Signal Decomposition Using Adaptive Block Transform Packets

Jyhchau Horng[1] and Richard A. Haddad[2]

[1]Department of Electrical Engineering, Polytechnic University, Five Metrotech Center
Brooklyn, NY 11201 ; jhorng01@utopia.poly.edu
[2]Department of Electrical & Computer Engineering, New Jersey Institute of Technology
University Heights, NJ 07102 ; haddad@admin.njit.edu

Abstract

A Block Transform Packet (BTP) is an orthonormal block transform which is constructed from conventional block transforms and represents an arbitrary tiling of the time-frequency plane.[5] Unlike the progenitor transforms, the BTP has time-localizabilities and is capable of dealing with non-stationary signals. This paper describes procedures for signal decomposition using the BTP in an adaptive way. Three examples show the adaptive compression efficiency over DCT.

I. Introduction

Block transforms such as the DCT, DFT and Walsh-Hadamard transform are widely used as suboptimal solutions for signal decomposition. These transforms have orthogonal bases and fast computational algorithms. However, these basis functions result in a uniform time-frequency(T-F) resolution grid, as indicated in Fig.1(a), which is not suitable for processing time-localized and non-stationary signals. On the other hand, the Kronecker delta sequence has the time-localized T-F pattern, as shown in Fig.1(b), which obviously can not resolve frequency concentrated signals. In these diagrams, the time-frequency tile or cell of a particular basis function is the region in the time-frequency plane where most of that function's energy is concentrated[4]. These tilings represent 8×8block transforms.

The discrete Wavelet Transform (DWT) has the nonuniform dyadic T-F tiling pattern[1][3], shown in Fig.1(c) for three stages of decomposition. The more general wavelet packet (WP)[2][8] can realize arbitrary tilings,as in Fig.1(d). The DWT and WP tilings are based on frequency segmentation.

On the other hand, a block transform packet (BTP)[5] can be based on either frequency or time segmentation and is an efficient alternative to the WP. The BTP is an orthonormal block transform which represents an arbitrary time-frequency tiling pattern. Given desirable T-F tiling pattern, BTP can be synthesized from conventional block transforms in an optimum way, while maintaining the computational efficiency of the progenitor transforms. Unlike the time-varying tree structure in [4][8], here no boundary or transition filters are needed in the transitions and the tree structure associated with the desired tiling pattern could be time-varying. For example, Fig.1(e) is a BTP which can not be realized by an 8×8 time-invariant frequency concentrated decomposition tree structure.

II. Block Transform Packets

We define two classes of BTP's which are represented by Fig.2[5,6,7].
1) the time-localizable BTP(TLBTP) is based on a transformation of a frequency selective block transform (e.g. DCT) with a T-F tiling pattern as in Fig.1(a) into an orthonormal block transform with desired features as in Fig.1(e).
2) The frequency localizable block transforms packets is the conceptual dual of the TLBTP. It transform the time localized Kronecker delta basis sequences with T-F pattern as in Fig.1(b) into a desired pattern as in Fig.1(e). Details are found in [5].

In TLBTP, the original frequency-focused transform looks like a bank of band-pass filters with impulse responses extending over the entire time frame. As indicated in Fig.2, we partition the entire set of original basis sequences into subsets and then find the optimal unitary submatrices $\boldsymbol{A_k}$ such that the new basis sequences are time-localized in accordance with the desired T-F tiling pattern.

In accordance with the uncertainty principle[3], we trade frequency resolution for desired time localization. Consider a $\boldsymbol{l}^2$(Z) space of dimension N with a given set of orthonormal basis sequences, $\Phi(\boldsymbol{n}) = \{\ \phi_j(n)$: $j,n = 0,1,...,N-1\}$; we make partitions into orthogonal subspaces of dimension M_k spanned by an associated subset of M_k bases, $\{\Phi_{M_k} : \Sigma_k M_k = N\}$. Then we may find new orthonormal basis sequences,$\Psi(\boldsymbol{n}) = \{\ \psi_j(n)\ :\ j,n = 0,1,...,N-1\}$, as follows:

$$\begin{bmatrix} \psi_0(n) \\ \psi_1(n) \\ \vdots \\ \psi_{N-1}(n) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_0 & & & \boldsymbol{0} \\ & \boldsymbol{A}_1 & & \\ & & \ddots & \\ \boldsymbol{0} & & & \ddots \end{bmatrix} \begin{bmatrix} \phi_0(n) \\ \phi_1(n) \\ \vdots \\ \phi_{N-1}(n) \end{bmatrix} \qquad (1)$$

