Pipeline Architecture of 2d Dct for High Efficiency Video Coding

Dhanya R  
Department of Electronics and Communication Engineering  
Musaliar College of Engineering and Technology  
Pathanamthitta, Kerala, India

Nishi G Nampoothiri  
Asso. Prof  
Department of Electronics and Communication Engineering  
Musaliar College of Engineering and Technology  
Pathanamthitta, Kerala, India

Abstract— In this paper, a novel computation and energy reduction technique for High Efficiency Video Coding (HEVC) Discrete Cosine Transform (DCT) for all Transform Unit (TU) sizes is proposed. The existing system reduces the computational complexity of HEVC DCT significantly at the expense of slight decrease in PSNR and slight increase in bit rate by only calculating several pre-determined low frequency coefficients of TUs and assuming that the remaining coefficients are zero. It reduced the execution time of HEVC HM software encoder up to 12.74%, and it reduced the execution time of DCT operations in HEVC HM software encoder up to 37.27%. Currently different types of transform techniques are used by different video codes to achieve data compression during video frame transmission. Among them, discrete cosine transform (DCT) is supported by most of modern video standards. The integer DCT is an approximation of DCT. It can be implemented exclusively with integer arithmetic. Integer DCT proves to be highly advantageous in cost and speed for hardware implementation. Implementation of an efficient discrete cosine transform with reduced complexity and number of multiplications. Pipelining technique is introduced to reduce the processing time. The full pipeline variable block size transform engine with the efficient hardware utilization is proposed to handle the DCT/IDCT. 2D-DCT is computed by combining two 1D-DCT that connected by a transpose. So in proposed system use pipeline architecture to reduce the computational complexity of HEVC than existing system. In this paper, low energy HEVC 2D DCT hardware for all TU sizes is also designed and implemented using Verilog HDL. The proposed hardware, in the worst case, can process 53 Ultra HD (7680x4320) video frames per second. The proposed technique reduced the energy consumption of this hardware up to 18.9%. Therefore, it can be used in portable consumer electronics products that require a real-time HEVC encoder.

Index Terms — HEVC, Discrete Cosine Transform, Hardware Implementation, FPGA, Energy Reduction.

I. INTRODUCTION

A new international video compression standard called High Efficiency Video Coding (HEVC) is recently developed [1]-[6]. It has 50% better video compression efficiency than H.264 standard. It uses Discrete Cosine Transform (DCT) / Inverse Discrete Cosine Transform (IDCT) same as H.264 standard. However, H.264 standard uses only 4x4 and 8x8 Transform Unit (TU) sizes for DCT/IDCT. HEVC standard uses 4x4, 8x8, 16x16, and 32x32 TU sizes for DCT/IDCT. Larger TU sizes achieve better energy compaction. However, they increase the computational complexity exponentially. In addition, HEVC standard uses Discrete Sine Transform (DST) / Inverse Discrete Sine Transform (IDST) for 4x4 intra prediction in certain cases. Transform operations (DCT/IDCT and DST/IDST) are heavily used in an HEVC encoder [7]. DCT and DST have high computational complexity. DCT and DST operations account for 11% of the computational complexity of an HEVC video encoder. They account for 25% of the computational complexity of an all intra HEVC video encoder.

In this paper, a low energy HEVC 2D DCT hardware for all TU sizes is also designed and implemented using Verilog HDL. The proposed hardware calculates 4, 8, 16 and 32 DCT coefficients per clock cycle for 4x4, 8x8, 16x16 and 32x32 TU sizes, respectively. It, in the worst case, can process 48 Quad Full HD (3840x2160) video frames per second. In this paper, another low energy HEVC 2D DCT hardware for all TU sizes with higher hardware utilization is also designed and implemented using Verilog HDL.

