

Using Color and Texture Feature Extraction Technique to Retrieve Image

Santanu Kumar Sahoo, Asit Kumar Subudhi

Abstract— This paper proposes a content image retrieval using the texture and the color feature of the images. Although for extraction of texture feature, the “gray level co-occurrence matrix (GLCM) algorithm” is used and for extracting color feature the color histogram is used. The presented system is tested on the WANG database that contains a thousand color images with ten different classes by the help of three various type of distances.

Keyword: Texture, feature, Gray Level co-occurrence Matrix (GLCM), Color feature, Distance Formulas.

I. INTRODUCTION

A photo or an image comprises complete information regarding the captured scenery of the whole place. A picture prescription test is taken by some interviewers which asks for different views from each candidate. That shows the thinking and other capabilities of the candidate. The image retrieval is a method of taking images as an input and then showing similar images through a display. The image retrieval technique is most popular nowadays and consists many advantages over the text type image based retrieval method [1].

T. Huang [2] presents idea of a survey of the image qualities by using different processes and methods. He mentioned using of the CBIR for real time application and also does proposed his achievements in present and past work. The 3 databases that he used in the system are raw images database, text annotation databases that consists text and keywords and feature database that is used for storing the extracted features value. This interaction displays an efficient interaction in between machine and human wherein the text/keyword based system is connected to the content based system.

S. Dai shows in [3] that how efficiently the similarity of the images can be measured by using the texture similarities [4]. He also mentioned advantages of the CBIR over the text image retrieval. Texture based retrieval is calculated and then it depends on the property of pixel values for every grey level image.

II. BACKGRUOND

Content Based Image Search:

The "Content-based image retrieval" means searching is done on the basis of contents in image rather than searching

on any textual descriptions that is given in the image. Basic 3 main features of the CBIR are shape, color and the texture of the image.

Texture Feature(s):

The texture presents internal data value of images, texture feature is like surface pattern of the image. Secondly this describes properties of all the surfaces that has been used to represent all the visual patterns and each of them having the same similarity [5].

GLCM:

“GLCM or Gray Level Co-occurrence Matrix” is mainly used for searching and extracting texture feature. The co-occurrence of the matrix and the texture features are important for the GLCM. The GLCM of any image is calculated by using the displacement vector and orientation. Let number of grey level in the image be N, and the GLCM matrix of size N x N.

$$M_{i,j} = M(i,j) = \sum_{x=1}^N \sum_{y=1}^N I(x + \Delta x, y + \Delta y) = j; \begin{matrix} 1, \text{if } I(x,y) = i \text{ and} \\ 0, \text{otherwise} \end{matrix}$$

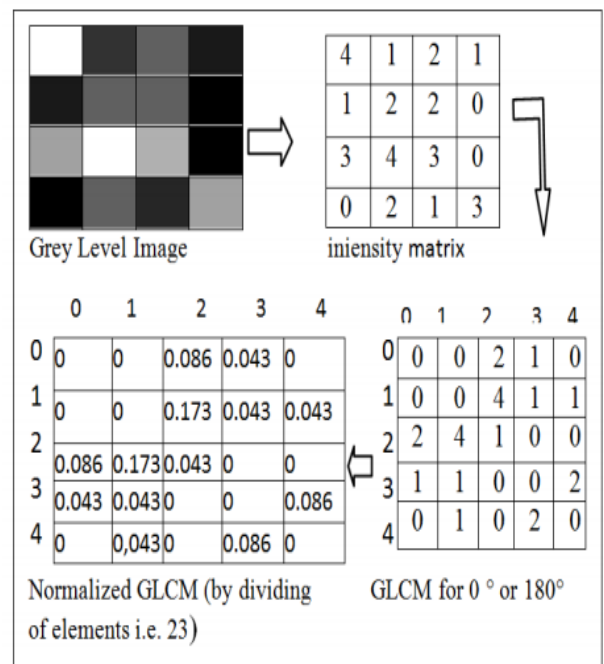


Fig. 1: GLCM matrix evaluation

Revised Manuscript Received on September 10, 2019.

Santanu Kumar Sahoo, Dept. of Electronics & Communication Engineering, Siksha O Anusandhan Deemed to be University, Odisha, India.

(E-mail: santanusahoo@soa.ac.in)

Asit Kumar Subudhi, Dept. of Electronics & Communication Engineering, Siksha O Anusandhan Deemed to be University, Odisha, India (E-mail: asitubudhi@soa.ac.in)

Contrast:

Contrast measures intensity of the gray level variation between multiple pixel and its neighboring:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-j)^2 P(i,j)$$

Where $P(i, j) = (i,j)$ th entry of the Normalized GLCM P.

Correlation:

The correlation is measurement of the multiple correlated variable pixels with their neighboring in the matrix.

