

Efficient transform coding of IR line-scan images based on spatial adaptivity

by D.Milovanovic, A.Marincic, B.Wiecek^(*), G.Petrovic, and Z.Barbaric

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

Phone: +381 11 3370147 Fax: +381 11 3248681 Email: emilovan@etf.bg.ac.yu

() Technical University of Lodz, Institute of Electronics, Stefanowskiego 18/22, 90-924, Poland*

Abstract:

Based on our systematic studies of IRLS (InfraRed Line-Scanner) image statistical properties and comparative analyzes of state-of-the-art transform based image coders, we propose coding performance improvements by spatial image segmentation and operational rate-distortion optimization. We present in this paper, a new approach for an adaptive coding of IRLS images which combines the advantages of the following items: hierarchical tree-like image segmentation along-scan direction and Lagrange optimization of a trade-off rate-distortion $D(R)$. The proposed discrete optimization algorithm is based on fast systematic search for the best possible image segmentation scheme, together with the best possible bit allocation in coding of each segment independently. We implemented proposed optimization algorithm on computer and obtained significant signal-to-noise (PSNR) improvement by spatial adaptivity and independent discrete wavelet transform (DWT) coding of each image segment as compared over the standard discrete cosine transform (DCT) coding of nonsegmented IRLS test images.

1. Introduction

The thermal image generated by an infrared line-scanner (IRLS) consists of a series of signals corresponding to each transverse line-scan as the aircraft moves forward along track. Digital line-scanner generates enormous instantaneous bit rates which could saturate transmission link, the real-time display and human observer. Furthermore, the scanning of a thermal image involves a significant scanning area overlapping at great scanning angles, which does not exist in a standard video signal [1].

Our systematic studies of techniques for infrared image statistical data analysis, compression and efficient transmission, begun with construction of a new analytical/program model [2, 3]. The obtained results were used as the guidelines for statistical analyses of real test images [4]. In the study of advanced transform based image coding techniques, we proved theoretically that the efficiency of DWT is comparable with the standard DCT transform [5] and afterwards experimentally that biorthogonal Doubechies wavelet outperforms improved DCT based coders in objective and subjective measures of IRLS decoded image quality [6]. In this paper, we propose a way to achieve an improved performance of transform based coding by spatial image segmentation in order to exploit observed statistical variations along the line-scan direction.

2. Spatial adaptive IRLS image coding

The goal of efficient image coding is reduction of enormous amount of image data for storage or transmission, preserving quality of reconstructed image under constraint of maximum bit rates, computational complexity and delay. In an ideal image coder one would like high fidelity of the reconstructed image at the receiver in combination with low bit-rate and low complexity of the coder and the decoder. Unfortunately, these requirements

contradicts with each other, and good design is reduced to finding optimal trade-offs. We focuses on the basic trade-off between average bit rate (the R bit-rate in terms of bits per pixel) and average distortion D (measured by mean-squared error between reconstructed and original image). The optimal trade-off between rate R and distortion D can be precisely defined in several equivalent ways: by minimizing the average distortion D , subject to a constraint on the average rate R ; by minimizing the rate subject to a constraint on distortion, or by an unconstrained minimization of the Lagrangian cost function $J=D+\lambda R$, where the Lagrange multiplier λ , provides the weighting of the relative importance of distortion and bit-rate.

Unfortunately, classical rate-distortion information theory does not provide us with an explicit method for constructing a practical optimum coder and decoder. Rate-distortion theory provides unbeatable lower bounds to the obtainable average distortion for a fixed average rate, or vice versa. It also promises that codes exist that approach these bound when the code dimension and delay become large. Nevertheless, it can give very important hints about the properties of an optimum practical coder/decoder. Recent work has made progress in this direction by bridging at least in part the gap between theory and practice [7]. The idea is to use standard optimization procedures such as Lagrangian methods to find *local optimal operating points* in a rate-distortion sense, under some assumptions about the image.

Our adaptive coding by image segmentation can be stated as a Lagrangian optimization problem: each segment generates an operational rate-distortion curve, and optimal allocation of bit-rate between segments are standard. Among all possible segmentation, Lagrange optimization allows one to choose the winning segmentation, and this can be done in an efficient tree-pruning manner by using recursive dynamic programming [7].

When applying joint image segmentation and coding, we have to solve 2 basic problems:

1. What is the best segmentation scheme?
2. How many bits do we have to allocate to coding each image segment and for what criterion?

