CRITICALLY SAMPLED GABOR TRANSFORM WITH LOCALIZED BIORTHOGONAL FUNCTION

Osama A. Ahmed, and Moustafa M. Fahmy, IEEE Fellow

King Fahd University of Petroleum and Minerals, 1192
Dhahran, 31261, Saudi Arabia
mmfahmy@dpc.kfupm.edu.sa

ABSTRACT

A new implementation of the critically sampled non-periodic real Gabor transform (GT) is presented for non-separable time-frequency (TF) plane sampling. In the propsed implementation, the quincunx sampling is used to sample the TF plane. This leads to a well localized biorthogonal function in both time and frequency. It thus overcomes the main problem of the previous implementations, which is the non-localization of the resultant biorthogonal function. A fast algorithm to compute the derived biorthogonal function is proposed.

1. INTRODUCTION

Gabor suggested representing a 1-D time signals in 2-D with time and frequency as coordinates[1]. He pointed out that there is a certain elementary signal which occupies the smallest possible area in the information diagram. He proved that the modulation product of a harmonic oscillation of any frequency with a Gaussian function is the only optimal elementary signal concentrated in the joint TF domain. Gabor defined two expansions, one for complex signals and the other for real signals. There are several approaches to implement complex GT like frame theory [2], filter-bank theory[3], and biorthogonal function theory[4]. Here the biorthogonal function theory is used since it gives a clear insight of the characteristics of the GT. Unfortunately, the biorthogonal function for complex GT is not localized in both time and frequency. As a result, the TF resolution gets worse. This forced the research to go to the over-sampling case. We are interested in the critical-sampling case which happened to be the most compact representation. Besides, it is the only case where the coefficients are linearly independent. In [5], an implementation of real GT was proposed which leads to a concentrated biorthogonal function in time. This permits a truncated version of the biorthogonal function to be used in near lossless signal analysis and synthesis. Its frequency response, however, is not localized. The above implementations are based on uniform separable sampling of the TF plane to obtain discrete GT. In [6][7], a generalization for the complex GT was proposed in which the TF plane is *arbitrarily* sampled (not just into rectangular lattice or grid). One of these general TF sampling schemes which gives better results is the quincunx-lattice. In this paper we give a simple way to implement the non-separable quincunx sampling and apply it for real GT. This results in a localized biorthogonal function in both time and frequency domains. We develop the matrix structure for the quincunx sampling set and we give a computational efficient algorithm to calculate the biorthogonal function. In section II, we review the real GT and show the problems resulting from the non-localization of the biorthogonal function. The proposed method is introduced in section III. In section IV, an efficient method is given to calculate the biorthogonal function. Throughout this paper we assume that the discrete signal $x(k)$ is of length L, the number of shifts of the modulated Gaussian pulse is M, and the number of frequency components in each shift is N. Here, $MN = L$.

2. REAL GABOR TRANSFORM AND THE BIORTHOGONAL FUNCTION

Gabor expansions of real continuous signal $x(t)$ is

$$x(t) = \sum_{m}\sum_{n} h(t - m\Delta_t)[a_{m,n}\cos(nt\Delta_\omega) + b_{m,n}\sin((n+.5)t\Delta_\omega)] \qquad (1)$$

where $\Delta_t = \frac{2\pi}{\Delta_\omega}$, and $h(t)$ is the Gaussian window. In [5], discretization and reformulation of (1) led to

$$x(k) = \sum_{m=0}^{M-1} h_m(k) \sum_{n=0}^{N-1} a_{m,n}\alpha \cos\frac{\pi n(k+.5)}{N} \qquad (2)$$

where $\alpha = \sqrt{\frac{1}{N}}$ for $n = 0$ and $\alpha = \sqrt{\frac{2}{N}}$ otherwise, $a_{m,n}$ is the expansion coefficients, and $h_m(k)$ is the discrete periodic version of the Gaussian window shifted to the center of the m^{th} interval of length N. $h_m(k)$ is given by

$$h_m(k) = \left(\sqrt{2}/N\right)^{\frac{1}{2}} e^{-\pi\left(\frac{k-mN-\frac{N-1}{2}}{N}\right)^2} \tag{3}$$

