International Journal of Latest Engineering Research and Applications (IJLERA) ISSN: 2455-7137

Volume – 02, Issue – 05, May – 2017, PP – 04-07

RESEARCH ON TEXTURE FEATURE EXTRACTION OF IMAGE RETRIEVAL ANALYSIS

Mr. Raghu Kumar Limgamallu,
M. Tech., Ph.D Scholar,
O.P.J.S University, Churu, Rajasthan State, India. Dept. of CSE

Dr. K. Naveen
M.Tech., Ph.D, HOD & Associate Professor, Dept of ECE,
Krishna Murthy Institute of Engg and Technology,
Ghatkesar, Hyderabad,

KRUSHIMA SOMA,
M.Tech, ASST PROFESSOR, KG REDDY COLLEGE OF ENGG & TECH,
MOINABAD

Abstract: This paper has a look for the in the lands ruled over of image processing, image mining is move-forward in the field of facts mining. image mining is the extraction of put out of the way knowledge for computers, association of image data and added good example which are quite not clearly able to be seen in field that has to do with, image processing, data mining, Machine Learning, not natural quick brains and knowledge-base. The lucrative point of image Mining is that without any before information of the designs it can produce all the important designs. This is the writing for a make observations done on the mixed, of different sort's image mining and knowledge for computers mining expert ways of art and so on. Facts mining says something about to the getting from of knowledge /information from a very great knowledge-base which is stored in further number times another heterogeneous knowledge-bases. Knowledge/ information is making an exchange of note through straight to or roundabout way of doing. These techniques join neural network, coding into groups, connection and association. This writing gives a first paper on the application fields of facts mining which is full of changes into telecommunication, making, Fraud discovery, and marketing and education part. In this way of doing we use size, feeling of a material and chief colour factors of an image. Grey Level Co-occurrence matrix (GLCM) point is used to come to a decision about the feeling of a material of an image. Points such as feeling of a material and colour are normalized. The image acts to get back point will be very sharp using the feeling of a material and colour point of image gave with the form point. For similar types of image form and feeling of a material point, weighted Euclidean distance of color point is put to use for getting back points.

Index Terms: Data Mining, Feature Extraction, Image recovery, Clustering, database, ray Level Co-occurrence Matrix, centroid, Weighted Euclidean Distance.

1.1. TEXTURE FEATURE EXTRACTION

Texture is a visual feature that refers to inherent surface properties of an object and their relationship to the surrounding environment. This section proposes a texture feature representation scheme based on image co-occurrence matrix. Co-occurrence matrix is widely used to extract texture feature in gray-scale image and has been shown to be very efficient. The color image will be converted to a gray-scale image and the number of the gray scale value is 256.

$$Y = 0.29 \times R + 0.589\times G + 0.14 \times B, \qquad (3)$$

Where Y is the gray-scale value and , G , B represent red, green, and blue components, respectively. The co occurrence probabilities provide a second-order method for generating texture features. These probabilities represent the conditional joint probabilities of all pair wise combinations of gray levels in the spatial window of interest given two parameters: inter pixel distance ( ) and orientation (θ). The probability measure can be defined as

$$P_r(x) = \left\{ C_{ij} \mid (\delta, \theta) \right\}, \qquad (4)$$

where C_{ij} (the co-occurrence probability between gray levels i and j) is defined as

$$C_{ij} = \frac{P_{ij}}{\sum_{i,j=0}^{G} P_{ij}}, \qquad (5)$$

where P_{ij} represents the number of occurrences of gray levels i and j within the given window, given a certain (δ , θ) pair; and G is the quantized number of gray levels. The sum in the denominator thus represents the total number of gray level pairs (i,) within the window. Statistics applied to the co-occurrence probabilities to generate the texture features are defined in

$$\text{Contrast} = \sum C_{ij}(i-j)^2, \qquad (6)$$

$$\text{Energy} = \sum \left(C_{ij}\right)^2, \qquad (7)$$

$$\text{Entropy} = \sum C_{ij} \log C_{ij}, \qquad (8)$$

$$\text{Correlation} = \sum \frac{(i-\mu_x)\left(j-\mu_y\right)C_{ij}}{\sigma_x \sigma_y}, \qquad (9)$$

$$\text{Local Stationary} = \sum C_{ij}\left|i-j\right|. \qquad (10)$$

The gray-scale quantification is made and the corresponding co-occurrence matrix of size 256 × 256 is obtained. The statistical properties such as contrast, energy, entropy, correlation, and local stationary are calculated using (6)–(10) to describe the image content. The texture features are extracted in the following five steps.

