On Low Bit-Rate Coding Using the Contourlet Transform

Ramin Eslami and Hayder Radha

ECE Department, Michigan State University, East Lansing, Michigan, USA
Emails: {eslamira, radha}@msu.edu , Tel: 517-432-9958.

***Abstract-* In this work, we study and analyze the contourlet transform for low bit-rate image coding. This image-based geometrical transform has been recently introduced to efficiently represent images with a spars set of coefficients. In order to explore the potentiality of this new transform as a tool for image coding, we developed a direct coding scheme that is based on using non-linear approximation of images. We code the quantized transform coefficients as well as the significance map of an image in the contourlet transform domain. Based on the proposed approach, we analyzed the rate-distortion curves for a set of images and concluded that this coding approach, despite its redundancy, is visually competitive with a direct wavelet transform coder, and in particular, it is visually superior to wavelet coding for images with textures and oscillatory patterns.**

I. INTRODUCTION

Mallat and Falzon [5] showed that the performance of a transform-based image coder at low bit-rate mostly depends on its non-linear approximation capability. However, good non-linear approximation results do not necessarily imply efficient coding schemes [1]. The contourlet transform is a new geometrical image-based transform (Fig. 1), which is recently introduced by Do and Vetterli [2][3]. They showed that, when compared to wavelets, the contourlet transform better approximates images in which textures and oscillatory patterns have a significant presence. However, it is a redundant transform with a redundancy factor of 4/3 that would increase the rate for a given distortion. Fig. 2 illustrates the contourlet transform of the image *Boats* using 3 LP subband levels and 8 directions at the finest level. In this study, we design and simulate a low bit-rate image coding scheme based on the contourlet transform and analyze its performance for a set of images.

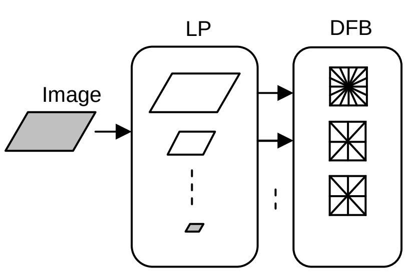


Fig. 1. A flow graph of the contourlet transform. The image is first decomposed into subbands by the Laplacian pyramid and then each detail image is analyzed by the directional filter banks.

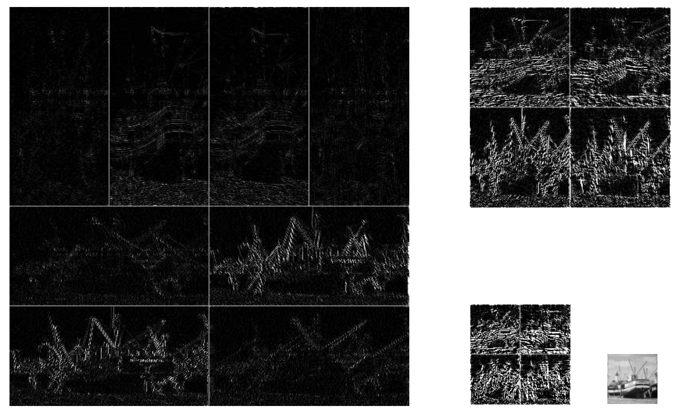
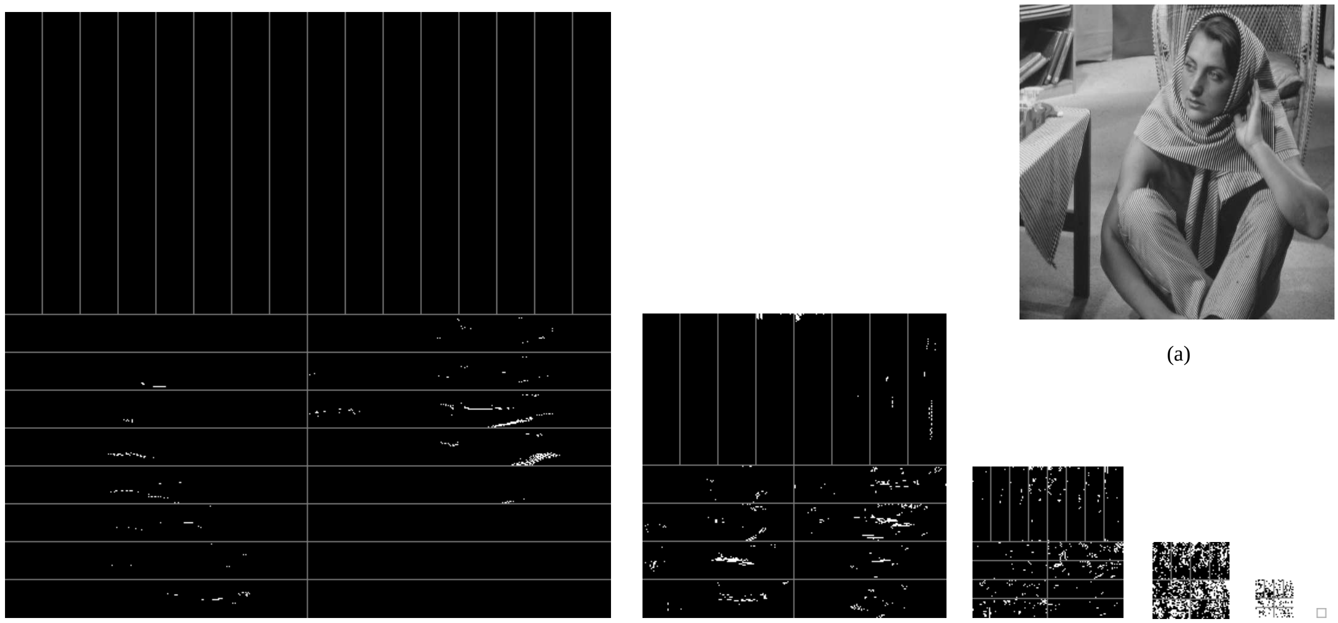


