Compression of Facial Images using Object Based SPIHT

K. Somasundaram*1 and N. Palaniappan*2

*1 Image Processing Lab, Department of Computer Science and Applications, The Gandhigram Rural Institute – Deemed University, Gandhigram – 624 302. India.
'somasundaramk@yahoo.com

'naapala@gmail.com

Abstract - There are a lot of security threats facing the countries in the world now. Governments are adapting various security measures to meet out the threats. The facial images of the suspected persons can help to identify the persons easily at the airports, railway stations and other public places. The facial images can be stored digitally in a central database and can be verified at different locations using networks. Compression of facial images helps to reduce the storage space and the transmission time. In this study we propose a method for compressing the facial images. We adapted a method from the work of Monica Penedo et al which uses object based SPIHT method for compression of mammogram images. The CDF9/7 wavelet filter is used to transform the image into wavelet coefficients. Then the coefficients of the uniform background are removed using a binary mask. Finally SPIHT encoding is applied. Run length encoding is used to decrease the required bpp further. This method gives a comparable PSNR value with Jpeg2000.

Keywords - Facial image, DWT, CDF9/7, SPIHT, RLE

I. INTRODUCTION

Threats to normal life are going on increasing now. In order to reduce the threats various security measures are taken by the governments all over the world. Photos of criminals and militants can help the security people to identify them in public gathering places like airports, railway junctions etc. The photos showing just the facial features will be enough to identify. A database of facial images cannot be maintained at every place. A central data server with the facial image database connected by network to all the required places can serve for this purpose. Compression of facial images can reduce the transmission time, processing time and storage cost.

To recognize people from facial image showing image object without the background details is enough. So, the lossy image compression techniques can be adopted for this purpose which can give less storage cost and less transmission time. Among lossy compression techniques image transform techniques give better compression rates than the other methods. Discrete Cosine Transform (DCT), Karhunen-Loeve Transform (KLT), Walsh-Hadamard Transform (WHT) and Wavelet Transform (WT) are some of the popular transforms used in lossy image compression. Among these transforms Wavelet Transform is used widely.

Discrete Wavelet Transform (DWT) gives better compression rate with less number of bits. In the method proposed by Monica Penedo et al [1], DWT with Set Partition in Hierarchical Trees (SPIHT) coding are used to compress digital mammograms. The background area which is not required to identify the micro calcifications are removed while compressing image using a binary mask. Like digital mammogram images, facial images are also having a uniform background which is not having much importance.

In the proposed method, the basic process of mammogram compression by Monica Penedo et al [1] with some modifications is adopted. The edges of the facial image with uniform background are traced and a binary mask is created. The image is transformed by DWT with Cohen-Daubechies-Feauveau 9/7 (CDF9/7) filter. The wavelet coefficients representing the background area are made as 0 by using the binary mask. After that the wavelet coefficients are encoded with SPIHT coding to output a bit stream. Run Length Encoding (RLE) is used to compress the bit stream to reduce the required bit rate further.

II. OVERVIEW OF THE METHODS USED

The methods used in this proposed method are Canny Edge detector, Wavelet transform, SPIHT Coding and Run Length Encoding. These methods are given briefly for continuity.

A. Canny Edge Detector

The edges of an image are the representations of the discontinuities of image intensity function. Edge detection is the process of detection of these discontinuities in the image. Canny edge detection method is one of the most popular edge detection methods. This is based on first-derivative of Gaussian and noise cleaning. [2]
B. Discrete Wavelet Transform (DWT)

DWT is a series of matrix operations. By multiplying the image matrix with wavelet filter matrix the transformed wavelet coefficient matrix is obtained. Many number of wavelet filters like Byelkin, Coifman, Symmlet and Daubechies are available [3]. Among these filters the Cohen-Daubechies-Feauveau 9/7 (CDF9/7) is used in the proposed method. This gives fine and very smooth results. The famous image compression standard JPEG2000 also uses this filter for compression [4].

The multi resolutional concept of DWT decompose the given image into various sub bands as LL, LH, HL and HH. When the level of decomposition increases the number of sub bands also increases.

