

Algorithm and Architecture for Bit- Level Implementation of the IDCT and IDST

Reeta Kumari Choudhury¹, Sudhansu Sekhar Nayak².

^{1,2}Department of Physics, Centurion University of Technology and Management Odisha, India.

Abstract: A 3-D bit-serial systolic architecture for VLSI implementation of prime – factor inverse discrete cosine transform (IDCT) and inverse discrete sine transforms (IDST) has been presented in this paper. Distributed arithmetic combined with bit – serial and bit – parallel data structure have been used to implement vector inner products concurrently. No extra hardware / time is necessary for transposition of the intermediate results. The RAM used for transposition operation is avoided. The proposed architecture consists of only memories, adders and registers. The proposed structure provides significant saving in hardware. It can easily meet the speed requirement of 14.3 MHz real-time operation.

Keywords: 3-D bit-serial systolic architecture, IDCT, IDST, bit-parallel data structure, RAC.

I. INTRODUCTION

Bit-level systolic arrays are more regular and require simpler PEs compared with word level structures. The basic PEs comprised of only gated full adders, are very small and an array of such PEs may be integrated on a small chip. The computation time of the bit-level PEs is very small compared with the word-level PEs, so that high throughput computation can be achieved by bit-level systolic arrays [1]. Therefore, bit-level systolic arrays have been used for efficient computation of the DST. Chakrabarti and Ja Ja [2] have suggested a bit-level architecture for prime-factor DHT which is computed via four temporary output. Besides, they have shown that prime-factor DCT can also be computed from the DHT structure by simple modification. The prime-factor approach of [2] involves significant hardware and time in the input and output interfaces for index mapping and storing of data matrices. Apart from that, structures for implementation of prime-factor algorithms require additional time, as well as, hardware for transposition of intermediate result for implementing in a planar architecture. Sun *et al.* [3] have presented a regular and efficient IC realization for DCT by concurrent architecture using distributed arithmetic and memory oriented structure. The ROM size of this architecture increases rapidly with the order of the DCT so that it may be useful for implementing the DCT of lower order only. It would not be suitable for implementing the DCT of higher order such as $N = 1024$ which is often required for image processing applications. Mc. Govern *et al.* [4] have suggested a bit-serial architecture for implementation of (8×8) point 2-D DCT, which is not modular and requires complicated interconnections. In Section II, we have presented a 3-D bit-serial systolic architecture for VLSI implementation of prime-factor IDCT and IDST. We have used distributed arithmetic combined with bit-parallel data structure to implement vector inner products concurrently. It is interesting to note that no extra hardware/time is necessary for transposition of the intermediate results. Thus the RAM used for transposition operation is avoided. Hardware and throughput considerations have been discussed in section III. Conclusion is given in Section IV.

II. BIT-LEVEL IMPLEMENTATION OF PRIME-FACTOR IDCT AND IDST

A. Algorithm for bit-level implementation of prime factor IDCT and IDST

The DCT of a sequence $\{x(n), n = 0, 1, 2, \dots, N - 1\}$ is defined as

$$X(k) = \frac{2}{N} \in(k) \sum_{n=0}^{N-1} x(n) \cos \left[\frac{\pi(2n+1)k}{2N} \right] \quad (1)$$

and the inverse discrete transform (IDCT) is given by

$$x(n) = \frac{2}{N} \sum_{k=0}^{N-1} \in(k) X(k) \cos \left[\frac{\pi(2n+1)k}{2N} \right] \quad (2)$$

For $k = 0, 1, 2, \dots, N-1$

$$\text{Where } \in(k) = \begin{cases} (2)^{-\frac{1}{2}} & \text{for } k = 0 \\ 1 & \text{for } 1 \leq k \leq N - 1 \end{cases}$$

Since $\epsilon(k)$ effects only the amplitude of $X(0)$ component, we shall take $\epsilon(k)$ as unity .