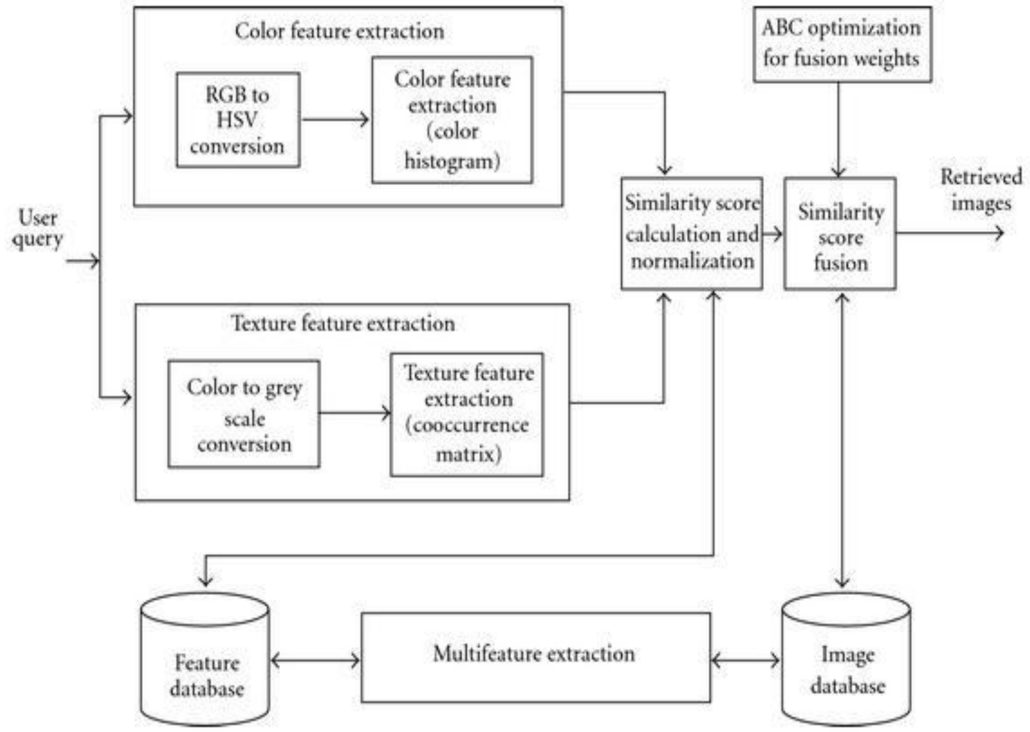
Step 1: The color image is converted to gray-scale image and the image co-occurrence matrix is derived using (4) and (5).

Step 2: The five statistical properties such as contrast, energy, entropy, correlation, and local stationary are calculated using (6)–(10) in four orientations such as 0°, 45°, 90°, and 135°, so that, totally 20 texture features are obtained.

Step 3: Mean and variance of the above five parameters are taken. The results are the ultimate texture features and are denoted as

$$T = (\mu 1, \mu 2, \mu 3, \mu 4, \mu 5, \sigma 1, \sigma 2, \sigma 3, \sigma 4, \sigma 5) \qquad .(11)$$

Step 4:The similarity value between the query image and that of the database images are calculated using Euclidean distance by (2); the closer the distance, the higher the similarity.



1.2 GRAY LEVEL CO-OCCURRENCE MATRICES:

- The statistical measures described so far are easy to calculate, but do not provide any information about the repeating nature of texture.
- A gray level co-occurrence matrix (**GLCM**) contains information about the positions of pixels having similar gray level values.
- A co-occurrence matrix is a two-dimensional array, **P**, in which both the rows and the columns represent a set of possible image values.
- A GLCM **Pd**[i,j] is defined by first specifying a displacement vector **d**= (dx,dy) and counting all pairs of pixels separated by **d** having gray levels i and j.

The **GLCM** is defined by:

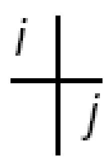
Pd(i,j)=nij

– where $\mathbf{n_{ij}}$ is the number of occurrences of the pixel values **(i,j)** lying at distance **d** in the image.

– The co-occurrence matrix $\mathbf{P_d}$ has dimension **n× n**, where n is the number of gray levels in the image.

