Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

More efficient ground truth ROI image coding technique: implementation and wavelet based application analysis*

KUMARAYAPA Ajith, ZHANG Ye‡

(*Department of Electronics and Communication Engineering, Harbin Institute of Technology, Harbin 150001, China*)

E-mail: ajithky@yahoo.com; zhye@hit.edu.cn

Received Dec. 19, 2006; revision accepted Mar. 12, 2007

Abstract: In this paper, more efficient, low-complexity and reliable region of interest (ROI) image codec for compressing smooth low texture remote sensing images is proposed. We explore the efficiency of the modified ROI codec with respect to the selected set of convenient wavelet filters, which is a novel method. Such ROI coding experiment analysis representing low bit rate lossy to high quality lossless reconstruction with timing analysis is useful for improving remote sensing ground truth surveillance efficiency in terms of time and quality. The subjective [i.e. fair, five observer (HVS) evaluations using enhanced 3D picture view Hyper memory display technology] and the objective results revealed that for faster ground truth ROI coding applications, the Symlet-4 adaptation performs better than Biorthogonal 4.4 and Biorthogonal 6.8. However, the discrete Meyer wavelet adaptation is the best solution for delayed ROI image reconstructions.

Key words: Smooth low texture remote sensing (SLTRS) image, Modified region of interest (ROI) image codec, Wavelet filters, Low bit rate (LBR), High bit rate (HBR), ROI coding

doi:10.1631/jzus.2007.A0835 **Document code:** A **CLC number:** TN919.81

INTRODUCTION

Remote sensing surveillance applications in military and natural disaster such as hurricane and tornado forming conditions, use smooth low texture remote sensing (SLTRS) images, where the low texture means there exists a majority of low frequency components (Hou *et al.*, 2004) representing on-ground or suburb aerial truth. Such images are massive in size, and it is necessary to progressively encode/decode the interested ROI initially at very low bit rate (LBR) coding in minimum onboard processing time for fast preview and then encode/decode the ROI image at high bit rate (HBR) with more time.

However, the remote sensing missions using observatory aircrafts (e.g. AVIRIS) (Zhang and Desai, 2000) or satellites (e.g. LandSat-5) encounter practical limitations such as onboard storage capacity and processor efficiency, limited downlink contact/transmission time and downlink bandwidth.

Therefore, there is a great need of low-complexity coding algorithm/technique which enables fast onboard SLTRS image compression which can facilitate LBR coding, efficient mixed coding of reversible ROI with respect to background of the acquired scene and alleviate progressive transmission. In this 2D SLTRS image coding research, we modified the set partitioning in hierarchical trees (SPIHT, Said and Pearlman, 1996). The following research literature supports our choice of the algorithm.

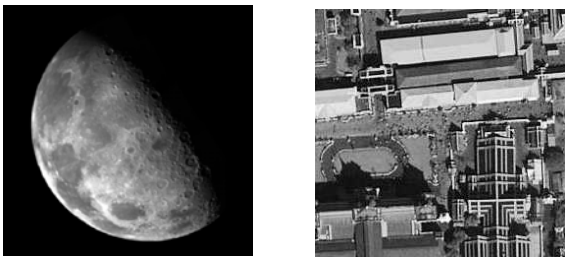
At LBR, tiling artifacts in the tile boundary of EBCOT: J2K coded images are annoying (Qin *et al.*, 2004). For multiple ROI coding, J2K: MAXSHIFT may cause overflow problems due to limited precision of the implementation (Chen and Wei, 2002). The onboard hardware implementation with J2K is more complex than the modified forms of SPIHT

‡ Corresponding author

* Project (No. 2004144013) supported by the Chinese Government Scholarship Council, China

algorithm (Huang *et al.*, 2003).

The ROI ground truth coding efficiency of our modified codec is verified using two 2D SLTRS test images with respect to the selected Orthogonal, Biorthogonal and discrete Meyer wavelet filter banks out of 110 variety.



(a) (b)

Fig.1 (a) Smooth low texture remote sensing image; (b) Remote sensing image with rich texture edge patterns

Based on the outcomes of our experiment result analysis, efficient state-of-the-art coding technology embedded with the use of different wavelet filters for specific 2D SLTRS image coding application environments can be chosen. Section 2 gives the convenient strategy for ROI specification, the consequent processing of the ROI and the theoretical formulation of the modified ROI ground truth codec. Section 3 gives experiment research and result analysis. Section 4 concludes the work.

