



Ain Shams Universty
Faculty Of Engineering
Electronics and Communications Engineering Department

Enhancing Inter Frame Prediction in Advanced Video Codecs

A thesis submitted in partial fulfillment of the requirement of the Degree of Doctor of Philosophy In Electronics & Communication Engineering

Submitted by

Nahed Ali Saleh Bahran

M.Sc Degree in Electronic and Communication,
Faculty of Engineering, Assiut University, Egypt, 2011

Supervised by

Prof. Dr. Abd El Halim Zekry

Electronics and Communications Engineering Department,
Faculty of Engineering, Ain Shams University

Prof. Dr. Ismael Mohamed Hafez

Electronics and Communications Engineering Department,
Faculty of Engineering, Ain Shams University

Cairo-Egypt
2017



**AIN SHAMS UNIVERSITY
FACULTY OF ENGINEERING
CAIRO-EGYPT**

Electronics & Communication Engineering Department

EXAMINERS COMMITTEE

Student: Nahed Ali Saleh Bahrn

Thesis: Enhancing Inter frame prediction in Advanced Video Codecs

Degree: Doctor of Philosophy in Electronics & Communication Engineering

Name, title and affiliation

Signature

1. Prof. El-sayed Mahmood El-rabaie

Electronics and Communications Engineering Department,
Faculty of Engineering, Menoufia University

.....

2. Prof. Wagdy Refa'at Anees

Electronics and Communications Engineering Department,
Faculty of Engineering, Ain Shams University

.....

3. Prof. Dr. Abd El Halim Abd Elnaby Zekry

Electronics and Communications Engineering Department,
Faculty of Engineering, Ain Shams University

.....

4. Prof. Dr. Ismael Mohamed Hafez

Electronics and Communications Engineering Department,
Faculty of Engineering, Ain Shams University

.....

STATEMENT

This dissertation is submitted to Ain Shams University for the degree of Ph.D. of Science in Electrical Engineering (Electronics and Communication Engineering).

The work included in this thesis was carried out by the author at the Electronics and Communication Engineering Department, Faculty of Engineering, Ain Shams University, Cairo, Egypt.

No part of this thesis was submitted for a degree or a qualification at any other university or institution.

Signature:

Name: Nahed Ali Saleh Bahran

Date: 24/ 12 /2017.

Researcher Data

Name	:	Nahed Ali Saleh Bahran
Date of birth	:	1982
Place of birth	:	Ibb- Yemen
Academic Degree	:	M. Sc Degree
Field of specialization	:	Electronic and Communication Engineering
University issued the degree	:	Ain Shams University
Date of issued degree	:	2011
Current job	:	Assisstant lecturer at Ibb University- Yemen

To the spirit of my father

To my beloved mother, brothers and sisters

To my wounded country "Yemen" which is suffered from a war

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor **Prof. Dr. Abd El Halim Zekry** for his constant support and assistance during the time of my Ph.D., I am greatly indebted to him for his great confidence and motivation he put in me since my first day in his work group. This work wouldn't have been possible without his guidance, involvement and encouragement, I have been benefited greatly from fruitful discussions with him which put me in the right way and helped me to overcome many difficulties during work.

Thanks also are extended to my supervisor **Prof. Dr. Ismael Mohamed Hafez**.

I want to give special thanks to **Dr. Rami Fathy, Dr. Mohammed Ebian** and **Dr. Ream Ibrahim**. I am really sincerely grateful to their guidance, advice and help rendered to me.

Additionally, I would like to thank the staff members of Electronic communication department for the great cooperation and support during my Ph.D work and as well as for all facilities they provided.

Finally, I would like to express my appreciation to the people who are the world to me, my mother, my brothers and sisters for their sacrifices, unceasing encouragement and support which helped me stay strong during these difficult years and inspired me to finish this work.

ABSTRACT

In real time applications such as video streaming and video conferencing it is important that the video encoding/decoding is fast. H.264/MPEG-4 Part 10 or AVC (advanced video coding) is currently one of the most widely used industry standards for video compression. It has been developed to achieve significant improvements, in the compression performance, over the existing standards.

In fact, the high compression performance is accomplished by the implementation of intra and inter prediction techniques that remove spatial and temporal redundancies. Most video coding standards depend heavily on the use of inter frame prediction technique to obtain high compression efficiency. Inter prediction technique is developed by motion estimation (ME) and motion compensation (MC) techniques.

It is known, that most of the complexity of the H.264/AVC compression standard lies in the H.264 encoder. Motion estimation (ME) introduces high computational complexity. It consumes about 70% of the time required for H.264 encoder.