Fig. 2. The contourlet transform of the *Boats* image using 3 LP levels and 8 directions at the finest level. For better visualization, the transform coefficients are clipped between 0 and 15.

In Section 2, we describe the algorithm of the proposed approach. In Section 3, we analytically examine this method and some of the numerical results are presented, and finally Section 4, briefly summarizes the main results of this study.

II. METHOD

In non-linear approximation, one retains the k largest transform coefficients of a signal and set the rest to zero. If we inverse transform these retained coefficients, we obtain the best k-coefficient approximation of the signal for an orthogonal transform. This is the underlying idea for a transform coder based on a non-linear approximation. However, the positions of the retained (significant) coefficients have to be coded as well. In the proposed contourlet-based image coder approach, and following similar procedures used for the wavelet transform, we save the positions of the significant coefficients in a binary significance map, where a 1 accounts for a significant coefficient. So, the total bits required for coding an image is $b = b_0 + b_1$, where b_1 is the number of bits needed for coding the quantized significant coefficients, and b_0 is the number of bits required for coding the significance map. To efficiently code the significance map, we first scan it and then run-length code the resulted vector. Fig. 3 shows the significance map of the *Barbara* image with $k = 4096$. It is noteworthy that the



(a)

(b)

Fig. 3. *(a)* The *Barbara* image (in half of the original scale). *(b)* The significance map of the contourlet coefficients of the *Barbara* image using k = 4096 largest coefficients (the white pixels correspond to the significant coefficients).

significant coefficients are close together in position; hence, compared to the wavelet coefficients, that leads to a smaller rate to code the significance map. In our method, we scanned the significance map column-wise.

Fig. 4 depicts the normalized histograms of the contourlet and wavelet coefficients of the image set: *Barbara, GoldHill, Mandrill,* and *Peppers* all with the same size equal to 512x512. Similar to wavelets, most of the contourlet coefficients are around zero; therefore, for a low bit-rate coder, we used a (scalar) quantizer with a zero bin that is twice as large as the other bins [4]. The other important point is that the histograms of the contourlet coefficients have less variance compared to those of wavelets. However due to the redundancy of the contourlet transform, we require more bits for coding the significant coefficients.

The contourlet transform that we used in our experiments decomposes an image into 6 subbands using the 7-9 biorthogonal Daubechies wavelet transform, where each subband is fed to the directional filter banks stage with 32 directions at the finest level. In addition, we implemented a direct wavelet transform coder using the fully decimated 7-9 biorthogonal Daubechies wavelet transform for comparison.

