CAMED: Complexity Adaptive Motion Estimation & Mode Decision for H.264 Video

Yong Wang

Prof. Shih-Fu Chang

Digital Video and Multimedia (DVMM) Lab, Columbia University

Outline

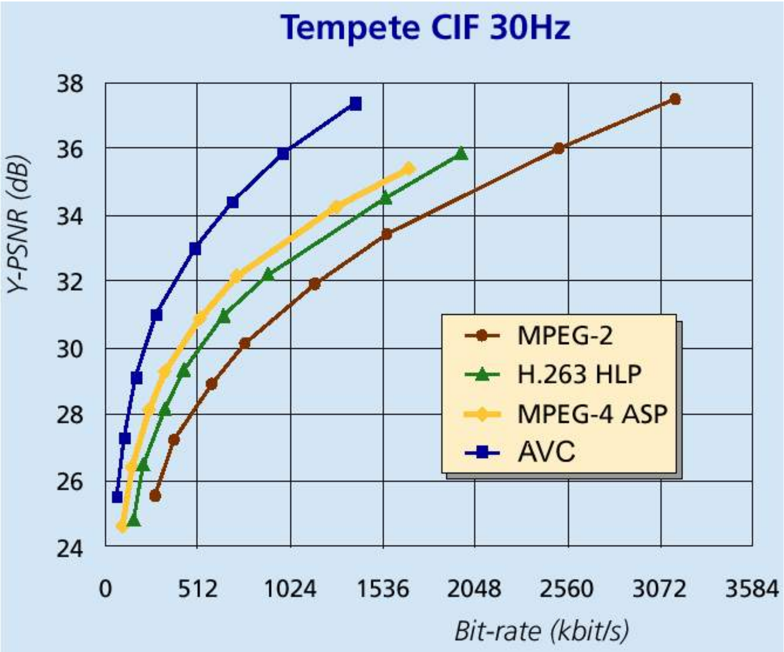
- Complexity aware H.264 coding
- CAMED
  - Rate-Distortion-Complexity video coding
  - Complexity cost function
  - Complexity control
- Experiment results
- Conclusions & future work

H.264: Emerging Video Standard

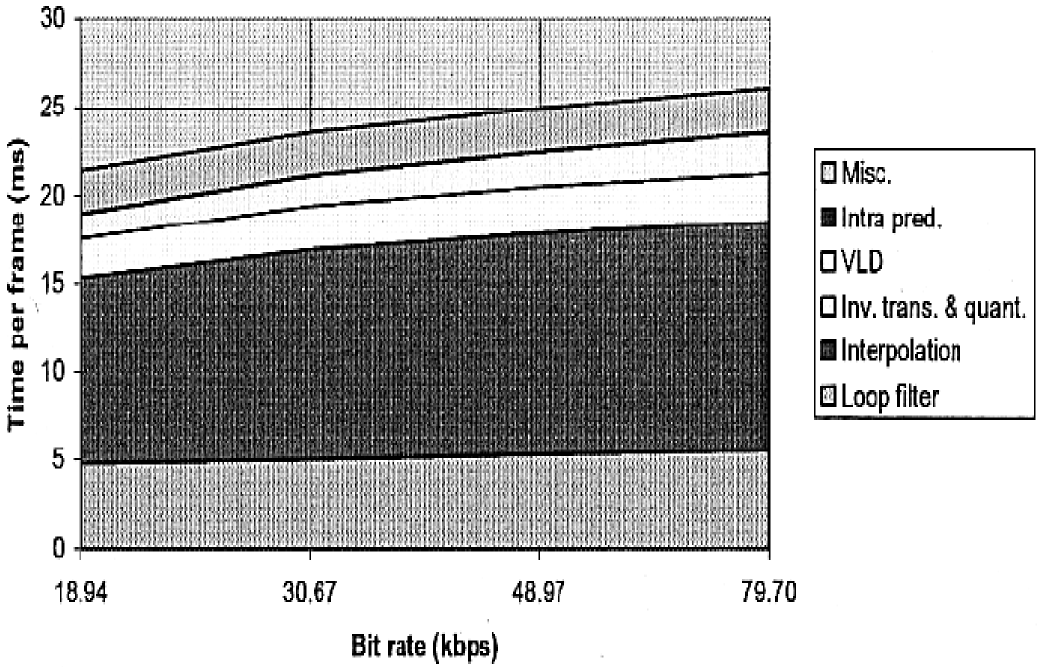
- More efficient motion prediction
  - Variable block size motion compensation (as small as 4 x 4)
  - Quarter pixel interpolation
  - Multiple reference frames (up to 5)
  - Decoupling of referencing order from display order
  - Directional spatial prediction for intra coding
  - SKIP/DIRECT Mode
- Improved coding efficiency
  - Small block-size transform (4 x 4 integer transform)
  - Lossless inverse transform
  - Context-based adaptive binary arithmetic coding (CABAC)
  - Switching P (SP) and switching I (SI) pictures

