2005 WSEAS Int. Conf. on REMOTE SENSING, Venice, Italy, November 2-4, 2005 (pp23-28)

On Sign Encoding and Magnitude Refinement of Still Images

MARIA BRAS-AMOROS, JORGE GONZALEZ-CONEJERO, PERE GUITART-COLOM,
JOAN SERRA-SAGRISTA, FERNANDO GARCIA-VILCHEZ
Department of Information and Communications Engineering, ETSE
Universitat Autonoma de Barcelona
Cerdanyola del Valles, E-08193 Barcelona
SPAIN

Abstract: A shape-adaptive search is defined based on the BISK scheme and it is applied to sign encoding and magnitude refinement of images. It can be generalized to a complete bitplane encoder whose performance is comparable to that of other state-of-the-art encoders.

Key-Words: Lossy Data Compression, BISK, Non-Regular Boundary Images, DWT, Remote Sensing, High Resolution Images, Geographic Information Systems

1 Introduction

In relation to Remote Sensing (RS) and Geographic Information Systems (GIS) applications, multispectral and hyperspectral images have been successfully used for, among many others, image classification and segmentation. Nevertheless, inherent to these images is their huge size, so that it seems reasonable to look for a compression approach both for storage and transmission scenarios.

This work is focused on the sign encoding and magnitude refinement of wavelet-based bit plane encoders. Some competitive techniques, e.g. SPIHT [9] or SPECK [6], have no specific method to encode the sign of recently found significant coefficients nor the magnitude refinement bits; other encoders, e.g. JPEG2000 [11], EZBC [3], use an adaptive contextual arithmetic coder. From another perspective, [1] propose an alternative method that uses the wavelet transform properties to encode the transformed coefficients sign. Here we propose to adapt the notion of shape-adaptive coding [2] to define new methods to encode both the sign and the refinement bits of the coefficients.

The aim of shape-adaptive coding is to compress an image with a non-regular boundary assuming that both the encoder and the decoder know this boundary. Usually the image is located within a larger rectangular frame; pixels belonging to the image are named *opaque pixels*, pixels inside the frame but not belonging to the image are named *transparent pixels*. Some of the bitplane encoders used for shape-adaptive coding consist of well known regular-shape bitplane encoders, but treating only those bits corresponding to the opaque zone. This is the case of OB-SPIHT and OB-SPECK [4, 5, 7]. On the other hand, BISK [2, 8] is a method based on SPECK, with the novelty that it alternates set partitioning with opaque zone shrinking. This shrinking step consists of reducing each partitioned subset to the minimum rectangular set containing all its opaque coefficients.

The approach we present here consists of encoding the sign bits and the refinement bits as if they were the coefficients of an irregular-shape image. For the case of sign encoding, from a whole image we consider opaques those coefficients that have been found significant in the last significance pass; then we encode the sign inside this opaque zone by using a BISK-based search. Similarly, to encode the refinement bits in a given bitplane, we split all the previously found significant coefficients in various opaque zones: two coefficients are placed in the same opaque zone if all the first bits of their binary representation (up to the bitplane previous to the one currently being encoded) are the same; then, for each of these opaque zones we encode the refinement bits using the BISK-based search again.

Although the proposed methods for encoding the sign and the magnitude refinement bits may be considered independently and may be integrated to other bitplane encoders, the search scheme suggests a new complete wavelet transform-based bitplane encoder defined by a Repeated BISK-based search (REBISK). We will see that for some experiments, REBISK may give similar or even better results than other state-of-the-art encoders.

2 Two-valued Shape-Adaptive Search

The framework for the two-valued shape-adaptive search problem (TVSAS) is an irregular-boundary image with only two possible values (1/0 or $+/-$) where the boundary is known by both the coder and the decoder. The aim is to define a coding method that determines the value of each point in the image. This is the case of determining the sign of recently found significant bits, or also the case of determining the refinement bits at a given bitplane.

