Image Recognition Method based on Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD)

Dr. Mohammad V. Malakooti (1), Zahed FerdosPanah (2) Dr. Seyyed Mohsen Hashemi (3)
(1) Faculty and Head of Department of Computer Engineering, Islamic Azad University, UAE Branch, Dubai
malakooti@iau.ae
(2) Graduate Student of Department of Computer Engineering, Islamic Azad University, UAE Branch, Dubai
ferdoszahed@yahoo.com
(3) Faculty of Department of Computer Engineering, Islamic Azad University, UAE Branch, Dubai, UAE
isrup@yahoo.com

Abstract: The search for one digital image among the vast quantity of images is a real dilemma when it is based on image content rather than metadata such as eigenvalues or frequency coefficients. The search results for most of the existing methods are satisfactory but still included irrelevant images for the target image. We have introduced a new Image recognition method based on Wavelet Transform (DWT) and Singular Value Decomposition (SVD) that is capable of retrieving most of the images similar to the target image. Our Method used DWT to transfer the target image from the spatial domain into frequency domain in which it is divided into four sub bands, Low-Low (LL), Low-High (LH), High-Low (HL) and High-High (HH) frequencies. We have applied three levels of 2-D DWT to concentrate the image illumination components into the third-level LL sub band and then applied SVD to extract its singular values as the reliable and robust features for the recognition. We also calculated the mean and variance of the low frequency LL sub-band for the input image to create a feature vector composed of mean, variance, and singular values of the LL sub band, as well as the coefficients of the LL sub band. Once the feature vector is formed it is compared with the feature vectors stored in our database to recognize and retrieve the images that are similar to target image based on the minimum error criteria.

Keywords: Image recognition, Image retrieval, DWT, SVD, feature extraction, Singular Values.

1. Introduction

The searches of images in the massive databases always were costly and time consuming and the result the researchers efforts in this area leads to the development of various image recognition systems. The earlier images were not classified in terms of their visual features but instead were classified in terms of the text annotation. Thus, the operator detects the concept of the image to be stored in the database as keywords of that image, and based on these keywords the users can access to the relevant image. This method of recognition is known as the text-based image recognition model. The test-based Image recognition is costly and requires more time and also depends on the user's understanding of the image. Since the concepts of images interpretation is totally different from one user to another, thus the attached annotations for the images are not exactly the same. To overcome the above image recognition problem, the content-based image retrieval system was introduced in the early 90's to extracts the images automatically on the basis of visual features such as color, texture and image layout [1,7]. When this model has not met the demand of the users, a new model is developed that enables users to input an image to the system and retrieve the similar images from the database. We used Mat-Lab simulation software to extract the features of the given image and compared with those features that stored in the database to find the images that are similar to given image based on the minimum error criteria. First we use the image read function, imread, to get the target image from the designated storage device. Then we apply three-level DWT to transfer the spatial target image into frequency domain that contains four sub-bands. Finally, the mean, variance, and singular values of the LL sub-band image are calculated to obtain the feature of the image required to build the feature vector for the recognition.

2. Singular Value Decomposition (SVD)

The SVD is one the strongest mathematical tools that can be used to decompose any square or non-square matrix, A, into the multiplication of two unitary matrices U and V and one Diagonal matrix D. SVD can be used to apply on any images to extract its useful features by decomposing the digital image matrix into three orthogonal matrices [2].The SVD of the matrix A can be written as following:

$$A = UDV^T \qquad (1)$$

Where the matrices of U and V are said to be unitary, the orthogonal matrices with length one, and D is a diagonal matrix comprised of the singular value of matrix A [2, 3].The columns of matrix U are called the left singular vectors, Orthogonal Eigenvectors of AA^T, and the columns of matrix V are called right singular vectors, Orthogonal Eigenvectors of A^TA, the diagonal elements of matrix D, the Square roots of the Eigen values of AA^T or A^TA , are called the Singular Values of the matrix A. The SVD represents an expansion of the original data in a coordinate system where the covariance matrix is diagonal [4].