Complexity Issue In H.264

- Much better quality, with much higher complexity



[Schäfer'03]



[Lappalainen'03]

Complexity Aware H.264 Coding

- Important for mobile video solutions
- H.264: Emerging coding standard
  - Much higher quality
  - Much higher complexity
- Complexity constrained video coding

$$
\begin{aligned}
&\min D(\mathbf{P}),\\
&s.t., R(\mathbf{P}) \le R_T, \quad \mathbf{P}: QP, MV, Mode, ?\\
&\quad C(\mathbf{P}) \le C_T
\end{aligned}
$$

Approaches for Comp. Reduction

- Hardware level
  - Hardware architecture (MMX, SIMD [Chen'04])
  - Low power chip (Advanced voltage scaling [ARM'03])
  - High capacity battery [GIT]

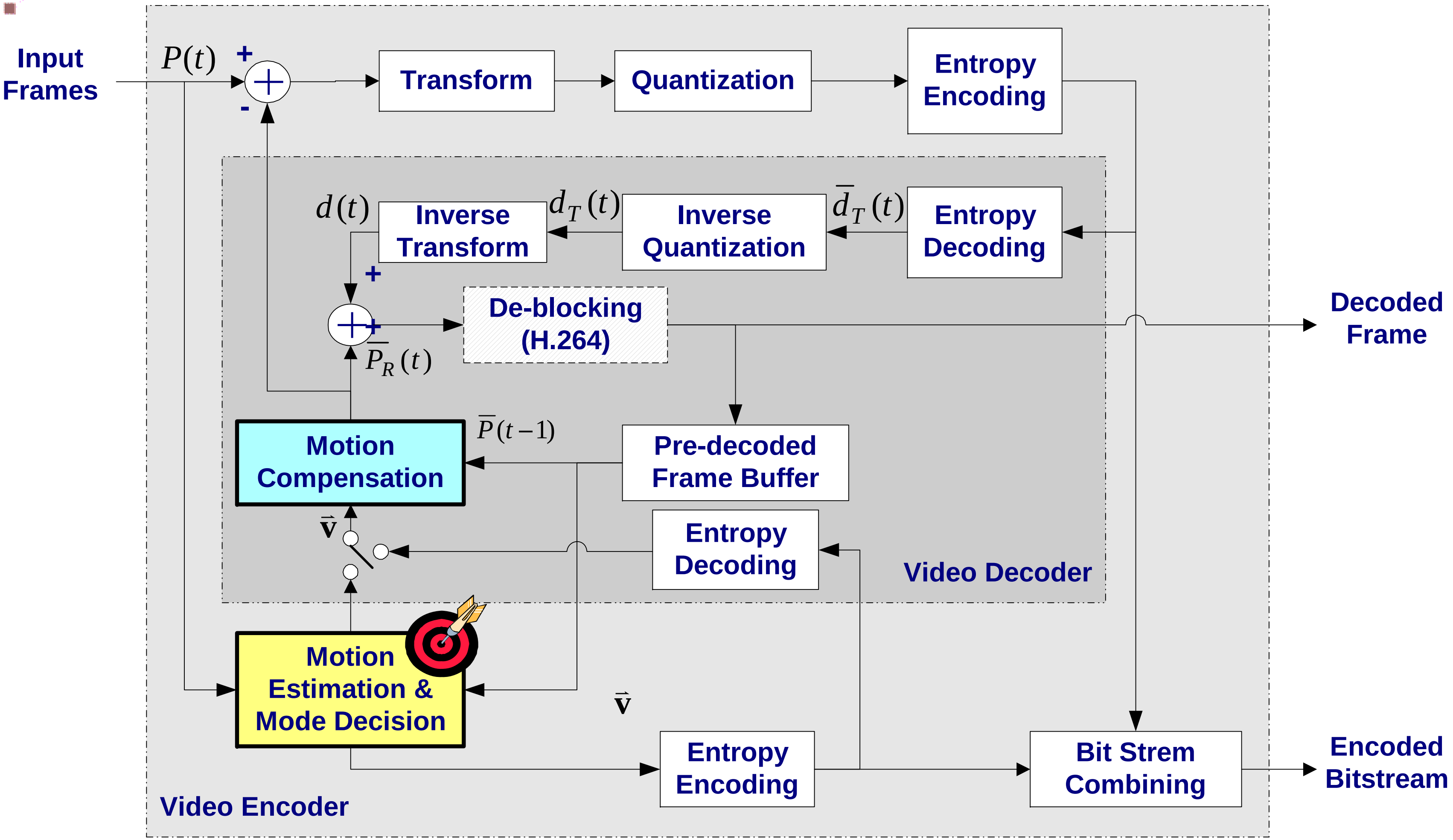
- Encoder complexity
  - Fast motion estimation (FME) [Li'05, Tourapis'03]
  - Fast mode decision (FMD) [Zhu'05, Peng'05]
  - Fast reference frame selection [Su'04]
  - Adaptive I-B-P GOP structure [Ray'04]
  - SAD (Sum of Absolute Difference) complexity modeling [He'04]

