Module
5
EMBEDDED WAVELET CODING

Lesson
13
Zerotree Approach.

Instructional Objectives

At the end of this lesson, the students should be able to:

1. Explain the principle of embedded coding.
2. Show the parent-child relationships between subbands of same orientation.
3. Define significance and insignificance of DWT coefficients with respect to a threshold.
4. Define zerotree and zerotree root.
5. Perform successive approximation quantization (SAQ) on DWT coefficients.
6. Perform dominant pass and subordinate pass for coding of DWT coefficients.
7. Implement a complete Embedded Zerotree Wavelet (EZW) encoder and study its performance on images.

13.0 Introduction

In lesson-12, we have studied Discrete Wavelet Transforms (DWT) and seen its applicability in images. We noted that DWT has excellent energy compaction capabilities and hence the coding technique must be well-designed to achieve significant image compression. In this lesson, we are going to study an embedded wavelet coding technique, known as Embedded Zerotree Wavelet (EZW) coding that effectively exploits the self-similarity between subbands and the fact that the high-frequency subbands mostly contain insignificant coefficients. First, we define the relationship between the subbands, based on the spatial locations and then define a data structure in the form of a hierarchical tree that includes spatially related coefficients across different subbands. The tree defines a parent-child relationship of DWT coefficients across subbands. The concept of a zerotree is introduced which identifies the parts of a tree that have all the DWT coefficients insignificant starting with a root. Since, DWT coefficients are generally insignificant at higher frequency subbands, occurrences of zerotrees are expected to be frequent and the zerotree roots can be encoded with a special symbol. The EZW algorithm is based on successive approximation quantization and this facilitates the embedding algorithm. Based on the concepts we are going to present in this lesson, the students should be able to design a complete wavelet coder, which can be suited to the desired bit-rate of the channel.

13.1 Embedded Coding

In embedded coding, the coded bits are ordered in accordance with their importance and all lower rate codes are provided at the beginning of the bitstream. Using an embedded code, the encoder can terminate the encoding process at any stage, so as to exactly satisfy the target bit-rate specified by the channel. To achieve this, the encoder can maintain a bit count and truncate the bit-stream, whenever the target bit rate is achieved. Although the embedded coding used in EZW is more general and sophisticated than the simple bit-plane coding, in spirit, it can be compared with the latter, where the encoding commences with the most significant bit plane and progressively continues with the next most significant bit-plane and so on. If target bit-rate is achieved before the less significant bit planes are added to the bit-stream, there will be reconstruction error at the receiver, but the “significance ordering” of the embedded bit stream helps in reducing the reconstruction error at the given target bit rate.

13.2 Relationship between subbands

In a hierarchical subband system, which we have already discussed in the previous lessons, every coefficient at a given scale can be related to a set of coefficients at the next finer scale of similar orientation. Only, the highest frequency subbands are exceptions, since there is no existence of finer scale beyond these. The coefficient at the coarser scale is called the parent and the coefficients at the next finer scale in similar orientation and same spatial location are the children. For a given parent, the set of all coefficients at all finer scales in similar orientation and spatial locations are called descendants. Similarly, for a given child, the set of coefficients at all coarser scales of similar orientation and same spatial location are called ancestors.

