AN EFFICIENT WAVELET BASED EMBEDDED COLOR IMAGE CODING TECHNIQUE USING BLOCK-TREE APPROACH

Athar Ali Moinuddin¹, Ekram Khan¹ and Mohammed Ghanbari²

¹Department of Electronics Engineering
Aligarh Muslim University, Aligarh, India
²Department of Electronic Systems Engineering
University of Essex, Colchester, UK
E-mail: aamoin@rediffmail.com, ekhan@lycos.com, ghan@essex.ac.uk

ABSTRACT

In this paper, we propose a simple and efficient technique for embedded color image coding. The proposed algorithm exploits inter- and intra-subband correlations of the wavelet transformed luminance (Y) and chrominance (U and V) planes as well as the interdependency among the coefficients of the three-color planes. The coefficients of the three-color planes are linked through the composite spatial orientation trees (CSOT) having root nodes in the Y-plane only. The CSOT is defined for blocks of coefficients rather than a single coefficient. Thus, the proposed algorithm combines the features of both zero-tree and zero-block based algorithm into a single algorithm due to the use of block-tree hierarchical structure. Simulation results show the improved performance of the proposed algorithm over other state-of-the-art coders such as MPEG-4, color SPIHT (CSPIHT), color SPECK (CPECK), and JPEG-2000.

Index Terms— wavelet transform, image coding, color image coding, embedded coding

1. INTRODUCTION

In recent years, wavelet based scalable image coding techniques [1]-[4] have attracted increased interest due to their excellent rate distortion performance, low encoding and decoding complexity, and rate scalability. The wavelet transformed images have both inter- and intra-subband correlations among the coefficients. Most of the wavelet based coding techniques use either zero-tree or zero-block to exploit these correlations. The zero-tree coding algorithms such as EZW [1] and SPIHT [2], exploit inter-subband correlations among the coefficients across the subbands only. On the other hand, zero-block coding algorithms like EBCOT [3] and SPECK [4] exploit intra-subband correlations among the coefficients with in the subbands only.

The above algorithms though primarily designed for coding of gray-scale images, but have also been extended for coding of color images and/or videos [4]-[6]. The trivial approach for coding of color images is to apply these algorithms independently on RGB or YUV color planes, and then serially concatenating the bit stream of each color plane. However, this would require proper bit allocation between each of the color planes and would not be fully embedded. The SPECK and SPIHT algorithms are extended to YUV color planes in [4] and [6] respectively to generate fully embedded bit streams. In both of these schemes, the list of insignificant pixels (LIP) and the list of insignificant sets (LIS) are initialized with the appropriate coordinates of the top-level subbands of the three-color planes serially one after the other. The appropriate algorithm (SPECK or SPIHT) is then used to code the three-color planes independently. However, it is a well-established fact that the three-color planes may be uncorrelated, but they are not independent. A color image coding algorithm exploiting the interdependency among color planes was first proposed in [5]. It uses EZW type of spatial orientation trees (SOT) with each luminance node having six children, four in the luminance plane and one in each of the two-chrominance planes. However, it suffers with the problem of early accumulation of chrominance nodes in LIP thereby reducing its coding efficiency. An efficient and embedded color image coding technique using SPIHT's SOT is proposed in [7], in which all three-color planes are linked through CSOT with root nodes in luminance plane only, but a root node has all of its descendents either in the luminance plane or chrominance planes. Kassim and Lee [8] also used the same SOT structure for 4:4:4 YUV color formats. These techniques efficiently exploit the interdependency of three-color planes and similarities of wavelet coefficients across the subbands only. However, in order to further increase the coding efficiency, there is need to exploit the intra-subband correlations of wavelet coefficients as well.
In this paper, we propose an efficient and embedded color image coding technique that combines the features of both zero-tree and zero-block coding algorithms, while exploiting the interdependency of the three-color planes. Images are assumed to be in 4:2:0 YUV CIF or QCIF format. The wavelet transformed color planes are divided into non-overlapping blocks of \( m \times n \) coefficients. To exploit the interdependency of the color planes, CSOT of blocks are used to link blocks of the three-color planes. Significant blocks are found using partitioning concept of SPIHT, whereas significant coefficients within each significant block are found using quad-tree partitioning of EBCOT or SPECK. A significant tree of blocks (block-tree) is recursively partitioned until significant coefficients are found.

