AN EFFICIENT ENERGY AWARE ADAPTIVE SYSTEM-ON-CHIP ARCHITECTURE FOR REAL-TIME VIDEO ANALYTICS

HISHAM AHMED ALI AHMED

UNIVERSITI SAINS MALAYSIA

2016

AN EFFICIENT ENERGY AWARE ADAPTIVE SYSTEM-ON-CHIP ARCHITECTURE FOR REAL-TIME VIDEO ANALYTICS

by

HISHAM AHMED ALI AHMED

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

October 2016

ACKNOWLEDGEMENTS

First and above all, I thank the Almighty Allah for his endless blessings.

This thesis would not have been possible without the support of many people over the years of my study. First of all, I would like to present my cordial gratitude to my supervisor Prof. Dr. Othman Sidek for his guidance and support. He always able to push ideas one level further, and shares his rich experience about the research process. His contagious enthusiasm, technical passion, work ethics, and qualities have greatly inspired me.

I am also grateful to the Universiti Sains Malaysia and Prof. Dr. Othman Sidek for supporting my research work through the Research Creativity and Management Office (RCMO) grant (1001/PCEDEC/854003) and the USM Fellowship.

Last but not the least, I would like to thank my parents, Ahmed and Suad, and my wife Arwa, and my two young daughters, Ragad and Razan, for their sincere prayers, unconditional love, and never-ending support over the past years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACP	Accelerator Coherency Port
ADAS	Advanced Driver Assistance System
AES	Advanced Encryption Standard
AFAS	Automated Fingerprint-based Authentication System
AMBA	ARM Advanced Microcontroller Bus Architecture
AMP	Asymmetrical Multiprocessing
ANPR	Automatic Number Plate Recognition
APSoC	All Programmable System-on-Chip
ASIC	Application Specific Integrated Circuit
ASIP	Application-Specific Instruction Set Processor
ASSP	Application-specific standard parts
ATLAS	Automatically Tuned Linear Algebra Software
AVS	Adaptive Voltage Scaling
AXI	Advanced eXtensible Interface
BFM	Bus Functional Model
BLAS	Basic Linear Algebra Subprograms
BRAM	Block RAM

BSCAN	Boundary-Scan Interface
BUFG	Global Clock Buffer
BUFR	Regional Clock Buffer
CCTV	Closed-Circuit Television
CLB	Configurable Logic Block
CNN	Convolutional Neural Network
CPLD	Complex Programmable Logic Device
CPU	Central Processing Unit
CPS	Cyber-Physical Systems
CRC	Cyclic Redundancy Check
DCP	Design Checkpoint
DDR3	Double Data Rate Type Three
DMA	Direct Memory Access
DPR	Dynamic Partial Reconfiguration
DRAM	Dynamic Random Access Memory
DSE	Design Space Exploration
DSP	Digital Signal Processor
DVFS	Dynamic Voltage and Frequency Scaling
ECC	Error Checking and Correction

EMIO	Extended Multiplexed Input/Output
FF	Flip Flop
FFT	Fast Fourier Transform
Flops	Floating-point operations
FMC	FPGA Mezzanine Card
FPGA	Field Programmable Gate Array
FPS	Frame Per Second
GP	General Purpose
GPU	Graphic Processing Unit
GUI	Graphic User Interface
HCD	Harris Corners Detection
HD	High-Definition
HDMI	High-Definition Multimedia Interface
HLS	High Level Synthesis
HMAC	keyed-Hash Message Authentication Code
HP	High Performance Port
HPC	High Performance Computing
HW	Hardware
ICAP	Internal Configuration Access Port

IDCT	Inverse Discrete Cosine Transform
IOB	Input/Output Block
IoT	Internet of Things
IP	Intellectual Property
ISERDES	Input Serializers/Deserializers
JPEG	Joint Photographic Experts Group
JTAG	Joint Test Action Group
LUT	Look Up Table
LZW	Lempel–Ziv–Welch
MCAP	Media Configuration Access Port
MGT	Multi-Gigabit Transceiver
MMCM	Mixed-Mode Clock Manager
MPSoC	Multiprocessor System-on-Chip
NMS	Non-Maximum Suppression
NRE	Non-Recurring Engineering
OI	Operational Intensity
OS	Operating System
OSERDES	Output Serializers/Deserializers
PCAP	Processor Configuration Access Port

PCIe	Peripheral Component Interconnect Express
PL	Programmable Logic
PLL	Phase Locked Loop
PMBus	Power Management Bus
PMM	Power Mode Management
PS	Processing System
QoS	Quality of Service
RISC	Reduced Instruction Set Computing
RM	Reconfigurable Modules
RP	Reconfigurable Partition
RTL	Register Transfer Language
SD	Secure Digital
SGEMM	Single-precision GEneral Matrix-matrix Multiplication
SIMD	Single Instruction Multiple Data
SIR	Sampling Importance Resampling
SLAM	Simultaneous Localization And Mapping
SMP	Symmetrical Multiprocessing
SoC	System-on-Chip
SUSAN	Smallest Univalue Segment Assimilating Nucleus

