HIGH EFFICIENT BIT-PLANE DECOMPOSITION MATRIX BASED VLSI FOR HEVC

[1]P.MANIKANTA, [2]A. RAGHAVA RAJU
[1]M.Tech Scholar, Dept. of E.C.E, Chebrolu Engineering College, Guntur DT, A.P, India
[2]Associate Professor, Dept. of E.C.E, Chebrolu Engineering College, Guntur DT, A.P, India

ABSTRACT: Basically, new very-large-scale integrated (VLSI) integer transform architecture is proposed for the High Efficiency Video Coding (HEVC) encoder. The architecture is designed based on the signed bit-plane transform (SBT) matrices, which are derived from the bit-plane decompositions of the integer transform matrices in HEVC. Mathematically, an integer transform matrix can be equally expressed by the binary weighted sum of several SBT matrices that are only composed of binary 0 or ±1. The SBT matrices are very simple and have lower bit width than the original integer transform in the form. The SBT matrices are also sparse and there are many zero elements. The sparse characteristic of SBT matrices is very helpful for saving the addition operators of SBT. In the proposed architecture, instead of the original integer transform in high bit width; the video data can be respectively transformed with the SBT matrices in lower bit width. As a result, the delay of the transform unit circuit can be significantly reduced with the proposed SBT.

KEY WORDS: Bit-plane matrix, High Efficiency Video Coding (HEVC), integer transform, ultra high definition (HD), very-large-scale integrated (VLSI) architecture.

I.INTRODUCTION

A picture is worth a thousand words. This expresses the essential difference among human ability to perceive linguistic information and visual information. For the same message, a visual representation tends to be perceived as being more efficient than the spoken or written words. The processing of language is inherently serial. Words and their meanings are recorded or perceived one at a time in a causal manner. In the mammalian visual system, this parallelism is evident from the retina right through to the higher-order Structures in the visual cortex and beyond. For example: video conferences, medical data transfer, business data transfer and so On, require much more image data to be transmitted and stored on-line. Due to the internet, the huge information transmissions take place. The processed data required much more storage, computer processor speed and much more bandwidth for transmission.

In the rapidly growing field of Internet applications, not only still images but also small image sequences are used to enhance the design of private and commercial web pages. Meeting bandwidth requirements and maintaining acceptable image quality simultaneously are a challenge. Wavelets are mathematical functions that provide good quality compression at very high compression ratios, because of their ability to decompose signals into different scales or resolutions. The standard methods of image compression come in numerous ranges. Most of the well-established compression schemes use the bi-variate Discrete Wavelet Transform (DWT) on wavelet-based image coding. At high compression rates, wavelet-based methods provide much better image quality in comparison with the JPEG (Joint Photographic Experts Group) standard, which relies on the discrete cosine transform (DCT). The good results obtained from DWT are due to multi-resolution analysis, which essentially brings out information about the statistical structure of the image data. The current most popular methods rely on removing high frequency components of the image by storing only the low frequency components (e.g., DCT based algorithms). Although some information loss can be tolerated in most of these applications, there is certain image processing applications that demand no pixel

difference between the original and the reconstructed image.

Fractal image compression is a lossy compression method, so there will be data losses in compressed image. For fractal coding, an image is represented by fractals rather than pixels. Each fractal is defined by a unique Iterated Function System (IFS) consisting of a group of affine transformations. Therefore, the key point for fractal coding is to find fractals which can best approximate the original image and then to represent them as a set of affine transformations. Standard fractal coding methods rise above many other image coding techniques in the sense that it maintains high image quality after decoding but presents high compression ratios during encoding. Rather than lossy compression with relatively high compression ratio, mathematical lossless compression techniques are favoured in this field. A lossless scheme typically achieves a compression ratio of the order of two, but will allow exact recovery of the original image from the compressed version.

II. BASIC IMAGE COMPRESSION SCHEMES

For a universal algorithm to compress images, a sequence of image pixels extracted from an image in the raster scan order is simply encoded. But, for a universal algorithm such a sequence is hard to compress. Universal algorithms are usually designed for alphabet of sizes not exceeding 28 and do not exploit directly the image data features. As images are 2-dimensional data, intensities of neighbouring pixels are highly correlated, and the images contain noise added to the image during the acquisition process.

