A novel image coding scheme using wavelet transform is presented. The developed coding scheme prevents undesirable classifications of wavelet coefficients which arise from conventional quadtree based image coding. This is accomplished by pre-processing the wavelet coefficients. This pre-processing is based statistical distributions of the wavelet coefficients in subbands and the human visual system. With these considerations, the wavelet coefficients are weighted with different values which are derived to maximize the performance of the developed coder according to their visual importance. This leads to a more compact symbol classification for a quadtree based image coding scheme. In experiments with test images, we demonstrate that the developed coding scheme produces better image quality.

I. INTRODUCTION

A fundamental goal of image processing techniques is to reduce the bit rate for transmission or storage while maintaining an acceptable fidelity or image quality [1]. This goal can be achieved by removing redundancy in image. Traditional image compression algorithms have been developed to exploit the statistical redundancy present within real world images. The discrete cosine transform (DCT), DPCM, and the entropy coding of subband images are all examples of this statistical approach [2]. However, removing redundancy can only achieve a limited amount of compression. So, for a high ratio compression, some of the non-redundant information must be removed. The statistical coders in this mode generate some bad effect such as blocking effect in DCT, resulting in visual degradation by producing errors in visually important parts of the image structure such as edges. By using multiresolution analysis (MRA) that closely mimics human visual system (HVS), compression can be accomplished in high ratios, taking the importance of each individual coefficient into account. So, wavelet transform which has multiresolution characteristics and shows good localization in both time and frequency domain is being widely researched for image coding [3].

Embedded tree coders, having their root in the self-similarity between wavelet coefficients of adjacent scales, are among the most successful wavelet-based image compression techniques. The self-similarity means that for a given image location, its wavelet coefficients in different subbands have certain predictable relationships. In the embedded zerotree wavelet (EZW) coder proposed by Shapiro [4], the self-similarity is defined as this, "If a wavelet coefficient x at a coarse scale is insignificant with respect to a given threshold T, i.e., |x| < T, then all wavelet coefficients of the same spatial orientation at finer scales are likely to be insignificant." The EZW used the quadtree to code the self-similarity of wavelet coefficients across subbands. Quadtree is one of the most widely adopted data structures for representation of graphical and imagery data, where a parent node together with its four children represent an image and its four equally divided quadrants [5].

Quadtree based image coding algorithm have been developed in past decade [6, 7]. In conventional schemes, however, quadtree is constructed only from the magnitude and sign of wavelet coefficients. In this paper, we add two concepts in quadtree. One is the statistics of wavelet coefficients in different subbands and the other is the human visual system (HVS). So, we construct quadtree data structure based on these concepts. By integrating the successive approximation quantization (SAQ) and the newly constructed quadtree, we develop a new image coding scheme.

The paper is organized as follows. Section II describes the basic idea related to the developed algorithm. Next, the newly developed coding scheme is presented in section III. Experimental results are given in section IV. Finally, section V draws conclusions.

II. BASIC IDEA

A. Conventional Coding Scheme and Its Problems

Conventional image coding schemes based on a multiresolution representation construct significant maps that are binary maps indicating whether a coefficient of a 2-D discrete wavelet transform (DWT) has a zero or nonzero quantization value. For example, the EZW algorithm constructs the significant maps by classifying wavelet coefficients into four symbols: POS (positive significant symbol), NEG (negative significant symbol), IZ (isolated zero), and ZTR (zerotree root). This classification of symbols, however, depends only on the magnitude and the sign of wavelet coefficients. Since edges or boundaries in images which carry more important information for the HVS are not decomposed into coefficients of high magnitude by the DWT, the conventional significant maps undesirably classify the wavelet coefficients corresponding to these regions into an insignificant class due to their low magnitude. Also, in scanning the wavelet coefficients, the scan begins at the lowest frequency subband, denoted as LLN, and scans subbands HLN, LHN, and HHN, at which point it moves on to scale N-1, etc. The scanning pattern for a 3-scale QMF-pyramid can be seen in Fig. 1. This scanning order assumes that the lower frequency subband is more important than the higher one. But, this assumption doesn't consider a statistical distribution of the wavelet coefficients in subbands. Above two facts result in the loss of useful information, especially edges which are usually decomposed into coefficients of low magnitude on the high frequency
subbands. Also, these facts mean that the conventional image coders do not fully reflect good localization of the wavelet transform in both time and frequency domain.

