Lecture notes for SSY150: Multimedia and video communications

Video compression (I, II)

(for lecture 5,7)

Irene Y.H. Gu, Dept. of Signals and Systems,
Chalmers Univ. of Technology, Sweden
March 31 + April 28, 2008

Contents

1. Fundamental of video compression: MC
2. Video coding modes
3. Block matching for motion compensation (MC)
4. I,P,B image frames, and uni-directional versus bi-directional MC
5. A typical video codec
6. Tradeoffs: select inter and intra modes
7. Video coding standards
8. Why video communications: brief introduction+ examples
9. Other applications of object-based video processing
10. Lab.3 descriptions and demonstration
11. References

1. Fundamental of video compression: how compression is achieved?

High redundancy in consecutive image frames:
only a small part has changed due to the motion

Only encode and send change parts to the channel
(using motion compensation techniques)

2. Video Coding modes

Switch between 2 modes:

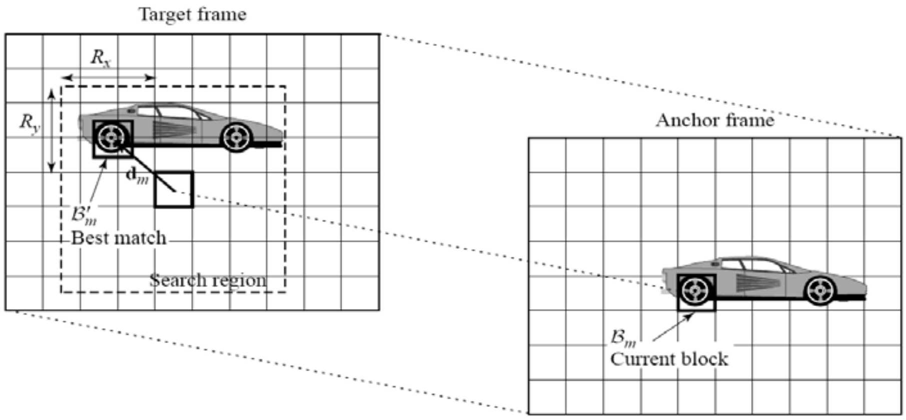
- **Intra** mode (= 2D still image coding):
  + (DCT) encoding of an image frame is independent of other frames.
  + Used in: 1st frame, refreshing frames (Intra or I frames)

- **Inter** mode:
  Block matching for motion compensation (MC), the simplest method

  Apply MC to P (predicted) frames and B (bidirectional predicted) frames

Inter mode: Block matching for MC

Find the "best" matching block within the searching region



(Courtesy from Dr. Y.Wang)

3. Block matching for MC

The block-matching scheme for Motion Compensation:

$$I_t(x,y) = I_{t-k}(x-d_x, y-d_y)$$

need to find the 'best' displacement vector (d_x, d_y) for the block

Frequently used objective criteria:

(a) **Mean Square Error (MSE)** criterion

$$MSE(d_x,d_y) = \frac{1}{M_x M_y} \sum_{(x,y)\in B_i} \left[I_t(x,y) - I_{t-k}(x-d_x, y-d_y) \right]^2$$

$$(d_x^*, d_y^*) = \arg \min_{(d_x,d_y)} MSE(d_x, d_y)$$

Block matching for MC (cont'd)

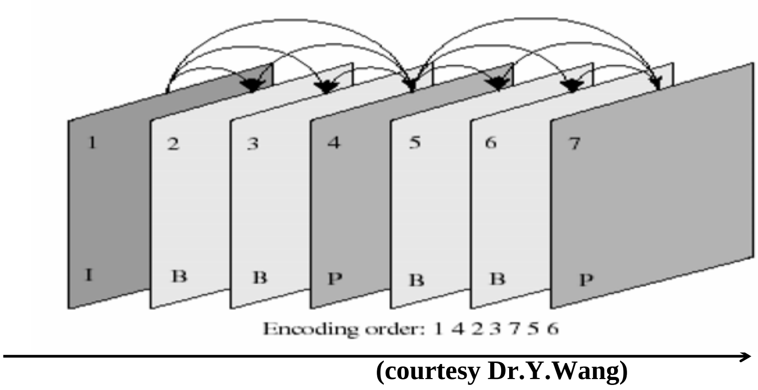
(b) **Mean Absolute Error (MAE)** criterion

$$MAE(d_x, d_y) = \frac{1}{M_x M_y} \sum_{(x,y)\in B_i} \left| I_t(x,y) - I_{t-k}(x-d_x, y-d_y) \right|$$

$$(d_x^*, d_y^*) = \arg \min_{(d_x, d_y)} MAE(d_x, d_y)$$

4. I,P,B image frames, uni-directional vs. bi-directional MC

I,B,P frames in video coding:



(courtesy Dr.Y.Wang)

MC can be done by forward and backward prediction using one or more frames

Unidirectional and bidirectional MC:

Unidirectional MC

MC is done by using a unidirectional predicted block from one previous frame:

$$I_t(x, y) = I_{t-k}(x - d_x, y - d_y)$$

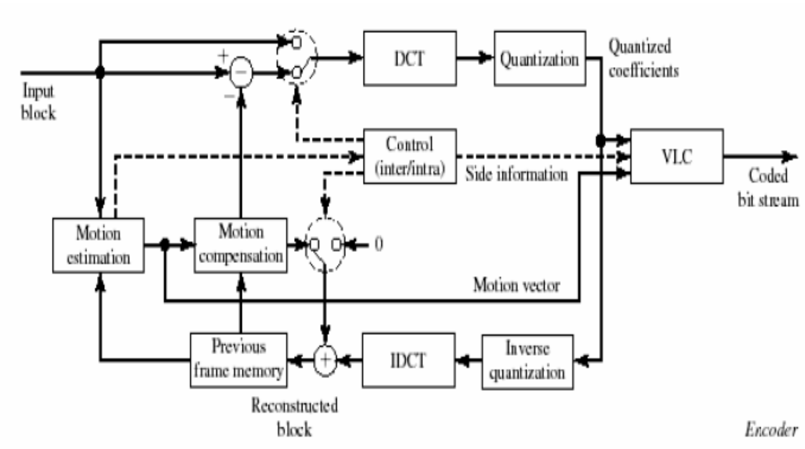
Bidirectional MC

MC is done by weighted sum of forward/backward predicted blocks from one/several previous frames and one/several future frames:

$$I_t(x, y) = \sum_k w_k I_{t-k}(x - d_{x_k}, y - d_{y_k})$$

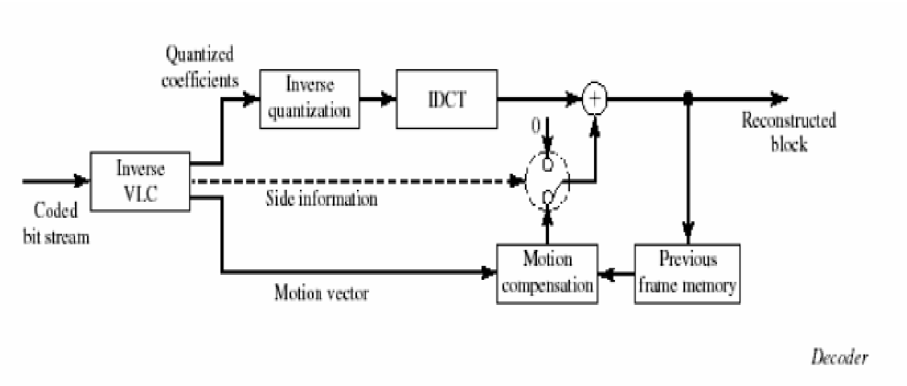
5. A typical video codec

A typical encoder using block matching MC:



(courtesy Dr. Y. Wang)

A typical decoder with block matching MC



(courtesy Dr. Y. Wang)

6 Tradeoffs: select inter and intra modes

Intra coding: apply to first frame and refreshing (I) frames;
low compression ratio ☹
purely dependent on the current frame ☺

Inter coding: apply to P frames and B frames,
encode motion vectors
(and residuals in the DCT domain);
high compression ratio ☺
motion errors may propagate through frames ☹

Inter and intra mode selection:
tradeoff between rate and performance

7. Video coding standards (brief)

H.26x (ITU-T video coding standards):
(**ITU-T**: the ITU Telecommunication Standardization Sector)

H261, H263, H.264/AVC (advanced video coding standard)

MPEG-x: (MPEG audio and video coding & compression standards)
(**MPEG**: Moving Picture Experts Group)

MPEG-1, 2, 4, 7 (for multimedia applications,
especially object based coding)

MPEG for multimedia applications

- **Content / object-based coding**

  Separate coding of foreground and background parts
  foreground: moving objects
  background: static

- **SPRITE video coding**

  A segmentation-based video coding algorithm

Basic features of video coding standards

	MPEG-1 (1993)	MPEG-2 (1995)	MPEG-4 (2000)	H.261 (1993)	H.263 (1995)	H.264/ MPEG-4 AVC (2002)
Transform	8x8 DCT	8x8 DCT	8x8 DCT	8x8 DCT	8x8 DCT	4x4
MC Block Size	16 x 16	16 x 16 8x16	8 x 8, 16 x 16	16 x 16	8 x 8, 16 x 16	16x16, 16x8, 8x8, 8x4, 4x4
MC Accuracy	_ -pel	_ -pel	_ -pel	1-pel	_ -pel	1/8 -pel
Additional Motion Prediction Modes	- B-Frames	- B-Frames - Interlace	- B-Frames - Interlace - GMC (Global MC) - SPRITE Coding	-	- B-Frames	- B-Frames - Long term frame memory - in-loop deblocking filter - CAVLC/CABAC