Modern gray-scale image compression algorithms employ techniques used in universal statistical compression algorithms. However, prior to statistical modelling and entropy coding the image data is transformed to make it easier to compress. To make the image data easily compressible, we use 2-dimensional image transforms, such as DCT or wavelet transform. In transform algorithms, instead of pixel intensities, a matrix of transform coefficients is encoded. Transforms can be used for both lossless and lossy compressions. Transform algorithms are more popular in lossy compression. Apart from lossy and lossless compressing and decompressing of whole images, transform algorithms deliver many interesting features such as progressive transmission, region of interest coding, etc. The usages of algorithms are dependent mostly on information content of images and types of application.

Lossless compression algorithms are often predictive in nature. In a predictive algorithm, the predictor function is used to guess the pixel intensities and the prediction errors are calculated. The prediction errors are differences between actual and predicted pixel intensities. To calculate the predictor for a specific pixel usually intensities of a small number of already processed pixels neighbouring it is used. Next, the sequence of prediction errors, called residium, is encoded. Prediction error distribution is close to Laplacian, that is, symmetrically exponential. Therefore, entropy of prediction errors is significantly smaller than that of pixel values. That is why; it is easier to compress residium. In respect to the lossless compression, better results in terms of computational speed are obtained by predictive algorithms.

a) Context Adaptive Lossless Image Coding (CALIC) Algorithm

CALIC obtains higher lossless compression for the continuous-tone images than other techniques reported in the literature. The nonlinear predictor adapts via an error feedback mechanism. The former estimation technique can afford a large number of modelling contexts without suffering from the sparse context problem. CALIC employs a two-step (prediction/residual) approach. In the

prediction step, CALIC employs a simple new gradient based non-linear prediction scheme called gradient-adjusted predictor (GAP), which adjusts prediction coefficients based on estimates of local gradients. Predictions are then made context-sensitive and adaptive by modelling of prediction errors and feedback of the expected error conditioned on properly chosen modelling contexts. The modelling context is a combination of quantized local gradient and texture pattern; the two features that are indicative of the error behaviour. The context-based error modelling is done at a low model cost. By estimating expected prediction errors rather than error probabilities in different modelling contexts, CALIC can afford a large number of modeling contexts without suffering from either context dilution problem or from excessive memory use.

b) Using Single Error Correction Codes to Protect Against Isolated Defects and Soft Errors:

The technology scaling process provides high-density, low cost, high-performance integrated circuits. To cope with defects in memory chips, many different techniques have been proposed, all of them based on the use of redundant elements to replace defective ones. For example, when all remaining defective cells are located in one half of the array, the other half can still be used as a memory with reduced capacity. This reduction is done by permanently setting the most significant bit of the addresses either to 0 or 1, depending on which part of the memory is to be used. However, in most cases, the remaining defective cells are evenly distributed across the whole array, and not clustered in one half of the array, making this technique useless. This change in the voltage level will change the state of the transistor, which will result in a change of the value in a memory cell. For example, if a memory cell holds "1," an SEU will force it to "0. Unfortunately, these techniques will fail in the appearance of multiple cell upsets (MCU). The most common approach to deal with multiple errors has been the use of interleaving in the physical arrangement of the memory cells, so that cells that belong to the same logical word are separated. As the errors in an MCU are physically close as discussed in, they will cause single errors in different words that can be corrected by the Single Error Correction-Double Error Detection (SEC-DED) codes. However, interleaving cannot be used, for example, in small memories or register files, and in other cases, its use may have an impact on floor-planning, access time, and power consumption.

III. EXISTED SYSTEM

Transform is a frequently used module when compressing video; thus, the complexity of a transform has an important effect on the whole complexity of the video encoder. Chen *et al.* derived the factorization relationship between $N \times N$ and $N/2 \times N/2$ DCT matrices by analyzing the periodic property of the cosine function. With the factorization relationship of DCT, the number of arithmetic operations of the transform can be reduced. Ahmed *et al.* decomposed the DCT matrix into sparse sub matrices where the multiplications are avoided by using the lifting scheme. Arai *et al.* proposed an Arai, Agui, and Nakajima (AAN) fast algorithm based on the common factor extraction algorithm in which the complicated common factors were moved from the transform kernel to the scale part. Only five multipliers are required in AAN's transform kernel. The multiplier is expensive against the adder in the integration circuit. Thus, the multiplication operation is usually replaced by adders in the circuit design. Tsui and Chan developed an efficient multiplier less fast Fourier transform (FFT)-like transform based on a recursive noise model that minimizes the hardware resources of the transform while maintaining the high performance. In, a multiplier less hardware implementation

using a second-order cone programming technique is presented, and the dynamic ranges of intermediate data are minimized through geometric programming

