EMBEDDED SLCCA FOR WAVELET IMAGE CODING

Junqiang Lan and Xinhua Zhuang

Multimedia Communication and Visualization Laboratory
Department of Computer Engineering & Computer Science
University of Missouri-Columbia, MO 65211

ABSTRACT

In 1997, we published a high performance scalable bit-rate-oriented wavelet image codec, Significance Linked Connected Component Analysis (SLCCA). In term of coding efficiency, SLCCA outperforms EZW, SPIHT and JPEG2000, which are each an embedded image codec. SLCCA is optimized for any given bit rate. Although scalable, its coded bit stream is not fully embedded. In the paper, we provide an embedded SLCCA, which improves our previous work[1] on embedded SLCCA. Like[1], we define a base layer and an enhancement layer for an image. The base layer encoded by SLCCA provides an acceptable image quality while achieving high coding efficiency and scalability. The enhancement layer provides more elaborate bit-rate adaptive granular picture quality enhancement for embedded image coding. The Peak Signal to Noise Ratio (PSNR) of the new embedded SLCCA (called ESLCCA) consistently outperforms JPEG2000 that can be as high as 0.5 dB.

1. INTRODUCTION

In recent years, we have seen several advanced wavelet based image codecs developed, such as EZW[2], SPIHT[3], MRWD[4], SLCCA[5], as well as the new image coding standard JPEG2000[6]. In these wavelet based image codecs, an image is decomposed into a hierarchical multi-resolution representation by Discrete Wavelet Transform (DWT). Due to the energy compaction property of wavelet transform, after quantization many wavelet coefficients in high frequency subbands become insignificant (zero). Thus, it would be more efficient to just organize and encode positions of those significant (non-zero) coefficients as well as their values. The difference between the above codecs lies in how the significance map and the coefficients are organized and encoded.

If both the significance map and the coefficients are organized and encoded progressively (for example, bit-plane-wisely), it is possible to produce an embedded bit stream, allowing the encoding to terminate at a point that meets precisely with a target bit rate or target distortion metric[2]. SLCCA represents a bit-rate-oriented coding technology. That is, for any pre-specified bit rate, the codec will produce a bit stream conforming to that rate through possibly a few iterations on quantization levels using the knowledge that the larger the quantization levels the lower the bit-rates in general. In bit-rate-oriented SLCCA, only the coefficients are organized and encoded progressively while the entire significance map is organized as significance-linked clusters and encoded once for ever. Although being non-embedded, the scalability provided by this bit-rate-oriented image coding technology works very well for a variety of both online and offline applications as far as a bit rate can be negotiated or specified in advance. For example, in time-critical wireless image transmission, the needed channel bandwidth can be negotiated, while for most of offline image storage applications the bit rate is fixed. However, there are certain time-critical progressive image transmission applications, where the bit rates cannot be specified or guaranteed in advance. Then, bit stream embedding is a useful and desired feature. For instance, when an estimated or specified bit rate, which is used to design SLCCA codec, is much lower than an actual bit rate, the scalability provided by SLCCA is less useful and the available resource becomes underutilized.

The paper is organized as follows. In Section 2, we give a brief review of SLCCA. In Section 3, the new embedded SLCCA (ESLCCA) algorithm is described in detail. The paper concludes with the performance evaluation in Section 4.

2. SLCCA

In SLCCA, significance map in each subband is consisted of irregular shaped clusters with significance linkage information embedded[5]. A cluster is constructed from a “seed”, followed by a recursive cluster expansion using a small (3x3, for instance) morphological structuring...
element for conditional dilation. In essence, a cluster takes a tree structure with all nodes being significant samples except leaves, which are insignificant samples (quantized zeros) or lie on subband border. The cluster or tree will be traversed during conditional dilation to form an initial string of three symbols POS/NEG/ZERO. When a significant sample points to a “seed” of a cluster in a child subband, a LINK symbol will be embedded into the initial string right after its significance symbol POS/NEG. SLCCA algorithm can be briefed as follows.

1. **Seed Encoding:** Scan a subband until a seed is found that is the first non-coded significant sample left in the subband. The seed position is encoded directly by its (x, y) coordinates or indirectly inferred by LINK from its significant parent sample.

2. **Cluster Expansion:** Apply morphological conditional dilation to expand the cluster until insignificant boundary zeroes or subband boundary is reached.

