A SPARSE IMAGE REPRESENTATION USING CONTOURLETS *

Ahmed Nabil Belbachir and Peter Michael Goebel

Pattern Recognition and Image Processing Group

Vienna University of Technology, A-1040 Vienna, Austria

Tel.: +43158801-18366, Fax.: +43158801-18392

e-mail: nabil@prip.tuwien.ac.at, goe@prip.tuwien.ac.at

Abstract

The JPEG 2000 norm uses the bi-orthogonal wavelet transform for image decomposition. Wavelets are well adapted to point singularities (discontinuities), however they have a problem with orientation selectivity. Therefore, they do not represent two-dimensional singularities (e.g. smooth curves) effectively. This paper introduces and presents an evaluation of the contourlet transform for image compression, which has good approximation properties for smooth 2D functions. The contourlet finds a direct discrete-space construction and is therefore computationally efficient. A comparative study is performed between the contourlet and the wavelet analysis in terms of result quality and compaction of energy. Furthermore, we combine advantages of both transforms for image compression into a hybrid approach; the results are shown on 3 selected images out of 100, with a minimal resolution of 1 Mega pixels; the results are promising; finally, directions for future efficient transform coding are given.

1 Introduction

During the past two decades, image compression has developed from a mostly academic Rate-Distortion field [19], into a highly commercial business. Various lossless and lossy image coding techniques have been developed [17]. Since the compression ratio obtainable from lossy compression can significantly exceed that obtainable from lossless compression, the primary trade-off concerns the need for reproducibility versus the storage and transmission requirements. Lossy compression mainly consists of decorrelation and quantization stages that reduce the image size by permanently eliminating certain information. The decorrelation stage of the image compression algorithm is usually done by a transformation from one space to another to facilitate compaction of energy. One approach is the use of multiresolution transforms, which are free from blocking effect artifacts such as in case of the **D**iscrete **C**osine **T**ransform (DCT), which is used in the JPEG (baseline) industry standard [17].

*This work was supported by the Austrian Space Agency.

1.1 The Wavelet Transform

By the use of the **W**a**V**elet **T**ransform (WVT), the corresponding coefficients of the different decomposition levels are correlated and show characteristic trends. The residual correlation is indicative for further compression potential. Some standard methods (e.g. JPEG2000) get profits from this potential. Wavelet-based methods have expanded in the field of still image and video compression [17]; they offer the advantage of a better trade-off between complexity, compression and quality over the traditional DCT-based methods. However, for image compression, WVT has a problem with the orientation selectivity because it provides local frequency representation of image regions over a range of spatial scales, and therefore, it does not represent two-dimensional singularities effectively. In a map of the large wavelet coefficients, one sees the edges of the images repeated at scale after scale. This effect means that many wavelet coefficients are required to reconstruct the edges in an image properly, reducing the number of coefficients will introduce artifacts on the edges of the reconstructed image [7].

1.2 From Ridgelets to Curvelets and from Curvelets to Contourlets

The **R**id**G**elet **T**ransform (RGT) [5] was developed over several years in an attempt to break an inherent limit plaguing wavelet denoising of images. This limit arises from the frequently depicted fact that the two-dimensional (2D) WVT of images exhibits a large number of wavelet coefficients to represent image edges. A basic model for calculating ridgelet coefficients is to view ridgelet analysis as a form of wavelet analysis in the Radon domain. It has been shown in [8] that ridgelet representations solve the problem of sparse approximation of smooth objects with straight edges. In [7], an attempt has been made to use RGT for image compression.

However, in image processing, edges are typically curved rather than straight and ridgelets alone cannot yield efficient representation. But, if one uses a sufficient fine scale to capture curved edges, such edges get almost straight, therefore ridgelets are deployed in a localized manner. As a consequence the **C**ur**V**elet **T**ransform (CVT) [4] has been introduced. CVT is based on multiscale ridgelets combined with a spatial bandpass filtering operation; it was initially developed in the continuous-domain via multiscale filtering followed by a block RGT on each bandpass image. Later, the authors proposed the second-generation CVT [6] that was defined directly via frequency partitioning without using RGT. Both curvelet constructions require a rotation operation for the frequency decomposition, which ensures the construction in the continuous-domain. However, for discrete images, sampled on a rectangular grid, the discrete implementation of the curvelet transform is very challenging.