When transform length $N = N_1 \times N_2$, N_1 and N_2 being relatively prime, the input index k in equation (2) may be mapped into (k_1, k_2) as

$$k = (N_2 k_1 + N_1 k_2) \bmod N \tag{3}$$

Equation (2) may be expressed as

$$x(n) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} y(k_1, k_2) \cos \left[\frac{\pi(2n+1)k_1}{2N_1} \right] \cos \left[\frac{\pi(2n+1)k_2}{2N_2} \right] \tag{4}$$

$$\begin{aligned} \text{Where, } y(k_1, k_2) &= X(k_1, k_2) - X(N_1 - k_1, N_2 - k_2) && \text{for } N_2 k_1 + N_1 k_2 < N \text{ and} \\ &= X(N_1 - k_1, k_2) + X(k_1, N_2 - k_2) && \text{for } N_2 k_1 + N_1 k_2 > N \end{aligned} \tag{5}$$

The output index n in equation (4) may be mapped into (n_1, n_2) as

$$n_i = \begin{cases} \bar{n}_i \text{ if } \bar{n}_i < N_i & \text{for } i = 1 \text{ and } 2 \\ 2N_i - 1 - \bar{n}_i & \text{otherwise where } \bar{n}_i = n \bmod 2N_i \end{cases} \tag{6}$$

Using equations (5) and (6), equation (4) can be written as

$$x(n_1, n_2) = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} y(k_1, k_2) \cos \left[\frac{\pi(2n_1+1)k_1}{2N_1} \right] \cos \left[\frac{\pi(2n_2+1)k_2}{2N_2} \right] \tag{7}$$

For DCT equation (7) may be expressed as

$$x(n_1, n_2) = \sum_{k_2=0}^{N_2-1} W(n_1, k_2) \cos \left[\frac{\pi(2n_2+1)k_2}{2N_2} \right] \tag{8}$$

Where

$$W(n_1, k_2) = \sum_{k_1=0}^{N_1-1} y(k_1, k_2) \cos \left[\frac{\pi(2n_1+1)k_1}{2N_1} \right] \tag{9}$$

For DST equation (8) may be expressed as

$$x(n_1, n_2) = \sum_{k_2=1}^{N_2} W(n_1, k_2) \sin \left[\frac{\pi(2n_2-1)k_2}{2N_2} \right] \tag{10}$$

Where

$$W(n_1, k_2) = \sum_{k_1=0}^{N_1} y(k_1, k_2) \sin \left[\frac{\pi(2n_1-1)k_1}{2N_1} \right] \tag{11}$$

Equations (8), (9), (10) and (11) imply that $[x(n_1, n_2)]$ is computed in two distinct stages. In the first stage, an intermediate result $[W(n_1, k_2)]$ is computed from each column of the matrix $[y(k_1, k_2)]$ of size $N_1 \times N_2$. In the second stage, $[x(n_1, n_2)]$ is computed from each row of the intermediate result $[W(n_1, k_2)]$.

Equation (9) and (11) may be expressed in vector form as

$$W(k, l) = C_k Y_l \tag{12}$$

$$W(k, l) = S_k Y_l \tag{13}$$

Where Y_l is the l th column vector of $[y(k_1, k_2)]$ and C_k / S_k is the k th row vector of the kernel matrix. Let the elements of the input data matrix $[y(k_1, k_2)]$ be represented by 2's complement code as

$$y_{ml} = -y_{ml}^{n-1} 2^{n-1} + \sum_{j=0}^{n-2} y_{ml}^j 2^j \tag{14}$$

Where y_{ml}^j is the j th bit of y_{ml} which has a value of either 0 or 1. n is the number of bits y_{ml} carries. y_{ml}^{n-1} is the sign bit. Substituting equation (14) in equation (12), we have

$$W(k, l) = - \sum_{m=0}^{N-1} C_{km} y_{ml}^{n-1} 2^{n-1} + \sum_{j=0}^{N-2} \sum_{m=0}^{N-1} C_{km} y_{ml}^j 2^j \tag{15}$$

$$W(k, l) = -F_{kl}(C_k, y_l^{n-1}) 2^{n-1} + \sum_{j=0}^{n-2} F_{kl}(C_k, y_l^j) 2^j \tag{16}$$

