PEAK TRANSFORM - A NONLINEAR TRANSFORM FOR EFFICIENT IMAGE REPRESENTATION AND CODING

Zhihai He

Department of Electrical and Computer Engineering
University of Missouri, Columbia, MO, 65203, USA
HeZhi@missouri.edu

ABSTRACT

In this work, we introduce a nonlinear geometric transform, called *peak transform*, for efficient image representation and coding. Coupled with wavelet transform and subband decomposition, the peak transform is able to significantly reduce signal energy in high-frequency subbands and achieve a significant transform coding gain. This has important applications in efficient data representation and compression. Based on peak transform (PT), we design an image encoder, called *PT encoder*, for efficient image compression. Our extensive experimental results demonstrate that, in wavelet-based subband decomposition, the signal energy in high-frequency subbands can be reduced by up to 60% if a peak transform is applied. The PT image encoder outperforms state-of-the-art JPEG2000 and H.264 (INTRA) encoders by up to 2-3 dB in PSNR (peak signal-to-noise ratio), especially for images with a significant amount of high-frequency components.

Keywords - Image compression, nonlinear transform, energy compaction, wavelet subband decomposition.

1. INTRODUCTION

The key in efficient image compression is to explore source correlation so as to find a compact representation of image data. Over the past decades, various spatial transforms, such as KLT, DCT (discrete cosine transform) and DWT (discrete wavelet transform) [1], have been developed to explore source correlation. These transforms mentioned here are linear which can be represented by matrices. These linear transforms, including KLT, DCT, and DWT, are designed to remove statistical source correlation such that the output components are statistically independent of each other. We refer to this type of statistical correlation explored by linear transforms as *linear correlation*. As we know, images (and videos) are a special type of data. They are not just 2-D arrays of pixels in a statistical sense. For example, if we randomly generate a 2-D array of data according to a given statistical distribution, the probability for this 2-D array of data to be a natural image is extremely low. This is because, besides statistical characteristics, natural images contains a lot of non-statistical perceptual image features, such as edges, contours, patterns, structures, and objects. In other words, in images and videos, besides linear correlation, there is a significant amount of *nonlinear* source correlation presented by these perceptual image features. This type of nonlinear correlation has been left largely unexplored by linear transforms, such as KLT, DCT, and DWT.

During the past decades, researchers have been designing efficient prediction scheme to explore the nonlinear source correlation which has been left largely unexplored by linear spatial transforms. For example, the cross-subband parent-children dependency has been observed and explored by EZW (embedded zero-tree wavelet), SPIHT, JPEG2000 [1], and many other wavelet-based image coding algorithms. Recently, several modified wavelet transforms which take edge flow or texture orientation into account have been developed. These transforms include curvelet and ridgelet transforms [4]. Another important type of techniques, called directional or orientation-adaptive wavelets which combine directional prediction and wavelet transform, have been developed in the literature [2, 3].

In this work, we propose to explore a new approach: developing a nonlinear geometric transform, called *peak transform*, to assist the wavelet transform in exploring nonlinear data correlation. Conceptually speaking, the proposed peak transform is able to convert a hard-to-compress signal into an easier-to-compress signal by exploring the nonlinear geometric source correlation within the input signal. We will study various design issues of the PT image encoder. Our experimental results demonstrate that the peak transform is able to significantly improve the transform coding again. The new PT encoder outperforms state-of-the-art image encoders, including JPEG2000 [1] and H.264 by up to 2-3 dB, especially for images with a significant amount of high-frequency components.

The rest of the paper is organized as follows. In Section 2, we will present the mathematical definition of peak transform and discuss its major properties from a data compression perspective. The PT image encoder design will be discussed in Section 3. The experimental results are presented in Section 4. Section 5 will discuss future research directions and conclude the paper.

2. DEFINITION AND PROPERTIES OF PEAK TRANSFORM

Definition: ***Curve Segment.*** A curve segment is a function $f(x)$ defined over a finite interval $[a, b]$.

