STILL IMAGE CODING USING A NEW TRANSFORM

Aldo Maalouf, and Mohamed Chaker Larabi

XLIM-SIC Laboratory, UMR CNRS 6172, University of Poitiers,
SP2MI-2 Bd Marie et Pierre Curie, PO Box 30179
86962 Futuroscope Chasseneuil, France
{maalouf, larabi}@sic.sp2mi.univ-poitiers.fr

ABSTRACT

In this paper, we propose an image coding scheme based on a wavelet-like transform derived from orthogonal polynomial basis. First, 2D non-separable wavelet functions are derived from a set of bivariate orthogonal polynomials. Then, a wavelet-like transform coding scheme using the proposed wavelet functions is proposed. The motivation behind using orthogonal polynomials is that they exhibit some properties related to the human visual system (HVS) [1]. After applying the proposed transform, the obtained coefficients are threshold coded using quantization and bit allocation as in JPEG baseline system. The performance of the proposed transform coding method is reported. The proposed coding scheme is also compared with other transform coding methods such as JPEG, JPEG2000 and JPEG-XR/HDPHOTO.

1. INTRODUCTION

Digital imaging has had an enormous impact on industrial applications and scientific projects. It is no surprise that image compression has been a subject of great commercial interest. In addition to being a topic of practical importance, the problems studied in image compression are also of considerable theoretical interest. The problems draw upon and have inspired work in information theory, applied harmonic analysis, and signal processing.
Efforts in this field of research can be categorized in two ways: lossless vs. lossy compression and predictive vs. transform coding.

1.1 lossless vs. lossy compression

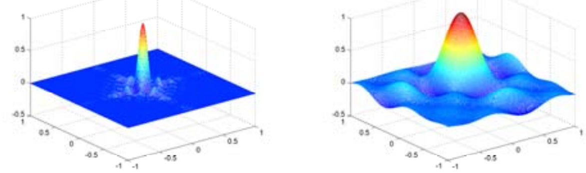
In lossless compression schemes, the reconstructed image, after compression, is numerically identical to the original image. However lossless compression can only achieve a modest amount of compression. An image reconstructed following lossy compression contains degradation relative to the original. Often this is because the compression scheme completely discards redundant information. However, lossy schemes are capable of achieving much higher compression.

1.2 predictive vs. transform coding

In predictive coding, information already sent or available is used to predict future values, and the difference is coded. Since this is done in the image or spatial domain, it is relatively simple to implement and is readily adapted to local image characteristics. On the other hand, transform coding first transforms the image from its spatial domain representation to a different type of representation using some well-known transform and then codes the transformed values (coefficients). This method provides greater data compression compared to predictive methods, although at the expense of greater computation.
A typical transform coding system consists of three closely connected components namely (a) Source Encoder, (b) Quantizer and (c) Entropy Encoder. Compression is achieved by applying a linear transform in order to decorrelate the image data, quantizing the resulting transform coefficients and entropy coding the quantized values.
A variety of linear transforms have been developed to this end, which include Hadamard Transform [17], Karhunan-Loeve Transformation [13], Discrete Cosine Transformation (DCT) [11], Wavelet Transformation [5] [9] and, recently, the lapped biorthogonal transform [14]. Each of these transforms has its own advantages and disadvantages but the DCT stands out to be the best and has been adopted in the JPEG still image compression standard [6].
However, the DCT requires a high computational complexity as it involves floating point operations. For that, there have been efforts work for finding simpler transforms[8]. Here we observe the work of Richardson [7] who defined the (4×4) DCT block transform that has become the core of JPEG which is currently the world wide standard for compression of digital images.
Nevertheless, DCT based codecs suffers from two major drawbacks: the decay rate of DCT coefficients is too slow, and generates infamous blocking artifacts. To cope these problems, the newest version of JPEG, the JPEG-2000 [2], adopts wavelets for coding standard.
Motivated by the fact that transformations play significant role in image data compression, a new bivariate non-separable orthogonal polynomials based transform coding technique for images is presented in this paper. The motivation behind using non-separable transform is that separable image representations have restrictions on maximizing coding gain because they have less freedom of parameters than non-separable ones [10] [16]. The important steps involved in the proposed coding work are: first, a class of wavelet functions obtained from bivariate orthogonal polynomials basis is proposed. Then, from the proposed wavelet functions, an image representation is proposed. After the proposed transformation, the resulting coefficients are threshold coded with a scalar quantization procedure and entropy coded with VLC tables as in JPEG.
It is to be noted that, recently, an approach for image coding that is based on orthogonal polynomials has been proposed in [15]. In fact, the authors of [15] presented a polynomial operator which is used to define a new transform coding approach. This polynomial operator is constructed from 1D orthogonal polynomials and is used to define an image transform. Our approach differs from this latter in that we do not define a polynomial operator. We rather have used bivariate orthogonal polynomials to construct 2D wavelet functions and to define a multiresolution wavelet-like image transform.
The rest of the paper is organized as follows: in section 2 a description of the wavelet functions that are defined from bivariate orthogonal polynomial basis is given. In section 3 we present the image decomposition using these wavelet functions and the compression scheme. Section 4 is devoted for experimental results. Finally, section 5 draws some concluding remarques.