- Decoder complexity
  - Smart inverse DCT [Ortega'00]
  - Interpolation approximation [Wang'05]
  - Implementation optimization (buffer reusage, assembly coding)
  - Complexity Adaptive ME & MD (CAMED)

Columbia University in the City of New York

dvmm DIGITAL VIDEO MULTIMEDIA LAB

CAMED: Philosophy



Columbia University in the City of New York

dvmm DIGITAL VIDEO MULTIMEDIA LAB

CAMED: Highlight

- Develop an encoding algorithm that achieves both high video quality and low decoding complexity
- Addressed the most significant module (Motion Compensation) in decoders, so that the complexity reduction is efficient.
- Nothing need to be done in decoders, and the extra cost in encoders is trivial
- Full standard compliant
- The proposed method can be a valuable complement for hardware-level and encoder-level solutions
- Very promising for mobile video applications.

Columbia University
in the City of New York

dvmm
DIGITAL VIDEO MULTIMEDIA LAB

R-D-C Optimization

- R-D-C Lagrange optimization

$$\mathbf{V}_C^*(B,M) = \underset{\mathbf{V}\in\sup\{\mathbf{V}\}}{\arg\min}\; J_{MOTION}^{R,D,C}(\mathbf{V}|B,M)$$

$$= \underset{\mathbf{V}\in\sup\{\mathbf{V}\}}{\arg\min}\left\{ J_{MOTION}^{R,D}(\mathbf{V}|B,M) + \gamma_{MOTION} C_{MOTION}(\mathbf{V}|B,M)\right\}$$

$$M_C^*(MB,QP) = \underset{M\in\sup\{M\}}{\arg\min}\; J_{MODE}^{R,D,C}\left(M|MB,QP\right)$$

$$= \underset{M\in\sup\{M\}}{\arg\min}\left\{ J_{MODE}^{R,D}\left(M|MB,QP\right) + \gamma_{MODE} C_{MODE}\left(M|MB,QP\right)\right\}$$

$$\gamma_{MOTION} = \sqrt{\gamma_{MODE}}$$

Complexity Cost Function

- Pixel-based interpolation operation

$$c_P(\mathbf{V}) = \begin{cases} 0 & \mathbf{V} \text{ is integer MV} \\ e_6 & \mathbf{V} \text{ is subpixel } b,h \\ e_6 + e_2 & \mathbf{V} \text{ is subpixel } a,c,d,n \\ 2e_6 + e_2 & \mathbf{V} \text{ is subpixel } e,g,p,r \\ 7e_6 & \mathbf{V} \text{ is subpixel } j \\ 7e_6 + e_2 & \mathbf{V} \text{ is subpixel } i,f,k,q \end{cases}$$

G	a	b	c	H
d	e	f	g	
h	i	j	k	m
n	p	q	r	
M		s		N

- Can be further simplified counting 7-tap interpolations.

