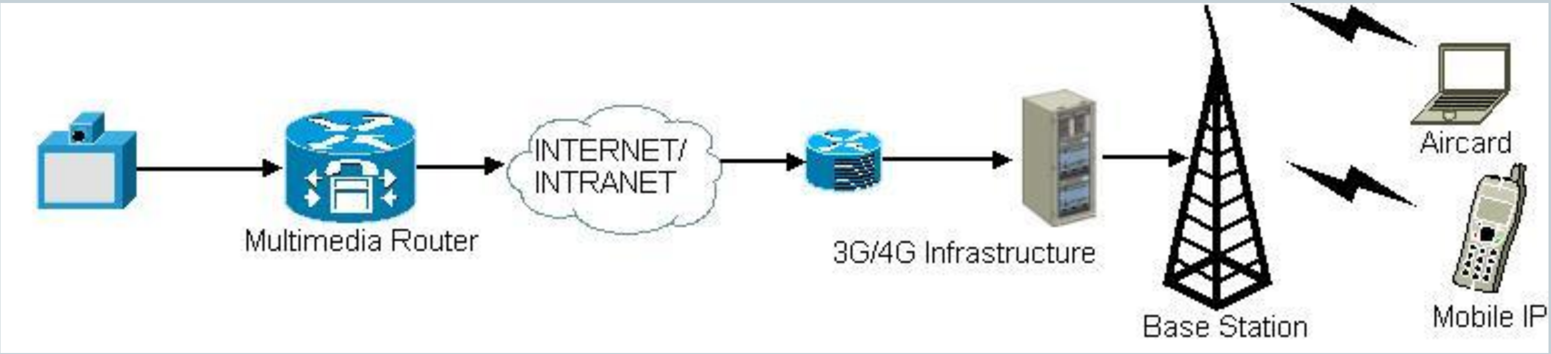
Error concealment techniques in H.264 video transmission over wireless networks

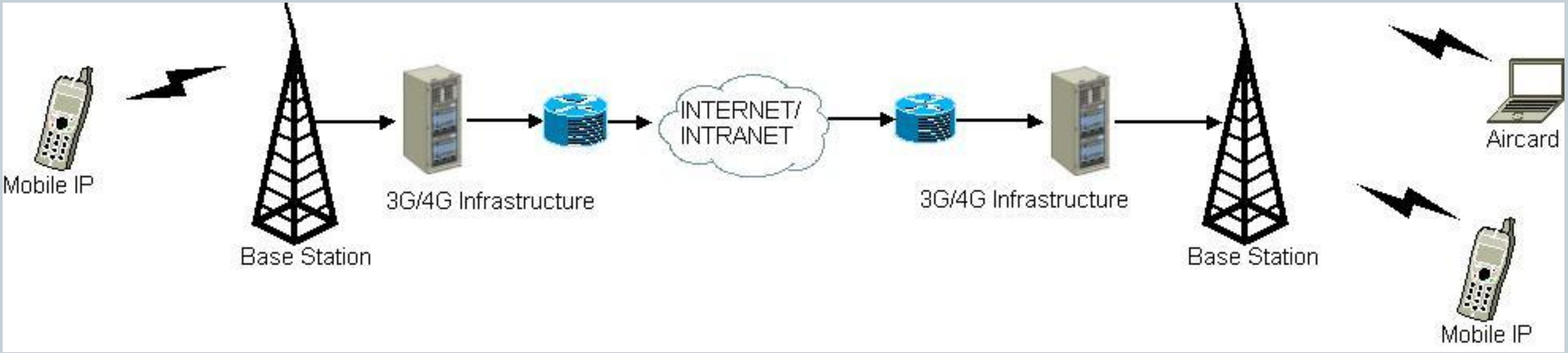
MULTIMEDIA PROCESSING (EE 5359)
SPRING 2011
DR. K. R. RAO
FINAL REPORT

Murtaza Mustafa Zaveri
murtaza.zaveri@mavs.uta.edu
(1000671952)

