

3D-DCT Chip Design for 3D Multi-view Video Compression

Yu-Cheng Fan, Shan-Ann Chen, Kuo-Gi Wu, and Jun-Lin You

Department of Electronic Engineering, National Taipei University of Technology, Taipei, 106, Taiwan

Corresponding author: Yu-Cheng Fan

Email: skystarfan@ntu.edu.tw

Received Aug. 15, 2011; Revised Sept. 12, 2011; Accepted Sept. 16, 2011

Published online:

Abstract: This paper describes the implementation of three dimensional (3D) data decreasing strategy. Due to the upcoming of 3D generation, the abundant data accompany with multi-view video. Therefore, video compression has become the most important process of dealing with 3D data. The purpose of this paper is to use 3D discrete cosine transform (DCT) to transfer data into frequency domain. In addition, it has the characters of gathering energy together in order to achieve the goal of data compression. First, we transform nine different views into grey to make data become easily to calculate. Second, we propose a disparity estimation algorithm to calculate the different views and block matching full search algorithm (BMFS) to find out the minimum of SAD. That is to say, this is the most similar block. According to the displacements of two images, we can find out its motion vector, and then use motion vector to compensate the present image. Third, we subtract compensated image from the original one and get the image difference of two images. By doing so, the similar part will be deleted, and it will also decrease the resource cost. Fourth, we transform the difference with 3D-DCT. Owing to 3D-DCT, which has the character of gathering energy; hence, all energy will be gathered in one corner. Finally, we quantize to the data, which had already been transformed, and delete the redundant part to achieve the effect of data compression. The result lead to the conclusion that 3D-DCT chip produces efficient results. The scheme speeds up 47% of the chip. Besides, the presented method decreases 14.4% logic gate count of this chip. Therefore, the power consumption decrease 16% at the same time. We adopt TSMC 0.18 um process to achieve the chip and perform 3D-DCT chip design for 3D multi-view video compression.

Keywords: 3D-DCT, 3D Stereoscopic, Compression, Motion Vector Estimation, Multi-view

1 Introduction

Three dimensional television technologies become more and more important today. Multi-view video adopts LCD based technology to provide well three dimensional visions and becomes main three dimensional television products in consumer market. However, multi-view video need huge storage space and is hard to be transmitted. In order to solve the above problem, some literature proposed multi-view compression methods to compress the data. Fan and You proposed a cubic memory based 3D discrete cosine transform for 3D image compression [1] and spatial

domain compression technology of 3D stereoscopic multi-view video [2]. However, the architecture needs larger hardware overhead and power consumption. P. An presented a multi-view video coding based on view prediction [3]. Sgouros adopted an adaptive 3D-DCT scheme for coding multi-view stereo images [4]. Ma developed an efficient compression of multi-view video using hierarchical B pictures [5]. However, hardware architectures are not be addressed. In order to solve the above problems, we proposed a memory-reduction-3D-DCT architecture to improve the

storage efficiency of traditional architecture. The chip has a great improved space on its speed and power because data in the chip are not sequential. The purpose of this paper is to propound a theory of accelerating the speed of 3D-DCT chip. In the meantime we would decrease the consumption of the power and components for this chip.

2 Three Dimensional Multi-view Video Compression

Three dimensional stereoscopic multi-view video capture methods often adopt annular focusing scheme or parallel focusing technique. In this paper, we use annular focusing to take nine view images and use them to perform multi-view compression. In the literature [1][2], three dimensional video compression flows contains six steps in figure 2.1 that include gray level transformation, finding disparity image between two images, getting compensated images, using compensated images to find different images, 3D-DCT processing, and 3D-DCT quantization. Based on the architecture, we design a new architecture to speed up the 3D-DCT architecture and reduce the power consumption in the proposed chip.