SVD can be used to approximate any matrices with its best rank approximation using the Frobenius norm, where the smaller singular values are set to zero and the matrix is reconstructed with a fewer singular values than its original singular values. The SVD can be used for the calculation of pseudo inverse of the matrix A, Least square curve-fitting, matrix approximation, and determining the range and null space of a matrix.

3. Discrete Wavelet Transform (DWT)

DWT is an orthogonal transform similar to the Discrete Cosine Transform that can be used for the audio and video compression, Speech recognition, feature extraction, finger print, Watermarking and many other applications in biomedical engineering. DWT has better time and frequency resolution than DCT and its coefficients can be calculated by performing the successive Low pass and High pass filter on the Discrete-Time samples. There are many applications of DWT in the image processing field, such as feature extraction, face recognition, watermarking and compression [5].The first level DWT divides an input image into four sub-band images., where each sub-band image contains one of high frequency bands and low frequency bands: LL, LH, HL, and HH, where LL denote a low frequency sub-band, LH a horizontal high frequency sub-band, HL vertical high frequency sub-band, and HH a diagonal high frequency sub-band. [6]Here the third-level DWT decomposition LL3 sub-band image is important for us to reduce the input image size and extract features to be saved in the feature vector. The three-level DWT is illustrated in Figure 1.

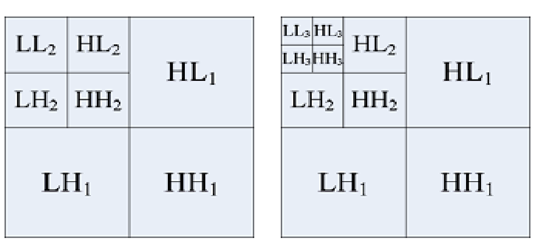

Figure 1- (a)-Two level DWT (b) Three-level DWT Figure 1-a shows two-level DWT and 1-b shows three-level DWT and its four sub-bands.

4. Proposed Image Recognition Algorithm

We have presented a new method for content based image recognition in which the features of every image stored in our data based are extracted and then stored into the feature vector. Once the feature vectors for all existing images are developed the new database consist of all feature vectors is formed and then stored inside our storage device. To retrieve all images that are similar to the target image, we must extract the features of the target image and compare it with all features vectors. Our algorithm calculates the distance between the image feature vector and all existing feature vectors and retrieves only those images that are similar to the target image based on our minimum distance criteria.

The image recognition steps are shown in figure (2).

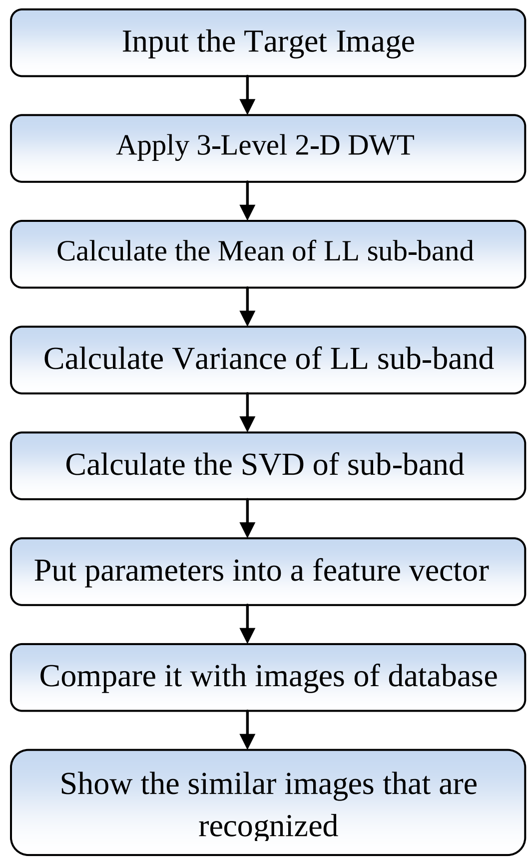

Figure 2: The Image Recognition Algorithm

4.1. Get the image file and form RGB Color Space

First we get the target image file and transfer it into RGB color space using the imread Matlab command. The imread function gets the image file and then returns a three-layer, color image matrices, represents the Red, Green, and Blue color matrices in which the

DWT and other transforms are applied on each matrix separately.

