

Bruno Zatt · Muhammad Shafique
Sergio Bampi · Jörg Henkel

3D Video Coding for Embedded Devices

Energy Efficient Algorithms and
Architectures

 Springer

3D Video Coding for Embedded Devices

Bruno Zatt • Muhammad Shafique
Sergio Bampi • Jörg Henkel

3D Video Coding for Embedded Devices

Energy Efficient Algorithms
and Architectures

 Springer

Bruno Zatt
Department of Computer Science
Karlsruhe Institute of Technology
Karlsruhe, Germany

Institute of Informatics
Federal University of Rio
Grande do Sul (UFRGS)
Porto Alegre, Brazil

Sergio Bampi
Institute of Informatics
Federal University of Rio
Grande do Sul (UFRGS)
Porto Alegre, Brazil

Muhammad Shafique
Karlsruhe Institute of Technology
Karlsruhe, Germany

Jörg Henkel
Department of Computer Science
Karlsruhe Institute of Technology
Karlsruhe, Germany

ISBN 978-1-4614-6758-8 ISBN 978-1-4614-6759-5 (eBook)
DOI 10.1007/978-1-4614-6759-5
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013934705

© Springer Science+Business Media New York 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Acknowledgements

This work was partly supported by the German Research Foundation (DFG) as part of the Transregional Collaborative Research Centre “Invasive Computing” (SFB/TR 89); <http://invasic.de>.

This work was also partly funded by the Brazilian Coordination for the Improvement of Higher Level Personnel (CAPES—“*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*”) and the Graduate Program on Microelectronics (PGMICRO), Institute of Informatics, Federal University of Rio Grande do Sul (UFRGS).

Contents

1	Introduction	1
1.1	3D-Video Applications.....	2
1.2	Requirements and Trends of 3D Multimedia	3
1.3	Overview on Multimedia Embedded Systems	5
1.4	Issues and Challenges.....	6
1.5	Monograph Contribution.....	7
1.5.1	3D-Neighborhood Correlation Analysis	7
1.5.2	Energy-Efficient MVC Algorithms	8
1.5.3	Energy-Efficient Hardware Architectures	9
1.6	Monograph Outline	9
2	Background and Related Works	11
2.1	2D/3D Digital Videos.....	11
2.2	Multiview Correlation Domains.....	14
2.2.1	Spatial Domain Correlation.....	14
2.2.2	Temporal Domain Correlation.....	15
2.2.3	Disparity Domain Correlation	16
2.3	Multiview Video Coding	16
2.3.1	MVC Encoding Process	18
2.3.2	Motion and Disparity Estimation	22
2.3.3	MVC Mode Decision	27
2.3.4	MVC Rate Control	28
2.4	3D-Video Systems.....	29
2.5	Multimedia Architectures Overview	30
2.5.1	Multimedia Processors/DSPs	30
2.5.2	Reconfigurable Processors for Video Processing.....	31
2.5.3	Application-Specific Integrated Circuits	32
2.5.4	Heterogeneous Multicore SoCs.....	33