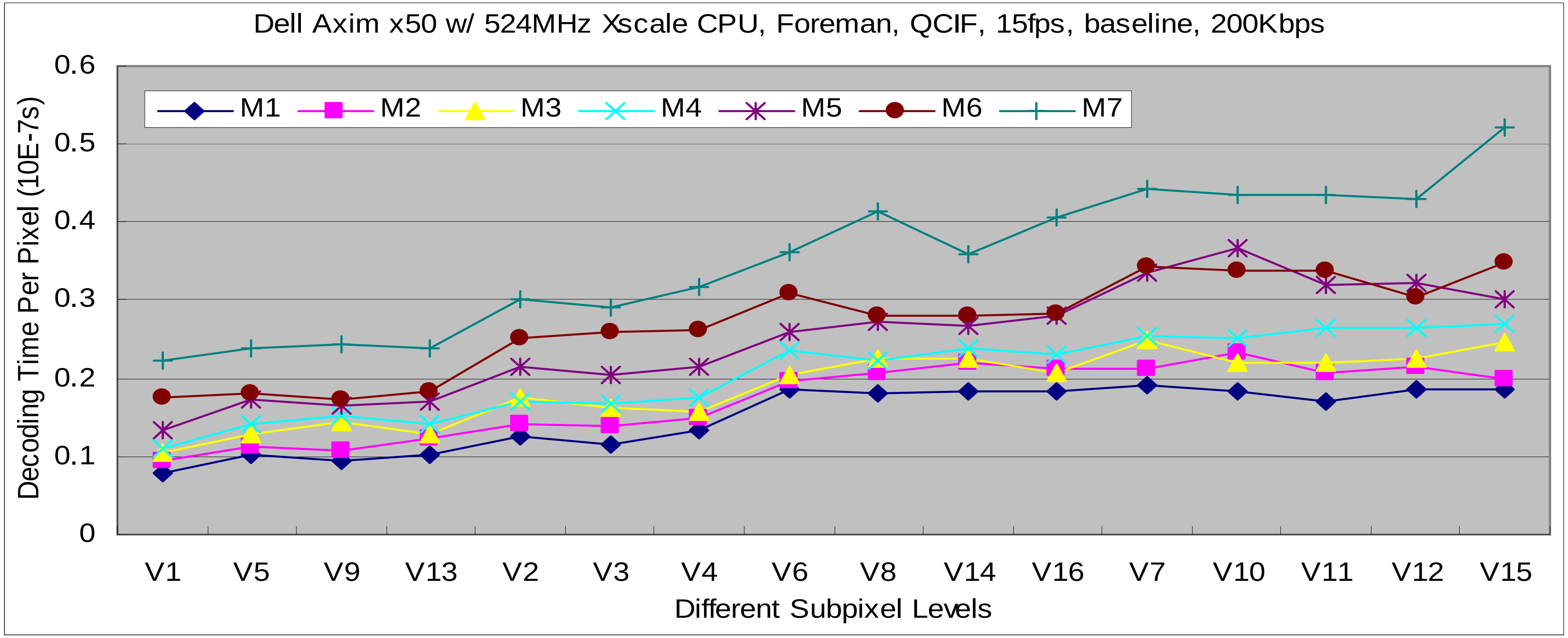
Complexity Cost Function (cont.)

- Block-based interpolation operation

Mode	Integer	1 6-tap	2 6-tap	7 6-tap
Skip/Direct	0	256	512	592
16x16	0	256	512	592
16x8	0	128	256	296
8x16	0	128	256	296
8x8	0	64	128	168
8x4	0	32	64	84
4x8	0	32	64	84
4x4	0	16	32	52
Intra	0	0	0	0

Complexity Cost Function (cont.)

- Complexity benchmark
  - Real time AVC decoder in Dell Axim x50 PDA



Columbia University in the City of New York

dvmm DIGITAL VIDEO MULTIMEDIA LAB

Complexity Cost Function (cont.)

- Modeling based on complexity benchmark

$$c(\mathbf{v}, M) = b_M^{-p}\left(n_{6,\mathbf{v}} c_6 + n_{2,\mathbf{v}} c_2 + c_l\right)$$

- Content-independent: the complexity is mainly decided by the arithmetic calculations
- Can be summarized into look-up table to reduce encoding cost
- Reflects the actual computing complexity, including not only interpolation, but also other cost rather than interpolation such as memory access and logic flow.

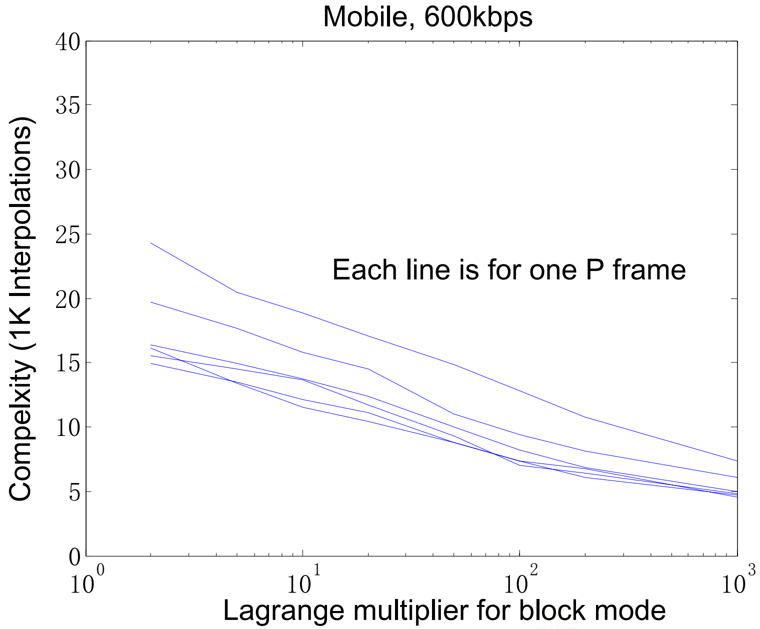
Complexity Control

- The procedure to adjust coding parameters so that the complexity level meets the target requirement.
  - Similar to rate control
  - Control parameter: Lagrange multiplier $\gamma_{MOTION}, \gamma_{MODE}$