Clock gating is used to reduce the energy consumptions of both hardware. Hcub Multiplierless Constant Multiplication (MCM) algorithm [9] is used to reduce number and size of the adders in both hardware. Hcub MCM algorithm reduced the energy consumption of the lower utilization (LU) hardware and the higher utilization (HU) hardware up to 5.9% and 13.1%, respectively. Finally, the proposed technique is used to reduce the energy consumptions of both hardware. It further reduced the energy consumption of the LU hardware and the HU hardware up to 17.9% and 18.9%, respectively. Ease of Use

II. PROPOSED COMPUTATION AND ENERGY REDUCTION TECHNIQUE

After forward transform and quantization, most of the forward transformed and quantized high frequency coefficients in a TU become zero. In addition, if the values of non-zero forward transformed and quantized low frequency coefficients in a TU are small, they have small impact on the inverse quantized and inverse transformed TU. Therefore, the proposed technique only calculates several pre-determined low frequency coefficients of TUs, and it assumes that the remaining coefficients are zero.
In this paper, a novel computation and energy reduction technique for High Efficiency Video Coding (HEVC) Discrete Cosine Transform (DCT) for all Transform Unit (TU) sizes is proposed. The existing system reduces the computational complexity of HEVC DCT significantly at the expense of slight decrease in PSNR and slight increase in bit rate by only calculating several pre-determined low frequency coefficients of TUs and assuming that the remaining coefficients are zero.

**EXISTING SYSTEM**

Currently different types of transform techniques are used by different video codes to achieve data compression during video frame transmission. Among them, discrete cosine transform (DCT) is supported by most of modern video standards. The integer DCT is an approximation of DCT. It can be implemented exclusively with integer arithmetic. Integer DCT proves to be highly advantageous in cost and speed for hardware implementation. Implementation of an efficient discrete cosine transform with reduced complexity and number of multiplications. In this paper, low energy HEVC 2D DCT hardware for all TU sizes is also designed and implemented using Verilog HDL. The proposed hardware, in the worst case, can process 53 Ultra HD (7680x4320) video frames per second. The proposed technique reduced the energy consumption of this hardware up to 18.9%. Therefore, it can be used in portable consumer electronics products that require a real-time HEVC encoder.

**Proposed System**

It reduced the execution time of HEVC HM software encoder up to 12.74%, and it reduced the execution time of DCT operations in HEVC HM software encoder up to 37.27%. Currently different types of transform techniques are used by different video codes to achieve data compression during video frame transmission. Among them, discrete cosine transform (DCT) is supported by most of modern video standards. The integer DCT is an approximation of DCT. It can be implemented exclusively with integer arithmetic. Integer DCT proves to be highly advantageous in cost and speed for hardware implementation. Implementation of an efficient discrete cosine transform with reduced complexity and number of multiplications. Pipelining technique is

<table>
<thead>
<tr>
<th>TU size</th>
<th>Org.</th>
<th>C.Set #1</th>
<th>Red. (%)</th>
<th>C.Set #2</th>
<th>Red. (%)</th>
<th>Coefficient Set #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>4x4</td>
<td>Add</td>
<td>224</td>
<td>84</td>
<td>62.5</td>
<td>147</td>
<td>34.4</td>
</tr>
<tr>
<td></td>
<td>shift</td>
<td>224</td>
<td>84</td>
<td>62.5</td>
<td>147</td>
<td>--</td>
</tr>
<tr>
<td>8x8</td>
<td>Add</td>
<td>2560</td>
<td>960</td>
<td>62.5</td>
<td>600</td>
<td>74.2</td>
</tr>
<tr>
<td></td>
<td>shift</td>
<td>2304</td>
<td>864</td>
<td>62.5</td>
<td>594</td>
<td>74.2</td>
</tr>
<tr>
<td>16x</td>
<td>Add</td>
<td>20992</td>
<td>7872</td>
<td>62.5</td>
<td>13776</td>
<td>34.4</td>
</tr>
<tr>
<td></td>
<td>shift</td>
<td>16896</td>
<td>6336</td>
<td>62.5</td>
<td>11088</td>
<td>34.4</td>
</tr>
<tr>
<td>32x</td>
<td>Add</td>
<td>18227</td>
<td>57600</td>
<td>62.5</td>
<td>46992</td>
<td>74.2</td>
</tr>
<tr>
<td></td>
<td>shift</td>
<td>15360</td>
<td>0</td>
<td>62.5</td>
<td>39600</td>
<td>74.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>3960</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>74.2</td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td>62.5</td>
<td>54.3</td>
<td></td>
<td></td>
<td>74.2</td>
</tr>
</tbody>
</table>
introduced to reduce the processing time. The full pipeline variable block size transform engine with the efficient hardware utilization is proposed to handle the DCT/IDCT. 2D-DCT is computed by combining two 1D-DCT that connected by a transpose. So in proposed system use pipeline architecture to reduce the computational complexity of HEVC than existing system.