Typical 3G/4G Wireless Architecture



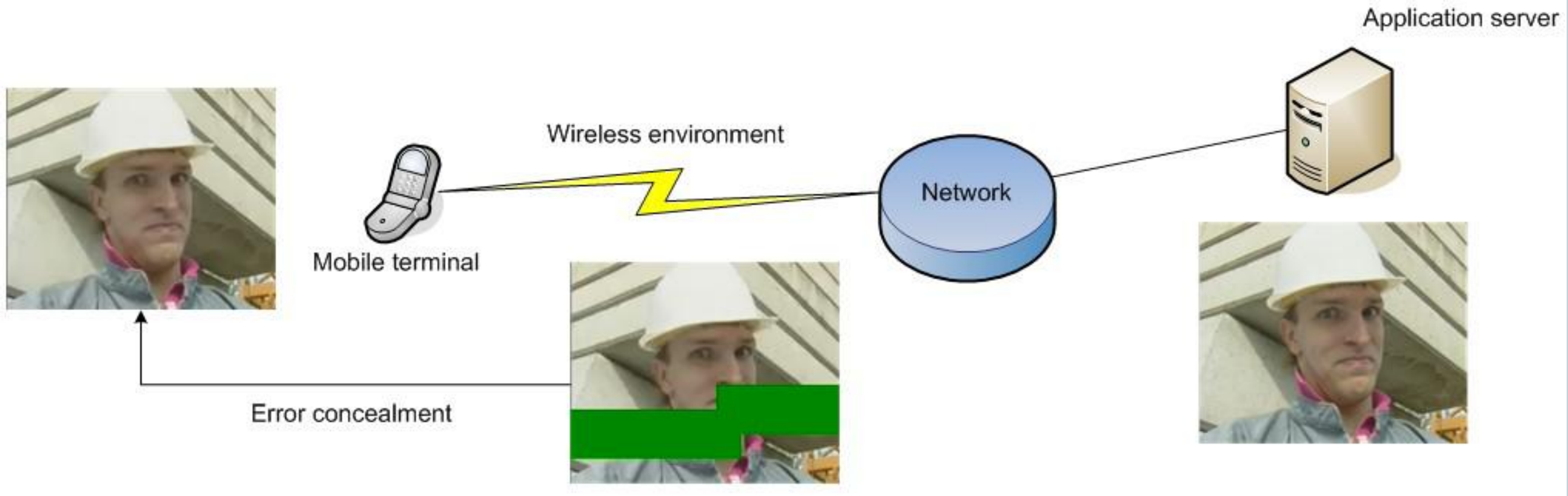
Land to Mobile[17]



Mobile to Mobile[17]

Video Transmission over Wireless Networks

- Major categories for video services over the wireless systems:
  - Multimedia messaging services (MMS)
  - Packet-switched pre-coded streaming service
  - Circuit-switched and packet-switched conversational services.
- When fading level or shadowing level is severe, up to 3% packets are lost



Typical situation in 3G/4G cellular telephony [20]

Error Propagation

- Imperfections in the communication channel, often result in packet loss, which in turn lead to frame loss or corrupted areas in the decoded frame.
- As H.264 employs predictive coding, this kind of corruption spreads spatio-temporally to the current and consecutive frames
- H.264 is thus susceptible to error propagation due to channel noise leading to a considerable degradation in the video quality [1]

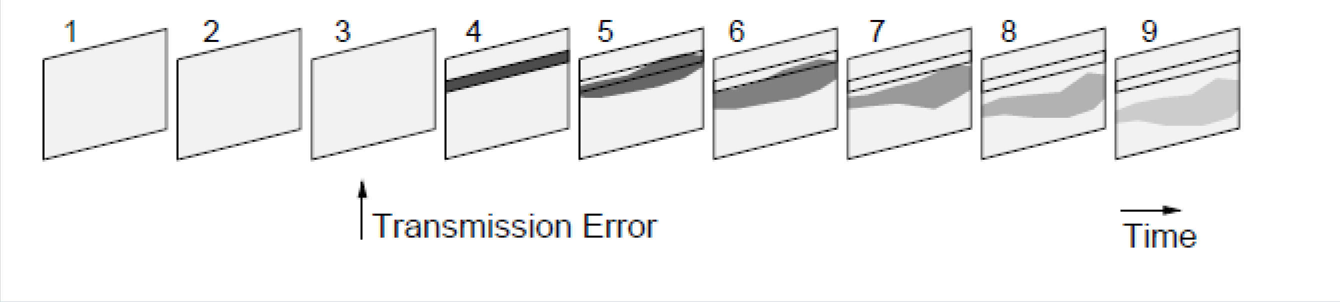


Illustration of spatio-temporal error propagation [14]

