THE DISCRETE COSINE TRANSFORM

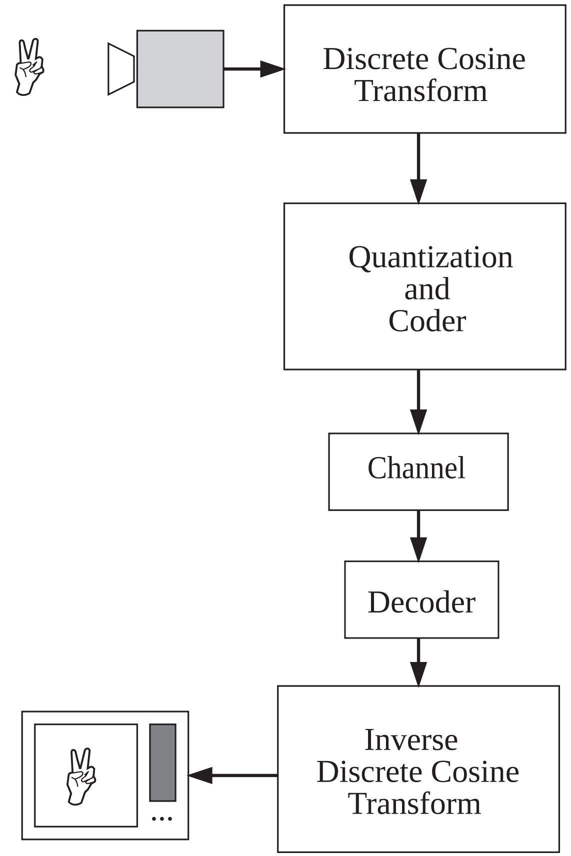
IMAGE CODING

Efficient image coding techniques are required in many applications, e.g. HDTV (High Definition TV) where the aim is to reduce the number of bits to be transmitted over the transmission channel.

Modern coding techniques can compress images by a factor of 10 to 50 without visibly affecting the image quality.

Ordinary TV signals require a transmission capacity of 216 Mbit/s of which 166 Mbit/s are allocated for the video signal.

The remaining bits are used for the sound signal and for synchronization purposes.



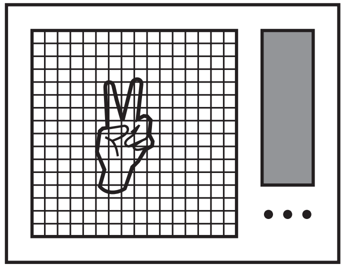
The required bit rates are roughly five times larger for HDTV.

In a typical transform coding scheme, an input image is divided into non-overlapping subframes or blocks.

The subframes are linearly transformed by using the discrete cosine transform into the transform domain.

The transform has the property that the signal is concentrated in relatively few transform coefficients compared to the number of samples in the original image.

Typical subframe sizes are in the range of 8×8 to 16×16 pixels.



DISCRETE COSINE TRANSFORMS

The *DCT (discrete cosine transform)* was first proposed by Ahmed et al. [2] in 1974.

There exist several types of DCTs: even, odd, symmetric, and the modified symmetric DCT and some more used in for example MP3.

EDCT (Even Discrete Cosine Transform also called DCT-II)

The EDCT (*even discrete cosine transform*) is defined as

$$X(k) \triangleq \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} c_k x(n) \cos\left(\frac{\pi(2n+1)k}{2N}\right) \qquad k = 0, 1, \ldots, N-1$$

The denominator of the cosine term is an even number.

The IEDCT (*inverse EDCT*) is

$$x(n) = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} c_k X(k) \cos\left(\frac{\pi(2n+1)k}{2N}\right) \qquad n = 0, 1, \ldots, N-1$$

where

$$c_k = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } k = 0 \\ 1 & \text{for } k = 1, 2, \ldots, N-1 \end{cases}$$

The forward and inverse transforms differ. Hence two different algorithms must be implemented.