2.6	Energy-Efficient Architectures for Multimedia Processing	33
2.6.1	Video Memories	34
2.6.2	SRAM Dynamic Voltage-Scaling Infrastructure.....	34
2.6.3	Dynamic Power Management for Memories.....	35
2.6.4	Energy Management for Multimedia Systems	36
2.6.5	Energy-Efficient Video Architectures	37
2.7	Energy/Power Consumption Background	38
2.8	Energy-Efficient Algorithms for Multiview Video Coding.....	40
2.8.1	Energy-Efficient Mode Decision	40
2.8.2	Energy-Efficient Motion and Disparity Estimation.....	42
2.9	Video Quality on Energy-Efficient Multiview Video Coding.....	45
2.9.1	Control Techniques Background	46
2.10	Summary of Background and Related Works	50
3	Multiview Video Coding Analysis for Energy and Quality.....	53
3.1	Energy Requirements for Multiview Video Coding.....	53
3.1.1	MVC Computational Effort.....	57
3.1.2	MVC Memory Access.....	59
3.1.3	Adaptivity in MVC Video Encoder	60
3.2	Energy-Related Challenges in Multiview Video Coding	62
3.3	Objective Quality Analysis for Multiview Video Coding	63
3.4	Quality-Related Challenges in Multiview Video Coding.....	65
3.5	Overview of Proposed Energy-Efficient Algorithms and Architectures for Multiview Video Coding	66
3.5.1	3D-Neighborhood.....	67
3.5.2	Energy-Efficient Algorithms	68
3.5.3	Energy-Efficient Architectures.....	69
3.6	Summary of Application Analysis for Energy and Quality	71
4	Energy-Efficient Algorithms for Multiview Video Coding.....	73
4.1	3D-Neighborhood Correlation Analysis	74
4.1.1	Coding Mode Correlation Analysis.....	74
4.1.2	Motion Correlation Analysis.....	82
4.1.3	Bitrate Correlation Analysis.....	84
4.2	Thresholds	87
4.3	Multilevel Mode Decision-based Complexity Adaptation	90
4.3.1	Multilevel Fast Mode Decision	90
4.3.2	Energy-Aware Complexity Adaptation	95
4.3.3	Multilevel Fast Mode Results.....	100
4.3.4	Energy-Aware Complexity Adaptation Results	105
4.4	Fast Motion and Disparity Estimation.....	107
4.4.1	Fast Motion and Disparity Estimation Algorithm.....	107
4.4.2	Fast ME/DE Algorithm Results	109

- 4.5 Video-Quality Management for Energy-Efficient Algorithms..... 111
 - 4.5.1 Hierarchical Rate Control for MVC..... 111
 - 4.5.2 Frame-Level Rate Control..... 113
 - 4.5.3 Basic Unit-Level Rate Control..... 119
 - 4.5.4 Hierarchical Rate Control Results..... 121
- 4.6 Summary of Energy-Efficient Algorithms for Multiview Video Coding 126
- 5 Energy-Efficient Architectures for Multiview Video Coding..... 127**
 - 5.1 Motion and Disparity Estimation Hardware Architecture 127
 - 5.1.1 SAD Calculator 130
 - 5.1.2 Programmable Search Control Unit 131
 - 5.1.3 On-Chip Video Memory..... 133
 - 5.1.4 Address Generation Unit..... 134
 - 5.2 Parallelism in the MVC Encoder and ME/DE Scheduling 136
 - 5.2.1 Parallelism in the MVC Encoder..... 136
 - 5.2.2 ME/DE Hardware Architecture Pipeline Scheduling..... 137
 - 5.3 Dynamic Search Window Formation 140
 - 5.3.1 ME/DE Memory Access Pattern Analysis 140
 - 5.3.2 Search Map Prediction 142
 - 5.3.3 Dynamic Search Window Formation 143
 - 5.4 On-Chip Video Memory..... 145
 - 5.4.1 On-Chip Memory Design..... 145
 - 5.4.2 Application-Aware Power Gating 146
 - 5.5 Hardware Architecture Evaluation 148
 - 5.5.1 Dynamic Window Formation Accuracy..... 148
 - 5.5.2 Hardware Architecture Evaluation..... 148
 - 5.6 Summary of Energy-Efficient Algorithms for Multiview Video Coding 150
- 6 Results and Comparison..... 151**
 - 6.1 Experimental Setup 151
 - 6.1.1 Software Simulation Environment 151
 - 6.1.2 Benchmark Video Sequences 152
 - 6.1.3 Fairness of Comparison..... 155
 - 6.1.4 Hardware Description and ASIC Synthesis 155
 - 6.2 Comparison with the State of the Art..... 156
 - 6.2.1 Energy-Efficient Algorithms 156
 - 6.2.2 Video Quality Control Algorithms..... 161
 - 6.2.3 Energy-Efficient Hardware Architectures..... 163
 - 6.3 Summary of Results and Comparison..... 166