SW	Software
Tcl	Tool Command Language
UART	Universal Asynchronous Receiver/Transmitter
UAV	Unmanned Aerial Vehicle
UHD	Ultra High Definition
VCA	Video Content Analysis
VGA	Video Graphics Array
VHDL	Very High Speed Integrated Circuit Hardware Description Language
VLIW	Very Long Instruction Word
WMSN	Wireless Multimedia Sensor Network
XADC	Xilinx Analogue-to-Digital Converter
XDC	Xilinx Design Constraint
XSDK	Xilinx Software Development Kit

SENI BINA BOLEH SUAI BAGI SISTEM-ATAS-CIP PEKA-TENAGA CEKAP BAGI ANALITIK VIDEO MASA-NYATA

ABSTRAK

Aplikasi analitik video yang kebanyakannya ada pada peranti terbenam semakin kerap digunakan kini. Pertumbuhan pesat yang ditunjukkan ini menyebabkan perlunya Sistem-atas-Cip (SoC) dibangunkan untuk menjalankan pemprosesan terbaik pada cip tunggal berbanding pada komponen diskret. Penglihatan terbenam tertakluk kepada keperluan yang ketat, iaitu prestasi masa-nyata, tenaga yang terhad, dan kemudahsuaian untuk mendepani evolusi piawaian. Tambahan pula, untuk mereka bentuk SoC yang sedemikian kompleks, khususnya SoC Boleh Atur Cara Semua Zynq, pendekatan reka bentuk yang selari untuk perkakasan/perisian tradisional bergantung pada pemprofilan perisian untuk menjalankan pemetakan perkakasan/perisian tidak mampu lagi menjalankan tugas ini kerana pemprofilannya tidak dapat meramal prestasi aplikasi pada perkakasan. Oleh itu, satu model yang menghubungkan ciri-ciri kepada prestasi platform adalah sangat penting untuk dibangunkan. Untuk menghantar prestasi masa-nyata bagi resolusi video yang pantas berkembang sambil menjaga kelenturan seni binanya pada pemproses, Unit Pemprosesan Grafik, Pemproses Signal Digital, dan Litar Bersepadu Aplikasi-Spesifik, ia tidak dapat dibuat. Selanjutnya, dengan penskalaan teknologi semikonduktor, dijangka bahawa pelesapan kuasa akan meningkat kerana kapasiti bateri dijangka tidak akan meningkat dengan mendadak. Model prestasi bagi Zynq dibangunkan dengan menggunakan kaedah analitis dan digunakan dalam reka bentuk selari bagi perkakasan/perisian adalah untuk membantu pemetaan algoritma bagi perkakasan. Selepas itu, SoC bagi analitik video masa-nyata direalisasikan pada Zynq

dengan menggunakan algoritma pengesanan sudut Harris. Analisis yang teliti terhadap algoritma tersebut dan penggunaan yang cekap pada sumber Zynq menghasilkan seni bina yang bukan sahaja terselari dan tertalipaipan, malah melebihi prestasi algoritma yang paling terkini. Dengan menjalankannya pada SoC peka-kuasa yang boleh ubah serta dibangunkan dengan menggunakan pengkonfigurasian semula separuh dinamik, aplikasi penjadual konfigurasi yang peka-konteks akan mengikut konteks operasi dan menukarkan resolusi video dengan penggunaan kuasa bagi menampung masa operasi yang lama ketika menghantar prestasi masa-nyata. Pengesanan sudut masa-nyata pada 79.8, 176.9, dan 504.2 bingkai sesaat tercapai, iaitu masing-masing bagi HD1080, HD720, dan VGA. Ketiga-tiga bingkai sesaat berjaya mengatasi prestasi kajian terdahulu dengan gandaan 31 kali lebih baik bagi HD720 dan 3.5 kali bagi VGA. Penjadual berfungsi pada ketika proses konfigurasi berjalan. Pada ketika itu, perkakasan yang sesuai digunakan di mana ia dapat memenuhi konteks operasi dan halangan yang didefinisikan oleh pengguna; dalam kalangan pemecut yang dibangunkan sebagai contoh piawaian video HD1080, HD720, dan VGA menggunakan tenaga yang rendah. Kaedah penyesuaian diri berjaya mencatatkan tempoh masa operasi yang lebih panjang berbanding dengan teras parameter IP untuk kadar kapasiti bagi bateri yang sama iaitu sebanyak 1.77 kali. Di samping itu, lebihan pengkonfigurasian tenaga boleh diabaikan bagi kaedah ini. Kesan pada kelewatan masa bagi pengkonfigurasian masa separuh diperhatikan, contohnya, hanya dua bingkai video diturunkan bagi HD1080p60 ketika masa pengkonfigurasian semula. Pemudahan proses reka bentuk dengan model analisis, dan penggunaan sumber Zynq serta keputusan adaptivity diri dalam peka-tenaga SoC dengan cekap, ini menyediakan prestasi masa-nyata untuk video analitik.