FAST APPROXIMATE DCT: BASIC–IDEA, ERROR ANALYSIS, APPLICATIONS

Abdulnasir Hossen *Ulrich Heute*

Institute for Network and System Theory
Christian–Albrechts University
D–24143 Kiel, Germany
abh@techfak.uni-kiel.d400.de uh@techfak.uni-kiel.d400.de

ABSTRACT

The discrete cosine transform (DCT) has a variety of applications in image and speech processing. The idea of the subband–DFT (SB–DFT) [1], [2] is applied in [3] to the DCT. In this paper the basic idea of the SB–DCT is discussed which is based on subband decomposition of the input sequence. Approximation is done by discarding the computations of bands of little energy. The complexity of this fast approximate method is examined in comparing it with a fast cosine–transform method [4] in terms of program running-time. New accurate analysis of the errors due to the approximation is presented for any number of decomposition stages. New applications of the SB–DCT in speech cepstrum analysis and in echo detection are also included by using the SB–DCT instead of the full–band FFT in calculating the real and complex cepstra.

1. INTRODUCTION

The DCT of an N–point sequence $x(n)$, with $n \in \{0, 1, ..., N-1\}$ is defined as:

$$C(k) = \sum_{n=0}^{N-1} 2x(n)\cos(\frac{\pi k(2n+1)}{2N}) \quad k \in \{0, 1, ..., N-1\} \tag{1}$$

A fast cosine–transform given by Makhoul [4] can be computed for an N–point real signal by an N–point DFT of a reordered version of the original signal according to the following procedure:

1. Compute $v(n)$ from $x(n)$ using:

$$\begin{aligned} v(n) &= x(2n) \quad 0 \le n \le [\frac{N-1}{2}] \\ &= x(2N-2n-1) \quad [\frac{N+1}{2}] \le n \le N-1 \end{aligned} \tag{2}$$

   where $[a]$ denotes integer part of a.

2. Find the DFT $V(k)$ of $v(n)$.
3. Multiply $V(k)$ with $2\exp(\frac{-j\pi k}{2N})$.
4. Find the real part of the result of the above step.

For the purpose of image coding the SB–DCT is introduced in [3]. In the SB–DCT the original signal is decomposed into two frequency–bands. Depending on the signal and the application, it may be acceptable to calculate only a band of adjacent points approximately but with a higher speed. The paper is organized as follows:
In the next section the idea of the SB–DCT is introduced with its error analysis and computational complexity. Section 3 presents two new applications of the subband–DCT in speech cepstrum analysis and echo detection. Concluding remarks are given in section 4.

2. SUBBAND–DCT

2.1. Basic–Idea

In Fig.1 the length–N data sequence $x(n)$ is decomposed into two subsequences of length $N/2$:

$$\begin{aligned} g(n) &= 1/2[x(2n) + x(2n+1)] \\ h(n) &= 1/2[x(2n) - x(2n+1)], \end{aligned} \tag{3}$$

where $g(n)$ and $h(n)$ are the down–sampled versions of the

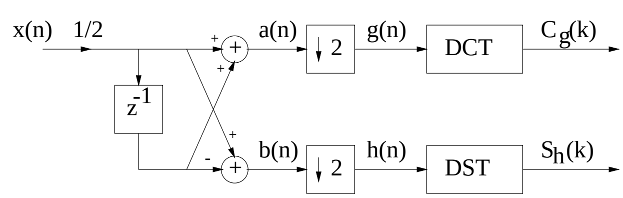

Figure 1: Two–band decomposition of the subband DCT

low–pass filtered sequence $a(n)$ and the high–pass filtered sequence $b(n)$ respectively. Eq.(1) can be written as

$$\begin{aligned} C(k) &= \sum_{n=0}^{N/2-1} 2x(2n)\cos(\pi k(4n+1)/(2N)) \\ &+ \sum_{n=0}^{N/2-1} 2x(2n+1)\cos(\pi k(4n+3)/(2N)). \end{aligned} \tag{4}$$

