CONTENT BASED IMAGE RETRIEVAL USING GRAY LEVEL CO-OCCURRENCE MATRIX WITH SVD AND LOCAL BINARY PATTERN

D.Kishore[1] , S.Srinivas Kumar[2], Ch.Srinivasa Rao[3]

1 *Research scholar, ECE Dept., JNTUK, Kakinada, A.P, India.*
2 R&D Director, Professor of ECE, JNTUK, Kakinada , A.P, India.
3 *Professor, ECE Dept, JNTUCE, Vizianagaram, A.P, India.*

ABSTRACT

In this paper, gray level co-occurrence matrix, gray level co-occurrence matrix with singular value decomposition and local binary pattern are presented for content based image retrieval. Based upon the feature vector parameters of energy, contrast, entropy and distance metrics such as Euclidean distance, Canberra distance, Manhattan distance the retrieval efficiency, precision, and recall of the images are calculated. The retrieval results of the proposed method are tested on Corel-1k database. The results after being investigated shows a significant improvement in terms of average retrieval rate, average retrieval precision and recall of different algorithms such as GLCM, GLCM & SVD, LBP with radius one and LBP with radius two based on different distance metrics.

Keywords

CBIR, GLCM, SVD and LBP.

1. INTRODUCTION:

The image database is increasing fast recently. The capturing of useful images is difficult in the large image database and we obtain frequently useful images, which is widely extending through the Internet. It is not easy to change the situation by using existing methods, so it is a new way to find retrieve images accurately as urgent as possible. Thus image retrieval becomes an essential topic in pattern recognition domain and image processing. Images can be retrieved in three ways generally: content based, text based and semantic based[1–4]. Text based retrieval approach is widely used, such as Google and Baidu, and images can retrieved by using keywords that are annotated on images. However, with this method we have obtained that images are not related to our expected results. The two drawbacks of this approach are firstly, images in the database were manually annotated by annotators and secondly, the retrieval results are poor, because they are similar to the understanding of the query images. The ordinal approach, example, content based image retrieval (CBIR) had been proposed in the early 1990's [4–7]. This approach is based on retrieving the images by using low level features like colour, shape and texture that can represent

an image. By using this method, we first extract the low level features on an example image, and then compute the similarity between the database images and the query image, finally images are sorted by similarity and the top images will be displayed. CBIR has been shown more subjectively and effectively compared to text based approach [8–9]. The final approach is semantic based method. CBIR has also failed to describe the semantic concepts, so some methods are proposed by researchers for image retrieval by using relevance feedback algorithms. The relevance feedback algorithms capture user's preferences and bridge the semantic gap [10], and the results are closer to human perception. In CBIR, color, shape and texture are the most significant features.

2. FEATURE EXTRACTION METHODS:

2.1 Gray Level Co-occurrence Matrix (GLCM):

The approach, which is the easiest, for texture description is to use statistical moments of the histogram based on intensity of region or image. Gray Level Co-occurrence Matrix is a one of the statistical approach, help to give important information about the relevant position of the neighboring pixels in an image. The gray level co-occurrence matrix P can be defined for an image I of size NxN as P(i, j) = Nx=1 Ny=1 , if I(x, y) = i and I(x + Δx, y + Δy) = j, 0 otherwise. Here, the offset (Δx, Δy), is representing the distance between the neighbor and its pixel of interest. Note that the offset (Δx, Δy) parameterization makes the gray level co-occurrence matrix sensitive to rotation. Choosing an offset vector, such that the rotation of an image is not equal to 180 degrees, will result in a different gray level co-occurrence matrix for the same (rotated) image. This can be avoided by forming the gray co-occurrence matrix using statistical features of Co-occurrence matrix, of offsets sweeping through 180 degrees at the same distance parameter Δ to achieve a degree of invariant rotation (i.e., [0 Δ] for 0◦: P horizontal, [-Δ, Δ] for 45◦: P right diagonal, [-Δ 0] for 90◦: P vertical, and [-Δ − Δ] for 135◦: P left diagonal). Figure 1 illustrates the details to generate the four co-occurrence matrices using Ng =5 levels for the offsets {[0 1], [-1 1], [-1 0], [-1 -1]} that are defined as one neighboring pixel in the possible four directions. We can observe that two neighboring pixels (2, 1) of the input image is reflected in PH concurrence matrix as 3, because there are 3 occurrences of the pixel intensity value 2 and pixel intensity value 1 adjacent to each other in the input image. The neighboring pixels (1, 2) will occur again 3 times in PH, which makes these matrices symmetric. In the same manner, the other three matrices PV, PLD, PRD are calculated.
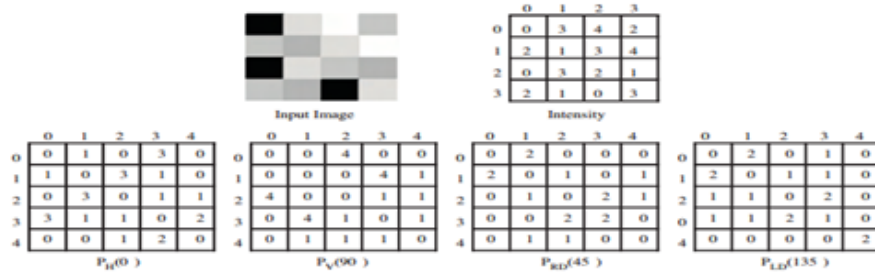
Figure1. Co-occurrence matrix generation for Ng =5 levels and four different offsets: PH
(0◦), PV (90◦), PRD (45◦), and PLD (135◦).