The rest of the paper is organized as follows. The proposed algorithm is described in Section 2. Simulation results and discussions are presented in Section 3 and finally the paper is concluded in Section 4.

2. PROPOSED TECHNIQUE

In the proposed algorithm, each of the three-color planes (4:2:0 YUV) is wavelet transformed using \( N \) levels of decomposition for Y-plane and \( (N-1) \) levels for each of U- and V-planes. Since in 4:2:0 color format, the resolution of chrominance planes are one-quarter to that of the luminance plane, wavelet decomposition of chrominance planes by one level less than that of the luminance planes will result in the LL-subband of each transformed planes of the same dimensions. This will simplify the child-parent relationship to link the three-color planes together. For example, CIF and QCIF resolution images with four and three levels of decompositions respectively result in the top-most (LL) subband of Y-plane of size (22x18), whereas for each of U- and V-planes, it will be of size (11x9). This dimension of U- and V-planes is not suitable for SPIHT-type child-parent relationship as it requires even dimensions of LL-subband. Kim et al. have solved this problem by extending the chrominance planes before the 2-D spatial transformation [6]. However, this increases the complexity of encoder and decoder and slightly reduces the coding efficiency. In order to retain even dimensions of LL-subbands in the chrominance planes and to facilitate quad-tree partitioning, we suggest to use non-equal level of wavelet decomposition in our algorithm.

After dyadic wavelet decomposition, coefficients in each color planes are divided into blocks of \( m \times n \) coefficients. In each plane, blocks are linked together using child-parent relationship and CSOT of blocks (block-tree) similar to SPIHT [2]. Additionally, to exploit the interdependency of the color planes, the block-trees of three planes are linked together through composite tree of blocks. Fig. 1 shows the child-parent relationship of root and other nodes, where all the blocks of LL-subband of U and V planes are linked with the root blocks in the LL-subband of Y-plane only.

Significant information is stored in three ordered lists: a list of insignificant blocks (LIB), a list of insignificant block sets (LIBS), and a list of significant pixels (LSP). At the initialization step, blocks in the LL-subband of only Y-plane are added to LIB and also to LIBS as type ‘A’ entries. The LSP starts as an empty list.

The proposed coding technique like SPIHT and SPECK is also a bit-plane based and comprises two main stages, sorting and refinement passes, within each bit-plane. The coding process starts with the most significant bit-plane and proceeds toward the finest resolution. At every bit-plane, the encoder goes through the three lists in order, starting with LIB followed by LIBS and then to LSP. For each block in LIB, one bit is used to describe its significance. If a block is not significant, then it is a zero-block and a ‘0’ is send and it remains in LIB. Otherwise, if a block is significant, a ‘1’ is sent and it is partitioned into four adjacent blocks (quad-tree partitioning). The partitioning is repeated recursively until no further division is needed or the smallest possible block size (individual coefficient) is attained. At this stage four coefficients and their significance are tested individually. For an insignificant coefficient, ‘0’ bit is appended in the bitstream and its address is added to LIB as single coefficient block. On the other hand if a coefficient is found to be significant, a ‘1’ is send followed by its sign bit and the coefficient is added to the LSP. After testing all the four individual coefficients in the block, the current block is deleted from LIB.

Similarly, each set in LIBS requires one bit for significance information. Insignificant sets remain in LIBS while significant set will be partitioned into subsets. A significant type ‘A’ set will be partitioned into a type ‘B’ set and four offspring blocks. The type ‘B’ set is added to the end of LIBS while the four blocks are immediately examined for significance as is done in LIB. A significant type ‘B’ set will be partitioned into four type ‘A’ sets; all of them are added to the end of LIBS. Since all newly generated insignificant sets are added to the end of LIBS, they will be processed in the same manner at the same threshold until each one of them is examined. After each sorting pass, each coefficient in LSP, except those just added in the current bit plane, is refined with one bit. The algorithm then repeats the above procedure by decreasing the current threshold by a factor of two until the desired bit rate is achieved.

