Error Resilience and Concealment in Video Coding

With focus on H.261/3

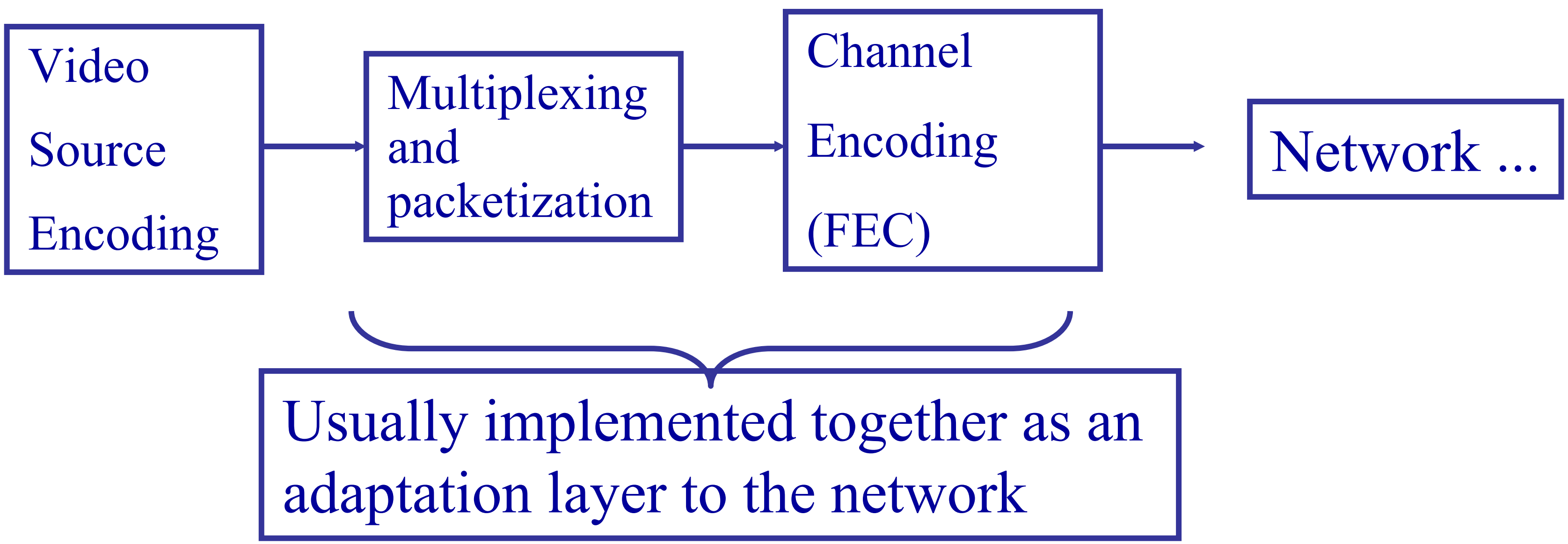
Resilience: ability to recover from errors

Concealment: error hiding

Nimrod Peleg
Update: Jan. 2003

Importance and Approach

A typical video transmission system:



<u>Note</u>: For a guaranteed error-free transmission no channel encoding needed !

Data Loss

- Data packets may be lost or corrupted due to :
  - Traffic congestion (צפיפות)
  - Impairment of the physical channel (bit errors)
- Immediate solution: retransmission (e.g. Automatic Repeat Request - ARQ)
- Immediate problem: Unacceptable delay for real-time applications (Broadcast …)

Transmission Error Types

- *Random bit errors*: imperfection of physical channel (bit inversion /insertion/ deletion).
  - Effect: depending on coding method - range from negligible to objectionable …
- *Erasure errors*: packet loss (in networks), burst errors (in storage) , short time system failure etc.
  - Effect: much more destructive

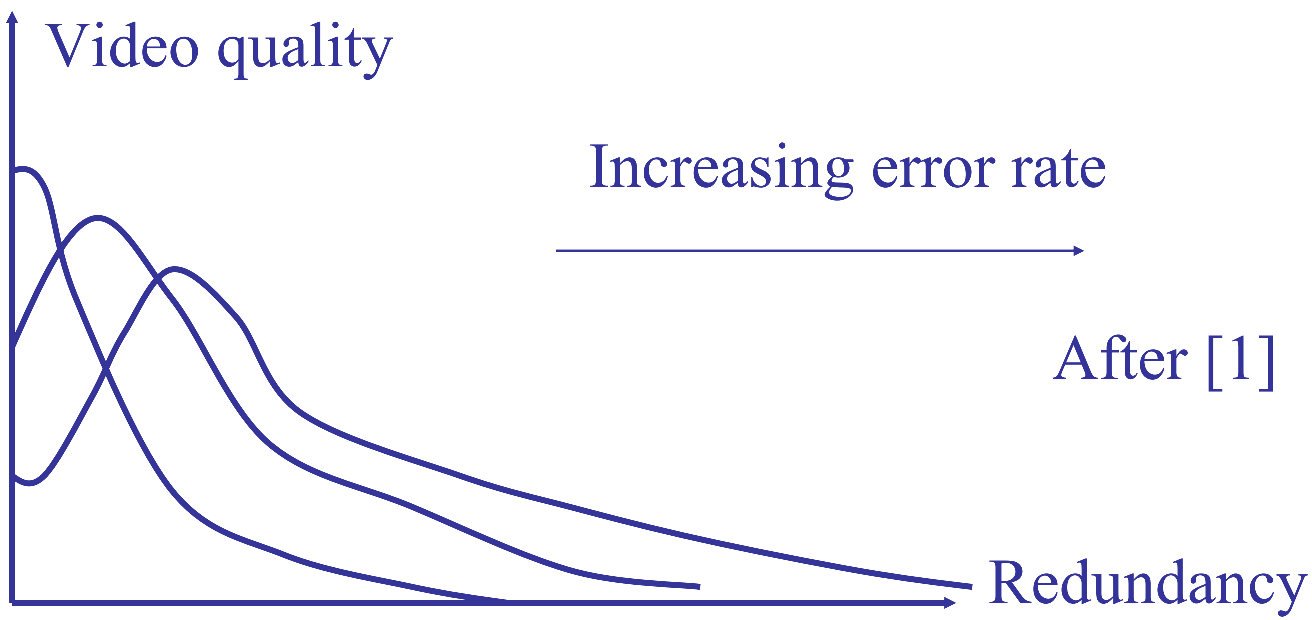
Since most video coding techniques are VLC, there is no much difference between the two types ...