Therefore a new method was introduced: the **C**on**T**ourlet **T**ransform (CTT) [9]; initially described in the discrete-domain, the authors proved its convergence to an expansion in the continuous-domain. Thus, a discrete-domain multiresolution and multidirectional expansion is constructed, in the same way as wavelets are derived from filter banks, but using non-separable ones. Due to the fast-iterated filter bank algorithm, the construction results in a flexible multiresolution, local and directional image expansion using contour segments. However, CTT has the adverse property of showing other types of artifacts due to the discrete approach.

1.3 Organisation of the Paper

This paper investigates an assessment of the CTT for image compression, by a combination between CTT and WVT. In order to achieve improved image compression performance and since image size is growing up to the order of several million pixels, sparse representations are highly required, which considers anisotropy in the image. CTT can be applied efficiently to capture smooth contours at larger resolutions ($\geq$ 256x256 pixels-resolution), while WVT can be used at lower resolution images for further compaction of energy. The image quality is evaluated using this combined approach for decorrelation compared to that with wavelets, too. The potential of CTT for energy compaction on high resolution images is demonstrated. In Section 2 of this paper, CTT is summarized. Section 3 contains a comparative study between WVT and the CTT-based applied technique on 3 selected high resolution images. Finally, the obtained results are discussed and future directions for image compression are proposed.

2 The Contourlet Transform

Efficient representations of signals require that coefficients of functions, which represent the regions of interest, are sparse. Wavelets can pick up discontinuities of one dimensional piecewise smooth functions very efficiently and represent them as point discontinuities. 2D WVT obtained by a tensor product of one-dimensional wavelets are good to isolate discontinuities at edge points, but cannot recognize smoothness along contours. Numerous methods were developed to overcome this by adaptive [16], Radon-based [4], or filter bank-based techniques [13].

2.1 The Pyramidal Directional Filterbank

Do and Vetterli [13] proposed the **P**yramidal **D**irectional **F**ilter-**B**ank (PDFB), which overcomes the block-based approach of CVT by a directional filter bank, applied on the whole scale, also known as CTT. It has been developed to offer the directionality and anisotropy to image representation that are not provided by separable WVT. CTT is a multiscale and directional decomposition of a signal using a combination of a modified **L**aplacian **P**yramid (LP) [3, 11] and a **D**irectional **F**ilter **B**ank (DFB) [14]. Note, that PDFB allows for different number of di-

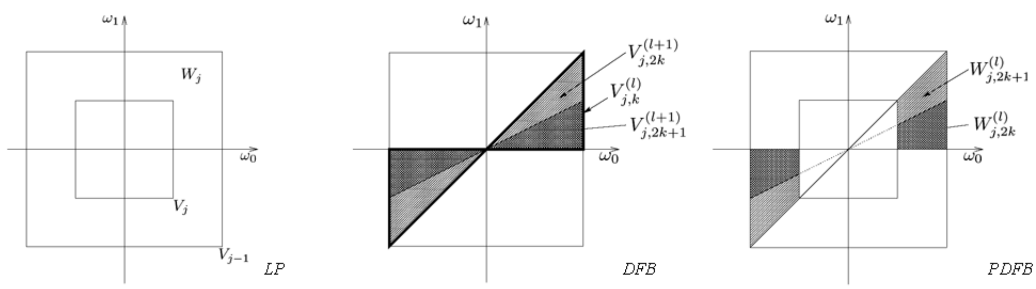

Figure 1: Generation of subspaces by the LP (left), DFB (middle) and PDFB (right)

rections at each scale/resolution to nearly achieve critical sampling [10]. As DFB is designed to capture high frequency components (representing directionality), the LP part of the PDFB permits subband decomposition to avoid ”leaking” of low frequencies into several directional subbands, thus directional information can be captured efficiently. The subspace splitting is illustrated in Figure 1.: V_j is a subspace, defined on a uniform grid with intervals 2^j x 2^j. The difference image in the LP carries the details necessary to increase the resolution from V_j to V_{j-1} on an image approximation; index k runs to all 2^l directions. As stated above, CTT offers the ability to choose the number of directions independently.