$$\Psi(\boldsymbol{n}) \quad = \quad \mathbf{A} \quad \Phi(\boldsymbol{n})$$

The $N\times N$ matrix $\mathbf{A}$ is a diagonal block matrix which maps the entire block transform bases set, $\Phi(\boldsymbol{n})$, into the entire TLBT bases set, $\Psi(\boldsymbol{n})$. The $M_k \times M_k$ coefficient matrix A_k is constrained to be a unitary matrix, *i.e.* $\boldsymbol{A_k}\, \boldsymbol{A_k}^* = \boldsymbol{I}$, where * is the hermitian operator. In each subspace, the new orthonormal basis sequence, $\Psi_k(\boldsymbol{n}) = \{\psi_{k_i}(n) :\ i=0,1,...,M_k-1,\quad n=0,1,...,N-1\}$, is a linear combination of the M_k original bases $\{\ \phi_{k_0}(n), \phi_{k_1}(n), \cdots, \phi_{k_{M_k-1}}(n)\}$ such that each $\psi_{k_i}(n)$ maximally concentrates its energy in the time interval $I_i=\{i(N/M_k) \leqslant n \leqslant (i+1)(N/M_k)-1\}$, and $\{I_i\}$ span the entire transform frame of length N. Equivalently, we minimize the energy outside I_i, expressed as

$$J_i' = \sum_{n \notin I_i} |\ \psi_{k_i}(n)\ |^2 \qquad (2)$$

The result is that u_i, the i^{th} row of $\boldsymbol{A_k}$, is the eigenvector of the matrix

$$E_i = \sum_{n \notin I_i} \Phi_k(n)\Phi_k^*(n) \qquad (3)$$

For computation efficiency, we can use the inverse block transform matrix as the approximation solution for $\boldsymbol{A_k}$. From Eq.(1), the location and shape of the resolution cell determine the location and size of localizing matrix $\boldsymbol{A_k}$.

As an sample example, we can convert a 64-point DCT with Fig.1(a) localized tiling pattern into the TF pattern shown in Fig.3(a). The energy distribution of the basis functions for resolution cells 1 and 3 are shown in Fig.3(b)(c).

III. Optimal Decomposition in T-F Plane

Since we have a procedure to determine optimally concentrated basis functions from T-F pattern, the next concern is how to determine the best T-F decomposition pattern for a given signal. In this paper, we develop a procedure to determine the optimal T-F decomposition for a given frame of signal. From this tiling pattern, the associated BTP is generated as described herein. This procedure can be adapted from frame to frame.

A resolution cell is a rectangle of constant area and a given location in the time-frequency plane. The tiling pattern is the partitioning of the time-frequency plane into contiguous resolution cells. This is a feasible partitioning. Each coefficient of the new transform represents the signal strength associated with a resolution cell. We want to find the tiling pattern corresponding to maximum energy concentration for that particular signal. From energy compaction point of view, the tiling pattern should be chosen such that the energies concentrate in as few coefficients as possible.

A. Microcell Approach

The Kronecker delta sequence resolves the time domain information and the frequency selective block transforms provide the frequency information. Combining these two characterizations together gives the energy sampling grid in the time-frequency plane. Let $\boldsymbol{x}_i$ represent the amplitude square of the function *f(n)*, $0 \le n \le N-1$, at time t_i and $\boldsymbol{y}_i$ be the magnitude square of the coefficient of the frequency selective block transform (e.g. DCT) at frequency slot f_i. Take outer product of these two groups of samples, $\boldsymbol{P_{ij}} = \boldsymbol{x}_i\ \boldsymbol{y}_j$, i,j=0,1,...,N-1, and each $\boldsymbol{P}_{ij}$ represents the energy strength in the corresponding area in the time-frequency plane. The area corresponding to each $\boldsymbol{P}_{ij}$ is called a "microcell". $\boldsymbol{P}=\{\boldsymbol{P}_{ij}\}$ is the microcell energy pattern for a given signal. Totally we have N^2 microcells and each resolution cell is composed of N microcells. Therefore, our task here is to regroup the microcells such that the tiling pattern has the maximum energy concentration.