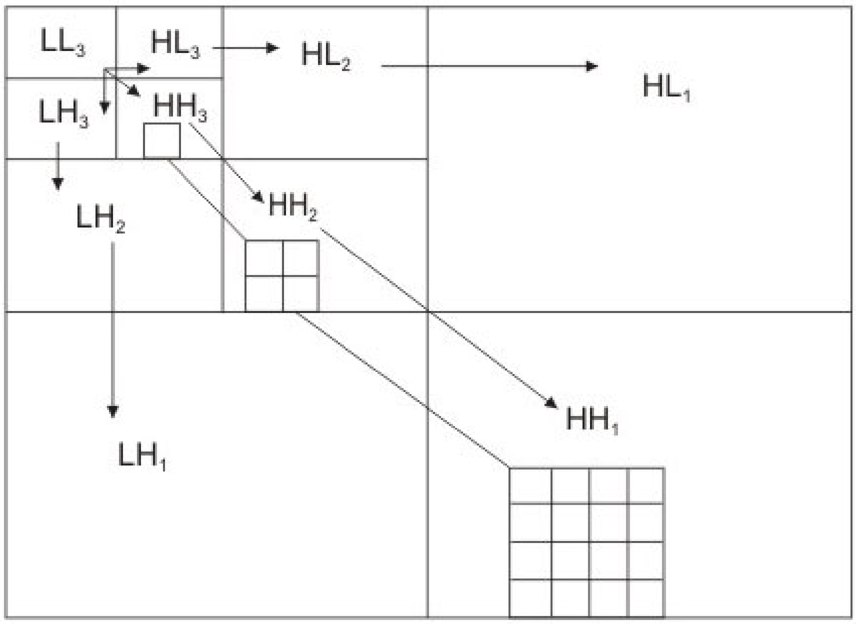

Fig 13.1: *Parent-child dependencies of sub bands*

Fig.13.1 illustrates this concept, showing the descendants of a DWT coefficient existing in HH3 subband. Note that the coefficient under consideration has four children in HH2 subband, since HH2 subband has four times resolution as that of HH3. Likewise, the coefficient under consideration in HH3 subband has sixteen descendants in subband HH1, which in this case is a highest-resolution subband. For a coefficient in the LL subband, that exists only at the coarsest scale (in this case, the LL3), the hierarchical concept is slightly different. There, a coefficient in LL3 has three children – one in HL3, one in LH3 and one in HH3, all at the same spatial location. Thus, every coefficient at any subband other than LL3 must have its ultimate ancestor residing in the LL3 subband.

The relationship defined above best depicts the concept of space-frequency localization of wavelet transforms. If we form a descendant tree, starting with a coefficient in LL3 as a root node, the tree would span all coefficients at all higher frequency subbands at the same spatial location.

13.3 Significance of DWT coefficients

Before we can exploit the hierarchical subband relationship concept for efficient encoding of DWT coefficient, it is necessary to introduce a very simple concept of significance. We say that a DWT coefficient of magnitude $|X|$ is *significant* with respect to a given threshold *T* if $|X| > T$ and is *insignificant* otherwise. In the embedded coding adopted in EZW, the significance of DWT coefficients are first examined with the highest value of threshold in the first pass and then progressively, the threshold is decreased by a factor of 2 in subsequent passes. Before we start, all coefficients are assumed to be insignificant and progressively, more and more significant coefficients will be detected and by the end of the final pass, all coefficients would assume significance at some pass. At each pass, there is a significance map that tells about the significance of the DWT coefficients and this map requires to be encoded efficiently. The significance map has an entry of zero if the coefficient is insignificant with respect to a threshold and is one if significant. It should be noted that the significance is decided only with respect to the magnitude and hence the sign of the significance (positive or negative) must be included in the encoding process.

