

CITY UNIVERSITY OF HONG KONG
香港城市大學

Fast Mode Decision and Rate Control for H.264/AVC and
SVC extension
视频编码的快速算法和码率控制

Submitted to
Department of Computer Science
電腦科學系
in Partial Fulfillment of the Requirements
for the Degree of Master of Philosophy
哲學碩士學位

by

HU Sudeng
胡速登

September 2010
二零一零年九月

Abstract

In this thesis, a Fast Inter-Mode Decision algorithm is proposed for H.264/AVC and Rate Control(RC) algorithms are proposed for temporal and spatial layer Scalable Video Coding (SVC) respectively.

Firstly, a new fast mode decision (FMD) algorithm is proposed for the state-of-the-art video coding standard H.264/AVC. Based on Rate-Distortion (RD) cost characteristics, all inter modes are classified into two groups, one is Skip mode (including both Skip and Direct modes) and all the other inter modes are called non-Skip modes. In order to select the best mode for coding a Macroblock (MB), minimum RD costs of these two mode groups are predicted respectively. Then for Skip mode, an early Skip mode detection scheme is proposed; for non-Skip modes, a three-stage scheme is developed to speed up the mode decision process. Experimental results demonstrate that the proposed algorithm has good robustness in coding efficiency with different Quantization parameters (Qp) and various video sequences and is able to achieve about 54% time saving on average while with negligible degradation in Peak-Signal-to-Noise-Ratio (PSNR) and acceptable increase in bit rate.

Secondly, for temporal scalable video coding, a novel frame-level RC algorithm is presented in this thesis. By introducing a linear quality dependency model, the quality dependency relation between a coding frame and its references is investigated for the hierarchical B-picture prediction structure. Linear Rate-Quantization (R-Q) and Distortion-Quantization (D-Q) models are introduced based on different characteristics of temporal layers. According to the proposed quality dependency model and R-Q and D-Q models for each temporal layer, an adaptive weighting factor is derived to

allocate bits efficiently among temporal layers. Experimental results on not only traditional QCIF/CIF but also Standard Definition (SD) and High Definition (HD) sequences demonstrate that the proposed algorithm achieves excellent coding efficiency as compared to other benchmark RC schemes.

Thirdly, for spatial layer of Scalable Video Coding, a novel rate control algorithm is presented in this thesis. A new best initial Qp model is proposed based on the power R-Q model. By applying the proposed sequence complexity measurement, the proposed model can provide proper initial Qp before encoding. Then the relationship between the best initial Qps of different layers is investigated and determination of the best initial for multiple Qps layer is introduced. Meanwhile by introducing a two stage RC scheme, a novel frame complexity estimation method is proposed. The dependency of the parameters in the RQ model is investigated to improve the model accuracy. The experimental results demonstrate that the proposed RC scheme and best initial Qp perform excellent coding efficiency and accurate bit achievement.

Contents

Abstract	i
Acknowledgement	iii
List of Publications	vii
List of Abbreviations	viii
List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 Background	1
1.2 Related Works	5
1.3 Thesis Organization	9
2 Fast Inter-Mode Decision In H.264/AVC	12
2.1 Fast Inter-Mode Decision Algorithm	12
2.1.1 Mode Classification Based on RD Cost Characteristics	12
2.1.2 Early Skip Mode Detection	15
2.1.3 Best Non-skip Mode Decision	19
2.1.4 Overall Algorithm	21
2.2 Experimental Results	24
2.3 Conclusions	25
3 Rate Control For Temporal Scalable Video Coding	27

3.1	R-D Dependency in Hierarchical Structure	27
3.1.1	Hierarchical B-picture Structure	27
3.1.2	Linear Quality Dependency	29
3.1.3	Propagation of Quality Dependency	31
3.2	Bit Rate Dependency	34
3.3	Linear R-D Model for Temporal Layer	37
3.3.1	R-Q Model	37
3.3.2	D-Q Model	40
3.4	R-D Optimized Frame-Level Rate Control	40
3.4.1	Adaptive Weighting Factor for Bit Allocation	40
3.4.2	GOP Level Bit Allocation	44
3.4.3	Frame Level Rate Control	44
3.5	Experimental Results On Temporal Layer	46
3.5.1	Simulation Setup	46
3.5.2	Accuracy of Bit Rate	48
3.5.3	R-D Performance	48
3.6	Conclusions	55
4	Initial Qp Determination for Spatial Scalable Video Coding	56
4.1	Best Initial Qp	56
4.2	Prototype of Best Initial Qp	57
4.3	Best Initial Qp for Multiple Layer	60
4.4	Experimental Results	62
4.5	Conclusions	65
5	Rate Control For Spatial Scalable Video Coding	66
5.1	Improved R-Q Model for Spatial Layer	66
5.1.1	R-D Optimization(RDO) in SVC	66
5.1.2	Frame Complexity	67
5.1.3	The Model Parameter	70

5.1.4	Bit Allocation Scheme	75
5.2	Experimental Results On Spatial Layer	77
5.2.1	Simulation Setup	77
5.2.2	Bit Achievement Accuracy	77
5.2.3	Buffer Regulation	79
5.2.4	R-D Performance	81
5.3	Conclusions	81
6	Conclusions and Future Work	85
6.1	Principal Contributions	85
6.1.1	Fast Mode Decision	85
6.1.2	RC algorithm for SVC	86
6.2	Future Work	87
	Bibliography	90