With a simple mathematical reformulation and with the aid of Eqs.(3) and (4) this becomes:

$$\begin{aligned} C(k) &= 2\cos(\pi k/(2N)) \sum_{n=0}^{N/2-1} 2g(n)\cos(\pi k(2n+1)/N) \\ &+ 2\sin(\pi k/(2N)) \sum_{n=0}^{N/2-1} 2h(n)\sin(\pi k(2n+1)/N) \end{aligned} \tag{5}$$

or:

$$C(k) = 2\cos(\pi k/(2N))C_g(k) + 2\sin(\pi k/(2N))S_h(k), \tag{6}$$

where $C_g(k)$ and $S_h(k)$ are the $N/2$–point DCT and DST of $g(n)$ and $h(n)$ respectively.
Eq.(6) can be approximated by calculating only the first term :

$$\hat{C}(k) = 2\cos(\frac{\pi k}{2N})C_g(k) \quad k \in \{0, 1, ..., \frac{N}{2} - 1\}. \tag{7}$$

The decomposition process can be also repeated to the sequence $g(n)$ in Eq.(3) to decompose it into two further bands, and the same procedure can be followed to get

$$\hat{C}(k) = 4\cos(\frac{\pi k}{2N})\cos(\frac{\pi k}{N})C_{gg}(k) \quad k \in \{0, 1, ..., \frac{N}{4} - 1\}, \tag{8}$$

where $C_{gg}(k)$ is the discrete cosine transform of the low–low frequency band. The two partial DCTs $C_g(k)$ and $C_{gg}(k)$ can be calculated from the respective sub–sequences using the fast cosine–transform procedure given by [4].

2.2. Approximation Errors

Repeating the decomposition m times results in $M = 2^m$ subbands. Two main types of errors appear during the approximation: linear distortions and aliasing. These approximation errors depend on the input signal and the number and type of decompositions. Combining the two Eqs.(6,7) and by relating $S_h(k)$ with $C(N-k)$, we obtain:

$$\frac{2\hat{C}(k)}{\cos(\frac{\pi k}{2N})} = C(k) - \frac{\cos(\frac{\pi(N-k)}{2N})}{\cos(\frac{\pi k}{2N})}C(N-k). \tag{9}$$

The left–hand side of this equation which will be denoted $\hat{C}_c(k)$, is the approximated transform after a compensation of the linear distortion. The second term in the right–hand side is due to the aliasing error created by non–zero transform points $C(N-k)$. For $m = 2$ decomposition stages, we obtain:

$$\begin{aligned} \hat{C}_c(k) &= B_1C(k) + B_2C(N-k) \\ &+ B_3C(\frac{N}{2} - k) + B_4C(\frac{N}{2} + k) \end{aligned} \tag{10}$$

For $N = 16$ and $m = 2$ and if only the first band is calculated, Table 1 shows the aliasing effects of points $C(N-k)$ and $C(\frac{N}{2} - k)$ and $C(\frac{N}{2} + k)$ on points $C(k)$, assuming that the ratio of the transform points causing aliasing to the true components in the calcualted band is fixed to 0.1.

Calculated Points	Effect of $C(N-k)$	Effect of $C(\frac{N}{2}-k)$	Effect of $C(\frac{N}{2}+k)$
$C(0)$	0	0	0
$C(1)$	-0.0098	-0.0155	$+0.0127$
$C(2)$	-0.0199	-0.0351	$+0.0235$
$C(3)$	-0.0303	-0.0616	$+0.0329$

Table 1: Aliasing effects of other bands on the calculated band for $N = 16$ and $M = 4$

