Image Retrieval Using Color-Texture Features Extracted From Gabor-Walsh Wavelet Pyramid

Sajjad Mohammadzadeh*
Department of Electronics and Communications Eng., University of Birjand, Birjand, Iran
s.mohammadzadeh@birjand.ac.ir
Hasan Farsi
Department of Electronics and Communications Eng., University of Birjand, Birjand, Iran
hfarsi@birjand.ac.ir
Received: 22 Jul 2013 Accepted: 04 Nov 2013

Abstract
Image retrieval is one of the most applicable image processing techniques which have been extensively used. Feature extraction is one of the most important procedures used for interpretation and indexing images in Content-Based Image Retrieval (CBIR) systems. Effective storage, indexing and managing a large number of image collections are critical challenges in computer systems. There are many proposed methods to overcome these problems. However, the rate of image retrieval and speed of retrieval are still interesting fields of researches. In this paper, we propose a new method based on combination of Gabor filter and Walsh transform and Wavelet Pyramid (GWWP). The Crossover Point (CP) of precision and recall are considered as metrics to evaluate and compare different methods. The Obtained results show using GWWP provides better performance in compared to with other methods.

Keywords: Content-Based Image Retrieval (CBIR), Gabor Filter, Walsh Transform, Wavelet Pyramid Transform, Texture.

1. Introduction

Content Based Image Retrieval (CBIR) techniques are one of the most applicable and increasingly important topics in multimedia information systems [1]. An important building block in the image retrieval system is image indexing. Image indexing is one of the most important procedures used for interpreting and indexing images in the CBIR systems [3]. There are many proposed methods and approaches for classification, indexing, searching and retrieval of visual information based on analysis of low-level image features, like color, texture, and shape [4].

Effective storage, transmission, indexing, and managing a large number of image collection are serious challenges in computer systems [5]. Recently these challenges have been studied on different image databases and it has been attempted to solve these problems in computer vision [6] and image processing [7], [8]. The main goal of researchers in this field is to find a procedure to locate a desired image in a large and varied collection of image database. Traditional problems of image indexing methods such as taking a long time for manually indexing and huge required storage have led to rise of interest in retrieving images based on automatically derived features such as color, texture, and shape which is known as CBIR [9], [10]. Nowadays CBIR technology is known as a form of commercial products such as QBIC [7] and Virage [8] in marketplace. However, because of some practical issues like, absence of hard evidence on the effectiveness of CBIR, this technology has not been used on the significant scale [8]. Many CBIR technology applications have been identified [7]. Medical, industrial and internet applications are some important examples of the applications. Nowadays color, texture or shape features are intensively used in image indexing. The combination of these features also showed more efficient performance in image retrieval [7].

In this paper we combine color and texture features and define the average of red, green and blue planes as gray plane. Then we extract some texture features from Gray plane. There are many different proposed methods to describe image texture. Texture analysis methods are divided into four categories: signal processing, model-based, geometrical and statistical introduced by Tuceryan and Jain [11]. We proposed only signal processing method for extracting texture feature. Texture feature is useless in image discrimination, if the variations of image intensity are highly uniform or non-uniform. The size of feature vectors and speed of retrieval are important aspects in performance of image retrieval. In this paper a novel method for image retrieval is proposed by using Gabor filter, Walsh transform and Wavelet Pyramid (GWWP) and then applied in database.

This paper has been structured as follows. In section 2 overview of the proposed CBIR system has been explained. Section 3 has been discussed parts of feature extraction proposed method. Haar transform has been explained in section 4. Similarity measurement is defined in section 5. Image retrieval using GWWP method has been described in section 6. In section 7, the obtained results have been presented.

* Corresponding Author
2. Content based image retrieval

Texture, color and shape features are basic features which used in CBIR systems. Texture and color features are absolutely easy for computing similarity. Some CBIR systems have combined texture and color feature to provide better performance and automatically retrieve relevant images from a large image database [12]. The standard CBIR system involves two important parts [12]. The first part extracts image features. This includes generating feature vectors of image in the database and representing the content of image accurately. The size of feature vectors must be extremely smaller than that of primary image. Similarity measurement is the second part of CBIR system. This part computes a distance between the query image and each image of database by using feature vectors of query image and each image in the database to obtain similar images.