7 Conclusion and Future Works	169
7.1 Future Works	171
7.1.1 Remaining MVC Challenges	172
7.1.2 3D-Video Pre- and Post-processing	172
7.1.3 Next-Generation 3D-Video Coding.....	172
Appendix A: JMVC Simulation Environment	175
A.1 JMVC Encoder Overview	175
A.2 Modifications to the JMVC Encoder	178
A.2.1 JMVC Encoder Tracing	178
A.2.2 Communication Channels in JMVC	178
A.2.3 Mode Decision Modification in JMVC.....	179
A.2.4 ME/DE Modification in JMVC.....	179
A.2.5 Rate Control Modification in JMVC.....	179
Appendix B: Memory Access Analyzer Tool	181
B.1 Current Macroblock-Based Analysis	182
B.2 Search Window-Based Analysis	182
Appendix C: CES Video Analyzer Tool	185
References.....	189
Index.....	199

List of Figures

Fig. 1.1	Video scaling trend	4
Fig. 1.2	(a) Mobile systems performance trend and (b) Li-ion battery capacity growth.....	6
Fig. 2.1	Macroblocks and slices organization.....	12
Fig. 2.2	Multiview video sequence	13
Fig. 2.3	Multiview video capture, (de)coding, transmission, and display system.....	13
Fig. 2.4	Neighborhood correlation example	15
Fig. 2.5	Prediction comparison between simulcast and MVC.....	17
Fig. 2.6	MVC encoder block diagram	18
Fig. 2.7	MVC prediction structure example	19
Fig. 2.8	Nine prediction directions for intra-prediction 4×4	20
Fig. 2.9	Four prediction directions for intra-prediction 16×16	20
Fig. 2.10	Block processing order in the transform module	21
Fig. 2.11	Zigzag scan order for a 4×4 block.....	21
Fig. 2.12	Order of macroblock borders filtering	22
Fig. 2.13	Temporal and disparity similarities	23
Fig. 2.14	Motion and disparity estimation.....	24
Fig. 2.15	MVC rate control actuation levels.....	28
Fig. 2.16	SRAM voltage-scaling infrastructure.....	35
Fig. 2.17	Energy/power dissipation sources	39
Fig. 2.18	Fast mode decision example.....	41
Fig. 2.19	ME/DE search conceptual example.....	43
Fig. 2.20	Model predictive control (MPC) conceptual behavior	47
Fig. 2.21	Markov decision process (MDP).....	48
Fig. 2.22	Variance-based region of interest map (Flamenco2)	50
Fig. 3.1	MVC energy consumption and battery life	54
Fig. 3.2	MVC component blocks energy breakdown	55
Fig. 3.3	MVC energy breakdown for multiple search window sizes.....	55