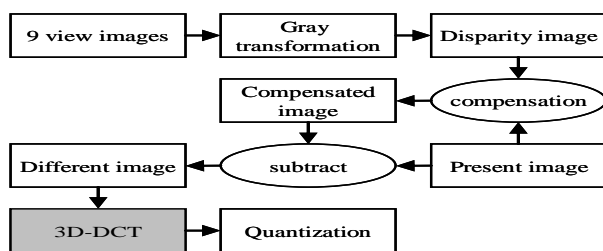


Figure 2.1: 3D Stereoscopic multi-view video compression flows

How to centralize the 3D video data is very important issue in 3D video compression field. In order to achieve the target, we adopt 3D-DCT architecture to raise video compression ratio. Traditionally, researchers adopt 2D-DCT algorithm to compress image and video data [6]. However, 3D video has different characteristics that contain multi-view images [1][2]. Based on the 3D-DCT concepts, we proposed a new architecture to reduce the processing time and power consumption to improve traditional architecture.

At first, we use two-view images as test images in figure 2.2 (a) and (b). Figure 2.2(a) is fifth view of nine-view video and figure 2.2(b) is sixth view of nine-view video. We adopt the two pictures as example to introduce the proposed method.

Next, we transform the color images to gray level images in figure 2.3(a) and figure 2.3(b). The transform function is as equation (2.1) [6].

$$Y = 0.299 \times R + 0.587 \times G + 0.114 \times B \quad (2.1)$$

where Y is luminance value of a pixel, R is red value of a pixel, G is green value of a pixel and B is a blue value of a pixel.

We regard fifth view image as a reference image and sixth view image as a present image. Then we use motion estimation algorithm to search the most similar block by calculating the minimum Sum of Absolute Difference (SAD) between present image and reference image. Then, the disparity image is obtained after calculation (Figure 2.4).



Figure 2.2: Two-view images: (a) 5th view image; (b) 6th view image



Figure 2.3: Gray transform: (a) 5th view image; (b) 6th view image

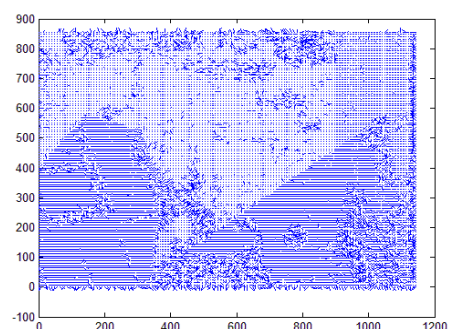


Figure 2.4: Disparity Image between 5th view image and 6th view image

The same strategies are performed to search eight disparity images in nine-view video. We calculate the disparity between first view and second view, second view and third view, third view and fourth view, fourth view and fifth view, sixth view and seventh view, seventh view and eighth view, eighth view and ninth view to gather the difference of each pair.

In order to observe vertical and horizontal change, we gather statistics for disparity images. The results are illustrated in figure 2.5 (a) and figure 2.5(b). The formula of SAD is specified as [6]

$$SAD(x, y) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} |f(n_1, n_2, s) - f(n_1 - x_1, n_2 - y, s - 1)| \quad (2.2)$$

In equation (2.2), $SAD(x, y)$, which is coordinates, shows the area is searched between present image and reference image [6]. n^1 and n^2 represent the dimensions of the block. s stands for Z direction.

Afterward, the compensated image is found by disparity image and reference image in figure 2.6(a) and figure 2.6(b). The more accurate disparity image is generated, the more data will be compressed. After producing compensated images, we subtract compensated images with present images to find different images in figure 2.7.

Next, the 3D-DCT is performed to compress the data of different images using 3D-DCT equation [1][2]. DCT has the character of gathering energy. The transform transfers spatial domain to frequency domain. In this paper, we adopt 3D-DCT because 3D-DCT is more concentrative than 2D-DCT in energy distribution in figure 2.8. Figure 2.8(a) describes the energy distribution of 3D-DCT coefficients and figure 2.8(b) present the energy distribution of 2D-DCT coefficients. According to the distribution of 3D-DCT and 2D-DCT, we could make sure 3D-DCT is more concentrative than 2D-DCT. Therefore, we use 3D-DCT technology to scatter needless data in order to decrease large number of data. Formula 2.3 describes the 3D-DCT equation [1][2][6]. X presents spatial domain signal of video or image [1][2]. Y expresses frequency domain signal [1][2]. Size of X and Y is $M \times N \times T$ [1][2]. The range of 3D-DCT from 0 to $T-1$, $M-1$ and $N-1$ for t , m and n [1][2]. We assume r , p , and q describe the index of frequency domain and a_r , a_p , and a_q present the 3D-DCT coefficients [1][2].