Several studies have been provided to reduce the complexity of the H.264 encoder and the compression time without affecting the quality of the video. In other words, many ME algorithms has been designed and implemented to overcome the problem of complexity and time consumption of the inter prediction technique.

In this thesis, a MATLAB design and implementation of Enhanced Predictive Zonal Search algorithm (EPZS) ME algorithm for H.264 AVC is presented.

This thesis also introduces a comparative study on the most popular ME algorithms using MATLAB. It presents that the EPZS algorithm is superior on the most common ME algorithms. From studying the ME algorithms utilized in the JM reference software “18.6”, EPZS ME provides the best results in terms of speed, video quality and data rate but there is no consideration about how it can be effectively implemented in hardware.

This thesis proposes, a novel hardware design and implementation of EPZS ME algorithm for H.264 AVC is presented. An efficient implementation of inter prediction process using EPZS algorithm as the motion estimation technique is realized. The resulting motion vectors are compared with that of the exhaustive search ME algorithm to verify its efficiency. The proposed design of EPZS is implemented in “VHDL”, simulated using “ModelSim SE 6.5”, synthesized using “Xilinx ISE Design Suite 13.3” and verified using “SIMULINK”. The proposed architecture provides a high speed and low hardware complexity implementation of H.264 encoder permitting its use in real-time applications of H.264 /AVC standard.

Also this thesis introduces two enhancement methods on the EPZS algorithm. These enhancements provide an improvement on the speed of the EPZS performance by about 18.05% without affecting the quality of the output frame.

Key words: Prediction, Motion Estimation, EPZS, VHDL, Codec.

Table OF CONTENTS

LIST OF FIGURES	iv
LIST OF TABLES	viii
LIST OF ABBREVIATIONS	ix
CHAPTER 1	1
INTRODUCTION	1
1.1 Motivation.....	1
1.2 The Objective.....	3
1.4 The Methods and Tools	4
1.5 Outline of The Thesis.....	5
CHAPTER 2	6
H.264 VIDEO CODING STANDARD	6
2.1 H.264/AVC Standard.....	6
2.2 H.264/AVC Performance and Features	7
2.3 H. 264/AVC Applications.....	8
2.4 Profiles and Levels.....	9
2.4.1 H.264 Profiles.....	9
2.4.2 H.264 Levels	11
2.5 H. 264/AVC Stream Structure	12
2.5.1 Network Abstraction Layer (NAL)	12
2.5.1.1 NAL Units	13
2.5.1.2 Parameter Sets	13
2.5.1.3 Access Units	14
2.5.2 Video Coding Layer (VCL)	14
2.5.2.1 Macroblocks, Slices, and Slice Groups	14
2.6 Slice Types.....	15
2.7 Error Resilience Features	15
2.8 H.264 Prediction	17
2.8.1 Intra Prediction	18
2.8.2 Inter Prediction.....	19
2.9 H.264 Transform Coding.....	22
2.10 Quantization.....	24

2.11	In-Loop Deblocking Filter	24
2.12	Entropy Coding	26
2.13	H.264 Decoder	27
CHAPTER 3		28
MOTION ESTIMATION TECHNIQUES FOR INTER PREDICTION IN H264/AVC		28
3.1	Introduction.....	28
3.2	Matching Criteria	29
3.3	Motion Estimation Algorithms	30
3.3.1	Full Search Algorithm (FS).....	30
3.3.2	Three Step Search (TSS) or (3SS).....	31
3.3.3	Cross Search (CSA)	32
3.3.4	Diamond Search (DS)	33
3.3.5	Unsymmetrical Multi Hexagon Search Algorithm (UMHexagonS).....	35
3.3.6	Motion Vector Field Adaptive Fast Motion Estimation (MVFAST).....	38
3.3.7	Predictive Motion Vector Field Adaptive Search Technique (PMVFAST)	40
3.3.8	Advanced Predictive Diamond Zonal Search (APDZS)	42
3.3.9	Enhanced Predictive Zonal Search (EPZS).....	44
3.3.9.1	Predictor Selection.....	44
3.3.9.2	Adaptive Early Termination	45
3.3.9.3	Motion Vector Refinement.....	46
3.4	Motion Estimation in JM Reference Software.....	47
CHAPTER 4		49
IMPLEMENTATION OF INTER PREDICTION TECHNIQUE USING EPZS MOTION ESTIMATION ALGORITHM		49
4.1	Introduction.....	49
4.2	MATLAB Implementation	50
4.2.1	MATLAB Simulation Results.....	50
4.3	Hardware Design and Implementation of EPZS	53
4.3.1	Design Architecture.....	54
4.3.1.1	CF Memory Module	55
4.3.1.2	RF Memory Module	58
4.3.1.3	Address Control Unit.....	59
4.3.1.4	SAD Module.....	62
4.3.1.5	Previous Frame Minimum SAD Memory Module (PF_MSAD_RAM)	63