Errors demo

Effect of transmission errors to a compressed video stream using the H.263 standard for a selected frame. Upper left: no transmission errors were present and the picture quality is as high as the bitrate allows; Upper right: 3% packet loss; lower left: 5% packet loss; lower right: 10% packet loss. From [3].

Formulation of the problem

- To enable error detection and concealment, *Redundancy* is needed.
- Could be at the : source coder, entropy coder or channel coder.



Video source model ...

- The (theoretic) "error concealment" problem could be some how formulated as source coder (+decoder) and channel coder (+decoder) pair, for a minimum signal distortion at the decoder.
- BUT…even if channel bandwidth and error characteristics are known, the "video source model" is very non-stationary (for natural video), and makes such solutions impractical.

Three main attitudes

- *Forward error concealment*: source coder and/or transport-control are designed to minimize the effect of transmission errors, with no need for concealment at the decoder.
  - FEC, joint source-channel coding, layered coding.

- *Postprocessing*: recovery from errors at the decoder side, by estimation and interpolation: no need for additional "side" information.
  - Spatial / Temporal smoothing, filtering etc.

Three main attitudes (Cont'd)

- *Interactive concealment*: Based on feedback from the decoder and the channel to the encoder, to change "something" in the encoding process, in order to overcome errors.
  - ARQ, and various modern techniques.

Which is better ?

- Reconstructed image quality
- Required delay (important for multipoint, and for two-way transmission)
- Side information overhead (redundancy …)
- Complexity

- priority of the above changes with application !

How to detect errors?

- *At the channel encoder/decoder*: mainly headers and payload fields.
  - Sequence number etc. (see H.223 for packet loss)
- *At the video decoder*: FEC - error correction bits allows detection and sometimes correction (See H.223 FEC)
  - In H.261: 18bit FEC per 493bit transport frame.
- Another option: exploit natural video characteristics: difference pixel/line/DCT values (spatial/temporal) etc.

Solutions for Real-Time

- Video encoding/decoding schemes that make the compressed bit-stream resilient to transmission errors
- A proper interfacing between the codec and the network, so the codec can adjust its operation based on the network condition

Summary:
the challenges in error concealment

Challenge #1: Coding Nature

<u>Predictive Coding:</u>

Due to spatial and temporal prediction, a single wrong recovered sample can lead to more errors in the same and following frames

<u>Variable Length Coding:</u>

Due to the nature of VLC, a single bit error can cause the decoder to lose synchronization

Effect of Transmission Errors

- Error Propagation:

- An error in a high dynamics region of a video

- An error in a static region of a video

Challenge #2: Time Varying

- Video source <u>and</u> network conditions are typically time-varying:

Hard (almost impossible) to derive optimal solution based on statistical models (of source and/or channel)

Challenge #3: Complexity

video sources have very high bit rate - high complexity operations are not acceptable for real-time applications.

Principal Solution: Redundancy

- In order to detect and correct errors, one must add redundancy into the bitstream.
- <u>Shanon Information Theory:</u>

One can separately design the source and channel coders, to achieve error free delivery, if the source is encoded by a rate below channel capacity.

- But…A FEC with <u>infinite delay</u> is needed !

Redundancy

- All error resilient encoding methods intentionally make the source coder less efficient than it can be.

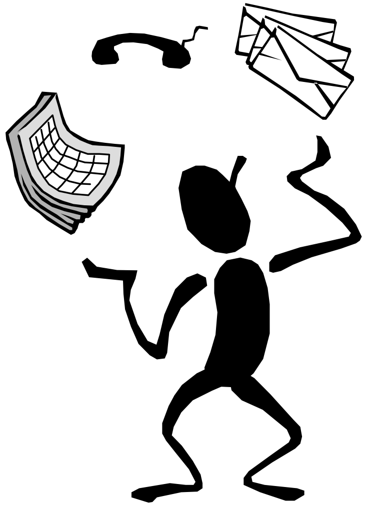
- This is usually accomplished by designing both the *predictive coding loop* and the *VLC coder* to limit the error propagation.