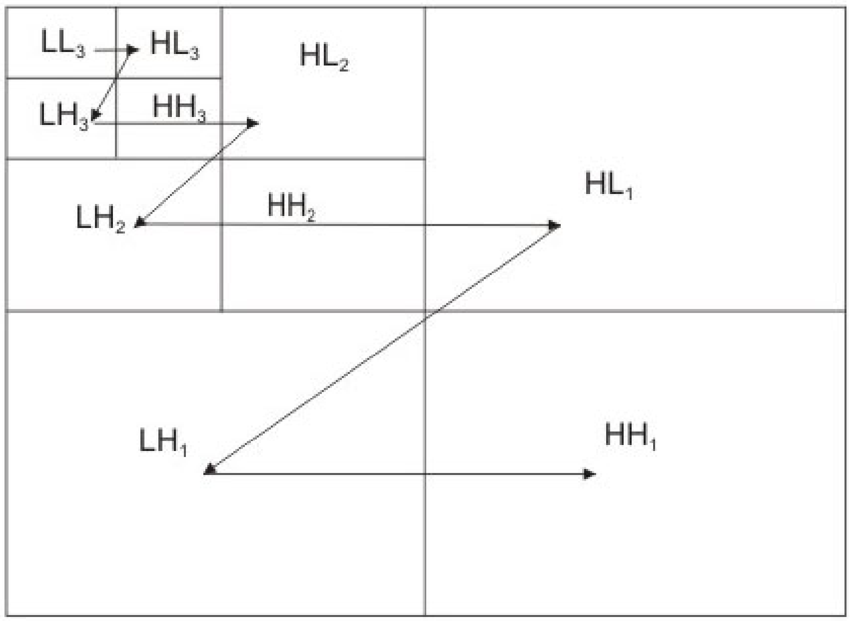

Fig 13.2 *Scanning order of sub bands*

The coefficients are scanned for significance in a manner illustrated in fig.13.2 for a 3-level subband decomposition. It starts with the lowest frequency subband, designated as LL_N where N s the number of levels. Following the scanning of all the coefficients in this subband, all the coefficients in subband HL_N are scanned. This is followed by LH_N and HH_N. Then the scanning proceeds to the next finer level N-1 in the same order HL, LH and HH. It continues till the highest frequency subbands are covered. This ensures that no child node is scanned before its parent.

13.4 Encoding the Significance map

We are now going to examine how to efficiently encode the significance map at any pass. For this, the hierarchical relationship of coefficients presented in Section-13.2 is utilized. A data-structure, called zerotree is defined as a tree-like data structure that includes an insignificant coefficient into it, provided all the descendants of that coefficient are also insignificant. A zerotree must therefore have a root, which itself is insignificant, but its parent is significant at that threshold. If all the ancestors till the coarsest frequency LL subband form the zerotree, then the ancestor at LL subband is declared as the zerotree root. The zerotree concept is based on the hypothesis that if a DWT coefficient at a coarse scale is insignificant with respect to a given threshold, then all its higher frequency descendants are likely to be insignificant with respect to the same threshold. Although, this may not be always true, but these are generally true.

It may however be noted that all insignificant coefficients may not be a part of zerotree. It is possible that a coefficient is insignificant, but has some significant descendants. These coefficients are called *isolated zero.* Four symbols are used to encode the significance map, namely

- Zerotree root (ZTR).
- Positive significance (PS).
- Negative significance (NS).
- Isolated Zero (IZ).

The encoding of the coefficients into one of the above four symbols is illustrated in fig.13.3. We shall explain this with an example, later on in this lesson.

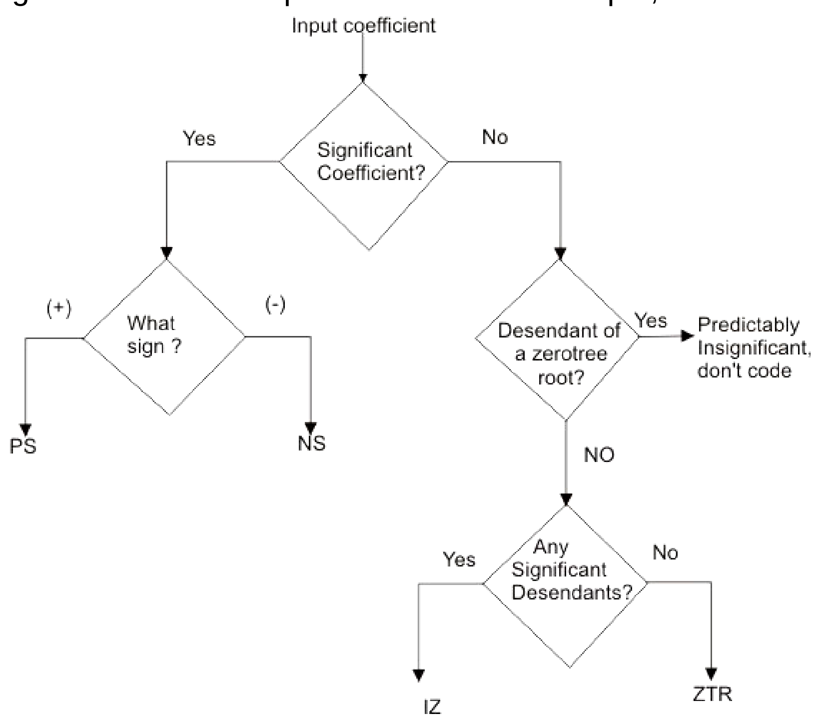

Fig 13.3 *Flowchart for encoding significance map*

Zerotree coding reduces the cost of encoding the significance map using self-similarity. Even though DWT essentially decorrelates the coefficients, occurrences of insignificant coefficients are not independent events. It is easier to predict insignificance, rather than predicting significant details across the scales and zerotree coding exploits the redundancies that the insignificant coefficients offer.

