18-899 Special Topics in Signal Processing

Multimedia Communications:
Coding, Systems, and Networking

Prof. Tsuhan Chen
tsuhan@ece.cmu.edu

Lecture 5

H.263, H.263 Version 2, and H.26L

Very Low Bit Rate Video Coding

- Started around Nov 1993 by ITU-T SG 15
  - PSTN and mobile network: 10 to 24 kbit/s
- Near-term: 1994-date
  - H.263 and H.263 Version 2 (“H.263+”)
  - Enhancements of H.261
- Long-term
  - H.26L
  - Fundamentally different from H.261

H.263

- H.261 combined with MPEG-like features optimized at very low bit rates
- Compared to H.261
  - More picture formats, different GOB structures
  - Half-pel motion compensation, no loop filtering
  - 3-D VLC table
  - Four negotiable options
- Performance
  - 3~4 dB better PSNR than H.261 at <64 kbit/s
  - 30% bit rate saving compared with MPEG-1

Picture Formats and Chrominance Sampling

- More allowable formats than H.261
- 4:2:0 format: Same as H.261

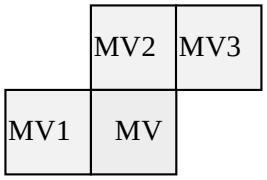
GOB Structures

Half-Pel Prediction

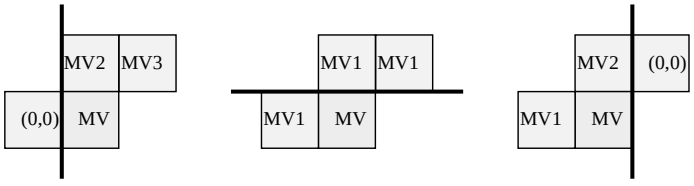
- Resolution of MVs is half-pel
- MV range: [-16,15.5]

- Bilinear interpolation

Motion Vector Prediction



MV: Current motion vector
MV1, MV2, MV3: predictors
prediction = median(MV1,MV2,MV3)



Picture boundary or GOB boundary

3D VLC Table

- A symbol is (last,run,level)
  - Last =1: indicates the last nonzero coefficient
  - No need for EOB

Negotiable Options in H.263

- Unrestricted Motion Vector Mode
- Advanced Prediction Mode
- PB-Frame Mode
- Syntax-based Arithmetic Coding Mode

- Usage
  - The decoder signals the encoder which of the options it has the capability to decode. If the encoder supports some of these options, it may enable them

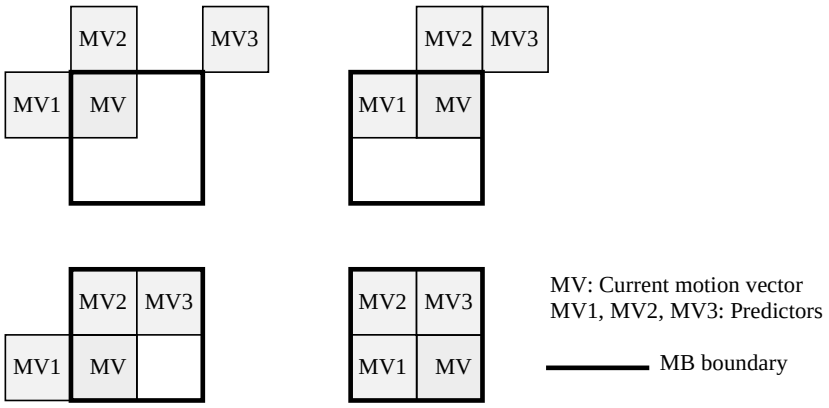
Unrestricted Motion Vector (UMV) Mode

- Motion vectors may point outside the picture
  - Edge pels are repeated
  - Significant gain for movement along the edge of the pictures, especially for smaller picture formats
- Extension of the motion vector range
  - [-31.5, 31.5] instead of [-16, 15.5]
  - Especially useful for 4CIF, 16CIF
- Good for camera movement and background motion