Fig. 3.4	MVC energy for distinct mode decision schemes	56
Fig. 3.5	ME/DE energy breakdown	56
Fig. 3.6	MVC vs. Simulcast complexity	57
Fig. 3.7	MVC computation breakdown	58
Fig. 3.8	Memory bandwidth for 4-views MVC encoding	59
Fig. 3.9	Frame-level energy consumption for MVC	61
Fig. 3.10	Memory requirements for motion estimation at MB level	61
Fig. 3.11	Objective video quality in relation to coding modes	64
Fig. 3.12	Energy-efficient Multiview Video Coding overview	66
Fig. 4.1	Coding mode distribution	74
Fig. 4.2	Visual analysis of the coding mode correlation	75
Fig. 4.3	Coding mode hits in the 3D-neighborhood	77
Fig. 4.4	Variance PDF for different coding modes	78
Fig. 4.5	(a) PDF for RDCost difference (between the current and the neighboring MBs) for SKIP <i>hit</i> and <i>miss</i> ; (b, c) Surface plots of RDCost difference for the SKIP coding mode hit and miss; (d) RDCost prediction error for spatial neighbors	79
Fig. 4.6	PDF of RDCost for different prediction modes in Ballroom sequence	80
Fig. 4.7	Average RDCost prediction error for spatial neighbors in Vassar Sequence	81
Fig. 4.8	MVC prediction structure and 3D-neighborhood details	82
Fig. 4.9	MV/DV error distribution between predictors and optimal vector (Ballroom, Vassar)	83
Fig. 4.10	View-level bitrate distribution (Flamenco2, QP=32)	85
Fig. 4.11	Frame-level bitrate distribution for two GGOPs (Flamenco2, QP=32)	86
Fig. 4.12	Basic unit-level bitrate distribution (Flamenco2, QP=32)	86
Fig. 4.13	PDF showing the area of high probability as the shaded region	87
Fig. 4.14	PDF of RDCost for SKIP MBs	88
Fig. 4.15	Threshold curves for RDCost	89
Fig. 4.16	PDF of variance for different prediction modes	89
Fig. 4.17	Overview of the multilevel fast mode decision	91
Fig. 4.18	Early SKIP threshold curves for (a) RDCost and (b) Variance	92
Fig. 4.19	Evaluation of thresholds for early termination (Ballroom, QP=32)	93
Fig. 4.20	Early termination threshold plots for Relax (<i>blue</i>) and Aggressive (<i>red</i>) complexity reduction	94
Fig. 4.21	MVC coding structure for asymmetric coding	96
Fig. 4.22	Energy-aware MVC complexity adaptation scheme	97
Fig. 4.23	Pseudo-code of mode decision for different QCCs	98

Fig. 4.24 Probability density function for RDCost..... 98

Fig. 4.25 Run-time complexity adaptation state machine 99

Fig. 4.26 Average tested modes (QP= {22,27,32,37,42}, GOP=8, Views=8) 102

Fig. 4.27 View-level time savings and Δ PSNR comparison of Relax and Aggressive levels (Exit sequence, QP=32) 102

Fig. 4.28 Average tested modes for all sequences 102

Fig. 4.29 Detailed number of evaluated modes for (a) *Relax* and (b) *Aggressive* (Exit Sequence) 103

Fig. 4.30 Frame-wise PSNR loss comparison of Relax and Aggressive levels (Exit, QP=32)..... 104

Fig. 4.31 Frame-wise time saving comparison of Relax and Aggressive levels (Exit, QP=32)..... 104

Fig. 4.32 Overhead of our scheme 105

Fig. 4.33 Flow diagram of the adaptive fast ME/DE 108

Fig. 4.34 Rate-distortion comparison with full search 110

Fig. 4.35 View-level execution time savings compared to TZ Search..... 110

Fig. 4.36 Comparison of the number of SAD operation..... 111

Fig. 4.37 Hierarchical rate control system diagram..... 112

Fig. 4.38 MPC-based RC horizons 115

Fig. 4.39 Frame-level rate control diagram 115

Fig. 4.40 Basic unit-level rate control diagram..... 119

Fig. 4.41 View-level bitrate distribution (Flamenco2)..... 124

Fig. 4.42 Bitrate and PSNR distribution at frame level (GOP #8)..... 125

Fig. 4.43 Bitrate distribution at BU level (GOP #8) 126

Fig. 5.1 ME/DE hardware architecture template 129

Fig. 5.2 SAD Calculator architecture..... 131

Fig. 5.3 Programmable Search Control Unit (a) FSM and (b) program memory 132

Fig. 5.4 On-chip video memory organization..... 134

Fig. 5.5 On-chip video memory cache tag..... 134

Fig. 5.6 Address Generation Unit (AGU) for dedicated video memory 135

Fig. 5.7 MVC prediction structure in our fast ME/DE algorithm..... 136

Fig. 5.8 Pipeline processing schedule of our ME/DE architecture..... 137

Fig. 5.9 GOP-level pipeline schedule 138

Fig. 5.10 MB-level pipeline schedule for (a) TZ module and fast ME/DE module in (b) Ultra fast and (c) Fast operation modes..... 139

