Overlapping in Video Signal Processing

Jiro Katto

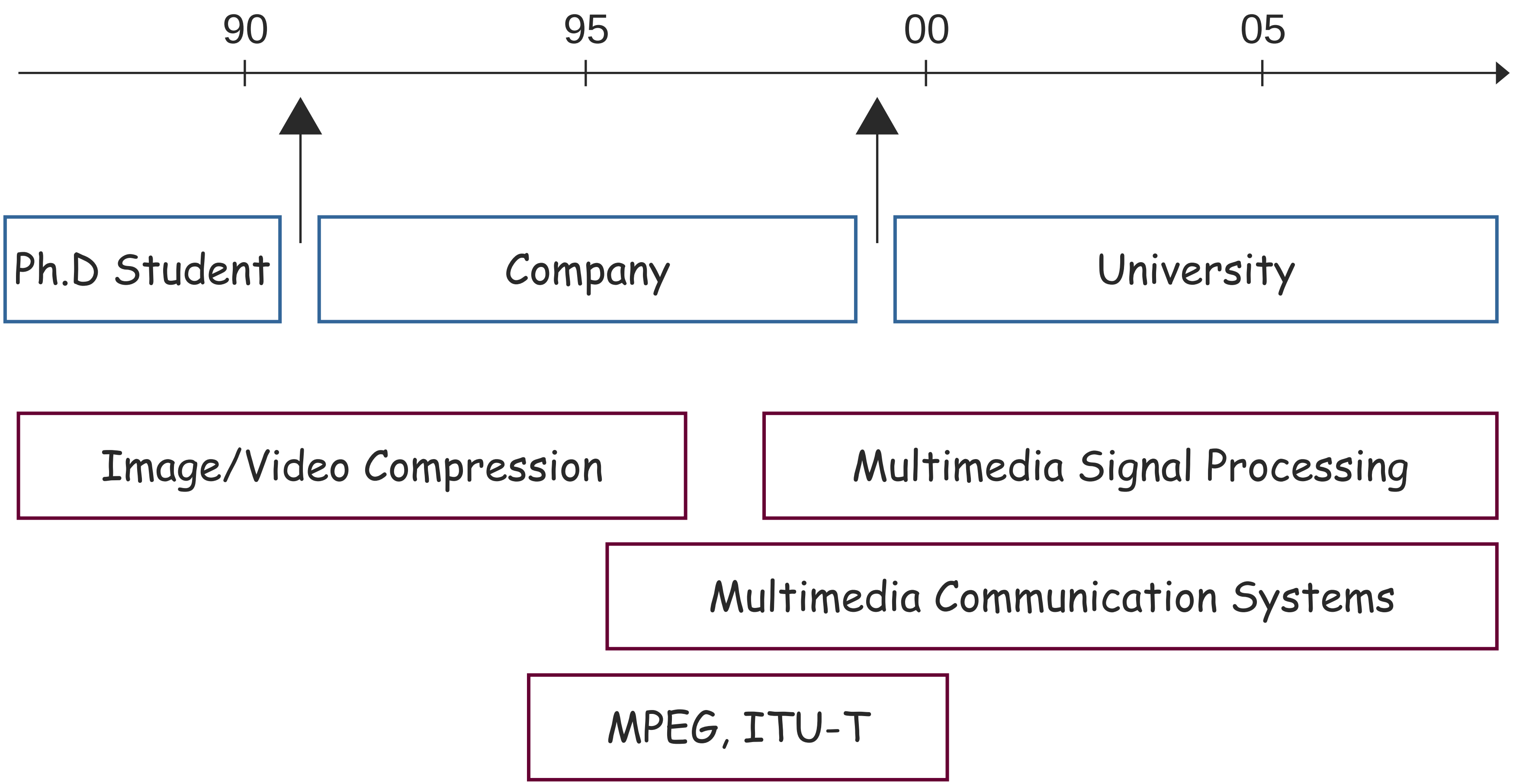
Dept. of Computer Science, Waseda University

E-Mail: katto@waseda.jp

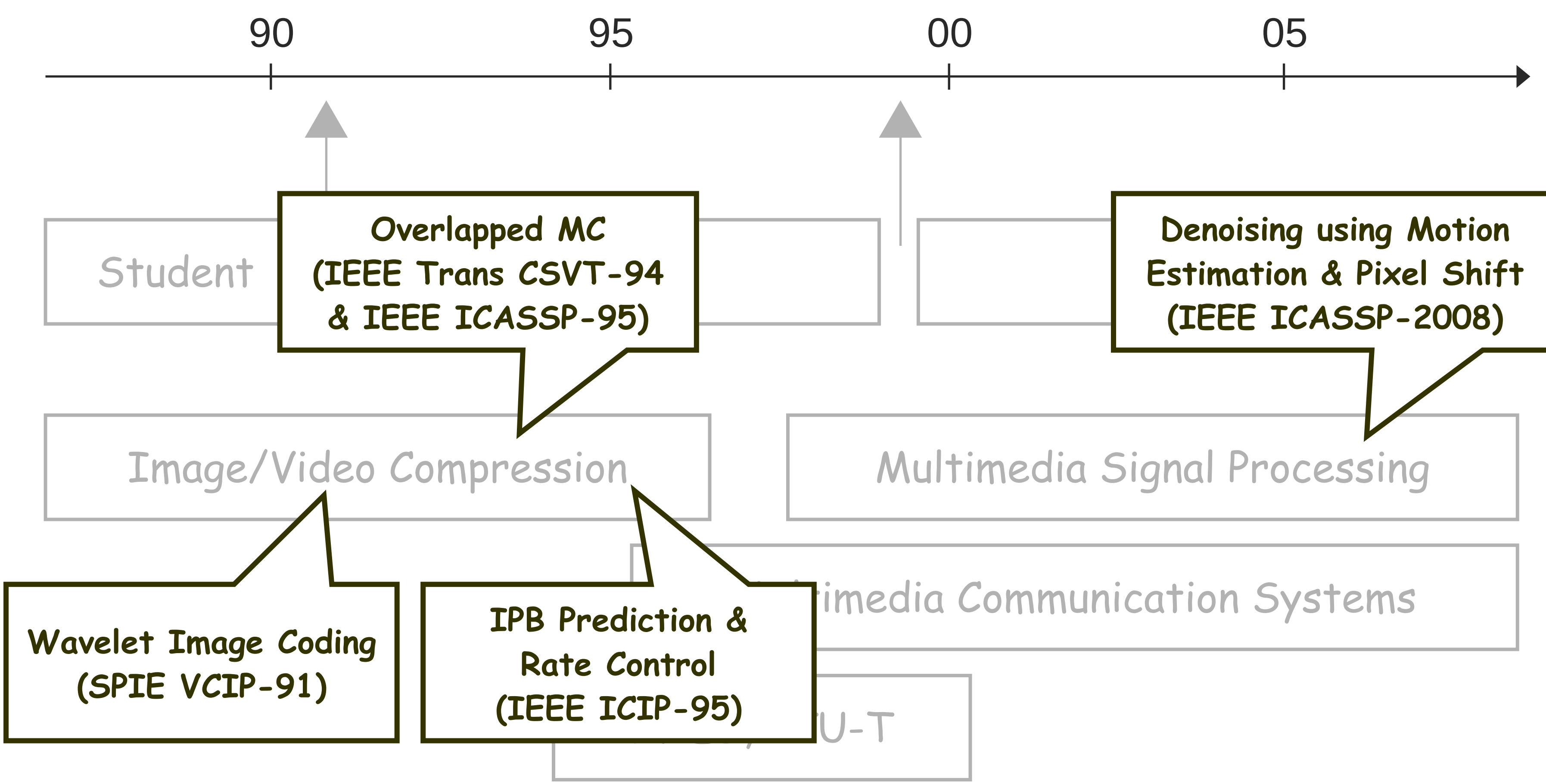
Outline

- Introduction
- “Overlapped“ Transform
- “Overlapped“ Motion Compensation
- “Overlapped“ Prediction
- Denoising by “Overlapping“
- Conclusions

What I am ...