MPEG-4: multimedia applications, object-based coding

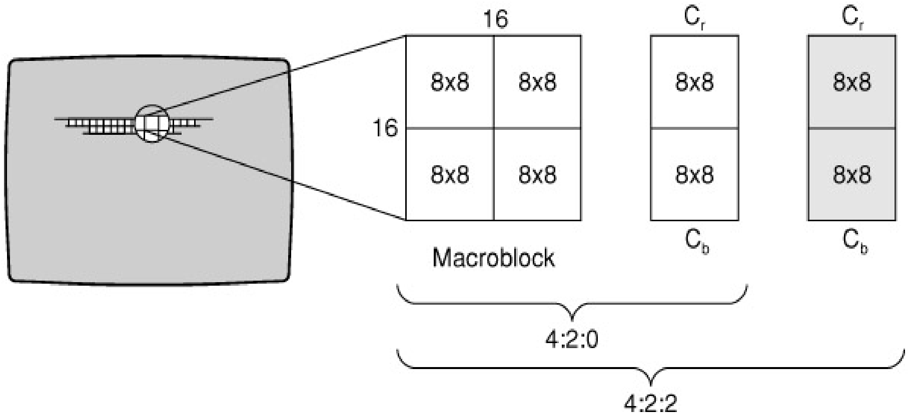
Ref: T.Sikora, proc.IEEE, Vol.93, No.1, 2005

ITU-T Multimedia

Media	**Narrow-band (H.320)**	**lowbitrate (H.324)**	**ISO Ethernet (H.322)**	**Ethernet (H.323)**	**ATM (H.321)**	**High Res. ATM (H.310)**
video	H.261	H.261, H.263	H.261	H.261, H.263	H.261	MPEG-2 H.261
audio	G.711, G.722 G.728	G.723	G.711,G.722 G.723,G.728	G.711,G.722 G.723,G.728 G.729	G.711,G.722 G.728	MPEG-1 MPEG-2
data	T.120	T.120,T.84, T.434, …	T.120	T.120	T.120,H.281 (H.224)	T.120
Multiplex	H.221	H.223	H.221	H.221	H.221	H.222.1 H.221
signaling	H.230 H.242	H.245	H.230, H.242	H.230,H.245 H.225.0	H.230 H.242	H.245
Multi-point	H.243	NA	H.243	NA	H.243	NA
Encryption	H.233 H.234	H233/324, H.234	H.320	TBD	H.233 H.234	NA

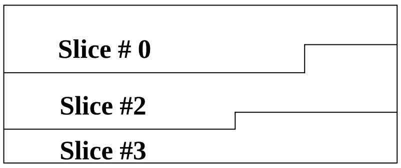
(from wikipedia, the free encyclopedia)

Macro-blocks and blocks in MPEG



Video Compression: H.264/MPEG-4 AVC standard

- **Macro-blocks, slices, slice groups**



divide image into slices

Slice types: I slice / P slice / B slice ...

- **Integer transform** (4x4 blocks, **to replace 8x8 DCT !**)

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

Zigzag and alternative scanning

Two types of DCT and two types of scan pattern:

- **Frame DCT**: divides an MB into 4 blocks for Lum, as usual
- **Field DCT**: reorder pixels in an MB into top and bottom fields.

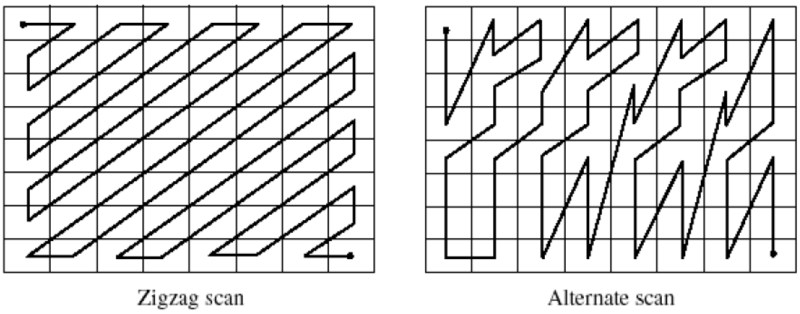
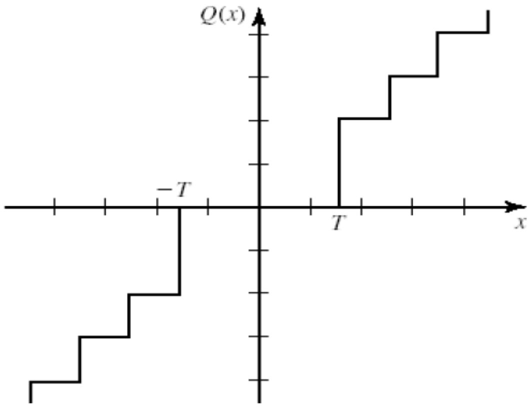


Figure 13.19 The zigzag scan as known from H.261, H.263, and MPEG-1 is augmented by the alternate scan in MPEG-2, in order to code interlaced blocks that have more correlation in the horizontal than in the vertical direction.