4.2 Apply Three-Level, Two-Dimensional DWT

To reduce the redundancy in the input image and decrease the number of elements in the feature vector we applied three-level 2-D DWT on the input image. For each level of decomposition, the input image is transformed into four bands of frequency, LL, LH, HL, and HH. We have selected the low frequency LL band because it contains the most useful features of the input image while the other three bands of frequencies only contain the details and can be ignored. We have calculated the size of LL sub-Band from equation (2).

$$LL\_Size = \frac{Original\_Image\_Dim}{8} \quad (2)$$

Each time we apply DWT on the input image, the LL sub-band will be one-half of the input image. Thus when we apply three-level DWT on the input image the size of LL sub-band would be 1/8 of the input image dimension. Figure-3 clearly has illustrated the process of three -level DWT, level by level.

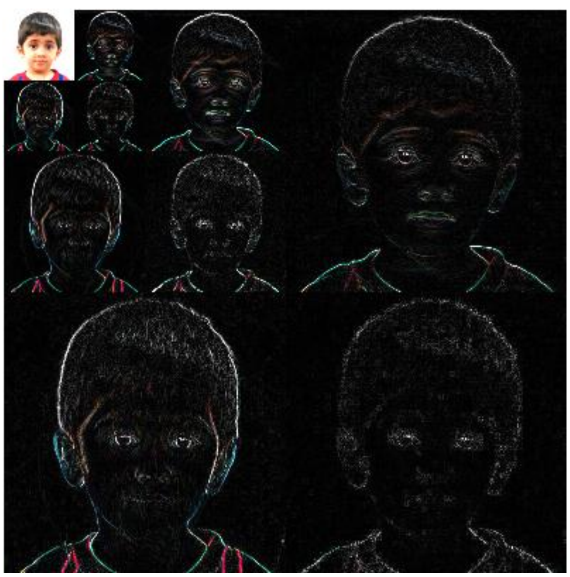

Figure 3: Three -Level DWT of the Original Image.

Once the input image is decomposed into four sub-bands then the LL sub-bands of Red, Green, and Blue matrices will be converted from 2-D matrix into 1-D array respectively. The elements of these arrays along with the mean and variance of the input image will be used to form three feature vectors, Red, Green, and Blue for the image recognition based on the DWT. The Matlab code for the conversion is shown as following;

```
for i:1 :LL_Size
  for j:1:LL_Size
Feature_Vect_R(i*j)=R_LL_S_band(i,j);
Feature_Vec_G(i*j)=G_LL_S_band(i,j);
Feature_Vect_B(i*j)=B_LL_S_band(i,j);
  end;
end;
```

4.3 Mean and Variance of Sub-band Matrices

The mean of Red, Green, and Blue low frequency sub-bands matrices can be calculated as following:

$$mean = \frac{1}{m*n}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{i,j} \quad (3)$$

Where m and n are equal to the size of low frequency sub-band matrix, LL_Size, and $I_{i,j}$ values are the elements of the low frequency LL sub-bands. The size of all three sub-band matrices, Red, Green, and Blue are the same and above formula can be applied for all of them.

Once the mean of each low frequency LL sub-band is calculated from equation (3) the result will be appended at the end of corresponding feature vector.

The Variance of Red, Green, and Blue low frequency sub-bands matrices can be calculated as following:

$$Var = \frac{1}{m*n}\sum_{i=1}^{m}\sum_{j=1}^{n}(I_{i,j} - mean)^2 \quad (4)$$

Where m and n are equal to the size of low frequency sub-band matrix, LL_Size, and $I_{i,j}$ values are the elements of the low frequency LL sub-bands, and mean is the mean value of corresponding matrix, Red, Green and Blue. The size of all three sub-band matrices are the same and above formula can be applied for all of them. Once the variance of each low frequency LL sub-band is calculated from equation (4) the result will be appended at the end of corresponding feature vector.

4.4 Apply Singular Value Decomposition (SVD

We have applied the Singular Value Decomposition on low-frequency LL sub-bands of the input image, using the Equation (1), but only extracted the singular values of Red, Green, and Blue Low-frequency LL sub bands[8].