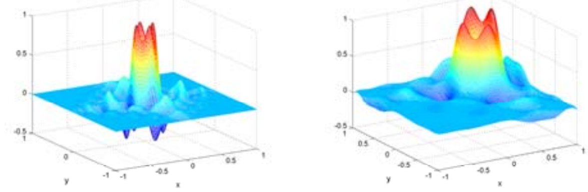
2. ORTHOGONAL POLYNOMIALS BASED WAVELET FUNCTIONS

In [1], Blaivas showed, by investigating retinal receptive fields, that orthogonal polynomials have certain properties that coincide with the HVS. Within the framework suggested by Blaivas, visual analysis in the retina can be regarded as a process of expansion in orthogonal polynomials basis. Motivated by this property we propose in this section a 2D non-separable wavelet function. The wavelet



(a) (b)

Figure 1: kernel functions using: (a) Legendre polynomial of degree 32, $x_0 = 0, y_0 = 0$, (b) Hermite polynomial of degree 32, $x_0 = 0.3, y_0 = 0.2$



(a) (b)

Figure 2: Legendre wavelet functions: (a)$\varphi_{32,32}(x,y,0,0)$ and (b) $\varphi_{16,16}(x,y,0.5,0.5)$

functions presented here are an extension of the wavelets proposed by Fischer et al. in [4] for the case of 2D. First we describe how to generate bivariate orthogonal polynomials and then we define the 2D wavelet functions.

2.1 Bivariate polynomials

A bivariate polynomial can be defined by

$$P_{K,L}(x,y) = \sum_{k=0}^{K}\sum_{l=0}^{L}\gamma_{k,l}(x)^k(y)^l \quad (1)$$

where $K+L$ is the polynomial degree. Then we consider the scalar product

$$\langle f_1, f_2\rangle = \int\int_{\Omega} f_1(x,y)\, f_2(x,y) w(x,y)\, dxdy \quad (2)$$

with w the weighting function and Ω the image domain.
Using the three terms recurrence, we can create an orthogonal polynomials basis by using an orthogonalization procedure:

$$\begin{cases} P_{-1,j}(x,y) = 0, \quad P_{i,-1}(x,y) = 0 \;, \quad P_{0,0}(x,y) = 1 \\ P_{i+1,j}(x,y) = (x-\lambda_{i+1,j})\, P_{i,j}(x,y) - \mu_{i+1,j}P_{i-1,j}(x,y) \\ P_{i,j+1}(x,y) = (y-\lambda_{i,j+1})\, P_{i,j}(x,y) - \mu_{i,j+1}P_{i,j-1}(x,y) \end{cases} \quad (3)$$

where λ and μ are two coefficients defined by

$$\begin{array}{ll} \lambda_{i+1,j} = \frac{\langle xP_{i,j},P_{i,j}\rangle}{\langle P_{i,j},P_{i,j}\rangle} & \lambda_{i,j+1} = \frac{\langle yP_{i,j},P_{i,j}\rangle}{\langle P_{i,j},P_{i,j}\rangle} \\ \mu_{i+1,j} = \frac{\langle P_{i,j},P_{i,j}\rangle}{\langle P_{i-1,j},P_{i-1,j}\rangle} & \lambda_{i,j+1} = \frac{\langle P_{i,j},P_{i,j}\rangle}{\langle P_{i,j-1},P_{i,j-1}\rangle} \end{array} \quad (4)$$