B. Search for the Most Energetic Resolution Cell

The most energetic resolution cell in $\boldsymbol{P}$ is the rectangular region which is composed of N microcells and has the maximum energy strength. Our objective is to search $\boldsymbol{P}=\{\boldsymbol{P}_{ij}\}$, the pattern of N^2 energy microcells in the T-F plane, to find the feasible pattern of N resolution cells $\boldsymbol{Z}_i$, $0 \le i \le N-1$, such that the signal energy is optimally concentrated in as few cells as possible. We can perform an exhaustive search of $\boldsymbol{P}$ using rectangular windows of size N to find the most energetic resolution cell, and then the second most energetic resolution cell, and so on. With some assumptions, we can improve the search efficiency as follows. Assume that the most energetic microcell $\boldsymbol{P}^*_{ij}$ is included in the most energetic resolution cell $\boldsymbol{Z}^*_i$. We search the neighborhood of $\boldsymbol{P}^*_{ij}$ to find the rectangular cluster of microcells with the most energy. That cluster defines the most energetic resolution cell $\boldsymbol{Z}^*_i$. Therefore, starting from the most energetic microcell, we regroup the microcells to find the most energetic resolution cell. The procedure is as follews:

1) Rank order $\boldsymbol{P}_{ij}$, and put the rank ordered index (i,j) in *idx*(.). idx(1) is the index of the most energetic microcell.
2) Calculate the rectangular area A_1 specified by *idx*(1) and *idx*(2). $A_1 = (i_1-i_2+1)\ (j_1-j_2+1)$. If $A_1 \le N$, then both microcells are included within or on the border of a resolution cell. If $A_1 > N$, these two microcells can not be included in the same resolution cell.
3) Test the third ranked microcell. If it is inside A_1 , fine. If it is outside A_1 , calculate A_2 which includs A_1 and *idx*(3). Test A_2 .
4) Repeat test until A_{last}=N. The location of A_{last} is the most energetic resolution cell $\boldsymbol{Z}^*_i$.

Repeat this search for next most energetic resolution cell and a complete optimal T-F tiling pattern can be obtained. This procedure is tedious and not practical for large transforms. In next section, we will describe a more efficient way, an adaptive approach.

IV. Adaptive Approach

The objective of the proposed method is to expand our signal in terms of a BTP basis functions in a sequential fashion, i.e., find one resolution cell from a succession of N T-F tiling patterns rather than N cells from one T-F pattern. Fig.4 suggests the following adaptive scheme:

1) Start at the stage q=1. We construct $\boldsymbol{P}_1$ from *f(n)* and use the microcell and search algorithm described in Section III to find the most energetic resolution cell $\boldsymbol{Z}_1$ with its associated basis function $\psi_1(n)$ and block transform packets $\boldsymbol{T}_1$. The projection of *f(n)* onto $\psi_1(n)$ gives the coefficient β_1 and our first approximation

$$\hat{f}_1(n) = \beta_1\psi_1(n) \qquad (4)$$

2) Take the residual $\tilde{f}_1(n)$ as the input to the next stage where

$$\tilde{f}_1(n) = f(n) - \hat{f}_1(n) = f(n) - \beta_1\psi_1(n) \qquad (5)$$

3) Repeat (1) and (2) for q>1 where the residual signal $\tilde{f}_i(n)$ at ith stage is

$$\tilde{f}_i(n) = \tilde{f}_{i-1}(n) - \hat{f}_i(n) = \tilde{f}_{i-1}(n) - \beta_i\psi_i(n) \qquad (6)$$

and $\psi_i(n)$ is the most energetic basis function corresponding to tiling pattern $\boldsymbol{P}_i$ and BTP $\boldsymbol{T}_i$.

In general, the basis functions $\psi_i(n)$ are not othonormal to each other. However, in each stage the BTP is an unitary transform and therefore, $\|f(n)\| > \|\tilde{f}_1(n)\|$ and $\|\tilde{f}_{i-1}(n)\| > \|\tilde{f}_i(n)\|$. Thus the norm of the residual $\tilde{f}_i(n)$ monotonically decreases and converges to zero.

Because this representation is adaptive, it will be generally concentrated in a very small subspace. As a result, we can use a finite summation to approximate the signal with a residual error as small as one wishes. The approximated signal can be expressed as

$$\hat{f}(n) = \sum_{i=1}^{L} \beta_i \psi_i(n) \tag{7}$$

The error energy for that frame using L coefficients is

$$\Omega_L = \sum_{n=0}^{N-1} | \tilde{f}_L(n) |^2 \tag{8}$$

For long-length signal, this scheme can be adapted from frame to frame.

