II.2 Error Control for Compressed Video: Error Resilience Coding & Error Concealment

Prof. Chia-Wen Lin (林嘉文)
Department of CSIE
National Chung Cheng University
886-5-272-0411 ext. 33120
http://cs.ccu.edu.tw/~cwlin
cwlin@cs.ccu.edu.tw

Outline

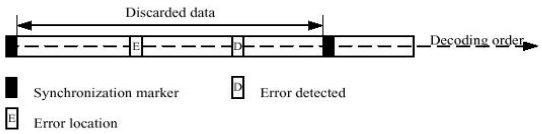
- Introduction
- Error Propagation
- Distortion Measures
- Error Resilience Coding
- Error Concealment
- Conclusions

Introduction

- The coding and transmission of compressed video over existing and future communication networks with non-guaranteed QoS presents many challenges
- In an error-prone environment, video-optimized error resilience techniques are necessary to accommodate the error/loss-sensitive nature of compressed video bitstreams
  - A single bit error in VLC can cause loss of sync
  - Motion compensation causes error propagation

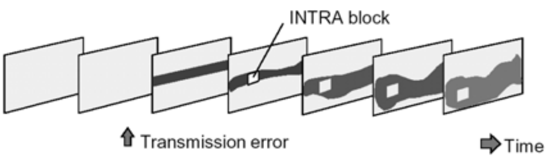
Effects of Bit Errors

- Since the decoder loses synchronization with the bitstream and fails to locate the bit errors, data between two synchronization words is usually discarded
- The impact of dropping a video segment that contains bit errors is equivalent to that of the loss of the complete segment

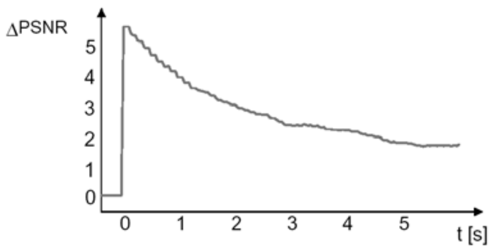


Error Propagation (1/2)

- The use of VLCs and predictive techniques in video coding leads to error propagation
- A single bit error can propagate to many bits
- MV prediction causes spatial error propagation
- Motion compensation causes temporal error propagation



Error Propagation (2/2)



- Single burst covering 1/3 of a frame
- Previous frame concealment
- Average over many trials
- No Intra

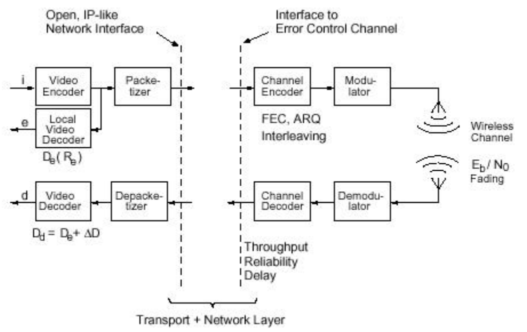
Effects of Packet Losses

- Packet sizes, spatiotemporal location of packets, and the extent of predictive coding will determine the impact of packet losses
- Depending on the size of packets, a packet loss may affect a small region of a frame, or lead to the loss of complete frame(s)
- In a predictive coding system, proper action is necessary at the encoder to limit the effects of error propagation

Trade-off between Source and Channel Coding

- The classical goal of *source coding* :
  - To achieve the lowest possible distortion for a given target bit rate
- The classical goal of *channel coding* :
  - To deliver reliable information at a rate that is closed as possible to the channel capacity
- A joint optimization:
  - Keep the source and channel coder separate
  - But optimize their parameters jointly
  - A key problem of this is the bit allocation between the source and channel coder

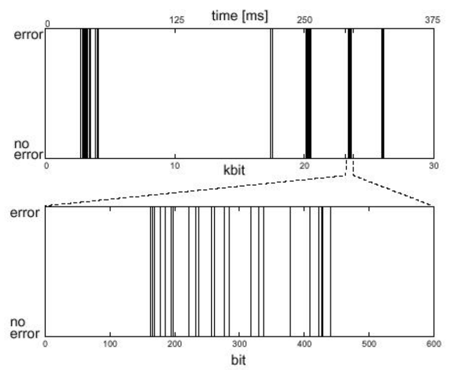
Components of a Wireless Video System



Transmission Errors in Wireless Channels

- Characteristics of the mobile radio channel:
  - It is a hostile medium
  - The propagation of electromagnetic waves is influenced by :
    - Absorption, reflection, diffraction, and scattering
  - It must cope with time-varying channel conditions
    - Large scale fading: by the *path loss*
    - Small scale fading: caused by *multipath propagation*
  - Errors are not limited to single bit errors but tend to occur in bursts
  - In severe fading situations the loss of synchronization may cause an intermittent loss of the connection

Burst Errors Encountered for Rayleigh Fading Channel



Distortion Measures (1/3)

- D_e: the video signal distortion introduced by the source encoder
- D_d: the distortion at the output of the video decoder
- In practice, the most common distortion measure is mean-square error (MSE)
- Hence, we define the distortion at the encoder as follows:

Distortion Measures (2/3)

- The MSE at the decoder is :

$$D_d = \frac{1}{XYTL}\sum_{x=1}^{X}\sum_{y=1}^{Y}\sum_{t=1}^{T}\sum_{l=1}^{L}(i[x,y,t]-d_l[x,y,t])^2$$

- The distortion due to source coding is described by D_e
- The distortion caused by transmission errors is described by $\triangle D$

$$\triangle D = D_d - D_e$$

Ref: B. Girod and N. Färber, "Feedback-based error control for mobile video transmission," *Proc. IEEE*, vol. 10, pp. 1707-1273, Oct. 1999.

