

A comparative study of advanced frequency-domain coding techniques in compression of infrared line-scan images

by D. Milovanovic¹, B. Więcek², A. Marincic¹, G. Petrovic¹
and Z. Barbaric¹

¹ Faculty of Electrical Engineering, P.O.Box 816, 11001 Belgrade, Yugoslavia

² Technical University of Łódź, Institute of Electronics, 18/22 Stefanowskiego St, 90-924, Poland

Abstract

A comparative study of transform coding techniques in data compression of IRLS (InfraRed Line-Scanner) images is described in this paper. In performance evaluation of block transform and wavelet image coders, we first carefully formed two sets of "good" basis functions: 4 types of generalized overlapped block transform (LOT) and 3 types of biorthonormal discrete wavelets transform (DWT). We found by extensive computer simulation on IRLS representative test-images, type and order of transform in the most efficient scalar quantized-entropy coder in each set for compression ratio from 10:1 to 50:1, based on objective and subjective quality measures of decoded test images. Finally, in performance comparison of the most efficient generalized block transform and wavelet coder we found that biorthogonal Doubechies DWT based coder outperforms standard DCT and optimized LOT coders. Additionally, DWT coder gives possibility for further research on spatial image segmentation and terrain/altitude adaptive IRLS image coding.

1. Introduction

InfraRed Line-Scanner (IRLS) is a mechanic-optoelectronic device for high-resolution image acquisition in the thermal IR spectral range. The terrain is sampled by a transverse scan as the aircraft moves forward along track. Consequently, IRLS operates as a linear angle scanning system, giving angular scan and linear displacement along track as the two dimensions of the still image [1]. The need for digital IRLS image compression arises from the fact that scanner generates enormous instantaneous bit rates which can easily saturate transmission link, the real-time display and human observer. The goal of image compression is the reduction of waste amount of image data for storage and/or transmission, preserving the quality of the reconstructed image under constraint of computational complexity and delay. One of the most important IRLS image compression requirements is the preservation of the amplitude and phase of the edges as well as a good visual quality of the image background. Therefore, IRLS image have only recently become candidates for image compression [2].

Statistical image compression techniques take advantage of the intrinsic features of IRLS image [3,4] as well as of their relation to the final human observer to eliminate redundancy before transmission. Compression techniques are divided into two classes: spatial-domain methods and frequency-domain methods. Spatial-domain techniques directly code pixel values into a binary stream, and frequency-domain techniques transform or filter pixel values before further coding. Spatial-domain methods have the advantage of exploiting the spatial correlation in images, while frequency-domain methods have the advantage of exploiting uneven energy and probability distribution in transform domain by applying different coding schemes for different frequency components. Predictive methods and vector quantization are representative examples of spatial-domain methods. Typical frequency-domain methods are transform coding (TC) and subband coding (SBC). In general, frequency-domain methods are preferred to spatial-domain methods in many aspects such as rate-distortion performance and computational complexity.

A way to study and optimize transform-based compression techniques is to make software simulation of coder and test it on representative images. The first part is the comparison of block transform and wavelet coders based on unified theory of perfect-reconstruction filter-banks. The second part is the experimental comparative study of considered transform coders. In the final part of the paper the important results are pointed out.

2. Comparison of block transform and wavelet coders

Statistical properties of IRLS images depend highly on the characteristics of scanner and the effective ground resolution. As a result, the correlation between adjacent pixels is usually very high, that implies a high degree of redundancy in the raw acquired image [3,4]. The *block orthogonal transform* is used to eliminate redundant information and to provide a suitable representation for an efficient individual quantization of coefficients. The idea is to form a decomposition of images into weighted sums of linear independent transform *basis functions*:

$$\mathbf{c} = \mathbf{A}\mathbf{x}, \quad \mathbf{x} = \mathbf{B}\mathbf{c} \tag{1}$$

where $\mathbf{x}^T = [x_0, x_1, \dots, x_{N-1}]$ is input signal vectors, $\mathbf{c}^T = [c_0, c_1, \dots, c_{N-1}]$ is transform coefficient vectors, \mathbf{A} forward transformation matrix, $\mathbf{B} = \mathbf{A}^{-1} = \mathbf{A}^T$ inverse transformation matrix, $\mathbf{b}_j = \{j \text{ columns of } \mathbf{B}\}$ are basis vectors. An efficient transformation should satisfy four conditions: (1) all transform coefficients ought to become almost statistically independent; (2) the energy of the image in the transform domain is compacted into a minimum number of coefficients which ought to be concentrated in a minimum frequency transform region; (3) a perfect reconstruction is required in the absence of quantization, and (4) one fast computational algorithm is needed for efficient transform implementation [2].