Example 3.7

Determine if the basis functions of the EDCT are symmetric or antisymmetric and if the DC component leaks into the higher order frequency components. For the sake of simplicity, use a DCT with $N = 4$.

For the EDCT we get

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}}\cos\left(\frac{\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{3\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{5\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{7\pi}{8}\right) \\ \frac{1}{\sqrt{2}}\cos\left(\frac{2\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{6\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{10\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{14\pi}{8}\right) \\ \frac{1}{\sqrt{2}}\cos\left(\frac{3\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{9\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{15\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{21\pi}{8}\right) \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \end{bmatrix}$$

We get after simplification

$$\begin{bmatrix} X(0) \\ X(1) \\ X(2) \\ X(3) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}}\cos\left(\frac{\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{3\pi}{8}\right) & \frac{-1}{\sqrt{2}}\cos\left(\frac{3\pi}{8}\right) & \frac{-1}{\sqrt{2}}\cos\left(\frac{\pi}{8}\right) \\ \frac{1}{\sqrt{2}}\cos\left(\frac{2\pi}{8}\right) & \frac{-1}{\sqrt{2}}\cos\left(\frac{2\pi}{8}\right) & \frac{-1}{\sqrt{2}}\cos\left(\frac{2\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{2\pi}{8}\right) \\ \frac{1}{\sqrt{2}}\cos\left(\frac{3\pi}{8}\right) & \frac{-1}{\sqrt{2}}\cos\left(\frac{\pi}{8}\right) & \frac{1}{\sqrt{2}}\cos\left(\frac{\pi}{8}\right) & \frac{-1}{\sqrt{2}}\cos\left(\frac{3\pi}{8}\right) \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \end{bmatrix}$$

The basis functions (rows) are either symmetric or antisymmetric.

The DCT can be viewed as a set of FIR filters where the coefficients in each row represent the impulse responses. All of the filters except the first one should either be a highpass or a bandpass filter, i.e., they should have a zero at $z = 1$ in order to suppress the DC component.

In this case we do have a zero at $z = 1$, since the sum of the coefficients for all rows is zero, except for the first row.

The EDCT is suitable for image coding applications.

Department of Electrical Engineering
Linköping University

larsw@isy.liu.se
http://www.es.isy.liu.se/

ODCT (Odd Discrete Cosine Transform)

The ODCT (*odd discrete cosine transform*) is defined as

$$X(k) \overset{\Delta}{=} \frac{2}{\sqrt{2N-1}} \sum_{n=0}^{N-1} c_k c_n x(n) \cos\left(\frac{2\pi nk}{2N-1}\right) \qquad k = 0, 1, \ldots, N-1$$

The denominator of the cosine term is an odd number.

The IODCT (*inverse ODCT*) is

$$x(n) = \frac{2}{\sqrt{2N-1}} \sum_{k=0}^{N-1} c_k c_n X(k) \cos\left(\frac{2\pi nk}{2N-1}\right) \qquad n = 0, 1, \ldots, N-1$$

where

$$c_k = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } k = 0 \\ 1 & \text{for } k = 1, 2, \ldots, N-1 \end{cases}$$

The forward and inverse transforms are identical, but it can be shown that the basis functions are neither symmetric nor antisymmetric.

The DC component appears in the other components.

Hence, the ODCT is unsuitable for image coding applications.

SDCT (Symmetric Discrete Cosine Transform)

The SDCT (*symmetric discrete cosine transform*) is defined as

$$X(k) \triangleq \sqrt{\frac{2}{N-1}} \sum_{n=0}^{N-1} c_k c_n x(n) \cos\left(\frac{\pi n k}{N-1}\right) \qquad k = 0, 1, \ldots, N-1$$

The ISDCT (inverse SDCT) is

$$x(n) = \sqrt{\frac{2}{N-1}} \sum_{k=0}^{N-1} c_k c_n X(k) \cos\left(\frac{\pi n k}{N-1}\right) \qquad n = 0, 1, \ldots, N-1$$

where

$$c_k = \begin{cases} \dfrac{1}{\sqrt{2}} & \text{for } k = 0 \text{ or } k = N-1 \\ 1 & \text{for } k = 1, 2, \ldots, N-2 \end{cases}$$

Only one algorithm needs to be implemented.