The existing transform architectures consider how to reduce the number of arithmetic operators, such as addition and multiplication, more than the data bit width in the transform. In fact, the data bit width is also an important factor impacting on the circuit speed and area of VLSI architecture. A circuit with a large bit width needs a larger number of fan-in or fan-out of logic gate, and more MOS devices are required in the logic gate circuit. Thus, the capacitive load and resistance of the logic gate all increase with widening bit width. According the first-order resistance and capacitance (RC) circuit model theory, the delay of the circuit is related with RC. Large RC leads to long circuit delay. The circuit delay varying with the increasing input bit width in two typical CMOS processes (SMIC40nm and GF28nm). As for the adder, the carry chain is the critical path for the circuit delay, which is also dependent on the input and output bit width.

Thus, aside from the number of arithmetic operations, the bit width is the other optimization factor for fast transform architecture. In this brief, we propose a new VLSI architecture for the integer transforms of the HEVC standard for reducing the bit widths of data. The integer transform matrix is decomposed into several signed bit-plane transform (SBT) matrices that are used in the proposed architecture. Moreover, a number of adders are reused based on the redundant property of elements of bit matrices. With the bit matrix-based transform algorithm, the proposed VLSI transform architecture can process 32 pixels/cycle data throughput maximally with very high working frequency and proper area.

a) Bit-Plane Decomposition of Integer Transform in signed bit matrix-based transform algorithm

In order to narrow the bit width of intermediate transformed data, we propose the bit decomposition algorithm which decomposes the integer transform matrix into several SBT matrices. Applying the existed SBT algorithm to the transform architecture, instead of the integer transform matrix circuits, the SBT matrix circuits are implemented and the input data are transformed with each SBT matrix circuit, respectively. Due to the simple elements of SBT matrices, the bit widths of intermediate transformed data and output data are significantly reduced. The bit width of output data should be $n + \_\log N2 \_$ maximally. Taking the 32 × 32 1-D integer transform as an example, the increasing bit width of output data is only 5 b with the SBT algorithm, compared with the 11-b increasing of the straightforward integer transform. The bit widths of SBT increase slowly as the intermediate data are processed stage by stage, which shortens the circuit delay and constrains the clock cycle to be smaller. Although the delay of the integer transform circuit is reduced based on the proposed bit transform algorithm, more adders are required due to more SBTs.

However, the bit widths of adders used in SBT are also so low that the addition operation is also very fast. Additionally, it can be observed from (6) that many zero elements are in the SBT matrix. The number of actually required addition operations is seldom due to the sparse SBT matrix according to the rule of matrix multiplication. The sparse characteristic of the SBT matrices can benefit for reducing the addition operations in the transform process.

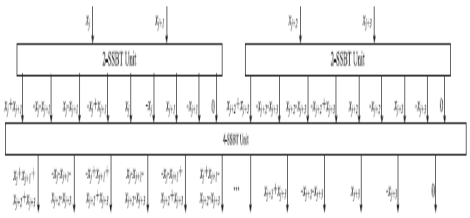

FIG. 1: HIERARCHICAL STRUCTURE OF SBT

There are many addition operation redundancies in SBT. Thus, we propose the adder reuse method based on the element redundancy characteristic of SBT matrices for reducing the number of adders in the next section.