Advanced Prediction Mode (AP)

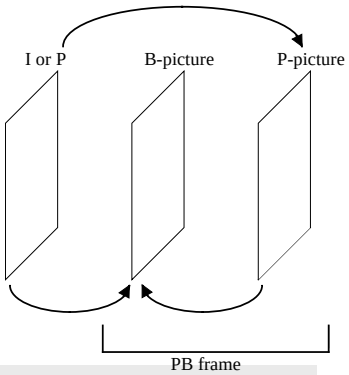
- Overlapped block motion compensation (OBMC)
  - Less blocking artifacts
- Four 8x8 vectors instead of one 16x16 vector are used for some macroblocks
  - Four vectors use more bits, but give better prediction. Encoder decides
- Motion vectors may point outside the picture as in UMV Mode

MV Prediction for 8x8 Mode



PB-Frame (PB) Mode

- A PB-frame consists of two pictures
  - P-picture: Predicted from the last decoded picture
  - B-picture: Predicted from both the last decoded picture and the P-picture currently being decoded



• MVs of B-picture derived from MVs of P-picture

• Decoding: One MB in P-picture followed by one MB in B-picture

Syntax-Based Arithmetic Coding (SAC) Mode

- Arithmetic coding is used instead of VLC
- Less bits at the same SNR
- Average bit rate saving ~5%
  - Inter frames: 3~4%
  - Intra blocks and frames: ~10%

Test Model Near-Term (TMN) 6

- Encoder specifications
  - Overlapped block motion compensation
  - Choice of 8x8 or 16x16 motion vectors
  - Syntax-adaptive arithmetic coding
  - Use of PB-Frame Mode

H.263 Version 2

- Also known as H.263+
- Enhancements of H.263
  - as opposed to H.26L
- Work plan

Development of H.263+

- 22 Key technical areas (KTA) were identified
- 12 KTAs were adopted in Nov 1996
  - Extended source formats
  - Negotiable options
    - 16 negotiable options in H.263 Version 2
  - Forward error correction
  - Supplemental enhancement information

Custom Source Formats

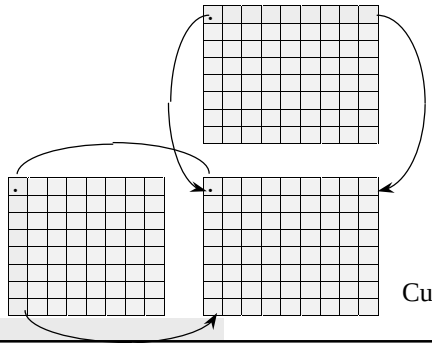
- Higher picture clock frequency (PCF)
- Custom picture formats
- Custom pixel aspect ratios (PAR)

New Negotiable Options

- Coding efficiency
  - Advanced Intra Coding Mode
  - Alternate Inter VLC Mode
    - Use intra table for inter DCT
  - Deblocking Filter Mode
    - Depending on quantization step size
  - Modified Quantization Mode
    - More flexible changes of quantization step sizes
    - Finer quantization for chrominance
    - Extended DCT range
  - Improved PB-Frame Mode
    - Forward, backward, or bi-directional

Advanced Intra Coding

- A separate VLC table for intra DCT
- Modified inverse quantization
- Spatial prediction of DCT coefficients
  - DC only, vertical DC & AC, horizontal DC & AC



New Negotiable Options (cont.)