Block diagram of the proposed CBIR system is shown in Figure 1. In this method, before extracting GWWP feature and generating feature vector, all images are resized to $256 \times 256 \times 3$ and gray plane with size of $256 \times 256$ is generated by averaging red, green and blue planes.

3. Feature extraction method

We have combined Gabor filter, Walsh transform and Wavelet Pyramid (GWWP) to extract texture features of image. The flowchart of GWWP method is represented in Figure 2 and the steps are represented as follows:

a. Applying Gabor filter (see section 3.1) on the gray plane with size of $N \times N$ separately. Gabor filter detects edges and lines of plane (image) and gives a maximum and suitable response at the edge. Orientation and frequency 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 [13].

b. Performing wavelet transform with size of $N \times N$ on the filter output with size of $N \times N$ to generate approximation (Low-Low), horizontal (Low-High), vertical (High-Low) and diagonal (High-High) components which are explained in section 3.2. We used approximation component for next step.

c. To construct modified approximation component by applying Walsh transform (see section 3.3) on the approximation component in step ‘b’. Walsh transform can be reduced to subtraction and addition operations (no division or multiplication). This allows the use of simpler hardware and low complexity to calculate the transform and increases the speed of retrieval.

d. Applying inverse wavelet transform with modified approximation component and zeroing horizontal, vertical and diagonal components. This method is wavelet pyramid transform which is completely explained in section 3.2. New image is constructed by using inverse wavelet whereas some information are lost due to removing horizontal and vertical components. In next level, we need to new image to construct the new approximation and diagonal components.

e. To take alternative rows and columns by down-sampling the output in step ‘d’ with size of $N/2 \times N/2$. Down-sampling reduces the size of feature vector which is very important for increasing the speed of retrieval.

f. To construct GWWP of level-p by repeating steps ‘b’ to ‘e’, ‘p’ times on each plane.

We consider approximation component level-p in step ‘c’ as GWWP feature of each image and store this feature as feature vector of image.

3.1 Gabor filter

Gabor filter or Gabor wavelet is a method used to extract features of image through analysis of the frequency domain rather than the spatial domain [13]. Eq. (1) represents the Gabor filter, where $x$ and $y$ shows the position of pixel in the spatial domain, $\omega_0$ is radial center frequency, $\theta$ shows the orientation of the Gabor direction and $\sigma$ is defined as the standard deviation of the Gaussian function along with the $x$ and $y$ axes where $\sigma_x = \sigma_y = \sigma$ [13]:

$$\psi(x, y, \omega_0, \theta) = \frac{1}{2\pi \sigma_\omega} \exp\left(-\frac{(xcos\theta + ysin\theta)^2}{\sigma^2} + \frac{(-xsin\theta + ycos\theta)^2}{2\sigma^2}\right) \times \exp\left[i \left(\omega_0 x cos\theta + \omega_0 y sin\theta\right) - e^{-\omega_0^2 \sigma^2/(2)}\right]$$

(1)

According to Eq. 1 Gabor filter can be decomposed into real and imaginary parts, which are given by:

$$\psi_r(x, y, \omega_0, \theta) = \frac{1}{2\pi \sigma_\omega} \exp\left(-\frac{(x^2+y^2)}{2\sigma^2}\right) \cos\omega_0 \theta$$

(2)

$$\psi_i(x, y, \omega_0, \theta) = \frac{1}{2\pi \sigma_\omega} \exp\left(-\frac{(x^2+y^2)}{2\sigma^2}\right) \sin\omega_0 \theta$$

(3)

Where:

