based image retrieval". IEEE proceedings of the 17th international conference on pattern recognition, 2004. ICPR 2004 Vol. 2. Issue, 23-26 Aug 2004.
- [7] Hwei-Jen Lin, Yang-Ta Kao, Shwu-Huey Yen, Chia-Jen wang." A study of shape based image retrieval,"24th International Conference on distributed computing systems workshops-W1: MNSA (ICDCSW'04),2004.icdcsw.pp.118-123.



- Le Bourgeois, F. "Content based image retrieval using gradient [8] color fields," Proceedings of 15th International conference on pattern recognition 2000. Volume 1 pages 1027-1030.
- [9] M. Swain and D. ballard, Color Indexing," Int. J.Comput Vis Vol. 7 pp 11-32, 1991.
- H.B.Kekre, Swapna Borde "color and texture based image [10] retrieval," Proceedings of international Conf. On advances in computer vision and Info. Tech. at BAMU, Aurangabad.
- [11] D. Balasubramaniam, V.Jeyabalaraja R. Boopathy, S. Thirumal-"Searching and Navigating for semantic web Techniques".
- [12] C.Halaschek-Wiener, B.Parsia, and E.Sirin. Description logic reasoning with syntactic updates. In Proc. of Int. Conf. on Ontologies, databases and App. of Semantics 2006.
- [13] Ramegowda, Karisiddappa, K.C.Krishna," RGB Image retrieval based on eigen values and eigen vectors".
- [14] R.O.Duda and P.E.Hart," Pattern Classification and Scene analysis", Ch6.John Wiely and sons, Inc.
- L.Rabiener and B.H.Juang," Fundamentals of Speech Recognition." Ch. 5, Englewood Cliffs, New Jersey : Prentice [15] Hall, 1993.
- [16] G.Salton and M.J.McGill," Introduction to Modern information Retrieval", McGraw-hill Book company, 1983.
- [17] A.M.Ferman and A.M.Tekalp," Multiscale content extraction and representation for video indexing." in Multimedia Storage and Archival systems Nov. 1997.
- [18] P.Geetha," Adaptive -threshold clustering of content based video retrival.
- [19] Dengsheng Zhang and Guojun Lu."A Comparative study of shape retrival using fourier descriptors with different shape signatures" Journal of visual communication and image representation, 2003.
- [20] S.Nandagopalan, Dr.B.S.Adiga, and N.Deepak," A Universal Model for Content-based image retrieval".
- [21] Alberto Amato, Vicenzo Di Lecce," Edge detection techniques in image retrieval: The semantic meaning of edge,"
- Minyoung Eom, and Yoonsik Choe," Fast Extraction of Edge [22] Histogram in DCT domain based on MPEG7"
- [23] Bin Luo and Jocelyn Chanussot" Unsupervised Classification Of Perspectral Images By Using Linear Unmixing Algorithm "2009 IEEE
- [24] Yo-Ping Huang, Tsun-Wei Changand Li-Jen Kao Using Fuzzy SOM Strategy for Satellite Image Retrieval and Information Mining " 2008.
- Hu Min, Yang Shuangyuan' Overview of Image Mining Research'' The 5th International Conference on Computer [25] Science & Education Hefei, China. August 24-27, 2010.
- [26] H. Bunke, Dr. U.-V. Marti, R. Messerli : Writer Identification
- Using Text Line Based Features, Proceedings of 6th ICDAR '01, B.S Manjunath and W.Y.Ma," Texture Features for Browsing and Retrieval of Image Data", IEEE Transactions on Pattern [27] Analysis and Machine Intelligence, Vol. 18, 8 Aug 1996.

AUTHOR'S PROFILE

Divakar Singh

BUIT, BU, Bhopal divakar.singh@buit.ac.in

R. C. Jain SATI Vidisha dr.jain.rcd@gmail.com