In *block transform coding*, an input image is subdivided into blocks of $N \times N$ pixels; each block is projected onto particular set of bases by means of an orthogonal transform \mathbf{A} ; and the coefficients of the transformation are quantized, entropy coded and transmitted. At the receiver, the coefficients are reconstructed by decoding the transmitted bit stream, and the inverse transformation is applied, so that an *approximation* to the original block is obtained (Fig.1). The discrete cosine transform (DCT) is most widely employed, because it is, for a wide class of highly correlated images, a good approximation to the statistically optimal Karhunen-Loeve transform (KLT). Recently, in order to overcome the problem of visible block boundaries, a new class of *lapped orthogonal transform* (LOT) is used. In the LOT transform (Fig.2a), each block of the size $N=2 \times M$ is mapped into a set of M basis function, each one being long $2 \times M$ [5]:

$$\mathbf{A}_b = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \mathbf{P}_1 & \mathbf{P}_0 & & \\ & \mathbf{P}_1 & \mathbf{P}_0 & \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}, \quad \mathbf{P} = [\mathbf{P}_0 \ \mathbf{P}_1 \ \dots \ \mathbf{P}_{L-1}], \quad \sum \mathbf{P}_m \mathbf{P}_{m+l}^T = \delta(l) \mathbf{I}, \quad 0 \leq l \leq L-1 \tag{2}$$

The equivalent uniform M -band perfect reconstruction filter-bank has impulse responses $h_k(n)$ equal to k -th row of matrix \mathbf{P}^T (Fig.2b).

The *subband image coding* is based on the decomposition of the input image into M narrow subband signals, where each subband is then maximally/critically decimated by M and coded separately. Pairs of filters-banks for decomposition (analysis filters-bank $\mathbf{H}(z)$) and reconstruction (synthesis filters-bank $\mathbf{G}(z)$) designed in such way that aliasing caused by decimation is canceled by the reconstruction filters and the *perfect reconstruction* is obtained $\mathbf{H}(z)\mathbf{G}(z) = \mathbf{I}$, where \mathbf{I} is the identity matrix. Subband decomposition can be made in both uniform and non-uniform size subbands. Non-uniform size subbands are usually

octave-bands (the size of subbands doubles as the frequency increases). Octave-subbands (recursively filtering low-pass subband by perfect reconstruction M=2 band banks with regular (smooth) filters) can be considered as coefficients of dyadic discrete *Wavelet transform* (DWT Fig.3). Wavelets have an advantage to adjust length of the basis function, which enables better perceptual image quality. Also, wavelets lead to the multi-resolution decomposition of an image, which is desirable property in progressive image transmission.

It is very important (from theoretical and practical view point) that the transform coder can be seen as a special case of subband coder. The direct transformation is equivalent to an analysis FIR filter-bank when decimation factor, number of subbands and filter lengths are the same N. Further, transform coefficients corresponding to a certain frequency component, taken one coefficient from each transformed block, constitute a subband (Fig.1b). Generally, *para-unitary* filter-banks $\mathbf{H}_p(z) \mathbf{G}_p(z) = \mathbf{I}$, $\mathbf{G}_p(z) = \mathbf{H}_p^{-1}(z) = \mathbf{H}_p^T(z^{-1})$ ($\mathbf{H}_p, \mathbf{G}_p$ are polyphase

transfer matrices: $H_k(z) = \sum_{l=0}^{M-1} z^{-l} E_{k,l}(z^M)$, $e_{k,l}(n) = h_k(nM+l)$) have orthonormal impulse responses $h(n)$ and generate a generalized LOT bases: the reconstruction filter (synthesis-inverse transform) are the same as the analysis filters (with time reversal) $g_k(n) = h_k(N-n)$.

If impulse responses are also symmetrical, resulting filter-bank has a linear phase, which is important property in image coding. However, in important case of M=2, the linear phase paraunitary filter-bank doesn't exist. The solution are perfect reconstruction filter-bank with nonidentical analysis/synthesis filters. Their impulse responses have different length and generate *biorthogonal* bases (orthogonality to dual bases) [5].