The rows and columns are either symmetric or antisymmetric.

The DC component will appear in some of the higher frequency components, thus rendering the SDCT unsuitable for image coding applications.

MSDCT (Modified Symmetric Discrete Cosine Transform)

The MSDCT (*modified symmetric discrete DCT*) was developed by Sikström *et al*. Not the same as used in MP3”

Its distinguishing feature are:

The basis vectors are symmetric or antisymmetric

The DC component is suppressed at the expense of a slightly non-orthogonal transform.

The forward and inverse transforms are identical.

The MSDCT (*modified symmetric discrete cosine transform*) is defined as

$$X(k) \triangleq \sqrt{\frac{2}{N-1}} \sum_{n=0}^{N-1} c_n x(n) \cos\left(\frac{\pi n k}{N-1}\right) \qquad k = 0, 1, \ldots, N-1$$

The IMSDCT (*inverse MSDCT*) is

$$x(n) \overset{\Delta}{=} \sqrt{\frac{2}{N-1}} \sum_{k=0}^{N-1} c_k X(k) \cos\left(\frac{\pi n k}{N-1}\right) \qquad n = 0, 1, \ldots, N-1$$

where

$$c_k = \begin{cases} \frac{1}{2} & \text{for } k = 0 \text{ or } k = N-1 \\ 1 & \text{for } k = 1, 2, \ldots, N-2 \end{cases}$$

It is necessary to implement only one algorithm.

The odd rows (basis vectors) are symmetric and the even rows are anti-symmetric

Two-Dimensional DCTs

The *two-dimensional MSDCT* is defined as

$$X(p, q) \triangleq \frac{2}{N-1} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} c_n c_k x(n, k) \cos\left(\frac{\pi np}{N-1}\right) \cos\left(\frac{\pi kq}{N-1}\right), \; p, q = 0, 1, \ldots, N-1$$

where

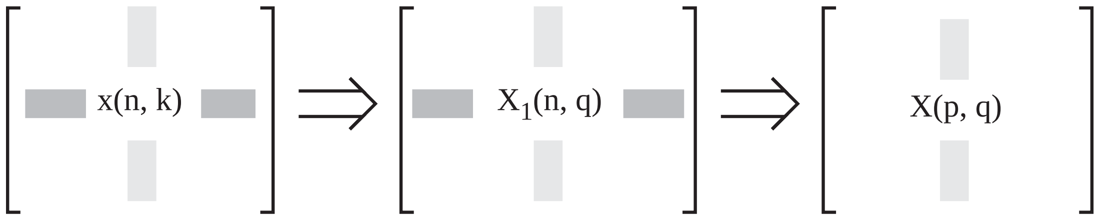
$$c_k = \begin{cases} \frac{1}{2} & \text{for } k = 0 \text{ or } k = N-1 \\ 1 & \text{for } k = 1, 2, \ldots, N-2 \end{cases}$$

First, an intermediate data array is computed using N steps by computing one-dimensional DCTs of the rows, i.e.,

$$X_1(n, q) = \sqrt{\frac{2}{N-1}} \sum_{k=0}^{N-1} c_k x(n, k) \cos\left(\frac{\pi kq}{N-1}\right), \; n, q = 0, 1, \ldots, N-1$$

The final result is obtained by computing the DCTs for the columns in the intermediate array.

$$X(p, q) = \sqrt{\frac{2}{N-1}} \sum_{n=0}^{N-1} c_n X_1(n, q) \cos\left(\frac{\pi n p}{N-1}\right), \quad p, q = 0, 1, \ldots, N-1$$



$2N$ one-dimensional DCTs and one matrix transposition of the intermediate data array are needed. A computation, based on one-dimensional DCTs, requires $2N^3$ multiplications per block.