$$Y_{rpq} = a_r a_p a_q \sum_{t=0}^{T-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{tmm} \cos \frac{\pi(2t+1)r}{2T} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N} \quad (2.3)$$

$$a_r = \begin{cases} \sqrt{\frac{1}{T}}, & r=0 \\ \sqrt{\frac{2}{T}}, & \text{else} \end{cases} \quad a_p = \begin{cases} \sqrt{\frac{1}{M}}, & p=0 \\ \sqrt{\frac{2}{M}}, & \text{else} \end{cases} \quad a_q = \begin{cases} \sqrt{\frac{1}{N}}, & q=0 \\ \sqrt{\frac{2}{N}}, & \text{else} \end{cases}$$

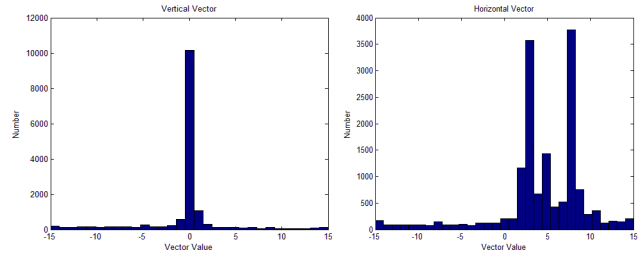


Figure 2.5: Statistic Images: (a) vertical change; (b) horizontal change



Figure 2.6: Compensated Image between 5th view image and disparity image: (a) reference image; (b) compensated image



Figure 2.7: Different image between compensated image and present image

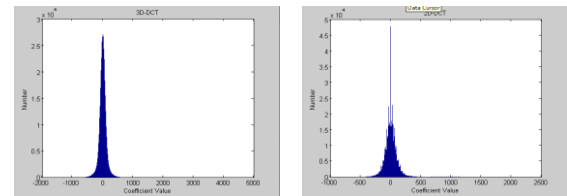


Figure 2.8: Comparison between 3D-DCT and 2D-DCT: (a) 3D-DCT; (b) 2D-DCT

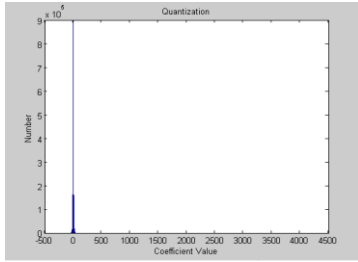


Figure 2.9: Quantization after 3D-DCT

After 3D-DCT processing, quantization technique is adopted to quantize the images to remove unnecessary data. The equation of quantization is defined as

$$C_{Qr,p,q} = \text{round}\left(\frac{C_{r,p,q}}{Q_{r,p,q}}\right) \quad (2.4)$$

Formula 2.4 expresses quantization function [6]. $C_{Qr,p,q}$ expresses quantization coefficient, r, p and q are coordinates, C represents the coefficient after DCT, Q is quantization table [6]. The coefficient value around zero point will be eliminated by quantization table [6]. Quantization is affected by the quality of DCT and it is illustrated in figure 2.9 [6].

3 3D-DCT Chip Design

3D-DCT is an important technology for 3D video compression. We focus on the topic to perform a chip implementation. We adopt a set of three 1D-DCT circuits to take the place of the single 3D-DCT circuit. The integration of circuits for this chip becomes more simplified than that of the previous chip.

The chip adopts 8×8 block with eight-bit input port and fourteen-bit output port. Before improvement, the chip which set to process sequentially causes data cannot connect tightly. After improvement, this chip, which adopts parallel process, cause data can output successively, this method can make output more efficiently and it can decrease a great quantity of components in this chip. There are other things to note, as the chip is consisted of three 1D-DCT, every 1D-DCT will handle the different dimension of data, as a result, how to choose memory addresses to store are extremely important [7][8].

We should inquire to the mode of accessing memory. At first, data are entered in the chip from (X_0, Y_0, Z_0) to (X_7, Y_0, Z_0) . Afterward, the next data will be entered from (X_0, Y_1, Z_0) to $(X_7, Y_1,$

$Z_0)$. Next, the circuit progresses by regular data as the above systematic way in Figure 3.1.