**B. Successive Approximation Quantization (SAQ) with Weighted Wavelet Coefficients**

In this paper, to avoid undesirable classifications and prevent a loss of useful information, wavelet coefficients are pre-processed by giving different weights to the wavelet coefficients. This weighting process is performed on two considerations. One is subband-dependent weighting and the other is region-dependent weighting.

Subband-dependent weighting considers a statistical distribution of the wavelet coefficients in each subband. Fig. 2 shows plots of the point statistics (histogram) for different subbands of a typical wavelet decomposition. These statistics are reasonably well fit [8, 9] by distributions of the form:

\[ P_X(x) \propto e^{-|x|^v} \]  
(1)

The distribution is zero-mean and symmetric, and the parameters \( \{u, v\} \) are directly related to the second and fourth moments. By modeling the statistics in this simplistic fashion, we assume both stationarity and independence of the wavelet coefficients. Nevertheless, the model provides a reasonable fit to the statistics of many natural images [10]. From this modeling, we can approximate the probability of wavelet coefficients. Let \( X_M \) be a wavelet coefficient in M subband and \( P_X^M \) be a probability density function modeled in (1) in M subband. Then, we conclude that if \( |X_{M-1}| > |X_M| \),

\[ P_X^{M-1}(X > |X_{M-1}|) < P_X^M(X > |X_M|) \]  
(2)

implies that the wavelet coefficients in M-1 subband is less probable to have the same magnitude as those in M subband. So, it is reasonable to give a priority to coefficient, \( X_{M-1} \) which less occurs. This means that it is needed to consider this statistics in quantizing the wavelet coefficients. In conventional SAQ, however, \( X_M \) and \( X_{M-1} \) are quantized to the same value and reconstructed. To apply subband statistics to SAQ, we weight the wavelet coefficients in different subband with different weighting value before scanning the wavelet coefficients and then use a conventional SAQ. The magnitudes of the weighting values are

\[ W_M > W_{M-1} > W_{M-2} > \ldots > W_1 \]  
(3)

where \( W_M \) is a weighting value for M subband.

Region-dependent weighting is considered to prevent a loss of useful information such as edges and boundary. This loss arises from the undesirable classifications of the wavelet coefficients. To avoid this loss of information we extract visually important regions and then give a priority to those region. These are performed through two steps. One is the visually important region extraction (VRE) step and the other is to give a priority to the extracted region by means of weighting the corresponding wavelet coefficients. Because the VRE step performs on the wavelet transform domain, useful features of the DWT are fully reflected and we can process this step without the increase of the processing time through parallel processing to the DWT. Then, referring to the result of the VRE step, the wavelet coefficients are weighted. By weighting operation, more compact classification of the wavelet coefficients is realized because we change significant coefficients corresponding to the visually less important regions into insignificant ones. In the EZW, the insignificant coefficients are classified into zerotrees. So, growth of zerotrees translates into a corresponding gain in compression efficiency. In other words, the significant map is constructed with fewer number of symbols. By the help of region-based weighting, the zerotrees are grown for the visually less important regions and the symbol classification based on the human visual system (HVS) is realized.