A PROPOSED HEVC 2D DCT HARDWARE

A. Proposed HEVC 2D DCT Lower Utilization Hardware

The proposed HEVC 2D DCT lower utilization (LU) hardware for all TU sizes including clock gating, Hcub MCM algorithm, and the proposed technique with coefficient set 3 is shown in Fig. 3. Input splitter is used to select the proper DCT inputs for each TU size. Output multiplexers are used to select the proper DCT outputs for each TU size. Column and row clip modules are used to scale the outputs of 1D column DCT and 1D row DCT to 16 bits, respectively. Column clip shifts 1D column DCT outputs right by 1, 2, 3 and 4 for 4x4, 8x8, 16x16 and 32x32 TU sizes, respectively. Row clip shifts 1D row DCT outputs right by 8, 9, 10 and 11 for 4x4, 8x8, 16x16 and 32x32 TU sizes, respectively.

Since HEVC DCT algorithm allows performing an N-point 1D DCT by performing two N/2-point 1D DCTs with some preprocessing, the proposed hardware performs N-point 1D DCT transforms by performing two N/2-point 1D DCT transforms with an efficient butterfly structure. It performs 2D DCT by first performing 1D DCT transform on the columns of a TU, and then performing 1D DCT transform on the rows of the TU. After 1D column DCT, the resulting coefficients are stored in a transpose memory, and they are used as input for 1D row DCT.

The butterfly structure used for column transforms is shown in Fig. 4. For 4x4 TUs, only 4x4 butterfly operation is used. For 8x8 TUs, 8x8 and 4x4 butterfly operations are used. For 16x16 TUs, 16x16, 8x8 and 4x4 butterfly operations are used. For 32x32 TUs, all butterfly operations (32x32, 16x16, 8x8, 4x4) are used.

One 4x4 datapath is used for 4x4 TU size. Two 4x4 datapaths are used for 8x8 TU size. Two 4x4 datapaths and one 8x8 datapath are used for 16x16 TU size. All datapaths (two 4x4, one 8x8 and one 16x16) are used for 32x32 TU size. In order to reduce the power consumption of proposed hardware, data gating is used for the inputs of 4x4, 8x8 and 16x16 column and row datapaths. The inputs of these datapaths are stored into registers. If a datapath is not used for a TU, its input registers are not updated. This prevents unnecessary switching activities in this datapath.

DCT multiplications are performed in the datapaths using only adders and shifters. In order to reduce number and size of the adders in the proposed hardware, Hcub MCM algorithm [9] is used for implementing multiplications with constants. Hcub algorithm tries to minimize number and size of the adders in a multiplier block which multiplies a single input with multiple constants using shift and addition operations. Hcub algorithm determines necessary shift and addition operations in a multiplier block.