13.5 Successive Approximation Quantization (SAQ)

Successive Approximation Quantization (SAQ) performs encoding of magnitudes of DWT coefficients in successive stages. An initial threshold T_0 to examine the significance is first set up such that $T_0 > |X_{\max}|/2$, where $X_{\max}$ is the maximum of all DWT coefficients. In each stage of encoding, it reduces the threshold by half and examines the significance once more. The sequence of thresholds that get applied in successive stages are $T_0, T_1, T_2, \cdots\cdots, T_{N-1}$ where *N* is the number of passes and $T_i = T_{i-1}/2$ for $i = 1, 2, \cdots\cdots, N-1$. Each stage consists of two passes – a *dominant pas*s and a *subordinate pass*.

A dominant pass is used to encode those coefficients that have not yet (that is, till the previous stage of encoding) been found to be significant with respect to a threshold T_i. The significant coefficients identified during this pass in the same scanning order, as illustrated earlier in fig.13.2 are encoded in zerotree structures, discussed in Section-13.4 and their magnitudes are appended to a list, known as subordinate list. At the same time, the coefficient in the DWT array is set to zero such that during the next dominant passes at lower thresholds, the coefficient is treated as insignificant and can be included as a part of zerotree.

A dominant pass is followed by a subordinate pass in which the coefficients found to be significant in the subordinate list are scanned and their magnitudes are refined with an added bit of precision, splitting the uncertainty region of encoding into two halves. For each magnitude in the subordinate list, this refinement can be encoded using a binary symbol, “0” if it falls in the lower half of the uncertainty region and “1” if it is in the other half. The string of symbols generated from during the subordinate pass is entropy coded. After the completion of a subordinate pass, the magnitudes on the subordinate list are sorted in decreasing amplitude, to the extent that the decoder also should be able to carry out the same sorting. The encoding process alternates between dominant pass and subordinate pass and the threshold is halved after each dominant pass. The encoding stops when some target bit rate is achieved. The ability to truncate the encoding or decoding anywhere is extremely useful in systems that are rate-constrained or distortion-constrained.

13.6 An encoding example

127	69	24	73	13	5	-8	5
-37	-18	-18	8	-6	7	15	4
44	-87	-15	21	8	-11	14	-3
55	18	29	-56	0	-2	3	7
34	38	-18	17	3	-9	-2	1
-27	-41	11	-5	0	-1	0	-3
6	17	5	-19	2	0	-3	1
32	26	-7	5	-1	-5	7	4

Fig 13.4 : Example DWT coefficient array for 3-levels on an 8x8 images.

The basic principles of EZW coding described so far can be best understood by considering an example array of DWT coefficients, as shown in fig.13.4. The example shows a 3-level DWT coefficient array of an 8 x 8 image, split into 10 subbands. It may be observed that the magnitude of the highest DWT coefficient is 127. The initial threshold may be set anywhere in the range (63.5, 127]. We set the initial threshold T_0 as 64. Before we begin the first dominant pass, all the coefficients in this array were treated to be insignificant. With respect to the initial threshold, the dominant pass picks up the following significant coefficients in the scanning order illustrated in fig.13.2:

- Coefficient value 127 in LL3. This will be encoded as “PS”, since the coefficient is of positive value. After decoding this signal, the decoder knows that the coefficient lies in the interval [64,128) and its reconstruction value is the centre of this interval, i.e, 96.

- Coefficient value 69 in HL3. This will also be encoded as “PS”. As before, its reconstruction value is also 96.
- Coefficient value 73 in HL2. This will also be encoded as “PS” with a reconstruction value of 96.
- Coefficient value -87 in LH2. This will be encoded as “NS”, since the coefficient is of negative value. The decoder knows that the magnitude of the coefficient lies in the interval [64,128) and its reconstruction value will be -96.

All remaining coefficients are insignificant in the first dominant pass. The first dominant pass scanning will identify the following zerotree root (coded as “ZTR”) and isolated zeros (coed as “IZ”):

- Coefficient value of -37 in LH3 is insignificant, but it has significant coefficient value of -87 in its descendants in LH2. Thus, this coefficient will be encoded as isolated zero (IZ).
- Coefficient value of -18 in HH3 is insignificant and all its descendants in HH2 and HH1 are insignificant. Thus, this coefficient qualifies to be a zerotree root (ZTR) and will be encoded accordingly.
- The reader may verify that the following coefficients are also zerotree root (ZTR):
    - 24, -18 and 8 in HL2.
    - 44, 65 and 18 in LH2.
- Also observe that the following coefficients are zeros, but not a part of any zerotree root:
    - -8, 5,15 and 4 in HL1.
    - -18,17,11 and -5 in LH1.