Once the singular values of each low frequency LL sub-band is calculated from equation (1) the result will be appended at the end of corresponding feature vector and the process of the feature extraction is completed.

5. Comparison of the Input Image Feature Vectors with the Database Feature Vectors

Our proposed algorithm extracted all features of the input image and inserted them into the feature vector. Then, these feature vectors were compared with the features in the database feature vectors which were stored previously. Our comparison criteria is based on the Euclidean distance calculation between the feature vector of the input image and each feature vector previously stored in the database for the Red, Green and Blue feature vectors. Once the Euclidean distances are calculated the from Equation (5) the images corresponding to the closest feature vector are retrieved from the database and then displayed into the recognition window.

$$\mathrm{d}(\mathbf{p},\mathbf{q}) = \mathrm{d}(\mathbf{q},\mathbf{p}) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \qquad (5)$$

Where p_i and $\mathbf{q_i}$ are the input image feature vector and database feature vector, respectively. The number of the similar images to the input image will be affected by the threshold value for the distance measure.

It is obvious that whenever the threshold number is large, the number of irrelevant images that are retrieved is increased but whenever it is small the number of irrelevant images that are retrieved is decreased. Thus with smaller threshold the result the recognition accuracy is higher, and only those images with high similarity to the input image are retrieved.

In our proposed method we have combined the features obtained from both DWT and SVD algorithms to get a better result over the existing methods for the input image recognition. Our feature vector consists of the mean, and variance of the low frequency LL sub-band, the coefficients of LL sub-band, as well as the singular values of the low frequency LL sub-band. The experimental results clearly shown that our proposed method has retrieved those images that are very close to the input image with a high degree of similarity and the numbers of irrelevant retrieved images are decreased.

6. Experimental Results

To be able to compare the performance of existing methods with our model the size of all images in the database are selected to be 200 by 200.The images retrieve from the database when DWT is applied and only coefficients of the DWT are selected to build the feature vector is shown in Figure-4a.

Figure 4a: Retrieved Images based on DWT

When the SVD algorithm is applied on the input image and only singular values of the low frequency LL sub-band are used to build the feature vector the retrieved images are illustrated in Figure-5a.)

Figure 5a: Retrieved Images based on SVD

The Figure-4b has shown the graph of similarity measures for 10 closer images when the DWT technique is applied on the input image.

The Figure-5b has shown the graph of similarity measures for 10 closer images when the SVD technique is applied on the input image.

ISBN: 978-0-9853483-3-5 ©2013 SDIWC

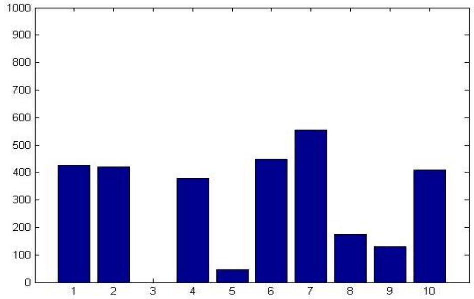

Figure 4b: Graph of Image Similarity based on DWT

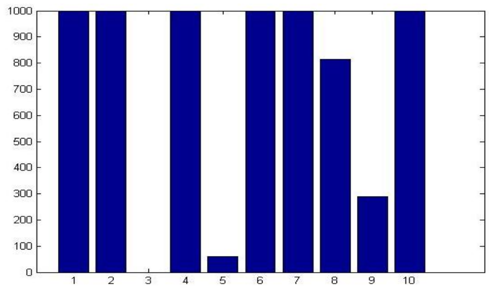

Figure 5b: Graph of Image Similarity based on SVD

But when both DWT and SVD are applied on the input image and mean and variance of the Low frequency LL sub-band are calculated, then feature vector would have useful information image retrievals as depicted on figure- 6a.