2.2 Wavelet functions

Let $V_{K,L}$ be the space formed by the set of orthogonal polynomials $P_{K,L}$, i.e. $V_{K,L} := span\{P_{0,0}, P_{0,1}, \ldots\ldots, P_{K,L}\}$. For a given fixed position in Ω: $(x_0,y_0)\in\Omega$, a kernel function $\Gamma_{k,l}(x_0,y_0)$ can be defined by using the Christoffel-Darboux formula:

$$\begin{aligned}\Gamma_{k,l}(x_0,y_0) &= \frac{2\left[P_{k,l}^T(x_0,y_0)A_{k+1,l}P_{k+1,l}(x,y) - P_{k+1,l}^T(x_0,y_0)A_{k+1,l}^T P_{k,l}(x,y)\right]}{(x-x_0)-(y-y_0)} \\ &= \sum_{k=0}^{K}\sum_{l=0}^{L} P_{k,l}^T(x_0,y_0)\, P_{k,l}(x,y)\end{aligned} \quad (5)$$

where $A_{k,l} = \langle P_{k-1,l}, P_{k,l}\rangle$.
Figure 1 shows some examples of kernel functions derived from Legendre polynomial basis (weight function $(w(x,y) = 1)$) and Hermite basis ($w = e^{(-x^2-y^2)}$). These basis are obtained by using equations (2), (3) and (4).

Equation (5) indicates that the kernel polynomials are localized around (x_0,y_0). Motivated by this property we define scaling functions as 2D kernel polynomials, i.e.

$$\varphi_{k,l}(x,y,x_0,y_0) = \Gamma_{k,l}(x_0,y_0) = \sum_{k=0}^{K}\sum_{l=0}^{L} P_{k,l}^T(x_0,y_0)\, P_{k,l}(x,y) \quad (6)$$

It is to be noted that $\varphi_{k,l}$'s form a basis for $V_{K,L}$. In fact it has been shown in [3] that kernel functions are the fundamental polynomial of Lagrange interpolation. As a result, they are linearly independent.
Let

$$W_{2K,2L} = V_{2K,2L} - V_{K,L} = \{P_{K+1,L+1}, P_{K+2,L+2}, \ldots, P_{2K,2L}\} \quad (7)$$

Then, a wavelet function can be defined as follows:

$$\psi_{K,L}(x,y,x_0,y_0) = \varphi_{2K,2L}(x,y,x_0,y_0) - \varphi_{K,L}(x,y,x_0,y_0) \quad (8)$$

Figure 2 shows some examples of kernel functions derived from Legendre polynomial basis.

We represent in the following section an image representation by using the bivariate polynomial based wavelet and scaling functions.

3. BIVARIATE POLYNOMIAL-BASED IMAGE TRANSFORM

The purpose of this section is to describe reconstruction and decomposition algorithms using functions defined in the previous section. The schemes are based on the space representation $V_{2K,2L} = V_{K,L} \oplus W_{K,L}$. A repeated application of this step would result in a multiresolution of a weighted L^2-space. To decompose an image $I(x,y)$ in $V_{2K,2L}$ one has first to approximate I by a suitable function $I_{2K,2L} \in V_{2K,2L}$ as follows

$$I_{2K,2L} = \sum_{k=0}^{K}\sum_{l=0}^{L} a_{2K,2L}\varphi_{2K,2L}(x,y) \quad (9)$$

where

$$a_{2K,2L} = \frac{\langle I, \varphi_{2K,2L}\rangle}{\langle \varphi_{2K,2L}, \varphi_{2K,2L}\rangle} \quad (10)$$

similarly, wavelet coefficients can be obtained by

$$b_{2K,2L} = \frac{\langle I, \psi_{2K,2L}\rangle}{\langle \psi_{2K,2L}, \psi_{2K,2L}\rangle} \quad (11)$$

In the next subsection, the inter-scale relationship is developed.