IV.PROPOSED SYSTEM

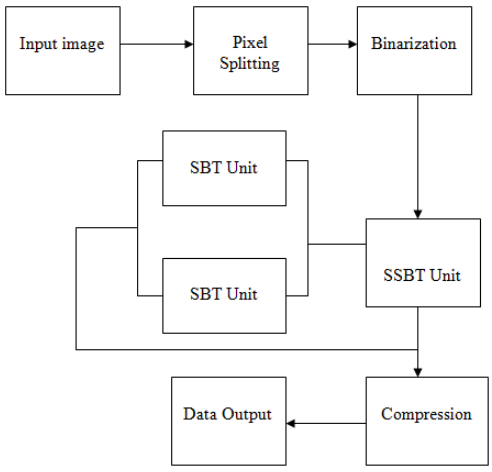

FIG. 2: PROPOSED SYSTEM

The above figure (1) shows the architecture of proposed system. HEVC was developed with the goal of providing twice the compression efficiency of the previous standard, H.264 / AVC. The Discrete Cosine Transform (DCT) has been widely applied in the area of image compression and video compression, such as JPEG, MPEG-2/4 and H.263. Its popularity is attributed to its ability to décor relate data of spatial domain into data of frequency domain. Data will become more compact after being transformed, redundant information can be further removed. The DCT is considered as the closest to K-L transform, which is the ideal energy compaction transform. However, the matrix elements of DCT contain real numbers presented by finite number of bits, which inevitably leads to the possibility of drift (mismatch between the decoded data in the encoder and decoder). In order to eliminate mismatch between encoders and decoders and to facilitate low complexity implementations, latest video standards like H.264, VC-1 and AVS begin to adopt integer transform. High Efficiency Video Coding (HEVC) is the newest standard for high definition video processing. It is considered as the successor of H.264. Main goal of HEVC is to achieve 50% higher coding efficiency than H.264. In order to achieve this goal, HEVC adopts lots of state of the art coding tools including 4/8/16/32 integer transform.

Development of efficient image compression techniques continues to be an important challenge to us, both in academia and in industry. In multiplier based DCT's were implemented, later to reduce area ROM-based DA was applied for designing DCT. Then knowing the advantage of ROM-based, DA-based multipliers using ROMs were implemented to produce partial products together with adders that accumulated these partial products. In DA-based computation partial products words are shifted and added in parallel. We need to reduce truncation error that error is introduced if the least significant part indirectly truncated. In order to reduce truncation error effect several error compensation bias methods have been presented based on statistical analysis of relationship between partial product and multiplier-multiplicand.

Distributed arithmetic is a bit level rearrangement of a multiply accumulate to hide the multiplications. It is a powerful technique for reducing the size of a parallel hardware multiply accumulates that is well suited to FPGA designs. The Discrete cosine transform (DCT) is widely used in digital image processing for image compression, especially in image

transform coding. Despite the many advantages of digital representation of signals compared to the analogy counterpart, they need a very large number of bits for storage and transmission. For example, a high quality audio signal requires approximately 1.5 megabits per second for digital representation and storage. The storage requirement for upcoming high-definition television (HDTV) of resolution 1280 x 720 at 60 frames per second is far greater. A digitized one-hour colour movie of HDTV-quality video will require approximately 560 gigabytes of storage. A digitized 14 x 17 square inch radiograph scanned at 70 pm occupies nearly 45 megabytes of storage. Transmission of these digital signals through limited bandwidth communication channels is even a greater challenge and sometimes impossible in its raw form.

The main advantage of compression is that it reduces the data storage requirements. It also offers an attractive approach to reduce the communication cost in transmitting high volumes of data over long-haul links via higher effective utilization of the available bandwidth in the data links. This significantly aids in reducing the cost of communication due to the data rate reduction. In systems with levels of storage hierarchy, data compression in principle makes it possible to store data at a higher and faster storage level (usually with smaller capacity), thereby reducing the load on the input output channels. According to the relationship between M-SSBT and M/2- SSBT, the SBT can be implemented in a hierarchical way. An M-SSBT can be implemented through jointing two M/2- SSBTs. The hierarchical structure of 4- SSBT as an example is illustrated where the output of two 2-SSBT units is input to a 4- SSBT unit for 4-SSBT computation. It can be summarized that the number of adders, used in the proposed adder reuse scheme for SBT can be calculated according to the expression.