V. Examples

Three examples are given to show the energy concentration property of the adaptive BTP. The BTP is constructed from a DCT base with block size 32. In each example the signal length is 1024. The data sequence is partitioned into 32 frames consisting of 32 samples per frame. For each frame, we compute the residual $\tilde{f}_i(n)$ and the corresponding error energy Ω_i, $1 \le i \le 4$. The average of these Ω_i's over 32 frames is plotted in Fig.5.

Fig.5(a) shows the energy concentration property in terms of number of coefficients for a narrow band gaussian signal S_1 with bandwidth = 0.2rad and central frequency $(\frac{5\pi}{6})$. Due to the frequency localized nature of the signal, BTP does not have much improvement over DCT.

The signal used in Fig.5(b) is a narrowband gaussian signal S_1 plus time-localized white gaussian noise with 10% duty cycle S_2. Basically, it is a combination of frequency-localized and time-localized signals and therefore, it can not be resolved in time or frequency domain. Fig.5(b) shows the result for power ratio(S_1/S_2) = -2dB and Fig.5(c) is the case for -8dB. Both figures demonstrate that BTP is a more efficient compression engine over DCT.

VI. Conclusions

We have described an adaptive procedure for signal decomposition in the T-F plane. Taking one frame of signal, we use the microcell approach to search the location and shape of the most energetic rectangular resolution cell and generate the corresponding basis function. After that, we take the residual signal as the next stage input signal and repeat this procedure until the residual error converges to as small as one wish. Three examples show the adaptive compression efficiency over DCT. Other applications such as excision of interference signal in spread spectrum communication system and adaptive tracking of most energetic resolution cell from frame to frame are under study.

VII. Reference

[1] A. N. Akansu, R. A. Haddad, '*Multiresolution Signal Decomposition*', Academic Press, 1992

[2] I. Daubechies, '*Orthogona Basis of Compactly Supported Wavelets'*, Comm. Pure Appl. Math, vol. XLI, pp. 909-996, 1988.

[3] R. A. Haddad, A. N. Akansu and A. Benyassine, '*Time-Frequency Localization in Transforms,Subbands and Wavelets: a Critical Review*', Optical Engineering, vol.32, No.7, pp. 1411-1429, July 1993.

[4] C. Herley, J. Kovacevic, K. Ramchandran, and M. Vetterli, '*Tiling of the Time-Frequency Plane: Construction of Arbitrary Orthogonal Bases and Fast Tiling Algorithm*', IEEE Trans. on Signal Processing, Vol. 41, No. 12 , pp.3341-3359, Dec. 1993.

[5] J. Horng and R. A. Haddad, 'Time-Frequency Tilings Using Block Transform Packets', Proc. of 5th NJIT symposium on Wavelet, Subband and Block Transforms in Communications, March 1997.

[6] J. Horng and R.A. Haddad, '*Variable Time-Frequency Tiling Using Block Transform*', in Proc. 1996 IEEE Digital Signal Processing Workshop, pp.25-28, Norway, September 1996.

[7] K. Y. Kwak and R. A. Haddad, '*New Family of Orthonormal Transforms with Time Localizability Based on DFT*', in Proc. of Visual Communications and Image Processing, SPIE, Vol. 2308, Oct. 1994.

[8] K. Ramchandran and M. Vetterli, '*BestWavelet Packet Bases in a Rate-Distortion Sense*', IEEE Trans. on Image Processing, Vol.2 No.2, pp.160-175, April 1993.

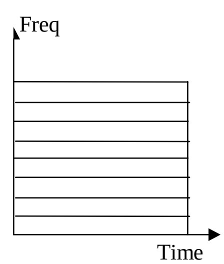

Fig.1(a)

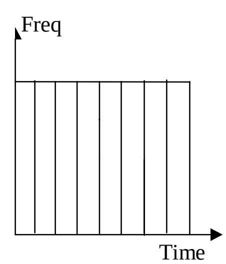

Fig.1(b)

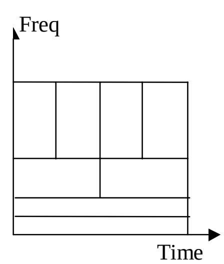

Fig.1(c)

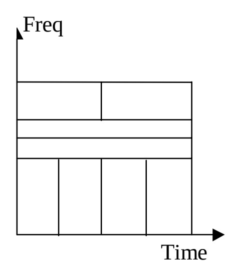

Fig.1(d)

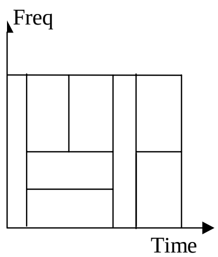

Fig.1(e)

Fig.1. Tiling pattern for resolution cells in (a) frequency localized BT (b) time localized BT (c) discrete WT (d) WP (e) BTP.