3.1 Inter-scale relationship

We should notice first that moving from one scale to another (from $V_{K,L}$ to $V_{2K,2L}$) is performed by adding the polynomials

$P_{K+1,L} \ldots P_{2K,L+1} \ldots P_{2K,2L}$ to the space $V_{K,L}$.
Since $\varphi_{2K,2L} = \varphi_{K,L} + \psi_{K,L}$, equation (9) can be written as:

$$a_{2K,2L} = \sum_{k=0}^{K} \sum_{l=0}^{L} a_{K,L} \varphi_{K,L} + \sum_{k=K+1}^{2K} \sum_{l=L+1}^{2L} b_{K,L} \psi_{K,L} \qquad (12)$$

Let the coefficients $a_{K,L}$ and $b_{K,L}$ be given, then the reconstruction can be obtained by using equation (12) and the decomposition can be obtained by

$$a_{K,L} = \frac{\langle a_{2K,2L}, \varphi_{K,L} \rangle}{\langle \varphi_{K,L}, \varphi_{K,L} \rangle}, \qquad b_{K,L} = \frac{\langle b_{2K,2L}, \psi_{K,L} \rangle}{\langle \psi_{K,L}, \psi_{K,L} \rangle} \qquad (13)$$

In the next subsection, some properties of this decomposition are presented.

3.2 Properties

In this subsection, we relate the transform presented previously to the classical concept of multiresolution analysis due to Mallat and Meyer. We should notice here that the major difference between the decomposition approach presented here and the classical wavelet transform is that, in our case, the scaling and wavelet functions are not the same as we move from one scale to another. Remember that moving from one scale to another is accomplished by doubling the order of the highest polynomial in the space V and then computing the kernel and the wavelet functions. However, some of the multiresolution properties can still be identified:
Property 1: if $f(x,y) \in V_{K,L}$ then $f(2x,2y) \in V_{2K,2L}$.
This in fact follow up from the property of the kernel function that are the fundamental polynomial of Lagrange interpolation. So dilating the kernel function by 2 enlarges the details by 2 and guarantees that it defines an approximation at the coarsest resolution.
Property 2: $V_{K,L} \subset V_{2K,2L}$.
This is evident since $V_{K,L} = \{P_{0,0}, P_{0,1}, \ldots, P_{1,0} \ldots P_{K,L}\}$ and $V_{2K,2L} = \{P_{0,0}, P_{0,1}, \ldots P_{n_1,n_2} \ldots\ldots, P_{2n_1,2n_2}\}$, then $V_{n_1,n_2} \subset V_{2n_1,2n_2}$.
Property 3: $\bigcup_{-\infty}^{+\infty} V_{n_1,n_2}$ is dense in L^2.
In other words, when the resolution goes to infinity, this property imposes that the signal approximation converges to the original signals. This also holds for our case since the kernel functions are interpolation polynomials and, therefore, the regularity of the approximations increases as we double the degree of the highest polynomial so that at infinity the approximation converges to the original signal.
Property 4: $\bigcap_{-\infty}^{+\infty} V_{n_1,n_2} = \{0\}$.
In other worlds, when the resolution tends to 0 implies that we lose all the details of the signal. In our case when the resolution tends to zero the scaling function will be constant (the order of the polynomial is null) and therefore, the projection of the signal onto V is a constant. As a result we lose the signal details.

3.3 Quantization and bit allocation

After computing the proposed 2D orthogonal polynomials based transform, the transform coefficients are coded. To this end, we use quantization to reduce the number of bits needed to store the obtained coefficient by reducing the precision (this step is performed on both wavelet and approximation coefficients). The quantization is implemented using a quantization matrix, whose formula, as in JPEG is given below:

$$Quantized\ value(x,y) = round\left[\frac{T_{K,L}(x,y)}{Quantum(x,y)}\right] \qquad (14)$$