![Sub bands decomposed by DWT at 2 levels](image)

C. SPIHT Coding

Many encoding methods are used for encoding the transformed wavelet coefficients. Among them Embedded Zero Tree Wavelet (EZW) coding method by Shapirao [5] gives good compression rate and quality. This EZW method was modified by Said and Perlman [6] to a new method named Set Partition in Hierarchical Trees (SPIHT). This is an embedded coding method. The transmission can be stopped at any point during decoding to get the best quality image with the available bits.

The basic idea of SPIHT is based on spatial oriented trees. Three lists, List of insignificant pixels (LIP), List of insignificant sets (LIS) and List of significant pixels (LSP) are used for storing coefficients. There are two passes in SPIHT, sorting pass and refinement pass. In the sorting pass, the wavelet transformed coefficients C(i,j) are sorted according to their magnitudes and the number of iterations required to complete the coding is given by

\[ n = \log_2(\text{maximum}|C(i,j)|) \]  \hspace{1cm} (1)

In every iteration the coefficients which satisfy the condition \( 2^n \leq |C_{i,j}| < 2^{n+1} \) are selected and checked for significance. The coefficient \( C_{i,j} \) which satisfies \( |C_{i,j}| \geq 2^n \) becomes significant else insignificant. The iterations are processed with the thresholds \( 2^n, 2^{n-1}, 2^{n-2}, \ldots, 2^1, 2^0 \). When the threshold decreases the number of significant coefficients increases. The significant coefficients are stored in LSP.

The wavelet coefficients of C are separated into number of spatial oriented trees \( T_k \). To reduce the complexity of comparison, the sorting pass checks every \( T_k \) by performing the test:

\[ \text{Max } C(i,j) \in T_k, \ |C(i,j)| > 2^n \]  \hspace{1cm} (2)

According to the significance of \( T_k \), 1 or 0 will be sent to the decoder. Further, for every iteration \( T_k \) is divided into subsets. The process is repeated until the size of the subset is 1. The refinement pass sends the bits of the significant coefficients one by one from msb to lsb, for every iteration.

![Process chart showing the encoding and decoding methods of the proposed method](image)
D. Run Length Encoding

The value of pixels along a row of a digital image is given by a sequence of integers \( i_1, i_2, i_3 \) and so on. Then the gray level along this row may be represented by a set of order pairs \((i_n, n)\), where \(n\) is the number of consecutive pixels that have same value and \(i_n\) is the value of that pixel [7]. Hence the name Run Length Encoding is given to this method. By this method the required storage space can be reduced, if large run-lengths of pixels are available.

III. PROPOSED METHOD

Usually facial images will have a uniform background. The background is treated as non-ROI and the remaining part is treated as ROI. A binary mask is created which shows the background area by 0s and the remaining by 1s. The image is transformed by CDF9/7 wavelet transform into a matrix of wavelet coefficients. The wavelet coefficient matrix is compared with the binary mask. The elements of wavelet matrix are replaced with 0s, if the equivalent elements in the binary mask are 0 values. Then the wavelet matrix is quantized with a value \( q \) to reduce the storage space. The resultant wavelet matrix is encoded with SPIHT coding. This bit stream is further encoded with run length coding. The process chart of the proposed method is shown in Fig.2.

A. Pre-Processing

Facial images of size \( N \times N \) with uniform background are used for testing. The test images are pre-processed to have the facial features in the upper middle area. Facial image I is taken for processing.

B. ROI Mask Generation

The image area without the background is to be selected as the ROI area. The edges of the image are found by canny edge detector algorithm as shown in Fig.3. The area between the edges is filled and a binary mask is obtained. The ROI area is filled with 1s and the background is filled with 0s as shown Fig.4. The outer edges of ROI are expanded with few pixels to avoid ringing effects. The resultant binary mask is shown in Fig.5. The binary matrix is divided into blocks of \( 2 \times 2 \) pixels and each pixel in the block denotes a sub band as shown in the Fig.6.