Definition: ***Cascade of Curve Segments.*** Given two curve segments $f_1(x)$ and $f_2(x)$ defined over finite intervals $[a_1, b_1]$ and $[a_2, b_2]$ with $b_1 \geq a_2$, the cascade of these two curve segments yields a new curve segment $f(x)$ defined over $[a_1, b_1 + b_2 - a_2]$:

$$f(x) = \begin{cases} f_1(x), & x \in [a_1, b_1] \\ f_2(x) - f_2(a_2) + f_1(b_1), & x \in (b_1, b_1 + b_2 - a_2]. \end{cases} \quad (1)$$

We denote this cascading operation by

$$f(x) = f_1(x) \uplus f_2(x). \quad (2)$$

Physically, the new curve segment $f(x)$ is obtained by joining two curve segments $f_1(x)$ and $f_2(x)$ with proper shifting operations as illustrated in Fig. 1.

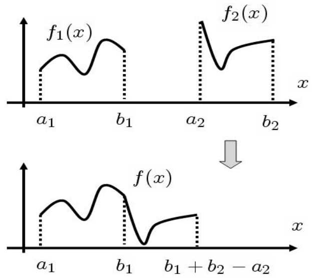

Figure 1: Cascade of two curve segments.

Now we are ready to define the peak transform.

Definition: *N-Point Forward Peak Transform.* A function $f(x)$ is defined over a finite interval $[a, b]$. This interval is partitioned into $N+1$ sub-intervals by N points, $a < x_1 < x_2 < \cdots < x_N < b$. For convenience, we write $x_0 = a$ and $x_{N+1} = b$. We refer to $\{x_n\}$ as peaks (or breaking points). The curve segment defined over interval $[x_{i-1}, x_i]$ is denoted by $f_i(x)$, $1 \leq i \leq N+1$. The N-point peak transform of $f(x)$, denoted by $\mathcal{P}(\{x_i\})[f(x)]$ is defined as

$$\mathcal{P}(\{x_i\})[f(x)] = g_o(x) \uplus g_e(x), \quad (3)$$

where

$$g_o(x) = f_1(x) \uplus f_3(x) \uplus \cdots \uplus f_{2\cdot\lfloor N/2 \rfloor + 1}(x), \quad (4)$$

and

$$g_e(x) = f_2(x) \uplus f_4(x) \uplus \cdots \uplus f_{2\cdot\lfloor N/2 \rfloor}(x) \quad (5)$$

are the cascades of all odd and even-numbered curve segments, respectively. Physically speaking, in peak transform, we first cascade all odd-numbered curve segments then all even-numbered curve segments and form a new curve. Fig. 2 shows an example of 5-point peak transform. It can be seen that the peak transform only changes the order of curve segments and is reversible. The backward transform can be done by simply cascading the curve segments according to their original order. We denote the backward peak transform operation by $\mathcal{P}^{-1}(\{x_n\})[\cdot]$.

2.1. Properties of Peak Transform

One important property of peak transform is that it is capable of converting a high-frequency signal into a low-frequency one if the peaks are properly selected. In the following example, we demonstrate this unique property of peak transform on image signals. We take the 330-th row of image *Barbara* as the input signal $f(x)$, which is shown in Fig. 3(A). The peaks used in peak transform are shown in diamonds. In this example, we first apply the forward peak transform to $f(x)$ and obtain $\mathcal{P}[f(x)]$. Fig. 3(B) shows the peak transform output. As in the toy example, we pass $\mathcal{P}[f(x)]$ to a Debauches (9, 7) filter bank and obtain the low and high-frequency components (subbands) of $\mathcal{P}[f(x)]$. We then apply backward peak transform to these two subbands. Figs. 4(A) and (B) show the low and high-frequency subbands without and with peak transform. Again, we can see that the output signal of peak transform has a much smaller amount of high-frequency components while their low-frequency subbands are very similar. According to our experiment, the energy of the high-frequency subband with peak transform is about 43% of that without peak transform.

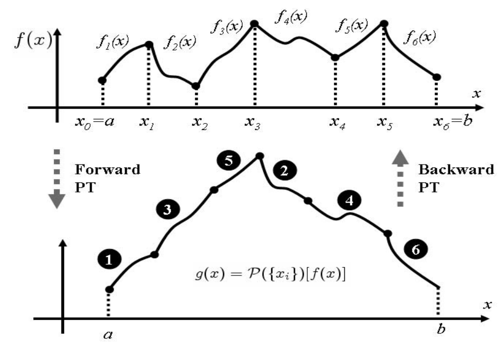

Figure 2: An example of 5-point peak transform.