- Complexity modeling

$$C(\gamma_{MODE}) = D\left(K_1 \ln(\gamma_{MODE}) + K_0\right)$$

$$\gamma_{MODE}(t) = \exp\left\{\frac{C(t) - K_0 D(t)}{K_1 D(t)}\right\}$$

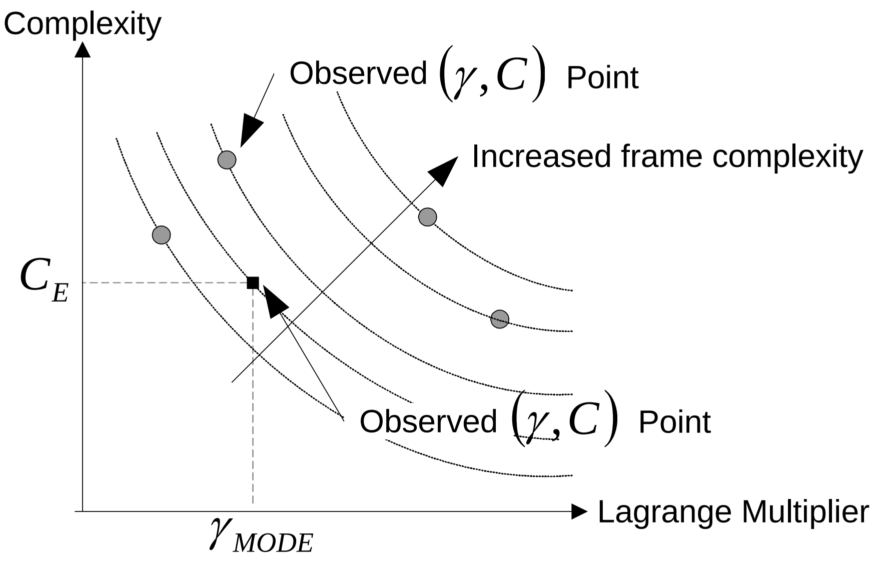


Columbia University in the City of New York

dvmm DIGITAL VIDEO MULTIMEDIA LAB

Complexity Control (cont.)

- Model estimation
  - Collect complexity stat using sliding window
  - Estimate model parameters through regression