List of Abbreviations

R-D	Rate-Distortion
RC	Rate Control
MV	Motion Vector
FMD	Fast Mode Decision
SVC	Scalable Video Coding
HBP	Hierarchical B-Picture
MAD	Mean Absolute Difference
Qp	Quantization Parameter
RDO	Rate Distortion Optimization
GOP	Group of Picture
BR	Bit Rate
PSNR	Peak Signal-to-Noise Ratio

List of Tables

2.1	Linear Correlation Coefficient between RD costs of inter modes.	14
2.2	Percentages of Skip mode to be the best.	16
2.3	Probability of current inter mode when M_0 is Skip and M_b is not Skip. .	17
2.4	Probability of Skip mode to be the best when $J_{skip} < 0.9 \cdot J_{non-skip}^p$. . .	19
2.5	Experimental results.	26
3.1	The number of 8×8 blocks using different reference indices.	29
3.2	Summary of Simulation Parameters.	47
3.3	Summary of results of four RC algorithms on QCIF sequences.	49
3.4	Summary of results of four RC algorithms on CIF sequences.	50
3.5	Summary of results of four RC algorithms on SD&HD sequences. . . .	53
4.1	Sequence complexity ratio between spatial layers.	61
4.2	Summary of R-D performance results.	63
5.1	Average R^2 of RQ model fitting	73
5.2	β prediction error	74
5.3	Summary of Simulation Parameters.	76
5.4	Error of bit achievement	78
5.5	Summary of R-D performance results for base layer (QCIF) and en- hancement layer (CIF)	82
5.6	Summary of R-D performance results for base layer (CIF) and enhance- ment layer (4CIF)	83

List of Figures

2.1	RD Costs of different inter modes.	14
2.2	RD curves of “Mobile” (CIF).	22
2.3	RD curves of “Coastguard” (CIF).	22
2.4	RD curves of “Foreman” (QCIF).	23
2.5	RD curves of “Silent” (QCIF).	23
3.1	The dyadic HBP structure with GOP size of 16.	28
3.2	Quality dependency in the sequence “Paris” with GOP size 16. (a) Quality dependency relation between the 2nd picture and its reference 4th picture for QCIF. (b) Quality dependency relation between the 2nd picture and its reference 4th picture for CIF. (c) Quality dependency relation between the 4th picture and its reference 8th picture for QCIF. (d) Quality dependency relation between the 4th picture and its reference 8th picture for CIF.	30
3.3	Propagation of quality change effects on the sequence “Mobile” (QCIF). With the GOP size of 16, the Q_p value of the 8th picture is changed from 21 to 35, meanwhile the Q_p values of the 1st, 2nd, and 4th pictures are set to 38, 37, and 36, respectively.	32
3.4	Relation of bit rate dependency in sequence “Akiyo” (QCIF) with GOP size of 16. (a) The 2nd picture and its reference 4th picture. (b) The 4th picture and its reference 8th picture.	35

3.5	Bit rate dependency in sequence “Football” (QCIF). (a) Q_p of Level 1 changes. (b) Q_p of Level 2 changes. (c) Q_p of Level 3 changes. (d) Q_p of Level 4 changes.	36
3.6	Relationship between R and $1/Q_{step}$ for the 16th picture in level 0, the 8th picture in level 1, the 4th picture in level 2, the 2nd picture in level 3 within the same GOP. (a) 1st GOP in “Coastguard” (QCIF). (b) 3rd GOP in “Coastguard” (QCIF).	39
3.7	The linear relationship of D and Q_{step} for 16th picture in level 0, 8th picture in level 1, 4th picture in level 2, 2nd picture in level 3 within the same GOP. (a) 1st GOP in “Soccer” (QCIF). (b) 3rd GOP in “Soccer” (QCIF).	41
3.8	Comparison of PSNR fluctuation vs. frames for different algorithms with target bit rate at 512 kbps. (a) “Foreman” (CIF). (b) “Highway” (CIF). (c) “Grandma” (QCIF). (d) “Silent” (QCIF).	51
3.9	R-D curves. (a) “Mother” (QCIF). (b) “Silent” (QCIF). (c) “Coastguard” (CIF). (d) “Stefan” (CIF).	54
4.1	The relationship between best initial Q_p and $\ln(B)$ for different sequences	58
4.2	The relationship between best initial Q_p and $\ln(B)/\alpha_a$ for different sequences.	59
4.3	Comparison of Y-PSNR fluctuation in both of base and enhancement layers.	64
5.1	Comparison of two frame complexity estimation. (a)-(d) in sequence “Soccer”, (e)-(h) in sequence “Akiyo” (BL@QCIF, EL@CIF). Q_{p1} equals 29 and Q_{p2} equals 27 for both layers. Actual MAD is calculated after RDO and α is calculated by Eq. (5.5) and β is fixed at -1.2.	69

5.2	RQ Curves of both basement layer and enhancement layer. For base layer BL: Qp=10:2:44 while EL: Qp=30. For enhancement layer, BL: Qp=30 while EL: Qp=10:2:44.	70
5.3	β_{BL} VS. β_{EL} . 161 frames are included in every sequence.	72
5.4	Bit achievement with fixed target bit.	77
5.5	Buffer status	80