where $T_{K,L}(x,y)$ is either $a_{K,L}$ or $b_{K,L}$. The quantum value matrix $Quantum(x,y)$ is obtained through the quality factor. The quality factor depends on the desired quality of the reconstructed image vis-à-vis the compression ratio.
In general, the human eye is good at seeing small differences in brightness over a relatively large area, but not so good at distinguishing the exact strength of a high frequency brightness variation. This allows to reduce the amount of information in the high frequency components. This is done by simply choosing a quality factor that can elegantly discard the obtained high frequency wavelet coefficients. In other words, when the quality is high, the quantum value corresponding to the higher frequency wavelet coefficient positions shall be high so that the quantized value is reduced to zero. Thus, the quantum value (ranging from 1 to 25) determines the step size.
After performing the quantization, a bit allocation scheme using variable length coding is performed on the obtained coefficients. For this purpose, the quantized coefficients are reordered using zigzag scanning to form a 1D sequence.
Since the approximation coefficients $a_{K,L}$ of the proposed orthogonal polynomials based coding scheme have high magnitudes, we perform a difference pulse code modulation (DPCM). Consequently, the first element of the zigzag sequence represents the difference pulse code modulated $a_{K,L}$ value and among the remaining wavelet coefficients $b_{K,L}$, the non-zero wavelet coefficients are Huffman coded using variable length code (VLC) that defines the value of the coefficients and the number of preceding zeros. For this purpose, we used the standard VLC tables of the JPEG baseline system.
Finally, the compressed image can be decompressed by using a look-up table. First, the rearranged array of transform coefficients is reordered into 2D block from the 1D regenerated zigzag sequence with dequantization. Then, we reconstruct the sub image under analysis by using the proposed transform defined in the previous subsection.
The performance of the proposed scheme is evaluated in the next section.

4. EXPERIMENTAL RESULTS

The proposed transform coding scheme has been experimented and compared with other image compression standards, namely, the JPEG (with arithmetic coding options), JPEG2000 (without visual weighting and visual masking), the JPEG-XR/HDPHOTO [14] and with the nonseparable 2D wavelet-based coding scheme proposed in [16] with entropy-constrained vector quantization (ECVQ). For Jpeg2000 and Jpeg-XR we have used the softwares or source codes available in [20] and [19] respectively. The JPEG software used is the one provided by the Independent JPEG Group [21]. Our compression scheme, including the quantization and entropy coding, was implemented using Matlab .
The set of image used for the experiments is from the ISO images originally used for JPEG-XR evaluation. Figure 3 shows the three images used for illustration in this paper. The Multiscale Structural Similarity Index (M-SSIM) [18] and the Visual Difference Predictor (VDP) [12] metrics are used to evaluate the quality of the decoded images. The M-SSIM judges how irritating image artifacts generated by the compression technology are, and VDP predicts whether these artifacts are detectable. Results are found in Fig. 4 for the three images of figure 3 and for three types of polynomials (Legendre, Tchebychev and Hermite). Both M-SSIM as well as VDP have been plotted in the logarithmic domain, i.e. the first set of graphs shows $-20log(1-MSSIM)$, the second $-20log(r)$ where r is the ratio of pixels the VDP standard observer would detect as different with a probability of $p \geq 75\%$. Note that a M-SSIM index of 1 indicates a perfect match, which will be mapped to ∞ in our plots. As we can see, our approach achieved higher quality results than JPEG baseline and JPEG-XR baseline. However, better quality is obtained by using JPEG2000. The approach proposed in [16] achieved better SSIM and VDP scores than our approach for the case of Chebychev and Legendre polynomials.
However, better scores are obtained with our approach with Hermite 2D polynomials.

