VLSI Implementation of Discrete Cosine Transform Approximation Recursive Algorithm

M.Deivakani¹, S.V.Sudheer Kumar², Naluguru Udaya Kumar³, E. Fantin Irudaya Raj⁴, V.Ramakrishna⁵

¹ PSNA College of Engineering and Technology, Dindigul, Tamilnadu, India, Email: mdeivakani82@gmail.com,
⁲, ⁵ SOET, Sri Padmavati Mahila Visvavidyalayam, Tirupati India, Email: 2 sudheer.ssk39@gmail.com, 5 ramakrishna.bharan@gmail.com
³ Marri Laxman Reddy Institute of Technology and Management, Hyderabad, India, Email: joyudaya@gmail.com
⁴ Dr. Sivanthi Aditanar College of Engineering, Tuticorin, Tamilnadu, India, Email: fantinraj@gmail.com
Corresponding author email: ramakrishna.bharan@gmail.com

Abstract. In general, the approximation of Discrete Cosine Transform (DCT) is used to decrease computational complexity without impacting its efficiency in coding. Many of the latest algorithms used in DCT approximation functions have only a smaller DCT length transform of which some are non-orthogonal. For computing DCT orthogonal approximation, a general recursive algorithm is used here, and its length is obtained using DCT pairs of length N/2 of N addition cost in input pre-processing. The recursive sparse matrix has been decomposed by using the vector symmetry from the DCT basis in order to achieve the proposed approximation algorithm that is highly scalable to enforce the highest lengths software and hardware by using a current 8-point approximation to obtain a DCT approximation with two-length power, N>8.

Keywords: 1-D DCT, DCT-SFG, VLSI, CADENCE, Recursive, Fusion

1. Introduction

DCT is used as a sum of cosine functions that oscillate with different frequencies to express data points in a finite sequence. It has wide-ranging applications in physics, engineering such as MP3 compressing audio and JPEG image compression with cosine compared with sine functions, but few cosine functions can be used for signal approximation in order to efficiently express boundary conditions for differential equations. DCT transform is similar to Fourier, and the Discrete Fourier transform (DFT), is used in reference to real numbers where DCT is associated with Fourier series coefficients with symmetrical and periodic chain, where DFT only extends periodically. DCT is almost twice the DFT’s length and fits with even actual data symmetry. Four are common forms of eight regular forms of DCT. Patil, U et al[1], showed pyramid decomposition, multi-state and multi-resolution processing image algorithms with hierarchic PCA. Patil, U et al. PCA is used to minimise dimension and extract image fusion functions. Qiang Wang et al. was clarified about the image fusion for different applications in image processing and also examined its efficiencies based on the results of these image fusion systems, by classifying them as arbitrary fusion, total fusion and Hierarchical
structure. Sruthy, S et al[2]. also proven that the image fusion technique is used by keeping the complete information from all images and its basic features as a high quality image to merge two or more image details in one image altogether. Desale, R.P et al[3] showed that image fusion, as a mixing process, is completely insightful as an image input by mixing the information from the group of images and resulting fusion images, as well as describe the formulation of flowchart and PCA algorithms, DCT and DWT-based image fusion procedures and listed results for comparison. DCT and PCA are normal inconvenient function methods, but DWT is more useful because it produces the most efficient image fusion results.

Prakash, C et al[4], has mentioned as image fusion is a basic route in which various images are combined as one resulted fused image which consists of higher information than input image and useful in medical images for diagnosing diseases. The medical image quality explained by Aribi, W et al[5]. is calculated with numerous technologies and the fusion information with multiple methods of imaging enables very detailed study.. Haozheng Ren et al[6] explored the fusion of images into a single image, suitable for machine or human vision used in further image processing, targeting, and wavelet- dependent data fusion. Yijian Pei et al[7] has presented a fusion algorithm which used an improved wavelet and obtained improved performance with high-frequency sub-band image region characteristics. Hai Huang and Liyi Xiao[8] proposed 8-point DCT structural design of high throughput, sophisticated modularity, less hardware complexity with good synchronization. Pragati Dahiya and Priyanka Jain[9] suggested an algorithm, which presents that if the input data of long length is processed truncation error is reduced because of the less computational cycles. Maurizio Masera et al[10], exploited folded and full parallel 2D DCT architectures and the from simulation results proposed design is having less distortion losses with low complexity. Algorithms used to synthesize useful information from each image obtained from the multi sensor in an efficient manner in which multi-focus and medical image fusion verified that the presented algorithm has more effective in image fusion and summarized by analysing the performance of the algorithm.

This paper is organized as section 2 discuss about the overview of DCT algorithm design, the next section 3 presents the design implementations and section 4 presents the simulation results and finally given the conclusion of the proposed design.