(Courtesy Dr. Y.Wang)

Scalar Quantization

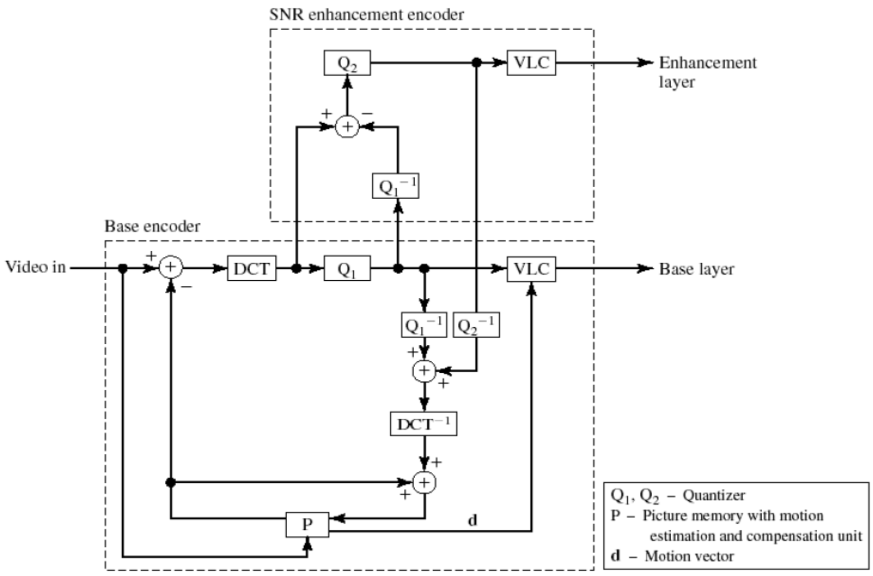


Uniform,
Stepsize=8
with a dead zone

Features in the codec

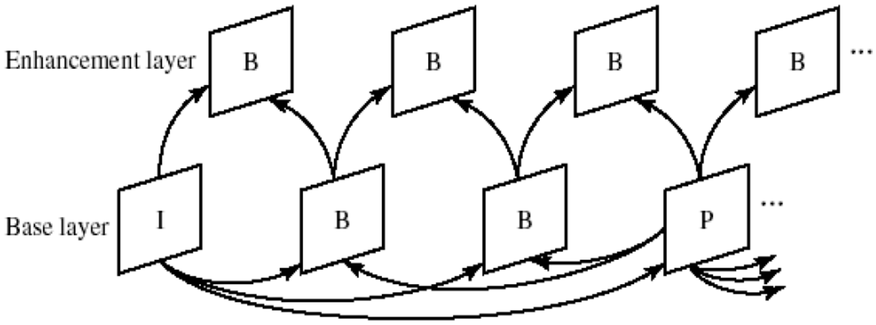
- SNR scalability
- Spatial scalability
- Temporal scalability

SNR (quality) Scalability



(Courtesy Dr. Y.Wang)

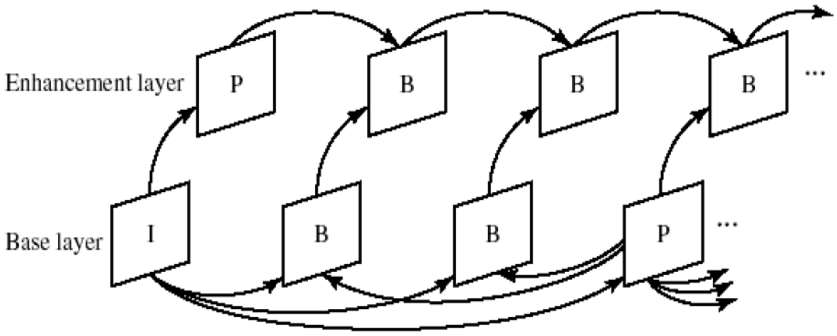
Temporal scalability: option-1



(courtesy Dr. Y.Wang)

- Use the base layer to predict the B images in the enhanced layer.
- Errors in the enhanced layer does not propagation.

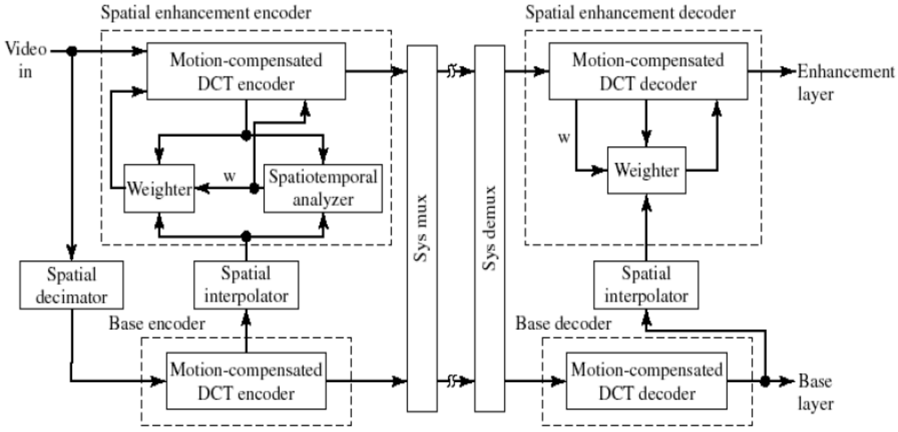
Temporal scalability: option-2



(courtesy Dr. Y.Wang)

