Packet Video Workshop, New York

Using RFC2429 and H.263+

Stephan Wenger

stewe@cs.tu-berlin.de

Guy Côté

guyc@ece.ubc.ca

Structure

- Assumptions and Constraints
- System Design Overview
- Network aware H.263 Video Coder and Decoder
- Packetization
- De-Packetization and Error Concealment
- Experimental results

Assumptions and constraints

- Complete standard compliance
  - For H.323 systems:
    - H.225/RTP as transport and
    - H.26x for video
- Low Delay necessary for interactive applications
  - No I-frames (!)
- Best effort network: Internet

Assumptions and Constraints

- Standard-compliant H.263+
  - mechanisms introduced are now part of TMN11
- RTP environment
  - Internet MTU size: 1500 bytes
    - ⇒ one packet per picture possible @ 100 kbit/s, 10 fps
  - Overhead per packet: ~40 bytes
    - ⇒ use as few packets per picture as possible
  - Low Delay
- No feedback channel available

System Design overview 1/2

- Video is coded using GOB headers
  - or Slices headers, one slice per line of macroblocks
  - this leads to a variable number of bits per GOB/slice
- (Optionally) use loss-aware RD optimization procedure
  - Select the coding mode that yields the best RD tradeoffs for the macroblock weighted with the probability of a loss/non loss of that macroblock
  - loss probability is estimated out of the packet loss rate (RTCP receiver reports)

System Design Overview 2/2

- Use GOB interleaving
  - even numbered GOBs into first packet
  - odd numbered GOBs into second packet
  - Overhead per picture is reasonable
    - 40 bytes for the additional packet headers
- Use motion vector error concealment (TCON model)
  - MV of a missing macroblock is taken from the spatially above macroblock

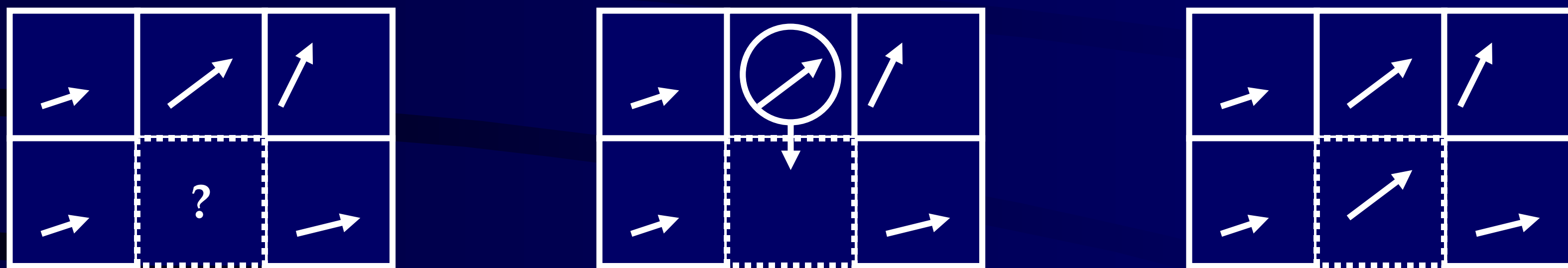
Network aware H.263+ video codec 1/3

- GOB headers are coded for each GOB
  - alternatively use out-of-order Slices
  - some people within Q.15 feel that this is more standard compliant with RFC2429

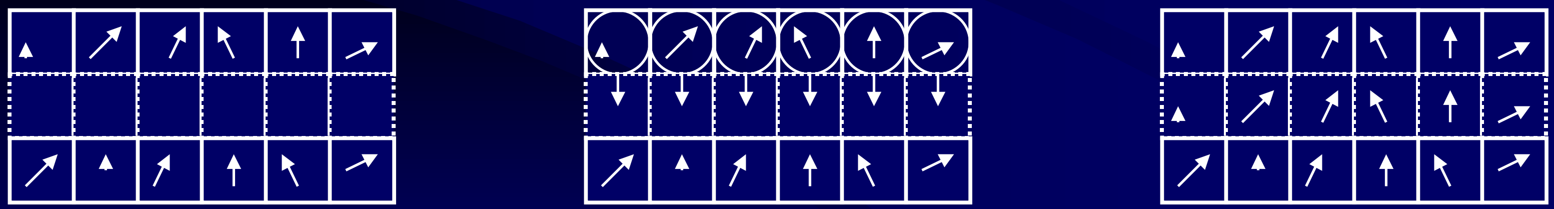
  - ~ 2%-10% overhead due to additional headers and in-picture prediction interruption
    - This will be further reduced during packetization

Network aware H.263+ video codec

- TCON Error Concealment (original concept)
  - take motion vector of missing MB from MB spatially above



- This also works great if a whole line of MBs is missing



Network aware H.263+ video codec 3/3

- Use loss-aware RD optimization
  - usual RD optimization: minimize Lagrangian $J=D+\lambda R$ for different coding modes
    - D: distortion, R: rate, $\lambda=0.85*(Q/2)^2$
  - loss aware RD optimization calculates two Distortion values
    - Dq: “usual” distortion (caused by quantization errors
    - Dc: distortion yield when MB is lost and concealed
      - this implies that the encoder knows the decoder’s EC mechanism
  - those values are wheighted with the loss probability p and the resulting Lagrangian is minimized:
    - $J=(1-p)*Dq + p*Dc + \lambda R$