2. DCT Algorithm

This section started with the 1D DCT and also discusses the design overview.

2.1. 1-D DCT

\[ T_X = \frac{1}{2} (T_0 + (-1)^X T_{M-1}) + \sum_{m=0}^{M-2} T_m \cos \left( \frac{n}{M-1} m \right) \cos (n+1) \frac{1}{2} ] \quad X = 0, 1, \ldots, N-1 \]  

In the above equation, multiply both \( T_0 \) and \( T_{M-1} \) terms by \( \sqrt{2} \) and also with \( T_0 \) and \( T_{M-1} \) with \( \frac{1}{\sqrt{2}} \) which forms matrix DC-I as orthogonal. Again multiply with scale factor; it breaks relation with real-even DFT. If DCT-I equals with scale factor 2 to DFT with real numbers of even symmetry where DCT-I for M less than 2 cannot be defined. The boundary conditions for \( T_m \) is even around \( m=0 \) also even around \( m=M-1 \) which is similarly for \( T_x \).

2-D DCT is popular for various image compression applications because of close optimal performance as compared with statistical transform. 2-D DCT is used for high speed, throughput, and short-latency computing architecture, because of high computation need, 2D-DCT processor [11] must be designed with a non-overlapping block like 16x16 or 8x8. Several 2D-DCT algorithms are developed for achieving reduced computational complexity thereby increases the speed of operation also throughput. 2D-DCT is assumed as the combination of two 1D-DCT along with each dimension. The basic definition is \( T(1,n) = (1) \)

\[ T_x = \sum_{m=0}^{M-1} T_m \cos \left( \frac{n}{M} (m+\frac{1}{2}) \right) \quad x = 0, 1, \ldots, M-1 \]
Where \( T(p,q) \) is pixel or value of image element located at the co-ordinates \( p,q \). \( T(1,n) \) is the 2D-DCT coefficient at \( (1,n) \) and can be calculated using above equation which uses values \( 1,n = 0,1,2,3 \ldots \), and the 2-D DCT multiplication base functions are defined.

Consider 8x8 image block at location \((0,0)\), squared frequency components gives the 2D-DCT product which contains DC coefficient \( T(0,0) \) at zero coordinates. Face recognition is important in vector quantization algorithms and variation methods are previously presented for extracting a feature vector from an image. PCA 3-D DCT[12] uses MPEG and JPEG for video, image compression methods which are not lossless. Due to the high compression ratio (CR) demand, the quality of the video sequence or output image is reduced with the removal of redundant data. Both methods use 3-D DCT for producing a spatial frequency spectrum in which the coefficient of DCT with low-frequency components are stored with less accuracy. MPEG video is considered as 2-D image series and all these images are compressed with 3-D DCT and also used to compress various images of the same size with 3-D DCT which differs with video cube of video elements \( MxMxM \).

2.2. Overview of the Design

DCT is used in video and image compression because various algorithms are developed to make DCT as computationally effective and implemented with 8-point DCT to reduce complexity by using rid of multiplication for approximation. Basis vector elements from 8-blocks in DCT are replaced with its sign and provided with better DCT estimation with replace of vector elements using \( 0, \frac{1}{2}, 1 \). Two transform elements of the kernel are derived as 0 and 1 and proved that they perform well especially with high, small CR ratio situations. For high DCT, the approximation is necessary due to the complexity of DCT growing is non-linear. High-Efficiency Video Coding(HEVC) method uses DCT with large blocks up to 32 to obtain high CR, but for extension method used for H264 AVC for large size transform, like 16, 32 points is not possible. Many applications of image processing like tracking required larger DCT.

DCT approximation consists the features such as it will be with low complexity, low error energy for providing better compression performance, and work for high length of DCT for supporting the latest video coding methods also other applications like Surveillance, encryption, tracking, encryption, and compression simultaneously. But, the present DCT algorithm cannot work efficiently for all the mentioned features. Some present methods perform efficiently for factors like scalable orthogonality and higher sizes. Maintaining of DCT approximation in orthogonality is based on two reasons. First, for orthogonal transform, finding its inverse, kernel matrix of inverse transform can be computed with the transpose of the kernel matrix of the forward transform. For orthogonal transform, some faster algorithm is used for both transforms.

3. Design Implementation:

3.1. 8-Point DCT Signal Flow

The Structure of the computation of DCT is given in figure 1. Approximate DCT coefficients are computed for the given input sequence. Two inputs for computation used with input adder, output permutation unit are shown in the figure 2 and the functions of the two blocks are shown. 16-Point structure can be stretched to get higher size DCT. 32-Point structure is formed with two 16-point DCTs, using input adder, output permutation units.

