An Efficient FPGA Implementation of HEVC Intra Prediction

Hasan Azgin, Ahmet Can Mert, Ercan Kalali, Ilker Hamzaoglu

Faculty of Engineering and Natural Sciences, Sabanci University

34956 Tuzla, Istanbul, Turkey

{hasanazgin, ahmetcanmert, ercankalali, hamzaoglu}@sabanciuniv.edu

Abstract— **Intra prediction algorithm used in High Efficiency Video Coding (HEVC) standard has very high computational complexity. In this paper, an efficient FPGA implementation of HEVC intra prediction is proposed for 4x4, 8x8, 16x16 and 32x32 angular prediction modes. In the proposed FPGA implementation, one intra angular prediction equation is implemented using one DSP block in FPGA. The proposed FPGA implementation, in the worst case, can process 55 Full HD (1920x1080) video frames per second. It has up to 34.66% less energy consumption than the original FPGA implementation of HEVC intra prediction. Therefore, it can be used in portable consumer electronics products that require a real-time HEVC encoder.**

Keywords—HEVC, Intra Prediction, Hardware Implementation, FPGA.

I. INTRODUCTION

An international video compression standard called High Efficiency Video Coding (HEVC) is recently developed [1]-[7]. HEVC standard provides 50% better video coding efficiency than H.264 standard. Intra prediction algorithm used in HEVC has higher computational complexity than the intra prediction algorithm used in H.264.

Intra prediction algorithm predicts the pixels of a block from the pixels of its already coded and reconstructed neighboring blocks. In H.264, there are 9 intra prediction modes for 4x4 luminance blocks, and 4 intra prediction modes for 16x16 luminance blocks. In HEVC, for the luminance component of a frame, intra prediction unit (PU) size can be from 4x4 up to 32x32 and number of intra prediction modes for a PU is 35 [1].

A computation and energy reduction technique for HEVC intra prediction is proposed in [8]. This technique reorganizes the HEVC intra prediction equations by utilizing the fact that the sum of the coefficients used in each HEVC angular intra prediction equation is 32. This reduces the amount of computations performed by 4x4, 8x8, 16x16 and 32x32 luminance angular prediction modes. It does not affect the PSNR and bit rate.

Xilinx FPGAs have built-in full-custom DSP blocks which can perform constant multiplications faster and with less energy than adders and shifters. A DSP block can be used to perform different constant multiplications by providing proper constant value to its input. Therefore, it is more efficient to implement constant multiplications using DSP blocks instead of using adders and shifters in an FPGA implementation.

In this paper, an efficient FPGA implementation of HEVC intra prediction for angular prediction modes of all PU sizes (4x4, 8x8, 16x16 and 32x32) is proposed. The proposed FPGA implementation uses the computation and energy reduction technique for HEVC intra prediction proposed in [8]. However, it implements intra angular prediction equations using DSP blocks in FPGA instead of using adders and shifters. In this way, one HEVC intra angular prediction equation is implemented using only one DSP block instead of using two DSP blocks and two adders.

The proposed FPGA implementation can work at 227 MHz in a Xilinx Virtex 6 FPGA. It, in the worst case, can process 55 Full HD (1920x1080) video frames per second. The proposed FPGA implementation has up to 15.97% less energy consumption than the FPGA implementation of HEVC intra prediction using the computation and energy reduction technique proposed in [8] and adders and shifters. The proposed FPGA implementation has up to 34.66% less energy consumption than the FPGA implementation of HEVC intra prediction using original prediction equations and DSP blocks.

Several HEVC intra prediction hardware are proposed in the literature [8]-[13]. In section III, they are compared with the HEVC intra prediction hardware proposed in this paper.

The rest of the paper is organized as follows. In Section II, HEVC intra prediction algorithm and the computation and energy reduction technique proposed in [8] are explained. The proposed HEVC intra prediction hardware is explained and the implementation results are given in Section III. Finally, Section IV presents the conclusions.