Packetization 1/2

- (At least) two packets per picture
  - Packet 1 contains all even GOBs
  - Packet 2 contains a redundant copy of the picture header and all odd GOBs
    - Packetization overhead is ~44 bytes for additional IP/UDP/RTP headers
    - RFC2429 headers pay for themselves (GBSC, SSC are not coded)
- Considering an MTU size of 1500 bytes, this scheme allows
  - 240 kbit/s @ 10fps or 720 kbit/s @ 30 fps (assuming similar picture sizes)
  - For higher data rates or smaller MTU sizes the scheme can be extended using more than 2 packets per picture

Packetization 2/2

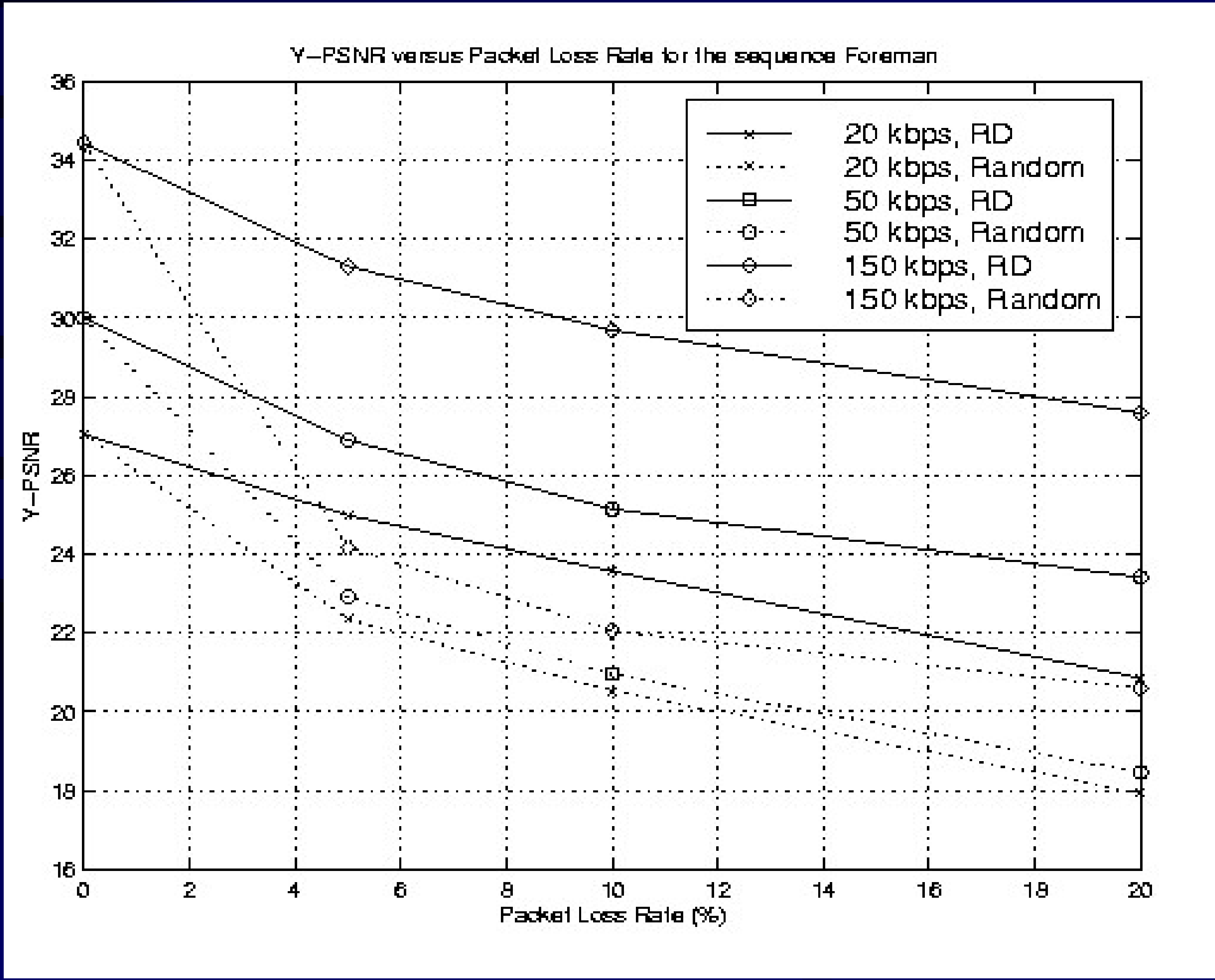
De-Packetization

- Collect all packets belonging to one picture
  - RTP provides the means for identification (timestamp, marker-bit)
- If all packets are available, decode directly
- If even packet is missing, then decode all even GOBs and conceal all odd GOBs
- If odd packet is missing, then use redundant picture header and odd GOB to start decoding
  - GOB 1 stays as from the last reconstruction
  - conceal all other odd GOBs