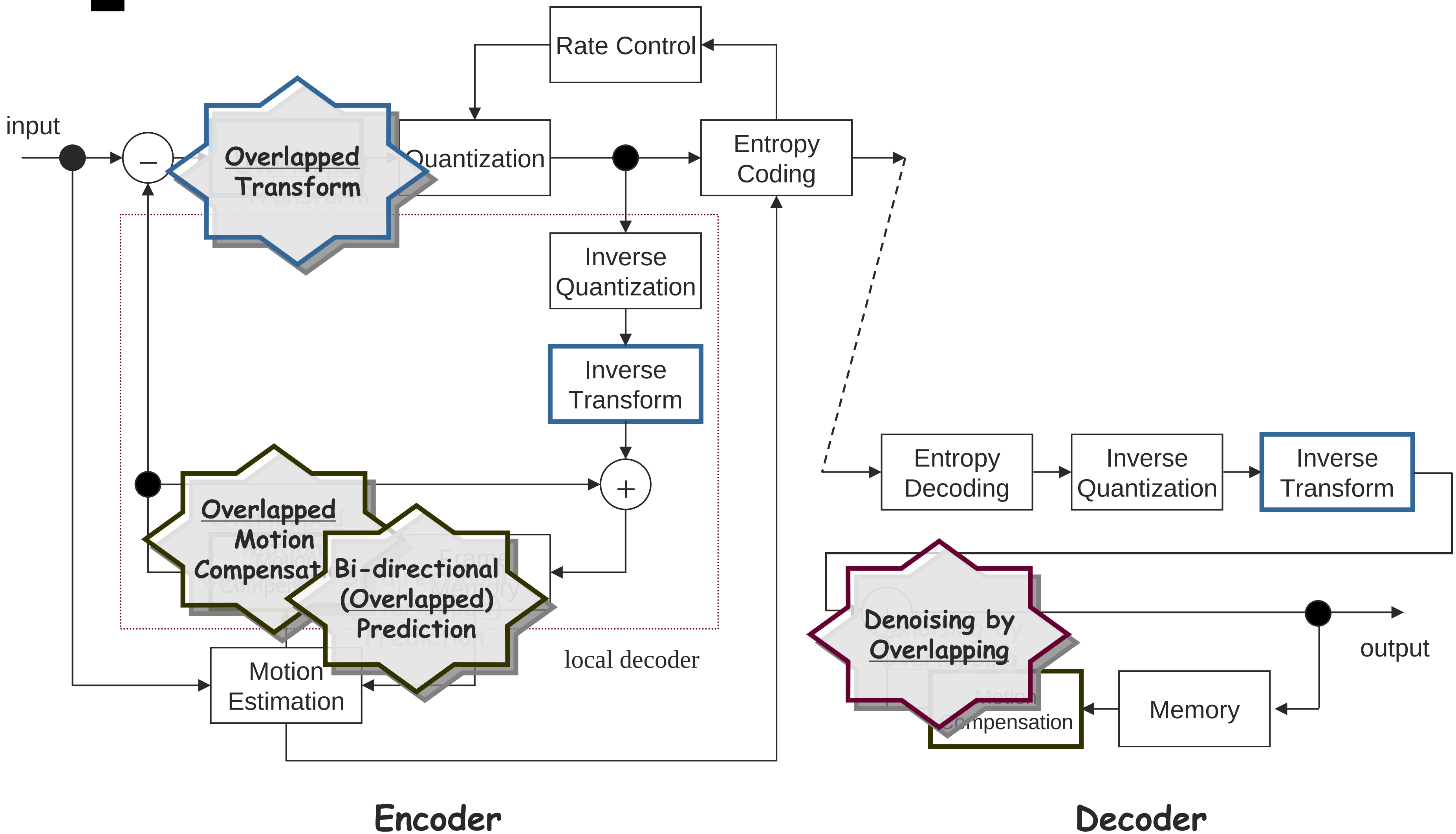

90
95
00
05
Ph.D Student
Company
University
Image/Video Compression
Multimedia Signal Processing
Multimedia Communication Systems
MPEG, ITU-T

What I did ...



90
95
00
05
Student
Overlapped MC (IEEE Trans CSVT-94 & IEEE ICASSP-95)
Denoising using Motion Estimation & Pixel Shift (IEEE ICASSP-2008)
Image/Video Compression
Multimedia Signal Processing
Wavelet Image Coding (SPIE VCIP-91)
IPB Prediction & Rate Control (IEEE ICIP-95)

Video Coder

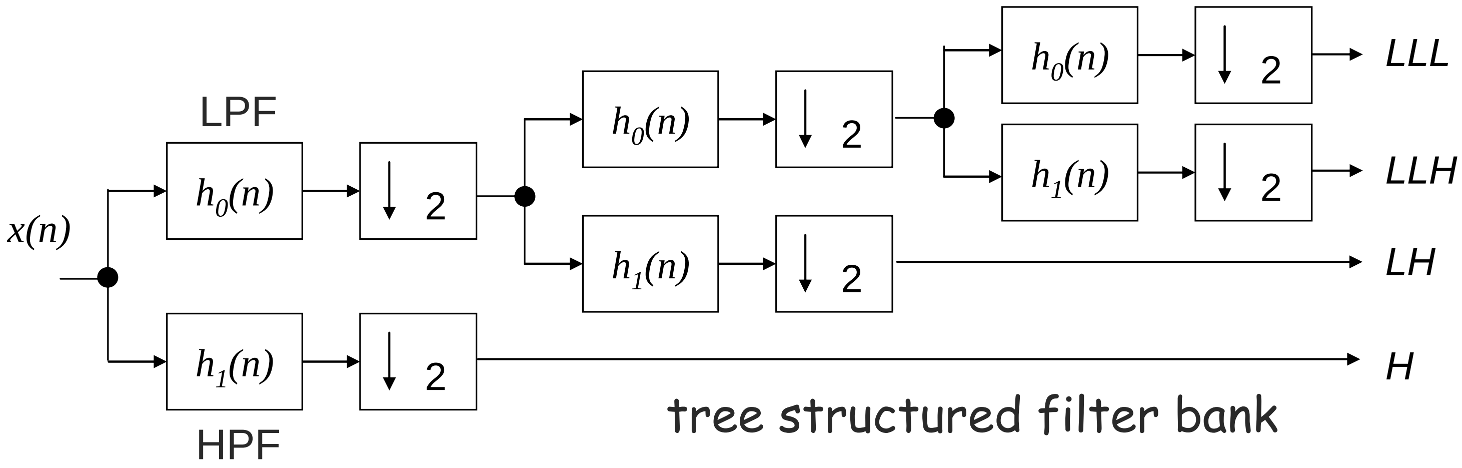

input
Rate Control
Overlapped Transform
Quantization
Entropy Coding
Inverse Quantization
Inverse Transform
Overlapped Motion Compensation
Bi-directional (Overlapped) Prediction
Motion Estimation
local decoder
Entropy Decoding
Inverse Quantization
Inverse Transform
Denoising by Overlapping
Compensation
Memory
output
Encoder
Decoder

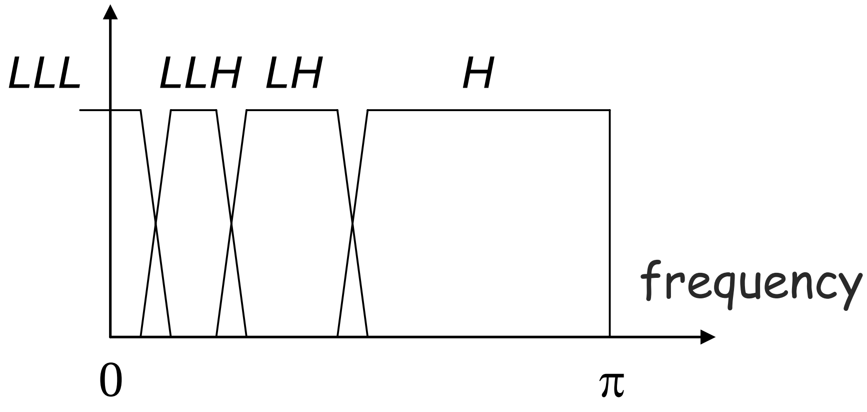
1. “Overlapped” Transform

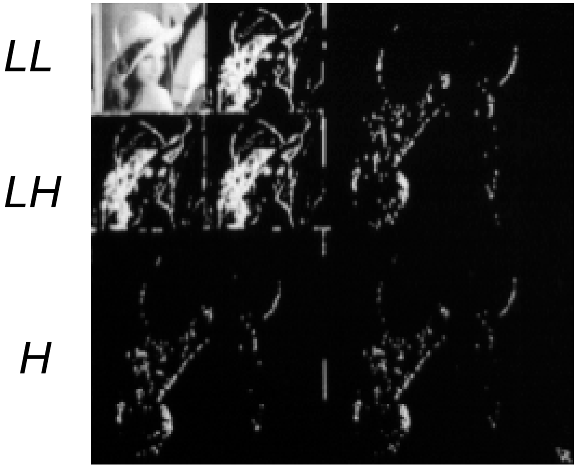
- Block Transform
  - DCT (Discrete Cosine Transform)
  - DFT (Discrete Fourier Transform)
  - KLT (Karhunen-Loeve Transform)
- Overlapped Transform
  - Wavelet transform / Subband coding
  - LOT (Lapped Orthogonal Transform)
  - MDCT (Modified DCT)

J.Katto and Y.Yasuda: "Performance Evaluation of Subband Coding and Optimization of Its Filter Coefficients", SPIE VCIP'91, Nov.1991.