Distortion Measures (3/3)

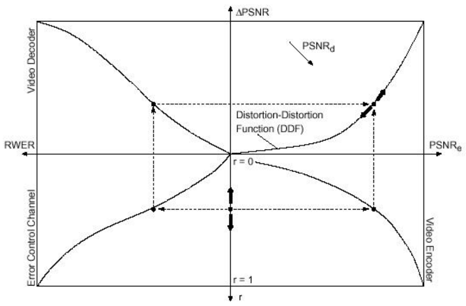
- MSE is commonly converted to *peak signal-to-noise ratio (PSNR)*
- PSNR is defined as 10 $\log_{10}(255^2/\text{MSE})$
- It is expressed in decibels (dB) and increases with increasing picture quality

$$PSNR_e = 10\log_{10}\frac{255^2}{D_e} \qquad PSNR_d = 10\log_{10}\frac{255^2}{D_d}$$

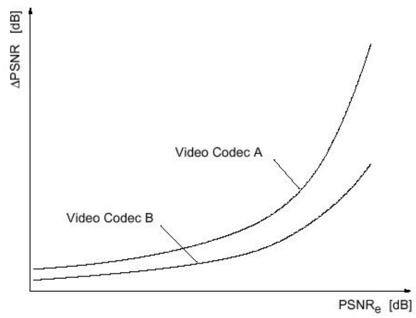
$$\triangle PSNR = PSNR_e - PSNR_d = 10\log_{10}\frac{D_e}{D_d} = 10\log_{10}\frac{D_e}{D_e + \triangle D}$$

Distortion-Distortion Function (DDF)

- The interaction of the various characteristics :



DDF of Two Typical Codecs



Error Resilience Coding

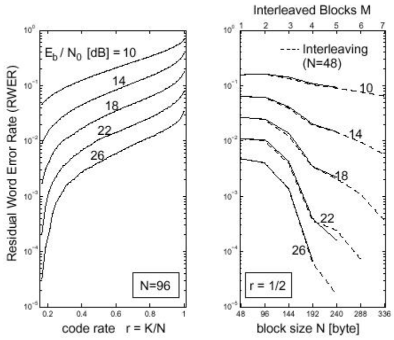
Outline

- FEC and ARQ
- Spatial Error Resilience Coding Techniques
- Temporal Error Resilience Coding Techniques
- Error Resilience Tools in Current Standards
- Summary

Forward Error Correction (FEC)

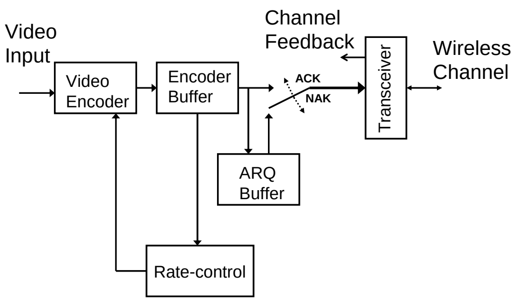
- Reduce effective channel bandwidth
- Increase delay with long block or interleaving
- Not adaptive to varying loss characteristic, best when packet loss-rate is stable
- Good for random errors with low BER, not suited for long burst errors
- Equal / Unequal Error Protection

RWER vs. Channel Code Rate & Block Size



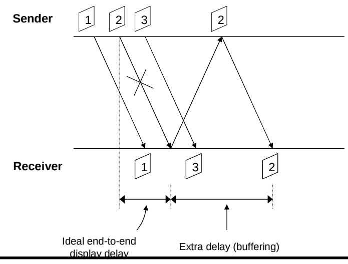
RWER: Residual word error rate

Automatic Repeat Request (ARQ) (1/2)



Automatic Repeat Request (ARQ) (2/2)

- Effective against burst errors and packet losses
- Very effective with large de-jittering buffer
- Cannot be used in video applications without feedback channel (e.g., TV broadcasting)
- Not realistic for delay stringent real-time applications



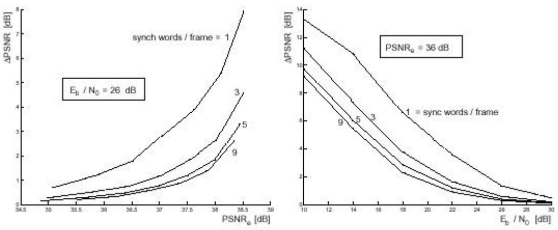
Spatial Error Resilience Coding Techniques

- Insertion of sync markers
- Data Partitioning
- Network-aware packetization
- Error resilience entropy coding methods
  - Reversible VLC
  - Error-resilient entropy coding
- Multiple-description (MD) coding

Synchronization Markers

- codewords that are uniquely identifiable in the bitstream (e.g., H.263 : use the 17-bit sync word “00000000000000001”)
- not only provide for bitstream synchronization, but also ensure spatial synchronization at the decoder
- data dependencies across slice boundary should be removed to contain the errors within a slice
- may be inserted at various locations to contain the errors to a small spatial region

