Reconfigurable Architecture for 1-D Discrete Cosine Transform of the HEVC Emerging Video Encoding Standard

Hecktheuer, Bruno; Conceição, Ruhan; Wrege, Gustavo; Souza, José Cláudio; Jeske, Ricardo; Agostini, Luciano; Mattos, Júlio C. B.

{bbhecktheuer, radconceicao, gwgoncalves, jcdsouza, rgjeske, agostini, julius}@inf.ufpel.br

Group of Architectures and Integrated Circuits – GACI
Technological Development Center - CDTEC
Federal University of Pelotas - UFPEL
Pelotas, Brazil

Abstract

This paper presents a reconfigurable architecture implementation for 1-D discrete cosine transform (DCT) of the emergent standard video coding – HEVC - High Efficiency Video Coding. The proposed architecture enables to execute all the variable size transforms (from 4x4 to 32x32) in a unique architecture by the hardware reconfiguration. Thus, this architecture is able to reuse the operations contained in different transforms sizes producing better results in terms of area consumption. The biggest size (32x32) of transform is used as base to the design implementation. The architecture was described in VHDL and synthesized using Xilinx ISE 10.1 tool and the Virtex 5 device.

1. Introduction

Nowadays, there are a number of devices that support high definition videos, such as cell phones, personal computers and so on. This market has been growing significantly through the last years due the internet and real time transmissions demand.

There is a huge growth of information generated by these devices, so it is necessary data compression to reduce the necessary space to store this information. Video encoding is one of the most important data compression techniques. The main objective of this technique is to minimize the volume of data required to represent videos while maintaining the quality [1]. For example, it is necessary 19GB to store a 10 minutes video - with 480p resolution, 30 frames per second, using 24 bits for pixel – making impossible to handle these videos. The data compression removes replicated information at the video frames, and also removes imperceptible information to the human eyes.

H.264/AVC [2] is the currently standard of video encoding. This standard was developed targeting up to 1080p. Nowadays, videos with higher resolution, like the QFHD (2160p), are being used. Thus, the JCT-VC (Joint Collaborative Team – Video Coding) starts the developed of a new video coding standard, the HEVC (High Efficiency Video Encoding) [3]. The goal of the JCT-VC is to increase video compression in 50% while maintaining the same computational complexity. This standard will be released in January, 2013, although several coding tools are available.

On HEVC, each frame is divided into a sequence of square units called treeblocks, which hold the information of chrominance and luminance. The chrominance blocks dimensions depend on the color sampling used. Each treeblock is composed of one or more basic Coding Units (CU). A CU can be recursively divided into four blocks of the same size starting from the treeblock and going all the way down to a minimum of 8x8 samples. This recursive process forms a quadtree composed of CU blocks, assuming dimensions that vary from 8x8 pixels to the size of the treeblock itself, in other words, 64x64.

The basic units for the transform and quantization operations on HEVC are called *Transform Units* (TU). Their format is always square and their dimensions can vary from 4x4 to 32x32 samples. As occurs with the CUs, TUs can be structured with quadtrees. Each CU can contain one or more TUs.

This work presents a reconfigurable architecture for a specific module of video coding. The transforms stage is one of the innovations proposed by HEVC ranging from 4x4 to 32x32 and the proposed architecture aims to execute all the variable size transforms (from 4x4 to 32x32) in a unique architecture by the hardware reconfiguration.

The reference software provided by JCT-VC group is used to develop the architecture. The proposed architecture is based on the 1-D DCT of 32 samples, thus it is possible to execute all the mathematical processes for the other sizes of DCT. The architecture is able to reconfigure the transform module dynamically, supporting all blocks size stipulated by HEVC.

The aim of the reconfigurable architecture implementation is to reduce the area consumption for the transform module, since this step is applied several times during the encoding process. With this unique

reconfigurable architecture, it is possible the reuse of operations contained in all size transforms (from 4x4 to 32x32) standardized by HEVC and it is not necessary the implementation of each individual transform.

The architecture was described in VHDL and synthesized using Xilinx ISE 10.1 tool and the Virtex 5 device. The results area presented in terms of area and frequency.

This paper is organized as follows: section 2 presents the state of the art and related works, section 3 presents the developed architecture, section 4 presents the results and, finally, section 5 presents the conclusions and future work.

2. State of the art in Video Encoding

The state of the art in video encoding is the standard H.264/AVC. The results obtained using this standard, in terms of compression, is better than others standards, such as MPEG-2 [1]. However as consequence, the H.264/AVC has more computational complexity, generating more data processing.

For H.264/AVC standard, there are several works in the literature that proposes hardware solutions for the encoding modules. Wang [4] proposed a reconfigurable VLSI architecture for multiples transforms, supporting MPEG – 2/4, VC-1, H.264/AVC and AVS standards. The architecture processes not only the forward, but also the inverse transform and it is also able to process Full HD videos in real time.