Wavelet Transform

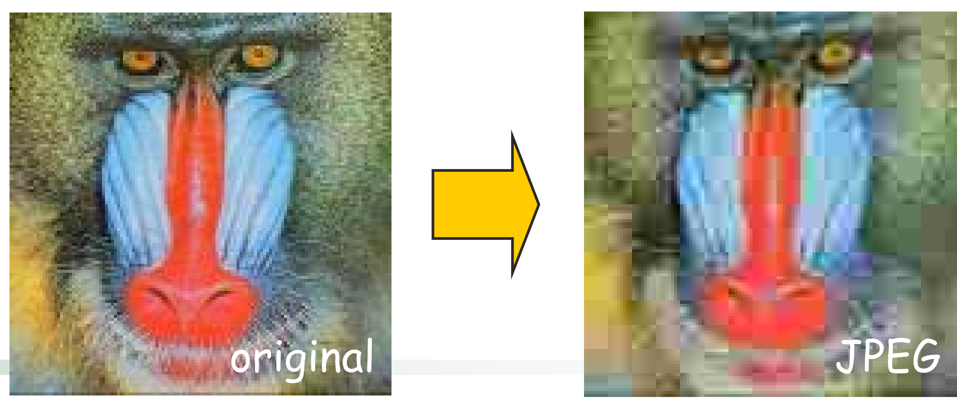






Problems (at that time)

- DCT was widely used in image/video compression (JPEG/MPEG) due to its suboptimal compression capability, but suffers from <u>blocking artifacts</u>.
- Wavelet transform is a promising alternative because it can alleviate the blocking artifacts due to its <u>overlapping properties</u>.
- However, there was <u>no theoretical compression performance measure</u> of Wavelet transform.



Optimum Bit Allocation Problem

minimize reconstruction error variance,

$$\sigma_r^2 \left(= f(\sigma_q^2) \right)$$

quantization error variance

under constant bit rate constraint

$$R \leq const$$

Methods: e.g. Lagrange multiplier method

N.S.Jayant and P.Noll: “Digital Coding of Waveforms”, Prentice Hall, 1984.

Coding Gain

$N \times N$ matrix (orthogonal)

optimum bit allocation problem

$$G_{TC} = \frac{\frac{1}{N}\sum_{k=0}^{N-1}\sigma_k^2}{\left(\prod_{k=0}^{N-1}\sigma_k^2\right)^{\frac{1}{N}}}$$

"Coding Gain"

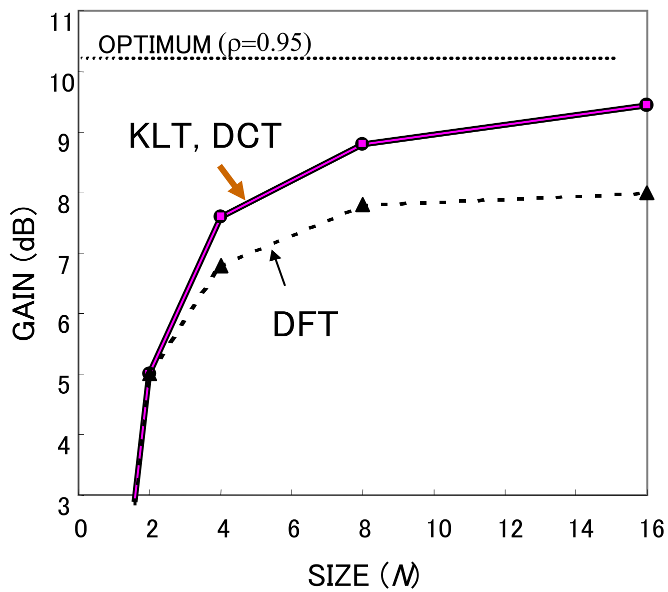
σ_k^2: coefficient variance

TC: transform coding

[Theory]

DCT can approximate KLT when $\rho \to 1$

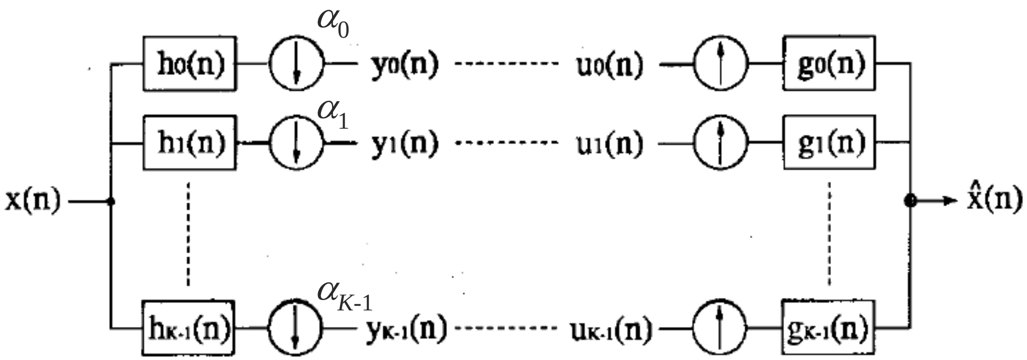
ρ : correlation factor



N.S.Jayant and P.Noll: "Digital Coding of Waveforms", Prentice Hall, 1984.

Coding Gain of Wavelet Transform ?

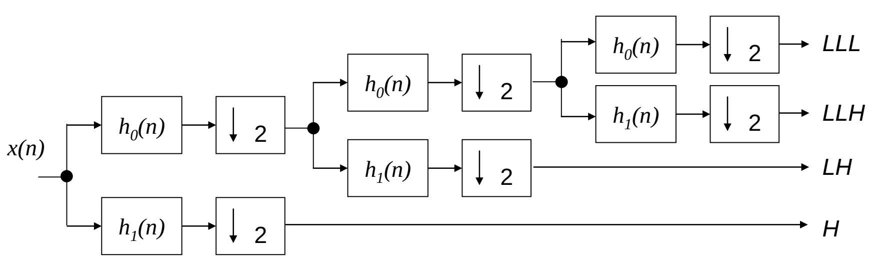
sampling ratio



multirate filter bank

optimum bit allocation problem

$$G_{wavelet} = \frac{1}{\prod_{k=0}^{K-1}\left(\frac{A_k B_k}{\alpha_k}\right)^{\alpha_k}}$$



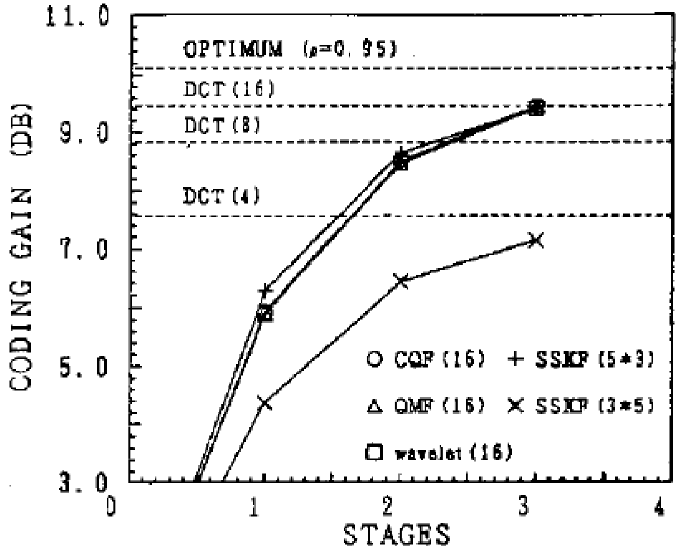
tree-structured filter bank

$$\sigma_{y_k}^2 = A_k \cdot \sigma_x^2 \quad \leftarrow \{h_k(n)\}$$

$$\sigma_r^2 = \sum_{k=0}^{K-1} B_k \cdot \sigma_{q_k}^2 \quad \leftarrow \{g_k(n)\}$$

Advantages & Evaluation Example

- Proposed coding gain <u>includes classical coding gain as its special case</u>, and enables integrated performance comparison with block transform.
- Proposed coding gain can be applied to <u>non-orthogonal filter bank</u> (e.g. bi-orthogonal case).