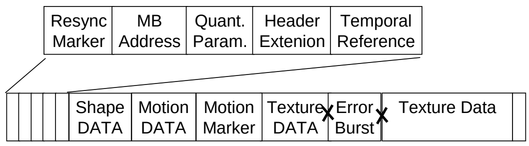
Video Quality vs. Number of Sync Words



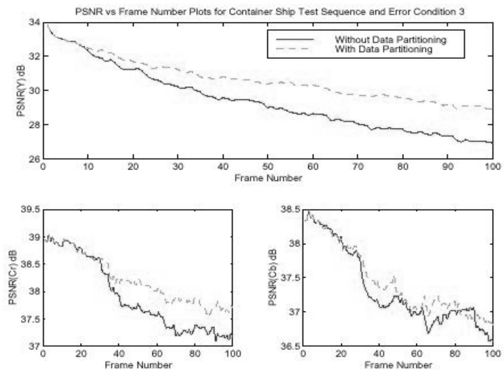
Sequence: Mother & Daughter

Data Partitioning

- The grouped MVs and DCT coefficients are separated by a boundary marker
- The boundary marker signals the end of the MV and the beginning of the DCT coefficient
- Advantages
  - errors can be localized to data of a certain type, and the unaffected data can be employed for video reconstruction
  - if undetected errors occur in a video packet, the received data can be considered invalid if the boundary marker is not detected



Performance of Data Partitioning

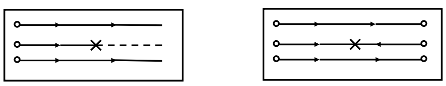


Network-Aware Video Packetization

- Staring a packet with a synchronization point, when possible, allows the independent coding of the packet
- The size of a packet is either fixed (e.g., ATM) or chosen subject to the network constraints (e.g., the Internet)
- The size of a coed video segment can be adapted to a target packet size by using slices
- Adding a redundant representation of the picture header to each packet can avoid dropping the whole picture when the first packet is lost (e.g., MPEG-4 HEC)

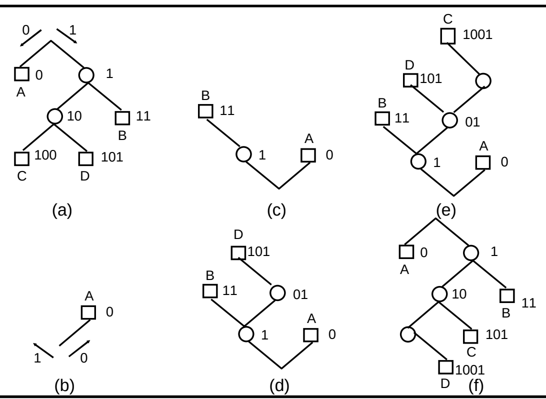
Reversible Variable-Length Codes (RVLC) (1/2)

- VLCs that have the prefix property in the forward and reverse directions => uniquely decodable in both directions
- Can be constructed to be symmetric or asymmetric
- Asymmetrical RVLCs provide better compression efficiency than symmetrical RVLCs
- Disadvantage : reduces the compression efficiency
- Usually employed with data partitioning



Forward decoding only — Forward and backward decoding

Constructing an Asymmetric RVLC

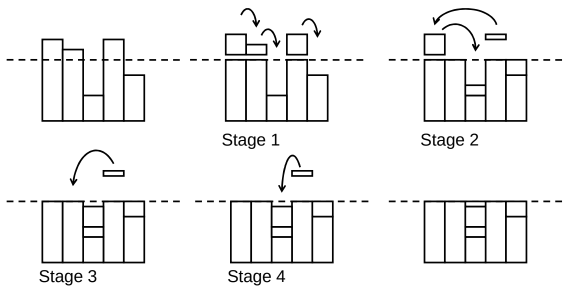


Error-Resilient Entropy Coding (EREC)

- Used to convert VLCs to fixed-length blocks of data
- Requires that some side information be transmitted on highly protected channels
- Does not guarantee image/video frame synchronization
- EREC bit reorganization algorithm :
  - allocate each block of VLC data to a corresponding EREC slot (either fully occupied or with some unused space)
  - at stage *n*, the remaining bits of VLC code at slot *i* are allocated to slot ($i + \phi_n$) mod *N*, if space is available => all the bits are placed within *N* stages
  - the remaining spaces in the EREC slots are filled with redundant bits

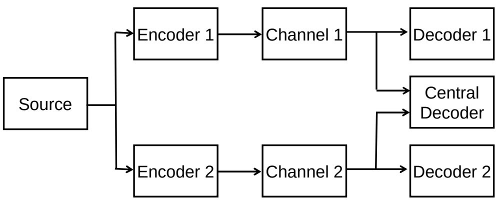
D. W. Redmil and N. G. Kingsbury, "The EREC: An error resilient technique for coding variable-length blocks of data," *IEEE T-IP*, vol. 5 pp. 565-574, Apr. 1996

EREC Bit Reorganization Procedure



Multiple Description Coding (MDC)

- MD coding allows a video decoder to extract meaningful information from a subset of the bitstream
- An encoder produces two descriptions (may be equally important) that are transmitted over two channels



Multiple Description Scalar Quantization [Vaishampayan 93]

- The input signal *x* is quantized to yield an integer index $I = q(x)$, where $q(\cdot)$ is a uniform quantizer
- Information about *I* is mapped to a pair of indexes $(i,j) = a(I)$
- The index *i* is transmitted on channel 1, while the index *j* is transmitted on channel 2
- If information for channel 1 or 2 only is received, the distortion level D_1 or D_2 will be incurred, respectively
- Receiving both information can obtain the full quality
- If only one index is received, it is possible to estimate the index *I* by choosing the central index in the row/column of the received index *i* or *j,* respectively
- Has been applied to intra coding of blocks in a DCT-based image/video coding framework