We say *classical search* (CS) when referring to the method that scans the whole image in a predefined order and just sends the value of each point. On the other hand, we consider a *BISK-based search* (BBS). Suppose that the two possible values are a and b, and that positions with value a are to be determined.

We use a First-In-First-Out (FIFO) structure whose nodes are image blocks. This FIFO structure of blocks, named BF, may be either initialized to an empty FIFO, or to a FIFO containing somehow selected blocks. The blocks in BF have to be evaluated and, if needed, partitioned. After the partitioning, each of the resulting parts containing at least one a is appended back to the BF. An empty block is inserted after the initial selected blocks to distinguish these blocks, which must be tested for significance, from those appended after partitioning a block. Now, following the BISK scheme which alternates the SPECK block partitioning with the shrinking step, we can define the BBS procedure written down in Figure 1.

Notice that for the BBS, each block can be partitioned into 1, 2, 3 or 4 blocks, while for the original BISK, each block is partitioned into 1 or 2 blocks. The elements in each set of 1, 2, 3 or 4 bits denoting the significance of the corresponding blocks are encoded together to save bits using an *ad hoc* mapping.

3 TVSAS Applied to Encoding of Coefficients

3.1 Magnitude Encoding

3.1.1 E-Sets and E-TVSAS

Consider a wavelet transformed image as a set I of coefficients. Let min, med and max be three values with $min \leq med \leq max$. Define the following subsets of I:

E–opaques(min, med, max)= $\{x \in I \mid min \leq |x| < max\}$.

E–significants(min, med, max)= $\{x \in I \mid med \leq |x| < max\}$.

E–insignificants(min, med, max)= $\{x \in I \mid min \leq |x| < med\}$.

Denote *E–TVSAS(min, med, max)* a TVSAS method that classifies E–significants and E–insignificants from the irregularly bounded set E–opaques. As examples of E–TVSAS methods, we consider:

E–CS: For each element in E–opaques, emit a 1 if it is in E–significants or a 0 otherwise.

E–SBBS: Determine the elements in E–significants among the ones in E–opaques by using the BISK-based search.

E–IBBS: Determine the elements in E–insignificants among the ones in E–opaques by using the BISK-based search.

E–comb: Shortest chain in $\{ 0|$*E–CS*, $10|$*E–SBBS*, $11|$*E–IBBS* $\}$.

Following the same ideas, one can define the sets E^+*-opaques*, E^+*–significants*, and E^+*–insignificants* (respectively E^-*–opaques*, E^-*–significants*, and E^-*–insignificants*) containing the positive (respectively negative) values of the analogous E–sets. Then E^+*–TVSAS* and E^-*–TVSAS* methods can be defined as for E–sets.

3.1.2 Significance Pass

Let MAX be the maximum absolute value among the coefficients in I and threshold $T = 2^{\lfloor log_2(MAX) \rfloor}$. The following procedures are consecutive significance passes: E–TVSAS(0,T,2T), E–TVSAS(0,T/2,T), E–TVSAS(0,T/4,T/2),...

Notice that if E–SBBS is used as the E–TVSAS method, and at each step the structure BF is initialized with the insignificant blocks from the previous bitplane, these significance passes are approximately the significance passes of SPECK. The unique difference is that SPECK uses its I sets while here we only consider rectangular sets.

3.1.3 Refinement Passes

An E–TVSAS(T, 3T/2, 2T) procedure after the second significance pass, namely E–TVSAS(0, T/2, T), gives the first magnitude refinement pass. Similarly, the three procedures E–TVSAS(T/2, 3T/4, T), E–TVSAS(T, 5T/4, 3T/2) and E–TVSAS(3T/2, 7T/4, 2T) after the third significance pass E–TVSAS(0, T/4, T/2) give the second magnitude refinement pass, and so on. Thus, the following algorithm yields a coding method for the absolute values in I.

```
for t from T to 1 by halving
   for min from 0 to 2T-2t by 2t
      E-TVSAS(min, min+t, min+2t);
```

Figure 2 shows the E–sets used for the significance pass and for the refinement passes.