Error Concealment Methods

- Even when a sample (or a block of samples) is missing, the decoder can try to estimate them based on surrounding received samples (spatial and temporal adjacent samples)
- Example: periodical restarts of prediction process prevent error propagation
- Interleaving of adjacent samples in different packets is another typical solution

Error Concealment (Cont'd)

- Error concealment, in contrast to Error Resilient source coding, does not employ any additional bit rate, but adds computational complexity.



Codec and Network Cooperation

- If some bits are more important than others, they should be assigned a more robust QoS parameters.
- The network may also provide a feedback channel to let the encoder know which part of the reconstructed signal (at the decoder) is damaged, and not use it for further prediction.

Means against transmission errors: a short summary

- In source and channel encoders: a more resilient bitstream
- Detection at the decoder, and concealment of the effect of errors
- Interaction between source encoder and decoder in order to allow the encoder adapt itself according to channel conditions (as detected at the decoder side)

Networks and Video Delivery

Practical Networks

- A basic assumption: The application environment does not regularly allow re-transmission of damaged or lost video data
  - Real time constraints
  - Broadcast transmission
- Active standardization organizations:
  - ITU-T (H series)
  - MPEG WG-11 (ISO committee)
  - IEC

Transport Systems for Video

- Public Switch Telephone Network (PSTN)
- Integrated Service Digital Network (ISDN)
  - Video-conferencing is primarily conducted utilizing PSTN & ISDN: very low BER
- Broadcast: Error free delivery of video within a certain service area
- Internet: Information can be lost during transmission.
  - Main applications: database retrieval, streaming

Wireless Video Communication

- 1st-Generation cellular systems
  - Analog technology, maximum rate of 14.4Kbps under VERY good channel conditions (long delay)
- 2nd-Gen. (Europe: GSM, US, Japan:CDMA)
  - Digital technology - higher rates.
  - Designed for voice data: higher rates accompanied with long delays - not useful for video
  - PHS (Japan): 4 slots of TDMA system provides 32Kbps unprotected channel, similar to DECT (Europe). Both limited to area near the transmitter.

3rd Generation Cellular Systems

- Mobile videoconferencing for business.
- A much greater bandwidth (new wireless technology and standards)
- Error conditions are still extremely severe: error prone transport system.

Error Characteristics of Mobile Channels

- AWGN (Additive White Gaussian Noise) model assumed as starting point, but not accurate enough for signal propagating near the ground.
- "Multipath" propagation cause random fluctuations in phase and amplitude of received signal ("mulitpath fading")
  - 2-G systems: 12-15dB drop in SNR: large variations in BER performance.
  - A "bursty" noise, depends on terrain type (urban, mountains …) and mobile speed

Error Characteristics of Wired Channels

- Circuit-Switched wire-line (POTS) BER is around 10^{-6} in worst case.
  - Applies to 28.8Kbps modem (V.34)
  - Causes only moderate disruption to video stream
  - Modem protocols provides further error protection, leading to practically “near zero” errors.
- Wired Packet-Switched networks suffer from packet loss due to network congestion.
  - TCP: reliable end-to-end connection (retransmission and heavy delay for loaded network)
  - UDP: unreliable, connectionless: no verification. Possible packet loss (after predefined time elapses -timeout).

Standard families for video transmission: ISDN

- ISDN Videophone (H.320)
  - Multiplex protocol (H.221)
  - Video coding standards (H.261, H.263)
  - Typical Video Bitrate: 64Kbps - 384 Kbps
  - Error free (no ER beyond MB refresh)
- H.320 is a protocol hierarchy used for video conference and telephony
- H.221 is the multiplex protocol, for practically error-free channel with fixed video bit rate

Standard families for video transmission: PSTN

- PSTN Videophone (H.324)
  - Multiplex protocol (H.223)
  - Video coding (H.263), Audio coding (G.723)
  - Typical Video Bitrate: 20Kbps
  - Packet size: ~100Bytes
  - Very few bit errors and packet loss
- Transmission over telephone networks and wire-line modems. Poor audio/video performance, but no ER needed since modem is trained and very low BER achieved.

Standard families …: Mobile (Wireless networks)

- Mobile Videophone (H.324 Wireless)
  - Multiplex protocol (H.223 with mobile extensions)
  - Video coding standards (H.263)
  - Typical Video Bitrate: 10Kbps - 300Kbps
  - Packet size: ~100Bytes
  - BER: 10^{-3} to 10^{-5}, losses of H.223 packets
- For mobile environment, Annex-C of H.324 provides:
  - A mandatory retransmission protocol (MSRP) for control information (adding sequence number eliminating ambiguity)
  - Setup procedure and optional dynamic changes of multiplex during call, allowing adaptive use of ER overhead (depending on channel conditions)

H.223: Multiplex for low bit rate

- Variable size MUX-PDUs (not available in H.221)
- Done by special 'high-level data link control' (HDLC)

Standard families … : Packet Network

- Videophone over packet network (H.323)
  - Multiplex protocol (H.225 / RTP/UDP/IP)
  - Video coding standards (H.261, H.263, MPEG-2)
  - Typical Video Bitrate: 10Kbps - 1000Kbps
  - Packet size: $\leq$ 1500 Bytes
  - BER:0, 0-30% packets losses
- Real-time transmission on private IP networks or Internet.

