

Low Power and Area Efficient DCT Architecture for Low Bit Rate Communication

Abstract. In this paper a low power and area efficient DCT (Discrete Cosine Transform) pipelined architecture using multiplier-less method is presented for low bit rate communications such as videoconferencing in mobile devices. The multiplier-less multiplication is implemented by minimum number of additions, subtractions and shifts using CSD (Canonic Signed Digit) representation for fixed point DCT coefficient. Power reduction is achieved by minimizing both the number of arithmetic operations and data-path width. The proposed DCT architecture was implemented on a XILINX FPGA (Field-Programmable Gate Array). The results from power estimation show that our design is capable of reducing the power dissipation 5.5 times compared to the other DCT architectures for video streaming/video conferencing in portable devices.

Streszczenie. W artykule zaprezentowano system komunikacji DCT (Discrete Cosine Transform) o małej przepływności bitów charakteryzujący się małą mocą, dobrym pokryciem obszaru. System ma strukturę potokową. Ograniczenie mocy osiągnięto przez zmniejszenie operacji matematycznych oraz szerokości ścieżki danych. System zaimplementowano z wykorzystaniem elementów FPGA. (Architektura komunikacji DCT o małej mocy do małej przepływności danych)

Keywords: Discrete Cosine Transform (DCT), Canonical Signed Digit (CSD), Multiplier-less

Słowa kluczowe: komunikacja DCT, przepływność danych

Introduction

Many of today multimedia applications such as videoconferencing, internet video streaming and video-over wireless are most bandwidth consuming modes of communication. Efficient hardware is needed for these types of applications. The key features of hardware design are to consume very low power and low area to implement in portable small devices such as mobiles phones or cameras.

Several international standards of image and video coders use compression techniques based on DCT which transform it in frequency domain. DCT algorithm has excessive numbers of multiplication and addition operations. In hardware implementation multiplication is costly and consumes large amount of power. To reduce the multiplication operation some mathematical manipulation techniques are used such as row-column efficient matrix form of DCT coefficient which reduces the complexity of the algorithm. In hardware fixed point constant multiplication is efficiently implemented by CSD. Furthermore, to decrease the arithmetic operations common sub-expression elimination technique is also used.

The proposed architecture is designed in such a way that the architecture consumes minimum power to compute the desired results with efficiently using the hardware resources. Pipelining is used to increase the frequency and throughput of the design. Hardware optimization is used to make resources more suitable and consumes less power with less complexity. The proposed design is implemented on XILINX FPGA. The comparison results show that proposed architecture proves to be efficient compared to other published results.

This paper is organized as follows. First section gives an overview of the video coding standards. Second section introduces the background of the DCT algorithm with efficient matrix approach. Third and Fourth sections briefly describe the CSD multiplication and common sub-expression elimination respectively. Fifth section presents the design of the proposed architecture and comparison results. Finally, conclusions are presented in last section.

Overview on Video Coding Standards

In emerging multimedia applications such as image and video compression, require the use of coders. There are many different video coding standards for low bit rate

communications such as H.261, H.263, H.263+, H.264 and many more [1, 2]. Normally for internet video streaming (@ 20-200 kbps) uses H.263, for video conferencing (@ 20-320 kbps) uses H.261 or H.263, and for video over 3G wireless (@ 20-100kbps) uses H.263 video standards.

Video is composed of frames of pictures, so the still image coding has to process one more dimension for video signals compression. To de-correlate the blocks of original pixels, DCT has been widely used in many multimedia compression international standards. Some low bit rate international video coding standards are given in Table 1.

Table 1. Comparison of Resources Usage

Coding Schemes	Bit Rate (kbps)	Applications
H.261 / H.263	20-320	Video Telephony/Video Conferencing
H.263 / H.264	20-100	Video over 3G Wireless
H.264 / H.264	20-200	Internet Video Streaming

Discrete Cosine Transform (DCT)