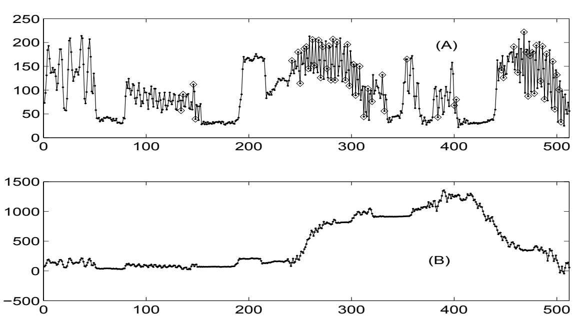

Figure 3: Peak transform of the 330-th row of image Barbara: (A) the original piece-wise linear function and the selected peaks shown in diamonds; (B) the peak transform output.

3. PEAK TRANSFORM BASED DATA COMPRESSION SYSTEM DESIGN

In this section, we will discuss how to design an data compression system based on peak transform. In peak transform based data compression, the peak transform is jointly used with wavelet transform and subband decomposition to

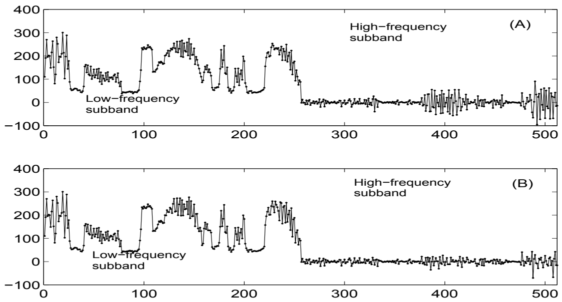

Figure 4: (A) the low and high-frequency subbands of the original signal; (B) the low and high-frequency subbands of the PT output.

minimize the signal energy in high-frequency subbands.

Let us start with 1-D input signals, which can be a row or a column of image pixels. Let $f(x)$ be an input signal of length M where $1 \leq x \leq M$. As illustrated in Fig. 5, an N-point forward peak transform with peaks $\{x_n\}$ is applied to the input signal $f(x)$. The peak transform output, denoted by $\mathcal{P}(\{x_n\})[f(x)]$, is then passed to a two-branch filter bank with a low-pass filter $H_0(z)$ and a high-pass filter $H_1(z)$. An N-point backward peak transform with peaks $\{y_n\}$, denoted by $\mathcal{P}^{-1}(\{y_n\})[\cdot]$, is applied to both low and high-frequency subbands. It should be noted that the sizes of low and high-frequency subbands due to down-sampling are both $\frac{M}{2}$. Therefore, the original peaks cannot be used any more. One possible approach is to force the location of each original peak x_n to be an even integer. During the backward peak transform of two subbands, we use $\{y_n = \frac{x_n}{2}\}$ as peaks. Certainly, we can choose other peaks. However, using $y_n = \frac{x_n}{2}$, the backward peak transform can bring the wavelet transform coefficients back to their initial pixel order as in the original signal. The above procedure, which decomposes the input signal into two subbands using peak transform (PT) and wavelet transform (WT), is called *PTWT subband decomposition*. The PTWT subband decomposition procedure can be repeated for the low and high-frequency subbands to further decompose the signal into more frequency subbands.

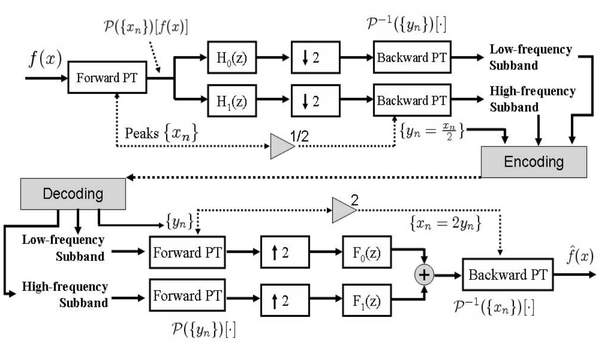

Figure 5: PTWT subband decomposition and synthesis.