4.3.1.6	Previous Frame MV Memory Module (PF_MV_RAM).....	64
4.3.1.7	Motion Vector Memory Module (MV_RAM)	66
4.3.1.8	Minimum SAD Memory Module (Min_SAD_RAM).....	67
4.3.1.9	EPZS ME Module	68
4.3.1.10	MC Module.....	73
4.3.1.11	Residual Memory (Res_Memory) Module.....	75
4.3.2	Hardware Simulation Results.....	76
4.3.2.1	Initialization of CF and RF Memory Units	76
4.3.2.2	Experimenting the SAD Computation.....	78
4.3.2.3	Complete Inter Prediction Implementation	81
4.3.3	Synthesis Results.....	98
4.3.4	Enhancement of the Simulated and Synthesized Results.....	98
4.3.5	Downloading on FPGA.....	103
CHAPTER 5	110
CONCLUSION AND FUTURE WORK	110
5.1	Conclusion	110
5.2	Future Work	111
REFERENCES	112
APPENDIX A	116
APPENDIX B	117
APPENDIX C	130

LIST OF FIGURES

Figure 1-1 General block diagram of video codec.....	1
Figure 1-2 Typical H.264 encoder	2
Figure 2-1 Structure of an H.264/MPEG4-AVC video codec [8].	7
Figure 2-2 Baseline, constrained baseline, extended and main Profiles.	10
Figure 2-3 Main and high profiles.	11
Figure 2-4 Hierarchical structure of H.264 encoder	13
Figure 2-5 The Division of a picture into slices	14
Figure 2-6 Structure of the slice syntax in H.264/AVC video encoder	16
Figure 2-7 An example of segmentation of an image into slice groups	16
Figure 2-8 Block diagram of the H.264 encoder	17
Figure 2-9 4×4 intra prediction modes [3].	18
Figure 2-10 MB partition for motion compensated prediction	19
Figure 2-11 Multi-picture MCP. In addition to the MV, reference indexes (Δ) are transmitted. The concept is similarly extended for B slices [12].	20
Figure 2-12 Interpolation of half- sample positions of luma MBs	20
Figure 2-13 Quarter- sample of luma positions interpolation.....	21
Figure 2-14 interpolation of eight positions chroma samples.....	22
Figure 2-15 Transmission order of all the transformed coefficients of the MB.	23
Figure 2-16 Edge filtering order in a MB.	25
Figure 2-17 Principle of the deblocking filter.....	26
Figure 2-18 Block diagram of the H.264 decode.....	27
Figure 3-1 Block Matching Algorithm.	28
Figure 3-2 Full Search algorithm a) Raster scan & b) Spiral scan.	31
Figure 3-3 Three Step Search Algorithm.	32
Figure 3-4 Cross Search Algorithm.	33
Figure 3-5 Large and small diamond search patterns used in the DS algorithm	34
Figure 3-6 Example of the DS algorithm.....	34

Figure 3-7 Typical process of the UMhexagonS algorithm.....	35
Figure 3-8 The four adjacent blocks to the current block in the reference frame.....	36
Figure 3-9 Sixteen points hexagon pattern (16-HP) used in UMhexagonS [25].....	37
Figure 3-10 ROS of the current MB consists of MB ₁ , MB ₂ and MB ₃	38
Figure 3-11 Example of the MV distribution corresponding to set V.	39
Figure 3-12 The search patterns of MVFAST algorithm.....	40
Figure 3-13 Example of PMVFAST algorithm.	42
Figure 3-14 Example of APDZS implementation.	44
Figure 3-15 Spatial predictors for the EPZS algorithm in the current frame, (a) median MV of SetA, (b) the spatial predictors of set B.	45
Figure 3-16 Motion vectors of the collocated MB and its surrounding blocks in the previous frame	46
Figure 3-17 Search patterns used in EPZS, (a) small diamond pattern, (b) square pattern	47
Figure 4-1 The block diagram of the implemented inter prediction technique.	50
Figure 4-2 EPZS ME algorithm flow chart.....	51
Figure 4-3 Pictures of the current, the reference, and the motion compensated residual Frame..	52
Figure 4-4 Block diagram of the designed inter prediction.	54
Figure 4-5 Schematic block diagram of the proposed implementation.	55
Figure 4-6 Segmentation of one QCIF frame into 16×16 MBs in raster scan order.	56
Figure 4-7 Pin out of the current frame memory module.	56
Figure 4-8 Structure of the designed memory.	57
Figure 4-9 Pin out of the reference frame memory module.....	58
Figure 4-10 rr & cc counters with a segmentation of QCIF frame in raster scan order.	60
Figure 4-11 ACU Module pinout.....	60
Figure 4-12 ACU with CF and RF memory blocks connections.....	61
Figure 4-13 Pinout of the SAD module.	62
Figure 4-14 Apart of the overall proposed implementation including ACU, CF Memory, RF Memory and SAD modules.	63
Figure 4-15 Pinout of the previous frame MSAD memory module	64
Figure 4-16 Combination of the x& y components of the MV in one memory location.	65
Figure 4-17 Pinout of the previous frame MV memory module.	65