Because of the great amount of segments available at a segmentation level n , it is inefficient to test all the possible configurations. Instead, we decided to optimize a distortion-rate trade-off obtained by a Lagrange optimization method. In order to achieve this, it is necessary to compute *operational distortion-rate curves* applying transform coder with varying bit rates. The aim is to compare the directly computed $D(R)$ curve of a given segment, with the one obtained by minimization of a cost criterion applied to its 2 "children" resulting from the binary tree-like segmentation (Fig.1b). The solution leading to the lowest distortion value is then chosen. Calling $D_n(R_n)$ and R_n , respectively, the distortion and the bit rate of a segment at n -th decomposition level, one obtains the following criterion

$$J_{n-1} = \sum_{k=1}^2 D_{n,k}(R_{n,k}) + \lambda \left(R_{n-1} - \sum_{k=1}^2 R_{n,k} \right) \quad (1)$$

The classical result of minimization of (1) is deduced after equating the derivative of the criterion to zero as shown by equation (2)

$$\min J \Rightarrow \frac{\partial J_{n-1}}{\partial R_{n,k}} = 0 \Rightarrow \frac{\partial D_{n,k}(R_{n,k})}{\partial R_{n,k}} = -\lambda \quad \text{for } k = 1,2 \quad (2)$$

There exist of course an infinity of solution (2) as long as the derivatives ranges overlap each other, otherwise there will be no solution at all. Complete description of recursive dynamic programming algorithm can be found in [7], where it is used in different coding framework. In our work, the result is an adaptive segmentation of the original image following a systematic tree-like searching and pruning, under the constraint of a maximum bit rate.

3. Experimental results

In computer simulation we used transform-based coder and described optimization algorithm implemented in MatLab 5.2 for Windows 98. The coder has fixed structure: 1) 1D separable discrete orthogonal transform DCT/LOT/DWT, 2) bit-allocation strategies based on variance of subbands, 3) uniform scalar quantizer, and 4) sequential baseline and Huffman entropy coding. In comparative study we used standard DCT $N=8$ block transform, improved LOT of order $N=M=8$ (overlapping factor $L=3$) and biorthogonal Doubechies wavelets DWT(6,10) ($L=4$ levels of multi-resolution decomposition) [6]. As objective measure of decoded image quality we used mean square error MSE (PSNR [dB]) between original and reconstructed images. Two types of IRLS test-images (512x512 pixels, 256 gray levels) acquired during night from low/high altitude aircraft at 1500/3000 feet, are used in our experiments: one is "City Area" and the other one is "Fields". Test images are segmented hierarchically (binary tree with $N=3$ levels of decomposition gives maximum 8 segments of equal size) in along-scan direction. Operational rate-distortion $D(R)$ curves of each segment are computed applying transform coder/decoder with varying bit rates in the range 0.2-1 bpp (compression ratio 32-8) and measuring PSNR of decoded image. At this point, we have all information to start with the optimization of coding.

For each type of transformation and chosen bit rate R (Table 1), we obtained by proposed algorithm optimum image segmentation and bit allocation (Figure 1d) which gives minimum distortion D (maximum PSNR). Using this information, we adaptively coded/decoded (Figure 1 c) test image (Figure 1 a) and measured PSNR. Table 1 shows measured PSNR gains over the standard DCT coding of nonsegmented image. For given test image "City Area – high altitude", it can be observed that the proposed method for adaptive DWT(6,10) coding outperforms the standard LOT and DCT coding of nonsegmented test image for $PSNR_{gains}=0.6-1$ dB.

References

- [1] BARBARIĆ, Ž., MARINČIĆ, A., PETROVIĆ, G. and MILOVANOVIĆ, D., "Thermal images generated by a line-scanning technique: Statistical properties", *Journal Applied Optics: Information Processing*, OSA 1994, vol.33 no.20, p 4416-4419.
- [2] BARBARIĆ, Ž., MARINČIĆ, A., PETROVIĆ, G. and MILOVANOVIĆ, D., "Thermal-image generation by line-scanning technique: A new computer model", *Journal Applied Optics: Information Processing*, OSA 1994, vol. 33 no.14, p 2883-2890.
- [3] MILOVANOVIĆ, D., MARINČIĆ, A., BARBARIĆ, Ž. and PETROVIĆ, G., "Statistical analysis of computer generated thermal images based on overall modeling of line-scanning process", QIRT 1994, Sorrento Italy, p 13-18.
- [4] MILOVANOVIĆ, D., WIECEK, B., MARINČIĆ, A., BARBARIĆ, Ž. and PETROVIĆ, G. "Statistical analysis techniques for aerial infrared images in wavelets transform domain", QIRT 1996, Stuttgart Germany, p 368-373.
- [5] MILOVANOVIĆ, D., MARINČIĆ, A., PETROVIĆ, G. and BARBARIĆ, Ž., "Efficiency of Wavelets transform in infrared thermal image coding: Statistical approach", *Journal Facta Universitatis*, vol.10, no.1, 1997, p 153-169.
- [6] MILOVANOVIĆ, D., WIECEK, B., MARINČIĆ, A., PETROVIĆ, G. and BARBARIĆ, Ž., "A comparative study of advanced frequency-domain coding techniques in compression of infrared line-scan images", QIRT 1998, Lodz Poland, p 135-139.
- [7] RAMCHANDRAN, K. and VETTERLI, M., "Best wavelet packet bases in a rate-distortion sense", IEEE Trans. on Image processing, vol.2, no.2, 1993, p 160-175.