**C. Effects of Weighting the Wavelet Coefficients**

In this part, we investigate effects of weighting operation on the wavelet coefficients. This investigation assume that the symbol classification of the EZW and the SAQ are applied. Firstly, the weighting operation leads to growth of zerotrees as explained before. This growth of zerotrees implies the increase of prediction on the wavelet coefficients and the prediction is translated into bit-saving for the representation of the symbols. This bit saving in \( k_{th} \) stage of the SAQ is expressed as

\[ \Lambda_k(w_m) = b_k N_k^F(F_k^w(T_k) - F_k(T_k)) \]  
(4)

where \( b_k \) is the number of bits for representation of each symbol, \( N_k^F \) means the number of significant symbol in \( k_{th} \) stage, \( T_k \) is threshold value for the symbol classification in \( k_{th} \) stage, \( F_k^w(\cdot) \) is cumulative density function of the wavelet coefficients and \( F_k^w(\cdot) \) is cumulative density function of the weighted wavelet coefficients. Although the bit-saving results in an efficient image coding, the side-effect from the weighting operation is a increase of a quantization error. Let the wavelet coefficient, \( x \) be between \( T_k \) to \( T_k/2 \) and \( \hat{x} \) be a reconstructed value for \( x \). Then quantization error, \( E_k \) in the SAQ and the EZW becomes

\[ E_k = |x - \hat{x}| = \sum_{\eta} \eta T_k \left( \frac{1}{2} \right)^{n+1} S_k \]  
(5)

where \( S_k \) is the number of the significant coefficients. With the weighting operation, the wavelet coefficients, \( x \) changes to \( w_m \times x = x_k \) and is reconstructed to \( \hat{x}_m / w_m \). So, the quantization error, \( E_k^w \) changes to
To determine the optimal weighting value, we define the benefit function, where

\[ E_k(w_m) = \left| \hat{x} - x^i / w_m \right| = \frac{1}{w_m} \left| w_m \hat{x} - \hat{x} \right| = \frac{1}{w_m} \sum w_m T_k \left( \frac{1}{2} \right) S_k \]  

(6)

So, the upper bound of the quantization error increases by the weighting value, \( w_m \) as follows:

\[ \sup \left| \hat{x} - x^i / w_m \right| = \frac{1}{w_m} \sup \left| \hat{x} - \hat{x} \right| \]  

(7)

As can be seen from above investigation, the performance of the weighting operation depends on the weighting value, \( w_m \). To determine the optimal weighting value, we define the benefit function, \( B(w_m) \) as

\[ B(w_m) = \sum \lambda_k (A_k(w_m) - E_k(w_m)) \]  

(8)

where \( \lambda_k \) is the average amount of a reconstructed wavelet coefficient per bit. We select \( w_m \) so that the benefit function is maximized.

### III. The Coding Algorithm

**A. Visually Important Region Extraction (VRE)**

Before weighting the wavelet coefficients, visually important regions are extracted. A visually important region is a region where intensity changes occur with respect to pixels and outstanding object is located[2]. This extraction step is realized through two steps. First, from low frequency characteristics overall image features are extracted. Then, high frequency components are weighted on those regions and highly weighted parts in those regions are selected as the visually important regions. After image is divided into two regions, a priority is placed on the visually important regions. In other words, the visually important regions are represented more accurately. This operation is accomplished by weighting the wavelet coefficients corresponding to the visually less important regions with lower weighting value. The VRE step is composed of four parts: successive region extraction (SRE), weighting region extraction (WRE), and generating weighting map (GWM).

Through multiresolution analysis, overall image feature can be extracted from low frequency subbands. In the SRE step, the regions which have low frequency characteristics are extracted and merged successively. The detail procedures are as follows.

i. Construct first stage decomposition using the DWT.

ii. Within subband \( LL_1 \), find maximum coefficient, \( X_{\text{max}} \). Set the initial region extraction threshold \( T_{\text{RE}} \) to \( X_{\text{max}} / 2 \).

iii. Select coefficients \( x_L \) such that \( \left| x_L \right| \geq T_{\text{RE}} \) through the \( LL_1 \) subband.

iv. From the coefficients selected in step (iii), first separate (+) coefficient regions and (-) coefficient regions. Second, merge regions if the same signed regions are adjacent.

