ITCT Lecture 10.3: Modified DCT : MDCT

Prof. Ja-Ling Wu

Department of Computer Science
and Information Engineering
National Taiwan University

1

MDCT is a Fourier-related transform based on the type-IV DCT, with the additional property of being lapped:

It is designed to be performed on consecutive blocks of a larger dataset, where subsequent blocks are overlapped so that the last half of one block coincides with the first half the next block.

2

- This overlapping, in addition to the entropy-compaction quality of the DCT, makes the MDCT especially attractive for signal compression applications, since it helps to avoid artifacts stemming from the block boundaries.
- MDCT is employed in MP3, AC3, and AAC for audio compression.
- MDCT also plays an important role in developing fast algorithms for DFT.

- In MP3, the MDCT is not applied to the audio signal directly, but rather to the output of a 32-band polyphase quadrature filter (PQF) bank. The output of this MDCT is post processed by an alias reduction formula to reduce the typical aliasing of the PQF filter bank.
- Such a combination of a filter banks with an MDCT is called a hybrid filter bank or a subband MDCT.

Definition: MDCT is a linear function $F : R^{2n} \rightarrow R^n$

The $2n$ real numbers $x_0, x_1, \ldots, x_{2n-1}$ are transformed into the n real numbers $f_0, f_1, \ldots f_{n-1}$ according to the formula:

$$f_j = \sum_{k=0}^{2n-1} x_k \cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(k+\frac{1}{2}+\frac{n}{2}\right)\right]$$

5

Inverse Transform: IMDCT

- Because there are different numbers of inputs and outputs, at the first glance it might seem that the MDCT should not be invertible.
- However, perfect invertibility is achieved by adding the overlapped IMDCTs of subsequent overlapping blocks, causing the errors to cancel and the original data to be retrieved; this technique is known as Time-domain aliasing cancellation (TDAC).

6

- The IMDCT transforms n real numbers $f_0, f_1, \ldots, f_{n-1}$ into $2n$ real numbers $y_0, y_1, \ldots, y_{2n-1}$ according to the formula:

$$y_k = \frac{1}{n}\sum_{j=0}^{n-1} f_j \cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(k+\frac{1}{2}+\frac{n}{2}\right)\right]$$

- IMDCT has the same form as the forward MDCT.

Computation:

- The direct application of the MDCT formula would required $O(n^2)$ operations, as in the FFT, it is possible to compute the same thing with only $O(n \log n)$ complexity by recursively factorizing the computation.

- As described below, any algorithm for the DCT-IV immediately provide a method to compute the MDCT and the IMDCT of even size.

Relationship to DCT-IV and origin of TDAC

For even n

$$f_j = \sum_{k=0}^{2n-1} x_k \cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(k+\frac{1}{2}+\frac{n}{2}\right)\right]$$

$$= \sum_{k=0}^{2n-1} x_k \cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left[\left(k+\frac{n}{2}\right)+\frac{1}{2}\right]\right], \quad (\text{Let } k' = k+\frac{n}{2})$$

$$= \sum_{k'=\frac{n}{2}}^{2n-1+\frac{n}{2}} x_{k'-\frac{n}{2}} \cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(k'+\frac{1}{2}\right)\right]$$

$$= \sum_{k=0}^{n-1} x'_k \cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right], \quad \text{where } x'_k \text{ is a folded sequence of } x_k.$$

9

- For even n the MDCT is essentially equivalent to a DCT-IV, where the input is shifted by $\frac{n}{2}$ and two n-blocks of data are transformed at once.

- Notice that

$$\cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(-k-1+\frac{1}{2}\right)\right] = \cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right]: \quad \text{even around } k=-\frac{1}{2}$$

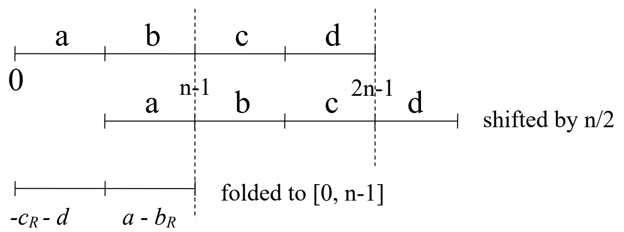
$$\cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(2n-k-1+\frac{1}{2}\right)\right] = -\cos\left[\frac{\pi}{n}\left(j+\frac{1}{2}\right)\left(k+\frac{1}{2}\right)\right]: \text{odd around } k=n-\frac{1}{2}$$

- Let X be an input array of length n and let X_R denote X in reverse order.

10

- Consider an MDCT with *2n* inputs and *n* outputs, where the *2n*-point input is divided into *4* blocks (a, b, c, d) each of size *n/2*.
- If we shift these by *n/2* (from the +*n/2* term in the MDCT definition), then (b, c, d) extent past the end of the *n* DCT-IV inputs, so we must "fold" them back according to the boundary conditions described above.

- The MDCT of *2n* inputs (a, b, c, d) is exactly equivalent to a DCT-IV of the *n* inputs: $(-c_R - d, a - b_R)$.



- In this way, any algorithm to compute the DCT-IV can be trivially applied to the MDCT.

- Similarly, the IMDCT formula above is precisely ½ of the (self) inverse DCT-IV, where the output sequence is shifted by n/2 and extended to a length *2n*.

- The inverse DCT-IV would simply give back the input ($-c_R - d$, $a - b_R$) from the above. When this is shifted and extended via boundary conditions, one obtains:

$$IMDCT(MDCT(a,b,c,d)) = (a-b_R,\ b-a_R,\ c+d_R,\ c_R+d)/2$$

(Half of the IMDCT outputs are thus redundant.)

Information Theory

13

TDAC:

- Suppose that one computes the MDCT of the subsequent, 50% overlapped, *2n* block (c, d, e, f). The IMDCT will then yield, analogous to the above ($c-d_R$, $d-c_R$, $e+f_R$, e_R+f) / 2. When this is added with the previous IMDCT result in the overlapping half, the reversed term cancel and one obtains simply (c,d), recovering the original data.

Information Theory

14

- The case of input data that extent beyond the boundaries of the logical DCT-IV causes the data to be aliased in exactly the same way that frequencies beyond the Nyquist frequency are aliased to low frequencies, expect that this aliasing occurs in the time domain instead of frequency domain. Hence the combinations $c\text{-}d_R$ are so on, which have precisely the right signs for the combinations to cancel when they are added.

Information Theory

15

References:

1. Princen and Bradley, "Analysis/synthesis filter bank design based on time domain aliasing cancellation," IEEE Trans. ASSP-34, pp. 1153-1161, 1986.
2. Johnson and Bradley, "Adaptive transform coding incorporating time domain aliasing cancellation," Speech Communication, pp. 299-308, 1987.

Information Theory

16