The two dimensional DCT is broadly used in still image compression techniques. It is widely recognized that performing DCT over 8 x 8 blocks represents a good trade-off between transform complexity and spatial correlation. The 8 x 8 2-D DCT transforms an 8 x 8 block sample from spatial domain $f(x,y)$ into frequency domain $F(k,l)$. DCT is defined mathematically as in Eq (1):

$$(1) \quad F(k,l) = \frac{1}{4} C(k)C(l) \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \cos \left[\frac{\pi k_x}{16} \right] \cos \left[\frac{\pi l_y}{16} \right]$$

$$k_x = (2x+1)k$$

$$l_y = (2y+1)l$$

$$C(k) = C(l) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } k,l=0 \\ 1 & \text{otherwise} \end{cases}$$

and k, l varies from 0 to 7

For direct computation of 2-D DCT according to the definition, it needs 4096 multiplications and 4032 additions. This computational complexity can be reduced by using the row-column approach [3]. In this row-column technique, 2-D DCT can be implemented by two successive 1-D DCT. First 1-D DCT is performed in row-wise and the second 1-D DCT is performed in column wise on the output of the first 1-D DCT. Eq (2) yields the mathematical expression of the 1-D DCT.

$$F(k) = \frac{1}{4} C(k) \sum_{x=0}^7 f(x) \cos \left[\frac{\pi k x}{16} \right]$$

$$(2) \quad k_x = (2x+1)k$$

$$C(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } k = 0 \\ 1 & \text{otherwise} \end{cases}$$

and k varies from 0 to 7

In this paper, the efficient DCT coefficient matrix form is used for computational and architectural optimizations [4]. The 8 x 8 DCT coefficient matrix form can be represented as:

$$\begin{bmatrix} c_4 & c_4 \\ c_1 & c_3 & c_5 & c_7 & -c_7 & -c_5 & -c_3 & -c_1 \\ c_2 & c_6 & -c_6 & -c_2 & -c_2 & -c_6 & c_6 & c_2 \\ c_3 & -c_7 & -c_1 & -c_5 & c_5 & c_1 & c_7 & -c_3 \\ c_4 & -c_4 & -c_4 & c_4 & c_4 & -c_4 & -c_4 & c_4 \\ c_5 & -c_1 & c_7 & c_3 & -c_3 & -c_7 & c_1 & -c_5 \\ c_6 & -c_2 & c_2 & -c_6 & -c_6 & c_2 & -c_2 & c_6 \\ c_7 & -c_5 & c_3 & -c_1 & c_1 & -c_3 & c_5 & -c_7 \end{bmatrix}$$

where

$$(3) \quad c(i) = \frac{1}{2} \cos \left(\frac{i\pi}{16} \right), \quad i = 1, 2, \dots, 6, 7$$

Since the coefficient matrix having constant values so minimum number of arithmetic operations will be possible on using CSD representation.

Multiplier-Less Implementation of Multiplication using Canonical Signed Digit

Multiplier-less method is widely used for VLSI (Very Large scale Integration) realization because of improvement in speed, area overhead and power consumption [5-6]. In normal multipliers-less multiplication of a constant number is implemented just by shift and add operations. In this method the number of adders are directly proportional to the number of ones '1' in a constant. The proposed scheme especially incorporates CSD method for more efficient hardware usage and reduces the hardware complexity significantly in multiplier-less implementation of DCT. In CSD form the numbers of ones are lesser than and equal to the normal binary representation of a number. Therefore CSD method decreases the number of add/sub operations

[7]. CSD form notation is:

$$(4) \quad S = \sum_{i=0}^{N-1} a_i 2^{-i}$$

Here a_i is in the set $\{-1, 0, 1\}$ for each i .

The DCT coefficients are implemented in fixed point CSD representation. According to IEEE 1180-1990, 12-bit precision is used in order to confirm the accuracy specifications [6]. CSD expression of the cosine coefficients are shown in Table 2.

Table 2. 12-bit Precision DCT Coefficient in CSD Form

Coefficient	Decimal Value	CSD Representation
c_1	0.4904	0.i00000101000
$-c_1$	-0.4904	0.T0000010i000
c_2	0.4619	0.i00010100100
$-c_2$	-0.4619	0.T00010i00100
c_3	0.4157	0.i0101010i010
$-c_3$	-0.4157	0.T0i0101010i0
c_4	0.3536	0.i0101010i000
$-c_4$	-0.3536	0.T0i0i0101000
c_5	0.2778	0.0i00i001000i
$-c_5$	-0.2778	0.0T00T00i000T
c_6	0.1913	0.0i01000i000T
$-c_6$	-0.1913	0.0T0i000T000i
c_7	0.0975	0.00i0100i000T
$-c_7$	-0.0975	0.00T0i00T000i

In this manner maximum addition latency is three for computing any multiple of fixed point DCT coefficient.