Standard families … : Cable/Satellite

- Cable/Satellite TV
  - Multiplex protocol (H.222)
  - Video coding standards (MPEG-2)
  - Typical Video Bitrate: 6-12 Mbps
  - Almost error free
- Used by many TV applications.
- Channel coders and MPEG-2 transport layer ensure almost error-free environment.
- No ER needed beyond.

Standard families… :Videoconference over ATM

- Videoconference over Native ATM (H.310,H.321)
  - Multiplex protocol (H.222)
  - Video coding standards (MPEG-2)
  - Typical Video Bitrate: 1-12 Mbps
  - Packet size: 53 Bytes (ATM cell)
  - Almost error free
- Much smaller use than expected a few years ago.
- Synchronization markers at the beginning of every cell is not effective.

Techniques for Error Resilience

For block based hybrid video encoding framework

Error Resilient Encoding: General Approach

- Transmission errors should not lead to unacceptable distortion in the reconstructed video.
- ER coders are typically less efficient: more bits for same quality (=redundancy).
- Design goal: maximum gain in error resilience with minimum redundancy.
- 3 basic techniques:
  - prevent error propagation
  - better error detection and concealment at the decoder
  - guarantee basic level of quality

I. ER encoding: robust entropy coding

- VLC is VERY sensitive to BER

1. Periodic resynchronization markers:

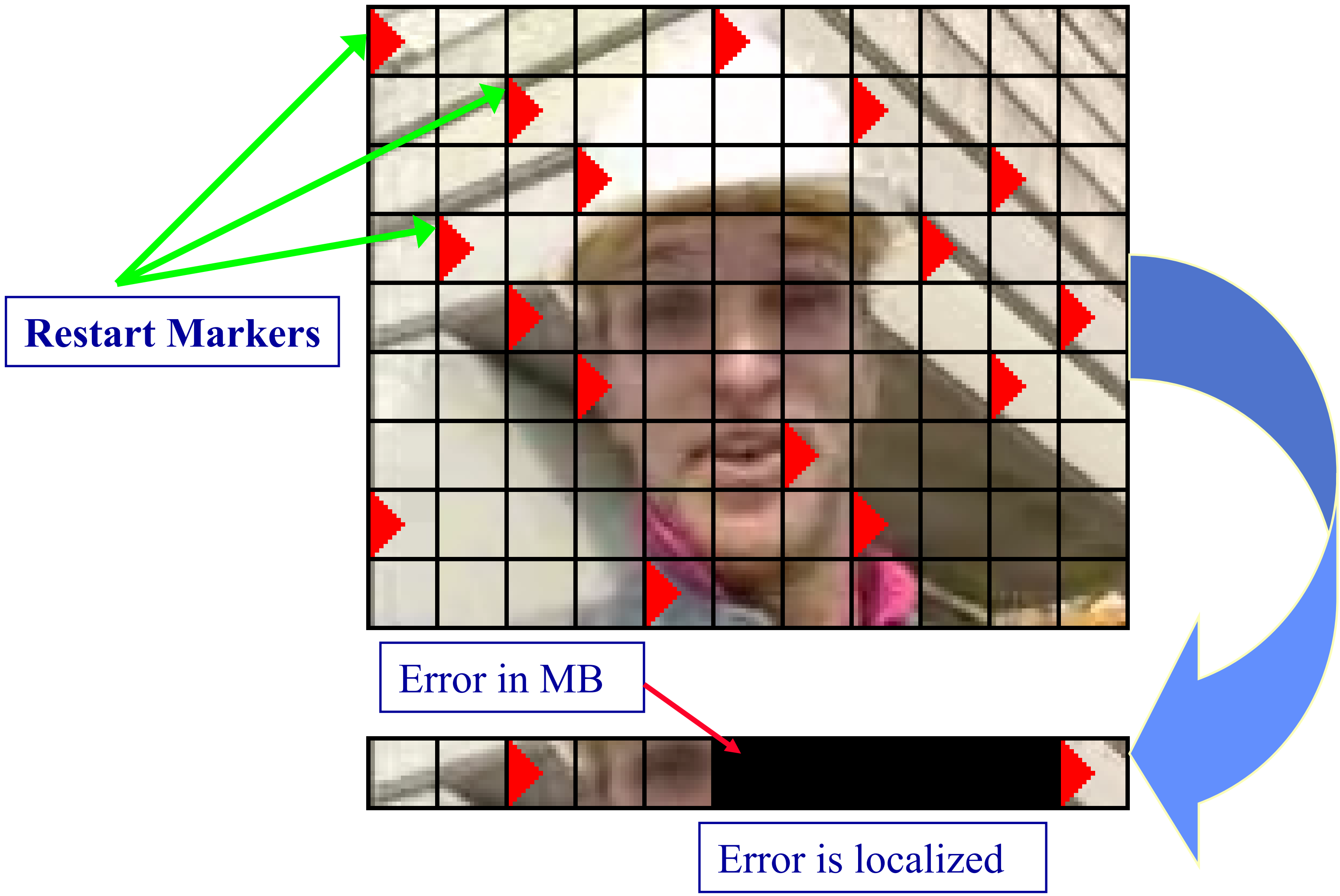
simple and effective.