2.2 Singular Value Decomposition:

Singular Value Decomposition (SVD) is one of the powerful mathematical tools in linear algebra used in many image processing applications such as image compression, object recognition [11], and image watermarking [12]. SVD is an optimal matrix decomposition technique in a least square sense that it packs the maximum signal energy into few coefficients. Singular value decomposition (SVD) of an image results in three matrices (U, D and V) of same size as that of the original image. U and V represents the left and right singular vectors of the image matrix and D is a diagonal matrix with singular values. Further, singular values extracted by SVD [11] have good performance for shape description. In the SVD of an image, each singular value in the SVD decomposed matrix (D) specifies the luminance of an image layer providing geometrical invariance, while the corresponding pair of singular vectors (U,V) specifies the geometry (structural details) of image layer. According to it, every real matrix A can be decomposed into a product of three matrices,

$$A=U\Sigma V^T \tag{1}$$

If A is represented in matrix format:

$$\left[u_1 \; u_2 \; u_N\right] \times \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & & 0 \\ 0 & 0 & 0 & \lambda_N \end{bmatrix} \times \left[v_1 \; v_2 \; . \; v_N\right]^T \tag{2}$$

The SVD of matrix A can be written as

$$A = \lambda_1 U_1 V_1 + \lambda_2 U_2 V_2 + ...\lambda_r U_r V_r \tag{3}$$

Where r is the rank of matrix A which is less than or equal to N. The singular values above rank r are zero. U and V are orthogonal matrices, satisfying the condition $U^T U = I$, $V^T V = I$, and $\Sigma = $ diag $(\lambda 1, \lambda 2, ...)$. The diagonal entries of Σ are called the singular values of A, the columns of U are

called the left singular vectors of *A*, and the columns of *V* are called the right singular vectors of *A*.

2.2.1 Properties of Singular Value Decomposition:

1. Existence and Uniqueness: Every matrix A, with real or complex values has singular valued decomposition. .If A is a square matrix and its singular values are distinct, than the left and right singular vectors are determined uniquely up to sign, i.e., a coordinate reflection of each set of singular values.

2. Energy: The largest singular values is related to spectral norm of the matrix A., i.e., $\sigma_1 = \|A\|_2$, while the square root of sum of its squared SVs is equal to Frobenius norm, i.e.

$$\|A\|_F = \left[\sum_{i=1}^{r}\sigma_i^2\right]^{\frac{1}{2}}$$

(4)

3. Stability: The SVs have a strong stability since the variation of both original and disturbed SVs cannot exceed 2-norm of the difference between the original and modified matrices.

4. Invariance to geometric distortions: The SVD exhibits the geometric invariance can be expressed in: Translational invariance: Both the matrix A and its translated counterpart has same singular values.

2.3 Local Binary Pattern:

The Local Binary Pattern initially appeared as a generic texture descriptor. The LBP operator gives a label for each pixel of an image by limiting a 3x3 neighborhood with the center pixel value and assuming the result as a binary number. In various publications, the circular 0 and 1 output values are read either clockwise or anti clockwise. In this research, the binary result will be computed by reading the values clockwise, starting from the top left neighbor, as shown below:
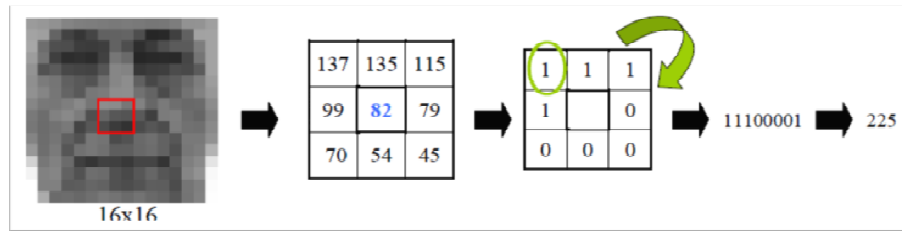

Figure2. LBP labeling: binary label is read clockwise starting from top left neighbor.