Error Propagation

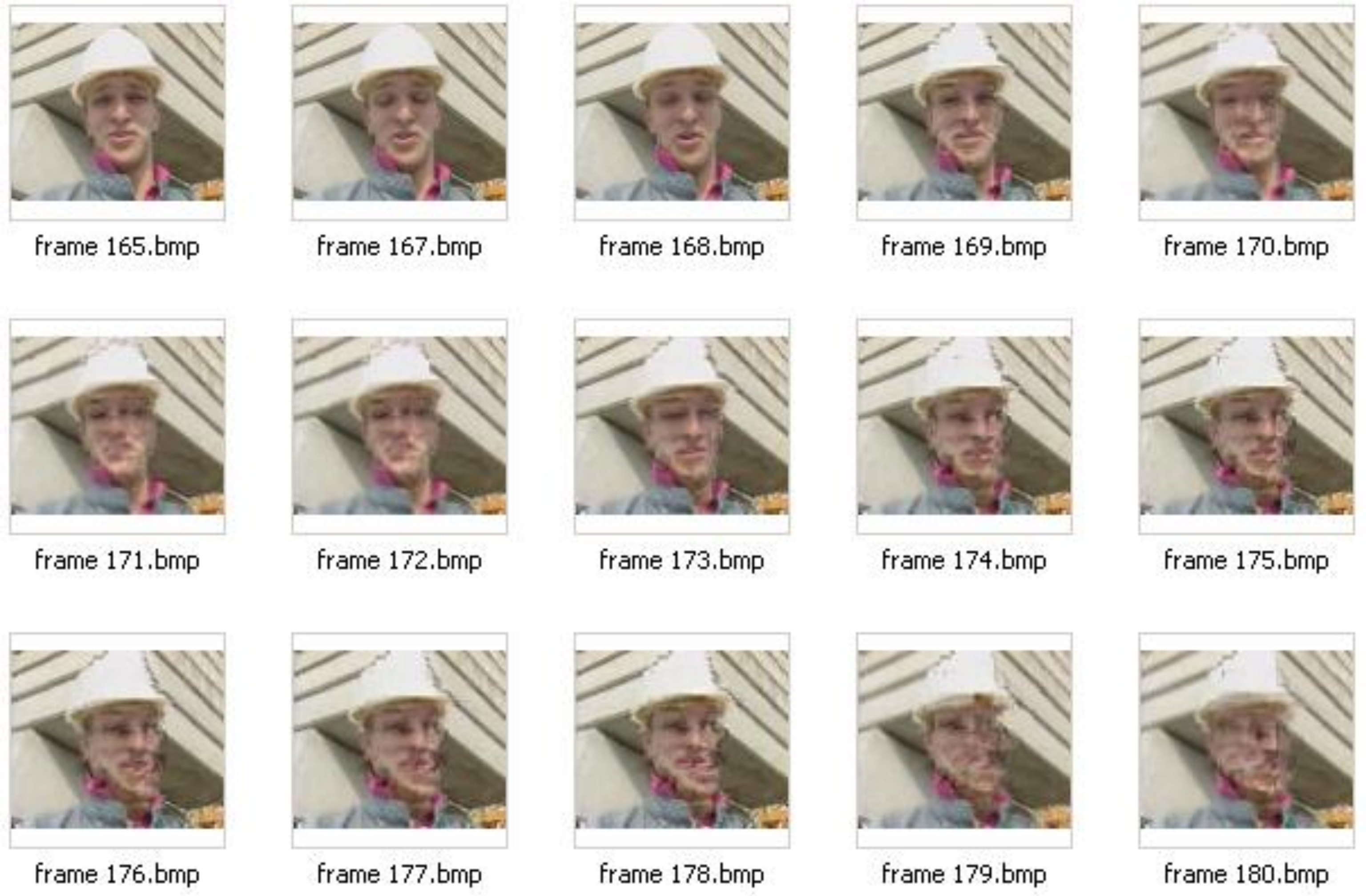
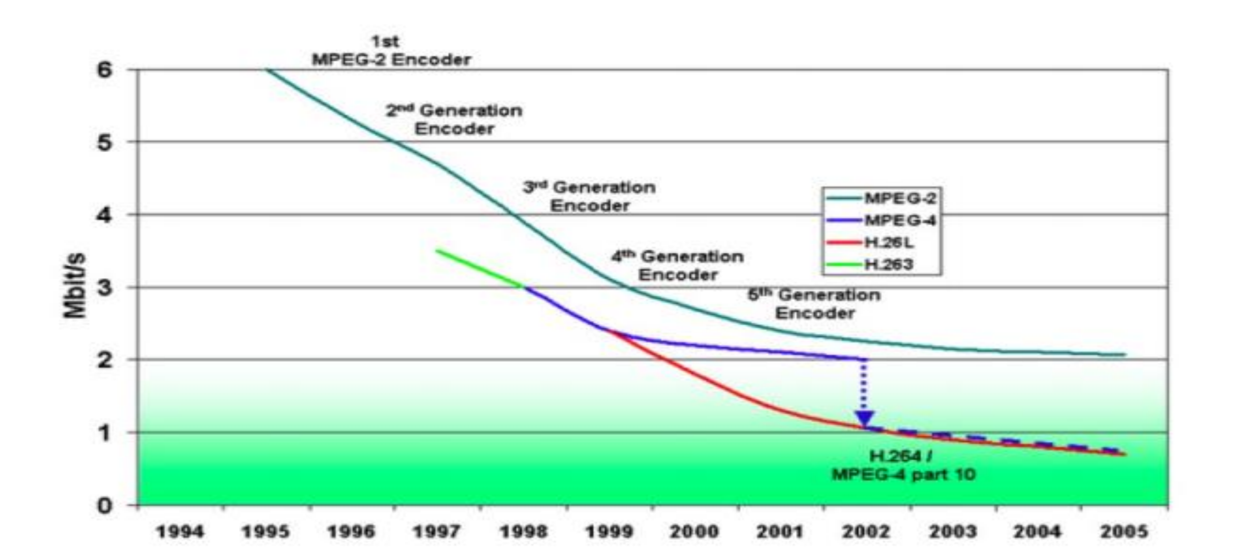


Illustration of error propagation [17]