☹ usually some header information added (temporal and spatial locations) : more bits

☹ Interrupts in-picture prediction (MV, DC coefficients etc.) ⇨ more bits for encoding

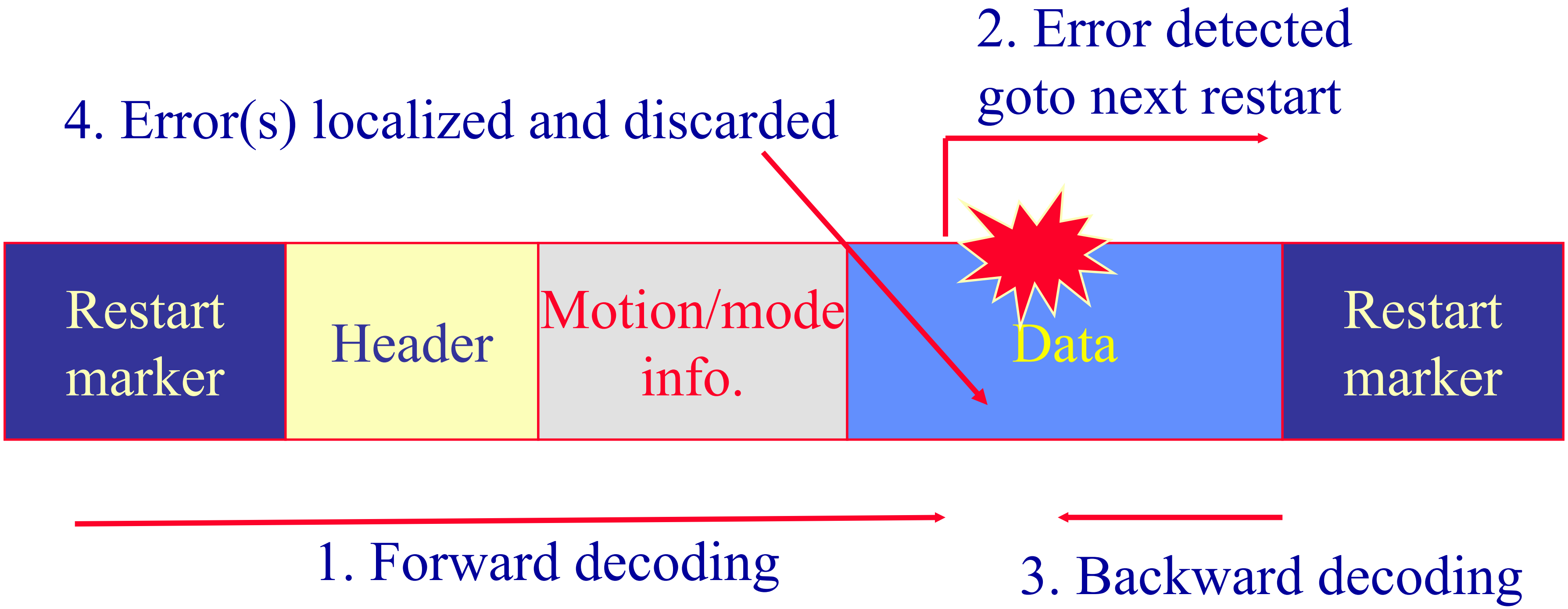
☺ More restarts: smaller regions affected by one error.

Restart codes



2. Reversible VLC (RVLC):

The decoder can not only decode the bits after the restart code, but also decode the bits before the next restart.



RVLC (Cont’d)

- RVLC techniques adopted by in both MPEG-4 and H.263, and so is insertion of restarts.
- No decrease in efficiency !
- In some ways it is more robust than regular VLC
- references:
  - Y.Takishima et al., ”Reversible Variable Length Coders”, *IEEE Proc. Trans. On Comm, Vol.43, No.2,3,4, 1995*
  - *ISO/IEC JTC 1/SC 29/WG 1 N1164 (JPEG2000)*

3. Error resilient prediction

- Temporal prediction is a major cause for sensitivity to errors: reconstructed frame (at the decoder) is different from the one assumed (in the encoder).
- Insertion of Intra data:
  - Insertion of Intra Macro-Blocks (or frames, but typically limited in real-time applications)
  - Intra MB number can be calculated according to channel conditions (as done in wireless comm.)
  - Intra MB placement can be random, periodical, in areas of high activity or according to feedback.

3. Error resilient prediction (cont'd)

- Independent segment prediction (ISD):

split data into several segments and make the prediction only within these areas.

- For both spatial and temporal.
- E.g.: separate even/off frames, GOB, Slice

- The term ISD is taken from H.263

4. Layered coding (LC)

- Also called scalable coding: base layer and one (or more) "enhancement" layer(s).
- Basic layer: acceptable quality
- Enhancement layers improve the basic quality: used for users with different bandwidth capacity.
- LC, together with different protection supplied to each layer (UDP: unequal error protection) serves as ER tool.

4. Layered coding (LC) (Cont’d)

- Not supported in any network protocol (possible in ATM)
- Possible implementation in application level !
  - Needs hierarchical, de-correlating decomposition.