3.2 Sign Encoding

As before, let min and max be two values with $min \leq max$ and define the subset *S–opaques(min,max)=*$\{x \in I \mid min \leq |x| < max\}$. We also denote *S–TVSAS(min,max)* a TVSAS method that classifies the positive and negative values in the (irregularly bounded) set S–opaques(min,max). As S–TVSAS methods, we consider:

S–CS: For each element in S–opaques, emit a 1 if it is positive or a 0 otherwise.

S–PBBS: Determine the positive values in S–opaques by using the BISK-based search.

S–NBBS: Determine the negative values in S–opaques by using the BISK-based search.

S–comb: Shortest chain in $\{0|S\text{–}CS, 10|S\text{–}PBBS, 11|S\text{–}NBBS\}$.

3.3 Image Coding

Now, the following algorithm yields a complete quality progressive image coding method.

```
for t from T to 1 by halving
    E-TVSAS(0, t, 2t);
    S-TVSAS(t, 2t);
    for min from 2t to 2T-2t by 2t
      E-TVSAS(min, min+t, min+2t);
```

The complete encoding algorithm that uses E^+/E^-–comb and S–comb as the E–TVSAS and S–TVSAS methods is named REBISK because of the Repeated E–BISK-based search.

4 Experimental Results

The lossy compression performance of REBISK is here compared to other coding systems for some images of the Landsat Corpus.

SPIHT, SPECK and REBISK results are produced with our frameworks. The three techniques are implemented in JAVA and do not use arithmetic coding. JPEG2000 results are produced with Kakadu [10], version v4.4. Kakadu is employed with the commonly lossy options {*Clevels=6 Creversible=no -precise*}. For all coding techniques, six levels of the 9/7 DWT are applied.

Evaluation of the different coding techniques is performed based on the trade-off between the compression ratio, given in bits per pixel (bpp), and the quality, given by the Peak Signal to Noise Ratio (PSNR), which is a measure accounting for the similarity between the original image I and the recovered image I^*, given in dB; for images with a B bpp bit depth, $PSNR = 10\,log_{10}\frac{(2^B-1)^2}{MSE}$, where the Mean Square Error (MSE) is given by $MSE = \frac{1}{N_x}\frac{1}{N_y}\sum_i^{N_x}\sum_j^{N_y}(I_{ij}-I^*_{ij})^2$.

4.1 Landsat Image

The Landsat Program is a joint effort of the U.S. Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA) to gather Earth resource data using a series of satellites. NASA was responsible for developing and launching the spacecrafts. The USGS is responsible for flight operations, maintenance, and management of all ground data reception, processing, archiving, product generation, and distribution. A primary objective of the Landsat Program is to ensure a collection of consistently calibrated Earth imagery. The Landsat Project is the longest-running enterprise for acquisition of moderate resolution imagery of the Earth from space. The Landsat 1 satellite was launched in 1972; the most recent, Landsat 7, was launched in 1999. The instruments on the Landsat satellites have acquired millions of images.

The chosen images taken for the experiments correspond to a Landsat 7 flight on 19 May 2002. The sensor producing these images is an Enhanced Thematic Mapper Plus (ETM+). The ETM+ instrument provides image data from eight spectral bands. The spatial resolution is 30 meters for the visible (band 1: blue; band 2: green; band 3: red), for the near-infrared (band 4), and for the mid-infrared (bands 5 and 7); resolution for the thermal infrared (bands 6 and 9) is 60 meters. The sensor also allows a panchromatic band (band 8) with 15 meters resolution. The satellite orbits at an altitude of 705 km and provides a 16-day, 233-orbit cycle.

The original images belong to path 197 and row 31. They are 11292 columns times 13350 rows of spatial size, but they have been cut off to 4096 times 4096 pixel scenes, with 8 bits per pixel resolution. Images have been ortocorrected and an atmospheric correction has also taken place. Compression experiments have been carried out on all eight spectral bands.