v. If the newly generated region and pre-formed region are the same signed and adjacent, merge two regions.

vi. Eliminate regions of which size is less than 2x2.

vii. The wavelet coefficients which are included in the extracted regions are set to zero so that these coefficients do not prevent the occurrence of new region on future processes at smaller thresholds.

viii. Threshold \( T_{\text{RE}} \) is chosen so that \( T_{\text{RE}} = T_{\text{RE}} / 2 \) and go to the step (iii) if \( T_{\text{RE}} \geq T_{\text{FRE}} \) (final region extraction threshold).

Fig. 3(b) and 3(c) show that the result of the SRE applying to Lena image.

Next step is the WRE. In this step, among regions extracted from the SRE step, the only regions which contain high frequency characteristics are selected as the visually important region, where a priority is given. Details are as follows.

i. Within high frequency subbands labeled \( HL_1, LH_1 \), and \( HH_1 \), find the significant coefficients \( x_H \) such that \( \left| x_H \right| \geq T_{\text{WR}} \) (weighting region threshold).

ii. Determine the degree of high frequency weighting in each region which is formed in SRE step. This degree is represented in terms of the density. This density is defined as follows:

\[ \text{Density} \Delta = \frac{\text{NUM}_{HH}}{A} \]

where \( \text{NUM}_{HH} \) is the number of the significant coefficients \( x_H \) within each region and \( A \) is the area of each region.

iii. The regions whose \( \text{Density} \geq T_D \) (density threshold) are selected as the visually important regions.

As can be seen in Fig. 3(d), hair, parts of hat, eyes, and edges in “Lena” image are extracted as the visually important regions. All of them are the visually important regions.

To address the location of the weighted wavelet coefficients, we need to construct the map which indicates whether a wavelet coefficient has a priority or not. This map is formed through the GWM step. The visually important regions which are generated through the SRE and the WRE are mapped on the binary image of which size is \( N/2 \times N/2 \) assuming the original image size is \( N \times N \). To inform this map of decoder, this map is transmitted through the channel. To reduce side information, this map should be reduced into an acceptable size for transmission. The detail descriptions of the GWM step are as follows.

i. Construct \( N/2 \times N/2 \) binary map, \( M_1 \) from the visually important region. Set the array contents to integer one for the visually important regions and fill the other contents with integer zero. Set \( k \) to integer one.

ii. Form the new binary map, \( M_{k+1} \) whose size is reduced by half, \( N/2^{k+1} \times N/2^{k+1} \) from the previous map, \( M_k \). The map
content \( M_{k+1}(i, j) \) is determined as follows:

\[
M_{\text{value}} = M_k(2i, 2j) + M_k(2i, 2j+1) + M_k(2i+1, 2j) + M_k(2i+1, 2j+1)
\]

if \( M_{\text{value}} \geq 2 \)

then \( M_{k+1}(i, j) = 1 \); masking region

else \( M_{k+1}(i, j) = 0 \); unmasking region

iii. Increase \( k \) by one and go to step ii) if \( N/2 \geq \) Target Size.

As an example, the Target Size is chosen at 32. The resulting map is given in Fig. 3(e).

### B. Selecting Weighting Value and Weighting Wavelet Coefficients

As shown in section II, the weighting value, \( w_m \) is selected to maximize (8). To maximize

\[
B(w_m) = \sum_k \lambda_k b_n N_k^\lambda (F_n^w(T_k) - F_n(T_k)) - \frac{1}{k} \sum_k \sum_n T_k \left( \frac{1}{2} \right)^n S_k
\]

we should know the \( N_k^\lambda (F_n^w(T_k) - F_n(T_k)) \) in the bit-saving function. But, it is difficult, if not impossible, to get a statistical distribution of all the wavelet coefficients. So, in implementing the developed algorithm, we use approximated version of the bit-saving function. The statistical part, \( N_k^\lambda (F_n^w(T_k) - F_n(T_k)) \) means the number of symbols reduced by the weighting operation. Therefore, \( N_k^\lambda (F_n^w(T_k) - F_n(T_k)) \) can be rewritten as