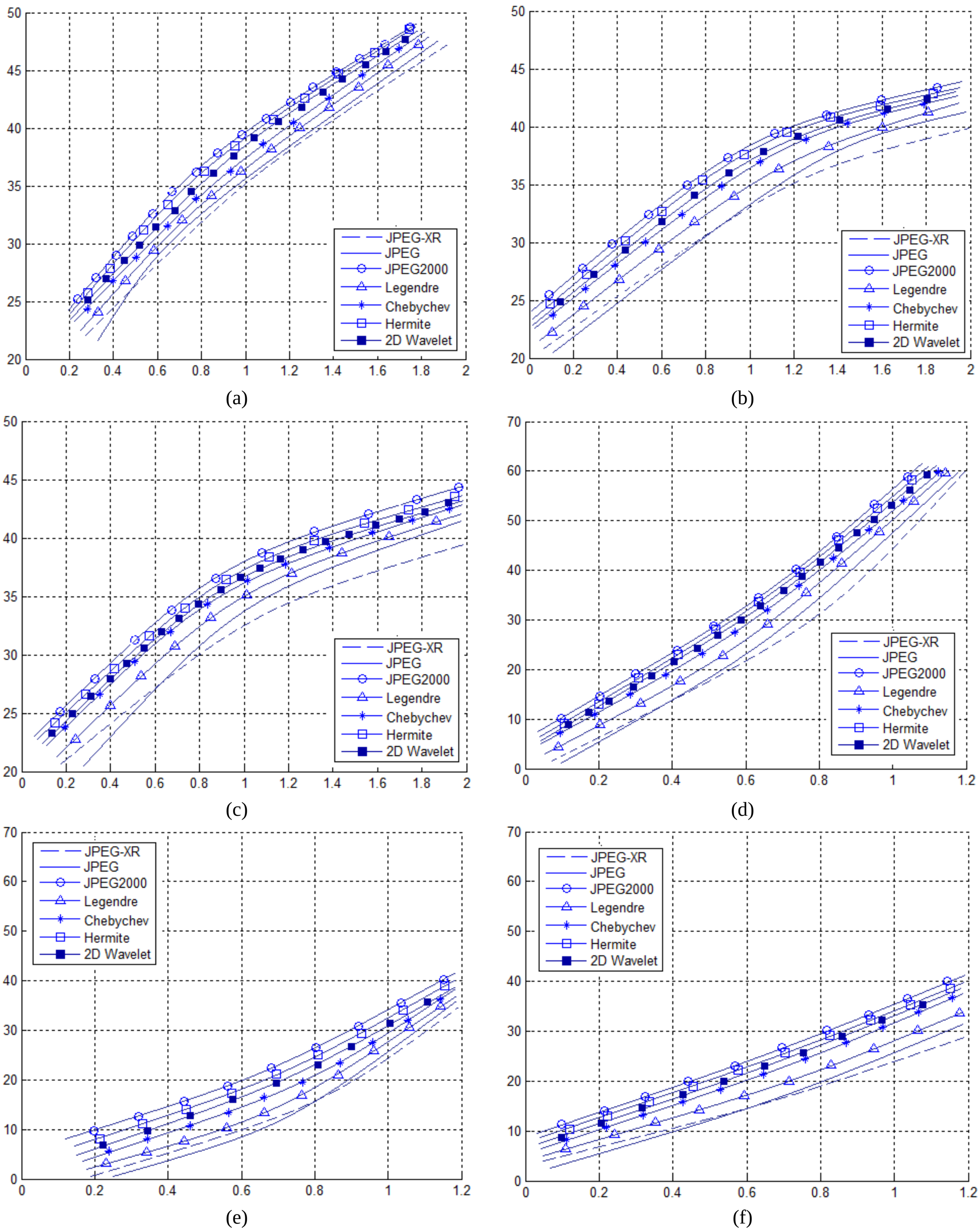


Figure 4: M-SSIM vs bpp graphs: (a), (b) and (c) and VDP vs bpp graphs: (d), (e) and (f) for the "honolulu cathedral", "oahu waimea2" and "waikiki at night" images respectively

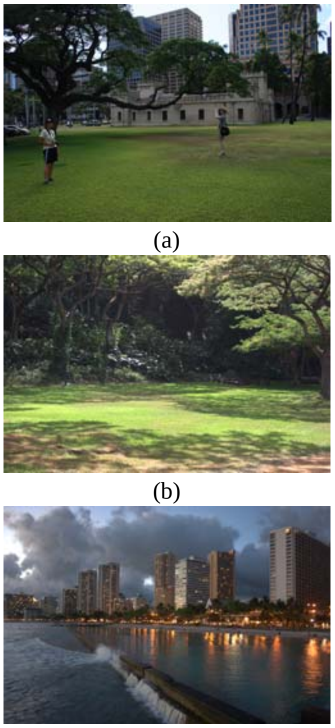

(a) (b) (c)

Figure 3: Test images: (a) "honolulu cathedral", (b) "oahu waimea2" and (c) "waikiki at night"

5. CONCLUSION

Bivariate orthogonal polynomials have been utilized to propose a new transform coding scheme. First we develop a wavelet-like transform from polynomial basis to propose a transform coding technique.
After applying the proposed transformation, the transform coefficients are scalar quantized and subjected to bit allocation scheme using variable length coding as in JPEG baseline system.
The performance of the proposed transform coding is reported by computing different quality assessment metrics and is also compared with other coding standard.