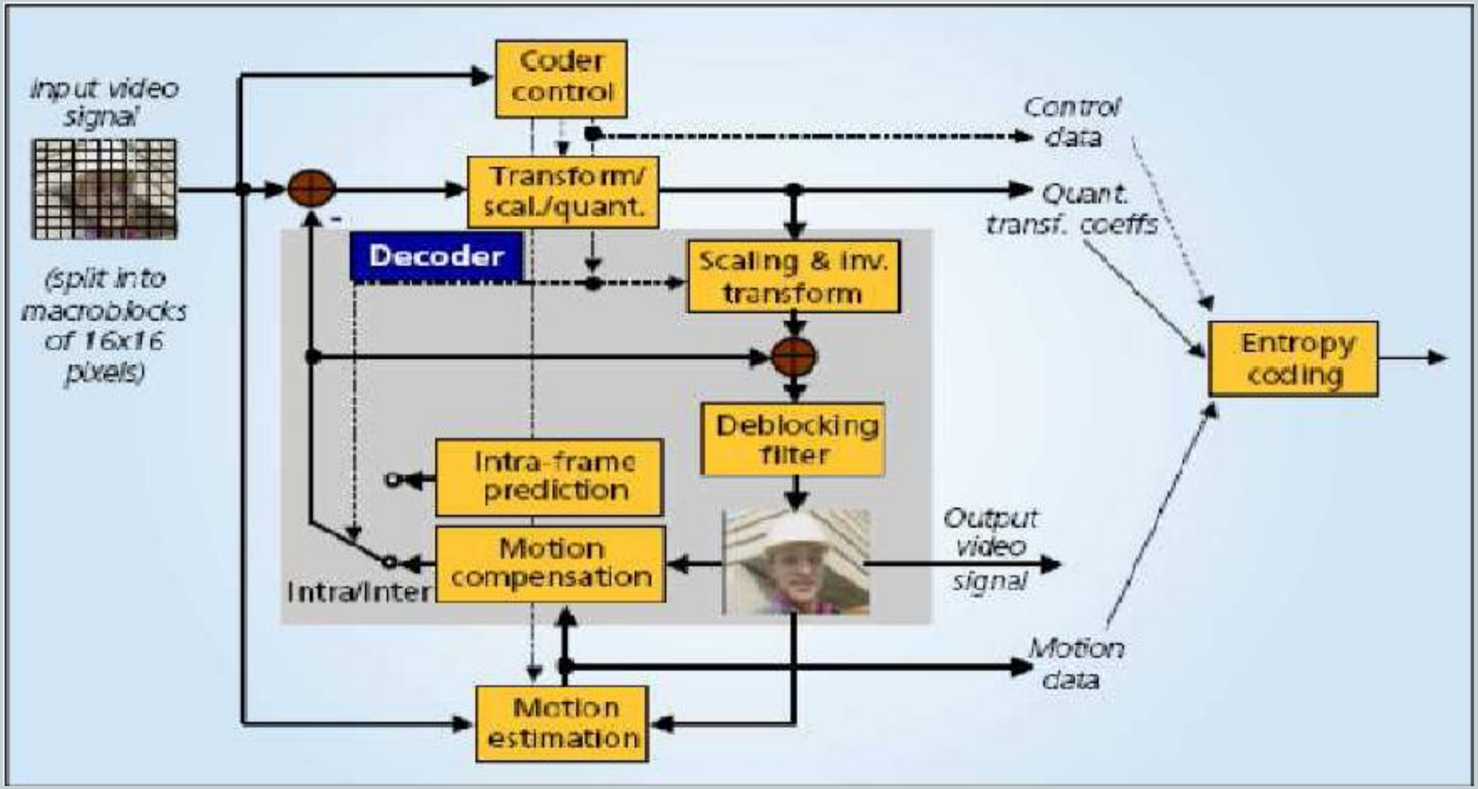
The H.264 Standard

- Newest entry in the series of international video coding standards.
- Developed by a Joint Video Team (JVT) consisting of experts from VCEG and MPEG
- Design provides the most current balance between the coding efficiency, implementation complexity, and cost
- Has improved coding efficiency by a factor of at least about two (on average) over MPEG-2 [6]



Position of H.264/MPEG-4 AVC standard [9]

The H.264 Standard (contd.)



The basic coding structure of H.264/AVC for a macroblock [7,14]

H.264 Encoder – Profiles

Seven prominent profiles :

- Baseline profile
- Main profile
- Extended profile
- High Profile
- High 10 Profile
- High 4:2:2 Profile
- High 4:4:4 Profile

H.264 Encoder – Profiles

- The Baseline profile was targeted at applications in which a minimum of computational complexity and a maximum of error robustness [8].

- The Main profile was aimed at applications that require a maximum of coding efficiency, with somewhat less emphasis on error robustness [8].

- The Extended profile was designed to provide a compromise between the Baseline and Main profile capabilities with an additional focus on the specific needs of video streaming applications, and further added robustness to errors and packet losses [8].