$x' = x \cos \theta + y \sin \theta$, $y' = -x \sin \theta + y \cos \theta$

We consider $\sigma = \pi / \omega_0$. Gabor features, $C_{\psi}$, can be achieved by using the convolution of image, $I$, and Gabor filter, $\Psi$, as represented in Eq. 4 [13]:

$$C_{\Psi_I} = I(x, y) \ast \psi(x, y, \omega_0, \theta)$$

(4)

As mentioned, Gabor filter can be written as the summation of real and imaginary parts shown in Eq. 2 and Eq. 3. Therefore we can compute real and imaginary parts of $\psi(x, y, \omega_0, \theta)$, represented by $C_{\psi_r}$ and $C_{\psi_i}$ respectively, by replacing Eq. 2 and Eq. 3 in the term of $\psi(x, y, \omega_0, \theta)$ in Eq. 4. Local properties of the image can be achieved using real and imaginary parts, which is given by:

$$C_{\Psi_I} = \left(\|c_r\|^2 + \|C_i\|^2\right)^{1/2}$$

(5)

A fast convolution method is utilized using scanning windows by applying a one-time convolution with Fast Fourier Transform (FFT), point-to-point multiplication and Inverse Fast Fourier Transform (IFFT). The values of radial
center frequencies and orientations used in this paper are shown in Eq. 6, where \( n \in \{0, 1, 2 \} \) and \( m \in \{0, 1, 2, \ldots, 7 \} \) [13]:
\[
\omega_n = \frac{\pi}{2 \times 2^n}, \quad \theta_m = \frac{\pi}{8} m
\] (6)

3.2 Wavelet pyramid transform

Wavelet transform is multi-level signal decomposition. It represents a signal as a basis function superposition called wavelets [14]. Wavelet transform has several interesting properties [14]. The principal property of the proposed wavelet feature analyzes the signal in at various frequency bands giving higher frequency resolution and lower time resolution at lower frequencies, lower frequency resolution and higher time resolution at higher frequencies which have been shown in Figure 3 [14]. For a given image with size of \( N \times N \), the two-dimensional Haar wavelet transform includes \( \log_2 N \) stages. The first stage provides four sets of coefficients known as, approximation coefficients \( c_A \), horizontal coefficients \( c_H \), vertical coefficients \( c_V \), and diagonal coefficients \( c_D \). These sets are computed by convolving columns or rows of image with the low-pass filter for approximation, and with the high-pass filter, which are followed by dyadic decimation (down-sampling). The length of these filters is \( 2n \) sample. Therefore, if the length of an image is \( N \), then the length of output signal using low-pass and high-pass filters will be \( N + 2n - 1 \) [15], [16]. Figure 4 describes a flowchart of the basic decomposition of wavelet transform for an input image. The next step splits the approximation coefficients \( c_A \), in two separate parts using the same method, described above, replacing input image (s) by \( c_A \), and producing \( c_A \), \( c_H \), \( c_V \), and \( c_D \), and so on [17]. As an example, three levels of wavelet pyramid is shown in Figure 5.

3.2.1 Haar transform

Haar transform was proposed in 1909 by Alfréd Haar [20]. Haar used this transform to give an example of a countable orthonormal system for the space of square-integrable functions on the real line [20]. The Haar wavelet is also one of the simplest possible wavelet. The principal property of Haar wavelet transform is one of the technical disadvantages of this method. However, this property can be considered as an advantage for analysis of signals with sudden transition like monitoring of tool failure in machines [21].

The Haar mother wavelet function \( \psi(t) \) can be described as:
\[
\psi(t) = \begin{cases} 
1 & 0 \leq t \leq \frac{1}{2} \\
-1 & \frac{1}{2} \leq t \leq 1 \\
0 & \text{otherwise}
\end{cases}
\] (7)

And its scaling function \( \varphi(t) \) is given as:
\[
\varphi(t) = \begin{cases} 
1 & 0 \leq t \leq 1 \\
0 & \text{otherwise}
\end{cases}
\] (8)

3.3 Walsh transform

Walsh transform matrix [18] is defined as a set of \( N \) rows, denoted \( W_j \), for \( j = 0, 1, \ldots, N - 1 \). The properties of Walsh transform matrix are described as:

a. \( W_j \) takes on the values +1 and -1.

b. \( W_j[0] = 1 \) for all \( j \).

c. \( W_j \times W_k^T = 0 \), for \( j \neq k \) and \( W_j \times W_k^T = N \), for \( j = k \).

d. \( W_j \) has exactly \( j \) zero crossings, for \( j = 0, 1, \ldots, N - 1 \).

