MPEG

CSC 790

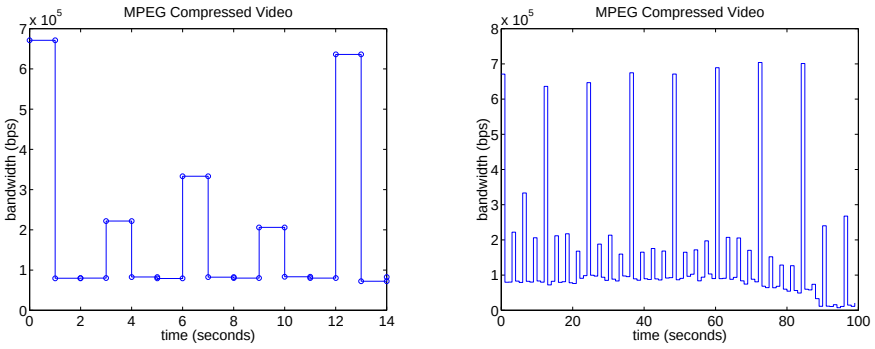
WAKE FOREST
UNIVERSITY

Department of Computer Science

Fall 2009

MPEG Frames

- Result of MPEG compression is a Variable Bit Rate (VBR) source



- High **peak-to-mean** ratios (18:1)
- MPEG has *self-similar behavior, slowly decaying autocorrelation*

  Why is this a problem for networks?

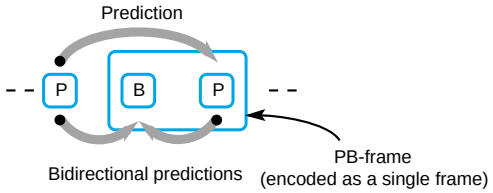
Error Concealment

- Estimate lost information to conceal an error that has occurred
  - Error concealment is performed at the decoder
- We know *video exhibits a significant amount of correlation along the spatial and temporal dimensions*
  - Assume single macroblock is lost or in error
  - Use **spatial interpolation**, **temporal interpolation (freeze frame)**, or **motion-compensated temporal interpolation**
- Spatial concealment, estimate missing pixels by extrapolating
  - Determine pixel values from surrounding pixels
  - Correctly recovering missing pixels is extremely difficult

- Temporal interpolation copies missing pixels from previous frame
  - Effective when there is no motion
- Motion-compensated interpolation copies pixels with movement
  - Can use coded motion vector, neighboring motion vector, or compute new motion vector
- These methods were designed for recovering one macroblock
  - *Over network, will usually lose more than one macroblock...*

H.263

- Low bit rate video compression standard by the ITU-T in 1995
  - Videoconferencing for PSTN, wireless and packet switched
  - Most Flash Video content (sites such as YouTube, Google Video, MySpace, etc...)[a]
- Uses I, P, and B-frames for higher compression than H.262
  - Neighboring pairs of P and B frames can be encoded together

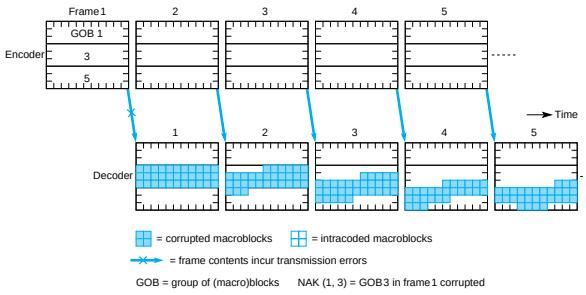


  - PB-frame consists of B and immediate succeeding P-frame

[a] TrueMotion VP6 is used by Flash8.