H.264 – Profiles

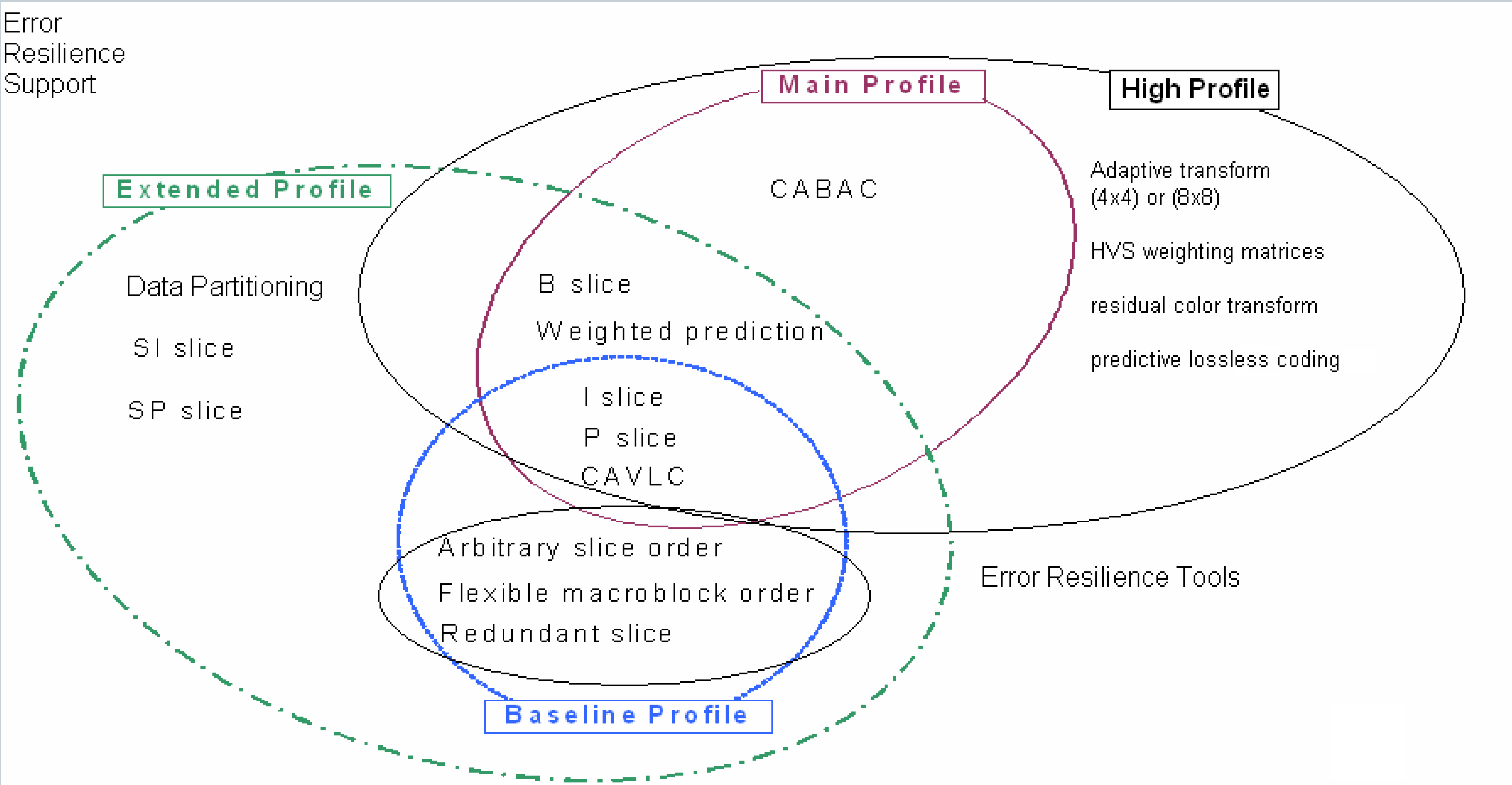


Illustration of H.264/MPEG4-AVC profiles [8]

H.264 Encoder – Profiles

The common coding parts for the profiles are listed below [14]:

- I slice (Intra-coded slice): coded by using prediction only from decoded samples within the same slice.
- P slice (Predictive-coded slice) : coded by using inter prediction from previously decoded reference pictures, using at most one motion vector and reference index to predict the sample values of each block.
- CAVLC (Context-based Adaptive Variable Length Coding) for entropy coding.