III. ANALYSIS AND DISCUSSION

We analyze our scheme using a similar approach to the one proposed by Mallat and Falzon [5]. They derived the rate-distortion function for low bit-rate transform-based coding using the following formula:

$$D(R) = (1+K)D_0\left(\frac{R}{r_1 + r_0}\right), \qquad (1)$$

where $r_1 + r_0 = \frac{b_1}{k} + \frac{b_0}{k}$; $R = \frac{b_1 + b_0}{N} = \frac{b}{N}$ is the rate; $D_0(\frac{k}{N}) = \sum_{|a[m] \leq T|} |a[m]|^2$ is the non-linear approximation error that results from retaining the most k significant transform coefficients ($|a[m]| > T, \quad m = 1, ..., N$) out of the total N transform coefficients, and $K = \frac{2\gamma_k - 1}{12\theta^2}$. For an orthonormal transform $\{g_m\}_{0 \leq m < N}$,

$$D_0(k/N) = \|x - x_k\|^2, \qquad (2)$$

where x is the input signal: $x = \sum_{m=0}^{N-1} a[m]g_m$, $a[m] = \langle f, g_m \rangle$, and x_k is the approximation of x using k significant transform coefficients: $x_k = \sum_{|a[m]|>T} a[m]g_m$. Similar to the wavelet coder, since the ratio of the zero-bin to the quantization step is 1, $\theta = 1$ for the contourlet coder. Also we assume that $\gamma_k = 1$,

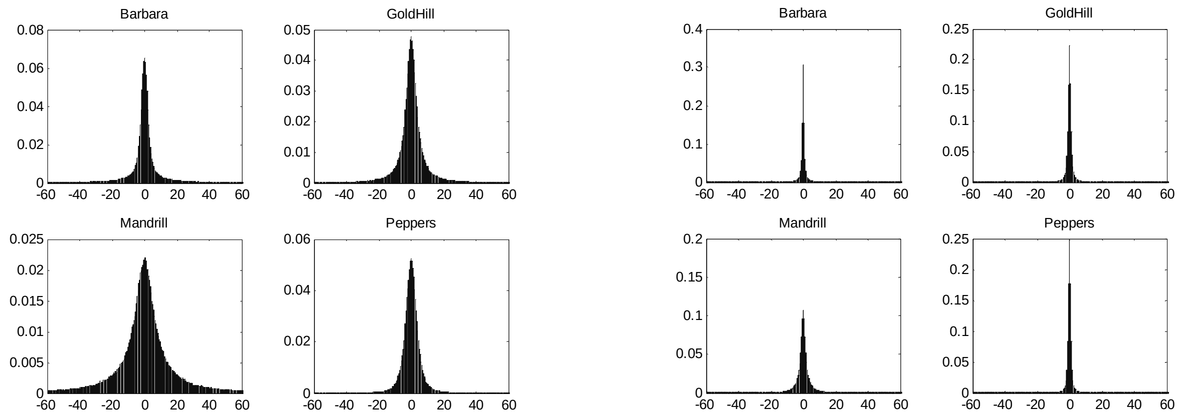


Fig. 4. The normalized histograms of the wavelet (*left images*) and contourlet coefficients (*right images*).

which yields a value of $K = \frac{1}{12}$. To estimate $r_1 + r_0$, we plot $\frac{b}{k}$ versus $\log_2\left(\frac{k}{N}\right)$. Fig. 5 shows the variation of $\frac{b}{k}$ for both the wavelet and contourlet coders. We see that the average $r_1 + r_0$ is about 6.5 for both of these coders. However since the contourlet transform is a redundant transform, (2) is not true and D_0 is more than that of an orthonormal transform. Therefore, and based on the proposed approach, analytically one expects lower performance for the contourlet coder when compared to wavelets. Fig. 6 depicts the PSNR curves of these transform coders for the tested image set.