According to the biorthogonal theory[8], the coefficients $a_{m,n}$ can be computed by

$$a_{m,n} = \sum_{k=0}^{L-1} x(k)\gamma^*(k-mN)\alpha\cos\frac{\pi k(n+.5)}{N} \tag{4}$$

where $\gamma(k)$ is the biorthogonal function to the analysis window function $h(k)$. Equation (4) shows that the GT is equivalent to short time Fourier transform (STFT) with the $\gamma(k)$ used as the STFT-window. It thus follows from the STFT theory that to maintain time and frequency resolution, the $\gamma(k)$ has to be also localized in both frequency and time. The implementation gives $\gamma(k)$ which is localized only in time domain, as in Fig. 1. As a result, its frequency resolution is destroyed. This

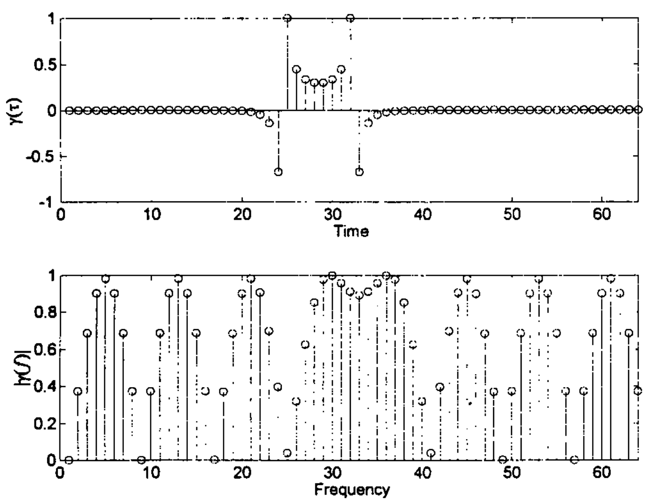

Figure 1: $\gamma(k)$ for real GT, $M = N = 8$

prohibits the practical use of real GT in many applications and results in an unstable transform.

In the next section, we will propose a generalization of this real GT for general TF plane sampling set. This proposed implementation gives a biorthogonal function that is localized in both time and frequency.

3. THE PROPOSED METHOD

In the previous section, the signal is expanded into windowed version of the discrete cosine transform (DCT-II). DCT-II is a sampled version of the discrete-space

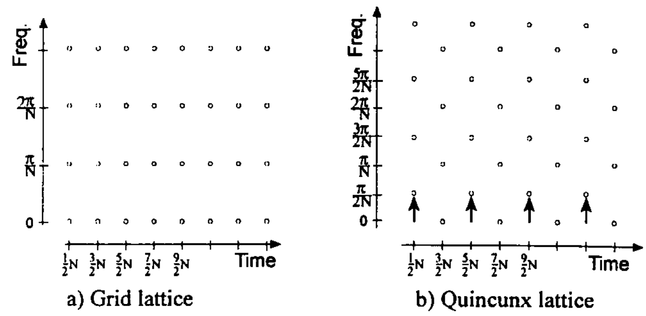

Figure 2: Two different Time-Freq. sampling set

cosine transform $C_x(\omega)$

$$C_x(\omega) = \sum_{n=0}^{\infty} 2x(k)\cos(\omega(k+.5)) \tag{5}$$

at $\omega = \frac{\pi n}{N}$ where $n = 0, 1, \ldots, N-1$. Thus, (2) is a uniform sampling of the TF plan at points $(t,\omega) = \left((m+.5)N, \frac{\pi n}{N}\right)$ where $m = 0, 1, \ldots, M-1$ as shown in Fig. 2-(a). To implement the quincunx sampling set shown in Fig. 2-(b), notice that, for every other column, i.e. for odd m, the sampling points are shifted upwards by the amount $\frac{\pi}{2N}$ i.e. the resultant discrete cosine transform is a sampled version of the discrete-space cosine transform at $\omega = \frac{\pi(n+.5)}{N}$. This is exactly what is known as the discrete cosine transform type IV or DCT-IV (see the appendix). Thus, expanding the signal according to quincunx sampling gives

$$x(k) = \sum_{\substack{m=0\\m=even}}^{M-1} h_m(k)\sum_{n=0}^{N-1} a_{m,n}\alpha\cos\frac{\pi\left(k+\frac{1}{2}\right)n}{N} + \sum_{\substack{m=0\\m=odd}}^{M-1} h_m(k)\sum_{n=0}^{N-1} a_{m,n}\beta\cos\frac{\pi\left(k+\frac{1}{2}\right)\left(n+\frac{1}{2}\right)}{N} \tag{6}$$