Next, we should describe the addresses of SRAM. At first, 1D-DCT (Y-dir) are handled by data that have passed in 1D-DCT (X-dir). Then, each SRAM are stored according to the order of data after 1D-DCT (Y-dir). For example, X_0 are stored in SRAM1 in figure 3.2. When all of SRAM have stored up already, data in every SRAM are read by CTRL signal and will be calculated by 1D-DCT (Z-dir).

For instance, addresses number of 0, 8, 16, 24, 32, 40, 48, 56 in SRAM1 and addresses number of 1, 9, 17, 25, 33, 41, 49, 57 will be read for 1D-DCT (Z-dir). Then, the circuit progresses by regular data as the above systematic way. Memory accessing steps and control flow are the most important procedure in the chip.

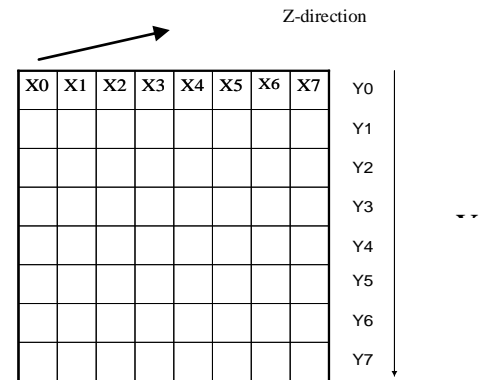


Figure 3.1: Data processing

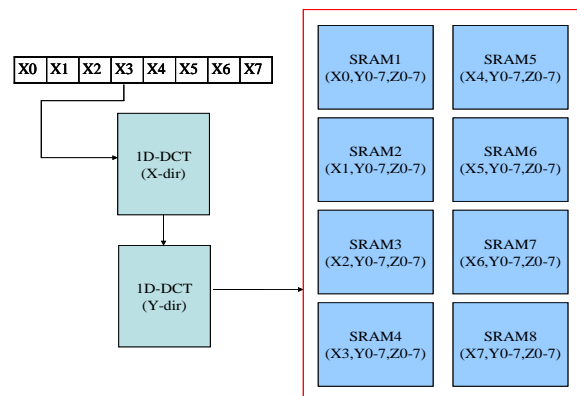


Figure 3.2: Memory accessing of the proposed architecture

4 Experimental Results

This chip, which contains 144 pins in total, has eight input ports with eight bits, eight output ports with fourteen bits in figure 4.1. The chip has to pass a succession of cell-based design flow because the chip is digital circuit. We use Verilog code to

design and the basis of the chip is the processing of TSMC 0.18um 1P6M Design kit v3.1 [7][8].

The chip have passed a series of processing, the usage rate about this chip carry out with success in ninety five percent and come to pass the resource of this chip effectively, the layout of this chip is in figure 4.2.

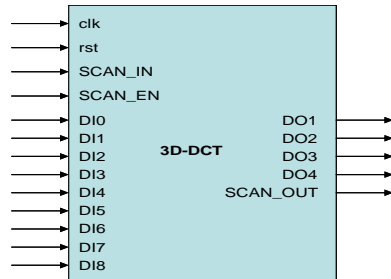


Figure 4.1: The chip of the proposed architecture

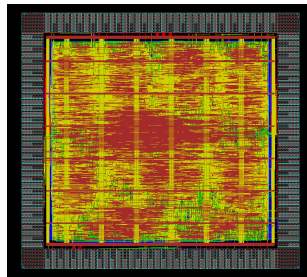


Figure 4.2: The layout of the proposed chip

5 Chip Comparisons

The comparison between the previous work and the proposed method is shown in Table 1. The main contribution of the proposed method is that we decrease the power consumption and quantity of components and raise a great quantity of speed in the chip.

Because we want to compare two methods, we adopt an 1152×864 size image as a standard test image. According to experimental results, the scheme speeds up 47% of the chip. Besides, the presented method decreases 14.4% logic gate count of this chip and the power consumption decrease 16% at the same time. The results prove the quality of the proposed chip.