For these highest frequency subbands, ZTR and IZ may be merged into a common symbol of “Zero” (Z).

At the end of the first dominant pass, the subordinate list will contain only the four significant coefficients identified. The first subordinate pass will refine the magnitudes of the significant coefficients and categorize them into one of the two uncertainty intervals, viz., [64,96) and [96,128). Thus, only the LL3 coefficient of magnitude 127 will belong to the latter interval and will be encoded with symbol 1, whereas the remaining three significant coefficients will belong to the former interval and encoded with symbol 0. The first coefficient will have a reconstruction value of 112 and the remaining coefficients will have a reconstruction value of 80 at the middle of the uncertainty interval. In this case, the first coefficient only is encoded as 1 and the remaining as 0s and no re-ordering in subordinate list is necessary.

The first dominant and the first subordinate pass complete the first stage of processing. Now, the second dominant pass starts with threshold set to $T_1 = T_0 / 2 = 32$. During this pass, the coefficients which are yet to be found as significant will only be scanned. All the coefficients previously found to be significant are set to zero so that they could be included as a part of zerotree in this, as well as latter passes. However, the subordinate list is still maintained and the second and subsequent passes will only append the significant coefficients found in that pass to the subordinate list.

As an exercise, the student is advised to complete the second dominant and second subordinate pass encoding. The processing alternately continues between the dominant pass and the subordinate pass and can be stopped at any time.

13.7 Order of importance in the bit-stream

The embedded bit-stream, which the EZW algorithm generates inherently, performs an ordering of bit-stream according to the importance. The importance follows the order of precision, magnitude, scale and spatial location according to the initial dominant list.

The first importance is assigned to the numerical precision of the coefficients. All the coefficients in a pass are encoded with the same numerical precision and it is only after a dominant pass that the numerical precision is refined by a factor of two.

The next in importance is the magnitude. Prior to a pass, all coefficients are assumed to be insignificant and the dominant pass picks up all significant coefficients, having magnitudes greater than those of the insignificant coefficients. During the subordinate pass, the magnitudes are sorted in a descending order of the centres of uncertainty intervals.

Scale is the next factor of importance. It follows the ordering of subbands on the initial dominant list. The coarser scales are covered before the finer or the high-frequency coefficients.

The final factor is the spatial location. It simply means that two coefficients, which cannot be distinguished by precision, magnitude and scale, have their relative importance decided arbitrarily by the initial scanning ordering of the two coefficients within a subband.

13.8 Summary of the EZW algorithm

The EZW algorithm, first proposed by Shapiro is one of the pioneering works in the embedded coding of wavelet coefficients and showed very promising results when applied on natural images. The embedded nature of the algorithm allows the bitstream to be truncated at any point when the target bit-rate is achieved. To summarize, we can say that EZW is based on three major concepts, namely,

a) Partial ordering of the coefficients by magnitudes, since the algorithm sorts the significant coefficients according to magnitudes in every subordinate pass. The sorting is not exact, but the coefficients are arranged in decreasing order of the centres of uncertainty intervals in magnitudes.
b) Ordered bit-plane transmission of refinement bits, in accordance with the SAQ approach described in Section-13.5
c) Exploitation of the self-similarity of coefficients across subbands of same orientation across higher scales through the zerotree data structures and efficient coding of zerotree roots.

However, the algorithm is at best sub-optimal and is not optimal. It is possible that the truncation of bit stream is done in the middle of a pass and only a part of the coefficients might have been encoded with refined precision and the remaining can only be reconstructed with the precision followed till the previous pass, that is, one half of the precision of the present pass. Moreover, the choice of initial threshold is also very important and may affect the result. It is seen that adjustment of initial threshold may be needed to make the EZW optimal for a given bit-rate.

Subsequent to Shapiro's work on EZW, several modified approach towards embedded coding of wavelet coefficients have been implemented. We are going to study these in the next lesson.