II. HEVC INTRA PREDICTION ALGORITHM AND THE COMPUTATION AND ENERGY REDUCTION TECHNIQUE

HEVC intra prediction algorithm predicts the pixels in PUs of a coding unit (CU) using the pixels in the available neighboring PUs [1]. For luminance component of a frame, 4x4, 8x8, 16x16 and 32x32 PU sizes are available. There are 33 angular prediction modes corresponding to different prediction angles for each PU size. There are also DC and planar prediction modes for each PU size. An 8x8 PU, four 4x4 PUs in it, and their neighboring pixels are shown in Fig. 1.
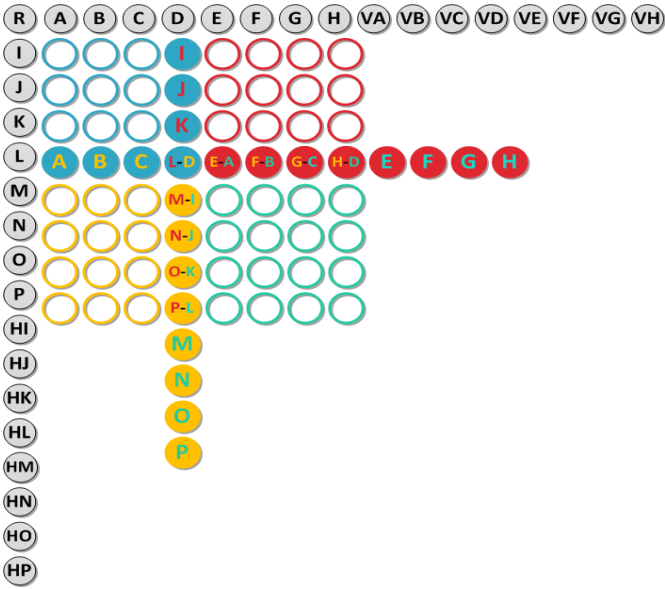
Fig. 1. Neighboring pixels of 4x4 and 8x8 PUs.

TABLE I
PREDICTION EQUATION REDUCTIONS BY DATA REUSE

	4x4 PU	8x8 PU	16x16 PU	32x32 PU	32x32 CU
# of P. Equations	528	2112	8448	33792	135168
# of P. Equations with Data Reuse	201	593	1507	3735	14848
Reduction (%)	61.93	71.92	82.16	88.94	89.02

In HEVC intra prediction algorithm, first, reference main array is determined. The pixels in the reference main array are used in the intra prediction equations. If the prediction mode is equal to or greater than 18, reference main array is selected from above neighboring pixels. However, first four pixels of this array are reserved to left neighboring pixels, and if prediction angle is less than zero, these pixels are assigned to the array. If the prediction mode is less than 18, reference main array is selected from left neighboring pixels. However, first four pixels of this array are reserved to above neighboring pixels, and if prediction angle is less than zero, these pixels are assigned to the array.

After the reference main array is determined, iIdx which is used to determine positions of the pixels in this array that will be used in the intra prediction equations and iFact which is used to determine coefficients of these pixels are calculated as shown in (1a) and (1b), respectively. If iFact is equal to 0, neighboring pixels are copied directly to predicted pixels. Otherwise, predicted pixels are calculated as shown in (2). All the intra prediction equations can be obtained from (2).

$$iIdx = \left((y + 1) * Angle\right) \gg 5 \tag{1a}$$

$$iFact = \left((y + 1) * Angle\right)\& \, 31 \tag{1b}$$

$$pred[x, y] = ((32 - iFact) * \tag{2}$$
$$refMain[x + iIdx + 1] + iFact *$$
$$refMain[x + iIdx + 2] + 16) \gg 5$$

$$x = 0 \; to \; (PU_{size} - 1), y = 0 \; to \; (PU_{size} - 1)$$

Data reuse technique is first used for reducing amount of computations performed by HEVC intra prediction algorithm in [8]. The same technique is also used in this paper. In HEVC, intra 4x4, 8x8, 16x16 and 32x32 luminance angular prediction modes have identical equations. There are identical equations between luminance angular prediction modes of different PU sizes as well. Data reuse technique calculates the common prediction equations for all 4x4, 8x8, 16x16 and 32x32 luminance angular prediction modes only once and uses the result for the corresponding prediction modes.