REFERENCES

[1] A. S. Blaivas, “ Visual analysis in unspecialized receptive fields as an orthogonal series expansion, ” *Neurophysiology, Springer New York*, 6, 2, pp. 1573–9007, 1974.

[2] Charilaos Christopoulos, A. Skodras, T. Ebrahimi, “The JPEG2000 still image coding system: an overview, ” in *IEEE Transactions on Consumer Electronics*, 46, 4, pp. 1103–1127, 2000.

[3] Charles F. Dunkl and Yuan Xu, *Orthogonal polynomials of several variables*, Cambridge University Press, 2001.

[4] Bernd Fischer and Jürgen Prestin, “ Wavelets based on orthogonal polynomials, ” *Mathematics of Computation*, 66, 220, pp. 1593–1618, 1997.

[5] G. Davis, A. Nosratinia, “ Wavelet based image coding: an overview, ” *IEEE Transactions on Image Processing*, 5, 1996.

[6] G.K. Wallace, “The JPEG still picture compression standard, ” *Communications of ACM*, 34, 4, pp. 31–44, 1991.

[7] Iain E.G. Richardson, *H.264/MPEG Ű 4, Part 10 White Paper: Transform and Quantization*, www.vcodex.com (2003).

[8] K.T. Lo, W.K. Cham, “Development of simple orthogonal transforms for image compression, ” in *IEE Proceedings on Vision Image and Signal Processing* 142, 1, pp. 22–26, 1995.

[9] M. Antonini, M. Barlaud, I. Daubechies, “ Image coding using wavelet transform, ” *IEEE Transactions on Image Processing*, 1, 2, pp. 205–220, 1992.

[10] M. Iwahashi, S. Fukuma, A. Hosotani, N. Kambayashi, “Two channel non-separable 2D subband coding and its optimization, ” in *International Conference on Image Processing (ICIP’97)*, 2, pp. 633–636, 1997.

[11] N. Ahamed, T. Natarajan, K.R. Rao, “ Discrete cosine transform, ” *IEEE Transactions on Computers*, C23, pp. 90–93, 1974.

[12] R. Mantiuk, S. Daly, K. Myszkowski and H. P. Seidel, “Predicting Visible Differences in High Dynamic Range Images-Model and its Calibration, ” in *Proc. of Human Vision and Electronic Imaging X, IS&T/SPIE’s 17th Annual Symposium on Electronic Imaging*, pp. 204-214, 2005.

[13] R.D. Dony, S. Haykin, “ Optimally adaptive transform coding, ” *IEEE Transactions on Image Processing*, 4, 10, pp. 1358–1370, 1995.

[14] S. Srinivasan, C. Tu, S. L. Regunathan, R. A. Rossi, and G. J. Sullivan, “HD Photo: a new image coding technology for digital photography, ” in *Applications of Digital Image Processing XXX, Proceedings of SPIE, San Diego, CA USA*, 6696, pp. 66960A, 2007.

[15] R. Krishnamoorthi and N. Kannana, “A new integer image coding technique based on orthogonal polynomials, ” in *Image and Vision Computing*, in press, 2008.

[16] Wajcer D., Stanhill D. and Zeevi Y., “Representation And Coding of Images With Nonseparable Two-dimensional Wavelet, ” in *International Conference on Image Processing (ICIP’98), Chicago, USA*, pp.336-340, 1998.

[17] W.K. Pratt, J. Kane and H.C. Andrews, “Hadamard transform image coding, ” in *Proceedings of IEEE*, 57, pp. 58–68, 1969.

[18] Z. Wang, E.P. Simoncelli and A.C. Bovik, “Multi-scale structural similarity for image quality assessment, ” in *IEEE Asilomar Conf. on Signals, Systems and Computers*, pp. 1398-1402, 2003.

[19] *ISO, Reference software implementation of JPEG XR ISO/IEC JTC1 SC29/WG1/N4560, April 2007*.

[20] *http://www.kakadusoftware.com/*.

[21] *http://www.ijg.org*.