Use both the base layer and the enhanced layer for image prediction in the enhanced layer

Spatial scalability



(courtesy Dr. Y.Wang)

Spatial scalability

Spatial scalability

6.5 kbps

133.9 kbps

21.6 kbps

436.3 kbps

(courtesy Dr. Y.Wang)

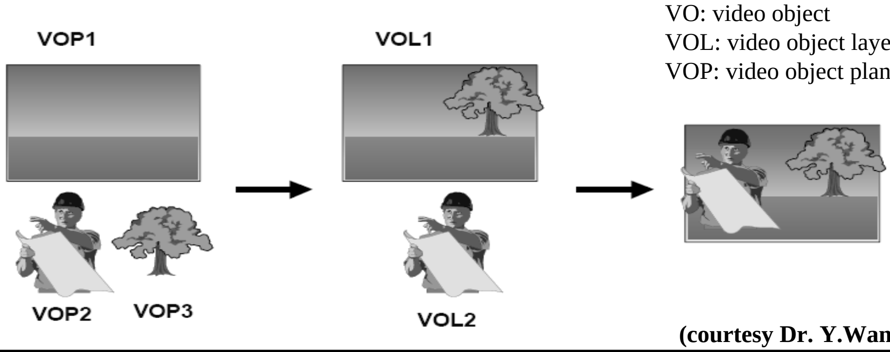
MPEG-4: Object-based coding

- Each image scene consists of multiple objects

  (require object segmentation –which is non-standardized)

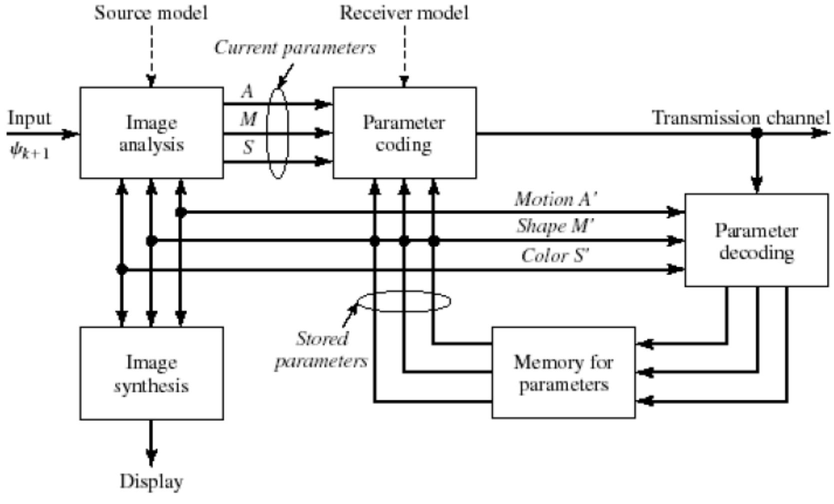
- Each object is specified by its shape, motion, texture (e.g. color)

VO: video object
VOL: video object layer
VOP: video object plane



(courtesy Dr. Y.Wang)

Object-based coding



(Courtesy Dr. Y.Wang)

Some features of MPEG-7

MPEG-7 allows one to find the content one needs!

– Enable multimedia document indexing, browsing, and retrieval

– Define the syntax for the metadata (e.g. index and summary) attached to the document

(Generation of index and summary is not in the standard)

Content description in MPEG-7

– Descriptor (D): describing low-level features

– Description scheme (DS): combining Ds to describe high-level features/structures

– Description definition language (DDL): define how Ds and DSs can be defined/modified

– System tools

Visual Descriptors

Color:

Histogram, dominant color, etc.