Fig. 5.11 ME/DE search pattern for TZ search and Log search 141

Fig. 5.12 Number of pixels accessed in external memory..... 141

Fig. 5.13 ME/DE search window wastage..... 142

Fig. 5.14 Memory usage variation within one video frame..... 142

Fig. 5.15 Search Map prediction for the Log search 143

Fig. 5.16	Algorithm for Search Map prediction and the dynamic formation of the Search window	144
Fig. 5.17	Analyzing the memory requirements for ME/DE of different MBs in Ballroom sequence	146
Fig. 5.18	Search window memory organization with power gating	147
Fig. 5.19	Search Map prediction accuracy and on-chip memory misses	148
Fig. 5.20	ME/DE hardware architecture block diagram.....	149
Fig. 6.1	Spatial–temporal–disparity indexes for the benchmark multiview video sequences.....	154
Fig. 6.2	Time savings comparison with the state of the art	156
Fig. 6.3	Time savings considering the multiple QPs	157
Fig. 6.4	Time savings distribution summary.....	157
Fig. 6.5	Rate-distortion results for fast mode decision algorithms.....	158
Fig. 6.6	Complexity adaptation for MVC for changing battery levels.....	159
Fig. 6.7	Complexity reduction for the fast ME/DE	160
Fig. 6.8	Average number of SAD operations.....	161
Fig. 6.9	Fast ME/DE RD curves.....	161
Fig. 6.10	Bitrate prediction accuracy.....	162
Fig. 6.11	Accumulated bitrate along the time.....	163
Fig. 6.12	Rate-distortion results for the HRC.....	164
Fig. 6.13	ME/DE architecture with application-aware power gating physical layout.....	165
Fig. 6.14	Memory-related energy savings employing dynamic search window technique.....	166
Fig. A.1	JMVC encoder high-level diagram.....	176
Fig. A.2	Mode decision hierarchy in JMVC.....	177
Fig. A.3	Inter-frame search in JMVC.....	177
Fig. A.4	Communication in JMVC	178
Fig. B.1	MVC viewer main screen	182
Fig. B.2	Current macroblock-based analysis screenshot.....	183
Fig. B.3	Output example: four prediction directions and their respective accessed areas.....	183
Fig. B.4	Current macroblock-based analysis screenshot.....	184
Fig. B.5	Output example: reference frame access index considering two block matching algorithms: full search and TZ search	184
Fig. C.1	CES video analyzer user interface.....	186
Fig. C.2	CES video analyzer features.....	186
Fig. C.3	Coding mode analysis using CES video analyzer	187
Fig. C.4	ME/DE analysis using CES video analyzer	187

List of Tables

Table 4.1	Predictors hit rate and availability	84
Table 4.2	Quality states	99
Table 4.3	Detailed results for Δ PSNR, Δ Bitrate, and time savings compared to the exhaustive RDO-MD	101
Table 4.4	Comparison between the Quality States (QS)	106
Table 4.5	Comparison of our fast ME/DE algorithm to TZ Search	109
Table 4.6	Variables definitions	114
Table 4.7	Comparison of frame-level HRC Bitrate accuracy	118
Table 4.8	Comparison of BD-PSNR	118
Table 4.9	Comparison of proposed HRC Bitrate accuracy	122
Table 4.10	BD-PSNR and BD-BR comparison	123
Table 5.1	Search Pattern Memory example	133
Table 5.2	Comparison of our fast ME/DE algorithm	149
Table 6.1	Video encoder settings	152
Table 6.2	Simulation infrastructure	153
Table 6.3	Benchmark video sequences	153
Table 6.4	Bjøntegaard PSNR and BR for fast mode decision algorithm	159
Table 6.5	Bjøntegaard PSNR and BR for the HRC	163
Table 6.6	Motion and disparity estimation hardware architectures comparison	164