Where

$$F_{kl}(C_k, y_l^j) = \sum_{m=0}^{N-1} C_{km} y_{ml}^j \text{ for } j = 0, 1, \dots, n-1 \tag{17}$$

Similarly for DST

$$F_{kl}(S_k, y_l^j) = \sum_{m=1}^{N-1} S_{km} y_{ml}^j \text{ for } j = 0, 1, \dots, n-1 \tag{18}$$

F_{kl} are functions of C_{km} and S_{km} and bit pattern of y_{ml} . Since C_{km} and S_{km} are fixed numbers, F_{kl} can be generated and stored in memory for all possible bit patterns of y_{ml} . So $W(k, l)$ can be calculated concurrently for $l=0, 1, \dots, N-1$ by shifts and adds of the values of F_{kl} stored in memory.

B. 3-D architecture for bit-serial implementation of IDCT and IDST

The proposed multilayer 3-D architecture for implementation of IDCT and IDST of size $N = N_1 \times N_2$ is shown in Fig.1. N_2 number of N_1 -point IDCT and IDST computation given by equations (9) and (11) are implemented in the first block comprising of N_1 number of parallel horizontal layers. Each layer of this stage consists of N_1 number of identical processors shown in fig.2. N_1 number of N_2 -point IDCT and IDST computation given by equations (8) and (10) are implemented in the second block comprising of N_1 number of parallel vertical layers. Each layer of this stage consists of N_2 number of identical processors. Each processor of the second block is identical to that of the first block except that it does not consist of Q shift registers. The transposition of the intermediate matrix is avoided by orthogonal alignment of the layers of the first block with respect to the layers of the second block, shown in Fig.1. The elements of the input sequence $y_{0n_2}, y_{1n_2}, \dots, y_{N-1n_2}$ are shifted sequentially in time into shift registers of the first block at input data rate of I/T . After N_1/T interval, the contents of Q shift registers are bit-parallelly loaded in to R shift registers concurrently in time. The data in R shift registers are taken bit serially shifted out concurrently in time, with the least significant bit first. The bits shifted out of the R shift registers are used to address the ROMs which contain the values of F_{kl} (equations 17 and 18). Detail structure of RAC (ROM and accumulator) is shown in Fig.3. The contents of ROM is read and added to the content of shift register 'SR' after 1-bit right shift. The same operations are repeated for $(n+1)$ times except for the last time (sign bit), when a subtraction instead of addition is performed. The outputs of the first block are bit-parallel loaded into R registers of the corresponding processors of the second block concurrently in time and then shifted bit-serially. For implementing in the planar board technology, the proposed structure may be stretched to two dimensions instead of three dimension.

III. HARDWARE AND THROUGHPUT CONSIDERATIONS

The 3-D bit-level systolic architecture proposed by us for VLSI implementation of prime-factor IDCT and IDST consists of only memories, adders and registers. The use of multipliers is eliminated. The transposition of the intermediate matrix is avoided by orthogonal alignment of the layers of the first block with respect to the layers of the second block. The RAM required for the transposition of the intermediate result is avoided. The size of the ROM required for the proposed structure is equal to $(2^{N_1} + 2^{N_2})$. The proposed structure provides significant saving in hardware over the structure of Sun *et al.*[3], which requires a ROM of size $N 2^N$ for computing the N -point DCT/DST, without any decrease in the throughput rate. The proposed multilayer 3-D architecture for implementation of IDCT and IDST of size $N_1 \times N_2$ requires $2N$ number of processors. The proposed architecture can easily meet the speed requirement of 14.3MHz real-time operation with CMOS technology.