- In MPEG-2 and H.263, optional enhancement layers are: Temporal, Spatial, SNR, Data Partitioning.

5. Multiple description coding (MDC)

- Source is also coded into several sub-steams (descriptions) which are correlated and have similar importance.
- Each description should provide basic quality, and together the provide an improved quality.
- Does not require special means from the network
  - E.g: Overlapping quantization, correlated prediction, correlating filter-banks, interleaved spatial-temporal sampling (Video redundancy coding in H.263)

II. Decoder Error Concealment

- Recovery (or estimation) of lost data.
- 3 types of losses:
  - texture information (pixel/DCT values)
  - Motion Vectors (MV) for P or B MB's.
  - MacroBlock's coding mode

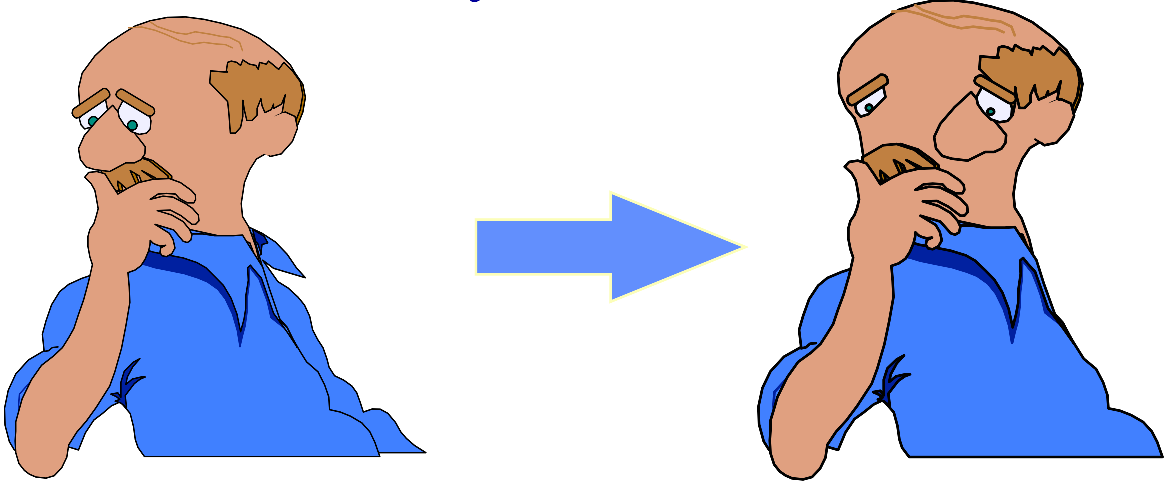
- Make use of smoothness property of natural images and motion.

1. Recovery of texture data

- 1.1 Motion compensated temporal prediction:

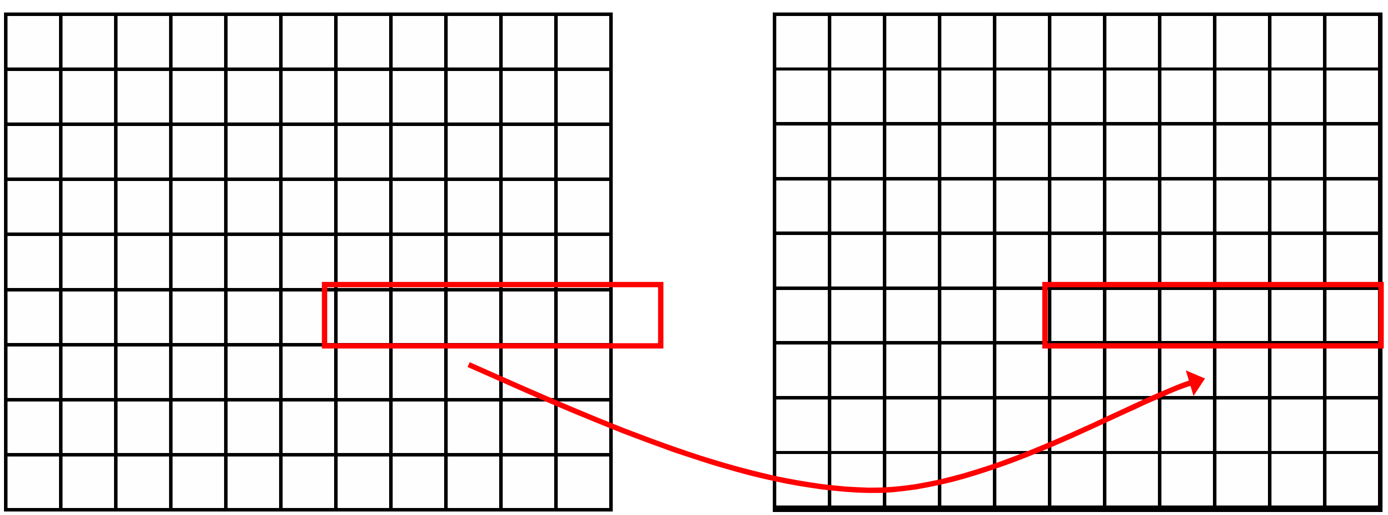
Copy corresponding MB from previous decoded frame, with same MV, estimated MV or Zero MV.