A direct computation requires $2N^4$ multiplications per block.

Fast Discrete Cosine Transforms

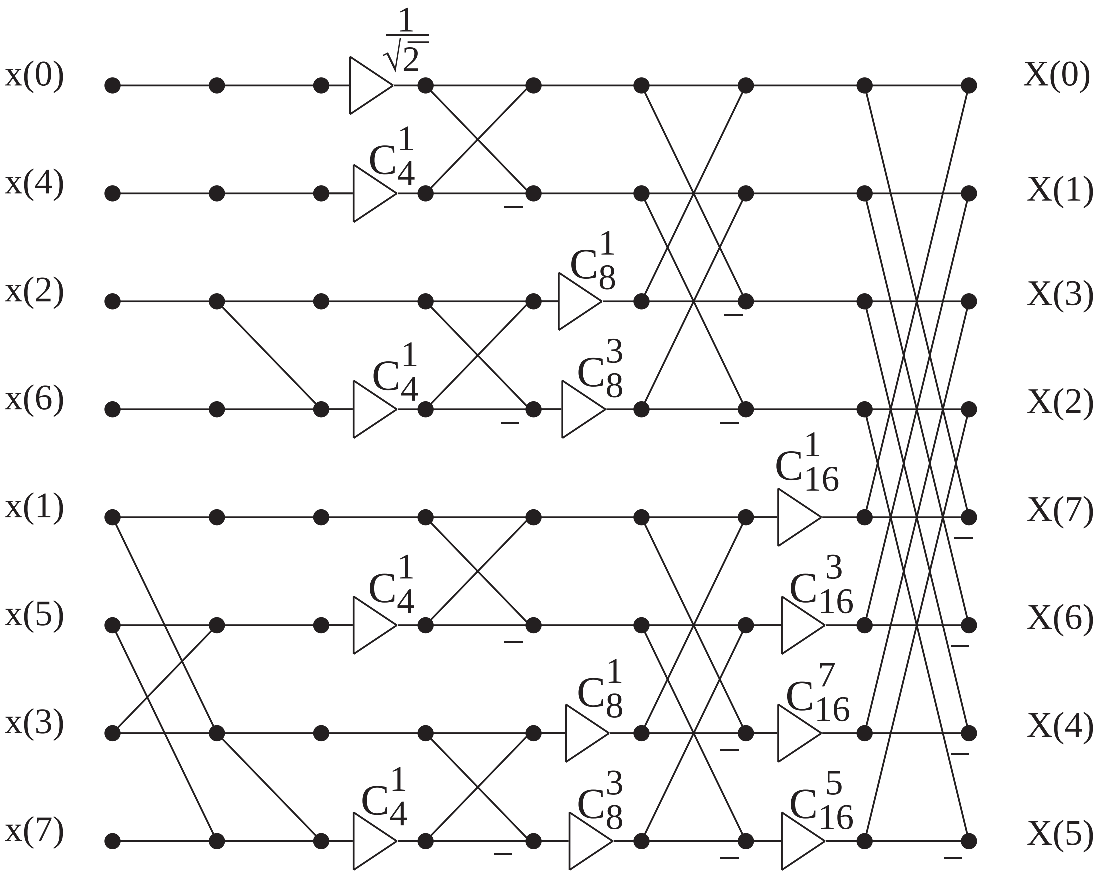
Fast algorithms with a computational complexity of $O(N \log_2(N))$, have also been developed for the DCTs by using the divide-and-conquer principle.

However, in practice, the asymptotic complexity is of little interest for image coding applications since only short sequences with $N = 4$ to 32 are used.