IV. CONCLUSION

In this paper we have developed efficient algorithm and architecture for bit-level VLSI implementation of IDCT and IDST. We have proposed a 3-D bit serial systolic architecture for VLSI implementation of prime-factor IDCT and IDST using distributed arithmetic combined with bit-serial and bit-parallel data structure to implement vector inner product concurrently. It has a highly regular structure. It consists of only memories, adders and registers. The use of multipliers is eliminated. The RAM required for the

transposition of the intermediate result is avoided. The size of ROM required for the proposed structure is equal to $N(2^{N_1} + 2^{N_2})$. The proposed structures provides significant saving in hardware over the structure of Sun *et al.*[3], which requires a ROM of size $N2^N$ for computing the N -point DCT/DST, without any decrease in the throughput rate.

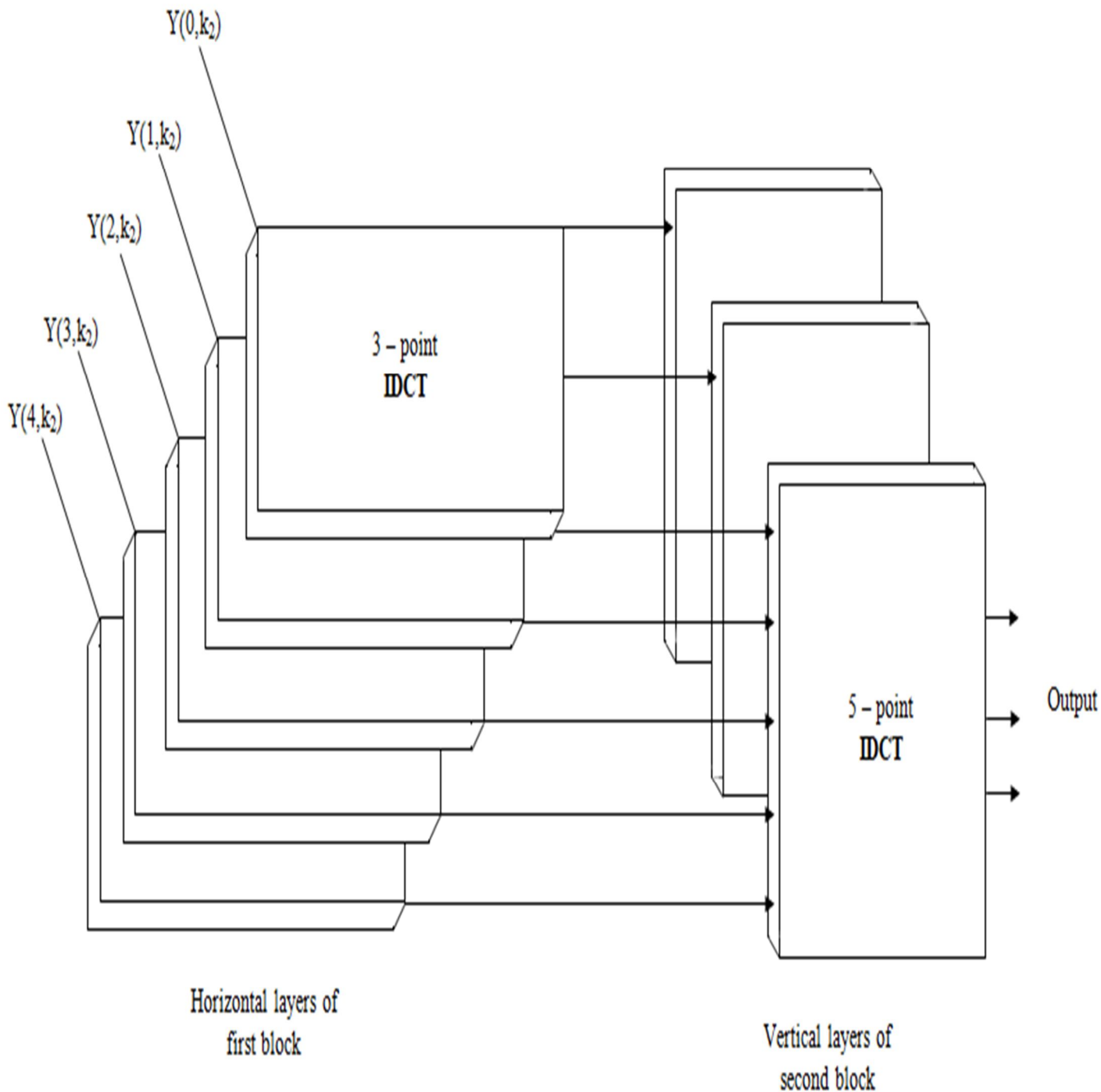


Fig.1: 3-D multilayer systolic architecture for computation for 3x5 IDCT / IDST

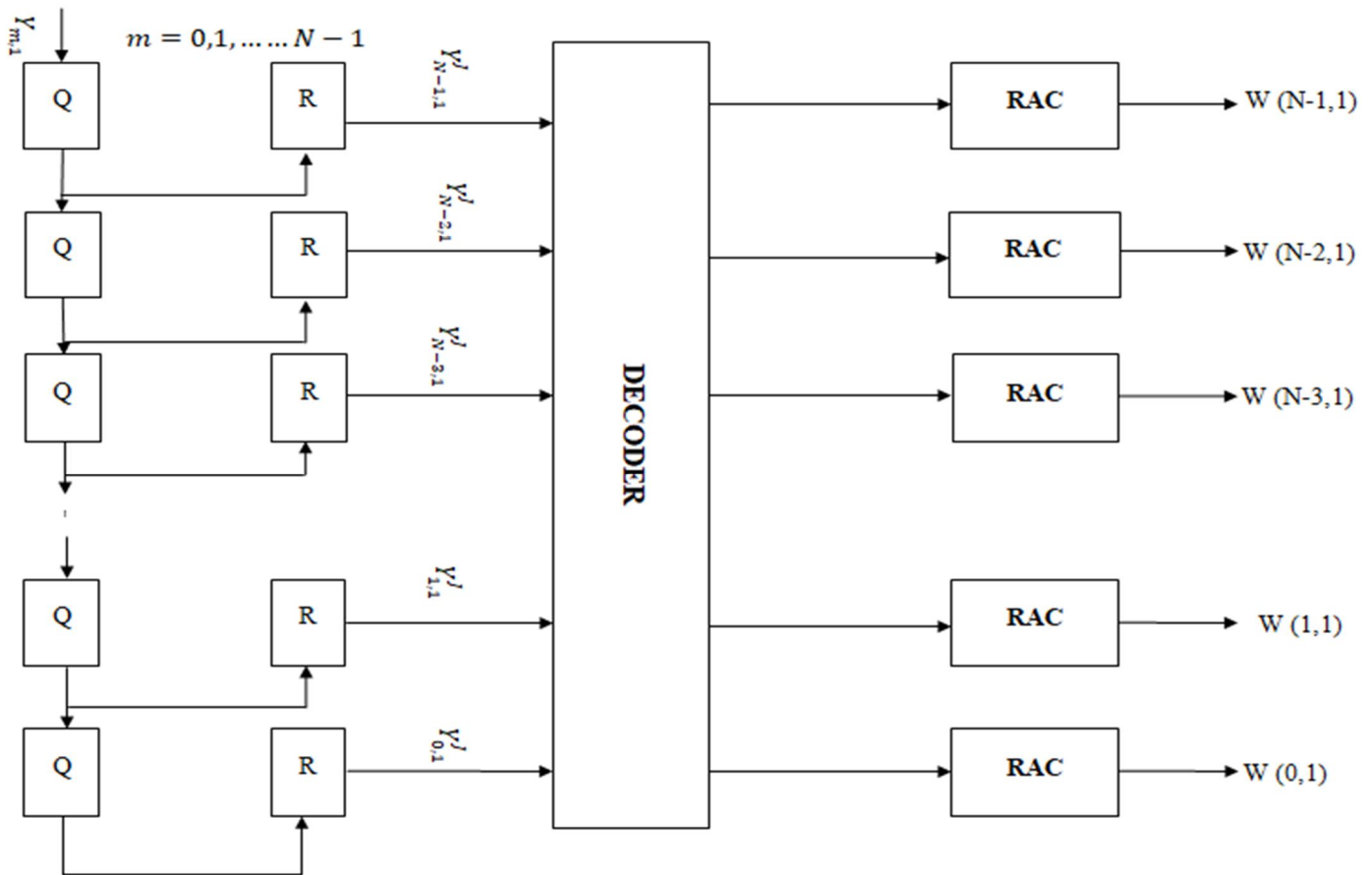


Fig. 2: Processors used in layers.

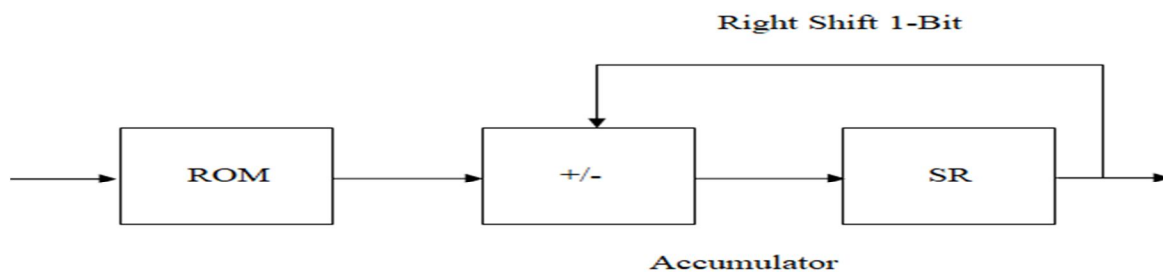


Fig. 3: Detailed structure of RAC.

REFERENCES

- [1] MC CANNY, J.V., MC WHIRTER, J.G., and KUNG, S.Y., "The use of data dependence graphs in the design of bit-level systolic arrays", IEEE Trans on Acoustics, Speech, & Signal Processing, vol. ASSP-18, No.5, pp. 787-793, May 1990.
- [2] CHAKRABARTI, C. and JA'JA', J., "Systolic architectures for the computation of the discrete Hartley and the discrete cosine transforms based on prime-factor decomposition", IEEE Trans, on computers, vol.39, No. 11, pp. 1359-1368, Nov.1990.
- [3] SUN, M.T., WU, L., and LIOU, M.L., "A concurrent architecture for VLSI implementation of discrete cosine transform", IEEE Trans. On Circuits & Systems, vol. CAS-34, No. 8, pp. 992-994, Aug. 1987.
- [4] MC GOVERN, F.A., WOODS, R.F., and YAN, M., "Novel VLSI implementation of (8x8) point 2-D DCT", Electron Lett., vol.30, No. 8, pp. 624-626, Apr. 1994.
- [5] NAYAK, S.S. and MEHER, P.K., "High throughput VLSI implementation of discrete orthogonal transforms using bit-level vector-matrix multiplier", IEEE Trans. on Circuits & Systems-II : Analog and Digital Signal Processing, vol.46, No. 5, pp. 655-658, May 1999.
- [6] REJU, V.G., SOO, N.K. and SOON, I.Y., "Convolution using discrete sine and cosine transforms", IEEE, Signal Processing Letters, vol.14, No.7, pp.445-448, July 2007.