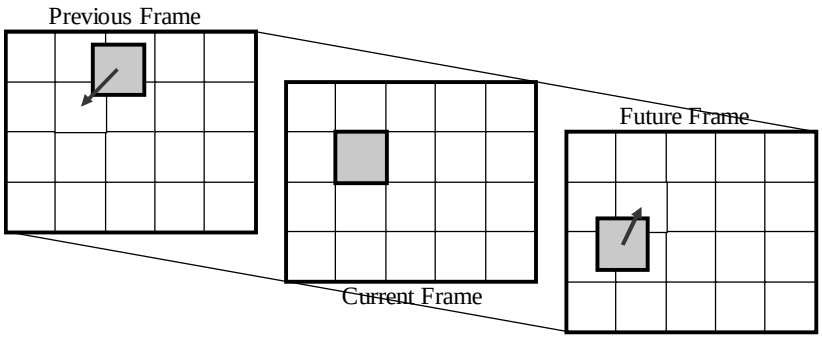
- Error robustness
  - Slice Structure Mode
    - Sequential or arbitrary
    - Rectangular or not
  - Reference Picture Selection Mode
    - Multiple reference pictures are stored
    - Avoid using pictures that contain errors
  - Independent Segment Decoding Mode
    - Prevent error propagation

New Negotiable Options (cont.)

- Scalability
  - Decode partial information from partial bitstream
  - To fit various bandwidth requirements
  - To fit terminals with different capabilities
- Scalability Mode
  - Temporal Scalability : Bi-directional prediction
  - Spatial Scalability
  - SNR Scalability

Bi-Directional Prediction

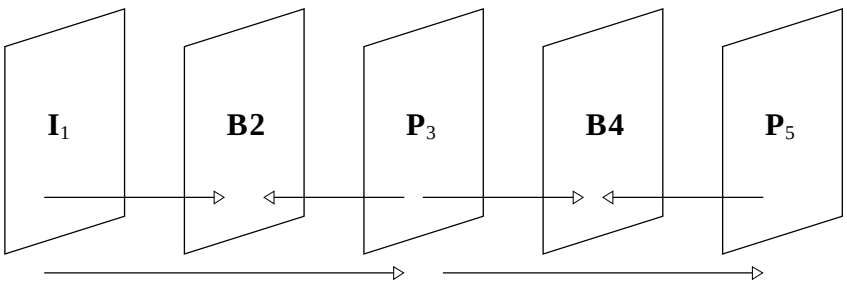
- Prediction from the previous frame, or the prediction from the future frame, or an average of both is used as the final prediction.
- The prediction error is then coded and transmitted



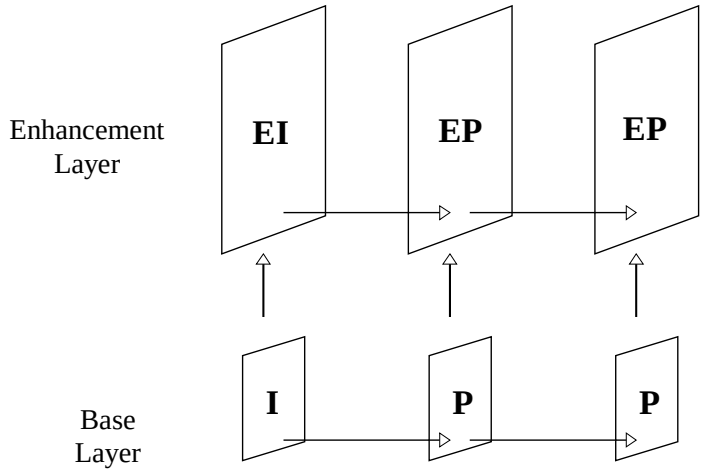
Bi-Directional Prediction (cont.)

- Advantages
  - Higher coding efficiency
  - No un-covered background problem
  - Increased frame rate with few extra bits
- Disadvantages
  - Two frames stores are needed at the decoder
  - More delay