### 3. ESLCCA Algorithm

As seen, to make SLCCA an embedded encoding algorithm we need a scheme for progressive encoding of the significance map. Following[1], we divide an image into two layers, the base layer and the enhancement layer. The base layer provides an acceptable picture quality, which is scalable, while the enhancement layer provides a fully embedded bit stream. The scalable base layer together with the embedded enhancement layer will generate a virtually fully embedded bit stream. The whole encoding scheme is shown in Fig 1.

![Fig 1. Encoding Algorithm of Embedded SLCCA](image)

The reason to define a base layer and use SLCCA to code the base layer is twofold: (1) a compressed image is valuable only if it delivers an acceptable picture quality, which would translate into a minimum bit-budget based on a specific coding scheme; (2) SLCCA is optimized for any specified bit rate and has better coding performance in the low bit rate range [1].

In [1], the embedded encoding of the enhancement layer essentially follows JPEG2000, which encodes each bit plane in three passes. The difference between JPEG2000 and SLCCA is that in JPEG2000, run length encoding is utilized to efficiently encoding the significant positions; while SLCCA algorithm is trying to only encode all the significant coefficients as clusters.

In ESLCCA, the embedded encoding of the enhancement layer is based on bit-plane-wise connected component analysis or morphological clustering, similar to SLCCA. That is, ESLCCA represents the enhancement layer of each subband by bit-plane-wise morphological clusters. However, those clusters are S-coefficients-based, where a coefficient in the current bit plane is defined an S coefficient if it has already been significant in a previous bit plane or becomes significant in the current bit plane. Notice that an S coefficient can take either binary 1 or 0 in the current bit plane. Thus the morphological clusters in the current bit plane are only composed of S coefficients and their insignificant neighbors. For each subband in the current bit plane, the S-coefficients-based clustering algorithm finds all the S-coefficients and forms connected S-coefficients into clusters. The cluster formation including two steps:

1. **Find Cluster Seed:** Scan a subband until a seed is found that is an S coefficient but has not been clustered yet in this bit plane.

2. **Cluster Expansion:** Apply morphological conditional dilation to expand the S-coefficients-based cluster until non-S-coefficients boundary zeroes or subband boundary is reached.

The encoding of the S-coefficients based clusters includes the encoding of seed position and each element of the cluster:

1. **Encoding seed position.** The seed falls into three exclusive categories, each category is coded differently.
   a. If the seed has already been a significant coefficient in the previous bit planes, there is no need to code such a seed since the decoder will find the same seed in the scanning process.
   b. If the seed just becomes significant in the current bit plane and its “parent”[1] is significant, the seed position is then encoded by a “LINK” symbol.
   c. If the seed just becomes significant in the current bit plane but its “parent” is insignificant, then the position value (x;y) of the coefficient is encoded.

2. **Encoding the elements of the cluster.** For every entry [x; y] in the cluster,
   a. If c[x; y] is not an S-coefficient, which means it is an insignificant neighbor of an S pixel, an “0” significant bit is encoded.
   b. If c[x; y] is an S-coefficient, but has already been significant in a previous bit
plane, a refinement bit \( r_{[x; y]} \) is encoded.

c. If \( c_{[x; y]} \) just becomes significant in the current bit plane, then its significant bit “1” followed by its sign bit is encoded.

Besides this progressive coding of the significance map, there are some other necessary modifications to be made on the original SLCCA codec.

### 3.1. Quantization

Enhancement layer is composed of the residuals between the original wavelet coefficients and the reconstructed wavelet coefficients in base layer. The quantization method used in SLCCA is shown in (1).

\[
\begin{align*}
    \text{level} &= \left\lfloor \frac{\text{abs}(x) - T}{q} \right\rfloor \\
    r &= \begin{cases} 
        0, & \text{if } s = 0; \\
        s \times (\text{level} \times q + T + q/2), & \text{otherwise}\n    \end{cases}
\end{align*}
\]

(1)

Where \( x \) is the original value, \( r \) is the reconstructed value, \( q \) is the quantization step, \( T \) (equals to \( q \) here) is the threshold of dead zone, \( s \) is the sign and \( \text{level} \) is the quantized value of \( x \). The reconstructed value \( r \) of this uniform quantizer lies in the middle of its two adjacent decision values, as shown in the upper half of Fig 2, where black dots are the reconstruction values and the vertical lines are the decision values. The sign of a residual can be either negative or positive. If a wavelet coefficient has already been significant in the base layer, before the encoding of its refined bits in enhancement layer, an extra sign of the residual needs to be encoded if this quantization is kept.