In other words, given a pixel position (xc, yc), LBP is defined as an ordered set of binary Comparisons of pixel intensities between the central pixel and its surrounding pixels. The resulting decimal label value of the 8-bit word can be expressed as follows:

$$LBP(x_c, y_c) = \sum_{n=0}^{7} s(l_n - l_c)2^n$$

(5)

Where lc corresponds to the grey value of the center pixel (xc, yc), ln to the grey values of the 8surrounding pixels, and function s (k) is defined as:

$$s(k) = \begin{cases} 1 & if \quad k \geq 0 \\ 0 & if \quad k < 0 \end{cases}$$

(6)

LBP can be extended as follows:

In order to treat textures at different scales, the Local Binary Pattern operator was extended to use neighborhoods of various sizes. By using the circular neighborhoods and bilinear interpolation of pixel values, any radius and number of samples in the neighborhood can be treated. Hence, the following notation is defined:

(P, R) which means **P** sampling points on a circle of radius **R**.

The following figure shows some examples of different sampling points and radius:

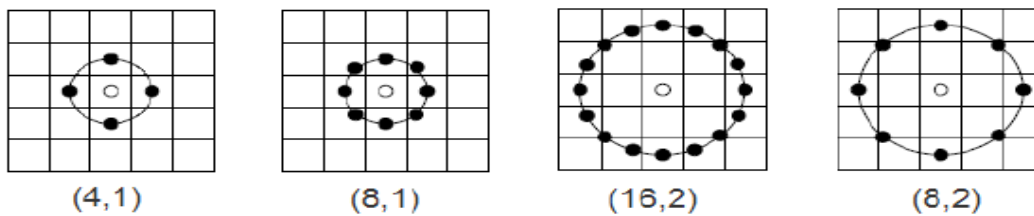(4,1)   (8,1)   (16,2)   (8,2)

Figure3. LBP different sampling point and radius examples.

In (4,1) LBP case, the four points selected correspond to vertical and horizontal ones, because that faces contain more vertical and horizontal edges than diagonal ones. When computing pixel operations considering the NXN neighborhoods at an image boundary, a portion of the NXN mask is off the image edge. In these cases, various padding techniques are generally used such as repeating border elements, zero-padding or applying a mirror reflection to define an image border. Also, in LBP operator case, the critical boundary, defined by the radius R of the circular operation, cannot be solved padding technique, so the operation is started at an image pixel (R, R). The benefit of this is the final LBP labels histogram will be not influenced by the borders, as the resulting LBP labels image size will be reduced to Width-R X Height-R pixels.

3. FEATURE VECTOR PARAMETERS:

Texture features are a sort of internal visual features which do not based on brightness or color. These features comprise the both surface information and surrounding environment of the image. Texture features can also describe the spatial information of an image quantitatively. Haralick et al. defined 14 feature parameters of gray level co-occurrence matrix for analyzing texture , the study found that only 4 features are not related, namely angular second moment also called energy, moment of inertia also called contrast, entropy and correlation. In our work we use 3 feature parameters; they are energy, entropy and contrast.

3.1 Energy:

Energy is also known as uniformity or the angular second moment. It provides the sum of squared elements in the Gray Level Co-occurrence Matrix. Energy is 1 for a constant image. The energy is given by:

$$\text{Energy} = \sum_i \sum_j Pd^2(i, j)$$

(7)

3.2 Entropy:

The entropy of an image can be calculated as

$$\text{Entropy } H = -\sum_{i=0}^{L-1} P_i \ln Pi$$

(8)

3.3 Contrast:

Measures the local variations in the gray-level co-occurrence matrix. It reflects the depth of texture grooves and image clarity. Contrast is 0 for a constant image. The contrast is given by

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

(9)

4. DISTANCE MEASURES:

In image analysis, the distance transform measures the distance of each object point from the nearest boundary and is an important tool in image processing and pattern recognition. In the distance transform, binary image specifies the distance from each pixel to the nearest non-zero pixel. The distance transform provides a metric or measure of the separation of points in the image. The distance metrics used in this retrieving of images are Euclidean distance, Manhattan distance and Canberra distance.