provides theoretical validity on JPEG-2000's 9x7 filter (bi-orthogonal filter)

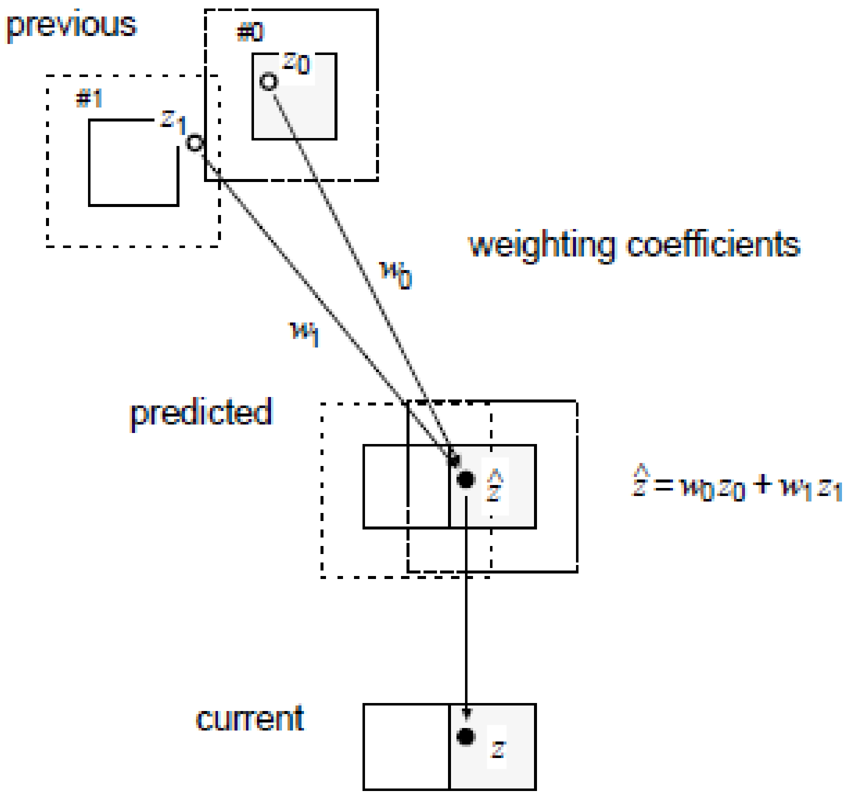
2. “Overlapped” Motion Compensation

- (Gradient Method: Optical Flow)
- Block Motion Compensation
  - 16x16 macroblock (until MPEG-4)
  - 4x4 ~ 16x16 macroblocks (H.264/AVC)
- Overlapped Motion Compensation
  - 32x32 macroblock
  - (Motion estimation may use 16x16 macroblock)

J.Katto, J.Ohki, S.Nogaki and M.Ohta: "Wavelet Codec with Overlapped Motion Compensation for Very Low Bit Rate Environment", IEEE Trans. on Circuit & Systems for Video Technology, June.1994.

J.Katto and M.Ohta: "An Analytical Framework for Overlapped Motion Compensation", IEEE ICASSP'95, May.1995.

Overlapped Motion Compensation



- Block MC:
  - Copy a macroblock in the previous frame to build a prediction frame.

$$\hat{z}(x,y) = z(x - d_x, y - d_y)$$

- Overlapped MC:
  - Overlap and add multiple macroblocks in the previous frame with weights.

$$\hat{z}(x,y) = \sum_i \sum_{m,n} w_i(m,n) z_i(x - d_{i,x}, y - d_{i,y})$$

M.Ohta and S.Nogaki: “Hybrid Picture Coding with Wavelet Transform and Overlapped Motion Compensated Interframe Prediction Coding", IEEE Trans. SP, Dec.1993.

H.Watanabe and S.Singhal: “Windowed Motion Compensation", SPIE VCIP-91, Nov.1991.

Problems (at that time)

- What is the optimum weighting coefficient for overlapped motion compensation ?
  - “Raised cosine” and “trapezoid” window functions were heuristically tried.
  - Only one work (below) exists.
- Can we theoretically compare overlapped motion compensation with fractional-pel (e.g. half-pel) motion compensation ?
  - At that time, MPEG-1 adopted half-pel motion compensation.
  - Later, overlapped motion compensation was adopted by H.263.

M.T.Orchard and G.J.Sullivan: “Overlapped Block Motion Compensation: An Estimation-Theoretic Approach”, IEEE Trans. IP, Sep.1994.

Problem Formulation

minimize $E\left[(z-\hat{z})^2\right]$

under
$$\begin{cases} \hat{z} = \sum_{k=0}^{N-1} w_k z_k \\ \sum_{k=0}^{N-1} w_k = 1 \end{cases}$$