\[
N_k^\lambda (F_n^w(T_k) - F_n(T_k)) = N_k^\lambda \left| w_m \right| - N_k^\lambda \left| w_m \right|_{m=1} = (S_k \left| w_m \right| - S_k \left| w_m \right|_{m=1}) - (ZT_k \left| w_m \right| - ZT_k \left| w_m \right|_{m=1})
\]

where \( N_k^\lambda \left| w_m \right| \) is the number of symbols when the wavelet coefficients are weighted with \( w_m \), \( S_k \left| w_m \right| \) is the number of the significant symbols when the wavelet coefficients are weighted with \( w_m \), and \( ZT_k \left| w_m \right| \) is the number of zerotrees when the wavelet coefficients are weighted with \( w_m \). In approximating the (10), we assume that the wavelet coefficients are uniformly distributed at least between adjacent threshold values for the symbol classification. With this assumption,

\[
\Delta S_k \left| w_m \right| = S_k \left| w_m \right| - S_k \left| w_m \right|_{m=1} = \left( \frac{2w_m - 1}{w_m} \right) S_k \left| w_m \right| + \left( 1 - \frac{w_m}{w_m} \right) S_k \left| w_m \right|_{m=1}
\]

and

\[
\Delta ZT_k \left| w_m \right| = ZT_k \left| w_m \right| - ZT_k \left| w_m \right|_{m=1} = \eta_k \left( \frac{1 - w_m}{w_m} \right) S_k \left| w_m \right|_{m=1}
\]

where \( \eta_k \) is the number of zerotrees generated by one zero tree root symbol. So, (9) is approximated to

\[
B(w_m) = \sum_k \lambda_k b_n \Delta N_k^\lambda - \frac{1}{k} \sum_k \sum_n T_k \left( \frac{1}{2} \right)^n S_k
\]

where \( \Delta N_k^\lambda \) is equal to

\[
\Delta N_k^\lambda = \Delta S_k \left| w_m \right| - \Delta ZT_k \left| w_m \right| = \left( \frac{1 - w_m}{w_m} \right) \left[ (1 + \eta_k) S_k \left| w_m \right|_{m=1} - S_k \left| w_m \right|_{m=1} \right]
\]

where \( S_k \) is equal to \( S_k \left| w_m \right|_{m=1} \). With approximated version (13), we can obtain \( w_m \) so that the benefit function is maximized.

Referring to the weighting map from the GWM step, the weighting operation is performed on the wavelet coefficients.

i. Start from subbands \( HL_m \) of which size is the same as the weighting map’s size. While referring to the masking map, give a priority to the coefficients which correspond to integer one in the map. This action is accomplished by multiplying the wavelet coefficients, which correspond to integer zero in the map, with \( 1/w_m \). Overall coding structure is shown in Fig. 5.

### C. Threshold Determination

In the developed coding algorithm, four variables should be specified for an efficient working. These variables are \( T_{\text{BRE}} \), \( T_{\text{WR}} \).
Among these, $T_{RE}$ and $T_{WR}$ are relatively less important. The recommended restriction is to prevent $T_{WR}$ from being selected as low value because with low $T_{WR}$, most of the high frequency component would be used for the weighting operation. Also, because the wavelet coefficients in $L_{LM}$ band for the M stage decomposition are generally larger than those in the other bands, $T_{WR}$ needs to be determined as $T_{RE}/2$. Considering these facts, fixing $T_{RE}$ and $T_{WR}$ with $X_{max}/16$ and $X_{max}/32$, respectively, will not affect coding efficiency significantly. When generating the weighting map, the Target Size is considered. The closer the Target Size is to the image size, the more the visually important region can be specified exactly. The side effect, however, is that many bits should be assigned to the weighting map due to its large size. A simple experiment shows that the weighting map of which size is 32x32 can be encoded with about 20 to 30 bytes using the arithmetic coding. This is an acceptable cost even for a low bit rate coding and 64x64 masking map is recommended for applications which need low compression ratios. The remaining threshold variable is $T_{D}$. More attention should be paid to selecting this value. Experimentally, it is observed that efficient coding is accomplished with $T_{D}$ which is selected as slightly less than the mean of density defined in section III-A.