Here, we choose to study the nonperiodic discretization of the GT[1], and thus neither the signal nor the window are assumed to be periodic. To put (6) in matrix notation, let $\mathbf{a}$ be a vectorizd form of the expansion coefficients $a_{m,n}$ defined by $\mathbf{a} \triangleq [\mathbf{a}_0, \mathbf{a}_1 \ldots, \mathbf{a}_{M-1}]^T$ where $\mathbf{a}_m \triangleq [a_{m,0}, a_{m,1}, \ldots, a_{m,N-1}]^T$ and $\mathbf{x}$ be a vector containing the discrete signal of length L defined by $\mathbf{x} \triangleq [\mathbf{x}_0, \mathbf{x}_1 \ldots, \mathbf{x}_{M-1}]^T$ where

$$\mathbf{x}_m \triangleq [x(mN), x(mN+1), \ldots, x(mN+N-1)]^T$$

[1]For GT, the assumption of periodicity is a more radical assumption than for the Fourier transform, as it involves a periodization of the window function as well as the signal[9].

Also, let $C_1 = [c_{n,k}]_{N\times N}$ denote the N point DCT-II transform matrix with $c_{n,k}$ given by

$$c_{n,k} = \alpha \cos \frac{\pi n\,(k+.5)}{N} \qquad n,k = 0,....,N-1 \qquad (7)$$

and $C_2 = [c_{n,k}]_{N\times N}$ denote the N point DCT-IV transform matrix with $c_{n,k}$ given by

$$c_{n,k} = \beta \cos \frac{\pi\,(n+.5)\,(k+.5)}{N} \qquad n,k = 0,....,N-1 \qquad (8)$$

Based on the DCT-IV properties (appendix-A), one can write

$$\mathbf{x}_0 = \sum_{\substack{m=0\\ m=even}}^{M-1} h_m(k) C_1^T \mathbf{a}_m + \sum_{\substack{m=0\\ m=odd}}^{M-1} h_m(k) C_2^T \mathbf{a}_m \qquad (9a)$$

$$\mathbf{x}_1 = \sum_{even} h_{m-1}(k) C_1^T J \mathbf{a}_m - \sum_{odd} h_{m-1}(k) C_2 J \mathbf{a}_m \qquad (9b)$$

$$\mathbf{x}_2 = \sum_{even} h_{m-2}(k) C_1^T \mathbf{a}_m - \sum_{odd} h_{m-2}(k) C_2 \mathbf{a}_m \qquad (9c)$$

$$\mathbf{x}_3 = \sum_{even} h_{m-3}(k) C_1^T J \mathbf{a}_m + \sum_{odd} h_{m-3}(k) C_2 J \mathbf{a}_m \qquad (9d)$$

where J is the $N\times N$ raw exchange matrix

$$J = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}$$

Based on (9), rewriting (6) in matrix form gives us

$$\begin{bmatrix} H_0 C_1^T & H_{-1} C_2 & \cdots & H_{-M+1} C_2 \\ H_1 J C_1^T & -H_0 J C_2 & \cdots & -H_{-M+2} J C_2 \\ H_2 C_1^T & -H_1 C_2 & \cdots & -H_{-M+3} C_2 \\ H_3 J C_1^T & H_2 J C_2 & \cdots & H_{-M+4} J C_2 \\ \vdots & \vdots & \ddots & \vdots \\ H_{M-1} J C_1^T & H_{M-2} J C_2 & \cdots & H_0 J C_2 \end{bmatrix} \times \mathbf{a} = \mathbf{x} \qquad (10)$$

where

$$H_m \triangleq \mathrm{diag}\left(h_m(0), h_m(1), \ldots, h_m(N-1)\right)$$

Equation (10) can be written as

$$\mathbf{H}\mathbf{C}^T\mathbf{a} = \mathbf{x} \qquad (11)$$

where

$$\mathbf{H} = \begin{bmatrix} H_0 & H_{-1} & \cdots & H_{-M+1} \\ H_1 J & -H_0 J & \cdots & -H_{-M+2} J \\ H_0 & -H_1 & \cdots & -H_{-M+3} \\ H_1 J & H_2 J & \cdots & H_{-M+4} J \\ \vdots & \vdots & \ddots & \vdots \\ H_{M-1} J & H_{M-2} J & \cdots & H_0 J \end{bmatrix}$$

and

$$\mathbf{C}^T \triangleq \mathrm{diag}\left(C_1^T, C_2, \ldots, C_1^T, C_2\right)$$

Equation (11) is the synthesis transform in matrix notation. Its inverse, i.e. the analysis transform, is

$$\mathbf{a} = \left(\mathbf{C}^T\right)^{-1} \mathbf{H}^{-1}\mathbf{x} = \mathbf{C}\mathbf{H}^{-1}\mathbf{x} \qquad (12)$$