6 Conclusions

In this paper, we proposed a memory-reduction-architecture of 3D-DCT for multi-view video coding. At first, we describe a 3D stereoscopic multi-view video compression completely. Then we

compare the difference between 2D-DCT and 3D-DCT. Afterward, we propound the method which is parallel structure that can make the chip more efficiently. What we wish to put a forward outline of this system in here regarding the 3D-DCT chip, having the three important and effective solutions that must be gotten through as the following. The proposed chip consists of a set of three 1D-DCT circuits, takes the place of the single 3D-DCT circuit. Besides, each 1D-DCT circuits should be handled the data, composed of different dimensions. The most important contribution of this paper is the memory reduction architecture. We propose an efficient architecture to solve the problem of 3D-DCT memory access and provide a well solution in this paper.

Table 5.1: Chip Comparison

Comparison	You's Chip [2]	Proposed Chip
Design Strategy	Sequential	Parallel
Package	128 CQFP	144 CQFP
Process	TSMC 0.18um	TSMC 0.18 um
Frequency	90MHz	58MHz
Power Consumption	219.85mW	182.75mW
Chip Size	2.7mmx2.7mm	2.75mmx2.75mm
Core Size	2.17mmx2.17mm	2.007mmx2.007mm
Circuit Run Time	218723328(ns)	115800192(ns)

Acknowledgements

This work was supported by the Taiwan E-learning and Digital Archives Programs (TELDAP) sponsored by the National Science Council of Taiwan under Grants NSC 100-2631-H-027-003-. The authors gratefully acknowledge the Chip Implementation Center (CIC), for supplying the technology models used in IC design.

References

- [1] Y. C. Fan, J. L. You and P. C. Chi, Cubic Memory Based 3D Discrete Cosine Transform for 3D Image Compression. *2010 International Asia-Pacific Data Storage Conference, APDSC 2010*, Hualien, Taiwan, Oct. 27~29, 2010.
- [2] J. L. You, *Spatial Domain Compression Technology of 3D Stereoscopic Multi-view Video*. Graduate Institute of Computer and Communication Engineering, National Taipei University of Technology, Taipei, 2010.
- [3] P. An, Q. Guo, T. Mi, L. Zhou, and Z. Zhang, Multi-view Video Coding Based on View Prediction. *International Conference on Audio, Language and Image Processing*, pp. 1481-1485, 2008.

- [4] N. P. Sgouros, S. S. Athineos, P. E. Mardaki, A. P. Sarantidou, M. S. Sangriotis, P. G. Papagerorgas, and N. G. Theogianous, Use of an Adaptive 3D-DCT Scheme for Coding Multi-view Stereo Images. *IEEE International Symposium on Signal Processing and Information Technology*, Dec. 2005, pp.180-185.
- [5] L. Ma, and F. Pan, Efficient Compression of Multi-view Video Using Hierarchical B Pictures. *International Conference on Multimedia and Ubiquitous Engineering*, Apr. 2008, pp. 118-121.
- [6] T. X. Chen, Z. C. Zhang, G. F. Huang, *Digital Image Processing*. Flag Publishing Co., Ltd, Taiwan, 2003, pp.1-50.
- [7] A. Sinha, A. Karmakar, K. Maiti, and P. Halder, A Reconfigurable Architecture for a Class of Digital Signal/Image Processing Applications, *IEEE Conference on Computers and Signal Processing*, Vol. 1, pp. 71-74, 2001.
- [8] S. Palnitkar, *Verilog HDL: A Guide to Digital Design and Synthesis*, Pearson Education, Inc, 2003.

(CCL), Industrial Technology Research Institute (ITRI). In 2006, Dr. Fan joined the Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan. Currently, he is an Associate Professor.



Shan-Ann Chen is currently a student in Electronic Engineering, National Taipei University of Technology. His research interests are in the areas of digital IC design, video compression.



Kuo-Gi Wu is currently a student in Electronic Engineering, National Taipei University of Technology. His research interests are in the areas of image processing and IC design.



Yu-Cheng Fan was born in Hsinchu, Taiwan, R.O.C., in 1975. He received his B. S. and M. S. degrees in Electrical Engineering from National Cheng Kung University in 1997 and 1999 respectively, his Ph.D. degree in Electrical Engineering from National Taiwan University in 2005. From 1999 to 2000, Yu-Cheng Fan was an IC design engineer at the Computer and Communications Research Laboratory



Jun-Lin You received the Master's degree in Computer and Communication Engineering from National Taipei University of Technology. His research interests are in areas of 3D stereoscopic multi-view video compression technology.