N: # of macroblocks to be overlapped

z_k : k-th mackroblock

Wiener-Hopf equation

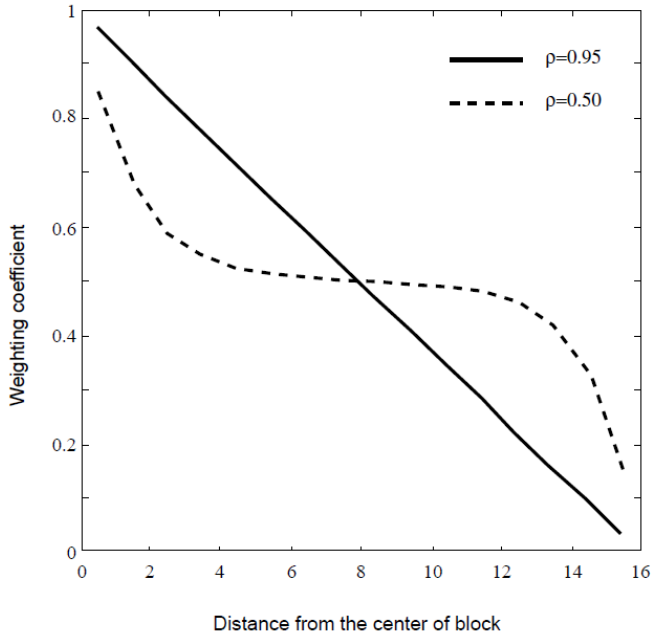
$$\mathbf{w}_{opt} = A^{-1}\mathbf{b}$$

A : correlation matrix

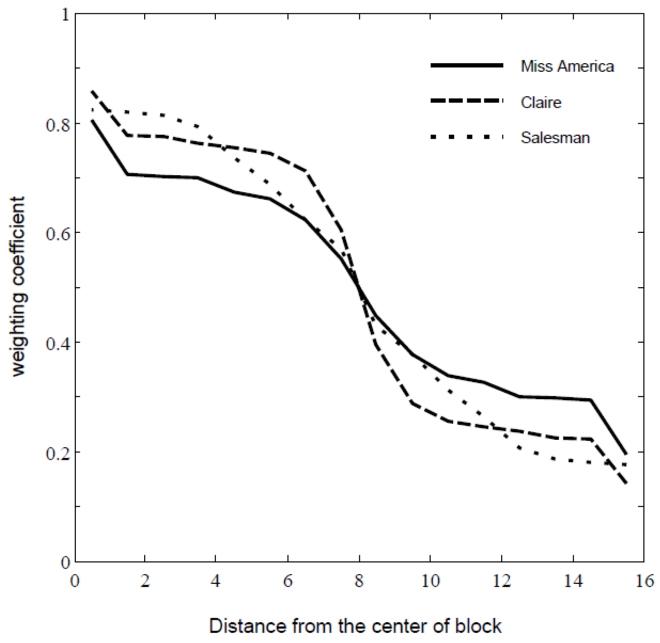
b : correlation vector

Optimum Windows

Theoretical: AR(1) model



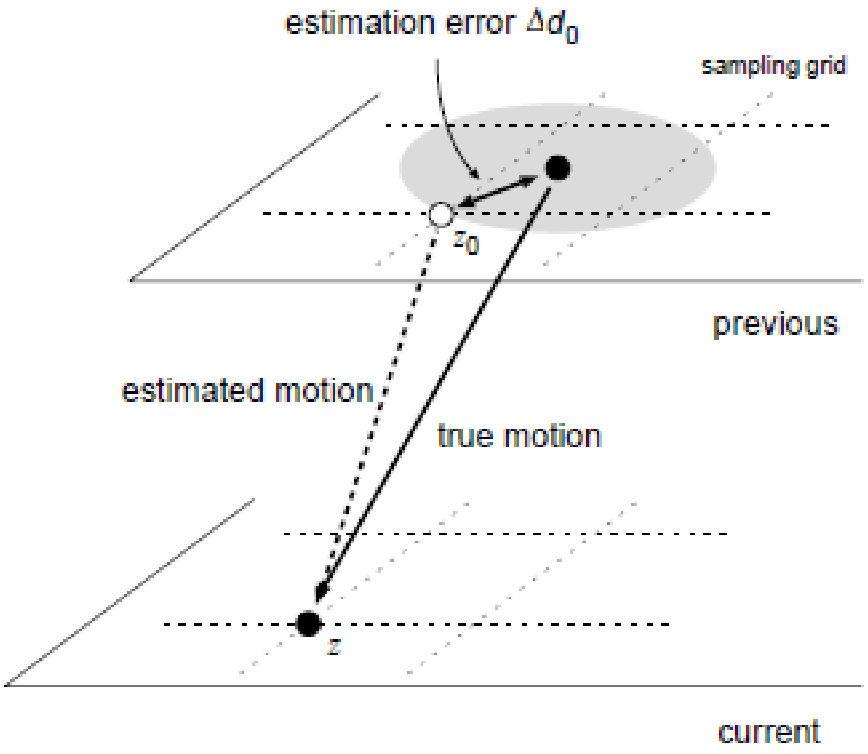
Experimental



Simulations

	Trapezoid	Raised Cosine	Proposed
Salesman	+0.90	+0.90	+0.92
Blue Jacket	+0.42	+0.41	+0.46

Motion Model (1)



Assumptions:

- corresponding pixels (between current and previous) do not change from frame to frame (though previous pixel may not be on sampling grid)

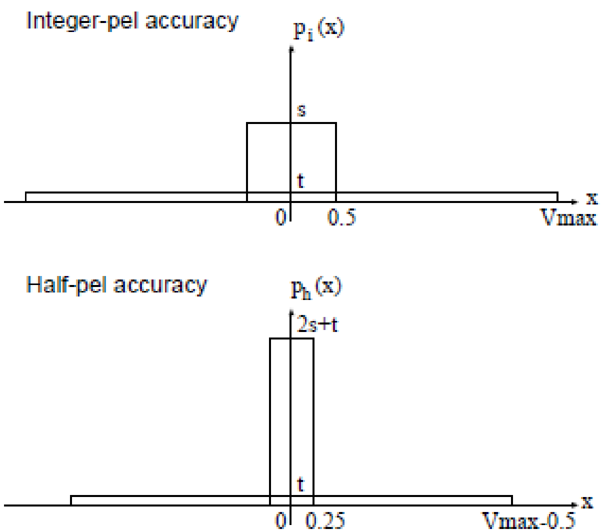
$$z(x,y) = z_0(x - d_{0,x} - \Delta d_{0,x}, y - d_{0,y} - \Delta d_{0,y})$$

- spatial correlation between pixels decreases exponentially

$$\frac{E\left[z_0(x,y) z_0(x - \Delta d_{0,x}, y - \Delta d_{0,y})\right]}{E\left[z_0(x,y)^2\right]} = \rho^{E[|\Delta d_0|]}$$

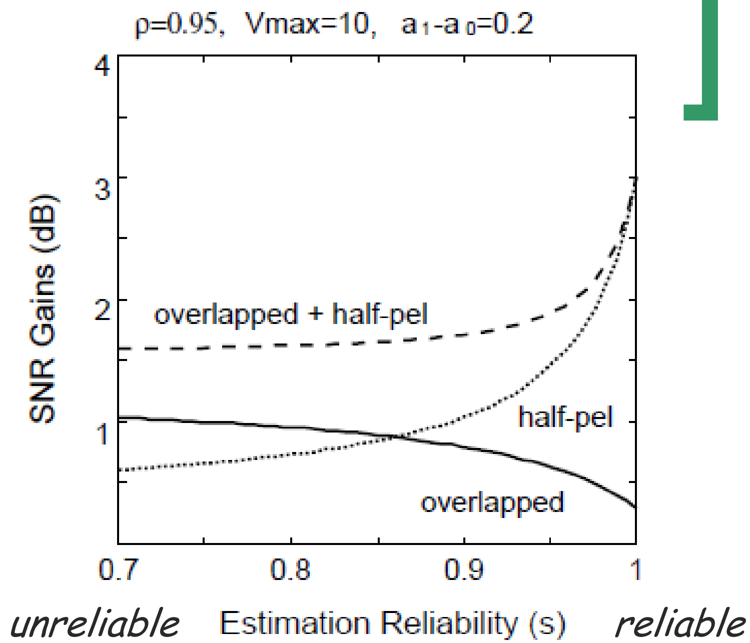
Motion Model (2)

Probability density function of motion estimation errors



$$a_{0,i} = E\left[\left|\Delta d_{0,i}\right|\right] = \frac{1}{4} + \frac{1-s}{2} \cdot V_{\max}$$

$$a_{0,h} = E\left[\left|\Delta d_{0,h}\right|\right] = \frac{1}{8} + \frac{1-s}{2} \cdot \left(V_{\max} - \frac{3}{4}\right)$$



overlapping effect

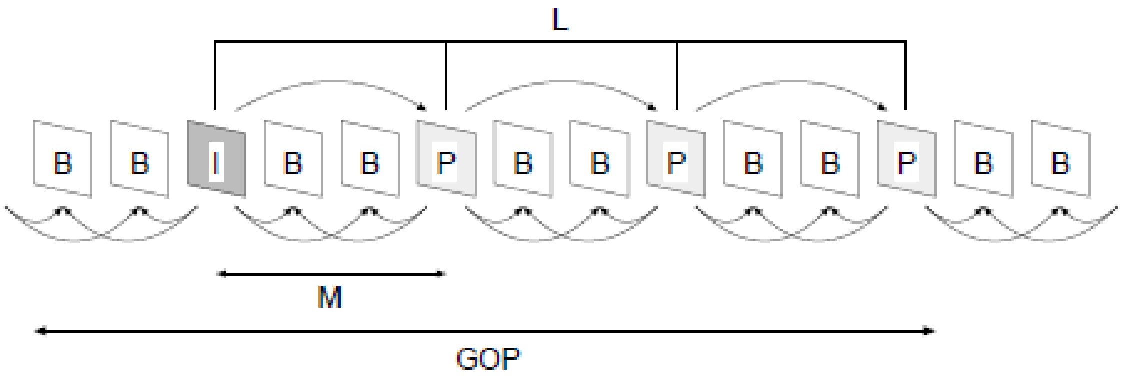
	$E\left[(z-z_0)^2\right] / E\left[z^2\right]$ **(prediction error reduction)**
Interger-pel	$1-\rho^{a_{0,i}}$
Half-pel	$1-\rho^{a_{0,h}}$
Overlapped	$1-\rho^{a_{0,i}} - \frac{1}{1-\rho^{(a_{0,i}+a_{1,i})/2}}\left[\frac{1-\rho^{(a_{0,i}+a_{1,i})/2}-\rho^{a_{0,i}}+\rho^{a_{1,i}}}{2}\right]^2$
Half-pel + Overlapped	$1-\rho^{a_{0,h}} - \frac{1}{1-\rho^{(a_{0,h}+a_{1,h})/2}}\left[\frac{1-\rho^{(a_{0,h}+a_{1,h})/2}-\rho^{a_{0,h}}+\rho^{a_{1,h}}}{2}\right]^2$

Results

- Overlapped motion compensation was originally proposed to reduce blocking artifacts in block motion compensation.
- Proposed theory validates the prediction error reduction effect of overlapped motion compensation.
- The theory proves that overlapped motion compensation performs well when motion estimation does not work well, though half-pel motion compensation works well when motion estimation works well. They can be combined and can help each other.

3. “Overlapped” Interframe Prediction

- IPB prediction structure
  - I-picture: Intraframe coding
  - P-picture: Interframe coding
  - B-picture: Bi-directional interframe coding
- Adopted in MPEG-1 standard



J.Katto and M.Ohta: "Mathematical Analysis of MPEG Compression Capability and Its Application to Rate Control", IEEE ICIP'95 Oct.1995.

Problems (at that time)

- Why does the IPB prediction structure work better than the IP prediction structure ?
  - MPEG-1 was indeed better than H.261. One reason is half-pel motion compensation, and another is IPB prediction. But, why ?

- TM5 (famous rate control for MPEG-2 standard at that time) is the best rate control method ?
  - Theory on IPB prediction might provide better rate control strategy.