H.263 Error Resilience

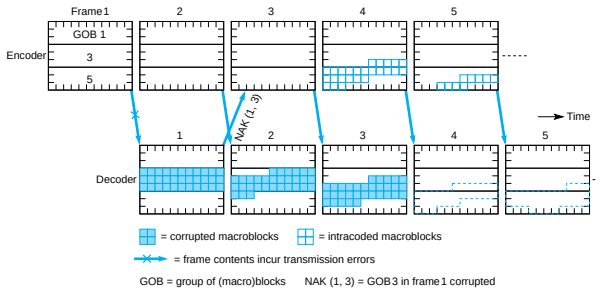
- H.263 is designed for PSTN and wireless environments
  - A **burst** of errors is possible (lose multiple macroblocks)
- H.263 implements error concealment, like H.261
  - *Use contents of corresponding GOB from previous frame*
  - However, errors in H.263 can propagate further than in H.261



- H.261 also includes **error tracking**, **independent segment decoding**, and **reference picture selection**

H.263 Error Tracking

- In video telephony, two-way communication is established
  - As a result, the decoder can inform the encoder of an error
  - Errors include, out-of-range motion vector, invalid codewords, out-of-range DCT, to many coefficients in macroblock
- In error prediction, encoder retains **error prediction** for each GOB
  - *Likely spatial and temporal effects on the macroblocks in the following frames if a specific GOB is in error*
- When an error is detected, decoder sends a NAK to the encoder
  - Indicates the frame and location of the GOB in error

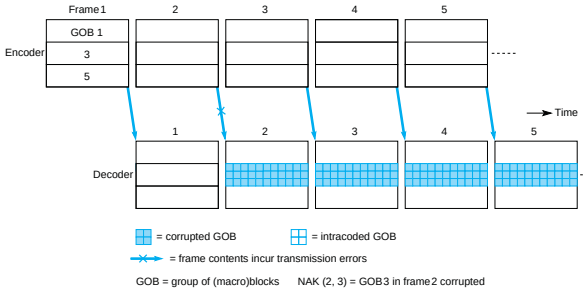


- Encoder uses the error prediction information about NAK frame
  - Identifies subsequent frames effected by the error
  - Transmits the macroblocks in their **intracoded form**

Fulp always said you don't have enough time to retransmit multimedia, is this proof he does not know anything?

H.263 Independent Segment Coding

- Assume each GOB is a *subvideo*, independent in the frame

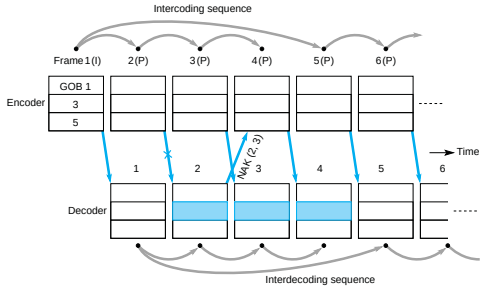


  - Motion estimation and compensation limited to within GOB
  - GOB in error only effects one GOB
- Decoder can NAK GOB, encoder can transmit intracoded form

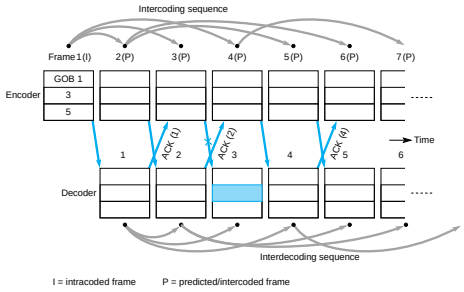
  What are the limitations?

H.263 Reference Picture Selection

- Attempts to stop errors propagating across frames
  - *Encoder selects one of several previous frames to use as a reference frame when encoding* (depends on the mode)
- NAK mode, decoder send NAK for GOB in error
  - Sends number of a previously-received, correctly-decoded reference frame that can be used as a reference for prediction