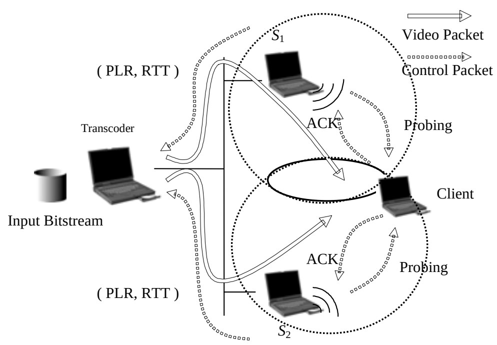
Example of Multiple Description Quantizer Assignment

i →

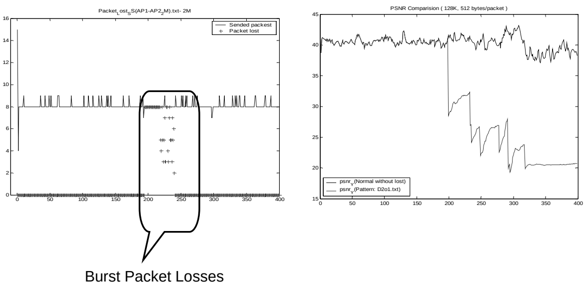
j ↓

1	3						
2	4	5					
	6	7	9				
		8	10	11			
			12	13	15		
				14	16	17	
					18	19	21
						20	22

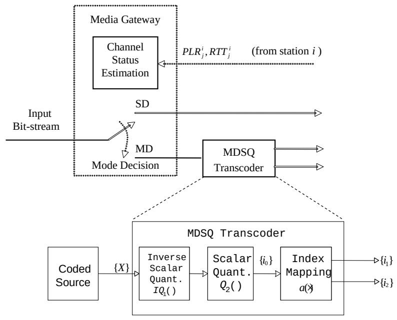
MDC with Diverse Routing: Roaming of Mobile Terminals in WLANs



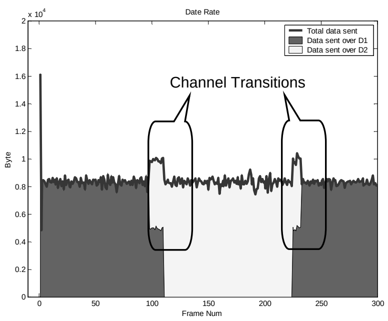
Quality Degradation due to Channel Switching



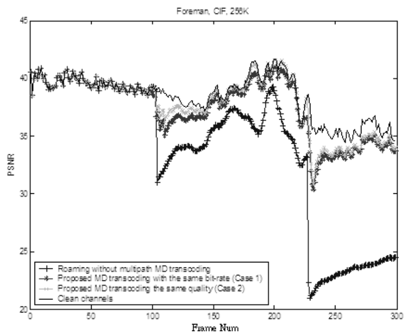
Channel-Aware Multiple-Description (MD) Transcoding



Channel Utilization with MD Transcoding



Performance Evaluation



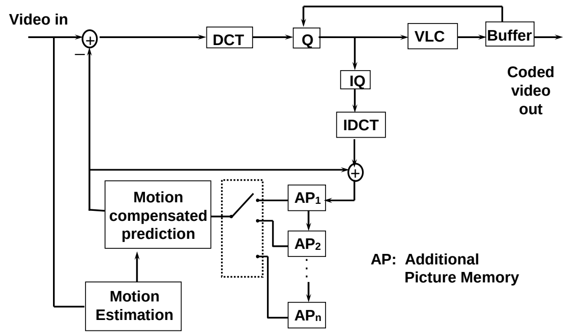
Temporal Error Resilience Coding Techniques

- Reference Picture Selection
- Video Redundant Coding
- Random Intra Coding
- Intra Coding Based on Feedback Information

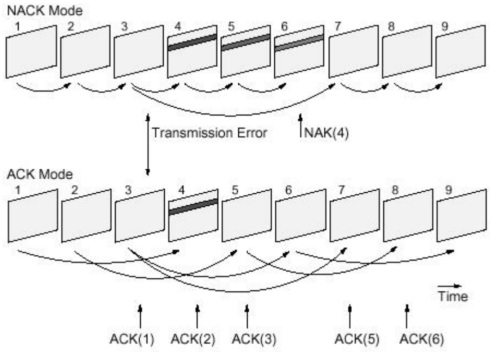
Reference Picture Selection (1/4)

- RPS dynamically replaces reference pictures in the encoder in response to an ACK of the decoder
- Two modes of operation are defined depending on the acknowledgment message
  - ACK mode
  - NACK mode
- Both methods are sensitive to errors in the feedback channel
- The NACK method is effective for low error rates and the ACK method is effective for high error rates

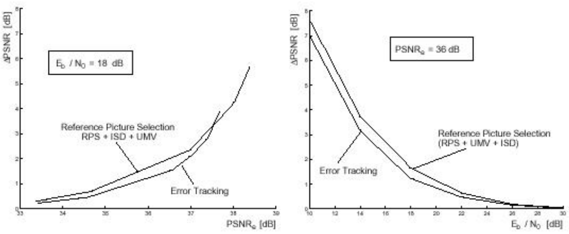
Reference Picture Selection (2/4)



Reference Picture Selection (3/4)

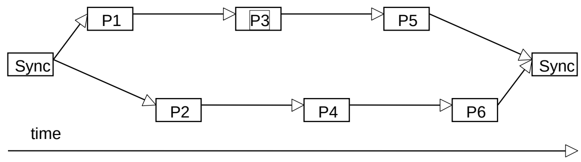


Reference Picture Selection (4/4)