- All kinds need memory of last decoded frame.
- Estimation of MV needs more complexity.
- Could cause many artifacts.



Zero Motion Vectors Demo

This method is easy to implement and gives good results in a "low motion" regions, where all MVs are close to zero



Original (No Errors)

Zero Motion Vector

Method 1 example
(mjackson.mpg:
1 Intra 7 B
1 P, 7 B , etc.)

Intra

Bidirectional

Predicted

Damaged Intra picture causes low quality reconstruction since the following pictures are dependent on it.

Other phenomena...

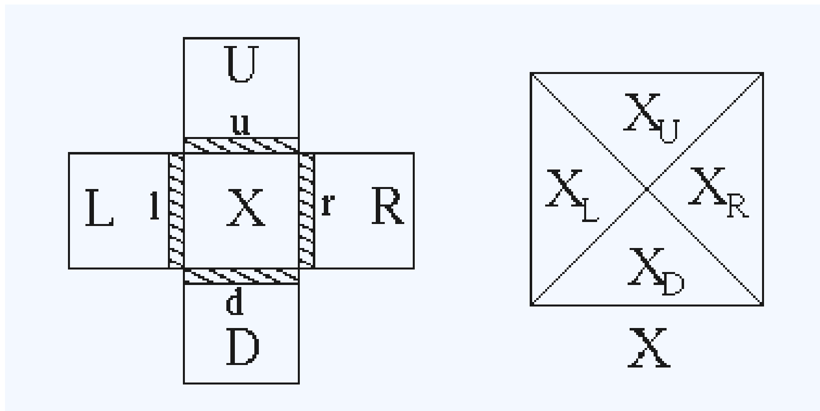
Other Techniques

- Spatial Concealment:

A smoothing operation that expresses each pixel of the missing block as a distance dependent linear combination of all the surrounding pixels.

- Spatial concealment, based on the current frame, gives comparative good (smoother) results in areas where there is motion.

Spatial Concealment



The missing block X, its adjacent blocks U,D,L,R, and the corresponding surrounding columns u,d,l,r

The four quadrants of X.

Spatial Concealment (Cont'd)

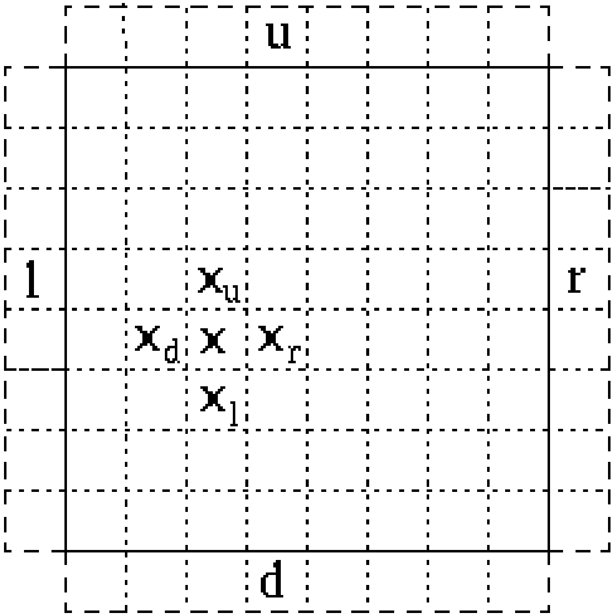
- If we detect motion in one of the adjacent blocks U,D,L,R, we mask the corresponding quadrant of X using a smooth estimate X^S of X computed

based on the

adjacent vectors

u,d,l,r.

$X=(x_u + x_d + x_l + x_r)/4$

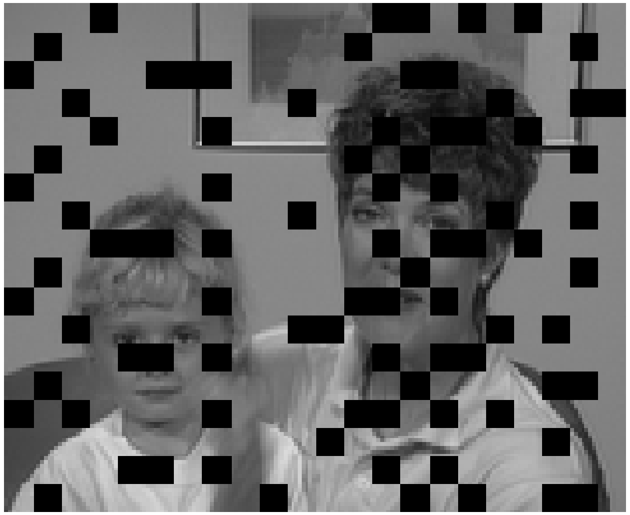


Temporal Concealment

- If there is no significant motion in that area of X, we replace that quadrant of X with the corresponding quadrant of the block P.

Mother & Daughter Example

Original Image



Damaged Image

Spatial Concealment

A Hybrid Technique

H.26x Critical Parameters

- Protection of most important information:
  - MBA: Address relative to previous MB. A lost MB will result wrong Addresses for all successive MB's (fixed offset...)