Figure 1. 8-point DCT Signal diagram
3.2. 16-Point DCT Signal Flow Graph (SFG)

For accessing the computation complexity of the present 16-Point DCT, the computational cost of matrices can be determined.

![Figure 2. SFG of 16-point DCT.](image)

3.3 Reconfigurable Structure of DCT Approximation for N= 8, 16

Arithmetic complexity of 16-point DCT approximation is 60, and for 32, 64-point DCT approximation is 152, and 368 additions respectively in which DCT complexity equals to additions. The presented DCT approximation with various lengths consisting of the number of arithmetic operations in addition to present approximations is shown and observed that the presented method requires lower additions without the requirement of shifting operations are not involved in combinational component, required rewiring in hardware installation. Shift-add operations increases bit width which makes the arithmetic unit to larger.

3.4 Reconfigurable Structure of DCT Approximation for N= 8, 16 and 32

DCT of various lengths like 16, 32 is needed to use in video coding applications as mentioned in HEVC is presented in the reconfiguration method. So, the DCT structure is effectively reused with various lengths instead of separate architectures with various lengths. Reconfigurable DCT architecture which can reuse for DCT computation with various lengths is presented for implementation unit that generates using 8 MUXs which selects either [a (0), a (1), .........., a(7)] or [x (0), x (1), .........., x(7)] based on 8 or 16-point DCT computation. Likely, input to second 8-point DCT unit is applied with 8 MUXs which selects [[b (0), b (1), .........., b(7)] or [x (8), x (9), .........., x(15)] based on 8 or 16-point DCT computation.

For Selecting and reordering the output permutation unit, it uses 14 MUXs based on DCT size and used as control input for MUXs for input selection and performing permutation based on the size of DCT to compute. Sel 16=1 enables 16-point DCT computation whereas sel 16=0 enables 8-point DCTs in parallel computation. Figure 3 allows computation of two 8-point DCTs or 16-point DCT in
parallel. Reconfigurable design for 8, 16, 32-point DCT computation is shown in figure 4 which performs the computation of four 8-point or two 16-point or 32-point DCTs in parallel.

![Figure 4](image)

**Figure 4. Block diagram for N= 32, 16 and 8**

Four 8-point DCT units or two 16-point input adder units or 32-point input adder unit is combined to form as architecture. Re-configurability is obtained with 3 control blocks composed using six 42:1 MUXs along with thirty 3:1 MUXs. For four 8-point, two 16-point parallel DCTs or 32-point DCT outputs are equal to {00}, {01} OR {11} respectively.

## 4. Simulation Results and Discussions

First, generated code and 32-bit DCT architecture is designed. Next, Verilog Code is written manually and Simulation is done with Functional Verification using the IUS tool (CADENCE) i.e., Native Code Compiler (NC) which is used for simulation purposes and using Register Transfer Level (RTL) Compiler (CADENCE) for Logical Synthesis purpose. Finally, Physical Design can be done using Cadence tool for Placement and Routing purpose in 180nm and 90nm technology and also implemented ECC processor based on RSD using Application-Specific Integrated Circuit (ASIC).

### 4.1 Design of 16-Point DCT

This section presents 16-point DCT design which is shown in figure 5, and also developed 32-point DCT and implemented it.

In view of this, a different integer transform is developed which applied to many block-length and proposed different 16x16 matrixes for 16-point DCT approximation and proved. Two different transforms are presented for 8-point DCT. Less complexity DCT depends on integer functions which are represented as a free form of DCT approximation and the first method is for 16, 32 lengths which depend on the extension of inter DCT. Systematic method used to develop binary DCT with sequence order.

![Figure 5](image)

**Figure 5. i) 16 point-DCT ii) Subsystem**
4.2. RTL Schematic of 16-point DCT

Below figure 6 shows the RTL schematic and simulation result of 16-point DCT. Implementation of the presented algorithm is scalable for software and hardware for various larger lengths easily by using the proposed best 8-point approximations. Proposed method achieved for fully reconfigurable, scalable, parallel structure for DCT computation. Presented algorithm of 32-point DCT is configured by using four 8-points or two 16-point DCTs for parallel computation and observed that this algorithm provides better performance than the present algorithm in terms of hardware and energy.

![RTL schematic of 16-point DCT](image1)

![Simulation of 16-point DCT](image2)

An Integrated logic analyzer uses various techniques to develop, synthesize, and validate both pipeline and non-pipelined designs. Validation can be done using 8-bit inputs and for 8-point transform used for 10, 11-bit outputs. Pipelined design is formed by adding registers at the input, output stages placing registers at every adder stage, and proved that pipelined designs provides maximum operating frequency and also proved that presented design consists 7%, 6%, 5% less area.