We observe that in spite of the lower rate required to code the significance map in the contourlet coder (when compared to the wavelet coder), due to the redundancy of the contourlet transform, the contourlet coder provides rate-distortion curves that have lower performance (based on the PSNR measure) than those resulting from the wavelet coder. However, since the contourlet transform is more efficient in representing textures and contours when compared to wavelets, the images coded by this approach preserve textures and fine details found in natural scenes. Therefore, for the *Barbara* image that mostly consists of textures and oscillatory patterns, the contourlet coder provides a rate-distortion curve that is competitive to the one resulting form the wavelet coder in a range of the rate. Meanwhile, Fig. 7 shows an example of the *GoldHill* image. The textures in this image are significantly better approximated and retained by the contourlet transform coding scheme. This superior visual performance for the contourlet performance is not captured adequately by the PSNR rate-distortion measure.

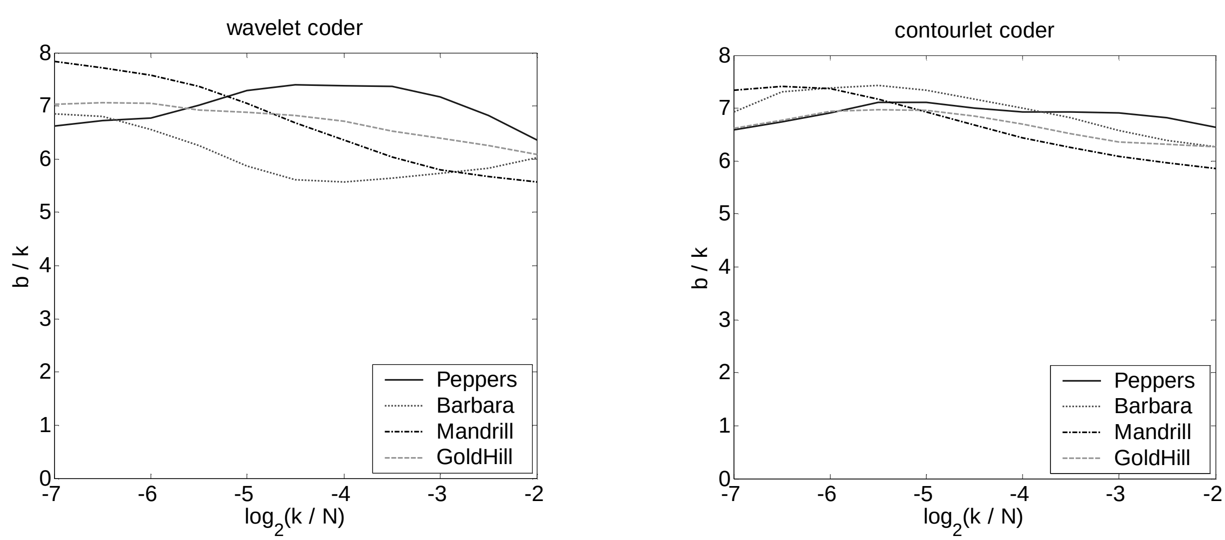


Fig. 5. $r_1 + r_0$ for the wavelet and contourlet transform coders for the image set.