Sequence: Mother & Daughter

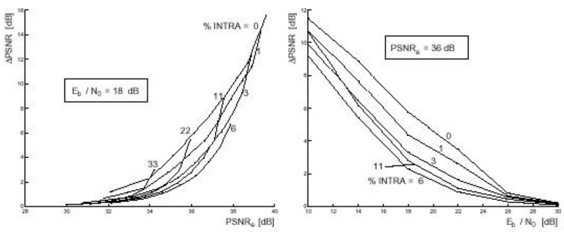
Video Redundancy Coding



Random Intra Coding

- Increasing the frequency of intra-coded frames/blocks can avoid error propagation while leading to an increase of bit rate
- The refresh rate of random intra coding can be chosen based on
  - the life expectancy of the errors [Haskell 92]
  - the block activities [Villasenor 96], [Hwang 00]
  - the block error rate [Kossentini 99]
- The frequency of intra updating can be approximated by $I_{freq} = 1/p$ [Kossentini 99]

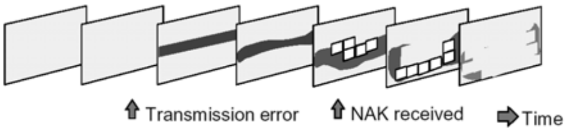
Performance of Intra Block Refresh



Sequence: Mother & Daughter

Feedback-Based Error Control

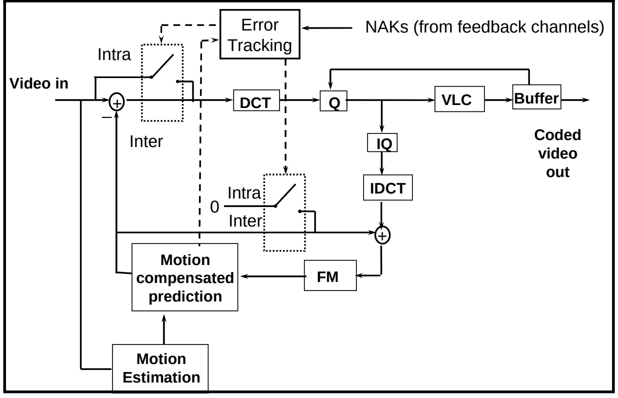
- Spatio-temporal error propagation can be reconstructed at the encoder using an Error Tracking algorithm and feedback from the decoder
- Feedback consists of sending Negative Acknowledgements (NAKs) for lost image parts
- Use INTRA-mode for macroblocks affected by transmission errors to stop error propagation



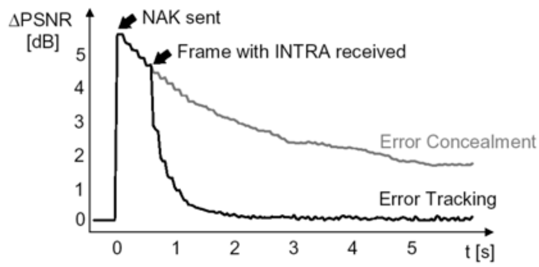
Error Tracking (1/4)

- Decoder
  - report addresses of erroneous blocks (NAKs) to encoder via a feedback channel
  - may also perform error concealment for unsuccessfully decoded GOBs
- Encoder
  - evaluate and reconstruct location and extent of propagated errors
  - INTRA refresh all or part of distorted MBs

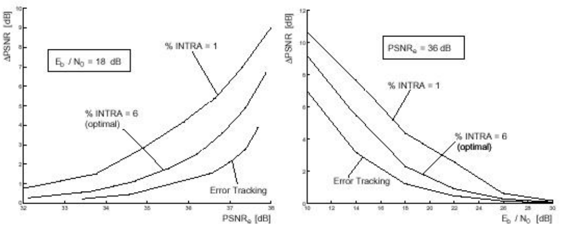
Error Tracking (2/4)



Error Tracking (3/4)



Error Tracking (4/4)



Sequence: Mother & Daughter

Rate-Distortion Optimized Intra Coding

- For RD optimized mode decision, the coding mode is chosen from three modes: skip, intra, inter, by minimizing

  $J_{mode}(n,(x,y)) = D(n,(x,y),\text{mode}) + \lambda_{mode}R(n,(x,y),\text{mode})$

  where $\lambda_{mode} = 0.45(Q/2)^2$
- For error prone environment, the distortion at the decoder is divided into two parts, so the cost function becomes:

  $J_{mode}(n,(x,y)) = [1-p(n,(x,y))]D_1(n,(x,y),\text{mode}) + p(n,(x,y))\, D_2(n,(x,y)) + \lambda_{mode}R(n,(x,y),\text{mode})$

  where D_1 represents the distortion of the received block, and D_2 is the concealment error when the block is lost

Rate-Distortion Optimized Intra Coding

The distortion D_1 can be obtained by

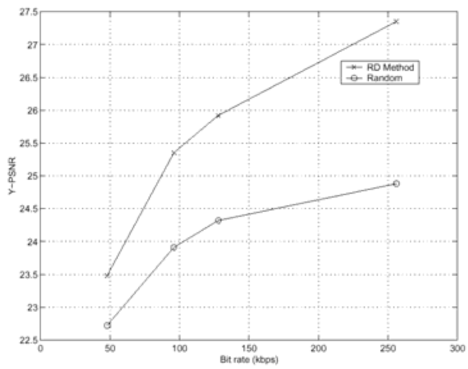
$$D_1(n,(x,y),\text{mode}) = D_q(n,(x,y),\text{mode}) + \sum_{k=1}^{N} p\left[n-k,(x+v_x,y+v_y)\right]D_2\left[n-k,(x+v_x,y+v_y)\right]$$

where

$$p\left[n-k,(x+v_x,y+v_y)\right]D_2\left[n-k,(x+v_x,y+v_y)\right] = \sum_{l=1}^{9} w(l)\,p\left[n-k,(x,y)_l\right]D_2\left[n-k,(x,y)_l\right]$$