Texture:

+ Homogeneity: energy in different orientation and frequency bands (Gabor transform)

+ Coarseness, directionarity, regularity

+ Edge orientation histogram

Motion:

+ Camera motion

+ Motion trajectory of feature points in a non-rigid object

+ Motion parameters of a rigid object

+ Motion activity

Shape:

Boundary-based vs. region-based

Summary

- H.261:
  - First video coding standard, targeted for video conferencing over ISDN
  - Uses block-based hybrid coding framework with integer-pel MC
- H.263:
  - Improved quality at lower bit rate, to enable video conferencing/telephony below 54 bkps (modems or internet access, desktop conferencing)
  - Half-pel MC and other improvement
- MPEG-1 video
  - Video on CD and video on the Internet (good quality at 1.5 mbps)
  - Half-pel MC and bidirectional MC
- MPEG-2 video
  - TV/HDTV/DVD (4-15 mbps)
  - Extended from MPEG-1, considering interlaced video

(Courtesy Dr. Y.Wang)

Summary

- MPEG-4
  - To enable object manipulation and scene composition at the decoder -> interactive TV/virtual reality
  - Object-based video coding: shape coding
  - Coding of synthetic video and audio: animation
- MPEG-7
  - To enable search and browsing of multimedia documents
  - Defines the syntax for describing the structural and conceptual content
- Newer standards
  - H.264: improved coding efficiency (by having more options for optimization)
  - MPEG-21: beyond MPEG-7, considering intellectual property protection, etc.

(Courtesy Dr. Y.Wang)

More information on H.264/MPEG-4 AVC and multimedia standards

Wikipedia, the free encyclopedia on H.264/MPEG-4AVC:

http://en.wikipedia.org/wiki/H.264

Multimedia Standards:

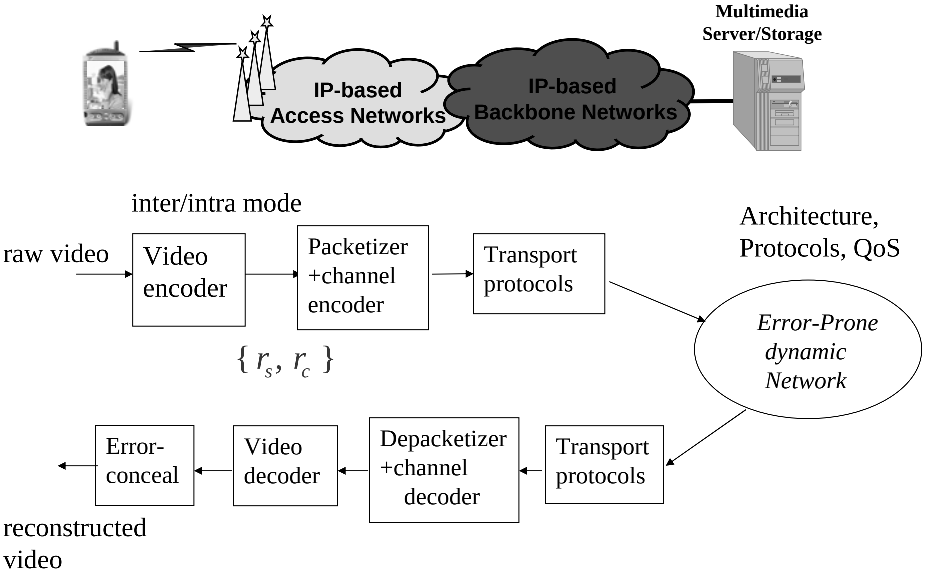
http://www.dtic.mil/ieb_cctwg/contrib-docs/CKing/CK-multi.html

8. Why video communications: brief introduction + examples

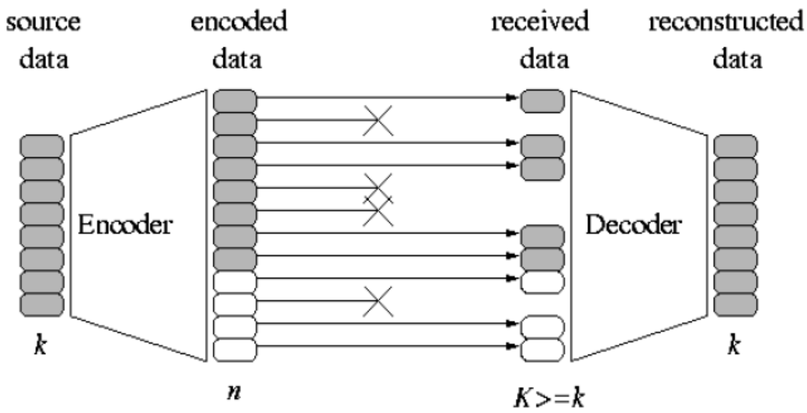
Applications:

- Video phone
- Video conferencing
- Video streaming
- ….

Compressed Video over Communication Networks



Channel Coding



(Courtesy of Dr. Dapeng Wu)

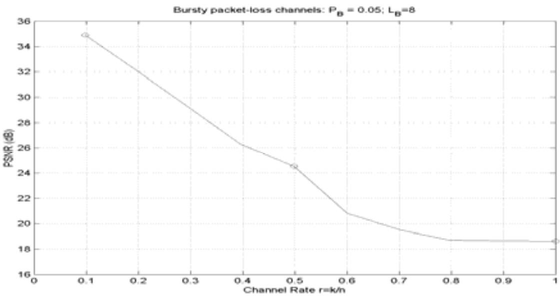
e.g: Reconst. video from bursty packet-loss channels