V. RESULTS

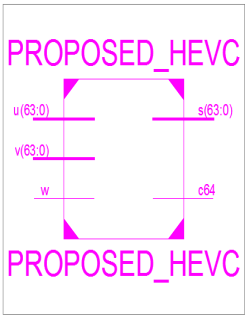

FIG. 3: RTL SCHEMATIC

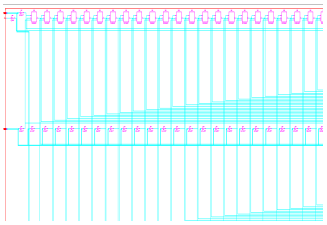

FIG. 4: TECHOLOGY SCHEMATIC

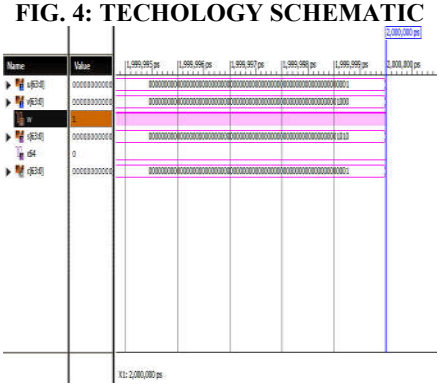

FIG. 5: OUTPUT WAVEFORM

HEVC Project Status (12/02/2018 - 07:56:47)			
Project File:	ET32.xise	Parser Errors:	No Errors
Module Name:	PROPOSED_HEVC	Implementation State:	Synthesized
Target Device:	xc3s100e-5vq100	• Errors:	No Errors
Product Version:	ISE 14.7	• Warnings:	1 Warning (1 new)
Design Goal:	Balanced	• Routing Results:	
Design Strategy:	Xilinx Default (unlocked)	• Timing Constraints:	
Environment:	System Settings	• Final Timing Score:	

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slices	74	960	7%
Number of 4 input LUTs	128	1920	6%
Number of bonded IOBs	194	66	293%

Detailed Reports					
Report Name	Status	Generated	Errors	Warnings	Infos
Synthesis Report	Current	Sun 2. Dec 07:56:45 2018	0	1 Warning (1 new)	0
Translation Report					
Map Report					

FIG. 6: REPORT

VI. CONCLUSION

A fast integer transform VLSI engineering based inadequate SBT is proposed for continuous ultra HD video coding acclimating to the HEVC standard. Considering the bit width impact on circuit delay, the bit width of the number change framework is upgraded in the proposed VLSI engineering. The whole number change lattice with high-piece width components is deteriorated into a few SBT

lattices with low-piece width components based on the proposed lattice bit-plane deterioration strategy. The change engineering with the SBT calculation can work more productively for the low piece width calculations. The circuit reuse system is especially proposed for the SBT to lessen the number of adders of the VLSI architecture. The proposed transform hardware architecture can process video data with higher speed and proper area compared with previous work. A large number of adders for the SBT is saved using the proposed circuit reuse strategy.

VII. REFERENCES

[1] N. Ahmed, T. Natarajan, and K. R. Rao, "On image processing and a discrete cosine transform,"IEEE Trans. Comput., vol. C-23, no. 1, pp. 90–93, Jan. 1974.

[2] G. K. Wallace, "JPEG still image data compression standard," Commun. ACM, vol. 34, no. 4, pp.30–44, Apr. 1991.

[3] "Generic coding of moving pictures and associated audio information Part 2: Video ITU-T andISO/IEC JTC1," in ITU-T Recommendation H.262, ISO/IEC 13818-2 (MPEG-2), Nov. 1994

[3]"Generic coding of moving pictures and associated audio information Part 2: Video ITU-T andISO/IEC JTC1," in ITU-T Recommendation H.262, ISO/IEC 13818-2 (MPEG-2), Nov. 1994.

[4] "Video coding for low bit-rate communication Version 1," in ITU-T Recommendation H.263, Nov. 1995.

[5] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC videocoding standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560–576, Jul. 2003.

[6] L. Yu, S. Chen, and J. Wang, "Overview of AVS-video coding standards," Signal Process., ImageCommun., vol. 24, no. 4, pp. 247–262, Apr. 2009.

[7] G. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency videocoding (HEVC) standard," IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp. 1649–1668,Dec. 2012.

[8] W.-H. Chen, C. H. Smith, and S. C. Fralick, "A fast computational algorithm for the discretecosine transform," IEEE Trans. Commun., vol. COM-25, no. 9, pp. 1004–1009, Sep. 1977.

[9] A. Ahmed, M. U. Shahid, and A. urRehman, "N-point DCT VLSI architecture for emergingHEVC standard," VLSI Design, vol. 2012, 2012, Art. no. 752024.

[10] Y. Arai, T. Agui, and M. Nakajima, "A fast DCT-SQ scheme for images," Trans. IEICE, vol.E71, no. 11, pp. 1095–1097, Nov. 1988