Critical Parameters... (Cont'd)

- Quant: GQuant Defines quantization step for whole GOB. MQuant does it in a MB.

If a MB with MQuant update is lost - a wrong quantization for subsequent MB's till a new MQuant comes

Critical Parameters... (Cont'd)

- MVD: Motion vectors are differential coded, so a lost MB (which includes MV) will result wrong MV decoding for the presceding MB's (fixed offset) till a new absolute MV

Packets and GOB's

- If a packet includes an integer number of GOBs, than in case of a lost packet - there will be no problem to decode it.
- Of course, there will be visual problems...
- In case of lost GOB - we can insert a "dummy" GOB just for bitstream legality
- Problem: what about small packets ? (ATM..)

Method 1: Inserting Empty GOB

- Pre-Decoder discovers a ‘missing’ GOB - it inserts an empty one (including only header)
- Effect on current image: ‘freezing’ in the area of the missing GOB (other GOBs move !)
- Effect on successive images: coming images include MV relative to a distorted image, and carry the error with them
- Needs no memory, minimal computation, and minimal quality...

Method 2: Copy Previous GOB

- Video Conference are ‘Head & Shoulders scenes: smooth movement. Using previous MV exploits movement continuity
- Effect: compensation done to previous image is also copied - very bad when lots of details and small movements
- Memory: 1 previous frame (compressed !)

Computation: minimal, Quality: scene dependent

Method 3: Copy Previous MVD's

- Copy only the MVD (without compensation)

prevents method 2 problem

- Technique: in case of Inter/Intra blocks (with no MC) - the pre-decoder skips them and causes image freezing

If the MB is MC - the pre-decodr copies MVD's (without decoing them)

- More computation needed (Huffman decoing and MVD identification)

Method 3 (Cont'd)

- <u>Better results </u>(than methods 1,2) with one exception: In case of fast movement the copied MVD is not close to the real one

Method 4: Averaging MVD's

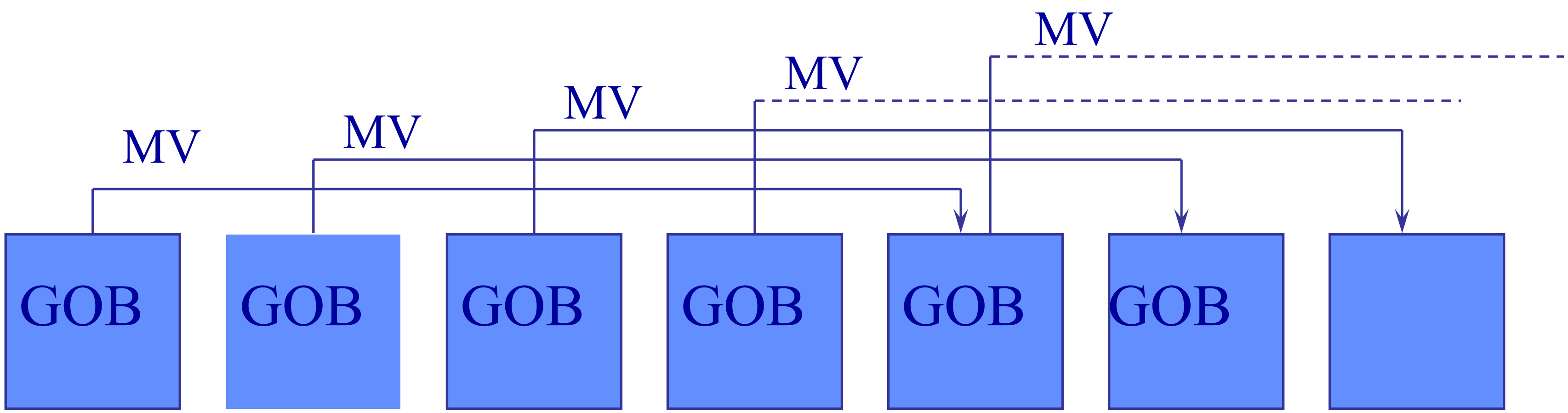
- Similar to method 3, but takes average MVD from previous and successive images
- Better results (not much !) , but needs 1 frame delay and 3 images memory buffer.
- Computation: large...

Method 5: MVD Prediction

- To avoid the delay of method 4, we can make linear extrapolation of 2 previous images and predict the new MVD
- Less quality (prediction from far image !)

Method 6: Data Partitioning

- Duplicate MVD data by taking them out of the code and adding them to far packet (about 2-6% overhead)



Assume GOB equals Packet !

Method 6 (Cont'd)

- Computation: same as method 3
- Memory: only for MVD of N GOB's

Original
(No Errors)

Data Partitioning
Emulation

Reference Literature

[1] Y.Wang and Q.Zhu, "Error control and concealment for video communication: A review", *Proc. IEEE, Vol.86, May 1998*

[2] M.R.Banham and C.J.Brailean, "Video Coding Standards: error resilience and concealment", in *Signal Recovery Techniques for Image and Video Compression and Transmission*, A.K.Katsaggelos and N.P.Galatsanos, Eds.Norwell, MA: Kluwer, 1998

[3] Y.Wang et al., "Error resilient video coding techniques", IEEE Signal Processing Mag.Vol.17, No.4, July 2000

Garden and Flower Example