2.2 The Hybird Approach

We did experiments with over 100 high resolution images, and it was proved, that the smoothness of the contours within an image is coupled with the spatial resolution of a desired scale. We found, that beyond a spatial resolution around 2^8 pixels the application of the CTT carries no advantage compared with WVT in terms of compaction of energy. Thus, in this work WVT has been used, instead the CTT for low resolution decomposition levels (scales). Therefore, the adopted image decomposition for compression purposes is then:

1. Four decomposition levels for image size $\leq$ 1024 x 1024 pixels, where two scales are CTT, (with l=16 directions) and the remaining two are WVT (l=3). (See Figure 2).

2. Five decomposition levels for image size $\leq$ 2048 x 2048 pixels, thus, there are three CTT levels (with l=16 directions) and two WVT levels (l=3);

3. six decomposition levels for image size $\leq$ 4096 x 4096 pixels, thus, there are four CTT levels (with l=16 directions) and two WVT levels (l=3).

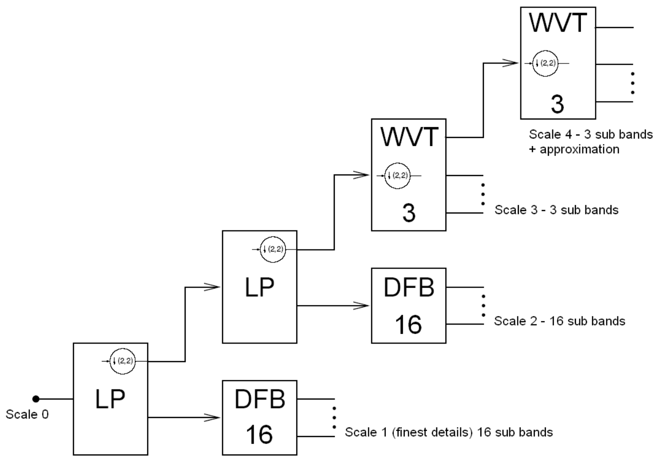

Figure 2: CTT with four decomposition levels, WVT is used for the two coarse levels.

In this work, images with size described from item 1. and 2. have been chosen for this investigation and the test results are reported in the next section.

3 Experiments

Compression results using WVT and CTT on 3 selected images (”Man”, ”Berry” and ”Art”), out of 100, are reported in this Section. The evaluation between these approaches is performed using the criterion PSNR and **P**otential **I**nformation **L**oss (PIL) [1]. PIL is a new metric, similar to the Kullback-Leibler distance [1], served for the evaluation of quality of the reconstructed image. It evaluates the probability density function of the image, so the histograms in Figure 3 of the original and reconstructed images are used for the calculation of PIL by means of the difference between both histograms (i.e. by computing the sum of the relative differences between all graylevel counts).