The frequency subbands will be quantized, entropy encoded, and transmitted to the decoder, as illustrated in Fig. 5. The peaks $\{y_n\}$ will be also compressed with lossless coding and sent as overhead information to the decoder. At the decoder side, the peaks $\{y_n\}$ will be decoded. A forward peak transform $\mathcal{P}(\{y_n\})[\cdot]$ with peaks $\{y_n\}$ will be applied to low and high-frequency subbands. After subband synthesis, the backward peak transform $\mathcal{P}^{-1}(\{x_n\})[\cdot]$ will be performed to obtain the reconstructed signal $\hat{f}(x)$. Here, $x_n = 2y_n$. We can see that the PTWT subband decomposition and synthesis illustrated in Fig. 5 guarantees perfect reconstruction and the only loss is caused by quantization, as in conventional wavelet-based data compression.

The 1-D PTWT subband decomposition can be extended to 2-D images. Fig. 6(B) shows a 3-level PTWT subband decomposition of image *Barbara*. The corresponding peak map is shown in Fig. 6(C). For comparison, we also shown the 3-level subband decomposition using DWT only in Fig. 6(A). It can be seen that the transform coefficients in (B) have much smaller magnitudes than those in (A). The subband data will be compressed using the SPIHT-like encoder developed in our previous work [6]. The corresponding peak map is shown in Fig. 6(C). The peak map will be encoded using context adapative arithmetic coding. [1]

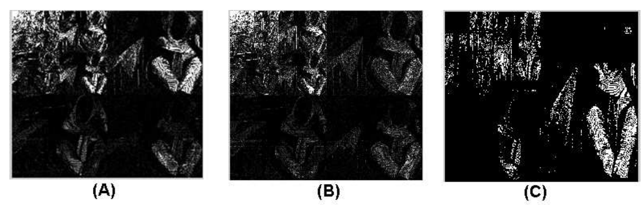

Figure 6: (A) 3-level subband decomposition of *Barbara*with wavelet transform; (B) 3-level PTWT subband decomposition; (C) the peak map.

4. EXPERIMENTAL RESULTS

In this section, we evaluate the data compression efficiency of the proposed PT image encoder and compare its performance with state-of-the-art image encoders, include JPEG-2000 and H.264 INTRA coding. The JPEG-2000 image encoder we used in this performance evaluation is the Jasper encoder. In H.264 encoding, we use JM 9.0 H.264 video encoder (INTRA frames only) with CABAC (context adaptive arithmetic coding). The test images (grayscale) of size 512×512 are shown in Fig. 7. To deal with grayscale images in H.264, we set the chrominance components (Cb and Cr) of each frame to be a constant 128. Figs. 8 shows the PSNR (peak signal-to-noise ratio) performance of the PT encoder in comparison with JPEG-2000 and H.264 image coding over all test images. It can be seen that the PT encoder consistently outperforms the other two image encoders on all test images. Fig. 10 shows the fraction of bits used for encoding peak map as a function of coding bit rate for each test image. We can see that, as the overall coding bit rate

[1] Due to page limitation, in this paper, we are not able describe in detail how to obtain the optimum set of peaks for peak transform. Please refer to our technical report for more detail.

increases, the fraction of peak map bits decreases. (However, it should be noted that, the number of peak map bits does increase.) Fig. 9 shows the images of *Barbara* encoded at 0.4 bpp by JPEG2000, H.264INTRA, and the proposed PT encoder. It can be seen that, with PT encoding, the image has a much better visual quality with enhanced edge information. This is because the PT encoder is able to efficiently preserve high-frequency image features using peak transform and peak map.

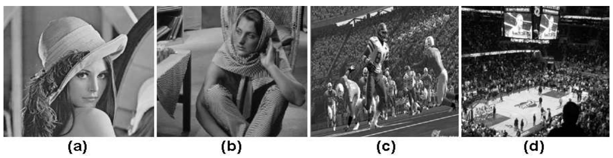

Figure 7: Test images: (a) Lena; (b) Barbara; (c) NBA2; (d) Football2.