The transpose memory is implemented using 32 Block RAMs (BRAM). 4, 8, 16 and 32 BRAMs are used for 4x4, 8x8, 16x16 and 32x32 TU sizes, respectively. In the figure, the numbers in each box show the BRAM that coefficient is stored. The results of 1D column DCT are generated column by column. For 32x32 TU size, first, the coefficients in column 0 (C0) are generated in a clock cycle and stored in 32 different BRAMs. Then, the coefficients in column 1 (C1) are generated in the next clock cycle and stored in 32 different BRAMs using a rotating addressing scheme. This continuous until the coefficients in column 31 (C31) are generated and stored in 32 different BRAMs using the rotating addressing scheme.

B. Proposed HEVC 2D DCT Higher Utilization Hardware

The proposed HEVC 2D DCT higher utilization (HU) hardware processes four 4x4 TUs or two 8x8 TUs in parallel. Same as the LU hardware, it uses two 4x4 datapaths and one 8x8 datapath for 16x16 TU size, and it uses all datapaths (two 4x4, one 8x8 and one 16x16) for 32x32 TU size. However, the HU hardware uses two 4x4 datapaths and one 8x8 datapath for 4x4 and 8x8 TU sizes. Since 4x4 and 8x8 column and row datapaths are used for all TU sizes, data gating is used only for the inputs of 16x16 column and row datapaths.

Same as the LU hardware, multiplier blocks in the first 4x4 datapath and 16x16 datapath multiply a single input with 3 and 16 different constants, respectively. However, in the HU hardware, multiplier blocks in the second 4x4 datapath and 8x8 datapath multiply a single input with 7 and 15 different constants, respectively. Because, in the HU hardware, the second 4x4 datapath and 8x8 datapath are used for all TU sizes.

In order to calculate each output of 1D DCT for 4x4, 8x8 and 16x16 TU sizes, an output from each multiplier block in both 4x4 datapaths and 8x8 datapath is selected, and these outputs are added or subtracted. Similarly, in order to calculate each output of 1D DCT for 32x32 TU size, 32 outputs from 32 multiplier blocks in all datapaths (two 4x4, one 8x8 and one 16x16) are added or subtracted.

Same as the LU hardware, transpose memory is implemented using 32 BRAMs. However, in the HU hardware, 8, 8, 16 and 32 BRAMs are used for 4x4, 8x8, 16x16 and 32x32 TU sizes, respectively.

IV. IMPLEMENTATION RESULTS

The proposed low energy HEVC 2D DCT LU and HU hardware for all TU sizes including clock gating (original hardware), including clock gating and Hcub MCM algorithm (MCM hardware), and including clock gating, Hcub MCM algorithm and the proposed technique with coefficient set 3 (proposed hardware) are implemented in Verilog HDL. The Verilog RTL implementations are verified with RTL simulations. RTL simulation results matched the results of 2D DCT implementation in HEVC HM software encoder [8]. The Verilog RTL codes are synthesized and mapped to an FPGA implemented in 40nm CMOS technology. The FPGA implementations are verified with post place & route simulations. Post place & route simulation results matched the results of 2D DCT implementation in HEVC HM software encoder [8]. The FPGA implementation results
given in Table VI show that Hcub MCM algorithm considerably decreased area, and the proposed technique slightly increased area. Power consumptions of the FPGA implementations are estimated using a gate level power estimation tool. Post place & route timing simulations are performed for Tennis, Kimono and ParkScene (1920x1080) videos at 100 MHz [19] and signal activities are stored in VCD files. These VCD files are used for estimating power consumptions of the FPGA implementations. The energy consumption results for the LU hardware and the HU hardware for one frame of each video are shown in Fig. 8 and Fig. 9, respectively. Hcub MCM algorithm reduced the energy consumption of the LU hardware and the HU hardware up to 5.9% and 13.1%, respectively. The proposed energy reduction technique further reduced the energy consumption of the LU hardware and the HU hardware up to 17.9% and 18.9%, respectively.