Weights $w(l)$	Value	Conditions			
		$v_x \ge 0$		$v_x < 0$	
		$v_y \ge 0$	$v_y < 0$	$v_y \ge 0$	$v_y < 0$
$w(1)$	$[(16-\lvert v_x\rvert)\times(16-\lvert v_y\rvert)]/256$	X	X	X	X
$w(2)$	$[(16-\lvert v_x\rvert)\times\lvert v_y\rvert]/256$		X		X
$w(3)$	$[v_x \times \lvert v_y\rvert]/256$		X		
$w(4)$	$[v_x \times (16-\lvert v_y\rvert)]/256$	X	X		
$w(5)$	$[v_x \times v_y]/256$	X			
$w(6)$	$[(16-\lvert v_x\rvert)\times v_y]/256$	X		X	
$w(7)$	$[\lvert v_x\rvert \times v_y]/256$			X	
$w(8)$	$[\lvert v_x\rvert \times (16-v_y)]/256$			X	X
$w(9)$	$[\lvert v_x\rvert \times \lvert v_y\rvert]/256$				X

Performance of R-D Optimized Intra Coding



Error Resilience Tools in Current Standards

- Forward Error Correction (FEC)
  - H.263: (511,493) BCH code (18 parity bits for 2-bit correction)
- Synchronization Words
  - H.263+, MPEG-4: GOB or slice (optional)
- Reversible Variable-Length Codes (RVLC)
  - H.263+: Annex D
  - MPEG-4: accompanied with data partitioning
- Data Partitioning
  - supported in H.263++ (version 3) & MPEG-4
- Independent Segment Decoding (ISD)
  - H.263+: a segment may be a slice, a GOB, or a number of consecutive GOBs
  - MPEG-4: similar error resilience gains can be achieved using video objects

Error Resilience Tools in Current Standards

- Reference Picture Selection (RPS)
  - H.263+ (Annex N): supports RPS and VRC. The methods can be applied to pictures or to individual rectangular picture segments
  - Will likely be include in MPEG-4 version 2
- Header Extension Code (HEC)
  - Allows the introduction of duplicate copies of important picture header information in the video packets
  - Supported in MPEG-4 and the RTP payload spec for H.263+

Standard	FEC	Sync words	RVLCs	Data Partitioning	Independent Segment Decoding	Reference Picture Selection	HEC
H.263	Yes	Yes	Yes	No [a]	Yes	Yes	No [b]
MPEG-4	No	Yes	Yes	Yes	No	No [c]	Yes

[a] Data partitioning will be included in Version 3 of H.263

[b] Although H.263 does not support this mode, the transport protocol for H.263 may support it, for example in RFC-2429.

[c] Reference Picture Selection will likely be included in Version 2 of MPEG-4

Summary

- Existing and future communication networks do not always guarantee error-free transmission; thus transmission of compressed video in such environments presents many challenges
- Error resilience coding methods allow reliable video communication
- Many of these error resilience methods are supported by video standards such as H.263 and MPEG-4
- Error resilience video coding is a relatively new area of research to explore

Error Concealment

Outline

- Introduction
- Error Detection
- Spatial Domain Error Concealment
- Temporal Domain Error Concealment
- Summary

Introduction

- Channel/network errors on coded video usually cause objectionable visual distortion when the compression ration is high
- Two key factors in error concealment
  - Redundancy which remains in the "coded" video
  - Human perception tolerance on video distortion
- Error concealment belongs to the general problem of image recovery and restoration
- Drawbacks of the use of prediction and VLC
  - Makes the video stream extremely sensitive to transmission errors
  - Prediction leads to error propagation
  - VLC makes it impossible to decode received bits following a single-bit error until a sync word is encountered

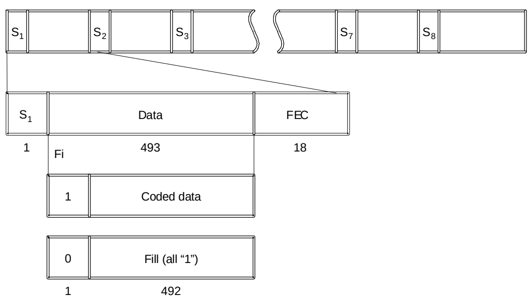
Introduction (Cont.)

- Depending on the frequency of sync codewords, transport packet size, and bit rate, a damaged region can range from part of a MB to an entire picture
  - High bit-rate transmission over small packet size (e.g., ATM)
    - A damaged block is typically surrounded by multiple undamaged blocks
    - Spatial and/or Temporal-domain error concealment
  - Low bit-rate application over relatively large packet size (e.g., 128 kbits/s over IP networks)
    - A lost packet results a large damaged region
    - Temporal-domain error concealment

Error Detection at Transport Level

- Sequence number is used to detect packet losses
  - Put in RTP for H.323, and in H.223 for H.324
- FEC code can be used to detect bit errors
  - Bitstream is divided into small pieces (frames)
  - FEC encoding is applied to each frame
  - H.223 uses FEC for error detection
  - In H.320-based systems, the video data encoded in H.261/263 are framed with an 18-bit FEC field in each 512-bit frame