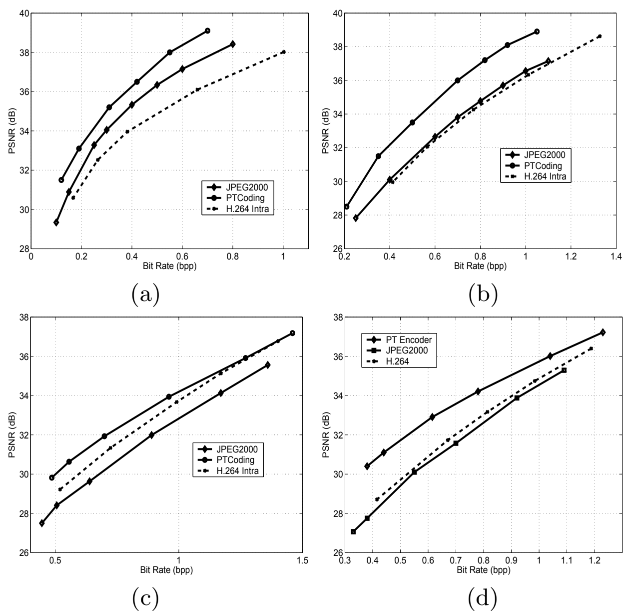

Figure 8: Compression performance comparison with JPEG2000 and H.264 (INTRA) on images: (a) Lena; (b) Barbara; (c) NBA2; (d) Football2.

5. CONCLUDING REMARKS, DISCUSSION, AND FURTHER RESEARCH DIRECTIONS

The major contribution of this work is that we have introduced a nonlinear geometric transform, called peak transform, which is able to convert high-frequency signals into low-frequency ones. Coupled with wavelet transform and subband decomposition, the peak transform is able to significantly reduce the signal energy in high-frequency subbands. This has significant applications in data compression of 1-D signals (e.g. speech and acoustic signals) and 2-D images. We have developed an fast and efficient dynamic

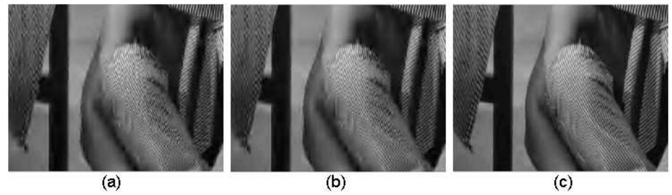

Figure 9: Subjective picture quality comparison on image *Barbara* coded at 0.4bpp: (a) JPEG2000; (b) H.264 INTRA; (c) PT encoding.

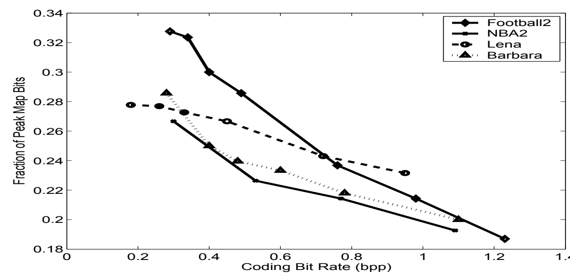

Figure 10: Fraction of peak map bits v.s. coding bit rate for each test image.

solution to find optimum (or sub-optimum) peak transform to minimize the high-frequency subband energy or maximize the transform coding gain. We have also studied how to design an image compression system based on peak transform. Our experimental results show that the proposed PT encoder outperforms the state-of-the-art image encoders, including JPEG2000 and H.264 (INTRA).

6. REFERENCES

[1] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*, Kluwar Academic Publishers, 2002.

[2] W. Ding, F. Wu, and S. Li, "Lifting-based wavelet transform with directionally spatial prediction," in *Proc. Picture Coding Symposium 2004*, San Francisco, CA, USA, Dec. 2004.

[3] D. Wang, L. Zhang, A. Vincent, and F. Speranza, "Curved wavelet transform for image coding," *IEEE Transactions on Image Processing*, vol. 15, No. 8, pp. 2413-2421, Aug. 2006.

[4] M. N. Do and M. Vetterli, "The finite ridgelet transform for image representation," *IEEE Trans. Image Processing*, vol. 12, no. 1, pp. 16-8, Jan. 2003.

[5] Peak Transform web page: http:// videonet. ece.missouri.edu / research / peaktransform.htm.

[6] Zhihai He, Tian-Hu Yu and S. K. Mitra, "Blockwise zero mapping image coding," *Proceedings of International Conference on Image Processing*, vol. 3, pp. 170-173, Sep. 2000.