Each row \( W_j \) is either even or odd with respect to its midpoint.

Hadamard matrix of order \( N \) is used to define Walsh transform matrix. The row of the Walsh transform matrix is the row of the Hadamard matrix determined by the Walsh code index, which is an integer in the range \( [0, \ldots, N - 1] \). For the Walsh code index equal to an integer \( j \), the respective Hadamard output code has exactly \( j \) zero crossings, for \( j = 0, 1, \ldots, N - 1 \).

Following stages show the Kekre’s Algorithm to generate Walsh Transform from Hadamard matrix [19]:

a. The \( N \) coefficients of Walsh matrix are arranged in a row and then the row is split to segments with length of \( N/2 \), one segment in forward order and the other part is written in reverse order:

\[
\begin{bmatrix}
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\
15 & 14 & 13 & 12 & 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\
\end{bmatrix}
\]

Now we have two rows, each of these rows are again split in \( N/2 \) parts and other part is written in reverse order below the upper rows as:

\[
\begin{bmatrix}
0 & 1 & 2 & 3 \\
15 & 14 & 13 & 12 \\
7 & 6 & 5 & 4 \\
8 & 9 & 10 & 11 \\
\end{bmatrix}
\]

This step continues, until a single column giving the order of Hadamard rows, is achieved. The result sequence is presented in following:

\[
[0, 15, 7, 8, 3, 12, 4, 11, 1, 14, 6, 9, 2, 13, 5, 10]
\]

b. Based on this sequence the Hadamard rows are arranged to generate Walsh transform matrix. The Walsh transform of the given image is calculated by product of Walsh matrix and image matrix.

The number of additions required to apply Walsh transform on an image with size of \( N \times N \) are \( 2N^2 \times (N - 1) \) [19].

4. Similarity measurements

Traditional Euclidean Distance (ED) is the most common metric used to compute match or similarity value in CBIR system to obtain relevant images.

If Feature Vector of Database (FVD) and Feature Vector of Query (FVQ) are two-dimensional feature vectors of database image and query image respectively, then the Euclidean distance of these feature vectors is obtained by:
\[
ED = \sqrt{\sum_{i=1}^{n} (FVD_i - FVQ_i)^2}
\]

(9)

5. Image retrieval using GWWP method

5.1 Feature extraction

We generate feature vectors for image database by applying GWWP level-1, level-2, ..., level-7 and store approximation components as feature vectors for each image. The size of feature vector is \(\frac{N}{2} \times \frac{N}{2}\) in GWWP level-1 and the size of feature vector in GWWP level-2 is \(\frac{N}{4} \times \frac{N}{4}\). Gray image is defined as the average of red, green and blue planes components and is used to generate Gray-GWWPs feature vector to obtain respective images for different levels.

5.2 Query execution

We use the proposed method to extract query feature and generate feature vector of query image. Then relevant images are retrieved by comparing feature vector of query with feature vectors of database in level-p using Euclidian distance as similarity measure.

The proposed method reduces the size of feature vector and computation time extremely in high level of GWWP and gives better precision and recall values. Complete GWWP needs \(2N^2 \times (N-1)\) additions and GWWP of level-p needs \(2(\frac{N}{2})^2 \times ((\frac{N}{2})^2 - 1)\) addition for image with size of \(N \times N\).