![Original Test Image](image.png)

Fig. 6. Original Test Image (a) “Lena”, 512x512 (b) “Closup”, 512x512

IV. EXPERIMENTAL RESULTS

All experiments were carried out by encoding and decoding an actual bit stream to verify the correctness of the algorithm. The entire bit stream is arithmetically encoded using a single arithmetic coder with an adaptive model [11]. The numerical evaluation of the developed coder’s performance is achieved by computing the peak signal-to-noise ratio (PSNR) of the original image and the coded image. The developed coder is applied to the standard black and white 8 bpp. test images, 512x512 “Lena” and the 512x512 “Closup” which are shown in Fig. 6(a) and 6(b). The coding parameters of $T_{RE}$, $T_{WR}$, and the Target Size are chosen at $X_{max}/16$, $X_{max}/32$, and 32, respectively. Six scales of QMF-pyramid are used.

Our experimental are focused on comparison with the standard EZW coder because Shapiro has shown that the EZW scheme excels other compression algorithms. For the test images, the image quality for the visually important regions such as edges, overall face of “Lena”, eyes, and nose is highly improved as can be seen from Fig. 7 to Fig. 10. The subjective quality of decoded image is evaluated by ordinary viewers. All the viewers make an agreement that the proposed coding algorithm highly improves the image quality and gives visually pleasing results for most images compared with conventional EZW. Even the PSNR of the proposed coding algorithm is improves 1dB to 3dB for the visually important region and by 0.5dB to 1dB for the whole image compared with the EZW. The whole results are presented in Table 1 and Table 2.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding Results for 512x512 “Lena” showing PSNR(dB) and Its Improvement Compared with EZW</td>
</tr>
<tr>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>Bytes</td>
</tr>
<tr>
<td>-----------------------------------------</td>
</tr>
<tr>
<td>16384</td>
</tr>
</tbody>
</table>
TABLE 2  
Coding Results for 512x512 “Closup” showing PSNR(dB) and Its Improvement Compared with Pure EZW

<table>
<thead>
<tr>
<th>Bytes</th>
<th>Ratio</th>
<th>Regional PSNR</th>
<th>Total PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>16384</td>
<td>16:1</td>
<td>33.961(3.219)</td>
<td>36.020(0.904)</td>
</tr>
<tr>
<td>8192</td>
<td>32:1</td>
<td>29.362(2.654)</td>
<td>32.083(0.639)</td>
</tr>
<tr>
<td>4096</td>
<td>64:1</td>
<td>25.675(1.433)</td>
<td>28.821(1.262)</td>
</tr>
</tbody>
</table>

Regional PSNR : PSNR for the visually important regions

( ) : Improvement when comparing with EZW

V. CONCLUSION

This paper is focused on pre-processing the wavelet coefficients from the DWT to prevent undesirable classifications in quadtree based image coding schemes. This pre-processing scheme is realized by giving different weighting to the wavelet coefficients. This weighting process is performed on two considerations. One is subband-dependent weighting which exploits subbands characteristics. The other is region-dependent weighting which takes the HVS into account. As shown in experimental results, this pre-processing scheme works very well. Although this result is obtained by integrating this scheme with the EZW and the SAQ, the concept of the developed algorithm can be applied to other coding scheme and quantization methods because the developed algorithm is modular form of pre-processing and is processed parallel to the DWT.

REFERENCES