EFFICIENT GROUND TRUTH ROI IMAGE CODING WITH MODIFIED TECHNOLOGY

Coding technology suitable for SLTRS images

The main attraction of the coding technique suitable for downlinking 2D SLTRS images from observatory aircraft or geo-orbiting satellite is the ability of the encoder to produce progressive bit stream. Moreover, the encoding process should be able to be stopped at any arbitrarily compressed file size while giving a recognizable image or it could be able to obtain high-fidelity 2D ground truth ROI image with the increase of downlinking time. The solution is to employ an efficient ROI coding algorithm with simplicity to be used for onboard hardware implementations, which also facilitates low to high bit rate coding and can be employed with many choices of efficient wavelet filters.

The modified ROI codec (based on SPIHT) satisfies the above-required features giving efficient resolution and SNR scalable results in terms of better subjective (Human Visual System, HVS) appearance with arbitrary bit rate coding.

The SPIHT original codec adopts a hierarchical quadtree (octave band decomposition) data structure on a wavelet decomposed image. The wavelet coefficients are encoded in four phases: initialization pass, sorting pass, refinement pass, and repetition from phase 2. Then the bit stream is transmitted (Said and Pearlman, 1996).

Convenient strategy for specifying ROI and consequent processing

First, the ROI is defined in the image domain by the expert observer (spending on average 1~2 s) while traveling in an aircraft or in the case of a satellite, with the approximate LBR on ground reconstruction using the progressive bit stream from the codec. This addresses the awareness problem of how to select the smooth low texture nature, which represents a majority of low frequency components in the interested remote sensing image.

Then the same position (GIS coordinate) is rescanned with the preset ROI for the coding of the 2D SLTRS image. For the task, only 2 coordinates are defined (in pixels) $\{(x_1,y_1), (x_1+W,y_1+H)\}$ as shown in Fig.2.

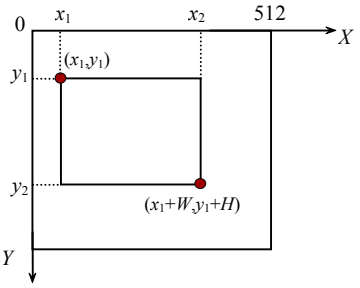


Fig.2 Initially specified ROI in spatial domain

For image containing ROI, applying an efficient form of discrete wavelet transformation (Li and Li, 2000; Chen *et al*., 2006) with the user predefined or automatically (a satellite) demarcated mask, the important features such as spatial correlation, locality and self-similarity across subbands are well preserved. Moreover, it keeps the number of transform coeffi-

cients needed to reconstruct ROI the same as the number of pixels in the image ROI.

Let the representation of two coordinates in the wavelet coefficient domain be (X_{C1},Y_{C1}) and (X_{C2},Y_{C2}). Then $X_{C1}=2^{-L}x_1$, $Y_{C1}=2^{-L}y_1$, $X_{C2}=X_{C1}+2^{-L}(x_2-x_1)$, $Y_{C2}=Y_{C1}+2^{-L}(y_2-y_1)$, where L is the decomposition level.

According to the shape of the ROI, the coding algorithm needs to keep track of the wavelet coefficient locations. To efficiently obtain the information on ROI coefficients, a rectangular shape mask image which specifies the ROI is decomposed by the same DWT used in wavelet decomposition of the image. In each decomposition stage, each subband with the decomposed mask contains ROI specifying information.

The number of trees denoting the ROI mask (N^{ROI}) is equal to the number of coefficients representing ROI in high-resolution lowest frequency subband (LL)

$$N^{ROI}\cong 2^{-L}(x_2-x_1)(y_2-y_1). \tag{1}$$

The largest energy gains are along LL subband. The consequent decomposition attributes of ROI merged in our new ROI codec can be illustrated as the rectangular blocks in Fig.3. With the decomposition process, the ROI mask representing spatial orientation tree (SOT) is spread as $Tree(x_i,y_j,b,p)$, where $p=[1,2,\ldots,L]$ and $b=[LL_p,LH_p,HL_p,HH_p]$. The information on ROI coefficients is obtained by successively decomposing the approximation coefficients for the number of decomposition levels. Once the ROI coefficients are identified, the modified coding algorithm (based on SPIHT) will be applied on these coefficients to create the embedded bit stream. Here an object-based non-balanced SOT (NBSOT) is used with the modified codec in order to efficiently code ROI coefficients by taking advantage of ROI information in the transform domain.