Coding Gain of IPB Prediction (1)

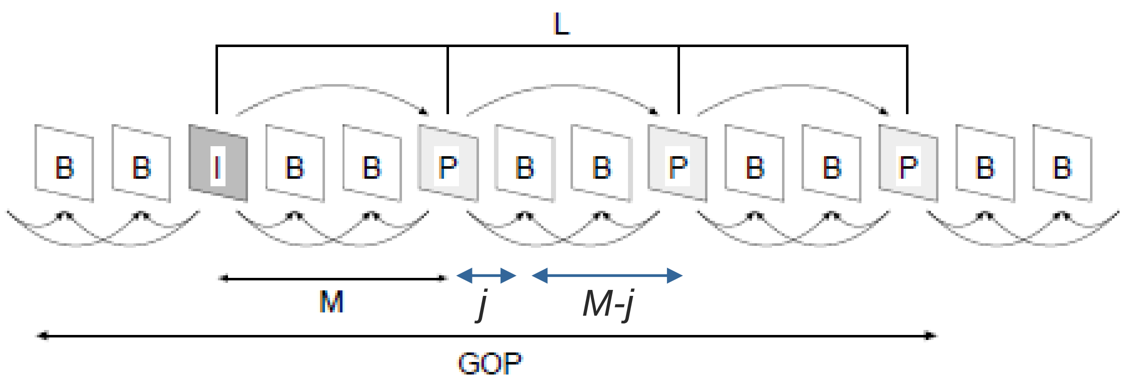
- Prediction error variances
  - P-picture

$$\sigma_P^2 = 2 \cdot \left(1 - cor(M)\right) \cdot \sigma_x^2$$

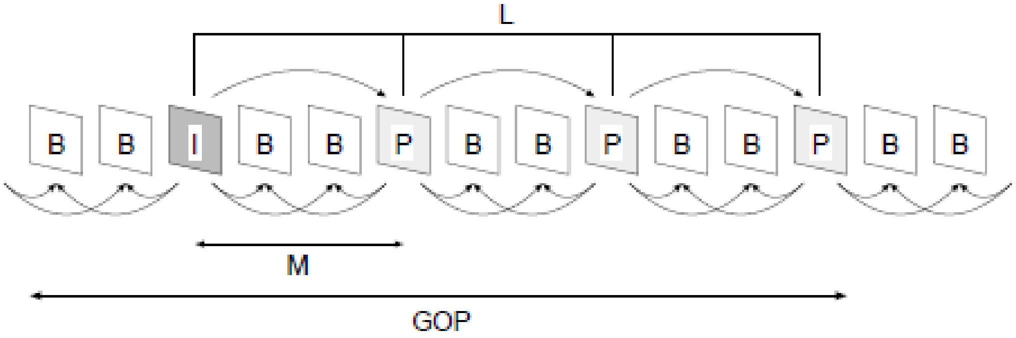
cor(*k*) : correlation between k-frame apart frames

  - B-picture

$$\sigma_B^2 = \left(\frac{3}{2} + \frac{1}{2} \cdot cor(M) - cor(j) - cor(M-j)\right) \cdot \sigma_x^2$$



Coding Gain of IPB Prediction (2)



IPB prediction structure

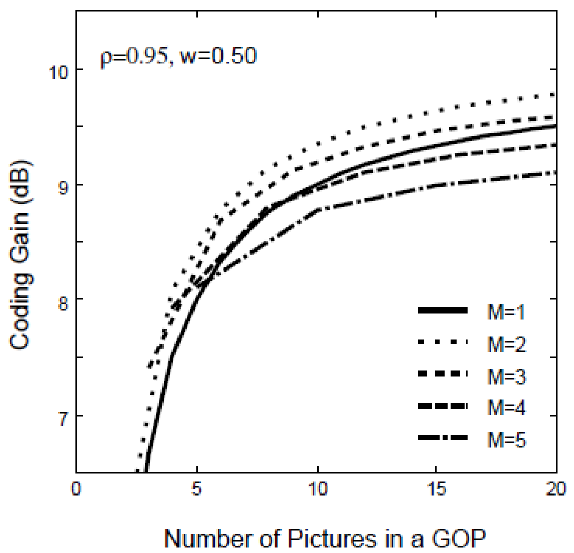
optimum bit allocation problem

$$G_{IPB} = \frac{1}{P(M)^{\frac{L-1}{LM}} \cdot \left[\frac{B(M)}{M-1}\right]^{\frac{M-1}{M}}}$$

$$cor(k) = \rho^{k^w}$$

$$P(M) = \sigma_P^2 / \sigma_x^2 = 2 \cdot (1 - cor(M))$$

$$B(M) = \sigma_B^2 / \sigma_x^2 = \left(\frac{3}{2} + \frac{1}{2} \cdot cor(M) - cor(j) - cor(M-j)\right)$$



Rate Control for IPB Prediction (1)

- Heuristics

$$\log R = a \cdot \log Q + b$$

R : rate

Q : quantization step size

a, b : constants

$$Q \cdot R^{\alpha} = X$$

Rate-distortion theory: $\sigma_q^2 = \varepsilon^2 \cdot 2^{-2R} \sigma_x^2$

→ difficult to apply to rate control directly ...

Rate Control for IPB Prediction (2)

- Optimum target bit assignment

  minimize average quantization step size

$$\frac{N_I \cdot Q_I^m + N_P \cdot Q_P^m + N_B \cdot Q_B^m}{N_I + N_P + N_B}$$

  under rate constraint

$$N_I \cdot R_I + N_P \cdot R_P + N_B \cdot R_B = const$$

N_I, N_P, N_B : # of pictures in GOP

Rate Control for IPB Prediction (3)

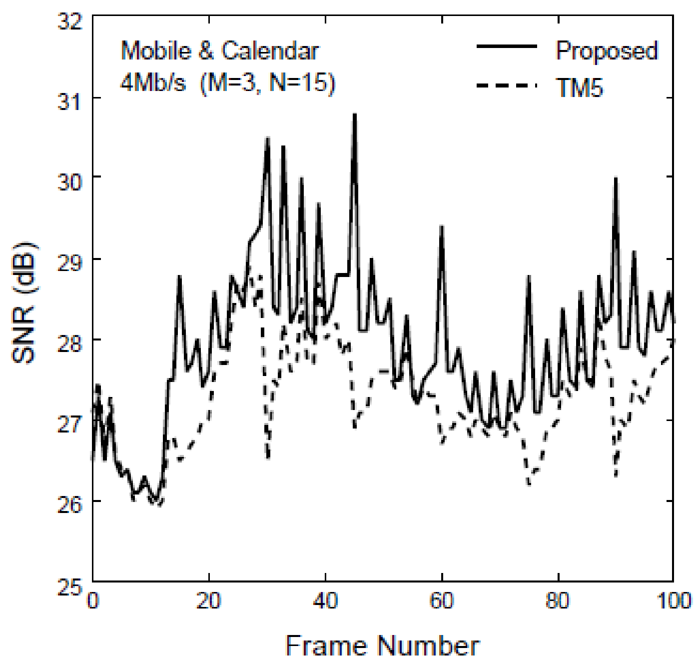
- Solution

$$\begin{cases} R_I = \dfrac{R}{N_I + N_P\left(\dfrac{X_P}{X_I}\right)^{\frac{m}{1+m\alpha}} + N_B\left(\dfrac{X_B}{X_I}\right)^{\frac{m}{1+m\alpha}}} \\ R_P = \dfrac{R}{N_P + N_B\left(\dfrac{X_B}{X_P}\right)^{\frac{m}{1+m\alpha}}} \\ R_B = \dfrac{R}{N_B + N_P\left(\dfrac{X_P}{X_B}\right)^{\frac{m}{1+m\alpha}}} \end{cases}$$

observable

$$\begin{cases} Q_I \cdot R_I^{\ \alpha} = X_I \\ Q_P \cdot R_P^{\ \alpha} = X_P \\ Q_B \cdot R_B^{\ \alpha} = X_B \end{cases}$$