The proposed algorithm has three distinct advantages over SPIHT. First, it will combine many clustered zero-trees of the SPIHT, which are likely to occur in the early passes, thus creating longer zero-trees. Second, intra-subband correlations can also be partially exploited. Third, because of block-based, memory requirement for storing the lists will be less as compared to pixel-based techniques.
3. SIMULATION RESULTS

The proposed algorithm is implemented in software and its performance is compared with other state-of-the-art color image coders. The color image is in YUV (4:2:0 format), which is the first frame of the standard MPEG-4 QCIF test sequence Akio (176x144). The 9/7-biorhormal-filter bank is used for 2-D wavelet decomposition. For Y plane, we used 3-levels of wavelet decomposition, and each of U and V planes is decomposed with 2-levels only. Initial block size is considered as 2x2. The test image is encoded at 1.5 bpp and decoded at different bit rates from the same embedded bit stream. The objective quality of the decoded image in terms of the PSNR of the three color components $X, X \in \{Y, U, V\}$ is defined as

$$PSNR_X = 10 \log_{10} \frac{255^2}{mse(X)}$$  (1)

where $mse(X)$ is the mean-square-error for the component $X$.

The overall PSNR is obtained as

$$PSNR = 10 \log_{10} \frac{255^2}{(mse(Y) + mse(U) + mse(V))/3}$$  (2)

The rate versus PSNR results obtained with the proposed technique is compared with CSPIHT [6], CSPECK [4], JPEG-2000 VM 8.0 [9], and the PEZW algorithm used in MPEG-4 for intra-frame coding. The results for other coders are taken from [4]. Coding results for three different bit budgets are summarized in Table 1. The PSNR readings are of the decoded images. The proposed technique outperforms MPEG-4 by 0.9-1.4 dB for Y plane, by 0.4-2.3 dB for chrominance planes and by 0.9-1.7 dB in terms of overall PSNR. Also it outperforms JPEG-2000 by 1.5-2.3 dB for Y plane, but is inferior for chrominance planes. Since human eye is more sensitive to the changes in the brightness, the gain for Y plane as obtained from the proposed coder yields visually better results. But it is worth mentioning here that the EBCOT coder used in JPEG-2000 is much more complex as compared to the proposed coder. When compared to CSPIHT, the proposed coder gives the gain of approximately 0.5 dB for Y plane and 0.3-0.5 dB in terms of overall quality. These improvements are obtained despite the fact that to simplify the coder design, the chrominance planes are decomposed one level less than the luminance plane. In comparison to CSPECK, the proposed coder has almost comparable performance in the Y plane, but has the gain of about 0.7-2.0 dB for chrominance planes. Also in terms of the overall quality, the proposed coder outperforms the CSPECK by approx. 0.5-0.8 dB.

Fig. 2 compares the subjective quality of the decoded images obtained from various coders. Due to non-availability of CSPECK coder, we could not reproduce its decoded image. From Fig. 2 it can be observed that the proposed coders yields visually better image. Although image coded with JPEG-2000 has relatively better color resolution, but has poor sharpness (which is apparent by comparing gray rectangular area near left shoulder and a rectangular box on the left of Akio) as compared to the images of CSPIHT and that of the proposed algorithm. However, reproduced image with the proposed coder has the best sharpness and minimal distortions near the shoulders and neck.

Additionally, the proposed algorithm initializes the LIBS and LIB with block addresses of Y-plane only, where as in both CSPIHT and CSPECK, lists are initialized with coefficients of all three planes. For 2x2 initial block size, the proposed algorithm reduces the initialization memory by a factor of six.