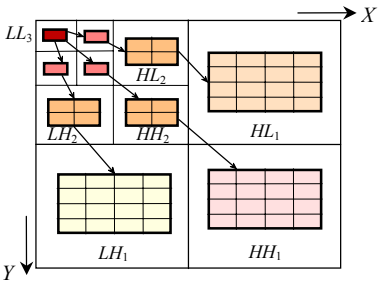


Fig.3 The new zero tree structure corresponding to the ROI attributes

When the SOTs are established in the initialization step of the codec, the ROI information obtained from decomposed ROI mask is used to mark the new NBSOT structure.

Modification to the algorithm

For our new modified codec, the accumulation of descendant wavelet coefficients from different subbands to the NBSOT can be given as

$$\rho(k)=\bigcup_{s\in O(k)}\{d(s)\cup\rho(s)\}. \tag{2}$$

Here, $O(k)$ is a descendant node in NBSOT corresponding to the ROI, $O(k)\in N^{ROI}$; $d(s)=1$ and s is the wavelet coefficient with index k corresponding to position (x_i,y_j) in image spatial domain.

As in Fig.3, the wavelet coefficients representing ROI image will be encoded after initializing with list of insignificant sets (LIS) and list of insignificant pixels (LIP). Let N_T^{ROI} be the set of total number of coefficients with trees denoting ROI. The 9-level decomposition process related to our modified new codec is defined as

$$LL_{9(LIP)}^{ROI}=\{k\mid k\in(LL_9\cap N_T^{ROI})\}, \tag{3}$$

$$LL_{9(LIS)}^{ROI}=\{k\mid k\in LL_9,\ s\in\rho(k)\}. \tag{4}$$

During codec initialization process the tree roots with indices $k\in LL_{9(LIS)}^{ROI}$ are used to initialize the LIS and the coordinates (X_{C1k},Y_{C1k}) and (X_{C2k},Y_{C2k}); $k\in LL_{9(LIP)}^{ROI}$ are used to initialize LIP.

Then, the algorithm proceeds to the modified step (Step 2) to process ROI image in our codec. The ROI coding modification introduced to the codec is given by the following algorithm:

Step 2:

For each descendant coefficient $d(i,j)\in\rho(k)$, do

If all $d(i,j)\notin N_T^{ROI}$, then TRIM the $d(i,j)$ from SOT;

If $d(i,j)\in N_T^{ROI}$, then

If $parent(i,j)\notin N_T^{ROI}$, then

Disjoin $d(i,j)$ from $parent(i,j)$;

Search a TRUE parent from neighborhood$\in p(i,j)$;

If find TRUE $parent(i,j)$, then

Join $d(i,j)$ to TRUE $parent(i,j)$;

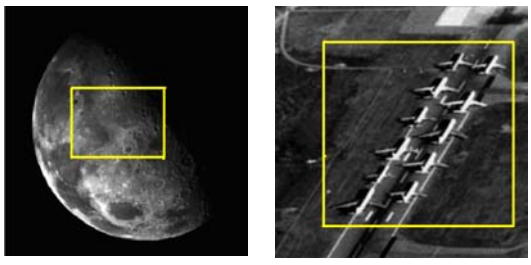
Else put $d(i,j)$ into LNP;

Then, the modified codec proceed with additional list of no parents (LNP) to the sorting phase and then to the refinement phase giving consequent output of well-embedded progressive bit stream representing 2D SLTRS image. After the efficient encoding process, the bit stream will be transmitted to the ground station, at which the inverse process of decoding ROI ground truth images and resulting object identifications and analysis tasks will take place.

EXPERIMENTS AND ANALYSIS

Ground truth ROI image coding power in PSNR (dB) and the onboard processing time efficiency with the new modified ROI codec are evaluated with respect to 8 wavelet analysis filters (Zhang *et al.*, 1993) representing Orthogonal (Daubechies-4 wavelets, Symlet-4, Haar), Biorthogonal wavelets (Biorth.1.1, Biorth.4.4, Biorth.6.8, Inv.Biorth.4.4) and discrete Meyer wavelet. Throughout the rest of this paper, unless otherwise stated, they are represented as Daub.4, Symlet4, Haar, Biorth.(1.1, 4.4, 6.8), Inv. Biorth.4.4 and D.Meyer, respectively.