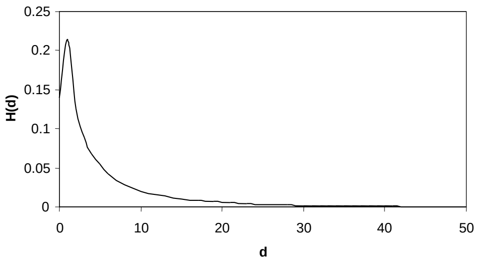
H.320 FEC Framing



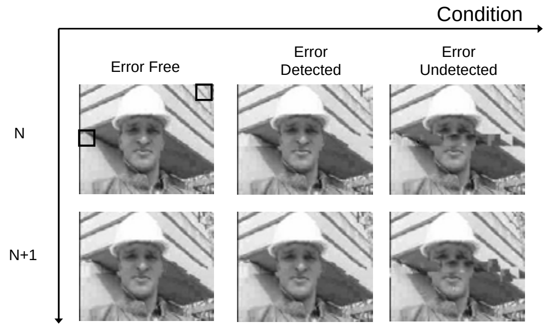
Error Detection in Video Decoder

- Codeword not included in the VLC table
- Invalid decoded sample and header info
  - Illegal quantization step-size (e.g., outside [1,31] for H.261/263)
  - The number of DCT coefficients exceeds 63
  - Loss of sync words
- Error detection in the pixel domain
  - Pixel differences between adjacent blocks
  - Average intersample difference (AID) (Leou *et al*. 98)
  - May require significant computation
- Mismatch in the number of encoded and decoded blocks
- In general, transport-level error detection is more reliable and requires less computation, but consumes more channel bandwidth

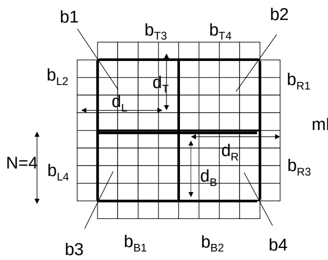
An Example of Histogram of Pixel Difference



Importance of Error Detection



Block-based Pixel Domain Interpolation



mb$_L$

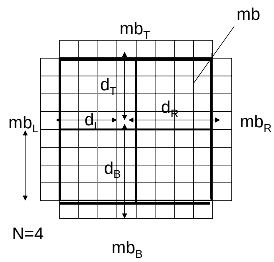
$$b_1(i,k) = \frac{d_T b_{L2}(i,N) + d_L b_{T3}(N,k)}{d_L + d_T}$$

$$b_2(i,k) = \frac{d_T b_{R1}(i,1) + d_R b_{T4}(N,k)}{d_R + d_T}$$

$$b_3(i,k) = \frac{d_B b_{L4}(i,N) + d_L b_{B1}(1,k)}{d_L + d_B}$$

$$b_4(i,k) = \frac{d_B b_{R3}(i,1) + d_R b_{B2}(1,k)}{d_R + d_B}$$

MB-based Pixel Domain Interpolation



$$mb(i,k) = \frac{1}{d_L + d_R + d_T + d_B}(d_R mb_L(i,2N) + d_L mb_R(i,1) + d_B mb_T(2N,k) + d_T mb_B(1,k))$$

Maximally Smooth Recovery

For a damaged image, the reconstructed image block can be described as

$$\hat{\mathbf{f}} = \mathbf{T}_r \tilde{\mathbf{a}}_r + \mathbf{T}_l \hat{\mathbf{a}}_l$$

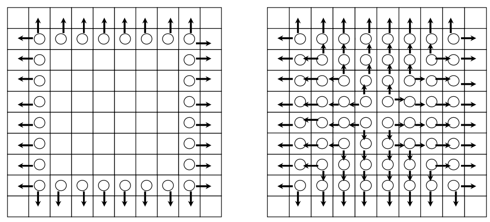
The optimal solution can be obtained by minimizing

$$\psi(\mathbf{a}_l) = \frac{1}{2}\left[(\|\mathbf{S}_w\hat{\mathbf{f}} - \mathbf{b}_w\|^2 + \|\mathbf{S}_e\hat{\mathbf{f}} - \mathbf{b}_e\|^2 + \|\mathbf{S}_n\hat{\mathbf{f}} - \mathbf{b}_n\|^2 + \|\mathbf{S}_s\hat{\mathbf{f}} - \mathbf{b}_s\|^2) \right]$$
$$= \frac{1}{2}\left[(\hat{\mathbf{f}}^T S \hat{\mathbf{f}} - 2b^T \hat{\mathbf{f}} + c) \right]$$

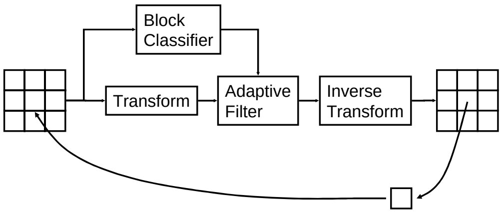
The optimal solution is

$$\hat{\mathbf{a}}_{opt} = (\mathbf{T}_l^T \mathbf{S} \mathbf{T}_l)^{-1} \mathbf{T}_l^T \left[\mathbf{b} - \mathbf{S}\mathbf{T}_r \tilde{\mathbf{a}}_r \right]$$

Two Smoothing Constraints

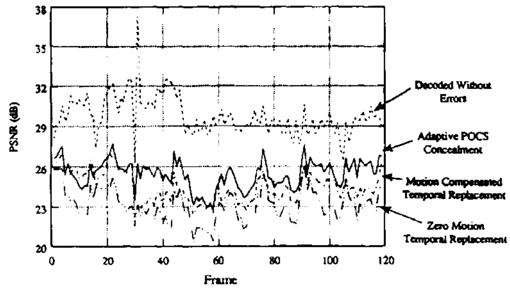


Projection onto Convex Sets (POCS) Method



Adaptive POCS iterative restoration process.