H.264/AVC supports resolutions until Full HD (1080p). But electronic gadgets manufactures has been developing devices able to obtain resolution higher than Full HD, such as QFHD (Quad Full High Definition – 3840 x 2160 pixels). For this, the new video encoding standard (HEVC) will be able to support resolutions higher than Full HD. HEVC intend to increase video compression in 50% compared to H.264/AVC, keeping the same computational complexity and image quality [3].

Nowadays, it is not found works related to reconfigurable transforms to HEVC. This fact occurs because this standard still under development.

3. Proposed Reconfigurable Architecture

The process of video encoding is composed of several steps, such as inter and intra prediction, entropy coding, transform calculus, quantization and so on. The input of the transform block - where our paper focuses - is the residuals generated by the prediction block [1], which will be transformed to the frequency domain. The developed architecture implements the DCT 1-D, present in transform block.

The methodology to develop the proposed architecture has started by the analysis of the reference software [7] (HEVC Model - HM) available by JCT-VC group. So, it was able to be seen that there are several calculations processed in the same way by the different transform sizes. Thus, it was proposed an architecture that is able to process the every transform sizes defined by the HEVC.

This first reconfigurable transform version has as an objective process all block sizes using the same architecture. The changing of the 1-D DCT size to be performed occur in process time, this way, the architecture is reconfigurable. So, the architecture proposed at this work process less data volume than if compared with all single architecture that process one size of transform. The 32 DCT 1-D defined by the standard is used as a base hardware to the reconfigurable architecture.

Process to calculate the transform 32 DCT 1-D, firstly are calculated 16 sums as follows: the firs position (number 0) of the vector is added to the latest (number 31) position; the second position (number 1) is added to the penultimate (number 30) and so on; until the number 15 and 16 are added. These sums will be named at this paper “mirrored sums”.

The sixteen results obtained by the mirrored sums will be stored in a second vector named “E”. The same way that the process is done by the sums, a “mirrored subtractions” are made. The results of the subtraction are stored in a vector named “O”. The vectors “E” and “O” are illustrated in figure 3.

The mirrored operations (sums and subtractions) are repeated only in vector generated from mirrored sums, until the resulting vector has two positions. For each sum stage, the resulting vector receives the name of its origin vector concatenated to the letter “E”. It also happens with the subtractions, but is concatenated to the letter “O”.

In this way, the sixteen values stored in the vector “E” generate the eight values of the vector “EE” (by the sums), and the eight values of the vector “EO” (by the subtractions). These mirrored operations are repeated in the vector “EE”, generating vectors with four positions named “EEE” and “EEO”. Vectors generated from subtractions are used to calculate the output of certain positions.

Figure 3 illustrates the dependences between the output vector and the intermediate vectors. It also illustrates the dependence between each intermediate vectors.

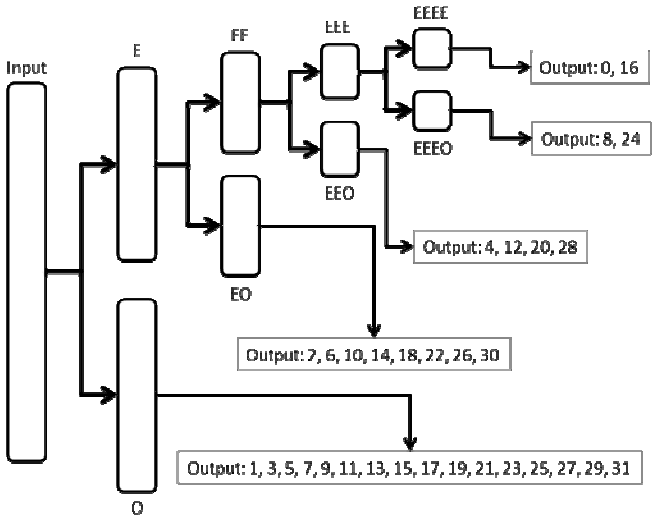

Fig.2 - Dependence between the vectors

It is able to be seen, the output uses all vectors generating from the subtractions, including the vector "EEEE". For each calculation, it is added all the positions of the related vector where each of them must be multiplied to corresponding coefficient matrix. For example, for each odd position of output vector, the result is obtained as presented in equation 1.

$$Out_n = \sum_{i=0}^{15} O_i \times CM_{n,i} \quad (1)$$

The 16 DCT 1-D performs the same process of mirrored operations, until obtain the vector of length 2. Also the coefficient matrix of size 32x32 has values of coefficient matrix of size 16x16 in the even lines. Thus, it is possible to shift the input values of 16 1-D DCT to the vector "E" of 32 DCT 1-D and the result will be in the even positions of the output vector.

This technique can be applied in each transform sizes. To perform the operations of the 8 DCT 1-D, the input is shifted to the vector "EE" and the output will be present in the positions multiple of four. The same happens to perform the 4 DCT 1-D, shifting the input to the vector "EEE", and the output will be present in the positions multiple of eight.