Initially, the LBR ROI coding power (at low bit rates) is evaluated using the 2 SLTRS test images: Moon and Airport-3. The goal is to observe the fairness of ROI codec results with respect to different filters for coding different scenes of SLTRS images (e.g. Moon, fairly smooth surface, and Airport-3, with medium smooth natural image without regular texture patterns). For the Moon image, a rectangular shape ROI is chosen to observe concave surfaces on its crust to identify volcanoes. For Airport-3, a rectangular shape ROI is chosen to clearly demarcate the interested ground truth of 11 aircrafts.



(a) (b)

Fig.4 Rectangular shape ROIs are chosen on the two test images 'Moon' (a) and 'Airport-3' (b)

PSNR of ROI and processor timing performance results in Table 1 demonstrate the fairness of considering different SLTRS images for LBR coding with our modified ROI codec. The performance results (averages obtained from 5 repetitions) for the two images with respect to different filters are almost similar.

Table 1 The low bit rate (0.05 bpp) ROI ground truth image coding results (PSNR: dB; time: s)

Filter	Moon		Airport-3	
	PSNR	Time	PSNR	Time
Biorth.1.1	31.759	20.015	25.266	20.982
Biorth.4.4	33.854	17.739	28.945	17.559
Inv.Biorth.4.4	32.740	20.647	27.300	21.298
Biorth.6.8	34.054	23.686	28.646	24.611
Haar	31.647	18.686	25.223	20.657
Symlet4	27.847	21.247	28.615	21.296
Daub.4	32.802	20.812	27.627	21.158
D.Meyer	33.847	22.626	28.976	22.212

Since the initial fast preview of the ROI at very low bit rate coding (e.g. at 0.04 bpp), medium (e.g. at 0.08 bpp) and delayed, nearly lossless ROI coding (e.g. at 1.03 bpp) reconstructions and ground truth object identification efficiency are very important in remote sensing. Secondly, we present the HVS subjective performance results (for Airport-3) of the ROI codec representing the above 3 stages of coding with respect to 8 wavelet filters. Fig.5 and Table 2 are the average evaluations by 5 human observers (HVS) using a display equipped with 3D view accelerator consisting of 256 MB VGA Hyper memory.

Table 2 The 5 human observer's (HVS) average results for ROI codec performance with respect to 8 wavelet filters at LBR, Medium and HBR coding

Filter	LBR*	Medium bit rate**	HBR***
Biorth.1.1	Poor	Poor	Poor
Biorth.4.4	Medium	Good	Good
Inv.Biorth.4.4	Poor	Medium	Medium
Biorth.6.8	Good	Medium	Good
Haar	Poor	Poor	Poor
Symlet4	Good	Medium	Medium
Daub.4	Medium	Medium	Medium
D.Meyer	Medium	Good	Good