There are 33792, 8448, 2112 and 528 prediction equations in 32x32, 16x16, 8x8 and 4x4 luminance angular prediction modes, respectively. As shown in Table I, using data reuse technique, the numbers of prediction equations that should be calculated for 32x32, 16x16, 8x8 and 4x4 luminance angular prediction modes are reduced to 3735, 1507, 593 and 201, respectively. The number of prediction equations that should be calculated for a 32x32 CU is reduced from 135168 to 14848.

The computation and energy reduction technique proposed in [8] is also used in this paper. This technique reorganizes the HEVC intra prediction equations by utilizing the fact that the sum of the coefficients used in each HEVC angular intra prediction equation is 32. This reduces the amount of computations performed by 4x4, 8x8, 16x16 and 32x32 luminance angular prediction modes. It does not affect the PSNR and bit rate.

Original version of each intra prediction equation requires two multiplications with constants. Both constants are between 1 and 31. Their sum is 32. Reorganized version of each intra prediction equation requires two multiplications with constants. One constant is always 32, and it can be implemented using shift operation. The other constant is between 1 and 16.

An HEVC intra prediction equation and its reorganized version are shown in (3a) and (4a), respectively. As shown in (3b), original intra prediction equation requires six addition and five shift operations. As shown in (4b) and (4c), its reorganized version requires three addition, one subtraction and three shift operations.

$$(9 * A + 23 * B + 16) \gg 5 \tag{3a}$$

$$\left(\begin{array}{c} A + (A \ll 3) + \\ B + (B \ll 1) + (B \ll 2) + (B \ll 4) + 16 \end{array} \right) \gg 5 \tag{3b}$$

$$(32 * B + 9 * (A - B) + 16) \gg 5 \tag{4a}$$

$$temp = A - B \tag{4b}$$

$$\left((B \ll 5) + (temp + (temp \ll 3)) + 16\right) \gg 5 \tag{4c}$$
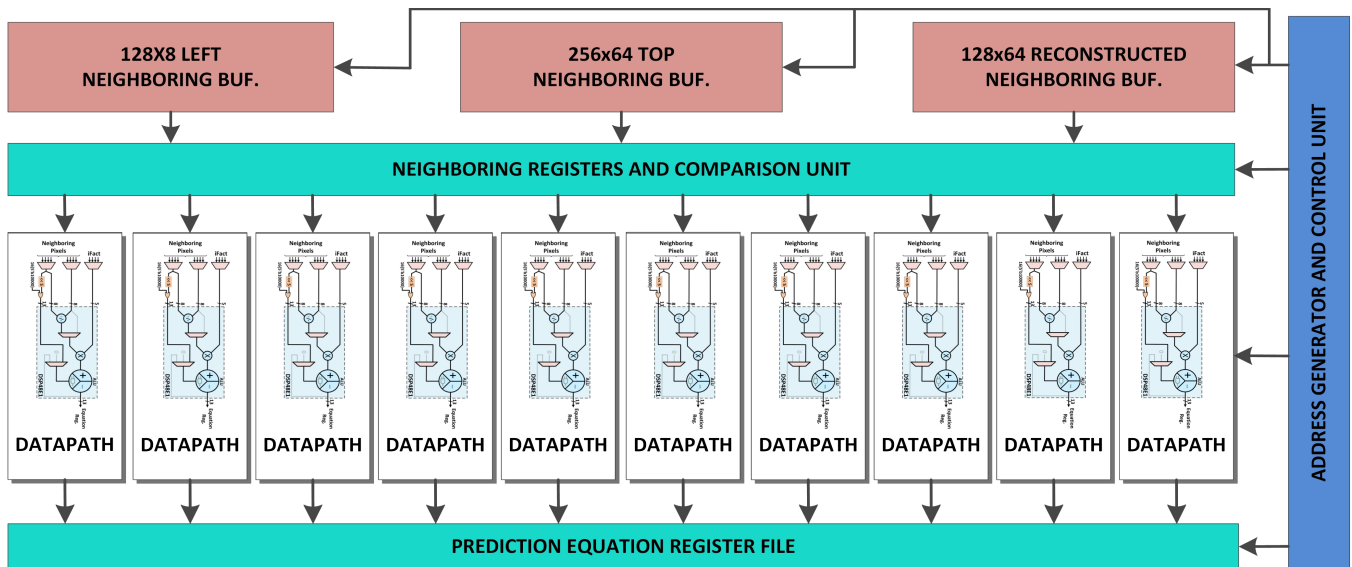
Fig. 2. Proposed HEVC intra prediction hardware.
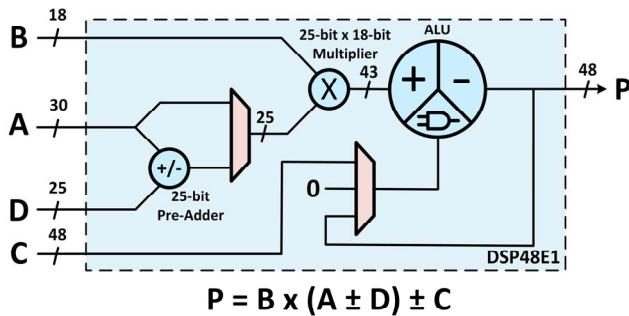