H.264 Encoder baseline profile

Primarily designed for –

- Low processing power platforms
- Error prone transmission environments

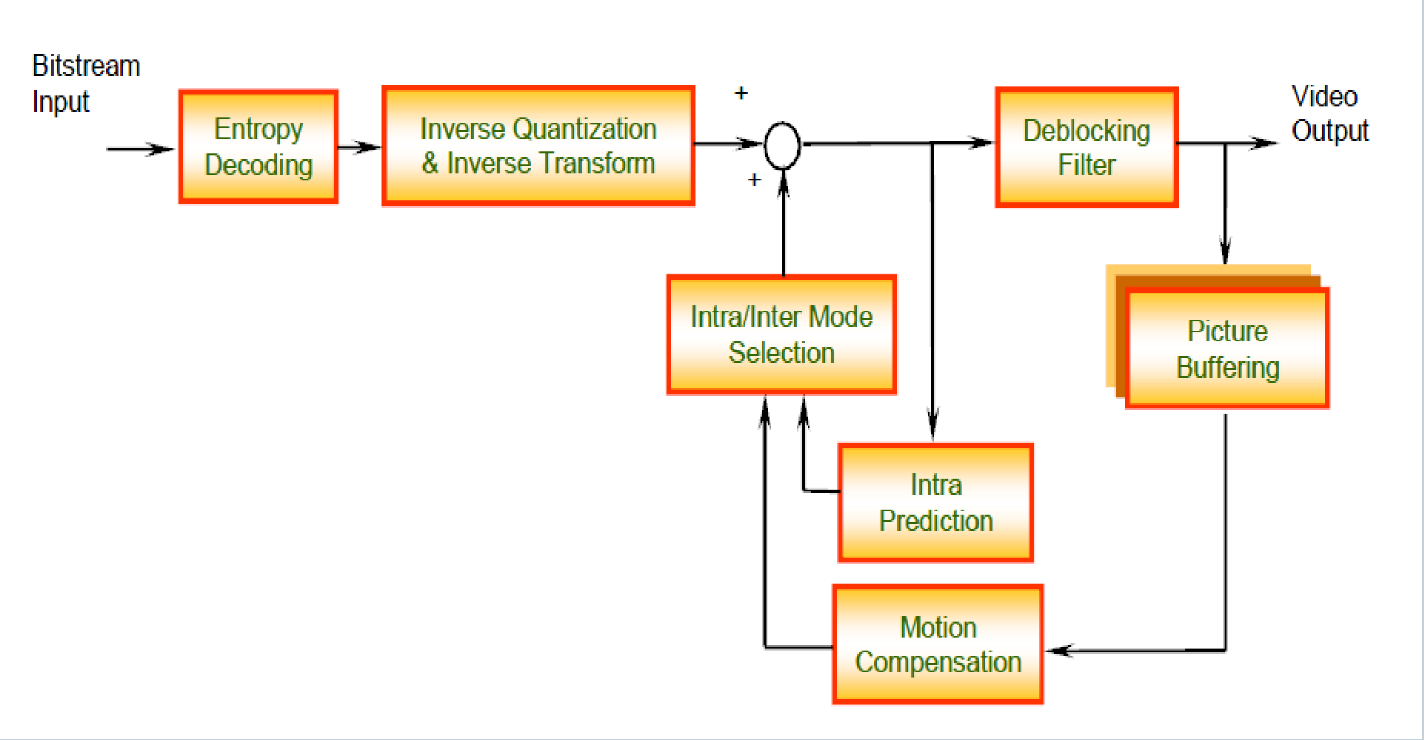
Features –

- Low on coding efficiency
- I- and P- slice coding
- Enhanced error resilience coding such as flexible macroblock ordering.
- Context adaptive variable length coding (CAVLC).

Features not included in baseline profile –

- B- slices, SI- or SP- slices.
- Interlace coding tools.
- Context adaptive binary arithmetic coding (CABAC).

The H.264 Standard (contd.)