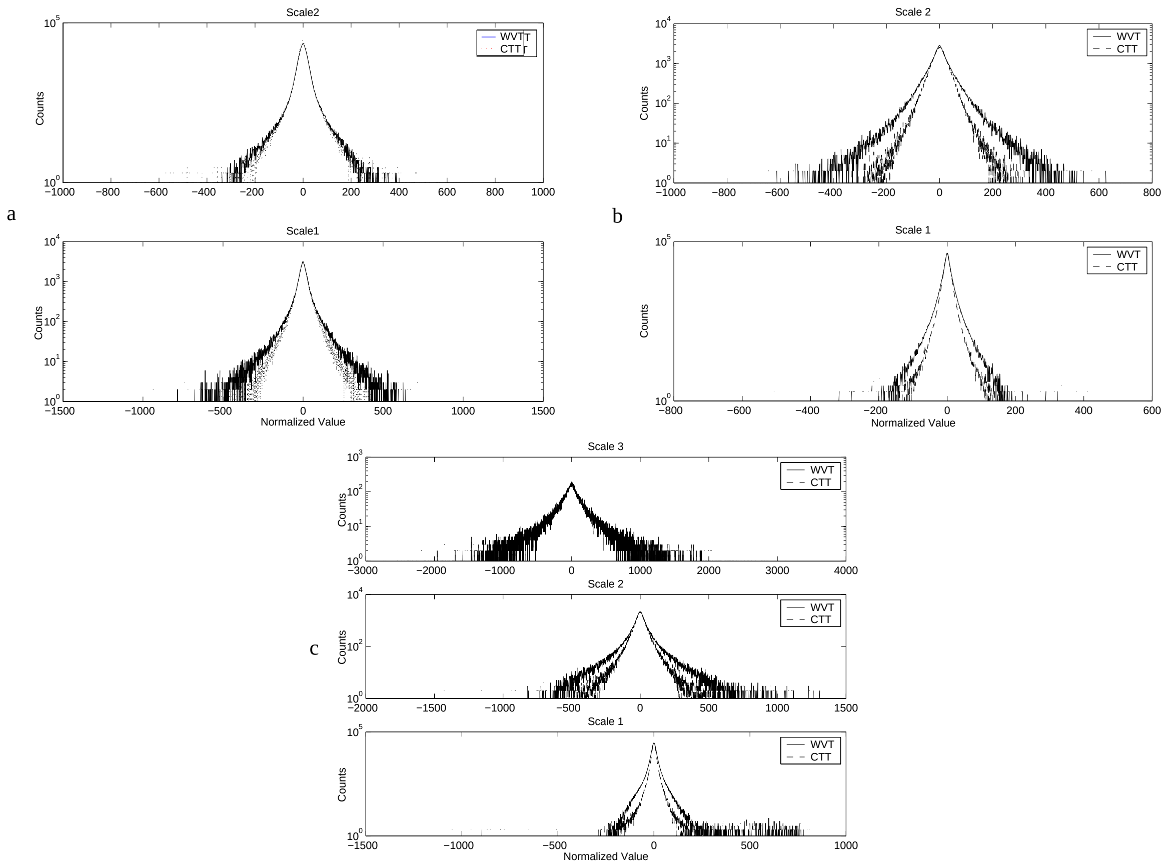

Figure 3: The histogram of the three images (a) ’Man’, (b) ’Berry’ and (c) ’Art’.

The original image is decomposed using CTT (see item 1 and 2 in Section 2) or WVT. Then, simple thresholding is performed over insignificant coefficients (with amplitude $<$ a threshold) for approximately the same PSNR, choosed accordingly to get acceptable image quality for both transforms. The image has been reconstructed from the remaining significant coefficients and the reconstruction error has been derived. The log energy entropy measures $-\sum p_i^2 \log(p_i^2)$ indicate the amount of compaction of energy for the desired transformation, where p_i is the probability of the transformed coefficients being in cell i of their measure space. An arithmetic coder [21] is used then to encode the remaining coefficients and the factor between the original image size and the encoded coefficients size is used to derive the compression ratio. Note that state-of-the-art coders (e.g. EBCOT contextual coding [12]) also exploit high order statistics. Thus, the used measures (entropy, number of coefficients) are an approximation that can only give a hint on the relative performance.

3.1 The Results

Experimental results are presented in Table 1. Using CTT, the number of maintained significant coefficients and the entropy is up to 20% smaller than that from WVT for an improved image quality of up to $1dB$ according to the PSNR and PIL. Therefore, the resulting compression ratio is slightly increased using CTT for an acceptable visual quality image.

Images	**Image Size**	**No. Maintained Coefficients**	**log energy entropy**	**PSNR dB**	**PIL**	**Compression Ratio**
”Man” by WVT	1k x 1k	48522	0.0277	24.5	0.62	2.8
”Man” by CTT	1k x 1k	43242	0.0234	25.5	0.42	3.2
”Berry” by WVT	1k x 1k	5869	0.0181	25.9	0.24	28.1
”Berry” by CTT	1k x 1k	4668	0.0163	26.4	0.20	29.3
”Art” by WVT	2k x 2k	16855	0.0101	23.0	0.32	31.4
”Art” by CTT	2k x 2k	15248	0.0079	24.1	0.30	33.1

Table 1: Number of maintained coefficients, entropy, PSNR, PIL and CR for the test images.