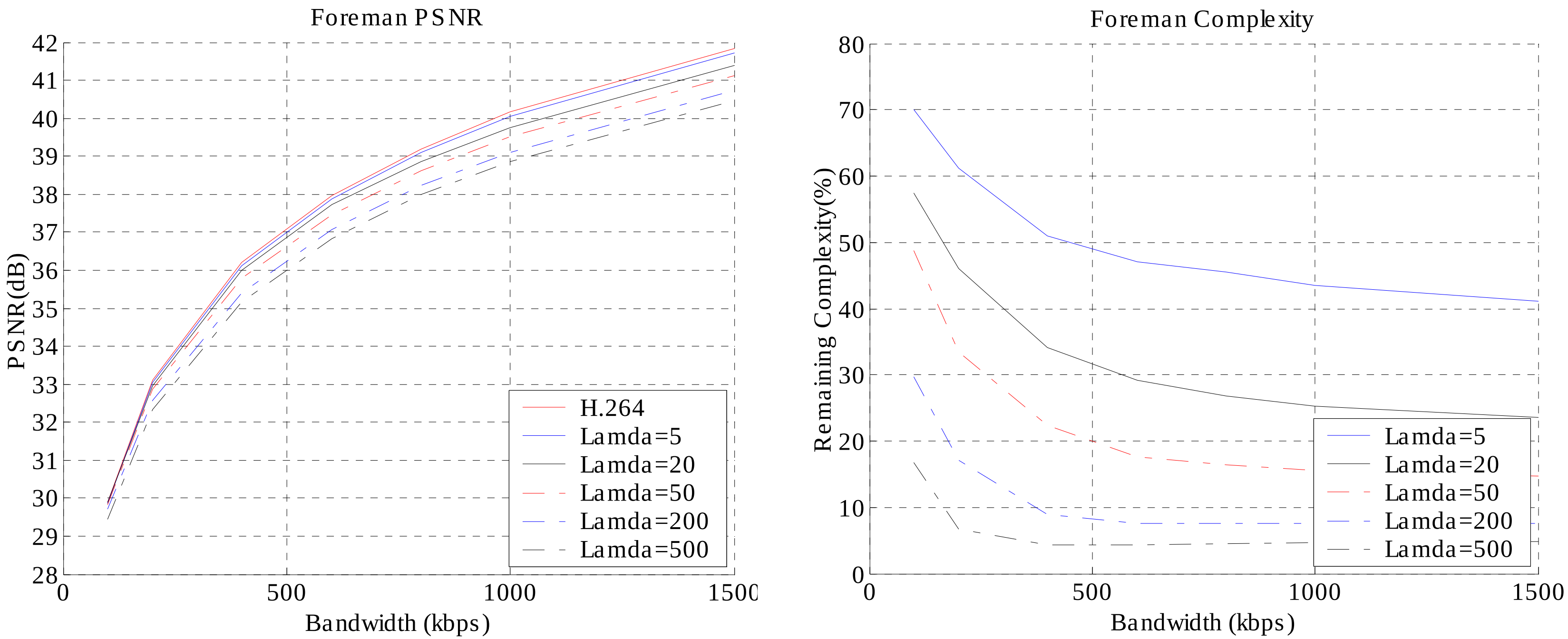


- Buffer management
  - Use complexity buffer to simulate the complexity usage status on the decoder side
  - Estimate of available resource, adjust the parameters, and avoid issues of buffer overflow or underflow

Experiment Setup

- Environment
  - Software: H.264 JM 8.2
  - Sequences: *Akiyo*, *Foreman*, *Stefan*, *Mobile*
  - Resolution: QCIF/CIF
  - Profile: Main/Baseline
  - Bit rate: 100kbps~1Mbps
  - Block mode: all on
  - Direct mode: temporal
  - Motion estimation search range: 32 pixels

R-D-C Optimization



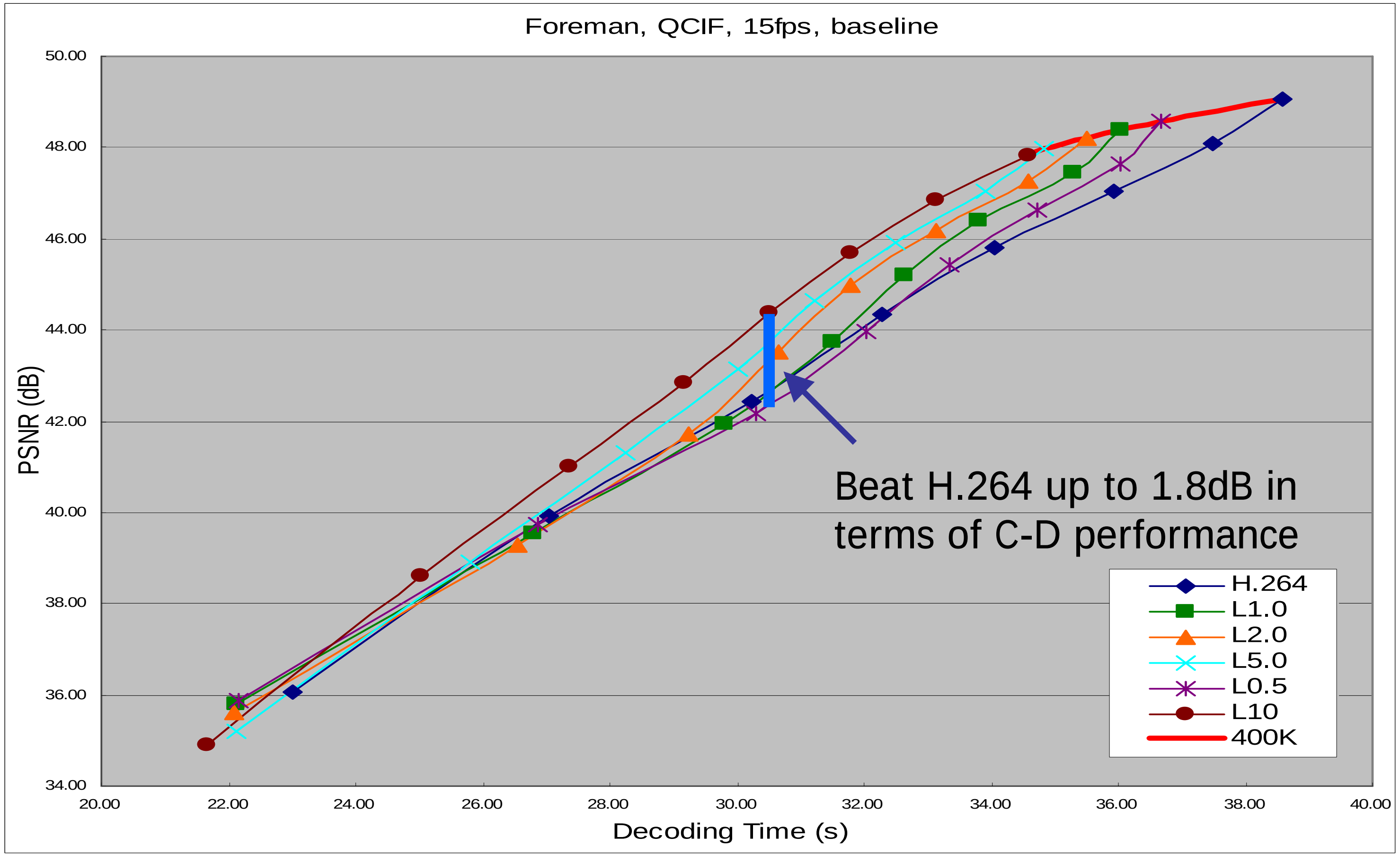
Up to 60% 6-tap interpolation can be saved while keeping the video quality almost intact (quality difference less than 0.2dB)