The inverse matrix $\mathbf{H}^{-1}$ has the structure

$$\begin{bmatrix} \Gamma_0 & \Gamma_1 & \cdots & \Gamma_{M-1} \\ J\Gamma_{-1} & -J\Gamma_0 & \cdots & J\Gamma_{M-2} \\ \Gamma_{-2} & -\Gamma_{-1} & \cdots & \Gamma_{M-3} \\ J\Gamma_{-3} & J\Gamma_{-1} & \cdots & J\Gamma_{M-4} \\ \vdots & \vdots & \ddots & \vdots \\ J\Gamma_{-M+1} & -J\Gamma_{-M+2} & \cdots & -J\Gamma_0 \end{bmatrix}^T$$

with $N\times N$ diagonal blocks

$$\Gamma_m = \mathrm{diag}\left(\gamma_m(0), \gamma_m(1), \ldots, \gamma_m(N-1)\right)$$

where $\gamma_m(k) \triangleq \gamma(k)|_{k=k+mN}$, where $\gamma(k)$ comprises the biorthogonal function. For Gaussian window, the resulting $\gamma(k)$ is plotted in Fig. 3. This figure shows the nice concentration of the resulting $\gamma(k)$ in both time and frequency compared with previous work.

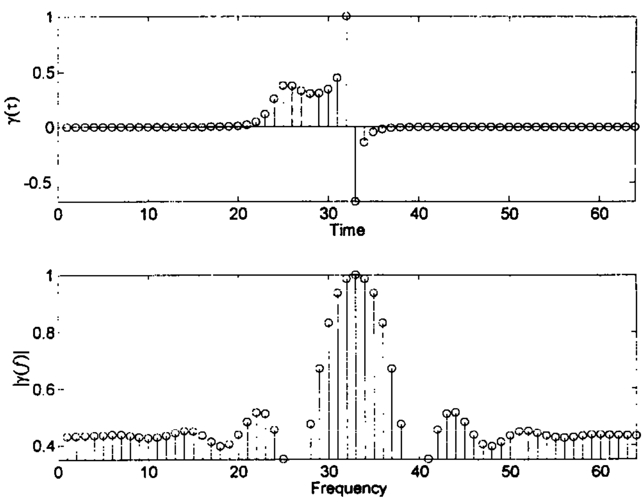

Figure 3: $\gamma(k)$ for propsed real GT, $M = N = 8$

4. CALCULATION OF $\mathbf{H}^{-1}$

The matrix $\mathbf{H}$ is an $MN \times MN$ matrix which is very costly to invert. We propose here a method to invert it utilizing its special structure. The matrix $\mathbf{H}$ is a block-Toeplitz matrix with diagonal or anti-diagonal blocks. We use row and column permutations to obtain a block diagonal matrix with N blocks. Each block is of dimension $M \times M$ and can be inverted separately as follows. Define a permutation $MN \times MN$ matrices[2] $\mathbf{P}_1$ and $\mathbf{P}_2$ whose encoding vectors[3] p_1 and p_2, for $k = 0, \ldots, MN-1$, are given by

$$
\begin{aligned}
p_1(k) &= \left\lfloor \frac{k}{M} \right\rfloor + N\,(k \bmod M) \\
p_2(k) &= l\,(N-1) + (-1)^l \left\lfloor \frac{k}{M} \right\rfloor + N\,(k \bmod M)
\end{aligned}
$$

where $l = k \bmod 2$ and $\lfloor x \rfloor$ is the integer part of x. The matrix $\mathbf{P}_2\mathbf{H}\mathbf{P}_1$ is block diagonal with $M \times M$ blocks $\mathbf{D}_n, n = 1, ..., N$. $\mathbf{D}_n$ is block-Toeplitz with 2×2 blocks. Efficient methods for inverting block-Toeplitz matrices exist [10] which take $\mathcal{O}(M^2)$, or even $\mathcal{O}(M \log M)$ as some iterative algorithms claim, i.e. the whole inversion process of $\mathbf{P}_2\mathbf{H}\mathbf{P}_1$ takes $\mathcal{O}(NM^2)$ or $\mathcal{O}(NM \log M)$. The inverse $\mathbf{H}^{-1}$ is given by

$$
\mathbf{H}^{-1} = \mathbf{P}_1 \; (\mathbf{P}_2 \; \mathbf{H} \; \mathbf{P}_1)^{-1} \, \mathbf{P}_2
$$

and the Gabor coefficient is given by

$$
\mathbf{a} = \mathbf{C} \; \mathbf{P}_1 \; (\mathbf{P}_2 \; \mathbf{H} \; \mathbf{P}_1)^{-1} \, \mathbf{P}_2 \; \mathbf{x}
$$

5. CONCLUSION

In this paper we have presented an implementation of the critically sampled real GT for nonseparable TF plane sampling. We showed that the resultant biorthogonal function is well localized in both frequency and time domains. An efficient method to calculate the biorthogonal function for any type of windows is presented.