Figure 4 depicts the original selected high resolution images ”Man”, ”Berry” and ”Art”. Figure 5, 6 and 7 present the resulting reconstructed images. All *(a)* labeled sub-figures are the WVT results. The WVT artifacts (image quality) can be seen as zoomed regions at the *(b)* labeled sub-figures. The *(c)* labeled sub-figures concern the CTT results. The CTT image quality can be seen as zoomed regions *(d)*. Additionally to the improved energy compaction using CTT, it can be seen from the zoomed images (d) the visual quality preservation (better representation of edges) compared to WVT results. A reason for that can be noticed from Figure 3 (b), such that CTT provides a sparse representation of the images at finest scales and therefore, compacts them in fewer coefficients. However, at coarse scales, usually for $\leq$ 256 x 256 pixel-images, WVT can outperform CTT as object geometry is simplified and therefore, contours are not smooth anymore.

4 Conclusion and Perspectives

This paper investigates the potential of the **C**on**T**ourlet **T**ransform (CTT) for compression of high resolution images ($\geq$ 1 million pixels). The experimental results show a promising perspective as CTT provides a more compact representation of the information compared to that of the **W**a**V**elet **T**ransform (WVT) used in the compression standard JPEG2000. This is figured out by the histograms of the coefficients at Figure 3, which mostly shows the CTT inside the WVT curves, according to the entropy measures. Furthermore, CTT shows less information loss and artifacts on the reconstructed images after simple thresholding of the transformed coefficients. Indeed, WVT exhibits a large number of coefficients for representing smooth contours, which are usually present in high resolution images.

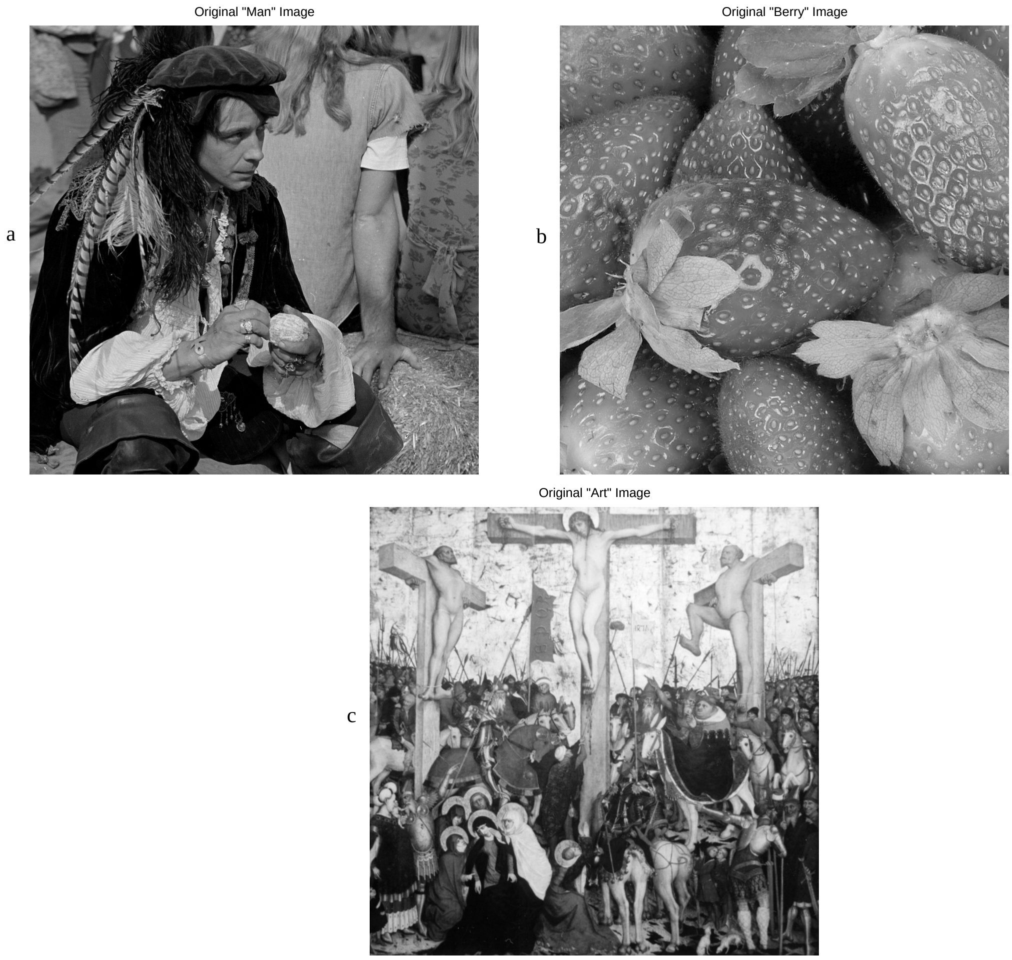

Figure 4: The original images: (a) ”Man”, (b) ”Berry” and (c) ”Art”