Demonstration

Columbia University in the City of New York

dvmm DIGITAL VIDEO MULTIMEDIA LAB

Results: Complexity Benchmark



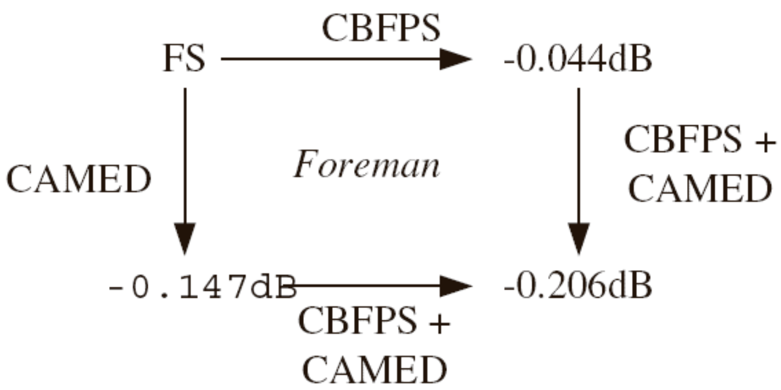
Dell Axim x50

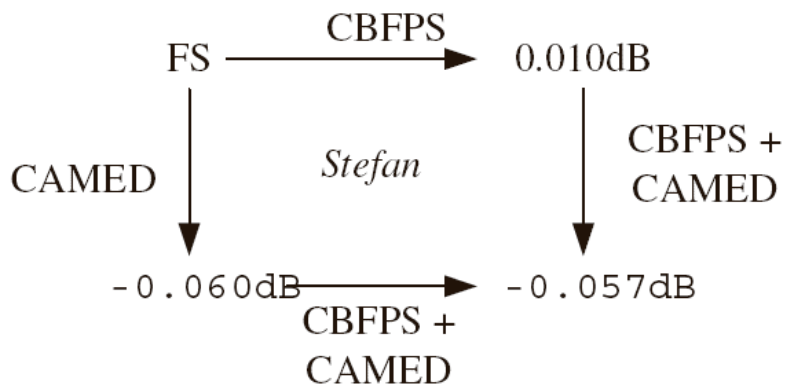
Columbia University in the City of New York