$$P = B \times (A \pm D) \pm C$$
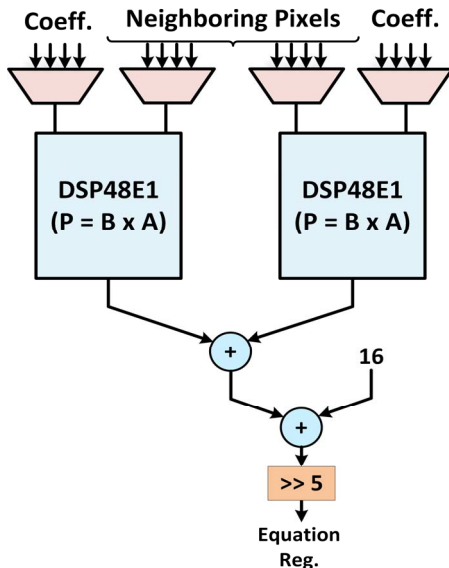
Fig. 3. Structure of a DSP48E1 block.

Fig. 4. Original HEVC intra prediction datapath.

III. PROPOSED HEVC INTRA PREDICTION HARDWARE

The proposed HEVC intra prediction hardware implementing angular prediction modes for all PU sizes (4x4, 8x8, 16x16 and 32x32) using the computation and energy reduction technique proposed in [8] and DSP blocks is shown

in Fig. 2. There are ten pipelined datapaths. Each datapath calculates the result of one intra prediction equation in each clock cycle. Therefore, ten parallel datapaths calculate the results of ten intra prediction equations in each clock cycle.

Three local neighboring buffers are used to store neighboring pixels in the previously coded and reconstructed neighboring PUs. After a PU in the current CU is coded and reconstructed, the neighboring pixels in this PU are stored in the corresponding buffers. These on chip neighboring buffers reduce required off-chip memory bandwidth. The predicted pixels are stored in the prediction equation register file.

A 32x32 CU has 528 neighboring pixels. Storing all 528 neighboring pixels in 528 registers would increase the hardware area. In order to reduce the hardware area, 32x32 CU is split into 8x8 blocks and prediction equations, regardless of their PU sizes, are divided into groups based on the pixels they use. The prediction equations using pixels from the same 8x8 block are grouped together. In this way, only neighboring pixels of current 8x8 block and corresponding four 4x4 blocks are stored in 42 registers. After these neighboring pixel registers are loaded in 16 clock cycles, ten parallel datapaths are used to calculate the prediction equations for current 8x8 block and corresponding four 4x4 blocks.

In an FPGA implementation, multiplication operations in the intra prediction equations can be implemented more efficiently using DSP blocks instead of using adders and shifters. Structure of a DSP48E1 block is shown in Fig. 3. If constant multiplications are implemented using adders and shifters, 10 adders and 10 multiplexers are necessary to implement one original intra prediction equation [8]. If constant multiplications are implemented using DSP blocks, as shown in Fig. 4, two DSP blocks and two adders are necessary to implement one original intra prediction equation.