However, WVT compacts better the energy for coarse resolutions (downto around 2^8 pixel-images) as the geometry, usually, gets simplified (e.g. staircase effects). Therefore, a combination CTT-WVT (CTT for finest resolution and WVT for coarse resolution) seems to be good candidate for a new compression codec. Also a metric should be developed to find the objective best scale for the turnover from CTT to WVT processing. Investigations are still needed to determine the adequate entropy coder for the resulting coefficients. Indeed, the hierarchical structure exploited by the JPEG 2000 codec is not explicitly valid for our combination of methods. In a future work the combination of the coefficients from both transforms into zero trees representing the interscale dependencies will be studied.

Figure 5: Reconstructed ”Man” image; most of the cuttlings in d), are missing in b).

References

[1] A.N. Belbachir and H. Bischof. On-Board Data Compression: Noise and Complexity Related Aspects. *Technical Report Number 75, PRIP, TU Vienna, 2003*.

[2] A.N. Belbachir, H. Bischof and F. Kerschbaum. A Data Compression Concept for Space Applications. *DSP-SPE'00, IEEE Digital Signal Processing Workshop in Hunt, TX, USA, October 2000*.

[3] P. J. Burt, E. H. Adelson. The Laplacian pyramid as a compact image coder. *IEEE Trans. Commun., Vol.31(4):532-540, April 1983*.

[4] E.J. Candes and D.L. Donoho. Curvelets - A Surprisingliy Effective Non-Adaptive Representation for Objects with Edges. *in Curve and Surface Fitting, A. Cohen, C. Rabut, and L.L. Schumaker, Eds. Saint Malo: Vanderbilt University, 1999*.

[5] E. J. Candes and D. L. Donoho. Ridgelets: a Key to Higher-Dimensional Intermittency? *Phil. Trans. R. Soc. Lond. A., pp. 2495-2509, 1999*.