The image is decomposed by CDF 9/7 wavelet transform for 5 levels. The resultant wavelet coefficient matrix \( M \) is rounded to integers. The binary mask area is the object to be compressed. The wavelet coefficient matrix \( M \) is compared with pixel by pixel with the binary matrix \( B \). The wavelet coefficients of \( M \) are replaced with 0s if the values of the equivalent pixels in the binary mask \( B \) are 0s.

\[
M(i,j) = 0, \quad \text{If } B(i,j) = 0 \tag{3}
\]

No change in \( M(i,j) \), \quad \text{If } B(i,j) = 1 \tag{4}

where \( M(i,j) \) is an element of wavelet coefficient matrix and \( B(i,j) \) is an element of binary mask matrix \( B \). When the above
process is over, more than 75% of the elements of matrix M are only 0s. The non-zero wavelet coefficients in the matrix M represents the object of the facial image to be compressed.

**D. Coding of Wavelet Coefficients**

The wavelet coefficient matrix M is encoded with SPIHT coding. Even though the whole matrix M is subjected to SPIHT coding, only the object part is encoded. Because SPIHT coding orders the wavelet coefficients from higher magnitude to lower magnitude. More number of 0s available in the matrix M makes the encoding easier. B is the encoded bit stream.

**E. Run Length Encoding**

Since long run lengths of 0s are available in bit stream B, run length encoding method is adopted to reduce the storage space further. RLB is the final bit stream.

**F. Decoding**

1. The resultant bit stream RLB is decoded with run length decoding to bit stream B’.
2. The bit stream B’ is decoded with SPIHT decoding to wavelet coefficient matrix M’.
3. The resultant matrix M’ is applied with inverse CDF9/7 wavelet transform to get the reconstructed image.

**IV. RESULTS AND DISCUSSIONS**

Facial images of size 256 x 256 pixels shown in Fig.9 are used to carry out the experiments. PSNR values at different bit rates were computed. For comparison standard SPIHT (without arithmetic coding) and JPEG2000 methods were used. Experiments were done at different bit rates 1.0, 0.08, 0.06 and 0.04. The PSNR values were computed and are given in Table 1. Fig.10 shows the results of the test image, Image2 at different bit rates by the three methods JPEG2000, SPIHT and Object Based SPIHT (OBSPIHT).

The proposed method was implemented in with Matlab 10. Jasper Software (Version 1.700.0) [8] was used to get the results of Jpeg2000. The proposed method gives better results than the JPEG2000 method. The variation in PSNR of the proposed method is slightly higher than the standard SPIHT method.

The SPIHT encoding method sorts the wavelet coefficient according to their magnitudes from higher to lower before encoding. The wavelet coefficients representing the background area are made as 0 in the proposed method. So, they are not encoded at low bit rates and the available bits can be used for coding the facial image area. This is the main advantage of the proposed method. In the method proposed by Monica Peneda et. al., the binary mask should be send along with the compressed data bits. This creates an additional overhead. There is no need to send the mask in the proposed method.

**TABLE 1**

<table>
<thead>
<tr>
<th>BPP</th>
<th>Image 1</th>
<th>Image 2</th>
<th>Image 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JPEG2K</td>
<td>SPIHT</td>
<td>OBSPIHT</td>
</tr>
<tr>
<td>0.10</td>
<td>29.99</td>
<td>32.00</td>
<td>33.07</td>
</tr>
<tr>
<td>0.08</td>
<td>29.07</td>
<td>31.32</td>
<td>32.34</td>
</tr>
<tr>
<td>0.06</td>
<td>27.36</td>
<td>30.07</td>
<td>30.93</td>
</tr>
<tr>
<td>0.04</td>
<td>24.20</td>
<td>28.75</td>
<td>29.83</td>
</tr>
</tbody>
</table>
V. CONCLUSION

In this paper we have proposed a method based on DWT and SPIHT to compress the facial images at low bit rates. Experimental results show that the proposed method gives better results in terms of PSNR than JPEG2000. Unlike the existing method there is no overhead data to be sent with the compressed data. This method can be used where fast access of facial image is needed.

REFERENCES