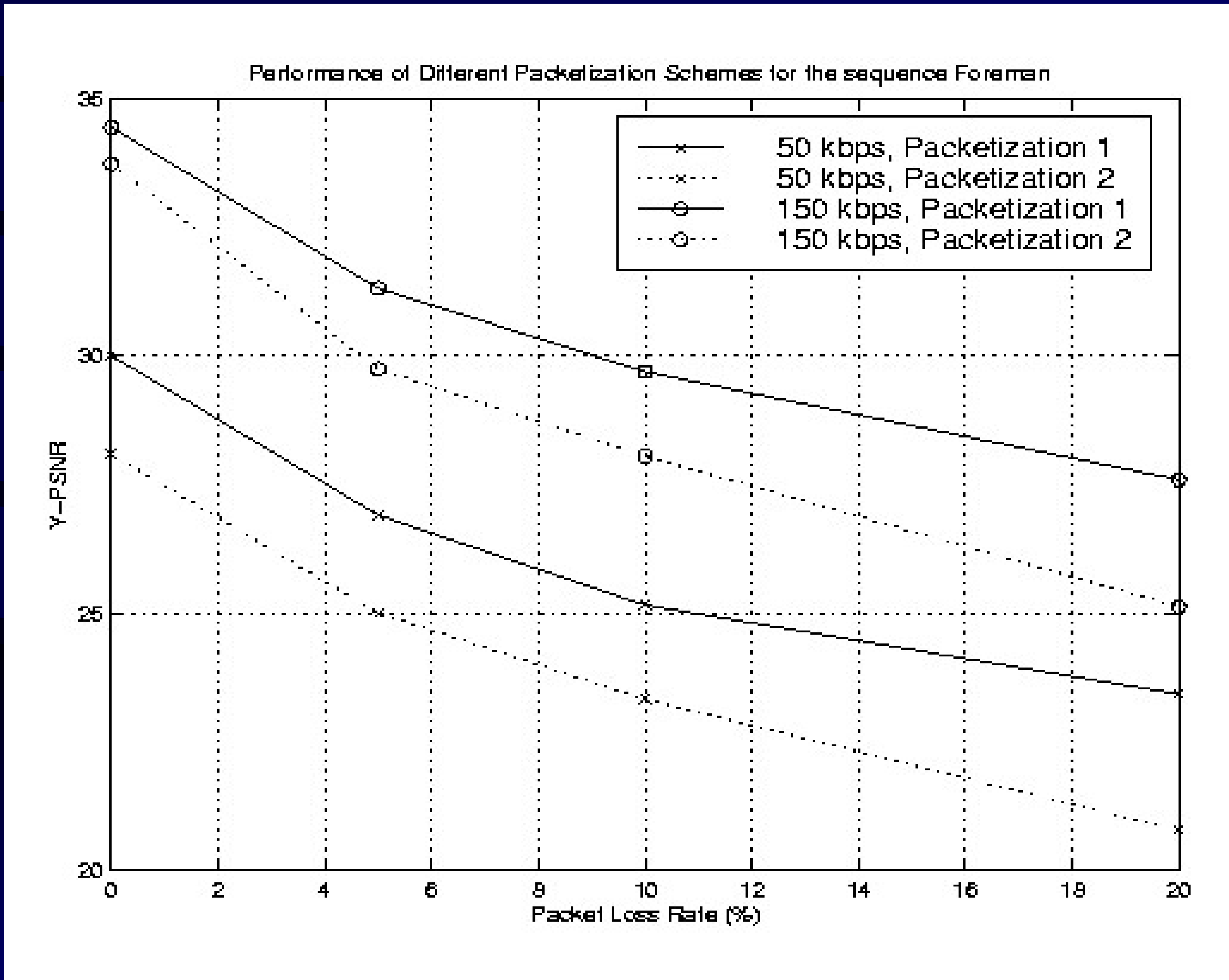
Experimental results 1/3

Transport bitrate	Total bitrate available for packet video	Pack.-Scheme	Packetization overhead @ 10 fps and QCIF	Bitrate for H.263+ video	PSNR at 0% PLR	PSNR at 20% PLR
Modem, 33 kbps	**20**	**1**	**6.4**	**13.6**	**27.1**	**20.9**
		2	**28.8**	**N/A**	**N/A**	**N/A**
ISDN, 64 kbps	**50**	**1**	**6.4**	**43.6**	**30.0**	**23.6**
		2	**28.8**	**21.2**	**28.1**	**20.7**
LAN, >150 kbps	**150**	**1**	**6.4**	**143.6**	**34.4**	**27.6**
		2	**28.8**	**121.2**	**33.7**	**25.1**

Experimental Results 2/3



Experimental Results 3/3



Conclusion

- The use of
  - state-of-the-art video coding (RD-optimized H.263+),
  - intelligent packetization schemes (RFC2429),
  - and a (simple) error concealment scheme
- yields exceptionable performance through
  - low packetization overhead,
  - use of all received data for decoding,
  - concealment of all unreceived parts of the picture to be reconstructed, and
  - low delay.