Test system: H263+, packetizing, RS channel codec, interleaving, bursty packet loss (P=0.05, L=8) network models,

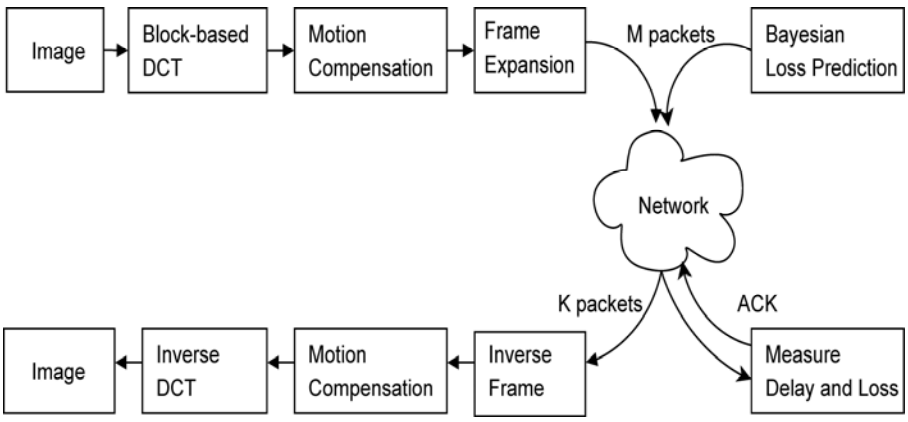
r=k/n=0.1, 30dB

r=k/n=0.5, 24dB

r=k/n=1.0, PSNR= 18dB



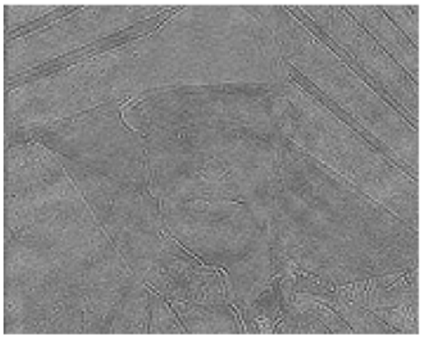
Example: Network-adaptive video coding



A proposed system for compressed video over IP networks

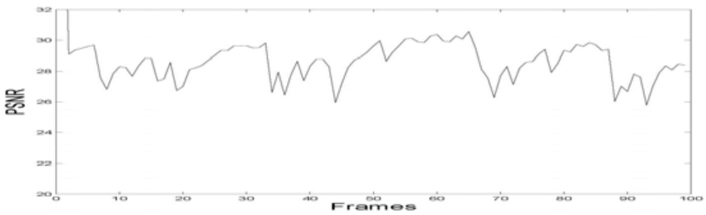
Foreman: frame #60,PSNR=26dB, 10% packet loss, 48bk/s

Reconstructed image



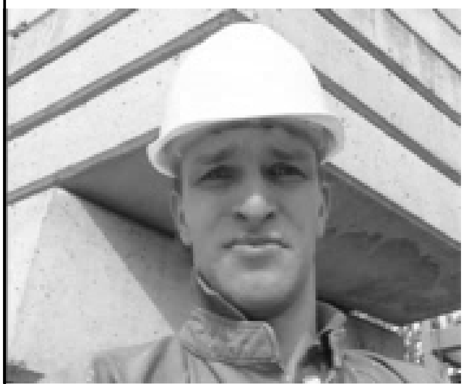
Error image

PSNR

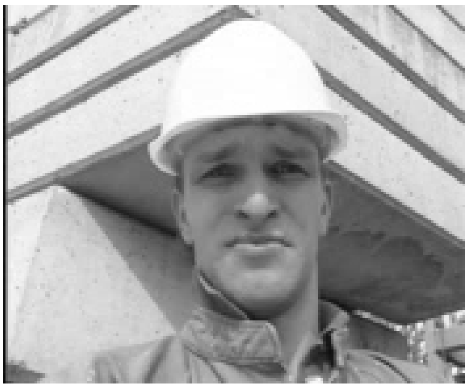


Frame No.

Results



20% losses, 70kbps

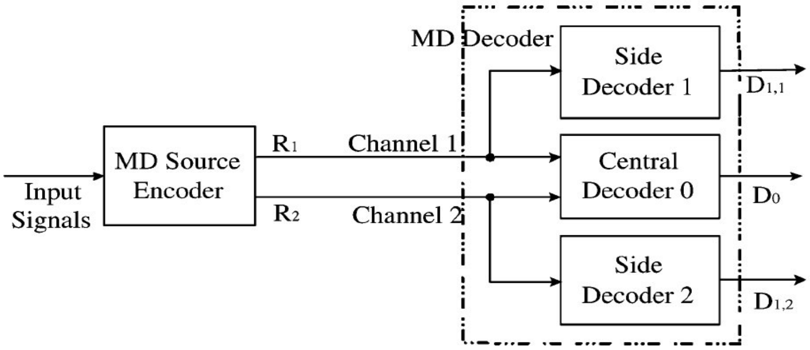


10% losses, 70kbps

(1 image: 99 macro-blocks, 1 packet: 25 macro-blocks, Macro-block size= 16x16 pixels, random packet losses)

Example: multiple description codes (MDC)

MDC is a source coding method for error prone network using: consisting of a central coder and a side coder



(Ref: Yao Wang, Proc. IEEE, vol.93, No.1, 2005).