6. Experimental Results

We have compared performance of the proposed method with precision and recall criteria. These two standard criteria are given by:

\[
\text{Precision} = \frac{\text{Number of Relevant Images Retrieved}}{\text{Total Number of Images Retrieved}}
\]

(10)

\[
\text{Recall} = \frac{\text{Number of Relevant Images Retrieved}}{\text{Total Number of Relevant Images in Database}}
\]

(11)

In order to evaluate the proposed method in image retrieval, we first provide an image database [22] including 1000 variable size images. These images are classified in 11 classes of human being, horse, elephant, flower, bus, man made thing and natural scenery. We have tried to collect a diverse database set to evaluate the proposed method. Figure 6 shows sample images from the database. After collecting a proper image database, we select 5 images from each class (55 random images) as query images to evaluate the proposed system with other methods. As described in previous sections, the features of each query image is extracted using GWWP method, and feature vectors of query images are constructed. Then Euclidian distance between each query feature vector of image and database feature vectors are computed by using Eq. 9. The obtained distances are sorted based on which images have minimum distance with query image to find the best relevant images in database. Then, the number of relevant images is computed and the precision and recall for each number of retrieved images for all 55 query images are obtained by using Eq. 10 and Eq. 11. We next consider the average of these 55 precisions and recalls for each number of retrieved images as the precision and recall for each method.

For example in Figure 7 sixteen closest images for the sample query image have represented by using the GWWP level-5 and computing Euclidian distances between query image and all images in database and sorting distances based on minimum distance. The average precision and recall for Haar wavelet, Walsh wavelet and GWWP methods (the proposed method) are obtained and plotted by grouping the number of retrieved images in Figures 8, 9 and 10, respectively. Note that Haar wavelet method has not used Walsh matrix and Gabor filter and Walsh wavelet has not utilized Haar transform and Gabor filter. According to these Figures it is obvious that the precision decreases and the recall increases by increasing the number of retrieved images. Also by increasing wavelet level until level-5, the precision and the recall increase but they decrease after level-5. For example the precision of GWWP level-1, 3 and 5 are 0.358, 0.366 and 0.382 in 90 of retrieved image respectively but the precision of GWWP level-7 is 0.337. Therefore precision and recall level-5 of wavelet methods are better than other wavelet levels. Moreover it is observed that precision of Haar wavelet and Walsh wavelet level-5 are 0.338 and 0.344 respectively, whereas the precision of GWWP level-5 is 0.384 in 88 of retrieved image. Based on these figures the proposed GWWP method provides better performance rather than Haar wavelet and Walsh wavelet.

The other important parameter for comparing CBIR techniques is the percent of CP of precision and recall [11] which the CP occurred in 91 of retrieved image in our database because the most classes have 91 images that is specified in Figure 8, 9 and 10 clearly. The percent of CP for different levels of Haar wavelet, Walsh wavelet, GWWP, dominant color descriptor (DCD) [23], scalable color descriptor (SCD) [23] methods and the size of feature vector have been represented in Table 1. As observed in Table 1, the size of feature vector decreases by increasing wavelet level but the CP only increases until level-5. However the size of feature vector is 4 in level-7 but the CP (31.21% of CP of GWWP) is less than level-1 with 16384 size of feature vector which CP of GWWP level-1 is 35.81%. It is obvious that level-5 of each three wavelet methods with 64 size of feature vector has higher CP of precision and recall than other levels in each method. In Table 1, it could be observed that the GWWP level-5 with CP of 38.84% and size of 64 feature vectors provides the best performance compared to other methods. However, SCD method with CP of 37.05% is nearest to GWWP level-5, but size of feature vector in SCD method is 121 which is 1.9 times of feature vectors of GWWP level-5. The size of feature vector is very important parameter for increasing speed of computing distance and retrieval and
for decreasing storage space of feature vector. However, as shown in Table 1, the proposed GWWP method in level-1 (CP = 35.81%) has even better performance than Walsh wavelet (CP = 34.03%) and Haar wavelet (CP = 33.60%) in level-5. Therefore GWWP technique can be considered as a more powerful method than Haar, Walsh wavelet, DCD and SCD. Moreover, the proposed system not only reduces the size of feature vector and storage space but also improves the performance of image retrieval.