In general, For M subbands, there are $M - 1$ aliasing terms that are uniformly distributed on the frequency axis with respect to the middle point $N/2$. For $M = 8$, Eq.10 becomes:

$$\begin{aligned} \hat{C}_c(k) &= B_1C(k) + B_2C(N-k) \\ &+ B_3C(\frac{N}{2} - k) + B_4C(\frac{N}{2} + k) \\ &+ B_5C(\frac{N}{4} - k) + B_6C(\frac{N}{4} + k) \\ &+ B_7C(\frac{3N}{4} - k) + B_8C(\frac{3N}{4} + k). \end{aligned} \tag{11}$$

The coefficients B_j of Eq.(11), which are multiplied by the transform points $C(l)$, are shown for m decomposition stages to be

$$B_j = s(j) \prod_{i=1}^{i=m} (\cos(\pi il/(2N))/\cos(\pi ik/(2N))), \tag{12}$$

where

$$\begin{aligned} s(j) &= 1 \qquad j \le \frac{M}{2} \\ &= -1 \qquad j > \frac{M}{2}. \end{aligned} \tag{13}$$

Fig.2 illustrates the normalized aliasing error $E(k)/e$, for $N = 128$ and different values of m, where only the low–pass branch is followed in all reduction stages and under the assumption of small components with equal amplitudes e outside the band of interest.

2.3. Computational Complexity

In [3] the complexity of the half–band SB–DCT (M=2) is studied in terms of the number of operations (additions and multiplications). In this section the complexity of the general narrow–band SB–DCT is examined by comparing it with a fast exact full–band DCT due to Makhoul[4]. Fig.3 shows the execution time in msec versus the number of decomposition stages m for four different numbers N of total input points. The execution–time values corresponding to $m = 0$ are the required times of running the full–band fast cosine–transforms of length N using the algorithm in [4]. The number of the calculated points is $(\frac{N}{M} = \frac{N}{2^m})$. The running–time measurements are done using a 486 processor with 66 MHz.

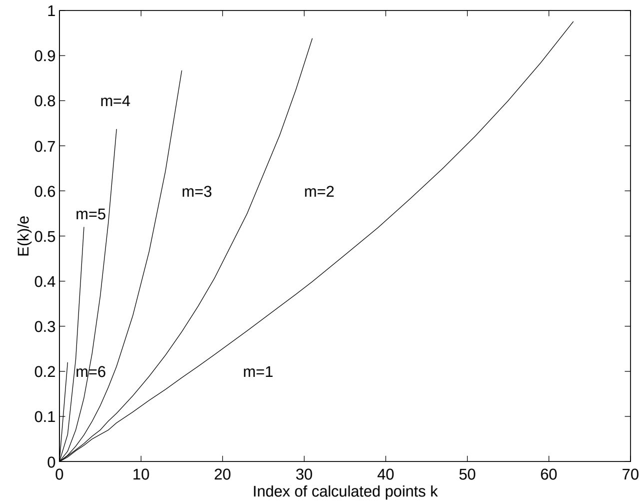

Figure 2: Normalized aliasing error for different number of stages

3. APPLICATIONS

The DFT–based real and complex cepstrum (RFC and CFC) of a signal x are defined as:

$$\begin{aligned} RFC &= Real(IFFT(ln(Abs(FFT(x))))), \\ CFC &= Real(IFFT(ln(FFT(x)))). \end{aligned} \qquad (14)$$

In [6], it is shown that using the DCT instead of the FFT does not degrade the information contained in the cepstrum while substantially reducing the computational complexity, so the DCT–based real and complex cepstra (termed RCC and CCC, respectively) according to [6] are

$$\begin{aligned} RCC &= Real(IDCT(ln(Abs(DCT(x))))), \\ CCC &= Real(IDCT(ln(DCT(x)))). \end{aligned} \qquad (15)$$

In this work the SB–DCT is used instead of the full–band DCT in computation of both the real DCT–based cepstrum and the complex DCT–based cepstrum. So Eq.(15) can be changed to

$$\begin{aligned} RSCC &= Real(IDCT(ln(Abs(SB\text{–}DCT(x))))), \\ CSCC &= Real(IDCT(ln(SB\text{–}DCT(x)))). \end{aligned} \qquad (16)$$

Comparing the last equation with the corresponding DCT–based cepstrum Eq.(15), reduction of computational complexity in both RSCC and CSCC is caused by the following facts:

1. A SB–DCT is calculated instead of a full–band DCT;
2. a smaller–size IDCT is needed instead of the full–band IDCT;
3. the ln and Abs functions are computed for smaller–size sequences.