4.1 Euclidean Distance:

The Euclidean distance of two images is given as:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2}$$

(10)

4.2 Manhattan Distance:

The Manhattan distance of two images is given as:

$$d(x, y) = \sum_{i=1}^{n} (x_i - y_i)^2$$

(11)

4.3 Canberra Distance:

The Canberra distance of two images is given as:

$$d(x, y) = \sum_{i=1}^{n} \frac{|x_i - y_i|}{|x_i| + |y_i|}$$

(12)

5. EXPERIMENTAL RESULTS:

The above mentioned three algorithms are validated through simulation experiments using MATLAB. The Corel image database is used for experimentation; choose 1000 images which are associated with busses, horses, elephants, buildings, mountains, flowers, African people, beaches, and food. Twenty images are randomly selected from each of the ten kinds of images regarded as query images. The above three methods are respectively used to retrieve images.

| AVERAGE VALUES OF IMAGES USING 3 DISTANCE METRICS				
Category	**Distance Metric**	**Retrieved Images**	**Precision**	**Recall**
GLCM	Euclidean	23.410	0.234	0.023
	Manhattan	23.410	0.234	0.023
	Canberra	13.070	0.133	0.013
GLCM & SVD	Euclidean	22.900	0.229	0.022
	Manhattan	23.610	0.234	0.023
	Canberra	22.960	0.229	0.022
LBP RADIUS 1	Euclidean	14.130	0.141	0.014
	Manhattan	14.480	0.141	0.014
	Canberra	20.280	0.201	0.020
LBP RADIUS 2	Euclidean	13.321	0.133	0.013
	Manhattan	13.380	0.216	0.021
	Canberra	19.433	0.194	0.019
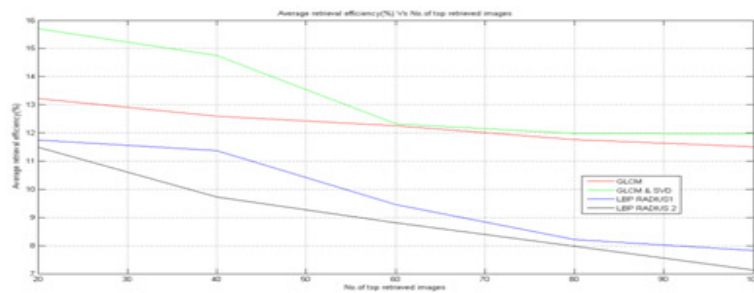
Figure 4. Percentage average retrieval efficiency on Corel 1K database using Euclidean Distance
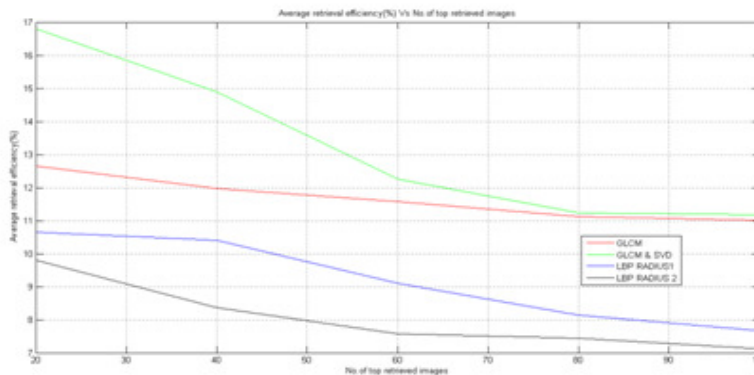

Figure 5. Percentage average retrieval efficiency on Corel 1K database using Manhattan Distance
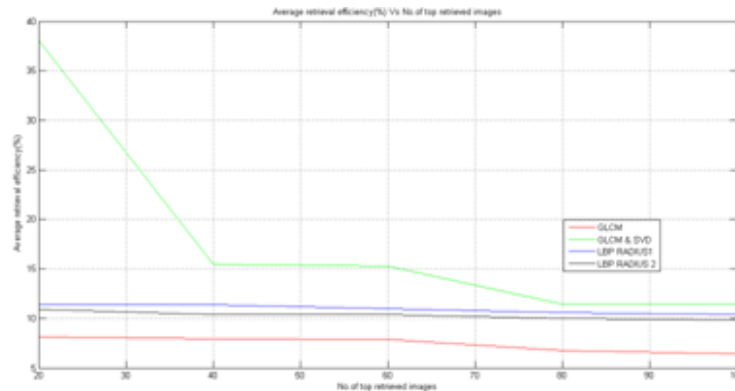
Figure 6. Percentage average retrieval efficiency on Corel 1K database using Canberra Distance

6. CONCULSION:

The CBIR technology can be used in several applications such as social networking sites, forensic labs and image search. CBIR system is a computationally an expensive task. Also it requires a lot of accuracy in the output. The accuracy in this paper has been increased by using various distance metrics such as Canberra distance, Manhattan distance and Euclidean distance on different algorithms such as GLCM, GLCM & SVD, LBP radius1 and LBP radius2 based on feature vector parameters such as energy, contrast, and entropy.