3. Experimental study

In order to study advanced transform coders, software implementations of IRLS image compression system were performed by computer. Our coder consists of several components: 1) perfect reconstruction analysis/synthesis filter-banks, 2) bit-allocation strategies based on variance of subbands, 3) uniform scalar quantizer, and 4) sequential baseline and Huffman entropy coding (Fig.1). As *objective* measure of decoded image

quality we used mean square error $MSE = \frac{1}{IJ} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} |x(i,j) - \hat{x}(i,j)|^2$, maximum absolute

error $MAE = \max |x(i,j) - \hat{x}(i,j)|$ and peak signal to noise ratio $pSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right)$ (where

$x(i,j)$ is an original, $\hat{x}(i,j)$ is reconstructed and $x(i,j) - \hat{x}(i,j)$ is error image. The *Subjective* criteria we used consist in categorization of the reconstructed images according to some predetermined scale based on overall impression (5-grade quality scale: bad, poor, fair, good, excellent) and on the reproduction of some significant details (impairments scale: blocking, ringing, blurring, texture loss, details loss). Two types of IRLS test-images (512x512 pixels, 256 gray levels) acquired during night from low altitude aircraft at 1500/3000 feet, are used in our experiments: one is "City Area", full of details of various size and contrast, and the other one is "Fields", representing a uniform-like terrain.

In a comparative study we used standard DCT block transform and Generalized LOT optimized linear phase paraunitary filter-bank of order N=M=8 and overlapping factor L (8 bands is a good compromise between the implementation complexity and performance). GenLOT is implemented in modular lattice structure, including DCT fast algorithm type II. In comparison of DCT L=1, LOT L=2, GenLOT L=3,6; we obtained the highest pSNR coding gain (decreasing from maximum of $pSNR_{diff} = 0.35dB$ at 0.2 bpp) over DCT by GenLOT L=3 (Fig.4a). In Discrete Wavelet Transform, we used biorthogonal Doubechies wavelets. Two different types of symmetric 2-band wavelets derived from the spectral factorization of

a maximally-flat filter are used: 1) DWT(6,10) the analysis filters have odd order (even length); the symmetric polarities low- and high-pass filters are symmetric and anti-symmetric, and 2) DWT(9,7), DWT(17,11) the analysis filters have even order (odd length), and both have symmetric impulse responses [5]. The highest pSNR gain (increasing from minimum at 0.5 bpp to maximum $pSNR_{diff} = 0.5dB$) is obtained with DWT(6,10) and DWT(17,11) wavelets in L=4 levels of multi-resolution decomposition (Fig.4b).

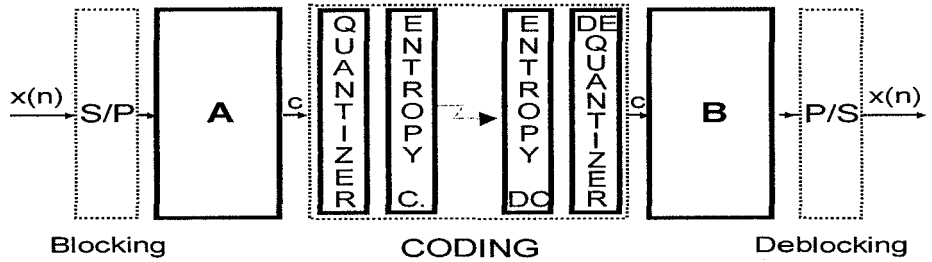
4. Conclusion

The most important experimental result in this paper is the proof that DWT based coder are more efficient than standard DCT and optimized GenLOT coders in objective (Fig.4c,d) and subjective quality (Fig.5) image measures for low and high bit-rates, in high demanded image compression of infrared line-scan data.

Considering two classes of IRLS test images, we obtained higher pSNR (about 5dB) for "Fields" class than "City Area" images. Further, taking into account various altitudes of image acquisition, we obtained higher pSNR (about 1.5dB) for high-altitude "City Area" images than low-altitude, and vice-versa on "Fields" images. These results give us possibility for research of terrain/altitude adaptive image coding. Further work can also be seen in additional performance improvements by spatial image segmentation along scan direction and adaptive transform coding.