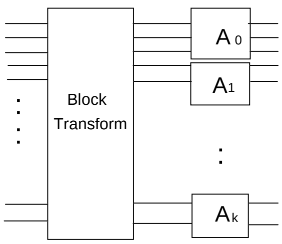

Fig.2. System diagram for BTP.

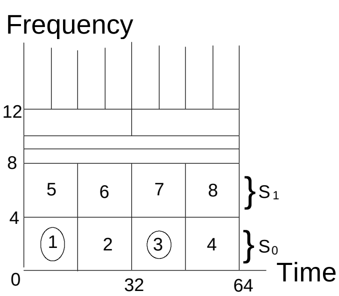

Fig.3(a)

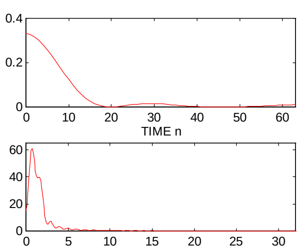

Fig.3(b)

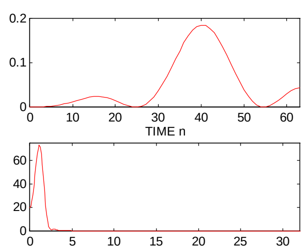

Fig.3(c)

Fig.3. Energy distributions of TLBT basis sequences in both time and frequency domain. (a) desired tiling pattern. (b) $\psi_0(n)$ for Cell 1 in Fig.3(a). (c) $\psi_2(n)$ for Cell 3 in Fig.3(a).

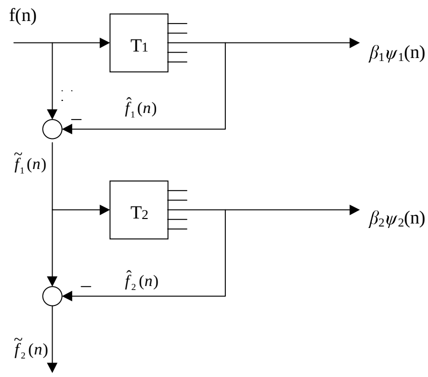

Fig.4. Adaptive Decomposition of signal *f(n)*.

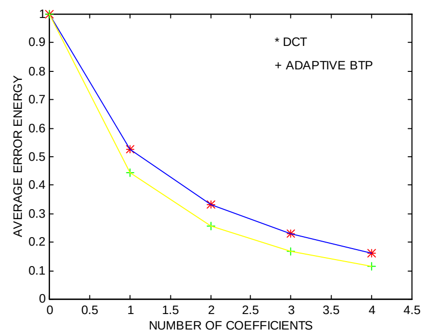

Fig.5(a)

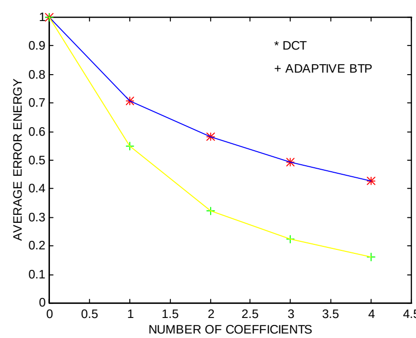

Fig.5(b)

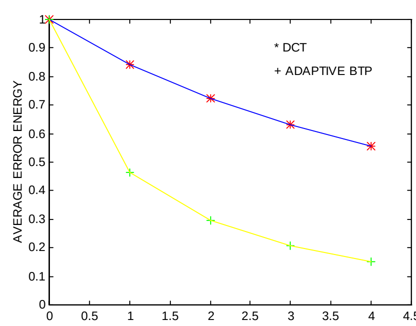

Fig.5(c)

Fig.5. Compression efficiency comparisons for (a) narrowband gaussian signal S_1 (b) S_1 plus time localized gaussian signal S_2 with power ratio S_1/S_2=-2dB (c) S_1+S_2 for -8dB.