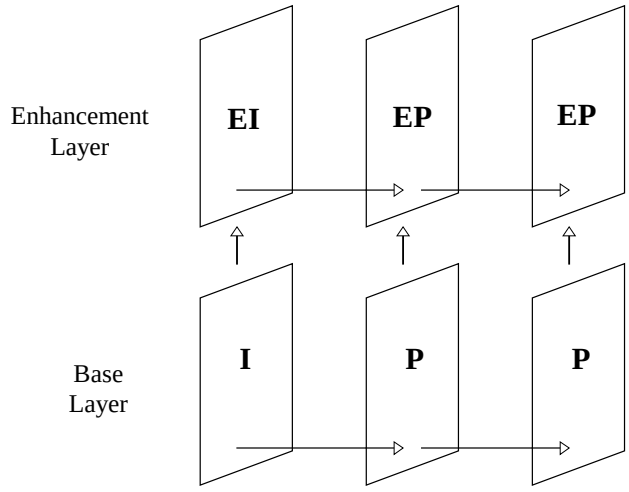
Temporal Scalability



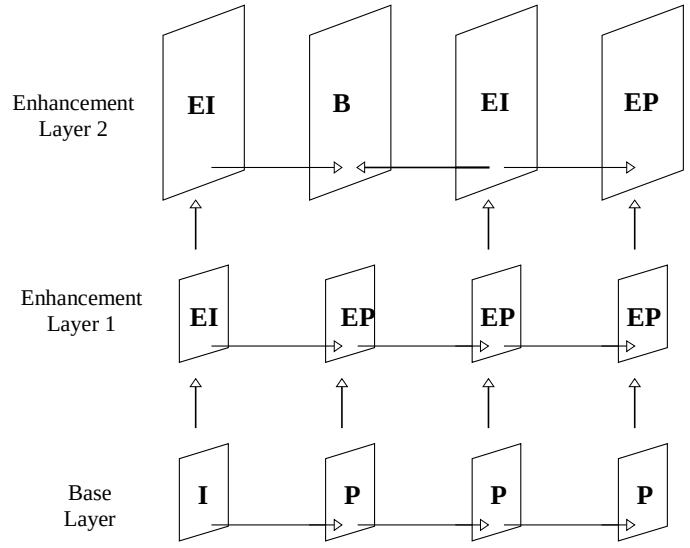
Spatial Scalability



SNR Scalability



Multilayer Scalability

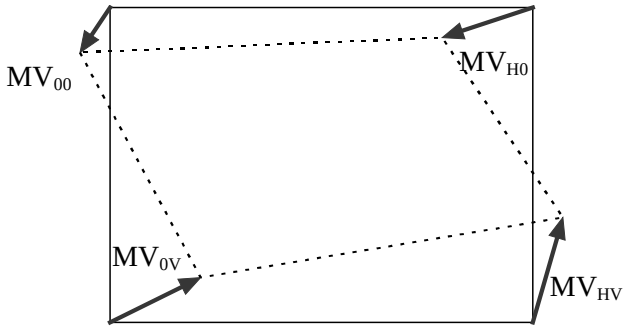


New Negotiable Options (cont.)

- Others
  - Reduced-Resolution Update Mode
    - Update a picture at a lower spatial resolution
    - To retain details in stationary background
  - Reference Picture Resampling Mode

Reference Picture Resampling

- For source format changes
- Global motion compensation
- Special-effect warping



Modifications to UMV Mode

- Single value VLC
  - Easy implementation
- Reversible VLC table
  - Better error resilience
- Larger motion vector range
  - Depending on the picture size
  - Up to [-256, 255.5]

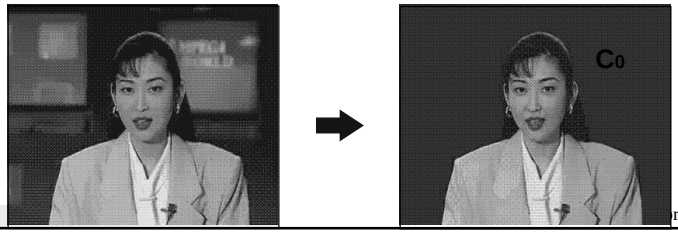
Supplemental Enhancement Information

- Can be discarded by the decoder
- Enhanced features
  - Picture freeze and release
    - Complete or partial picture
  - Tagging information
    - Snapshot
    - Video segment start/end
    - Progressive refinement start/end
  - **Chroma key**: to represent transparent pixels

Description

- A video sequence $f(x,n)$ with regions R_i
- For each region R_i, replace the non-object area with a special color C_0

$$g(\mathbf{x},n)=\begin{cases} f(\mathbf{x},n) & \text{if } \mathbf{x}\in\Re_i \\ C_0 & \text{if } \mathbf{x}\in\overline{\Re_i} \end{cases}$$



Description (cont.)