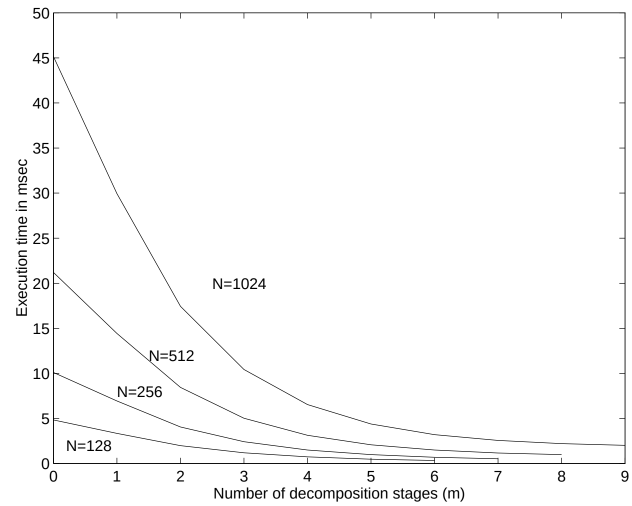

Figure 3: Running–time comparison

3.1. Approximate Speech Cepstrum

The real cepstrum of a voiced speech segment contains impulses at the multiples of the pitch period, while an unvoiced speech cepstrum contains no such impulses [5]. Fig.4 shows the RCC and the $RSCC$ for two different speech segments (voiced and unvoiced) of a signal sampled at 16 kHz. In all cases the speech signal is windowed by a Hamming window of the wanted segment size before applying the DCT or the SB–DCT. We can conlude from this figure that the SB–DCT determines correctly the mode of excitation. The pitch period is seen to be the same for the voiced segment by applying either RCC or $RSCC$.

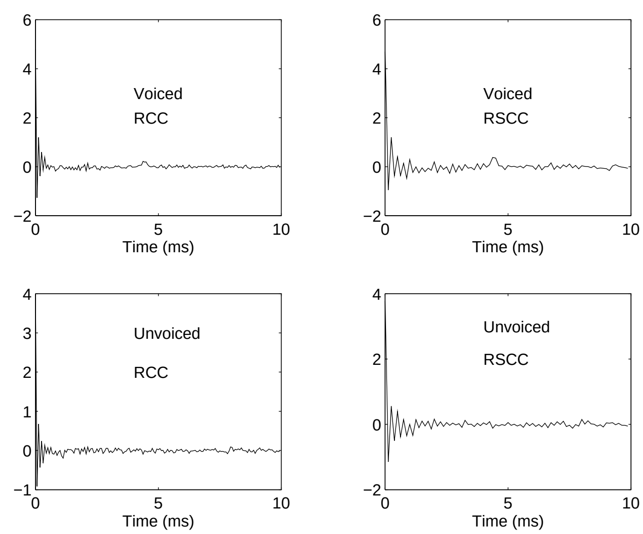

Figure 4: DCT and SB–DCT Speech Cepstra

3.2. Approximate Echo Detector

A new application of both fast full–band discrete cosine transform and SB–DCT is given in this subsection. A simulation test similar to that given by [7] for applying the complex cepstrum in echo detection is applied here for both the CCC and $CSCC$. In this test, a sine wave of frequency f is created with a sampling frequency f_s. An echo is added to the signal with an amplitude A and a position P seconds after the begining of the signal. Fig.5 shows the simulated signal for $f = 10\ Hz$ and $f_s = 200\ Hz$ with values of $A = 0.75$ and $P = 0.2$ seconds and also the results of finding the CCC and $CSCC$. The subband complex cepstrum is shown for the half–band ($M = 2$) and quarter–band ($M = 4$) cases.