Fig. 3 shows the rate distortion curves obtained for all three visible bands for bit rates running from 0.001 bpp (compression ratio 8000:1) to 1.5 bpp (compression ratio 5:1). Fig. 4 shows the rate distortion curves obtained for all three infrared bands for the same bit rates. Finally, Fig. 5 shows the rate distortion curves obtained for the thermal bands for the same bit rates. In these figures, the curve for SPECK is not plotted because its performance is equivalent to SPIHT.

5 Conclusions

In this paper we have introduced a new approach for encoding sign and refinement bits. Each refinement pass and each sign encoding procedure is seen as a two-valued shape-adaptive search. To proceed with each search we defined BISK-based schemes which combine the block partitioning of SPECK with the block shrinking of BISK. In addition, the significance pass can be treated also as a two-valued search and, in this way, the whole bitplane encoder REBISK has been defined by a repetition of the BISK-based scheme.

For all compression ratios, REBISK provides better coding performance than other state-of-the-art coding techniques as SPIHT and SPECK (see Fig 3, and Fig. 4 and Fig. 5). On the other hand, the comparison of the results obtained with JPEG2000 and the results of REBISK for the rate-distortion curves in Figs. 3, 4 and 5 shows that for all bit rates JPEG2000 achieves better results than REBISK although they are very close. Notice that the JPEG2000 coding scheme implements arithmetic codification in front of the no arithmetic coding strategy followed for the REBISK algorithm.

Acknowledgments: This work has been partially supported by the Spanish Government MCYT Grant TIC2003-08604-C04-01 and by FEDER.

References:

[1] Aaron T. Deever and Sheila S. Hemami. Efficient sign coding and estimation of zero-quantized coefficients in embedded wavelet image codecs. *IEEE Trans. Image Process.*, 12(4):420–430, 2003.

[2] James E. Fowler. Shape-adaptive coding using binary set splitting with K-D trees. In *IEEE International Conference on Image Processing*, volume 2, pages 1301–1304, Singapore, October 2004.

[3] Shih-Ta Hsiang. Embedded image coding using zeroblocks of subband/wavelet coefficients and context modeling. In *IEEE Data Compression Conference*, pages 83–92. IEEE, March 2001.

[4] Z. Lu and W. Pearlman. Wavelet coding of video object by object-based SPECK algorithm. In *Proc. of SPC*, pages 413–416, Seoul, Korea, April 2001.

[5] Gavin Minami, Zixiang Xiong, Albert Wang, and Sanjeev Mehrotra. 3-D wavelet coding of video with arbitrary regions of support. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(9):1063 – 1068, September 2001.

[6] W.A. Pearlman, A. Islam, N. Nagaraj, and A. Said. Efficient, low-complexity image coding with a set-partitioning embedded block coder. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(11):1219–1235, November 2004.

[7] M. Penedo, W.A. Pearman, P.G. Tahoces, M. Souto, and J.J. Vidal. Region-based wavelet coding methods for digital mammography. *IEEE Transacions on Medical Imaging*, 22(10):1288 – 1296, October 2003.

[8] J.T. Rucker and J.E. Fowler. Coding of ocean-temperature volumes using binary set splitting with K-D trees. In *IEEE International Geoscience and Remote Sensing Symposium*, September 2004.

[9] Amir Said and William A. Pearlman. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Transactions on Circuits and Systems for Video Technology*, 6:243–250, June 1996.

[10] D.S. Taubman. Kakadu software. http://www.kakadusoftware.com/, 2000.

[11] D.S. Taubman and M.W. Marcellin. *JPEG2000: Image Compression Fundamentals, Standards, and Practice*, volume 642. Kluwer International Series in Engineering and Computer Science, 2002.