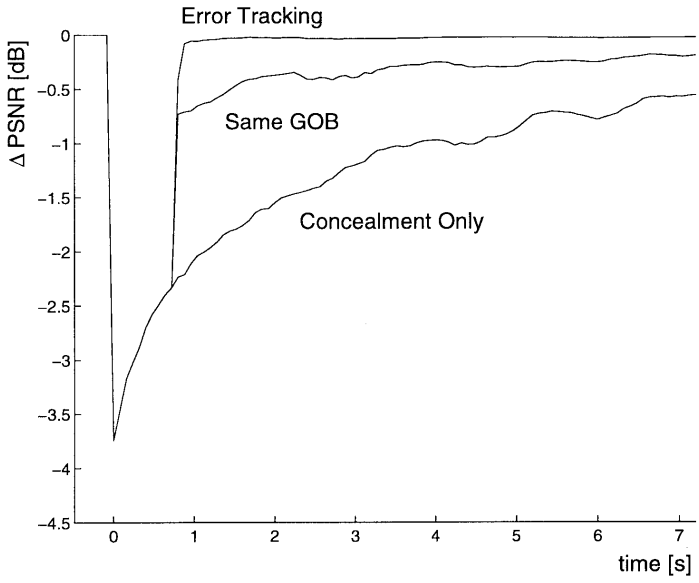
- ACK mode, decoder sends an ACK for all correct frames
  - Only frames that are acknowledged can be used for reference



- *Although error tracking can solve the problem of temporal error propagation, the use of intra mode coding can significantly increase the data rate; Reference Picture Selection (RPS) allows the encoder to select one of the several previously decoded frames as a reference for prediction*

Error Recovery Performance

- Consider the performance of concealment and error tracking



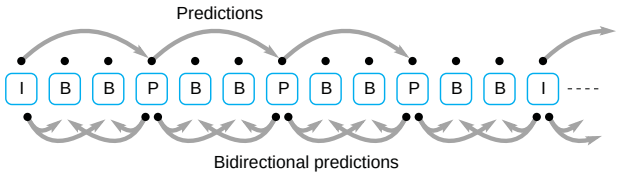
Error Recovery Performance

Error concealment of frame 75 and 90

Error tracking of frame 75 and 90

MPEG-1

- Developed for video compression, storage on CD-ROM
  - Compression is similar to H.261
- The standard allows I, B, and P-frames



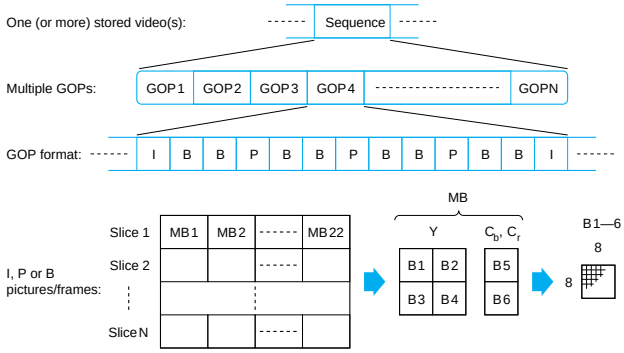
  - I-frames are used for random access (*minimum time 0.5 sec*)

    What?

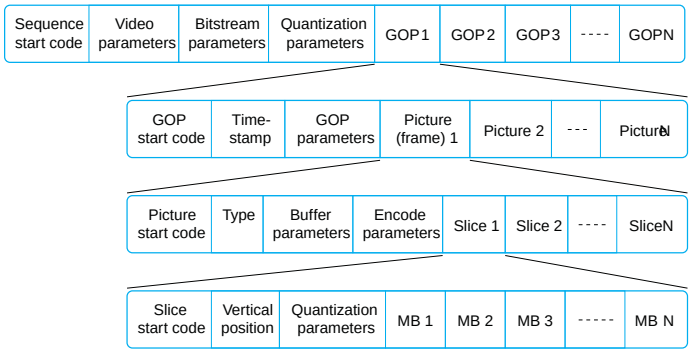
- Minimum access influences the separation of I-frames

MPEG-1 Hierarchy

- Complete video is called a **sequence**
- Sequence consists of GOP
- Each GOP consists of frames
- Each frame consists of **slices**
- Each slice consists of macroblocks (*like H.261*)