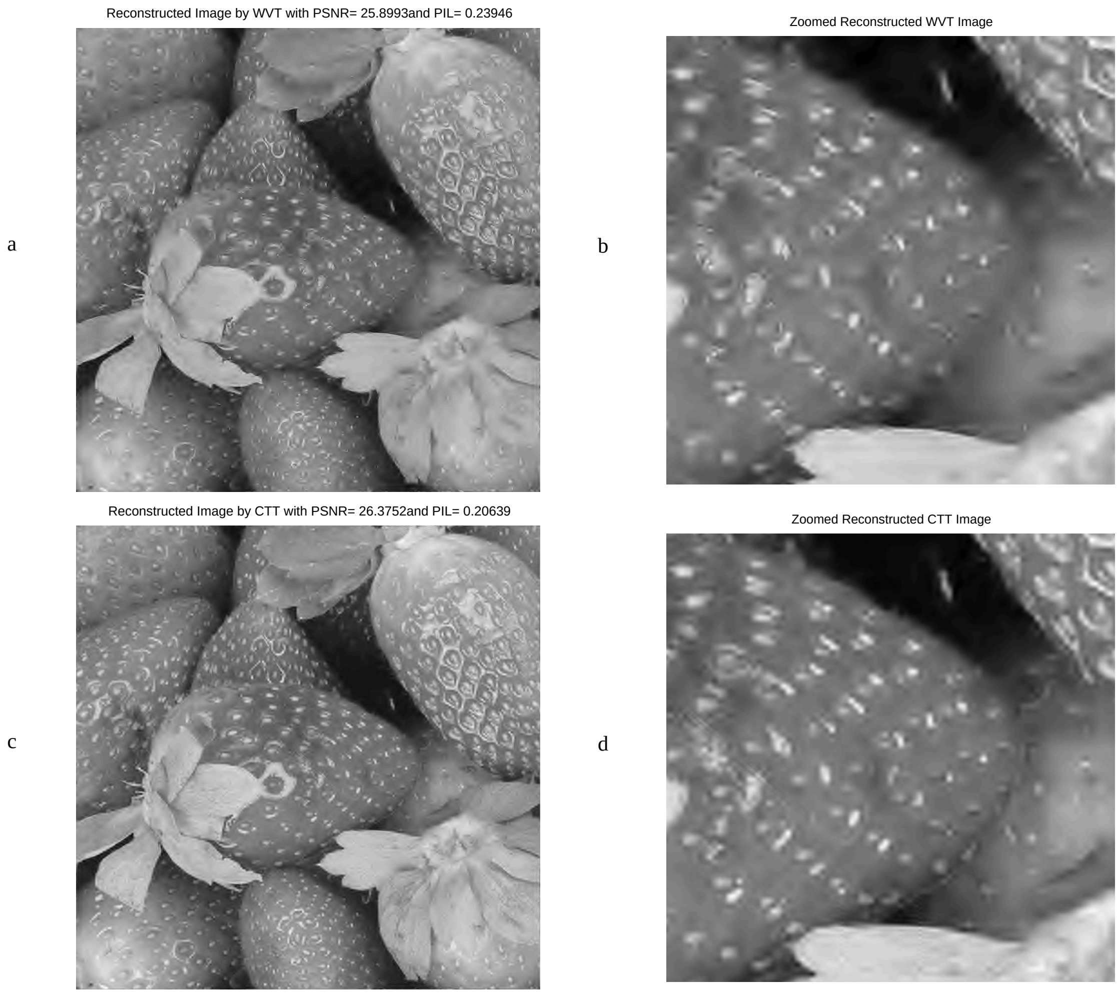

Figure 6: Reconstructed ”Berry” images; the berry in d) looks much sharper than that in b).

[6] E. J. Candes and D. L. Donoho. New Tight Frames of Curvelets and Optimal Representations of Objects with Smooth Singularities. *Department of Statistics, Stanford University, Tech. Rep., 2002*.

[7] M.N. Do and M. Vetterli. Orthonormal Finite Ridgelet Transform for Image Compression. *ICIP'2000, Vancouver,Canada, September 2000*.

[8] M.N. Do and M. Vetterli. The Finite Ridgelet Transform For Image Representation. *IEEE Transactions on Image Processing, Jan. 2003*.

[9] M.N. Do and M. Vetterli. The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. *IEEE Transactions on Image Processing*, Oct. 2003.

[10] Y. Lu and M. Do. CRISP-Contourlet: A Critically Sampled Directional Multiresolution Image Representation. *SPIE Conf. on Wavelets X*, San Diego, Aug. 2003.

[11] M. Do and M. Vetterli. Framing Pyramids. *IEEE Trans. on Signal Processing, VOL. 51, NO. 9, September 2003.*

[12] D. Taubman. High Performance Scalable Image Compression with EBCOT. *IEEE Trans. on Image Processing, VOL. 9, NO. 7, July 2000.*

[13] M. Do. Directional Multiresolution Image Representations. *Ph.D. Thesis, Department of Communication Systems, Swiss Federal Institute of Technology Lausanne, November 2001.*