Abbreviations

3D	Three-Dimensional
3DTV	Three-Dimensional Television
3DV	Three-Dimensional Video (future video standard)
ASIP	Application-Specific Instruction-Set Processor
AVC	Advanced Video Coding
BR	Bitrate
BU	Basic Unit
CABAC	Context-Based Adaptive Binary Arithmetic Coding
CAVLC	Context-Based Adaptive Variable Length Coding
CIF	Common Intermediate Format
CODEC	Coder/Decoder
DC	Direct Current
DCT	Discrete Cosine Transform
DDR	Double Data Rate
DE	Disparity Estimation
DF	Deblocking Filter
DMV	Differential Motion Vector
DPB	Decoded Picture Buffer
DPM	Dynamic Power Management
DSP	Digital Signal Processing
DV	Disparity Vector
DVS	Dynamic Voltage Scaling
EPTZ	Early Prediction Terminator Zone
FIR	Finite Impulse Response
FPGA	Field Programmable Gate Array
FPS	Frames Per Second

FRExt	Fidelity Range Extensions
FSM	Finite State Machine
FTV	Free-Viewpoint Television
GB	Giga Bytes
GDV	Global Disparity Vector
GGOP	Group of <i>Group of Pictures</i>
GIPS	Giga Instructions per Second
GOP	Group of Pictures
HBP	Hierarchical Bi-Prediction
HD1080p	High Definition 1920 × 1080 Progressive
HDTV	High-Definition Digital Television
HEVC	High Efficiency Video Coding
HRC	Hierarchical Rate Control
HVS	Human Visual System
IC	Integrated Circuit
IEC	International Electrotechnical Commission
IEEE	Institute of Electric and Electronics Engineers
IQ	Inverse Quantization
ISO	International Organization for Standardization
IT	Inverse Transform
ITU-T	International Telecommunication Union—Telecommunication
JM	Joint Model for H.264
JMVC	Joint Model for MVC
JVT	Joint Video Team
KIT	Karlsruhe Institute of Technology
MB	Macroblock
MC	Motion Compensation
MD	Mode Decision
MDP	Markov Decision Process
ME	Motion Estimation
MPC	Model Predictive Controller
MPEG	Moving Picture Experts Group
MSE	Mean of Square Errors
MV	Motion Vector
MVC	Multiview Video Coding
MVP	Motion Vector Predictor
PC	Personal Computer
PDF	Probability Density Function
PID	Proportional-Integral-Differential Controller
PMV	Predictive Motion Vector
POC	Picture Order Counter

PSM	Power-State Machine
PSNR	Perceptible Signal-to-Noise Ratio
Q	Quantization
QCC	Quality-Complexity Class
QCIF	Quarter Common Intermediate Format
QHD	Quad HDTV
QP	Quantization Parameter
QS	Quality State
RC	Rate Control
RD	Rate-Distortion
RDO	Rate-Distortion Optimization
RDO-MD	Rate-Distortion Optimized Mode Decision
RGB	Red, Green, Blue
RL	Reinforcement Learning
RoI	Region of Interest
RTL	Register-Transfer Level
SAD	Sum of Absolute Distances
SATD	Sum of Absolute Transformed Distances
SI	Switching I
SIMD	Single Instruction Multiple Data
SoC	System on Chip
SP	Switching P
SRAM	Static Random Access Memory
SSE	Sum of Square Errors
T	Transform
UFRGS	Universidade Federal do Rio Grande do Sul
UVLC	Universal Variable Length Code
VCEG	Video Coding Experts Group
VGA	Video Graphics Array
VHDL	VHSIC Hardware Description Language
VHSIC	Very High Speed Integrated Circuit
VLIW	Very Large Instruction Word
VP	Viewpoint
YUV	Luminance, Chrominance Component 1, Chrominance Component 2