Table 2 shows the efficiency of the SB–DCT cepstra in detecting echo signals in terms of the maximum signal to echo ratio or in other words in terms of the minimum detectable echo for different numbers of decomposition stage.

Method Used	Signal to Echo Ratio
Full–Band DCT	40 dB
SB–DCT ($m = 1$)	26 dB
SB–DCT ($m = 2$)	20 dB
SB–DCT ($m = 3$)	16 dB
SB–DCT ($m = 4$)	14 dB

Table 2: Maximum signal to echo ratio for correct echo detection

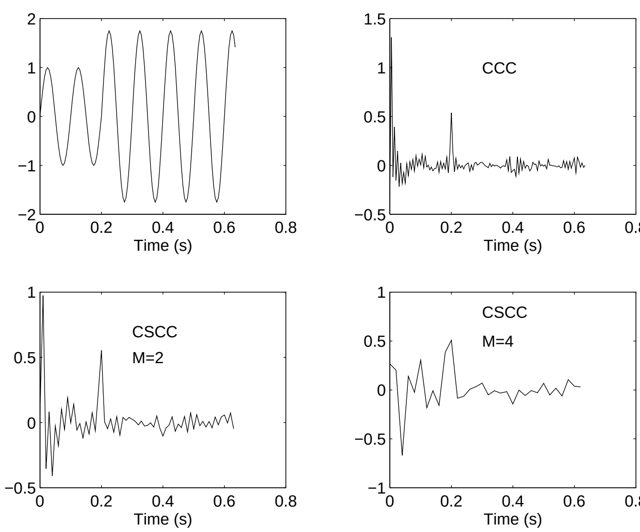

Figure 5: Echo Detection Examples

4. CONCLUSIONS

The subband–DCT method is investigated in this paper. Both linear distortion and aliasing errors occurring due to the approximation are analyzed. Aliasing–error coefficients are found for any number of decomposition stages. The SB–DCT is applied in cepstrum analysis and used as a detector for voiced/unvoiced mode of excitation and as a pitch estimator. No essential difference is found between the results of the SB–DCT cepstrum and the full–band DCT cepstrum. The approximate SB–DCT complex cepstrum is used for detecting echo signals. The algorithm shows a high efficiency for a wide range of amplitudes and positions of the echo.

5. REFERENCES

[1] Mitra, S. K.; Shentov, O.V.; and Petraglia, M. R., "A Method for Fast Approximate Computation of Discrete Time Transforms," Proc. of ICASSP'90, Albuquerque, NM, pp.2025-2028.

[2] Hossen, A. N.; Heute, U.; Shentov, O.; and Mitra, S.,"Subband DFT–Part II: Accuracy, Complexity, and Applications," Signal Processing, Vol. 41, no.3, Feb. 1995, pp.279–294.

[3] Jung, S.; Mitra, S.; and Mukherjee, D.," Subband DCT: Definition, Analysis, and Applications," IEEE Trans. on Circuits and Systems for Video Technology, Vol.6, N0.3, June 1996.

[4] Makhoul, J., " A Fast Cosine Transform in One and Two Dimensions," IEEE Transaction Vol. ASSP-28, No. 1, pp. 27-34, Feb. 1980.

[5] Oppenheim, A. V.; Schafer R. W., "Discrete–Time Signal Processing," Prentice Hall, 1989.

[6] Hassanein, H.; and Rudko, M.," On the Use of Discrete Cosine Transform in Cepstral Analysis," IEEE Trans. Vol.ASSP–32, No.4, Aug.1984.

[7] "Matlab Signal Processing Toolbox," The MathWorks 1994.