7. REFERENCES:

[1]   G.H.Liu, L.Zhang, Y.K.Hou, Z.Y. Li, J.Y. Yang, "Image retrieval based on multitexton histogram", Pattern Recognition vol.43, Issue No.7, pp. 2380–2389, 2010.

[2]   G.H.Liu, Z.Y.Li, L.Zhang, Y.Xu, Image retrieval based on micro-structure descriptor, Pattern Recognition Vol.44 , Issue 9, 2011.pp. 2123–2133.

[3]   R.Kwitt, P.Meerwald, A.Uhl, "Efficient texture image retrieval using copulas in Bayesian framework", IEEE Transactions on Image Processing Vol. 20 , Issue.7, 2011.

[4]   C.H.Lin, D.C.Huang, Y.K.Chan, K.H.Chen, Y.J.Chang, "Fast color-spatial featurebased image retrieval methods", Expert Systems with Applications V ol.38, 2011, pp. 11412–11420.

[5]   Z. Konstantinos, E. Kavallieratou, P. Nikos, "Image retrieval systems based on compact shape descriptor and relevance feedback information", Journal of Visual Communication and Image Representation Vol. 22 No.5, pp.378–390, 2011.

[6]   A. El-ghazal, O. Basir, S. Belkasim, "Invariant curvature-based Fourier shape descriptors", Journal of Visual Communication and Image Representation. Vol. 23, no. 4, pp. 622–633, 2012.

[7]   T. Celik, T. Tjahjadi, "Bayeesian texture classification and retrieval based on multiscale feature vector", Pattern Recognition Letters.Vol.32, Issue no.2, pp.159–67, 2011.

[8]   S. Murala, R.P. Maheshwari, R. Balasubramanian, "Local tetra patterns: a new feature descriptor for content-based image retrieval", IEEE Transactions on Image Processing Vol. 21, no.5, pp.2874–2886, 2012.

[9]   X.F. He, "Laplacian regularized D-optimal design for active learning and its application to image retrieval", IEEE Transactions on Image Processing. Vol. 19, pp.254–263, 2010.

[10] D.S.Zhang, M.M.Islam, G.J.Lu, "A review on automatic image annotation techniques", Pattern Recognition. Vol.45, No.1, pp.346–362, 2012.

[11] Z.Q.Hong, "Algebraic Feature Extraction of Image for Recognition", Pattern Recognition, Vol.24, No.3, pp.211-219, 1991.

[12] Ch.S.Rao and S.S.Kumar, "Content Based Image Retrieval using Singular Value Decomposition and Fuzzy Set Theoretic Technique," International Journal of Intelligent Information Processing, Vol.2, No.2, pp. 187-201, 2008.

[13] H.S.Prashanta, H.L.Shashidhara, and K.N.B.Murty, "Image Compression using SVD," Proc.of the international Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), Vol.03,pp.143-145,2007.

AUTHORS:

D. Kishore received the B.Tech degree from JNTU, AP, India in Electronics & Communication Engineering and Masters degree in Communications and Signal Processing from Bapatla Engineering College, Bapatla, AP, India in 2009. He is currently pursuing his Ph.D in Digital Image Processing at JNTUK, Kakinada.

Dr.S.Srinivas Kumar is currently working as Director (Research & Development) and Professor of Electronics & Communication Engineering at JNTUK, Kakinada, AP, India. He obtained his Ph.D from IIT, Karaghpur India. He received M.Tech degree from the same institute. He published 50 research papers in reputed International Journals and Conferences.

Dr. Ch. Srinivasa rao is currently working as Professor of ECE at JNTUK University College of Engineering, Vizianagaram, AP, India. He obtained his Ph.D. in Digital Image Processing area from University College of Engineering, JNTUK, Kakinada, A.P, India. He received M. Tech degree from the same institute. He published 20 Research papers in reputed International Journals and Conferences.