However, as shown in Fig. 5, one reorganized intra prediction equation can be implemented using only one DSP block. The DSP block is configured to perform multiplication and addition operations. For example, reorganized intra prediction equation shown in (4a) is implemented using a DSP block as follows. $\left(9 * (A - B)\right)$ is implemented using part of

the DSP block implementing $B * (A \pm D)$. One neighboring pixel is shifted left by 5 and ORed with 16 to implement $(32 * B + 16)$ and the result is given to C input of DSP block. Since the last 5 bits of $32 * B$ is zero, $(32 * B + 16)$ can be implemented by changing 5th bit of $32 * B$ from zero to one.

In this paper, an HEVC intra prediction hardware implementing angular prediction modes for all PU sizes (4x4, 8x8, 16x16 and 32x32) using the original intra prediction equations and DSP blocks is also designed for comparison. Both HEVC intra prediction hardware designs are implemented using Verilog HDL. The Verilog RTL codes are verified with RTL simulations. RTL simulation results matched the results of HEVC intra prediction implementation in HEVC HM software encoder [15].

The Verilog RTL codes are synthesized and mapped to a Xilinx XC6VLX75T FF1759 FPGA with speed grade 3 using Xilinx ISE 14.7. FPGA implementations are verified with post place and route simulations. Post place and route simulation results matched the results of HEVC intra prediction implementation in HEVC HM software encoder [15].

FPGA implementation results of HEVC intra prediction hardware using original intra prediction equations and adders and shifters (ORG_AS) [8], reorganized intra prediction equations and adders and shifters (REORG_AS) [8], original intra prediction equations and DSP blocks (ORG_DSP), reorganized intra prediction equations and DSP blocks (REORG_DSP) are shown in Table II.

Power consumptions of all FPGA implementations are estimated using Xilinx XPower Analyzer tool. Post place and route timing simulations are performed for Tennis and Kimono videos at 100 MHz [16], and signal activities are stored in VCD files. These VCD files are used for estimating power consumptions of FPGA implementations.

Energy consumption results of all FPGA implementations for one frame of each video quantized with three different quantization parameters (QP) are shown in Fig. 6. The proposed FPGA implementation of HEVC intra prediction using the computation and energy reduction technique proposed in [8] and DSP blocks has up to 15.97% less energy consumption than the FPGA implementation of HEVC intra prediction using the computation and energy reduction technique proposed in [8] and adders and shifters. The proposed FPGA implementation of HEVC intra prediction using the computation and energy reduction technique proposed in [8] and DSP blocks has up to 34.66% less energy consumption than the FPGA implementation of HEVC intra prediction using original prediction equations and DSP blocks.

Comparison of the proposed FPGA implementation of HEVC intra prediction using the computation and energy reduction technique proposed in [8] and DSP blocks with the FPGA implementations of HEVC intra prediction hardware proposed in the literature is shown in Table III. Area of the proposed FPGA implementation is smaller than the ones proposed in [8]-[12]. Power consumptions of the HEVC intra prediction hardware proposed in [9]-[12] are not reported. The proposed FPGA implementation consumes less power than the one proposed in [8]. Since the HEVC intra prediction hardware proposed in [13] performs intra prediction only for 4x4 and 8x8 PU sizes, it has smaller area and consumes less power than the one proposed in this paper.