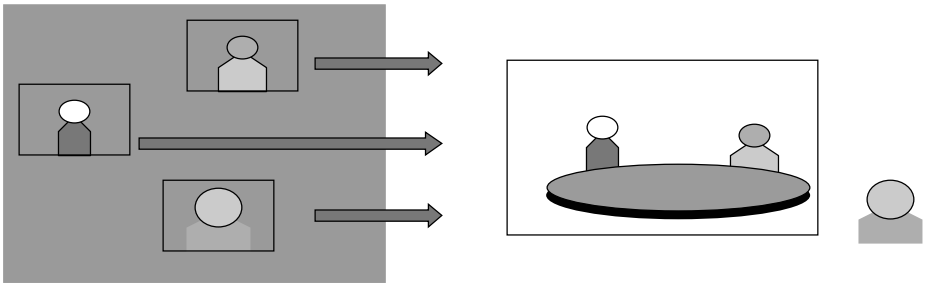
- The sequence $g(x,n)$ is then coded and transmitted
- The color C_0 and a threshold T are sent as side information to the decoder
- Chroma-keying
  - Based on C_0 and T , the decoder detects transparency and recovers the region boundary

Advantages of Chroma Keying

- Low complexity
- Efficiency
  - Minimal side information
  - Implicit shape coding
  - Solves “uncovered background”
- Minimal syntax change