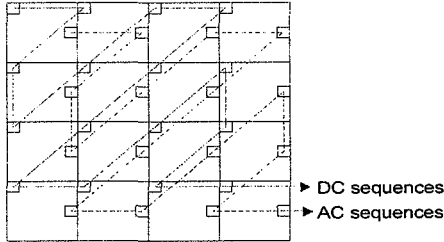
REFERENCES

- [1] MILOVANOVIĆ (D.), MARINČIĆ (A.), BARBARIĆ (Ž.) and PETROVIĆ (G.) - *Statistical analysis of computer generated thermal images based on overall modeling of line-scanning process*. Proc. QIRT 1994, Sorrento Italy, p.13-18.
- [2] MILOVANOVIĆ (D.), MARINČIĆ (A.), PETROVIĆ (G.) and BARBARIĆ (Ž.) - *Efficiency of Wavelets transform in infrared thermal image coding: Statistical approach*. Journal Facta Universitatis - Series electronics and energetics, Vol.10, No.1, 1997. Yugoslavia, p.153-169.
- [3] BARBARIĆ (Ž.), MARINČIĆ (A.), PETROVIĆ (G.) and MILOVANOVIĆ (D.) - *Thermal images generated by a line-scanning technique: Statistical properties*. Journal Applied Optics: Information Processing, Optical Society of America, 1994 Vol.33 No.20 p.4416-4419.
- [4] MILOVANOVIĆ (D.), WIĘCEK (B.), MARINČIĆ (A.), BARBARIĆ (Ž.) and PETROVIĆ (G.) - *Statistical analysis techniques for aerial infrared images in wavelets transform domain*. Proc. QIRT 1996, Stuttgart Germany, p.368-373.
- [5] STRANG (G.) and NGUYEN (T.) - *Wavelets and filters banks*, Cambridge Press, 1996.



a)

DCT subband decomposed images



b) Coefficients scanning prior entropy coding

DWT decomposed image $L = 3$

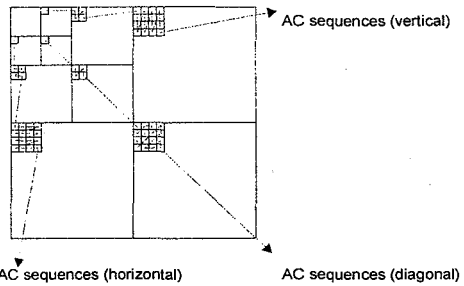


Fig. 1. The block scheme of transform image coder/decoder

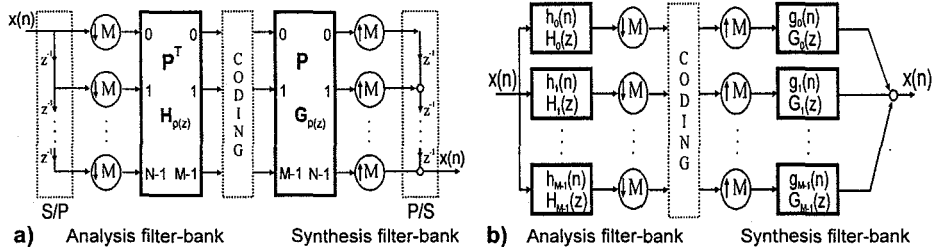
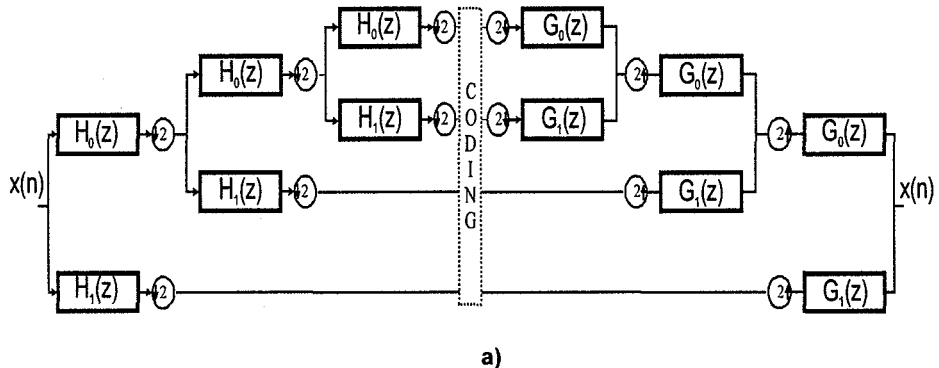


Fig. 2. a) Generalized LOT (paraunitary filter-bank) based coder, and b) equivalent uniform perfect reconstruction FIR filter-bank



a)