Some of the HEVC intra prediction hardware have higher performance than the one proposed in this paper at the expense of much larger hardware area. Frames per second performance of the HEVC intra prediction hardware proposed in [12] is not reported. Since the HEVC intra prediction hardware in [10] is proposed for an HEVC decoder, its frames per second performance for an HEVC encoder is not reported.
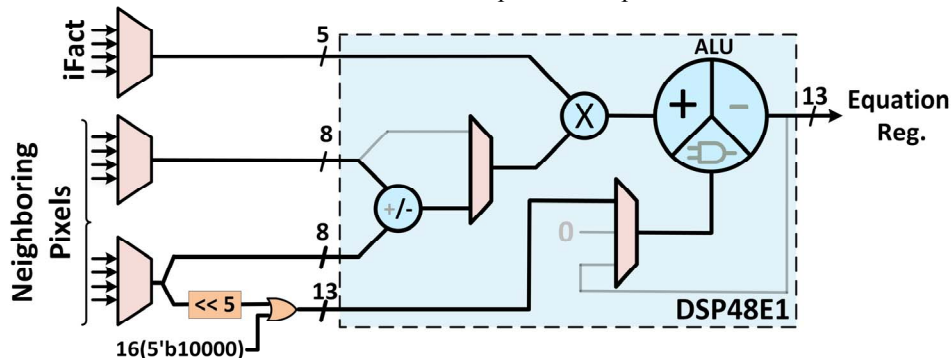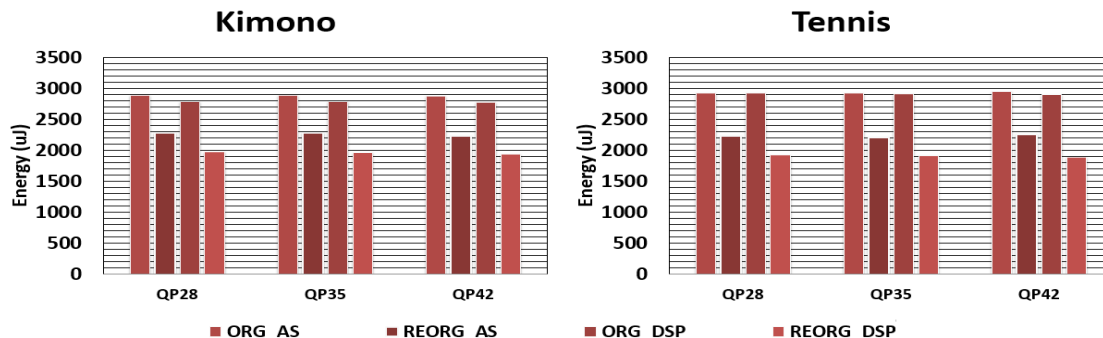

Fig. 5. Proposed HEVC intra prediction datapath.

Fig. 6 Energy consumption results.

TABLE II
IMPLEMENTATION RESULTS

	ORG_AS [8]	REORG_AS [8]	ORG_DSP	REORG_DSP
FPGA	Xilinx Virtex 6	Xilinx Virtex 6	Xilinx Virtex 6	Xilinx Virtex 6
DFF	2567	2006	1167	1168
LUT	5521	6013	4510	4425
BRAM	4	4	4	4
DSP48E1	---	---	20	10
Max. Freq. (MHz)	166	166	212	227
Frames per Second	40 1920x1080	40 1920x1080	52 1920x1080	55 1920x1080
PU Size	4, 8, 16, 32	4, 8, 16, 32	4, 8, 16, 32	4, 8, 16, 32

TABLE III
COMPARISON OF FPGA IMPLEMENTATIONS

	[8]	[9]	[10]	[11]	[12]	[13]	**Proposed**
FPGA	Xilinx Virtex 6	65 nm FPGA	Xilinx Zynq 7045	Xilinx Virtex 6	Altera Arria II GX	Xilinx Virtex 6	Xilinx Virtex 6
DFF	2006	5.5 K	22 K	110 K	6934	849	1168
LUT	6013	14 K	43 K	170 K	13409	2381	4425
BRAM	4	---	94	---	---	4	4
DSP48E1	---	---	---	---	8	---	10
Max. Freq. (MHz)	166	110	150	219	162	150	227
Frames per Second	40 1920x1080	30 3840x2160	---	24 3840x2160	---	30 1920x1080	55 1920x1080
PU Size	4, 8, 16, 32	4, 8, 16, 32	4, 8, 16, 32	4, 8, 16, 32	4, 8, 16, 32	4, 8	4, 8, 16, 32

IV. CONCLUSION

In this paper, an efficient FPGA implementation of HEVC intra prediction for angular prediction modes of all PU sizes is proposed. It uses the computation and energy reduction technique proposed in [8]. However, it implements intra prediction equations using DSP blocks instead of using adders and shifters. The proposed FPGA implementation, in the worst case, can process 55 Full HD (1920x1080) video frames per second. It has up to 34.66% less energy consumption than the original FPGA implementation of HEVC intra prediction.