4.3. RTL Schematic of 32-point DCT

Figure 7 shows the RTL schematic and simulation result of 32 point DCT. This paper presents an algorithm for DCT approximation to meet all three features and also to compute DCT approximation using recursive decomposition of the DCT matrix and observed as the presented algorithm consists of low arithmetic complexity than the present DCT approximation algorithm. The presented DCT approximations formed with various lengths are orthogonal which results in as low error energy as compared with present algorithms for DCT approximation and the decomposition method permits presented transform for high-size DCTs in the generalization process.

![RTL schematic of 32-point DCT](image3)

![Simulation of 32-point DCT](image4)
From the design and results of proposed algorithm it can be summarized that the proposed and presented method is having reduced hardware complexity because of the reconfigurable architecture and the error energy rate is low and also the transform is linear throughout the length and is orthogonal. Proposed algorithm is used in the applications like digital filters, image and video compression.

**Conclusion**

Here, this paper presented a recursive algorithm for computing orthogonal approximation of DCT in which DCT length will be obtained using DCT pair. So, it can be concluded that approximated DCT have many advantages like simple structure, low complexity, regularity, and scalable compared with previous methods, our proposed method is more effective in terms of image quality compression, consumption of hardware resources, and error-energy. This paper also presented approximate DCT reconfigurable architecture which is fully scalable in which 32-point DCT was structured with two 16-point or using four 8-point DCTs which can also applicable in tracking and survey lines. Here presented the implementation of 2-bit DCT and increased for higher bit lengths i.e., 64, 128, etc. and suggested that can also implement in 90nm and 45nm technology and can further design DCT by using exploiting image indexing techniques.

**References**

[1] Patil, Ujwala, and Uma Mudengudi 2011 Image fusion using hierarchical PCA IEEE int.conf.on image Info. Processing (ICIIP) pp 1-6
[2] Sruthy, S., Latha Parameswaran, and Ajeesh P. Sasi 2013 Image Fusion Technique using DT-CWT Int.Conf. on Automation, Computing, Comm., Control and Compressed Sensing (iMac4s) pp 160-164
[3] Desale, Rajendra Pandit, and Sarita V. Verma. 2013 Study and analysis of PCA, DCT & DWT based image fusion techniques IEEE Int.Conf. on Signal Proces.Image Proc. & Pattern Recog. (ICSIPR) pp 66-69
[4] Prakash, Chandra, S. Rajkumar, and P. V. S. S. R. Mouli. 2012 Medical image fusion based on redundancy DWT and Mandani type min-sum mean-of-max techniques with quantitative analysis IEEE Proc.of Recent Advances in Comp. and Software Systems (RACSS) pp 54-59
[5] Aribi, Walid, Ali Khalfallah, Med Salim Bouhlel, and Noomene Elkadri 2012 Evaluation of image fusion techniques in nuclear medicine proc.of Sciences of Electr., Tech. of Info. and Telecom pp 875-880
[6] Ren, Haozheng, Yihua Lan, and Yong Zhang 2011 IEEE Research of multi-focus image fusion based on M-band multiwavelet transformation Proc. of fourth int.workshop of Advanced Computational Intelligence (IACI) pp 395-398
[7] Pei, Yijian, Huayu Z, Jiang Yu, and Guanghui Cai. 2010 The improved wavelet transforms based image fusion algorithm and the quality assessment IEEE Proc.of Image and Signal Proc.vol 1 pp 219-223
[8] Hai Huang. Liyi Xiao. 2014 CORDIC based fast algorithm for power-of-two point DCT and its efficient VLSI implementation. Microeletr. Journal, vol 45(11) pp 1480-1488
[9] Pahniya, P., Jain, P.2019 Realization of Second-Order Structure of Recursive Algorithm for Discrete Cosine Transform. Circuits Syst Signal Process vol.38 pp 791–804
[10] M. Masera, G. Masera and M. Martina 2020 An Area-Efficient Variable-Size Fixed-Point DCT Architecture for HEVC Encoding IEEE Trans.on Circ.and Sys.for Video Tech vol 30(1) pp 232-242
[11] Saad Al-Azawi, Omar Nibouche, Said Boussakta, Gaye Lightbody. 2019 New fast and area-efficient pipeline 3-D DCT architectures Digital Signal Proce. vol. 84 pp.15-25
[12] W. Maleika and P. Forczmański 2020 Adaptive Modeling and Compression of Bathymetric Data With Variable Density IEEE Journal of Oceanic Engg, vol 45(4) pp1353-1369