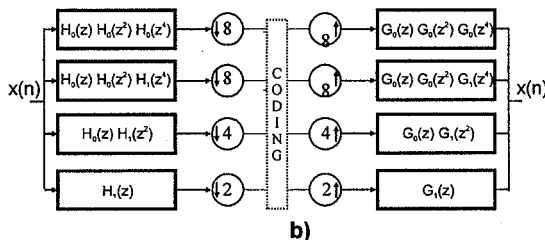
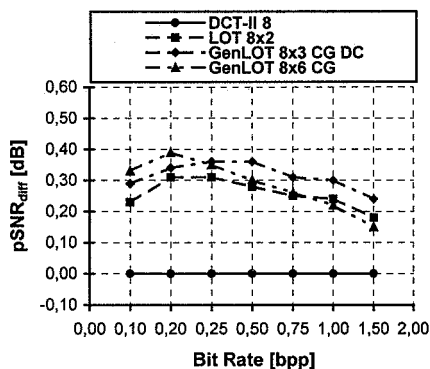
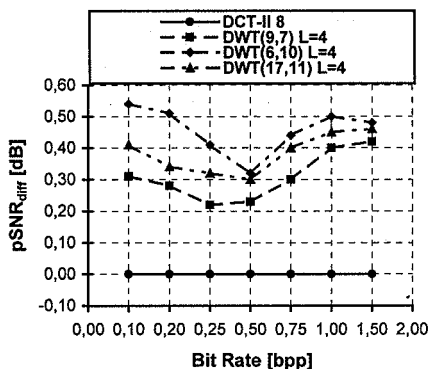


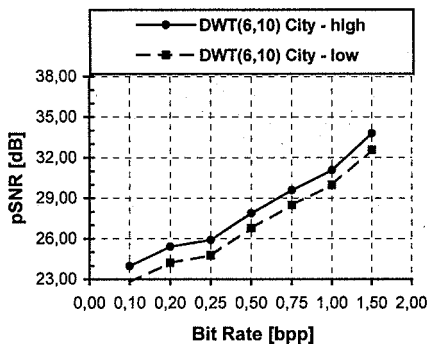
Fig. 3. a) Biorthogonal discrete Wavelet transform based coder (L=3 levels of recursive decomposition of low-pass band), and b) equivalent non-uniform perfect reconstruction FIR filter-bank



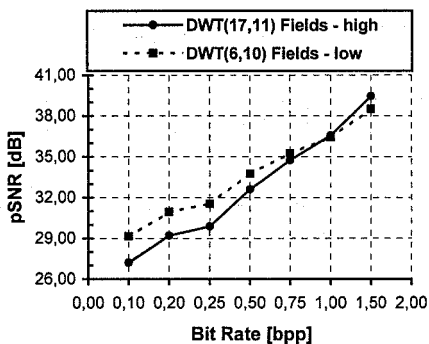
a) Test-image IRLS 200 ("City"-high altitude)



b) Test-image IRLS 200 ("City"-high altitude)

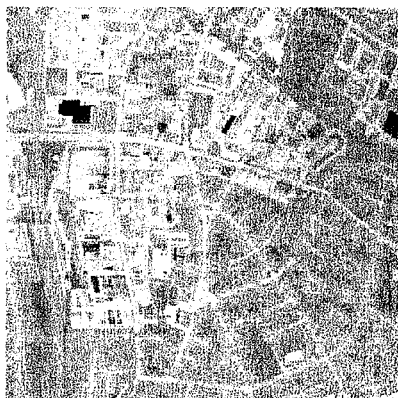


c) Test-images IRLS 200/100 ("City"-high/low altitude)



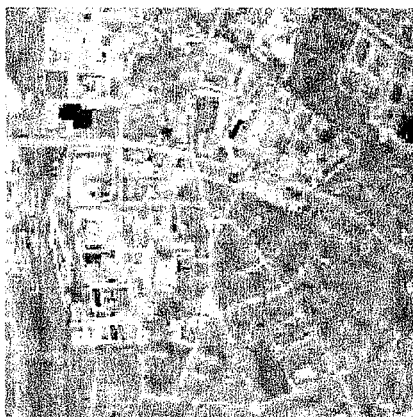
d) Test-images IRpLS 210/10 ("Fields" - high/low altitude)

Fig. 4. Experimental results: GeneralizedLOT and b) Biorthogonal DWT peak signal to noise ratio difference $pSNR_{diff}$ over DCT based coder; c) and d) $pSNR$ versus bit-rate of the best DWT coder for representative test images



Test image IRLS 200 ("City"-high altitude)

Test image IRLS 210 ("Fields"-high altitude)



Decoded image IRLS 200 DWT(6,10) L=4
coder: CR=32:1, BR=0.25bpp MSE=166.9,
MAE=82, PSNR=25.91dB

Decoded image IRLS 210 DWT(17,11) L=4
coder: CR=32:1, BR=0.25bpp MSE=67.2,
MAE=63, PSNR=29.86dB

Fig. 5. A comparative objective/subjective quality of IRLS test images (enlarged 2 times) in compression based on optimized discrete wavelet transform coder