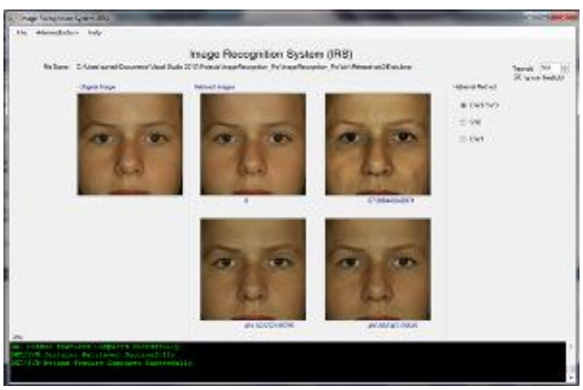

Figure-6a: Retrieved Images based on DWT and SVD as well as the Mean and Variance

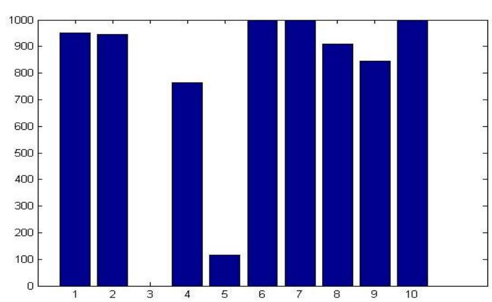

Figure-6b: Graph of Image Similarity based on SVD and DWT as well as the Mean and Variance

7. Conclusion:

In this paper we proposed a new content-based image retrieval method based on both DWT and SVD when they applied on the input image. We have applied three-level DWT on the input image to reduce the image redundancy and to divide the processed image into four sub-bands. Then we applied the SVD on the low frequency LL sub-band to extract its singular values. We also build a feature vector consists of the mean and variance of LL sub band, coefficients of low frequency LL sub-band, as well as its singular values to be used for a reliable and robust image recognition and retrieval. The combination of all these features empowered our system to retrieve similar images very accurately. To demonstrate the performance of our method over the existing ones we applied our model on a database of containing 500 images [9]. The retrieved images that are highly similar to the input image are obtained by calculation of the Euclidean distance between the input image feature vector and all existing feature vector in database. If the distance measure is zero the retrieved image will be exactly as the input image and also indicates that corresponding image was previously stored in the data. The retrieved images in Figure -6a clearly showed the superiority of our model over those in which the feature vectors are based on DWT, Figure- 4a, or the feature vectors are based on SVD, Figure-5a. Table 1 compares the similarity measures of the method based on DWT, SVD, or DWT/SVD.

Table 1- Similarity for the models based on SVD, DWT or both SVD and DWT

Picture	SVD	DWT	DWT & SVD
P1	1000	410	950
P2	1000	400	940
P3	0	0	0
P4	1000	370	750
P5	50	40	100
P6	1000	430	1000
P7	1000	520	1000
P8	800	160	870
P9	300	80	800
P10	1000	400	1000

References:

[1] Z. Mansoori, M.Jamzad."Content based image retrieval using the knowledge of texture, color and binary tree structure". 2009 IEEE.

[2] N.S.T. Sai , R.C. Patil."SVD Based Features for Image Retrieval" .IJCSAI Vol.2 Issue 2 2012 PP.10-17.

[3] XiangliXu, Libiao Zhang, Zhezhou Yu, Chunguang Zhou ."The Application of DWT and SVD in Image Retrieval".IEEE 2008 Congress on Image and Signal Processing

[4] Alter O, Brown PO, Botstein D, "*Singular value decomposition for genome-wide expression data processing and modeling*", Proc Natl. Academic Sci U S A, 97- 101, 2000.

[5] Ganesh Nagasai Prasad. V, Habibullah khan ,Bhavana.k , Muralidhar.Ch ,Tulasisai kiran.Ch. "Image enhancement using Wavelet transforms and Singular Value Decomposition (SVD)".IJEST Vol. 4 No.03 March 2012.

[6] Mohammad V. Malakooti, Mohammad Ali Neamatollahi, "Image Watermarking based on Discrete Wavelet Transform and Singular Value Decomposition Method", IAU Majlesi branch, 2011

[7] W. M. a. H. J. Zhang, "Content-based Image Indexing and Retrieval": CRC Press, 1999.

[8] Huirao Zheng, Shaolin Chen, Zhongxi Mo etal., "Numerical Calculation Method ". Wuhan: Wuhan University Press, 2002.

[9] http://gps-tsc.upc.es/ GTAV/ ResearchAreas/ UPCFaceDatabase/GTAVFaceDatabase.htm