Example Application

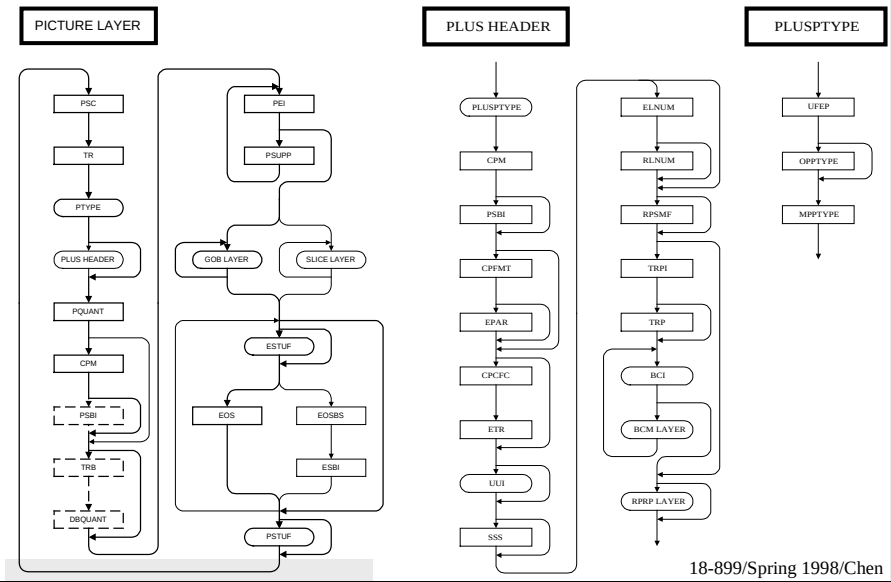
- Virtual video conference



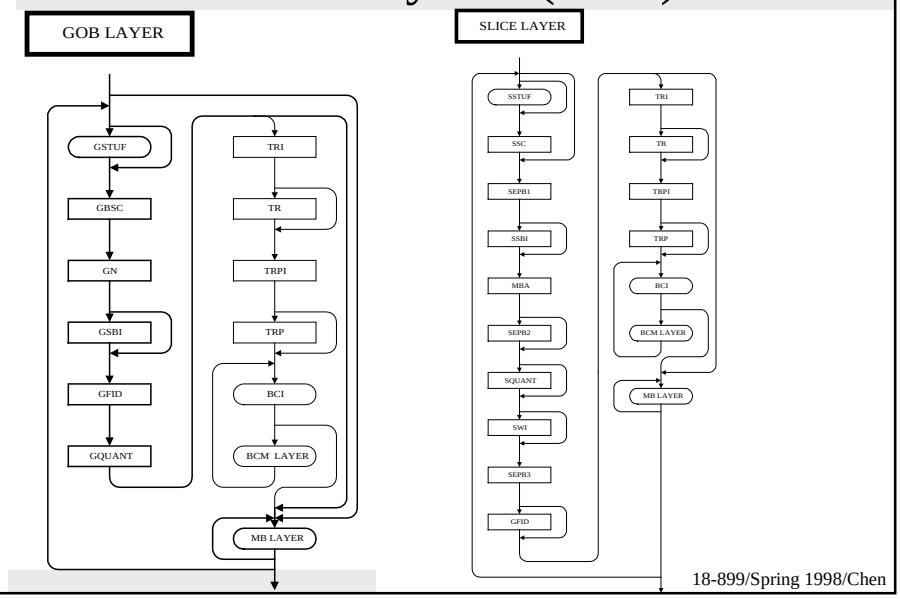
Test Model Near-Term (TMN) 8

- Encoder specifications
  - Rate control
    - Depending on buffer fullness
    - Skip pictures
    - Increase quantization step sizes

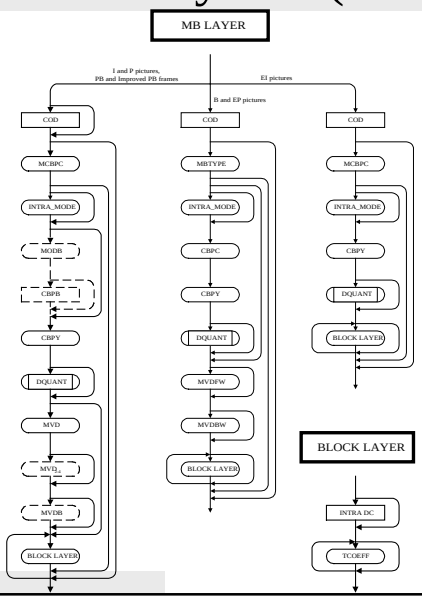
Bitstream Syntax



Bitstream Syntax (cont.)



Bitstream Syntax (cont.)



H.26L

- Better quality and more functionalities
- Call for Proposals, Jan 1998
  - Very low bit rates, real-time, low end-to-end delay
  - Low complexity permitting software implementations
  - Enhanced error robustness, including mobile networks
  - Adaptable rate control mechanisms
  - Variety of source materials

Applications of H.26L

- Real-time conversational services
- Internet video applications
- Sign language and lip-reading communication
- Video storage and retrieval services (e.g. VOD)
- Video store and forward services (e.g. video mail)
- Multi-point communication over heterogeneous networks

Tentative Time Schedule

Jan 1998	Call for proposals
Nov 1998	Evaluation of the proposals
Jan 1999	1st Test Model of H.26L (TML1)
Nov 1999	Final Major Feature Adoptions
Aug 2001	Determination
May 2002	Decision

H.263++

- Further extensions of H.263. Targeted for decision in Nov'2000
- Four KTAs
  - Data partitioning with reversible VLCs
  - 4×4 motion and DCT
  - Adaptive quantization
  - Long-term/Background frame store use

References

- ITU-T Q.15/16, Gary Sullivan, ed., "Draft Text of Recommendation H.263 Version 2 ("H.263+") for Decision," Sept. 1997
- ftp://standard.pictel.com/video-site/h263plus/
- Joan L. Mitchell et al., Sec. 19.3, *MPEG Video: Compression Standard,* Chapman & Hall, New York, NY
- Barry G. Haskell, Atul Puri, Arun N. Netravali, Sec 17.1, *Digital Video : An Introduction to MPEG-2*, Chapman & Hall, New York, NY