*: at 0.04 bpp; **: at 0.08 bpp; ***: at 1.03 bpp

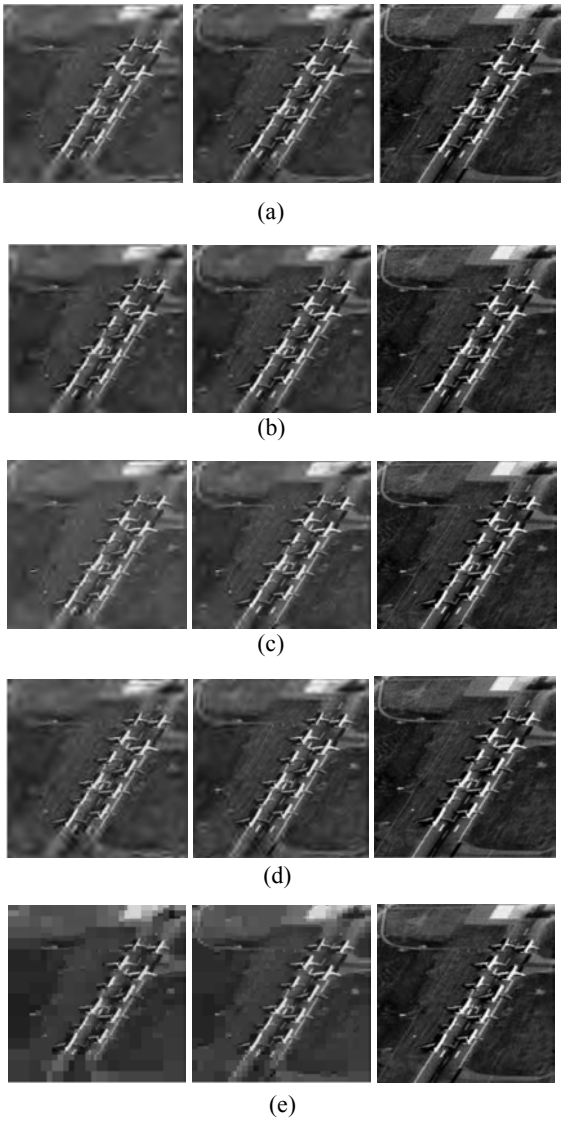


Fig.5 HVS subjective performance results for Airport-3 of the ROI codec with respect to 8 filters. This figure illustrates only 5 wavelet filters, 4 better and 1 worst (e), due to space limitation. (a) Biorth.6.8; (b) Biorth.4.4; (c) Symlet4; (d) D.Meyer; (e) Haar. In each subfigures, from left to right, the bit rates are 0.04 bpp, 0.08 bpp, and delayed reconstruction at 1.03 bpp respectively

Thirdly, the Airport-3 ROI image processing time and the relevant bit rate distortions are evaluated for a range of bit rates with respect to the wavelet filters. Results are given in Table 3, Fig.6 and Fig.7.

Final results of the research analysis

From above results, it can be concluded that for applications which need fast preview and then to reconstruct better quality SLTRS ROI, the Symlet-4 adaptation with our new modified codec gives better results, comparatively with Biorthogonal 4.4 and Biorthogonal 6.8. For delayed nearly lossyless ROI image reconstruction, the discrete Meyer wavelet adaptation is the best solution in terms of PSNR and HVS results, however, with the cost of high processing time of the encoder.

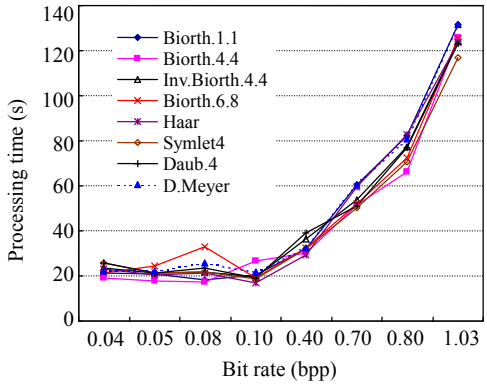


Fig.6 The Airport-3 ROI image processing time vs a range of bpp for the modified codec with respect to 8 wavelet filters

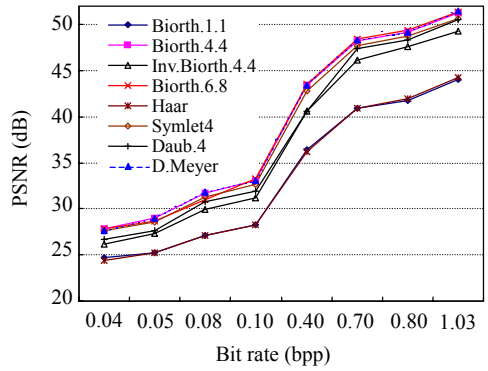


Fig.7 Applicable range of bit rates against the PSNR for Airport-3 with respect to 8 wavelet filters

CONCLUSION

In this research, aiming at fast ground truth ROI image signal coding applications, we introduced a novel concept of using 8 more efficient wavelet filter adaptations (we previously verified using 110 filters in another research) based on subjective (HVS observations combined with new 3D display technology) and objective performance analysis for the new modified ROI image codec. The fairness of our codec for LBR SLTRS image ROI coding was comparatively

Table 3 The Airport-3 ROI image processing time (unit: s) using Pentium 1.73 GHz CPU and the relevant bit rate distortions (PSNR, unit: dB) for a range of bit rates with respect to the 8 wavelet filters