4. CONCLUSIONS

In this paper, an efficient and embedded color image coding technique is proposed. The algorithm is based on the exploitation of inter- and intra-subband correlations of wavelet coefficients within each color plane and interdependency among three-color planes. It is observed that the proposed algorithm has better performance than algorithms based on either zero-tree (MPEG-4, CSPIHT) or zero-block (CSPECK, JPEG-2000). The proposed coder is especially attractive for very low-rate applications. Also, being block based, it reduces the memory requirements at the encoder and decoder.

5. REFERENCES


---

<table>
<thead>
<tr>
<th>Bits</th>
<th>Methods</th>
<th>Y</th>
<th>U</th>
<th>V</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>10256</td>
<td>MPEG4</td>
<td>32.3</td>
<td>34.2</td>
<td>36.9</td>
<td>34.1</td>
</tr>
<tr>
<td></td>
<td>CSPIHT</td>
<td>32.6</td>
<td>35.0</td>
<td>38.3</td>
<td>34.7</td>
</tr>
<tr>
<td></td>
<td>JPEG2000</td>
<td>30.8</td>
<td>37.4</td>
<td>38.8</td>
<td>34.2</td>
</tr>
<tr>
<td></td>
<td>CSPECK</td>
<td>33.5</td>
<td>34.4</td>
<td>36.6</td>
<td>34.6</td>
</tr>
<tr>
<td></td>
<td>Proposed</td>
<td>33.1</td>
<td>35.0</td>
<td>38.4</td>
<td>35.0</td>
</tr>
<tr>
<td>20816</td>
<td>MPEG4</td>
<td>37.5</td>
<td>39.1</td>
<td>41.0</td>
<td>38.9</td>
</tr>
<tr>
<td></td>
<td>CSPIHT</td>
<td>38.4</td>
<td>40.7</td>
<td>41.7</td>
<td>40.0</td>
</tr>
<tr>
<td></td>
<td>JPEG2000</td>
<td>36.9</td>
<td>43.3</td>
<td>43.8</td>
<td>40.1</td>
</tr>
<tr>
<td></td>
<td>CSPECK</td>
<td>39.1</td>
<td>38.7</td>
<td>39.8</td>
<td>39.2</td>
</tr>
<tr>
<td></td>
<td>Proposed</td>
<td>38.9</td>
<td>39.5</td>
<td>41.8</td>
<td>40.0</td>
</tr>
<tr>
<td>29240</td>
<td>MPEG4</td>
<td>40.8</td>
<td>41.5</td>
<td>42.6</td>
<td>41.6</td>
</tr>
<tr>
<td></td>
<td>CSPIHT</td>
<td>41.7</td>
<td>43.4</td>
<td>44.0</td>
<td>42.8</td>
</tr>
<tr>
<td></td>
<td>JPEG2000</td>
<td>40.6</td>
<td>46.1</td>
<td>46.6</td>
<td>43.5</td>
</tr>
<tr>
<td></td>
<td>CSPECK</td>
<td>41.9</td>
<td>43.1</td>
<td>43.6</td>
<td>42.8</td>
</tr>
<tr>
<td></td>
<td>Proposed</td>
<td>42.1</td>
<td>43.8</td>
<td>44.1</td>
<td>43.3</td>
</tr>
</tbody>
</table>

Table 1: PSNR (in dB) comparison of decoded ‘Akiyo’

---

Fig. 1: Child-parent relationship of nodes in the proposed algorithm (a) 2x2 root node blocks in LL-subband of Y and their offspring blocks (b) A node other than the root node and its offspring and (c) description of a node.

---

Fig. 2: Reconstructed Akiyo QCIF image compressed to 10256 bits: Original (top left), JPEG2000 (top right), color SPIHT (bottom left), and Proposed (bottom right).
Appendix A

Test Images and Videos

A.1 Test Images

Shown below are the test images used in this thesis.

![Lena](image.png)

*Lena*

![Barbara](image.png)

*Barbara*
Bike
Woman
Girls

Baboon
Goldhill

Toys
A.2 Test Videos

Shown below are the first frames of each of the test video sequences used in this thesis.

Hall-Monitor  Mother-Daughter

Salesman  Akiyo

Carphone