MPEG-1 Details



- Sequence starts with *sequence start code*
- GOP includes timestamp for synchronizing, then frames
- Frames include the frame type and slices
- Slices contain the macroblocks

MPEG-2

- Superset of MPEG-1 to support higher bit rates
  - Operates at different profiles and levels
- Profiles, a *subset of tools available for a family of applications*

Abbr.	Name	Frames	YCbCr	Streams
HP	High Profile	P, I, B	4:2:2	1-3
SP	Spatial Profile	P, I, B	4:2:0	1-3
SNR	SNR Profile	P, I, B	4:2:0	1-2
422P	4:2:2 Profile	P, I, B	4:2:2	1
MP	Main Profile	P, I, B	4:2:0	1
SP	Simple Profile	P, I	4:2:0 1	no interlacing

What is interlaced?

- Levels, *bounds of complexity of any profile*
  - Higher levels compatible with lower levels

Abbr.	Name	Pixel/line	Lines	fps	Bitrate (Mbps)
HL	High Level	1920	1152	30	80
H-14	High 1440	1440	1152	30	60
ML	Main Level	720	576	30	15
LL	Low Level	352	288	30	4

- For application, select a profile and level
  - For example, digital TV is MP@ML, HDTV is MP@HL

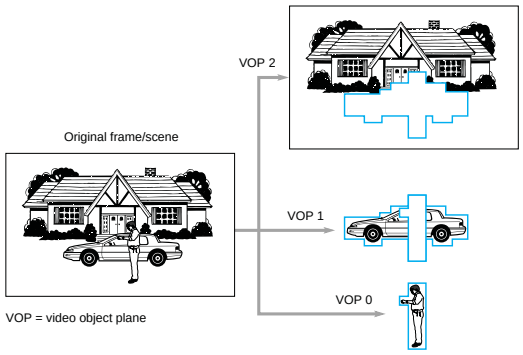
Prof@Level	Resolution	fps	Sampling	Mbps	Application
MP@HL	1920×1080	30	4:2:0	80	HDTV
MP@ML	720×480	30	4:2:0	15	DVD, SD-DVB
MP@LL	352×288	30	4:2:0	4	Set-top boxes (STB)

MPEG-4

- Designed for *interactive* multimedia (start/stop/play)
- MPEG-4 include several **content-based functionalities**
  - Before compression, picture fore- and background determined
  - Foreground objects are called **Audio-Visual Objects** (AVO)
  - Encoding of background and AVO's done separately
- **Scene descriptor** then defines the picture composition
  - Defines how AVO's are related to each other within picture
  - Picture divided into **Video Object Planes** (VOP)
  - Each VOP contains an AVO of interest

VOP

- VOP created such that is contains the minimum number of GOB

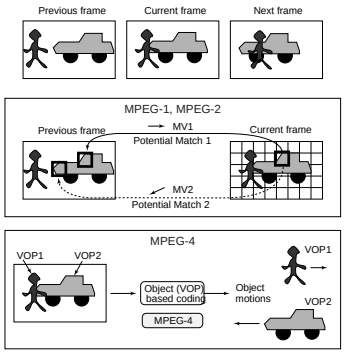


- VOP spatial coordinates determined relative to top left-hand
  - Motion of each VOP are encoded separately

    How do you identify an object?