REFERENCES

[1] High Efficiency Video Coding, ITU-T Rec. H.265 and ISO/IEC 23008-2 (HEVC), ITU-T and ISO/IEC, April 2013.

[2] E. Kalali, A. C. Mert, I Hamzaoglu, "A Computation and Energy Reduction Technique for HEVC Discrete Cosine Transform", *IEEE Trans. on Consumer Electronics*, vol. 62, no. 2, pp. 166-174, May 2016.

[3] E. Kalali, E. Ozcan, O. M. Yalcinkaya, I Hamzaoglu, "A Low Energy HEVC Inverse Transform Hardware", *IEEE Trans. on Consumer Electronics*, vol. 60, no. 4, pp. 754-761, Nov. 2014.

[4] E. Ozcan, E. Kalali, Y. Adibelli, I. Hamzaoglu, "A Computation and Energy Reduction Technique for HEVC Intra Mode Decision", *IEEE Trans. on Consumer Electronics*, vol.60, no.4, pp.745-753, Nov. 2014.

[5] E. Kalali, Y. Adibelli, I. Hamzaoglu, "A Reconfigurable HEVC Sub-Pixel Interpolation Hardware", *IEEE Int. Conf. on Consumer Electronics – Berlin (ICCE-Berlin)*, Sep. 2013.

[6] J. Jeong, S. Kim, J. M. Moon, Y. H. Kim, "Fast Intra Mode Decision by Estimating The Lower Bound on The Rate-Distortion Cost for HEVC", *IEEE Int. Conf. on Consumer Electronics (ICCE)*, Jan. 2017.

[7] J. Vanne, M. Viitanen, T.D. Hämäläinen, A. Hallapuro, "Comparative Rate-Distortion-Complexity Analysis of HEVC and AVC Video Codecs", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp.1885-1898, Dec. 2012.

[8] H. Azgin, E. Kalali, I Hamzaoglu, "A Computation and Energy Reduction Technique for HEVC Intra Prediction", *IEEE Trans. on Consumer Electronics*, vol. 63, no. 1, pp. 36-43, Feb. 2017.

[9] B. Min, Z. Xu, R. C. C. Cheung, "A Fully Pipelined Hardware Architecture for Intra Prediction of HEVC", *IEEE Trans. on Circuits and Systems for Video Technology*, July 2016.

[10] M. Abeydeera, M. Karunaratne, G. Karunaratne, K. De Silva, A. Pasqual, "4K Real Time HEVC Decoder on FPGA", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 236-249, Jan. 2016.

[11] F. Amish, E. B. Bourennane, "Fully Pipelined Real Time Hardware Solution for High Efficiency Video Coding (HEVC) Intra Prediction", *Journal of System Architecture*, vol. 64, pp. 133-147, March 2016.

[12] M. U. K. Khan, M. Shafique, M. Grellert, J. Henkel, "Hardware-Software Collaborative Complexity Reduction Scheme for The Emerging HEVC Intra Encoder," *Design, Automation and Test in Europe (DATE) Conference*, pp. 125-128, March 2013.

[13] E. Kalali, Y. Adibelli, I. Hamzaoglu, "A High Performance and Low Energy Intra Prediction Hardware for High Efficiency Video Coding", *Int. Conf. on Field Programmable Logic and Applications (FPL)*, pp. 719-722, Aug. 2012.

[14] A. C. Mert, E. Kalali, I. Hamzaoglu, "An FPGA Implementation of Future Video Coding 2D Transform", *IEEE Int. Conf. on Consumer Electronics – Berlin (ICCE-Berlin)*, Sep. 2017.

[15] K. McCann, B. Bross, W.J. Han, I.K. Kim, K. Sugimoto, G. J. Sullivan, "High Efficiency Video Coding (HEVC) Test Model 15 (HM 15) Encoder Description", *JCTVC-Q1002*, June 2014.

[16] F. Bossen, "Common test conditions and software reference configurations", *JCTVC-I1100*, May 2012.