Bit rate	Parameter	Biorth.1.1	Biorth.4.4	Inv.Biorth.4.4	Biorth.6.8	Haar	Symlet4	Daub.4	D.Meyer
0.04	PSNR	24.657	27.817	26.220	27.872	24.415	27.623	26.694	27.596
	Time	21.157	19.059	23.428	21.272	23.095	25.629	25.793	22.135
0.05	PSNR	25.266	28.942	27.300	28.646	25.223	28.615	27.627	28.976
	Time	20.982	17.559	21.298	24.611	20.657	21.296	21.158	22.212
0.08	PSNR	27.116	31.749	29.885	31.010	27.085	31.336	30.821	31.825
	Time	18.240	17.297	21.647	32.857	21.240	21.567	23.494	25.887
0.10	PSNR	28.254	33.075	31.204	33.264	28.215	32.622	31.947	33.037
	Time	20.311	26.620	19.669	19.027	17.001	18.456	19.145	21.809
0.40	PSNR	36.433	43.428	40.562	43.539	36.258	42.777	40.554	43.459
	Time	32.165	30.307	36.499	32.308	29.424	32.191	39.012	32.566
0.70	PSNR	40.935	48.248	46.195	48.416	40.950	47.665	47.399	48.300
	Time	60.378	51.584	53.865	52.063	59.679	50.339	51.034	60.225
0.80	PSNR	41.789	49.189	47.650	49.342	41.999	48.741	48.342	49.157
	Time	82.224	66.323	77.411	72.260	82.516	70.627	77.016	80.778
1.03	PSNR	44.016	51.282	49.319	51.387	44.249	50.676	50.579	51.468
	Time	131.459	125.823	123.839	125.296	123.612	116.863	123.019	131.334

evaluated initially using 2 test images. The performance of our ROI codec for the range of bit rates representing low to high bit rate coding applications is regular and stable. Moreover, our experiment gives the aspirations of designing further improved wavelet filters for the adaptation as well the design of low-complexity, efficient ROI codec hardware platform for our modified ROI coding algorithm suitable for smooth low texture remote sensing images with a majority of low frequency components/contents.

ACKNOWLEDGEMENT

This paper is dedicated to mother Mrs. Violet Subasinghe Kumarayapa, who gave up last minute wish to be with her only child for making the first author's PhD dissertation works in China cheerful and success.

References

Chen, M.J., Wei, C., 2002. Improved Region-of-Interest Image Coder and its Application. IEEE International Conference of Consumer Electronics, p.226-227.

Chen, H., Zhang, Y., Wang, L., 2006. Constraint Net Based Error Recovery for High-frequency Information of Remote Sensing Image in JPEG2000 Transmission. ISSCAA 2006, p.436-439. [doi:10.1109/ISSCAA.2006.1627659]

Hou, X.S., Liu, G.Z., Zou, Y.Y., 2004. SAR image data compression using wavelet packet transform and universal-trellis coded quantization. *IEEE Trans. on Geosci. Remote Sensing*, **42**(11):2632-2642. [doi:10.1109/TGRS.2004.834761]

Huang, W.B., Chang, Y.J., Su, A.W.Y., Kuo, Y.H., 2003. VLSI Design of DWT/Modified Efficient SPIHT Based Image Codec. Proc. Joint Conference of the 4th International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia, **1**:248-252. [doi:10.1109/ICICS.2003.1292453]

Li, S., Li, W., 2000. Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding. *IEEE Trans. on Circuits Syst. Video Technol.*, **10**(5):725-743. [doi:10.1109/76.856450]

Qin, X., Yan, X.L., Yang, C.P., Ye, Y., 2004. Tiling Artifact Reduction for JPEG2000 Image at Low Bit-Rate. IEEE Inter. Conf. on Multimedia and Expo, **2**:1419-1422.

Said, A., Pearlman, W.A., 1996. A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. on Circuits Syst. Video Technol.*, **6**(3):243-250. [doi:10.1109/76.499834]

Zhang, Y., Zhao, G., Yu, H., 1993. Wavelet Transform and its Application in Image Compression. IEEE Region 10 Conference of Computer Communication, Control and Power Engineering, p.418-421.

Zhang, Y., Desai, M.D., 2000. Hyperspectral image compression based on adaptive recursive bidirectional prediction/JPEG. *Patt. Recog.*, **33**(11):1851-1861. [doi:10.1016/S0031-3203(99)00180-6]