VOP vs Frame

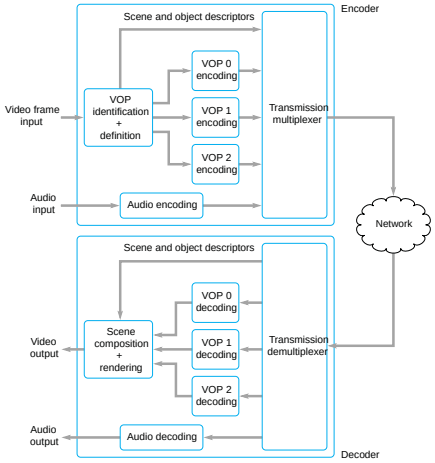
- MPEG-1 and 2 do **not** support VOP, thus called **frame-based**
  - Frame divided into macroblocks, motion compensation made
  - Concerned only in compression ratio, not visual objects



- MPEG-4 identifies VOP and provide motion compensation

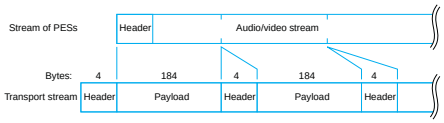
MPEG-4 Encoder/Decoder

- Video and audio encoded separately
  - Video frame divided into VOP, encoded, then multiplexed
  - Different audio encoders available, G.731, AC-3, etc...

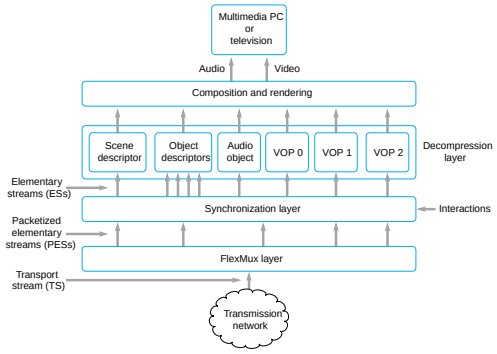


MPEG-4 Transmission Format

- **Transport Stream** consists of **Packetized Elementary Streams**

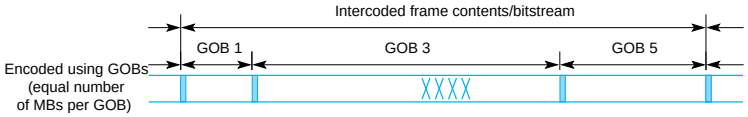


- **Elementary Streams** (ES) carry audio or video
  - PES contains header for *correct routing*



Error Resilience Problems

- H.261 and H.263 could detect errors at the GOB level
  - Could identify a GOB in error, discard entire structure (or try to recover some manner...)
  - *Amount of compressed data varies per GOB*

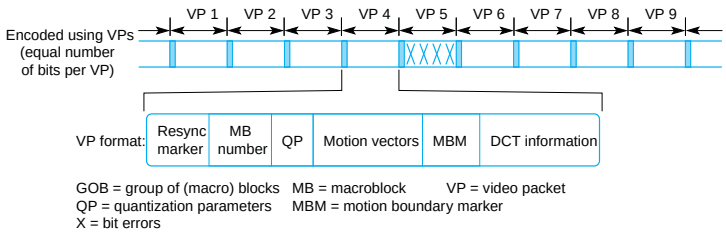


So what?

- MPEG-4 attempts to address this issue using **video packets**

MPEG-4 Video Packets

- MPEG-4 uses fixed length video packets separated by sync markers



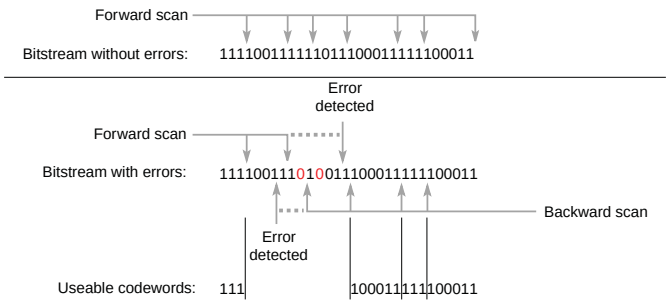
- If an error occurs a smaller set of macroblocks are impacted
- DCT is encoded using Variable-Length Code (*like Huffman*)
  - However with VLC, an error can cause loss of synchronization
  - MPEG-4 uses **reversible** VLC, can be read in either direction