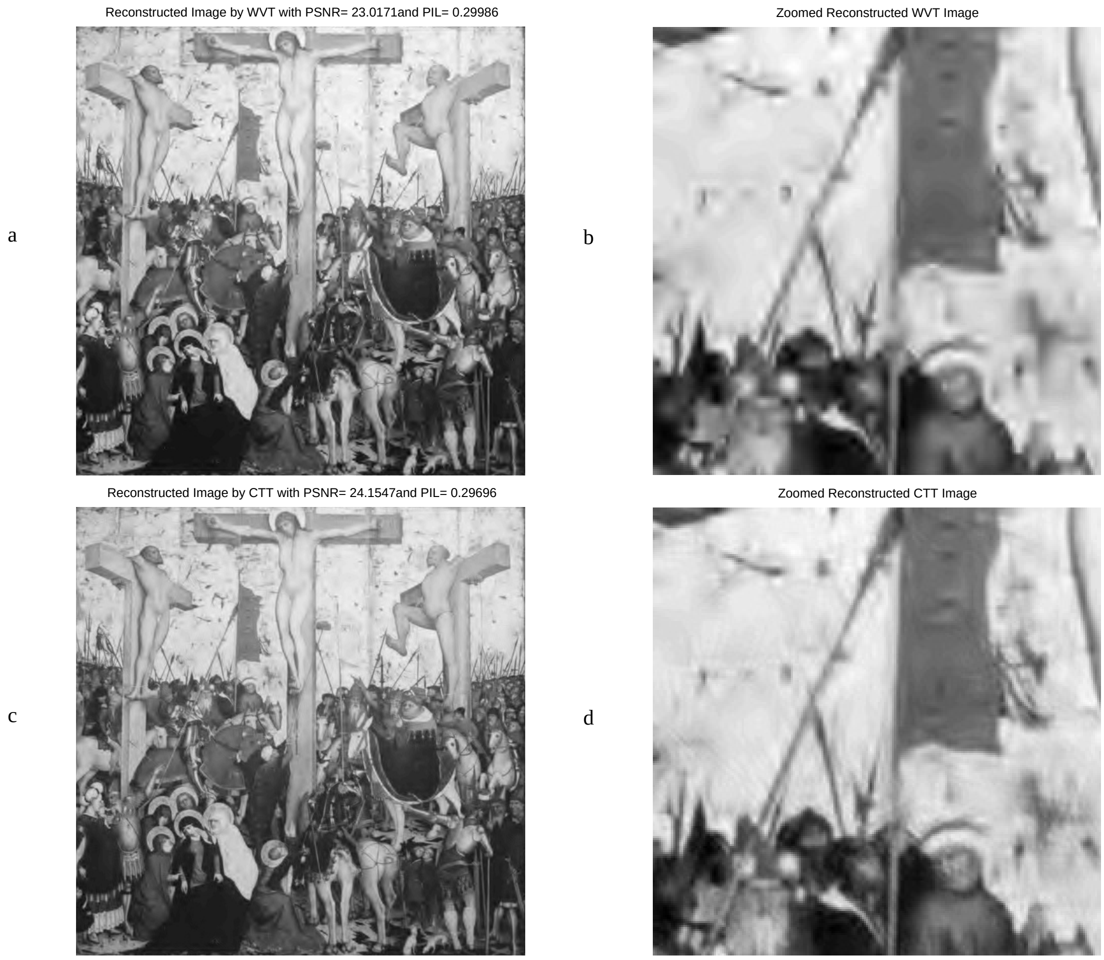

Figure 7: Reconstructed ”Art” images, notice the man’s head near the lower right corner.

[14] M. Do and M. Vetterli. Contourlets. *in: J. Stoeckler, G. V. Welland (Eds.), Beyond Wavelets, pp.1-27., Academic Press, 2002.*

[15] D.L. Donoho and M. Duncan. Digital Curvelet Transform: Strategy, Implementation, and Experiments. *in Proc. Aerosense 2000, Wavelet Applications VII, SPIE, 4056, 2000*.

[16] E.L. Pennec and S. Mallat. Image Compression with Geometric Wavelets. *IEEE International Conference on Image Processing, 2000.*

[17] K. Sayood. Introduction to Data Compression. *Second Edition, Morgan Kaufmann, 2000*.

[18] J.L. Starck, D.L. Donoho and E.J. Candes. Astronomical Image Representation by the Curvelet Transform. *Astronomy and Astrophysics 398, 785-800, 2003*.

[19] E. Tuncel and K. Rose. Computation and Analysis of the N-Layer Scalable Rate-Distortion Function. *IEEE Trans. on Information Theory, Vol.49, No.5, May 2003*.

[20] Z. Wang and C. Bovik. A Universal Image Quality Index. *IEEE Signal Processing Letters Vol. 9, no.3, pp.81-84, March 2002*.

[21] A. Moffat, R. Neal, I. Witten. Arithmetic coding revisited. *ACM Transaction on Information Systems, vol. 16, no. 3, pp. 256-294, 1998*.