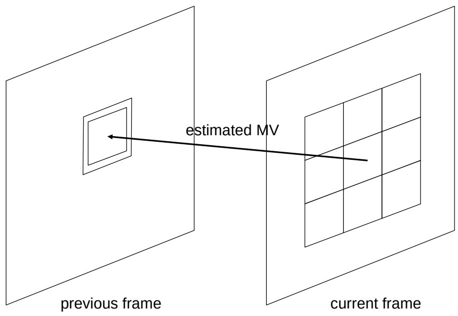
PSNR Comparison for POCS & Temporal Replacement Concealment



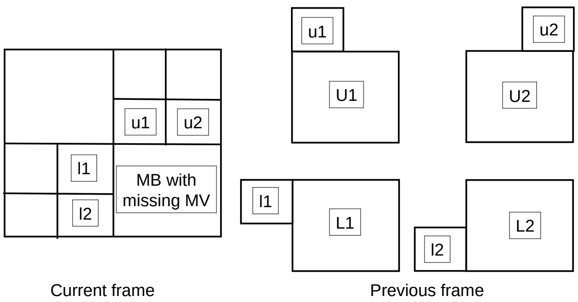
Motion-Compensated Prediction/Interpolation

- Copying algorithm
  - Replace a damaged MB by the spatially corresponding MB in the previous frame
  - Can produce adverse visual artifacts
- Motion-compensated error concealment
  - Better performance
  - Need motion information
- Chu & Leou 98: find the following candidate concealment blocks and choose the block with the smallest error function
  - The MC blocks obtained with the MVs of its neighboring blocks
  - The MC blocks obtained with the average or median of the MVs of its neighboring blocks
  - The average and median of the blocks obtained in set 1
  - All its undamaged/concealed neighboring blocks
  - The average and median blocks of blocks obtained in set 4
- Motion-compensated interpolation

Boundary/Side Matching for MV Recovery

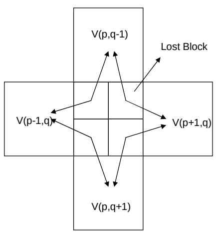


Temporal Estimation of Blocks with Missing MVs (Kossentini 98)



$$\hat{\mathbf{p}} = w_{u1}\mathbf{U1} + w_{u2}\mathbf{U2} + w_{l1}\mathbf{L1} + w_{l2}\mathbf{L2}$$

Overlapped Motion Compensation



Weighting Matrices

Recovered vectors

4	5	5	5	5	5	5	4
5	5	5	5	5	5	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	6	6	6	6	5	5
5	5	5	5	5	5	5	5
4	5	5	5	5	5	5	4

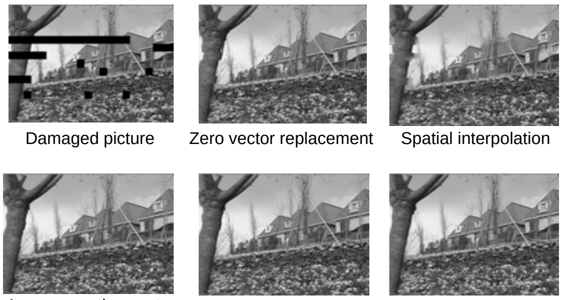
Top or Bottom

2	2	2	2	2	2	2	2
1	1	2	2	2	2	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	2	2	2	2	1	1
2	2	2	2	2	2	2	2

Left or Right

2	1	1	1	1	1	1	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	2	1	1	1	1	2	2
2	1	1	1	1	1	1	2

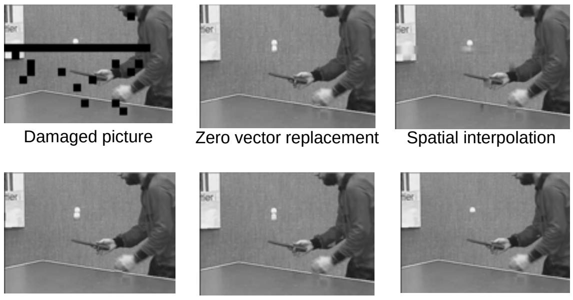
Subjective Picture Quality Comparison



Damaged picture | Zero vector replacement | Spatial interpolation

Average motion vector | SM+OMC | Genetic Algorithm

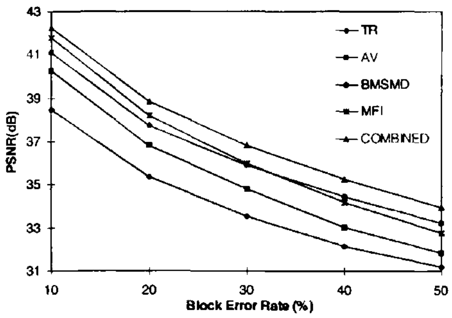
Subjective Picture Quality Comparison (Cont.)



Damaged picture | Zero vector replacement | Spatial interpolation

Average motion vector | SM+OMC | Genetic Algorithm

Performance Comparison of Various Concealment Schemes



Recovery of MVs and Coding Modes

The corrupted motion vectors can be interpolated from those in spatially and temporally adjacent blocks

A scheme for estimation coding modes for MPEG-2

Top MB (columns) / Bottom MB (rows)

MB Type	Forw	Intra
Forw	Forw	Forw
Intra	Forw	Intra

Top MB (columns) / Bottom MB (rows)

MB Type	Forw	Back	Inter	Intra
Forw	Forw	Inter	Inter	Forw
Back	Inter	Back	Inter	Back
Inter	Inter	Inter	Inter	Inter
Intra	Forw	Back	Inter	Intra

Conclusions

- Existing and future communication networks do not always guarantee error-free transmission; thus transmission of compressed video in such environments presents many challenges
- Error resilience coding methods allow reliable video communication
- Many of these error resilience methods are supported by video standards such as H.263 and MPEG-4
- Error resilience video coding is a relatively new area of research to explore