The architecture has two modules. The first is where are calculated de mirrored sums and subtractions and where are done the shifts to intermediate vectors. Addition to this module, the architecture developed has another module, which is responsible to generate the multiplications by the coefficient matrix. This module uses sums and shift operations instead of multiplying operations since the multipliers are very expensive in terms of hardware consumption.

The control unit present in the 32 DCT 1-D reconfigurable receives an external signal, indicating the size of transformed to be processed. Based on this signal the control unit generates the selection bits for the multiplexers, shifting the input for the intermediate vectors. The reuse of operations is done just by the first module, the second one is purely combinational.

4. Results Obtained

The developed architecture was described in VHDL (functional description) and synthesized using Xilinx ISE 10.1 tool [5] and Virtex 5 device [6]. The modules in the process have been tested and validated for different transform configurations.

The DCT 1-D proposed in this paper uses as base architecture the transform 32x32. Using this architecture as base, it is possible to perform all other transform sizes (from 4x4 to 32x32) standardized by the HEVC. This is possible because of replication of operations between the different sizes.

The reconfiguration is the best advantage of proposed architecture because it provides the way to reuse the architecture to different transform size. However, there is another advantage of this architecture, the larger reduced are by the use of sums and shifts instead of multiplying.

The designed architecture uses different sizes of adders and subtractors. Table 1 shows total number of adders and subtractors classified by its size produced by the ISE Tool. There are adders/ subtractors from 9 to19 bits.

In addition to using these functional units available by the device, an adder/subtractor was developed. This adder/subtractor is based on traditional "and" gates to realize the sums/subtractions and a "xor" gate to define the operation.

Tab.1 – Number of adders and subtractors

Size(bits)	9	10	11	10	13	14	15	16	17	18	19
Number	15	7	3	2	32	100	203	336	84	26	6

The design is divided in two parts. The first part (module 1) is responsible by the mirrored operations; the second part (module 2) is responsible by the operations with coefficient matrix. Table 2 shows the results in terms of area and frequency synthesized using the Virtex 5 device.

Tab.2 – Modules of the project in terms of area an frequency

Modules	***LUTs***	***Frequency (MHz)***
Module 1	1034	74.11
Module 2	19607	54.73
Complete	20855	37.74

The module 1 (responsible by the mirrored operations) is less expensive in terms of hardware (uses only 1034 LUTs) and achieves higher frequency (74.11MHz). The module 2 (responsible by the operations with coefficient matrix) is more expensive than the module one, since it realized every sums and shifts with coefficient matrix with different sizes (module 1 uses only sums and subtractions up to 13 bits). Synthesizing both modules together the total area is 20855 LUTs and the frequency can achieve 37.74MHz.

5. Conclusion

This paper has presented a reconfigurable architecture implementation for 1-D discrete cosine transform (DCT) of the emergent standard video coding – HEVC - High Efficiency Video Coding. The algorithm was studied and implemented following the HEVC Model. The DCT 1-D reconfigurable architecture was implemented using VHDL and synthesized and validated using ISE Xilinx 10.1 tool and Virtex 5 device.

The developed reconfigurable architecture enables to use only one module for all size DCTs. Thus, it is not necessary the implementation of different size transforms. Using the 32x32 transform size as the base architecture is it possible to perform all operations necessary for all sizes of DCT.

The reconfiguration of the architecture is basically the use of multiplexors. So, they change the data path. This way, it is able to calculates the transform of the all sizes (4x4, 8x8, 16x16, 32x32) using the same hardware.

The main advantage of this architecture is the reuse of the operations, reducing area. Moreover, the design was done using sum and shifts instead multipliers, saving more area.

For future work, we intend to implement an architecture implementation using pipelines and develop the module of DCT 2-D using the same engine of this work.

6. References

[1] Agostini, L., “Desenvolvimento de Arquiteturas de Alto Desempenho Dedicadas à Compressão de Vídeo Segundo o Padrão H.264/AVC”,Doutorado em Computação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

[2] Richardson, I. E. G., “H.264/AVC and MPEG-4 Video Compression – Video Coding for Next-Generation Multimedia”, Chichester: John Wileyand Sons.

[3] Joint Collaborative Team on Video Coding (JCT-VC), avaliable on <http://phenix.int-evry.fr/jct/index.php>, August, 2007.

[4] Wang, K., Chen, J. Cao, W., Wang Y., Wang, L., Tong, J. “A Reconfigurable Multi-Transform VLSI Architecture Supporting Video Codec Design”, IEEE Transaction on Circuits and Systems, vol. 58, no. 7, pages 432-436, July, 2011.

[5] Xilinx, “ISE Design Suite 10.1 Software Tutorials”, avaliable on < http://www.xilinx.com/>, November, 2011.

[6] Xilinx, “FPGA, CPLD, and EPP Solutions from Xilinx, Inc.”, avaliable on < http://www.xilinx.com/>, November, 2011.

[7] Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 – “HM3: High Efficiency Video Coding (HEVC) Test Model 3 Encoder Description”, 5th Meeting: Geneva, CH, 16-23 March, 2011.