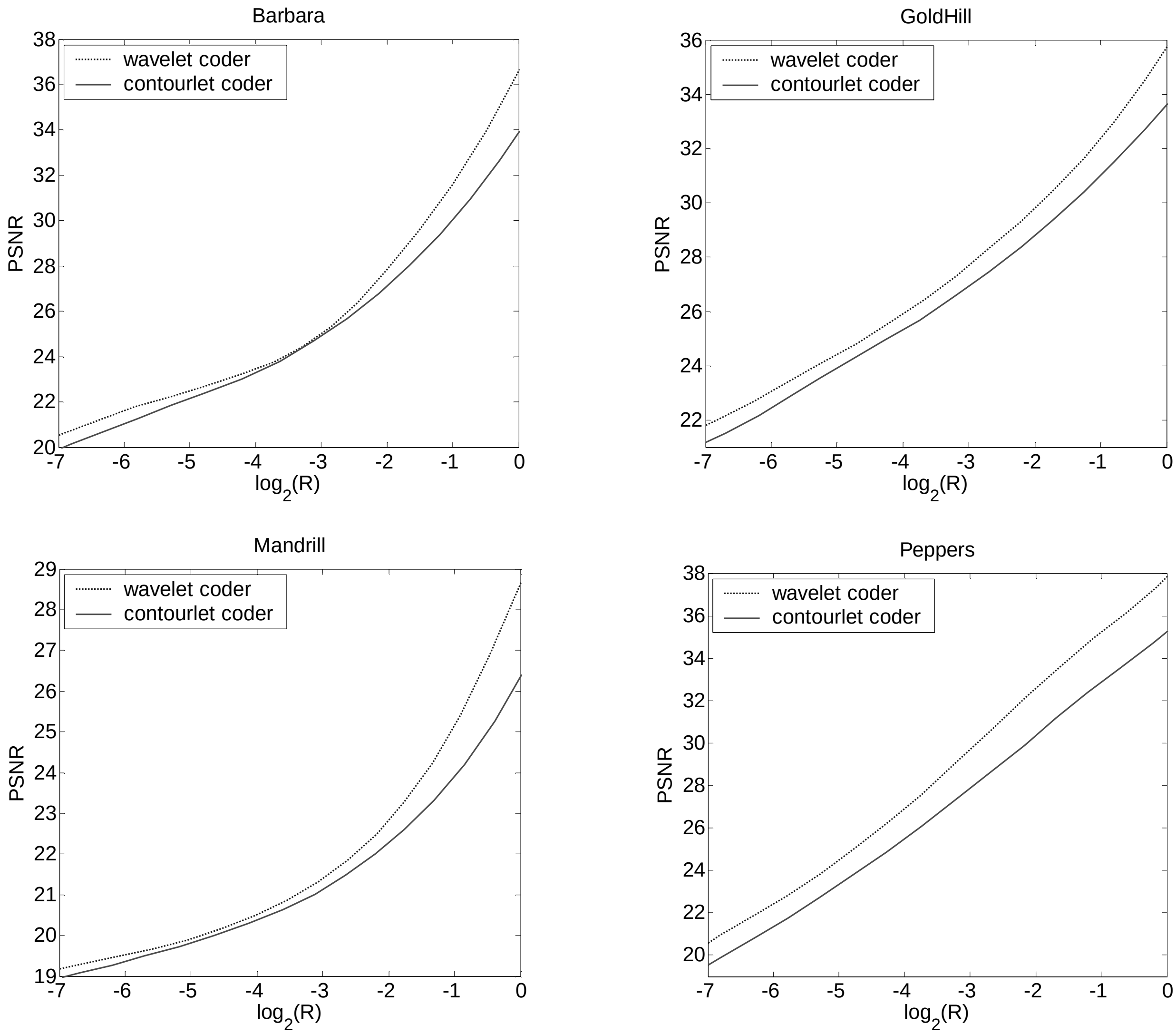


Fig. 6. The PSNR values vs. rate for the wavelet and contourlet coding schemes.

IV. CONCLUSIONS

In this study, we explored the capabilities of the new image-based transform, the contourlet transform, for image coding. We designed and analyzed a direct coding scheme and observed that although it provides a lower performance when compared to a direct wavelet transform coder in terms of the rate-distortion curves (due to the redundancy of the contourlet transform), the contourlet coder is capable of preserving and retaining textures and oscillatory patterns and achieves superior visual quality for images that contain mainly textures and fine details.

REFERENCES

[1] A. Cohen, I. Daubechies, O. Guleryuz, and M. Orchard, “On the Importance of Combining Wavelet-Based Non-linear Approximation with Coding Strategies,” *in IEEE trans. Information Theory*, vol. 48, no. 7, pp. 1895-1921, Jul. 2002.

[2] M. N. Do, *Directional multiresolution image representations.* PhD thesis, EPFL, Lausanne, Switzerland, Dec. 2001.

[3] M. N. Do and M. Vetterli, “Contourlets: A Directional Multiresolution Image Representation,” *in Proc. of IEEE Intl. Conf. on Image Processing*, Rochester, September 2002.

[4] S. Mallat, *A Wavelet Tour of Signal Processing.* Academic Press, 2nd Ed., 1998.

[5] S. Mallat and F. Falzon, “Analysis of low bit-rate image transform coding,” *IEEE trans. Signal Processing*, vol. 46, no. 4, pp. 1027-1042, Apr. 1998.

Fig. 7. The coded *GoldHill* image as well as a close-up at rate $R = 0.250$.