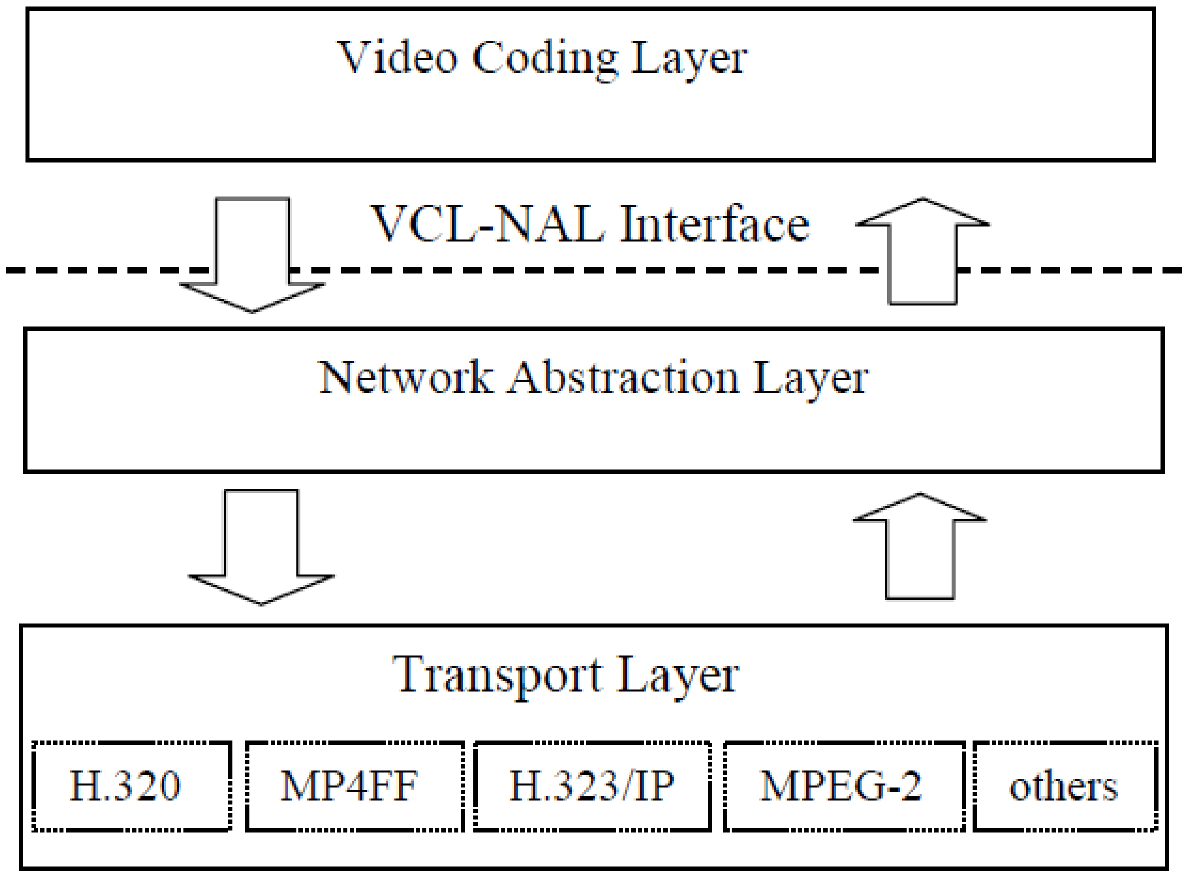


Block diagram of a H.264 decoder[14]

Error Resilience

- For better coding efficiency, the H.264 standard gives strong emphasis to error resiliency and the adaptability to various networks
- H.264/AVC has adopted a two-layer structure design containing a video coding layer (VCL), which is designed to obtain highly compressed video data, and a network abstraction layer (NAL), which formats the VCL data and adds corresponding header information for adaptation to various transportation protocols or storage media [14]

Error Resilience (contd.)



VCL/NAL layers of H.264 [14]

Error Resilient Video Coding

To handle the errors, the following stages are required in an error resilient decoder [14]:

- Error detection and localization
- Resynchronization
- Error concealment

Assumptions used in the implementation of error concealment techniques:

- The missing part of a video content is limited to one macroblock
- The location of the missing macroblock is known

Error Resilient Video Coding (contd.)

- *Error detection* is done with the help of video syntax and/or semantics
- When violation of video semantics/syntax is observed, decoder reports an error, and tries to *resynchronize* at the next start code
- H.264 test model is based on the assumption that the data recovery does not bring a significant advantage to the reconstructed frames. Therefore, the corrupted packets are simply discarded and the lost region of video frame is concealed
- The *error concealment* schemes try to minimize the visual artifacts due to errors

Error Resilient Video Coding

Generation of errors[16]

- Modify the function “decode_one_slice” which can be found in the “image.c” file of the decoder source code.
- The purpose of this function is as its name suggests: to decode one slice.
- The function takes a slice, reads macroblocks successively from the bitstream and decodes them by calling the function “decode_one_macroblock”. When the flag “end of slice” gets the value “TRUE” it exits the function until the next slice needs to be decoded.
- When a new slice is detected and it needs to be treated as erroneous, instead of calling the function ”decode one macroblock”, the selected error concealment method is used to conceal the slice.

Error Concealment

- The main task of error concealment is to replace missing parts of the video content by previously decoded parts of the video sequence in order to eliminate or reduce the visual effects of errors caused by corrupted areas in the decoded frame [2]
- Error concealment exploits the spatial and temporal correlations between the neighboring image parts within the same frame or from the past and future frames
- Typical parameters used to evaluate the quality of reconstruction include: peak signal to noise ratio (PSNR) and structural similarity index metric (SSIM)[18]

Error Concealment

Metric used: PSNR[1]

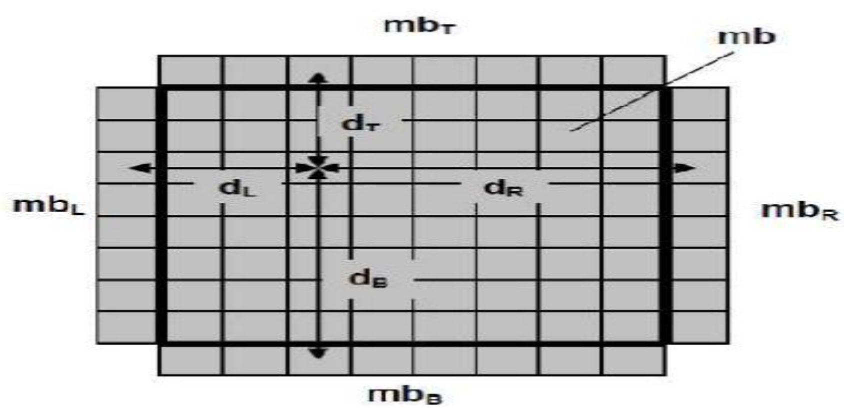
- It is common to evaluate the quality of reconstruction of a frame by analyzing its peak signal to noise ratio (PSNR).
- One way is to analyze is through a representation of a frame-by-frame versus PSNR graph.
- Another to give the information about PSNR in a table or graph for the overall PSNR of the sequence.
- The latter is used here as it gives a reliable result when all frames of the sequence are of similar character (spatial and temporal information)[16].