Simulation



Summary of Formulations

	1st (Coding gain)	2nd (Target bit assignment)
R-D function	$\sigma_q^2 = \epsilon^2 2^{-2R} \sigma_x^2$	$Q \cdot R^{\alpha} = X$
Cost function	$\dfrac{\sigma_{r,I}^2 + (L-1)\cdot\sigma_{r,P}^2 + L(M-1)\cdot\sigma_{r,B}^2}{LM}$	$\dfrac{N_I \cdot Q_I^m + N_P \cdot Q_P^m + N_B \cdot Q_B^m}{N_I + N_P + N_B}$
Constraint	$R_I + (L-1)R_P + L(M-1)R_B = LM \cdot R$	$N_I R_I + N_P R_P + N_B R_B = R$
Coding gain	$G_{IPB} = \dfrac{1}{P(M)^{\frac{L-1}{LM}} \cdot \left[\frac{B(M)}{M-1}\right]^{\frac{M-1}{M}}}$	*(not given in an explicit form)*
Bit allocation	$R_I = R - \dfrac{L-1}{2LM}\log_2 P(M) - \dfrac{M-1}{2M}\log_2 B(M)$ $R_P = R + \dfrac{L(M-1)+1}{2LM}\log_2 P(M) - \dfrac{M-1}{2M}\log_2 B(M)$ $R_B = R - \dfrac{L-1}{2LM}\log_2 P(M) + \dfrac{M+1}{2M}\log_2 B(M)$ *(fragile for practical use)*	$R_I = \dfrac{R}{1 + N_P\left(\frac{X_P}{X_I}\right)^{\frac{m}{1+m\alpha}} + N_B\left(\frac{X_B}{X_I}\right)^{\frac{m}{1+m\alpha}}}$ $R_P = \dfrac{R}{N_P + N_B\left(\frac{X_B}{X_P}\right)^{\frac{m}{1+m\alpha}}}$ $R_B = \dfrac{R}{N_B + N_P\left(\frac{X_P}{X_B}\right)^{\frac{m}{1+m\alpha}}}$ *(including TM5)*

Results

- Coding gain of IPB prediction quantitatively proves its efficiency against IP prediction.
  - (However, though omitted, coding gain based rate control is not practical)
- Rate control for IPB prediction based on heuristics between rate and quantization step size provides smarter bit assignment than TM5.
  - TM5 is included as a special case of the proposal.

4. Denoising by “Overlapping”

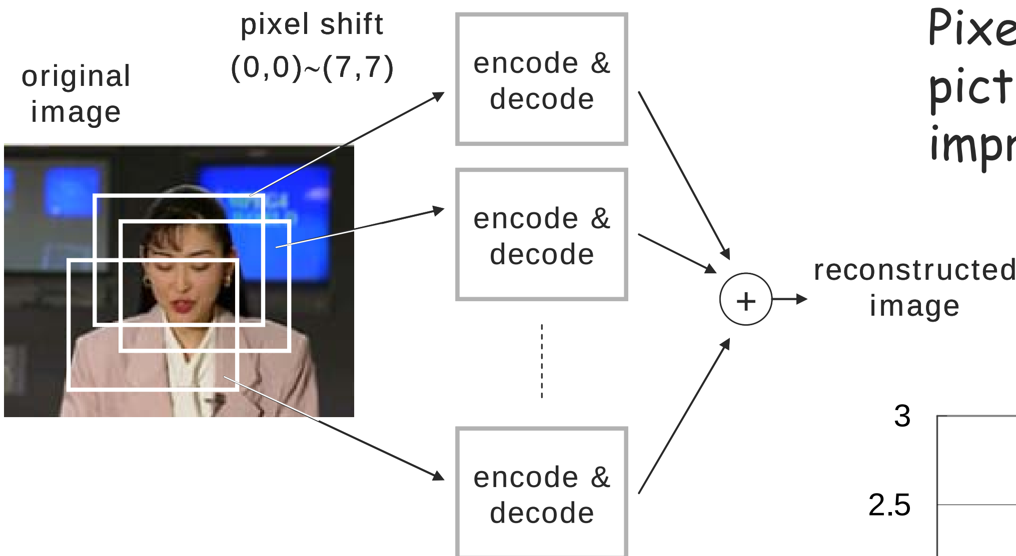
- Reconsiderations on video coding
  - Complexity shift to decoders
    - Distributed video coding
      - Application to wireless sensor networks
    - Multiple description coding
      - Application to heterogeneous networks
    - Super-resolution
      - SDTV/HDTV conversion at a decoder
  - Image restoration
    - Denoising
      - Reduction of blocking artifacts and flicker noises
      - Deblocking filters, motion compensated temporal filtering, re-application of JPEG, and so on ...

J.Katto, J.Suzuki, S.Itagaki, S.Sakaida and K.Iguchi: “Denoising Intra-Coded Moving Pictures using Motion Estimation and Pixel Shift”, IEEE ICASSP 2008, Apr.2008..

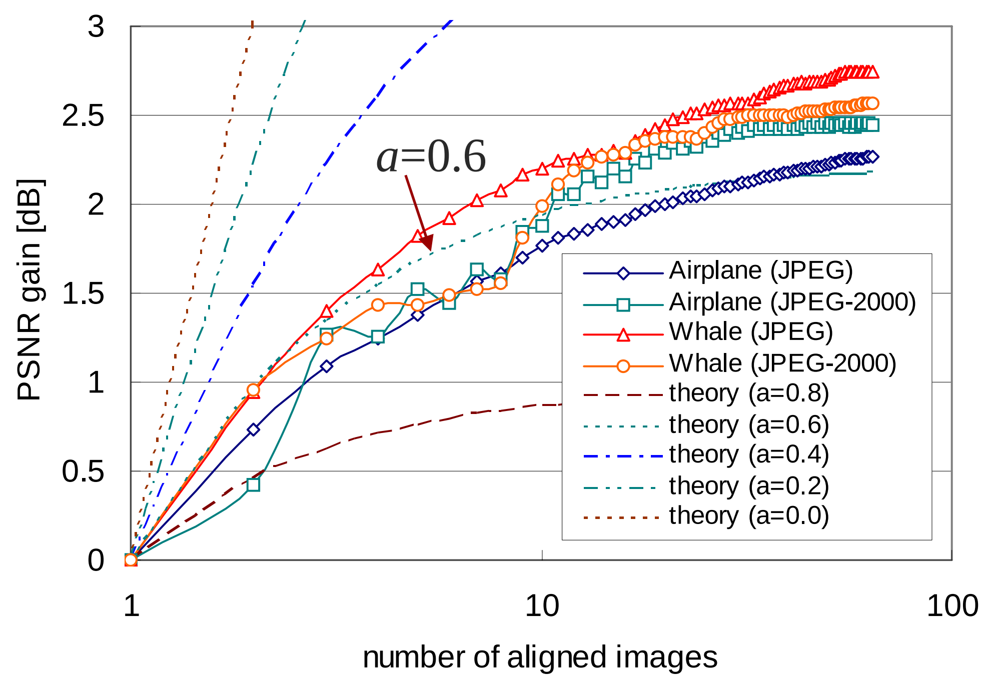
Problems

- Quality enhancement of compressed video streams becomes important
  - Enormously rapid increase of compressed visual contents over storages and networks
  - People prefer better quality contents if they don't pay much money for them
  - Decoders (PCs) become powerful
  - Smart quality improvement of old visual contents is impressive
  - Smart denoising might contribute to smarter compression in future

Auxiliary Experiment



Pixel shift, JPEG compression and picture "overlapping" brings PSNR improvement.



Formulation (1)

- Statistical reduction of quantization errors

minimize $E\left[(x-\tilde{x})^2\right] = E\left[\left(\sum_{k=1}^{K} w(k)q(k)\right)^2\right]$

where $\tilde{x} = \sum_{k=1}^{K} w(k)\hat{x}(k) = x + \sum_{k=1}^{K} w(k)q(k)$

under constraint $\sum_{k=1}^{K} w(k) = 1$

K : # of overlapped pictures

$q(k)$: k-th quantization error

$w(k)$: k-th weight

Formulation (2)

- Solution
  - variant of Wiener-Hopf equation

$$\mathbf{w}_{opt} = C \cdot R^{-1}\mathbf{u}$$

where $\mathbf{u} = (1,1,\cdots,1)^t$, $C = \dfrac{1}{\mathrm{sum}(R^{-1}\mathbf{u})}$ and

$$R = E\left[\mathbf{q}\cdot\mathbf{q}^t\right] = \begin{bmatrix} E[q(1)^2] & E[q(1)q(2)] & \cdots & E[q(1)q(K)] \\ E[q(2)q(1)] & E[q(2)^2] & & E[q(2)q(K)] \\ \vdots & & \ddots & \\ E[q(K)q(1)] & E[q(K)q(2)] & & E[q(K)^2] \end{bmatrix}$$

correlation matrix of quantization errors

Formulation (3)

- Special cases
  - when independent, i.e. $E[q(k)q(l)]=0 \quad (k \neq l)$

$$\begin{cases} w_{opt}(k) = \dfrac{1}{\sigma_{q(k)}^2 \cdot \sum_{l=1}^{K} \dfrac{1}{\sigma_{q(l)}^2}} \\ \sigma_{q,\min}^2 = \dfrac{1}{\sum_{k=1}^{K} \dfrac{1}{\sigma_{q(k)}^2}} \end{cases}$$