7. Conclusions

Image retrieval is an applicable technique finds relevant images in a large image database. By increasing the size of database, the challenge of average reduction in precision and recall and speed of retrieval become more serious. In this paper, we proposed a new method based on Gabor filter, Walsh transform and wavelet pyramid to improve the performance of image retrieval. We used an average of three planes of red, green and blue planes to generate gray plane to examine the proposed method. The obtained results show using Gabor filter before Walsh transform and wavelet pyramid can improve the performance of image retrieval in our database. Specially, wavelet in level-5 results in better performance in compared to other levels of wavelet pyramid. Moreover, the GWWP level-5 reduces the size of feature vectors and storage space and provides higher performance for image retrieval at the same time.

![Fig. 1. Block diagram of CBIR system.](image-url)
Fig. 2. Flowchart for generating GWWP of level-\(p\).

Fig. 3. Time frequency resolution of wavelet transform [14]
Fig. 4. Decomposition of the image or cA.

Fig. 5. Different levels of wavelet pyramid [12].

Fig. 6. Sample images of database.
Fig. 7. Image retrieved using GWWP level-5.

Fig. 8. Haar wavelet precision/ recall.

Fig. 9. Walsh wavelet precision/ recall.
Table 1. Size of feature vector and the present of CP of different methods

<table>
<thead>
<tr>
<th>Type</th>
<th>Size of feature vector</th>
<th>Crossover point</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haar wavelet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>16384</td>
<td>31.02%</td>
</tr>
<tr>
<td>Level 3</td>
<td>1024</td>
<td>31.60%</td>
</tr>
<tr>
<td>Level 5</td>
<td>64</td>
<td>33.60%</td>
</tr>
<tr>
<td>Level 7</td>
<td>4</td>
<td>28.33%</td>
</tr>
<tr>
<td>Walsh wavelet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>16384</td>
<td>31.02%</td>
</tr>
<tr>
<td>Level 3</td>
<td>1024</td>
<td>31.64%</td>
</tr>
<tr>
<td>Level 5</td>
<td>64</td>
<td>34.03%</td>
</tr>
<tr>
<td>Level 7</td>
<td>4</td>
<td>28.64%</td>
</tr>
<tr>
<td>GWWP wavelet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>16384</td>
<td>35.81%</td>
</tr>
<tr>
<td>Level 3</td>
<td>1024</td>
<td>36.67%</td>
</tr>
<tr>
<td>Level 5</td>
<td>64</td>
<td>38.34%</td>
</tr>
<tr>
<td>Level 7</td>
<td>4</td>
<td>31.21%</td>
</tr>
<tr>
<td>DCD</td>
<td>32</td>
<td>36.37%</td>
</tr>
<tr>
<td>SCD</td>
<td>121</td>
<td>37.05%</td>
</tr>
</tbody>
</table>

Fig. 10. Gray-GWWP method precision\recall.

References


Sajjad Mohammadzadeh received the B.Sc. degree in electrical engineering from Sistan & Balochestan, University of Zahedan, Iran, in 2010. He received the M.Sc. degree in communication engineering from South of Khorasan, University of Birjand, Birjand, Iran, in 2012. He is currently Ph. D student in Department of Electrical and Computer Engineering, University of Birjand, Birjand, Iran. His area research interests include Image Processing and retrieval, Pattern recognition, Digital Signal Processing and Sparse representation. His email address is: s.mohammadzadeh@birjand.ac.ir.

Hasan Farsi received the B.Sc. and M.Sc. degrees from Sharif University of Technology, Tehran, Iran, in 1992 and 1995, respectively. Since 2000, he started his Ph.D in the Centre of Communications Systems Research (CCSR), University of Surrey, Guildford, UK, and received the Ph.D degree in 2004. He is interested in speech, image and video processing on wireless communications. Now, he works as associate professor in communication engineering in department of Electrical and Computer Eng., university of Birjand, Birjand, IRAN. His Email is: hfarsi@birjand.ac.ir.