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{(1-\mu_x)(1-\mu_y)P(i,j)}{\sigma_x \sigma_y}$$

Where $\mu_x, \mu_y, \sigma_x, \sigma_y$ are the mean and standard deviations of P_x and P_y , calculated as equation (3.1) to (3.3) which are equal for symmetric GLCM.

$$\mu_x = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} i \cdot P(i,j) \text{ and}$$

$$\mu_y = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} j \cdot P(i,j)$$

$$\sigma_x = \sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - \mu)^2 P(i,j) \text{ and}}$$

$$\sigma_y = \sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (j - \mu)^2 P(i,j)}$$

$$P_x = \sum_{j=0}^{N-1} P(i,j) \text{ and } P_y = \sum_{i=0}^{N-1} P(i,j)$$

Energy:

The energy is used for measuring the sum of all squared elements in GLCM.

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j)^2$$

Color Feature(s):

The image color feature is very important feature that takes a big role in the process of retrieval. In this paper, the conversion of an image from RGB color model to HSV color model is done.

Conversion from RGB to HSV:

$$\text{Hue (H)} = \cos^{-1} \frac{0.5[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}$$

$$\text{Saturation (S)} = 1 - \frac{3[\text{MIN}(R,G,B)]}{R+G+B} \quad (7)$$

$$\text{Value (V)} = \frac{R+G+B}{3}$$

The normal value of the hue data varies from zero to one, wherein zero represents of being red. Saturation ranges from zero to one and more grey is present in lower saturation value. The Values are occur from zero to one and where zero is total dark value.

III. FEATURE EXTRACTION & RESULTS

WANG database is selected as shown below:

Class Name	Images		
African People			
Seas			
Monuments			
Buses			
Dinosaurs			
Elephants			
Roses			
Horses			
Mountains			
Food			

Figure 2: Wand database

A GLCM matrix is generated from the given database of the food class image.

The GLCM matrix is calculated and the matrix is normalized by dividing each and every element by summing all the elements.

Table 1 represents calculation of feature extraction values.

2349	1293	233	96	46	23	1	0
1377	19565	3624	699	305	158	42	3
200	3698	27476	1590	578	268	135	8
74	646	1807	3872	1541	469	277	19
25	239	528	1851	4864	1344	441	50
23	106	222	422	1596	4856	1337	64
2	32	98	215	429	1463	4939	153
0	2	4	17	28	55	191	9

Figure 3: GLCM matrix

Table 1: 4features extracted from Normalized GLCM

Property	Contrast	Correlation	Energy	Homogeneity
Value	0.7223	0.8663	0.1334	0.8314

Table 2: Histogram Color Values evaluated using LCH of a food class image

Partion no.	Red	Green	Blue	Yellow	Cyan	Magenta	Black	White
1	0.0458	0.0028	0	0.0195	0	0	0.0034	0
2	0.0497	0.0034	0.0001	0.0279	0.0001	0.0004	0.0043	0.0012
3	0.0824	0.0043	0.0008	0.0573	0.0002	0.0004	0.0049	0.0008
4	0.0770	0.0093	0.0013	0.0639	0.0144	0.0012	0.0057	0.0014

Table 2 shows the histogram color values by using the LCH

Table 3: Precision value for four different features with L=5, 10, 20,40,60,80,100 (using Euclidean distance)

Feature Name	Number of Retrieved images						
	5	10	20	40	60	80	100
Texture Feature	61.00	54.10	50.55	45.30	42.88	40.23	38.19
Color (GCH)	67.40	57.40	50.75	44.00	39.85	36.45	34.05
Color (LCH)	69.80	60.00	52.30	45.62	40.50	37.74	35.31
Color(LCH)+texture	82.20	74.20	65.75	57.00	51.73	47.70	44.33

Table 3 shows precise values with 4 different features.

IV. CONCLUSION

This paper proposes a system that extracts and combines features of the texture and color, wherein the textures and the color are extracted with the help of GLCM algorithm and the color features are extracted by the use of histogram values in the HSV model. Three distances that is manhattan distance, Euclidean distances and chessboard distance are used for measuring similarity between the retrieving images. From both the distances that is Euclidean and manhattan distance shows an efficient output when compared with the prior results of prior papers.

As of the future work, this experiment can be performed with additional features that is shape. Although the paper could use more classifiers for identifying and classifying.

V. REFERENCE

1. T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2002.
2. Y. Rui, T. S. Huang, and S. F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *J. Vis. Commun. Image Represent.*, 1999.
3. P. W. Huang and S. K. Dai, "Image retrieval by texture similarity," *Pattern Recognit.*, 2002.

4. T. R. Reed and J. M. H. du Buf, "A review of recent texture segmentation and feature extraction techniques," *Computer Vision and Image Understanding*, 1993.
5. D. C. He and L. Wang, "Texture Unit, Texture Spectrum, and Texture Analysis," *IEEE Trans. Geosci. Remote Sens.*, 1990.