6. APPENDIX

The Discrete Cosine Transform type IV (DCT-IV) of an N point real sequence $x(n)$, $n = 0, 1, \ldots, N-1$ is defined as:

$$
\begin{aligned}
X(n) &= \beta \sum_{n=0}^{N-1} x(k) \cos \frac{\pi\left(k+\frac{1}{2}\right)\left(n+\frac{1}{2}\right)}{N} \\
x(k) &= \beta \sum_{n=0}^{N-1} X(n) \cos \frac{\pi\left(k+\frac{1}{2}\right)\left(n+\frac{1}{2}\right)}{N}
\end{aligned}
$$

where $\beta = \sqrt{2/N}$. DCT-IV has found several applications in signal processing [11]. Several algorithms are available in the literature for efficient calculation of this transform [12]. It is easy to prove the following DCT-IV properties

$$
x(k+N) = -\beta \sum_{n=0}^{N-1} X(n) \times \cos \frac{\pi\left((N-1-k)+\frac{1}{2}\right)\left(n+\frac{1}{2}\right)}{N} \quad \text{(A1)}
$$

$$
x(k+2N) = -\beta \sum_{n=0}^{N-1} X(n) \cos \frac{\pi\left(k+\frac{1}{2}\right)\left(n+\frac{1}{2}\right)}{N} \quad \text{(A2)}
$$

$$
x(k+3N) = \beta \sum_{n=0}^{N-1} X(n) \times \cos \frac{\pi\left((N-1-k)+\frac{1}{2}\right)\left(n+\frac{1}{2}\right)}{N} \quad \text{(A3)}
$$

7. ACKNOWLEDGMENT

The authors would like to acknowledge the support of King Fahd University of Petroleum and Minerals.

8. REFERENCES

[1] D. Gabor, "Theory of communication", IEE, , no. 93, pp. 429–457, 1946.

[2] C. E. Heil and D. F. Walnut, "Contineous and discrete wavelet transforms", SIAM Rev., vol. 31, no. 4, pp. 628–666, Dec. 1989.

[3] R. A. Gopinath and C. S. Burrus, Wavelet transforms and filter banks, pp. 603–654, NY Academic, 1992.

[4] R. Orr, "Derivation of gabor transform relations using bessel's equality", Signal Processing, vol. 30, no. 2, pp. 257–262, Jan. 1993.

[5] L. C. Potter D. F. Stewart and S. C. Ahalt, "Computationally attractive real gabor transform", IEEE Trans. on Signal Proc., vol. 43, no. 1, pp. 77–83, Jan. 1995.

[6] H. G. Feichtiger and O. Christensen, "Group theoretical approach to gabor analysis", Opt. Eng., vol. 34, no. 6, pp. 1697–1703, June 1995.

[7] P. Prinz, "Calculating the dual gabor window for general sampling sets", IEEE Trans. on Signal Proc., vol. 41, no. 8, pp. 122–130, Aug. 1996.

[8] B. Bastiaans, "A sampling theorem for the complex spectrogram, and gabor's expansion of a signal in gaussian elementary signals,", Opt. Eng., vol. 20, no. 4, pp. 594–598, July 1981.

[9] N. Redding and N. Newsam, "Efficient calculation of finite gabor transform", IEEE Trans. on Signal Proc., vol. 44, no. 2, pp. 190–200, Feb. 1996.

[10] H. Akaike, "Block toeplitz matrix inversion", SIAM J. Appl. Math., vol. 24, no. 2, pp. 234–241, Mar. 1973.

[11] H. S. Malvar, Signal processing with lapped transforms, Norwood, MA: Artech House, 1991.

[12] N. R. Murthy and M. N. Swamy, "On the online computation of dct-iv and dst-iv transforms", IEEE Trans. on Signal Proc., vol. 43, no. 5, pp. 1249–1251, May 1995.

[2] Permutation matrix P is the identity matrix with rows reordered.

[3] Encoding vector p is the vector whose element $p(k)$ is the column index of the sole "1" in k^{th} row of $\mathbf{P}$.