Bit-rate R	PSNR DCT8	PSNR _{gain} LOT 8x3	PSNR _{gain} DWT ⁴ (6,10)
--------------	-----------	------------------------------	--

[bpp]	[dB]	[dB]	[dB]
0.20	24.90	0.60	1.01
0.25	25.50	0.65	0.81
0.50	27.57	0.63	0.61
0.75	29.15	0.65	0.66
1.00	30.57	0.87	0.98

Table 1. PSNR gain of LOT/DWT spatial adaptive coding of IRLS "City-high" test image over DCT coding of nonsegmented image.

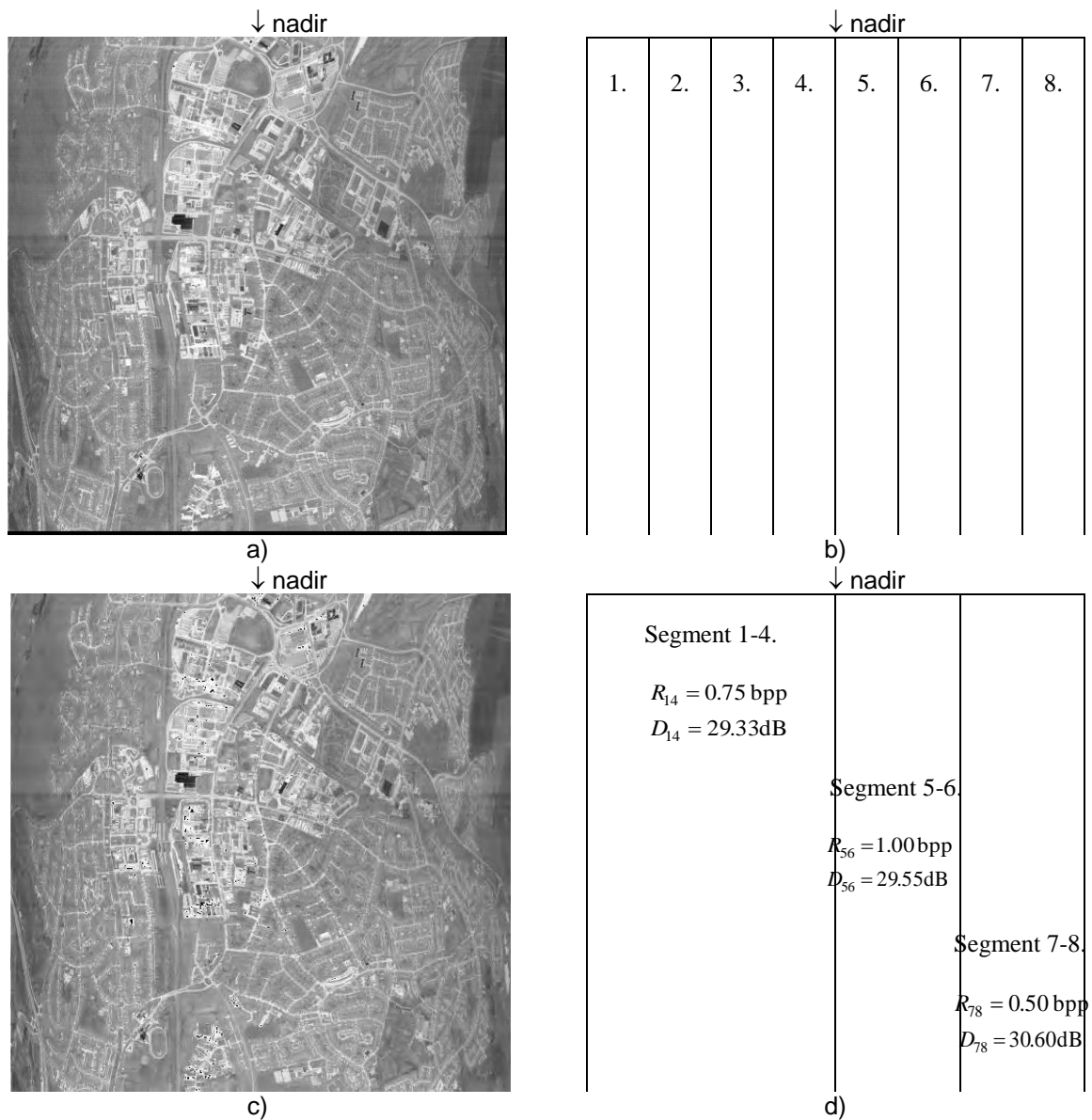


Figure 1. Transform coding of IRLS test image based on spatial adaptivity: a) "City-high" test image, b) N=3-levels binary segmentations (segments i=1-8), c) decoded test image ($DWT^4(6,10)$, $R=0.75$ bpp, PSNR=29.81dB) using proposed algorithm for optimal d) spatial segmentation and bit allocation R_i .