Common Sub-Expression Elimination

The fixed coefficients of a DCT in CSD format have some common sub-expressions. Common sub-expression means that some of the bit patterns occur more than once in any expression. Close observation on the fixed coefficients extracts some common sub-expressions which can be easily eliminated. For example, a bit pattern of $(10\bar{1})_{\text{CSD}}$ in coefficient c_3 comes twice so it is not needed to compute this two times. It is achieved just by computing it one time and using its shifted result. This characteristic makes it possible to reduce the number of arithmetic operations and the hardware complexity. c_3 is computed as:

$$t_1 = x \gg 1 - x \gg 3$$

$$t_2 = x \gg 5 + x \gg 7$$

$$c_3 = t_1 + t_2 + t_1 \gg 8$$

The given example shows that the common sub-expression technique eliminates one arithmetic operation. The first operation result is re-used with some shift in computing the final result. It provides the maximum addition latency of 3.

Proposed DCT Architecture

The proposed DCT architecture is inherently low power and low area because of minimum required arithmetic operations, therefore making it highly suitable to compact mobile applications e.g. video streaming, video conferencing and online video-gaming. The 5-stage pipelined structure is pretty simple, composed of 4 barrel shifters, 5 adders/subtractors and 1 multiplexer, shows in Fig. 1.

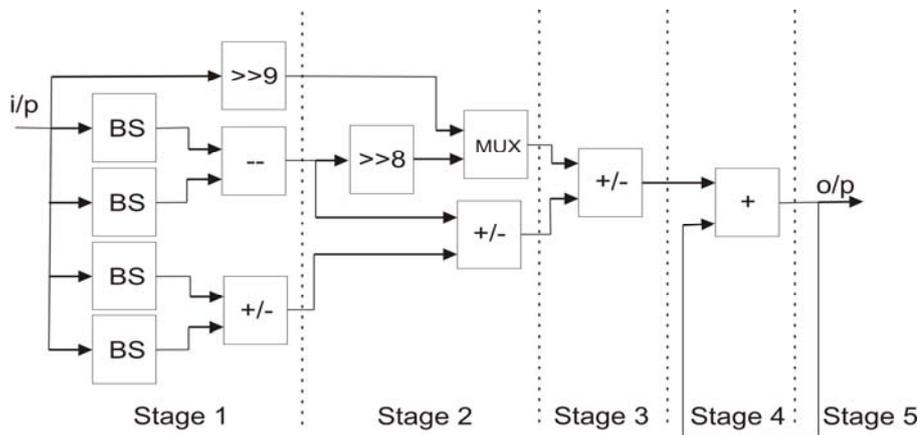


Fig. 1. Proposed DCT Architecture

First three stages are performing the multiplication operation. The 4th stage adder is adding the eight multiplication values which make one resultant value of the 1D-DCT. The barrel shifter shifts the input within a same clock cycle so there is no wasting of clock cycle which increases the overall speed of the architecture design. These barrel shifters perform specific shifting which occupies lesser area than the conventional barrel shifter. Pipelining is also increasing the overall speed of the architecture by decreasing the clock period.

This design uses different bit width for different stages resulting in a low power and low area architecture. In portable devices the major concern is area and power and this architecture is highly efficient related to these two features. Comparison of resources usage is given in Table 3. The proposed architecture uses no multiplier. Multiplier consumes more area than the adders so this design uses resources more efficiently.

Table 3. Comparison of Resources Usage

	[8]	Proposed
Adds/Subs	5	5
Multipliers	2	0

The comparison results with some other implementations are given in Table 4. In the proposed design the major improvements versus other architectures are in terms of power consumption that is most important requirement for mobile applications and other portable devices applications. Number of Configurable Logic Blocks (No. of CLB) is also very less which reduces the area of the design. The maximum frequency of the design is 249.719 MHz.