$\Rightarrow$

esp. when $\sigma_{q(k)}^2 = \sigma_q^2$

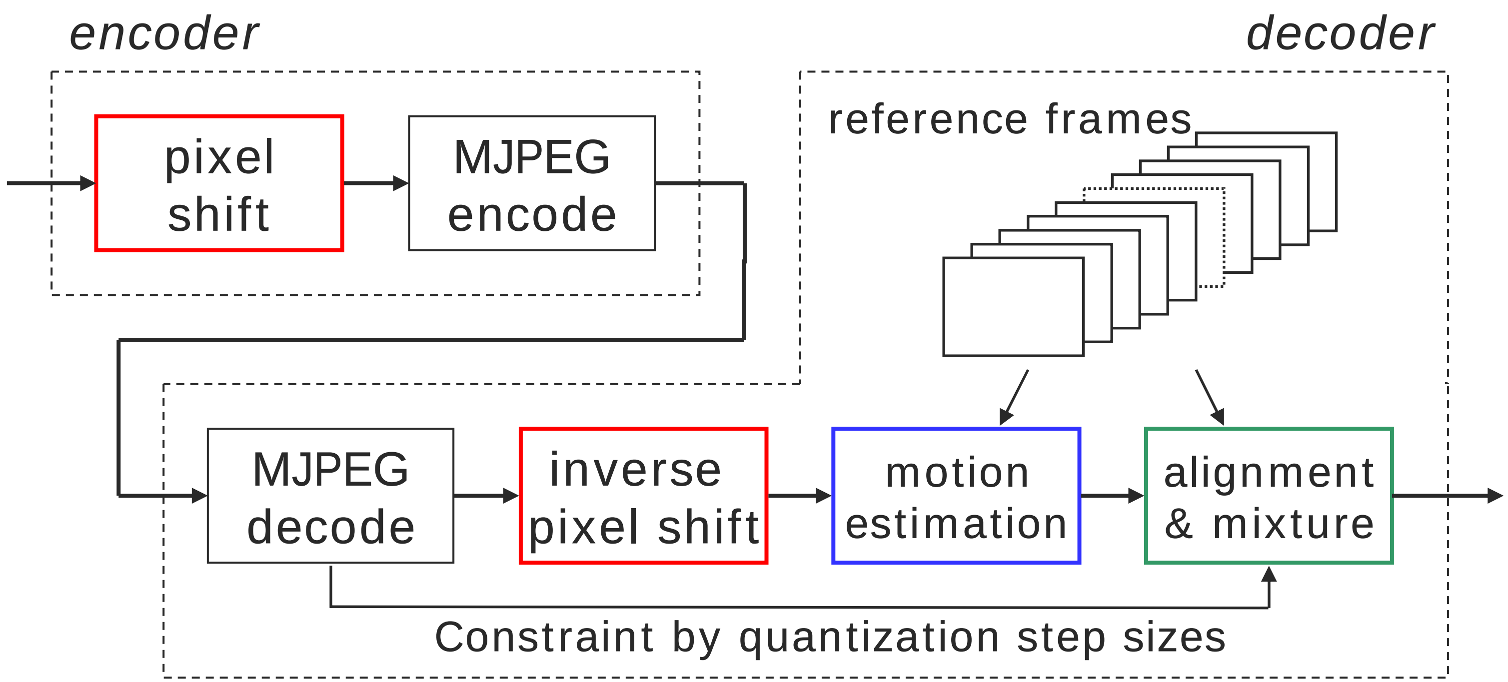
$$\begin{cases} w_{opt}(k) = \dfrac{1}{K} \\ \sigma_{q,\min}^2 = \dfrac{\sigma_q^2}{K} \end{cases}$$

Gain: $10\log_{10} K\,(dB)$

  - when completely dependent, i.e. $E[q(k)q(l)]=\sigma_q^2$
    - no gain (corresponding to no pixel shift)

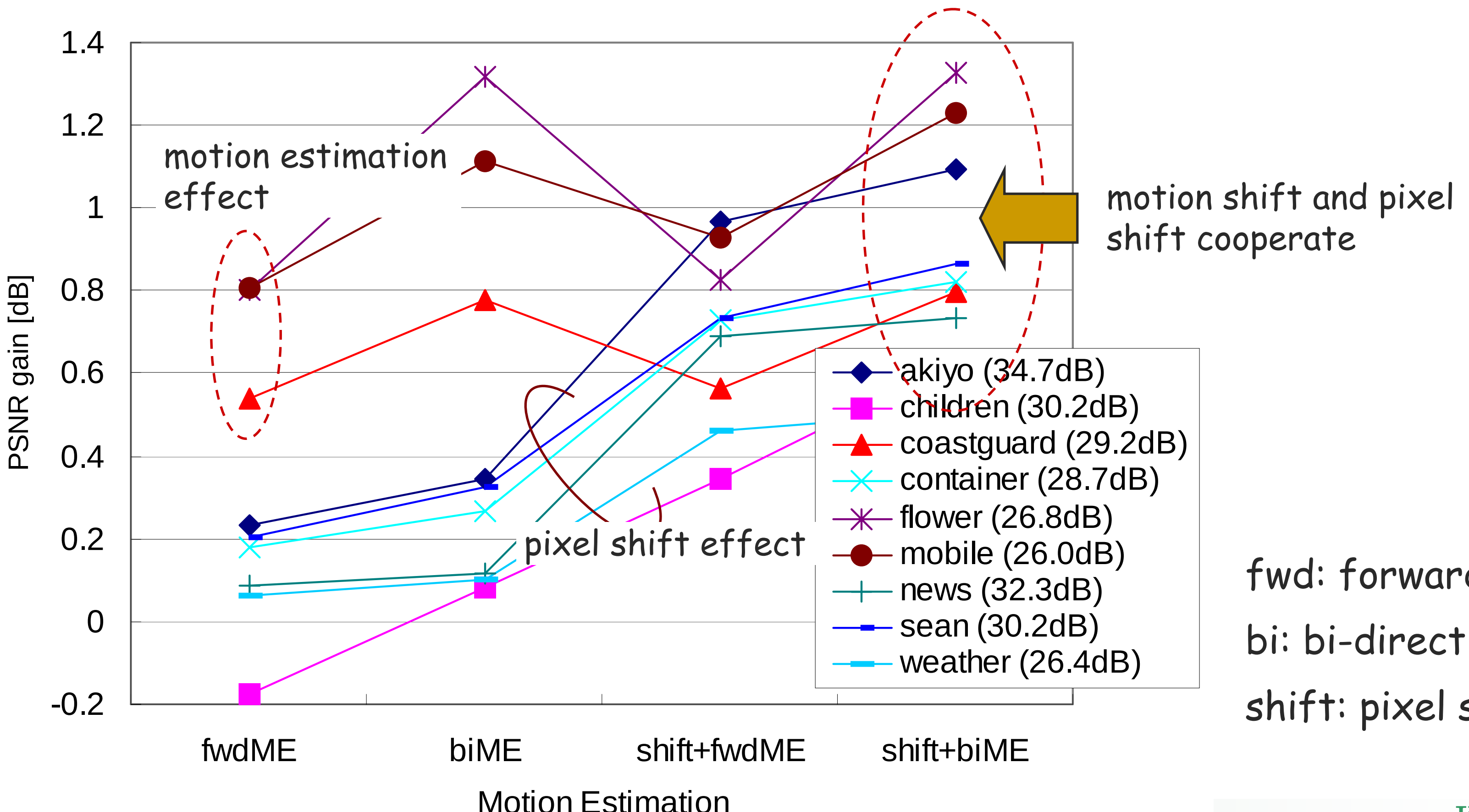
Motion JPEG Extension

- Motion estimation at a decoder
- Intentional pixel shift at an encoder and decoder (though breaking compatibility)



Experiments (1)

- Motion estimation works well for moving regions
- Intentional pixel shift works well for static regions



fwd: forward

bi: bi-directional

shift: pixel shift

Experiments (2)

- Subjective quality comparison

Motion JPEG

Proposed method

impressive reduction of blocking artifacts
and flicker noises

Results

- Overlapping similar regions with different quantization noises contributes to <u>statistical reduction of quantization noises.</u>
- <u>Subjective quality improvement</u> in blocking artifacts and flicker noises is also observed.

- Extension to interframe coding is ongoing.
- Extension to super-resolution is also ongoing.

Conclusions

- " Overlapping" had brought many performance improvements in video compression.
- "Optimum bit allocation" is one of important key formulations.
- "Wiener-Hopf equation" is another important key formulation.