Figure 1: Functions used by the BISK-based search.

```
procedure BBS(BF)
{
  if BF is empty then
    for each wavelet transf subband S
       SS=Shrink(S)
            #i.e. SS is the
            #minimal rectangle
            #containing S
       if SS is not empty
         append SS to BF

  append empty block to BF

  S = extract first block of BF
  while S is not empty do
    if a in S
       emit 1
       Partition (S,BF)
    else
       emit 0
    S = extract first element of BF

  while BF is not empty do
    S = extract first element of BF
    Partition (S,BF)
}
```

```
procedure Partition(S,BF)
{
  horizontal split S into S1 and S2:
    S1: size floor(y(S)/2) by x(S)
          #where
          #y(S) is the n. of rows of S
          #x(S) is the n. of columns of S
    S2: size (y(S)-floor(y(S)/2)) by x(S)
  Shrink(S1)
  Shrink(S2)

  if S1 is not empty then
    vertical split S1 into s1 and s2:
      s1: size y(S1) by floor(x(S1)/2)
      s2: size y(S1) by (x(S1)-floor(y(S1)/2)
    Shrink(s1)
    Shrink(s2)

  if S2 is not empty then
    vertical split S2 into s3 and s4:
      s3: size y(S2) by floor(x(S2)/2)
      s4: size y(S2) by (x(S2)-floor(y(S2)/2)
    Shrink(s3)
    Shrink(s4)

for i from 1 to 4
 if si is not empty
   if a in si then
     emit 1
     if si is not a single coefficient
       append si to BF
   else
     emit 0
}
```

Figure 2: E-sets for the signifi cance pass (left column) and for the refi nement pass (right column)

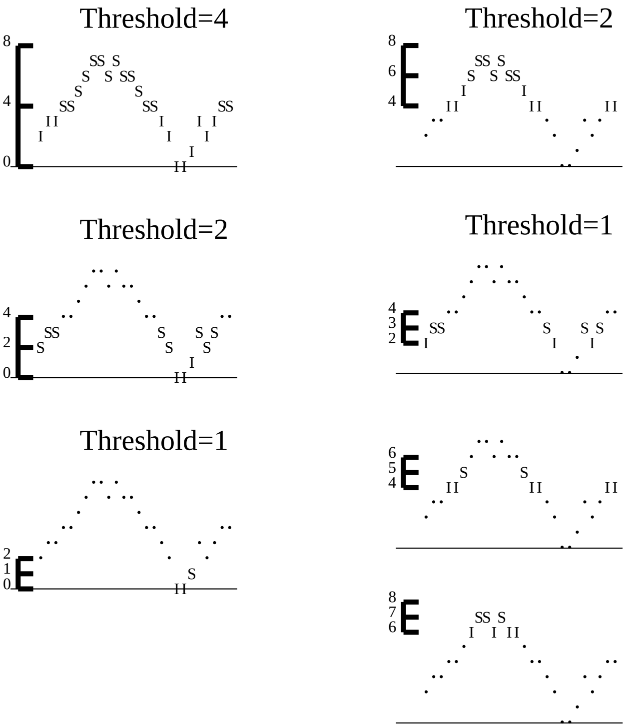

Figure 3: Rate-distortion curves for 4096×4096 Landsat Image. Visible bands.