For example, if **d**=(1,1)

2	1	2	0	1
0	2	1	1	2
0	1	2	2	0
1	2	2	0	1
2	0	1	0	1



$$P_d = \begin{array}{cc} \begin{bmatrix} 0 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 3 & 2 \end{bmatrix} & \begin{array}{c} 0 \\ 1 \\ 2 \end{array} \quad i \\ \begin{array}{ccc} 0 & 1 & 2 \end{array} & \\ j & \end{array}$$

there are 16 pairs of pixels in the image which satisfy this spatial separation. Since there are only three graylevels, **P[i,j]** is a **3×3** matrix.

1.2.1 Algorithm:

- Count all pairs of pixels in which the first pixel has a value ***i***, and its matching pair displaced from the first pixel by **d** has a value of ***j***.
- This count is entered in the i^{th} row and j^{th} column of the matrix $\mathbf{P_d[i,j]}$
- Note that $\mathbf{P_d[i,j]}$ is not symmetric, since the number of pairs of pixels having gray
- The elements of $\mathbf{P_d[i,\ j]}$ can be normalized by dividing each entry by the total number of pixel pairs.
- Normalized GLCM **N[i ,j],** defined by:

$$N[i,j]=\frac{P[i,j]}{\sum\sum P[i,j]}$$

which normalizes the co-occurrence values to lie between 0 and 1, and allows them to be thought of as probabilities.

Numeric Features of GLCM

- Gray level co-occurrence matrices capture properties of a texture but they are not directly useful for further analysis, such as the comparison of two textures.
- Numeric features are computed from the co-occurrence matrix that can be used to represent the texture more compactly.

1.3 COLOR CO-OCCURRENCE MATRIX:

Conventional color co-occurrence matrix represents three dimensional matrix where the colors of any pair are along the first and second dimension and the spatial distance between them along the third [5]. In this sense, conventional CCM is same as color correlogram [7]. In this paper, CCM is simplified to represent the number of color (hue) pairs between adjacent pixels in the image. For each pixels in the image, 4-neighbors (horizontal and vertical neighbors) are accounted.

Let I be an N×M image quantized to m colors, and p (x , y) is the color of the image pixel p(x,y) . Then, the simplified CCM is given by

$$H^I(i,j)=\eta((p(x,y),p(\mathcal{N}_{(x,y)}))==(i,j))$$

$$=\alpha\sum_{x=1}^{N}\sum_{y=1}^{M}C_i(x,y)\sum_{(x',y')\in\mathcal{N}_{(x,y)}}C_j(x',y') \qquad (1)$$

where η indicates the number of times (p(x,y),p(N(x ,y)) equals the value of the color indices (i,j) and N(x,y) indicates 4 neighbors of the pixel (x ,y),C_i(x, y)

$$c\mathbf{i}(x,y)=\begin{cases}1 & if\ p(x,y)=i\\ 0 & 0.w\end{cases}$$

And the normalization constant α is 1 4. NXM, for the total number of pixel pairs p(x, y), p (N(x ,y) is approximately. NxM by discounting the difference of boundary pixels. The simplified CCM is symmetric because the adjacent pixels pairs are neighbors of each other. In this paper, color was quantized to 16 colors, since empirically 16 colors (in hue model) are sufficient for proper color invariant object retrieval. Therefore, the dimension of simplified CCM is 16X16.

References

[1]. Baeza-Yates, R. and Ribeiro-Neto, B. (1999): Modern information retrieval. New York, Addison Wesley. Cooper, G. (1990):

[2]. Computational complexity of probabilistic inference using Bayesian belief networks. Artificial Intelligence 42:393-405.Cox, I. J., Miller, L., Minka, P., Papathomas, V. and Yianilos P. (2000):

[3]. The Bayesian image retrieval system, PicHunter: Theory, Implementation and Psychophysical Experiments. IEEE Transaction on Image Processing 9(1):20-37. Heckerman, D. (1995):

[4]. A tutorial on learning with Bayesian Networks. Technical Report MSR-TR-95-06. Microsoft Research, Redmond, Washingtons. Hong, P., Tian, Q. and Huang, T.S. (2000):

[5]. Incorporate Support Vector Machines to Content-Based Image Retrieval with Relevance Feedback, IEEE International Conference on Image Processing. Vancouver, Canada. Ishikawa,