Table 4. Comparison Results for 1-D DCT Core Implementation

Design	No. of CLB	Min Clk period(ns)	No. of cycles	Power Cons. (mW)
DAA	302	17	10	222
DSFG	599	23	9	470
Systolic	1088	40	8	334
[8]	504	12	7	186
Proposed	44	4	64+4	34

Acronyms in Table 4. are for Distributed Arithmetic Architecture (DAA) and Digital-Serial Flow Graph (DSFG).

This design after place and route process can operate at maximum frequency of 242 MHz. At this frequency it

enables an encoding rate of up to 68 PAL (Phase Alternating line) frames (720 x 576 pixels) per second and 280 CIF (Common Intermediate Format) frames (352 x 288 pixels) per second. However PAL frames are too large for portable device LCD, normally small frames such as CIF standards are used. Table 5. illustrates the dynamic power of the proposed design on different operating frequencies. The dynamic power has been obtained from Xilinx XPOWER tool.

Table 5. Results of Dynamic Power of the Proposed Design Operated at Different Frequencies

Frequency (MHz)	PAL (720x576) pixels frames/s	DCIF (528x384) pixels frames/s	CIF (352x288) pixels frames/s	Dynamic Power (mW)
242	68	140	280	5.47
200	56	116	232	4.52
150	42	87	174	3.39
106	30	61	123	2.40
44	---	25	51	0.99
22	---	---	25	0.50

The results show that this design consumes very low dynamic power. The quality of the image/video will remain high at low operating frequencies too as they are satisfying the coder standards requirements (frames throughput). These characteristics show that this architecture is very suitable in compact devices.

Conclusion

In this paper a low-complexity, power and area efficient pipelined DCT architecture is proposed which uses multiplier-less method for low bit rate applications. The proposed design uses CSD representation with common sub-expression elimination to achieve minimum number of addition, subtraction and shifting operations. This hardware uses minimum logic utilizing mere 44 CLBs thus achieving the minimum dynamic power of 0.50mW when operating at frequency of 22 MHz. Its low power consumption can be used in any portable devices or in mobile applications. It can be very effectively applied in H.261, H.263 or H.264 video coding standard schemes for internet video streaming, videoconferencing and many other video applications.

Acknowledgements

The authors would like to thank the COMSATS Institute of Information Technology, Islamabad, Pakistan, for using its labs and providing relative research material.

REFERENCES

- [1] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, Rate-Constrained Coder Control and Comparison of Video Coding, IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, pp. 688-703, 2003.
- [2] International Telecommunication Union, "ITU-T H.263 Video coding for low bit rate communication", ITU-T H-series recommendations, 1998.
- [3] M. El Aakif, S. Belkouch, N. Chabini and M. M. Hassani, Low Power and Fast DCT Architecture using Multiplier-Less Method, IEEE Faible Tension Faible Consomation, pp. 63-66, 2011.
- [4] K. Parhi, VLSI Digital Signal Processing Systems, New York: Wiley, 1999.
- [5] B. Kim, S. G. Ziavras, Low-power multiplierless DCT for image/video coders, IEEE 13th International Symposium on Consumer Electronics, pp. 133-136, 2009.
- [6] A. M. Shams, A. Chidanandan, W. Pan, M. A. Bayoumi, Neda: a low-power high-performance DCT Architecture, IEEE Transactions, Signal Processing, vol. 54, pp. 955-964, 2006.
- [7] R. Hartley, IEEE International Symposium on Circuits and Systems, vol. 4, pp.1992-1995, 1991.
- [8] G. Masera, A. Molino, G. Piccinini, and M. Zamboni, A Parametric DCT architecture for H.263+ mobile terminals video streaming, IEEE Midwest Symposium on Circuits and Systems, vol.2, II-91 - II-94, 2002

Authors: Prof. Ph.D. Raja A. Riaz, B.Sc. M. Faisal Siddiqui, and MSc. Syed Saud Naqvi, Electrical Engineering Department, COMSATS Institute of Information Technology, Park Road, Chak Shehzad Campus, 44000, Islamabad, Pakistan, email: rajaali@comsats.edu.pk, faisal_siddiqui@comsats.edu.pk saud_naqvi@comsats.edu.pk, <http://ciit.isb.edu.pk>