In the modified method, the reconstructed value \( r \) is changed to the lower point of the quantization bin, except the dead zone, as shown in (2).

\[
\begin{align*}
    r &= \begin{cases} 
        0, & \text{if } s = 0; \\
        s \times (\text{level} \times q + T), & \text{otherwise}\n    \end{cases}
\end{align*}
\]

(2)

The difference is taken between the original and the lower boundary point, as shown in the bottom half of Fig. 2. In the modification, the residue from the positive side remains positive and the residue from the negative side remains negative. Therefore in enhancement layer, for a coefficient already been significant in base layer, only the refinement bits needs to be encoded and the cost to encode an extra sign is saved.

### 3.2 Entropy Coding

In SLCCA, the entropy encoder adopted is the adaptive context based arithmetic encoder [7]. The context used to determine the conditional probability model of significant coefficient at \( [x; y] \) is related to the significance status of its parent and its eight neighbors. Let \( K_p[x; y] \) denote the significance status of the parent, i.e., \( K_p[x; y] = 1 \) if the parent pixel is significant, otherwise \( K_p[x; y] = 0 \). Let \( K_n[x; y] \) denotes the number significant coefficients in a 3x3 causal neighborhood of the current pixel \( [x; y] \). The adaptive context \( K[x; y] \) is selected by \( K[x; y] = K_n[x; y] + 9K_p[x; y] \), which yields a total of 18 possible models, as shown in Fig 3(a).

In ESLCCA, the old context is remained to encode the coefficients values and a new context model is introduced to code the map (significance and bits). To exploit the redundancy in the sign information, instead of using eight neighbors and its parent node, four immediate neighbors and its parent node are used, as shown in Fig 3(b). According to the available information concerning the signs of the immediate horizontal and vertical neighbors, nine contexts are generated. Let \( h \) equals 0 if both horizontal neighbors are insignificant or both are significant with different signs, 1 if at least one of the horizontal neighbors is positive and -1 if at least one of the horizontal neighbors is negative. Define in similar fashion for the vertical neighbors \( v \). Because both \( h \) and \( v \) have three possible values, totally there are nine new contexts for sign bit coding.

The context definition is shared between base layer and enhancement layer. The accumulated frequency in base layer will be used in enhancement layer directly, thus have more efficient lossless arithmetic encoding.

### 3.3. Small clusters
In SLCCA, extremely small clusters, which normally contain only 1-2 significant coefficients, are eliminated because they have little effect on the visual quality but need relatively high bit budget to encode. But in embedded SLCCA, quantization of base layer is quite coarse. Although the loss of such small clusters may not have big effects on base layer image quality, its impact on enhancement layer image will be non-negligible since the small clusters in base layer may turn out to be quite large clusters in enhancement layer. Such kind of loss can never be compensated no matter how many bits are spent in the enhancement layer, and thus the removal of small clusters is not permitted in ESLCCA.

4. PERFORMANCE EVALUATION

ESLCCA is evaluated on two natural 512x512 grayscale images, i.e., “Lena” and “Baboon”. Its coding performance is compared with both SLCCA and JPEG2000. As usual, the distortion is measured by peak signal-to-noise ratio (PSNR) defined as

$$PSNR[dB] = 20 \log_{10} \frac{255}{RMSE}$$ (3)

The bit rate is changing from 0.25 bpp (bit per pixel) to 1 bpp. As shown in Fig. 4, the coding performance of embedded SLCCA is in between JPEG2000 and SLCCA. And the subjective image quality comparison is shown in Fig 5. (a)- (d) respectively.

Figure 4. PSNR comparison of JPEG2000 ESLCCA and SLCCA

(a) Base layer of ESLCCA (b) JPEG2000 at 0.25bpp
(c) ESLCCA at 0.25bpp (d) SLCCA at 0.25bpp

Figure 5. (a): base layer of ESLCCA (with a bit rate of 0.0278bpp, with PSNR=25.65dB), (b)-(d): images coded by JPEG2000, ESLCCA and SLCCA respectively.

In conclusion, the ESLCCA sacrifice minor image quality to achieve embedding. Its performance is not as good as SLCCA, but still outperforms JPEG2000 and enjoys the fine scalability brought by the embedding in the enhancement layer.

5. References