Figure 4-18 Pinout of the current frame MV memory module.....	66
Figure 4-19 Pinout of the current frame MSAD memory module.....	67
Figure 4-20 16×16 pixels MB, the yellow pixel is the top-left pixel of the MB.	68
Figure 4-21 Pinout of the EPZS ME Module.	69
Figure 4-22 Schematic diagram of a subsystem including the connections between ACU, SAD and EPZS ME modules.	72
Figure 4-23 Connections between EPZS ME Module with the MSAD and MV pf the current and previous frames.	73
Figure 4-24 Pin out of the MC module.	73
Figure 4-25 Operations between MC unit with ACU, CF memory, RF memory, and EPZS ME units.	74
Figure 4-26 Pin out of the Residual memory (Res_Memory) Module.	75
Figure 4-27 Segmentation of a video source into a sequence of consecutive frames.....	77
Figure 4-28 Converting the CF and RF from RGB to YUV format.	77
Figure 4-29 SAD computation operations.	78
Figure 4-30 Example of SAD computations between the 1 st MBs of CF206 and RF204 using proposed implementation in VHDL.....	79
Figure 4-31 SAD computation of the 1 st MBs of CF206 and RF204 using “Excel”.....	79
Figure 4-32 SAD computation example between CB (3,6) in CF205 and RB (80,80) in RF204 using proposed implementation in VHDL.	80
Figure 4-33 SAD computation example between CB (3,6) in CF205 and RB (80,80) in RF204 using Excel.....	80
Figure 4-34 Complete inter prediction implementation including ME and MC of the proposed Implementation.	82
Figure 4-35 ME simulation results for MB number (70).....	83
Figure 4-36 ME simulation results of MB number (98) for EPZS ²	86
Figure 4-37 Pictures of the current, the reference, and the motion compensated residual Frame.89	
Figure 4-38 Block diagram of the designed model of FS ME using “Simulink”.	90
Figure 4-39 Parameter settings of the template matching block.....	91
Figure 4-40 SA specifications of the FS ME using Template Matching block in SIMULINK. .	92
Figure 4-41 i& j counters of the MBs in a QCIF frame.....	93

Figure 4-42 Deviation between the resultant MVs of EPZS and FS.	96
Figure 4-43 Deviation between the resultant MVs of TSS and FS.....	97
Figure 4-44 Deviation between the resultant MVs of EPZS ² and FS.....	97
Figure 4-45 The relation between the current frame and the previous frame MVs.....	99
Figure 4-46 ME simulation results for MB no. (98) and the total no. of clock cycles of the overall ME&MC processes.	100
Figure 4-47 Deviation between the resultant MVs of improved EPZS and FS.....	101
Figure 4-48 ME simulation results for MB no. (98) and the total no. of clock cycles of the overall ME&MC processes.	102
Figure 4-49 Deviation between the resultant MVs of the 2 nd improvement on EPZS and FS ...	103
Figure 4-50 A board photo of vertex-6 XC6VLX240T FPGA with its features.	104
Figure 4-51 RTL schematic of the implemented motion estimation technique.	104
Figure 4-52 RTL schematic of the new structure used to present the ME results on FPGA.....	105
Figure 4-53 The place and route to the FPGA is performed successfully.	106
Figure 4-54 Programming the bit file to the FPGA is carried out successfully.....	107
Figure 4-55 Motion estimation results for the 1 st MB.	108
Figure 4-56 Output MinSAD of MBno=0, (a) shows the LSB bits of the MinSAD, (b) shows the MSB bits of the output MinSAD.	109