dvmm DIGITAL VIDEO MULTIMEDIA LAB

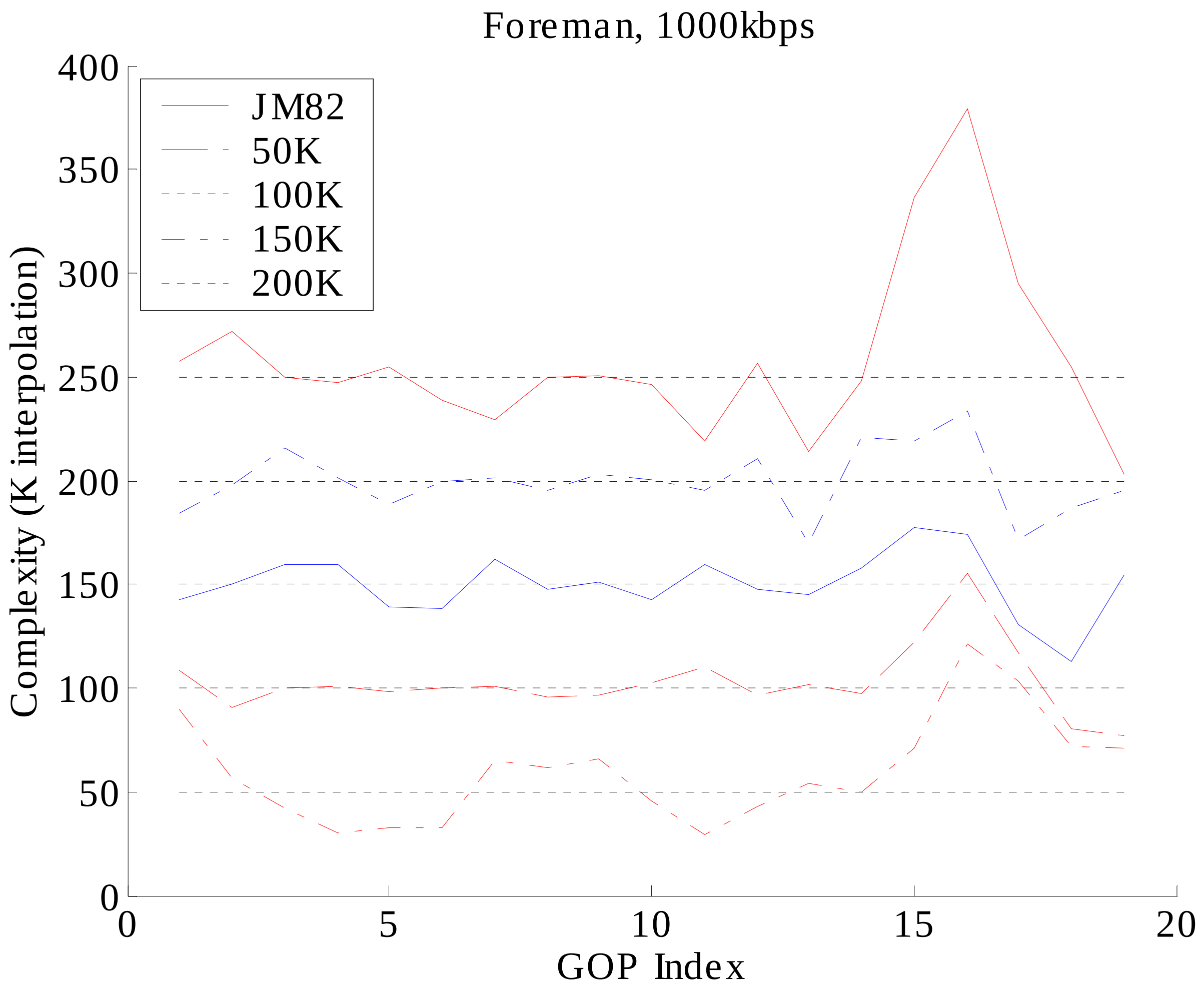
Results: Compatibility w/ FME

Foreman	PSNR(dB)†	Encoding Saving♮	Decoding Saving♯
FS	36.024	0.00%	0.00%
CBFPS	-0.044	61.80%	17.39%
CAMED $\gamma_{MODE} = 10$	-0.147	0.00%	70.08%
CBFPS+$\gamma_{MODE} = 10$	-0.206	64.75%	74.28%
CAMED $\gamma_{MODE} = 50$	-0.312	0.00%	85.63%
CBFPS+$\gamma_{MODE} = 50$	-0.390	66.09%	86.49%





Results: Complexity Control



Results: Complexity Control

Target Complexity (6-tap Interpolation)		5K	10K	20K	30K	40K
Foreman	Complexity control error (%)	37.07	5.00	1.50	3.38	3.09
	Complexity saving (%)	88.06	81.70	65.67	45.95	28.14
	Quality degradation (dB)	0.22	0.16	0.09	0.03	0.03
Stefan	Complexity control error (%)	39.38	19.45	0.32	0.26	0.37
	Complexity saving (%)	88.07	86.21	65.64	48.49	31.25
	Quality degradation (dB)	0.40	0.30	0.09	-0.03	0.06
Mobile	Complexity control error (%)	114.97	7.48	0.65	1.34	4.73
	Complexity saving (%)	81.89	81.89	66.08	50.13	35.79
	Quality degradation (dB)	0.40	0.40	0.17	0.14	0.06

CIF, 30fps, 100kbps

Conclusions & Future Work

- Proposed CAMED for complexity-efficient H.264 coding
- Applied rigorous R-D-C optimization framework
- Defined several complexity cost functions
- Proposed the algorithm of complexity control
- Performance Validated through extensive simulation experiment

- Future work
  - Theoretical analysis on the relationship between the complexity level and the control parameter (i.e., the Lagrange multiplier)
  - Considering hardware acceleration (such as MMX, SSE)
  - More efficient complexity modeling and complexity control

THANK YOU!

Prof. Shih-Fu Chang

sfchang@ee.columbia.edu

Yong Wang

ywang@ee.columbia.edu

Columbia University in the City of New York

DIGITAL VIDEO • MULTIMEDIA LAB