9. Other applications of object-based video processing

Multimedia applications: tracking moving objects in video

Example: Bayesian framework-based video object tracking

pdf: Mixture of Gaussians
Dynamic training and maintenance of background reference image

e.g: multi-person tracking in crowded environment

Each row shows a few video frames where persons are tracked, indicated by different colored boxes

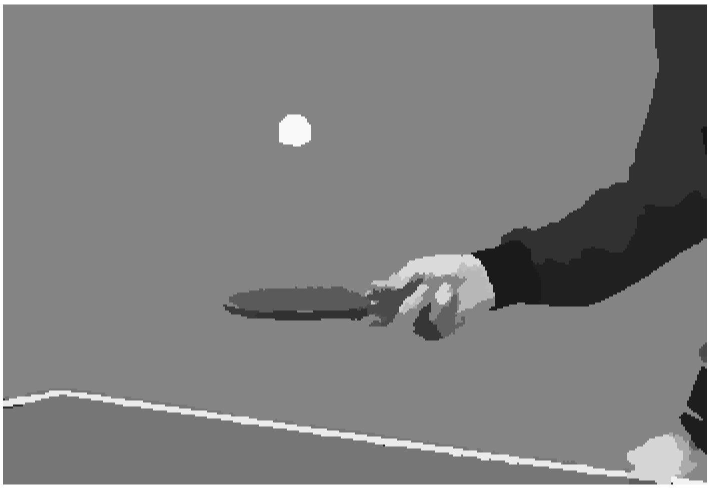
Foreground detection + sequential tracking

Track moving targets: this video shows 2 tracked cars in red and blue boxes (in poor weather conditions)

File name: ’TrackedCars’

Tracking results ...

Segmentation



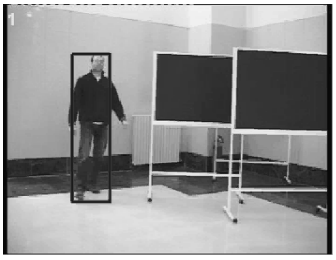
Tracking results...

Red box: proposed method

Tracking results ...

compare_result_prdestrian.avi

compare_result_hand.avi

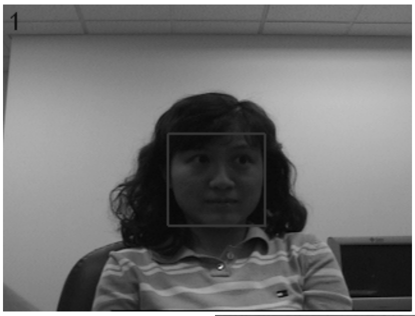


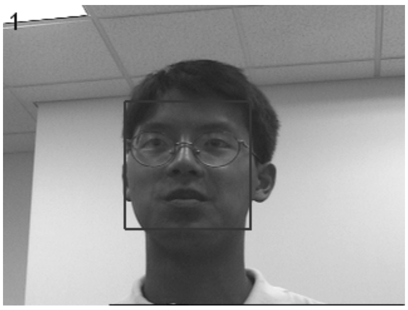
compare_result_occlussion.avi

compare_result_JAM.avi

Monitor eye movement: sleepy?

Tracking faces containing pose changes...





10: Lab.-3: Video compression using block-matching motion compensation

Task-1: Learn how to handle video data in Matlab

e.g. read /save *.avi video files, play movies,
extract video frames, convert image formats in Matlab.

Task-2. Obtain blocks containing motion

compute the difference of 2 image frames, and apply a threshold.

Task-3: Estimate motion vectors

apply block matching with the MAE criterion

Task-4: Inter coding through motion compensation

use matched blocks specified by motion vectors

Task-5: Intra coding

copy previous blocks for non-motion blocks

Task-6: Objective quality measure using PSNR

11. References

- **J.W.Woods, Subband image coding, Kluwer Academic Publisher,1991.**
- **D.S.Taubman and M.W.Marcellin, JPEG 2000, Kluwer Academic Pub.,2002.**
- **JVT /H.26L/ AVC Standard**
- **K.G.Nganm C.W.Yap, K.T.Tan,**
  Video coding for wireless communication systems, Marcel Dekker Inc., 2001
- **B.G. Haskell, A.Puri, A.N. Netraval,**
  Digital Video: an introduction to MPEG 2, Chapman & hall, 1997.
- **Proceedings of the IEEE, Vol.93, No.1, 2005:**
  Special Issue on advances in video coding and delivery