RVLC

- Choose a set of codewords with equal *Hamming weight*
  - For example, each codeword has the same number of 1's
  - Can be done by adding same prefix and suffix to VLC

VLC	RVLC
1	111
01	1011
001	10011
0001	100011

Maximum codeword length = 6 bits



MPEG-4 Profiles and Levels

Profile	Level	Size	Bitrate	Max Objects
Simple	1	176×144	64 Kbps	4
	2	352×288	128 Kbps	4
	3	352×288	384 Kbps	4
Core	1	176×144	384 Kbps	4
	2	352×288	2 Mbps	16
Main	1	352×288	2 Mbps	16
	2	720×576	15 Mbps	32
	3	1920×1080 (HDTV)	28.4 Mbps	32

MPEG-4 Parts

- **Part 1** (ISO/IEC 14496-1): Systems: Describes synchronization and multiplexing of video and audio. For example Transport stream.
- **Part 2** (ISO/IEC 14496-2): Visual: A compression codec for visual data (video, still textures, synthetic images, etc.). One of the many "profiles" in Part 2 is the Advanced Simple Profile (ASP).
- **Part 3** (ISO/IEC 14496-3): Audio: A set of compression codecs for perceptual coding of audio signals, including some variations of Advanced Audio Coding (AAC) as well as other audio/speech coding tools.
- **Part 4** (ISO/IEC 14496-4): Conformance: Describes procedures for testing conformance to other parts of the standard.
- **Part 5** (ISO/IEC 14496-5): Reference Software: Provides software for demonstrating and clarifying the other parts of the standard.
- **Part 6** (ISO/IEC 14496-6): Delivery Multimedia Integration Framework (DMIF).
- **Part 7** (ISO/IEC 14496-7): Optimized Reference Software: Provides examples of how to make improved implementations (e.g., in relation to Part 5).

- **Part 8** (ISO/IEC 14496-8): Carriage on IP networks: Specifies a method to carry MPEG-4 content on IP networks.
- **Part 9** (ISO/IEC 14496-9): Reference Hardware: Provides hardware designs for demonstrating how to implement the other parts of the standard.
- **Part 10** (ISO/IEC 14496-10): Advanced Video Coding (AVC): A codec for video signals which is technically identical to the ITU-T H.264 standard. P
- *Etc...*

H.264

- ISO/IEC MPEG and ITU-T developed H.264 (MPEG-4 Part 10)
  - Objective, *create a standard capable of providing good video quality at bit rates that are substantially lower (e.g., half or less) than previous methods (MPEG-4 advance simple)*
  - From low bit rate mobile video to high-definition TV
- Some *core* features of H.264
  - Entropy encoding based on VLC, statistically generated
  - More sophisticated motion compensation
  - In-loop deblocking filters to remove macroblock edges

Video Compression Standard

Standard	Application	Bit Rate
H.261	Video telephony over ISDN	$p\times$64 kbps
MPEG-1	Video on digital storage (CD-ROM)	1.5 Mbps
MPEG-2	Digital Television	2 - 20 Mbps
H.263	Video telephone over PSTN	$>$ 33 Kbps
MPEG-4	Object-based coding	variable

Encoding Methods

Media	Encoding Methods		
Video	MPEG-4 ASP	H.264/MPEG-4	Lossless
	3ivx, DivX, FFmpeg, MPEG-4, HDX4, Xvid	AVC, CoreAVC, HDX4, QuickTime, H.264, x264	CorePNG, CoreAVC, FFV1, Huffyuv, Lagarith, MSU Lossless, x264
	XviD, DivX, Nero Digital QuickTime libavcodec	*x264, libavcodec, Apple Computer, Sorenson*	
Audio	General	Speech/Voice	Lossless
	ADPCM, ATRAC, Dolby Digital, Musepack, TwinVQ, Vorbis, WMA	iLBC, IMBE, iSAC, QCELP, Speex	Dolby, TrueHD,FLAC, TTA, WavPack, WMA Lossless