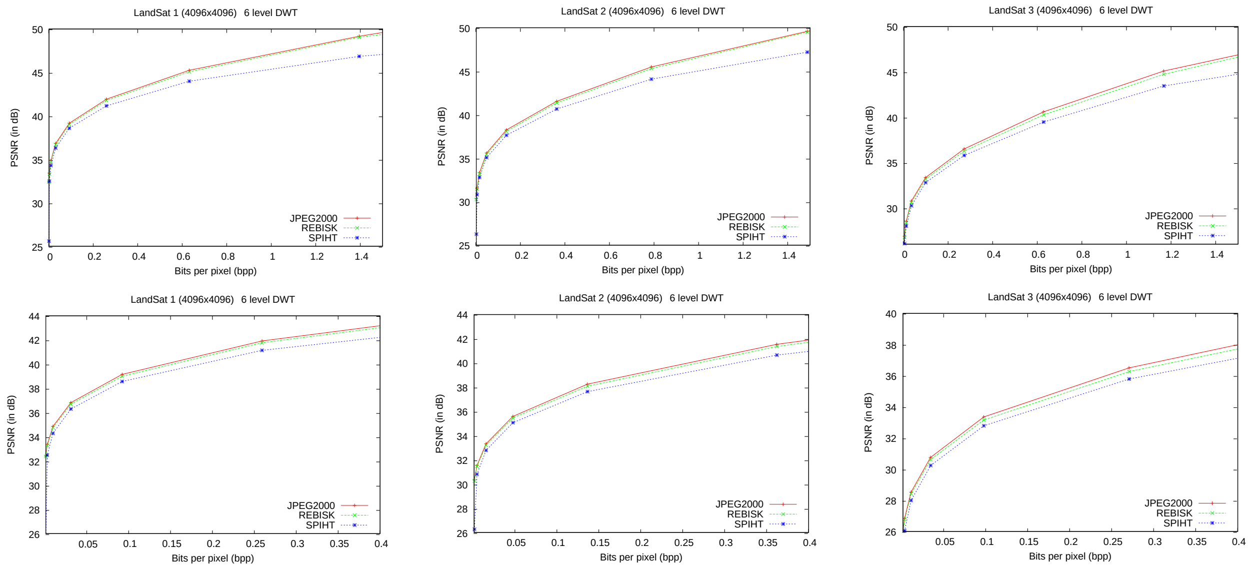

Figure 4: Rate-distortion curves for 4096×4096 Landsat Image. Infrared bands.

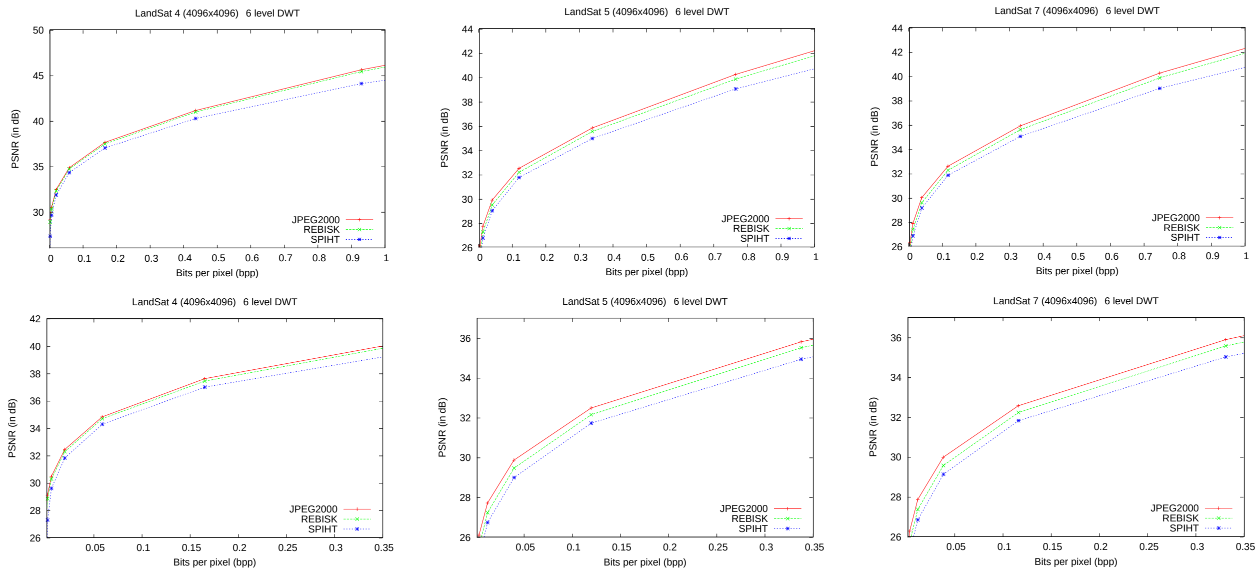

Figure 5: Rate-distortion curves for 4096×4096 Landsat Image. Thermal bands.