In order to compare the LU hardware and the HU hardware with the HEVC DCT hardware in the literature, their Verilog RTL codes are also synthesized to a 90nm standard cell library and the resulting netlists are placed and routed. The resulting ASIC implementations of the LU hardware and the HU hardware work at 140 MHz and 130 MHz, respectively. Gate counts of the LU hardware and the HU hardware are calculated as 175K and 197K, respectively, according to NAND (3x1) gate area excluding on-chip memory. The comparison of the LU hardware and the HU hardware with the HEVC DCT hardware in the literature is shown in Table VII.

The proposed 2D DCT hardware has smaller area and power consumption than the 2D DCT hardware proposed in [14]-[17]. The DCT hardware proposed in [18] only performs 1D DCT, and its performance is not given. Since the 2D DCT hardware proposed in [14] and [17] use multipliers, they have larger area than the proposed 2D DCT hardware. Since the 2D DCT hardware proposed in [16] performs DCT operations for several TUs in parallel for smaller TU sizes, it achieves higher performance than the proposed 2D DCT hardware at the expense of much larger area and power consumption. It has same performance as the proposed 2D DCT hardware with larger area.

V. CONCLUSIONS

In this paper, a novel computation and energy reduction technique for HEVC DCT for all TU sizes is proposed. The proposed technique reduced the computational complexity of HEVC DCT significantly at the expense of slight decrease in PSNR and slight increase in bit rate. In this paper, a low energy HEVC 2D DCT hardware for all TU sizes is also designed and implemented using Verilog HDL. The proposed hardware, in the worst case, can process 53 Ultra HD (7680x4320) video frames per second. The proposed technique reduced the energy consumption of this hardware up to 18.9%. Therefore, it can be used in portable consumer electronics products that require a real-time HEVC encoder.

Currently different types of transform techniques are used by different video codes to achieve data compression during video frame transmission. Among them, discrete cosine transform (DCT) is supported by most of modern video standards. The integer DCT is an approximation of DCT. It can be implemented exclusively with integer arithmetic. Integer DCT proves to be highly advantageous in cost and speed for hardware implementation. Implementation of an efficient discrete cosine transform with reduced complexity and number of multiplications. Pipelining technique is introduced to reduce the processing time. The full pipeline variable block size transform engine with the efficient hardware utilization is proposed to handle the DCT/IDCT. 2D-DCT is computed by combing two 1D-DCT that connected by a transpose. So in proposed system use pipeline architecture to reduce the computational complexity of HEVC than existing system.

### Table III: Timing summary of proposed and existing system

<table>
<thead>
<tr>
<th>System</th>
<th>Minimum period</th>
<th>Maximum frequency</th>
<th>Minimum input arrival time before clock</th>
<th>Maximum output required time after clock</th>
<th>Total memory usage is</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>2.333ns</td>
<td>428.604MHz</td>
<td>1.536ns</td>
<td>16.536ns</td>
<td>427172 kilobytes</td>
</tr>
<tr>
<td>Existing</td>
<td>5.044ns</td>
<td>198.238MHz</td>
<td>5.254ns</td>
<td>18.769ns</td>
<td>507620 kilobytes</td>
</tr>
</tbody>
</table>

**ADVANTAGES**

- Reused for any of the prescribed lengths.
- The proposed structure could be reusable for DCT of lengths 4, 8, 16, and 32 with a throughput of 32 DCT coefficients per cycle irrespective of the transform size.
- Less-area delay due to Parallel implementation.
- The proposed architecture could be pruned to reduce the complexity of implementation substantially with only a marginal affect on the coding performance.

**APPLICATIONS**

- It is used in Mobile Multimedia Devices.
- The proposed architecture is found to support ultrahigh definition 7680 × 4320 at 60 frames/s video, which is one of the applications of HEVC.
- Signal & Image Processing. Digital Cameras, HDTV

**FUTURE SCOPE**

- The proposed system can modified by reducing the Area and delay of the design in future.
- The fast algorithm for the 8-point DCT 2D architecture designed by applying the 1D DCT structure in the folded and full parallel 2D DCT architecture.

**REFERENCES**