Fortunately, the fast algorithms that have been developed are also efficient for short DCTs.

The number of real multiplications and additions (subtractions) for the fast DCT/IDCT are $0.5N \log_2(N)$ and $1.5N \log_2(N)–N+1$, respectively.

In 1991, a fast algorithm for the 2-D even DCT was proposed by Cho and Lee [10]. The algorithm is regular and modular and requires only $0.5N^2 \log_2(N)$ multiplications and $2.5N^2 \log_2(N)–2N+2$ additions.



$$C_{2N}^{(2n+1)k} = \frac{1}{2\cos\left(\frac{\pi(2n+1)k}{2N}\right)}$$

DCT PROCESSOR — CASE STUDY 2

The two-dimensional DCT processor is assumed to be a part of a larger integrated circuit and in this study we therefore neglect the I/O aspects of the processor.

The throughput requirement for such a DCT processor varies considerably between different transform coding schemes and applications. Currently there is no standard established for HDTV.

One of the candidates is called *CIL* (common image lattice).

The color signals, (R, G, and B) are mapped linearly to a luminance, (Y), and two chrominance signals, (U and V).

The discrete cosine transform is computed for these signals individually.

Most designs today are based on an 8×8 point DCT due to the stringent requirements.

The required number of 8×8 DCTs/s for the most stringent HDTV requirement with the largest screen size is about $1.95 \cdot 10^6$.

Field frequency [Hz]	60	50
Active lines/frame	1080	1152
Samples/active line	1920	2048
Sample frequency	74.25	74.25
Total number of lines	1125	1250
Samples/line	2200	2376
Scanning algorithm	2:1	2:1
Bits/sample	8	8
Gross bit rate [Mbit/s]	1188	1188
Active bit rate [Mbit/s]	995.3	943.72
Pixels/s [MHz]	124.4	118
16 x 16 DCTs/s	486000	461000

Specification

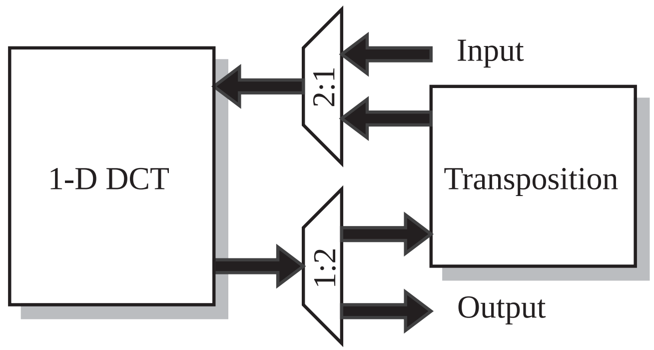
To meet these stringent requirements, the processor should be able to compute 486,000 two-dimensional (16×16) DCTs/s.

Typically, input data word lengths are in the range 8 to 9 bits and the required internal data word length is estimated to be 10 to 12 bits. The coefficient word length is estimated to be 8 to 10 bits.

System Design Phase

The 2-D DCT is implemented by successively performing 1-D DCTs.

In the first phase, the input array with 16 rows is read rowwise into the processing element, which computes a 1-D DCT, and the results are stored in the RAM.



In the second phase, the columns in the intermediate array are successively computed by the processing element.

This phase also involves 16 1-D DCT computations. Thus, 32 time units are required.

Theoretically, a transposition of the intermediate array is required, but in practice it can be realized by appropriately writing and reading the intermediate values.

The required number of arithmetic operations per second for the 2-D DCTs is

$$486{,}000 \cdot (16 + 16) \cdot 16 = 248.8 \cdot 10^6 \text{ Mult./s}$$

and

$$486{,}000 \cdot (16 + 16) \cdot 15 = 233.3 \cdot 10^6 \text{ Add/s}$$

Department of Electrical Engineering
Linköping University
larsw@isy.liu.se
http://www.es.isy.liu.se/