Spatial domain error concealment

Weighted Averaging

- Each missing pixel of the corrupted image part is interpolated from the intact surroundings pixels. [3]
- To implement spatial domain error concealment the pixel values within the damaged macroblock were interpolated from four next pixels in its four 1-pixel wide boundaries. This method is known as ‘weighted averaging’ [21]
- Missing pixel values can be recovered by calculating the average pixel values from the four pixels in the four 1-pixel wide boundaries of the damaged macroblock weighted by the distance between the missing pixel and the four macroblocks boundaries (upper, down, left and right boundaries).[22]

Spatial domain error concealment



$$mb(i,k) = \frac{1}{d_L + d_R + d_T + d_B}[d_R mb_L(i,2N) + d_L mb_R(i,l) + d_B mb_T(2N,k) + d_T mb_B(l,k) \quad [10]$$

where i,k=1,2,3..................,N.

d_L : distance between the interpolated pixel and the nearest pixel $mb_L = mb(i,0)$ in left boundary.

d_R : distance between the interpolated pixel and the nearest pixel $mb_R = mb(i, N+1)$ in right boundary.

d_T : distance between the interpolated pixel and the nearest pixel $mb_T = mb(0,j)$ in top boundary.

d_B : distance between the interpolated pixel and the nearest pixel $mb_B = mb(N+1,j)$ in bottom boundary.

NXN: Size of the block.

Temporal domain error concealment

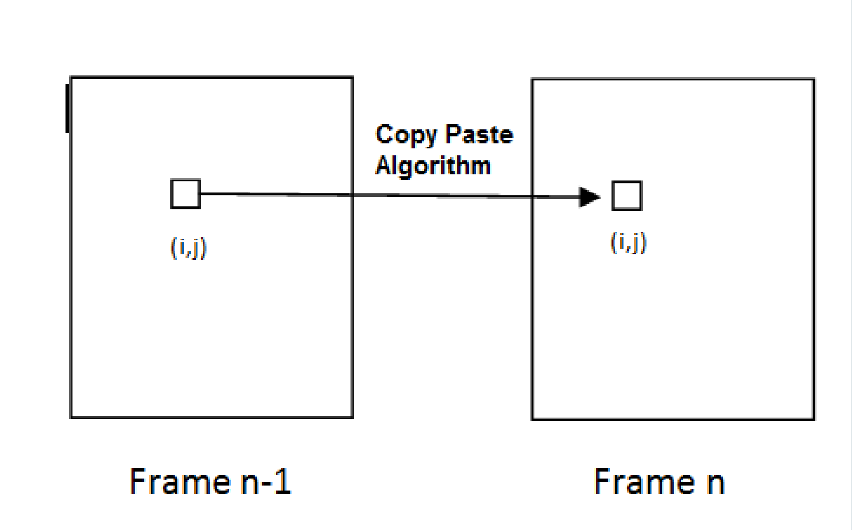
Copy-paste algorithm

- Replaces the missing image part with the spatially corresponding part inside a previously decoded frame, which has maximum correlation with the affected frame [4]
- It is easier to conceal linear movements in one direction because pictures can be predicted from previous frames (the scene is almost the same).
- If there are movements in many directions or scene cuts, finding a part of previous frame that is similar is going to be more difficult, or even impossible.
- The slower is the movement of the camera, the easier will be to conceal an error.

Temporal domain error concealment

Copy Paste Algorithm:

$$F_n(i,j) = F_{n-1}(i,j)$$

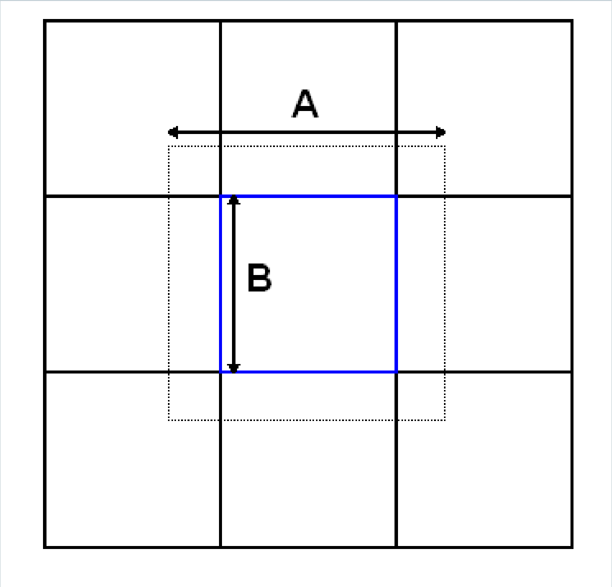


- This method only performs well for low motion sequences but its advantage lies in its low complexity[22]

Temporal domain error concealment

Boundary Matching Algorithm[16]

- Missing block of area B in the n^{th} frame F_n
- Area A is chosen in the $n\text{-}1^{th}$ frame F_{n-1} with its centre corresponding to the centre of B



Temporal domain error concealment

- A block to conceal block B is found in block A by using the sum of absolute differences(SAD) similarity metric
- The formula is as follows:

$$[\hat{x}, \hat{y}] = \arg \min_{x,y \in \mathbf{A}} \sum_{i,j \in \mathbf{B}} |\mathbf{F}_{n-1}(x+i, y+j) - \mathbf{B}(i,j)|.$$

where [^x, ^y] is the best match for (x, y) in B according to SAD

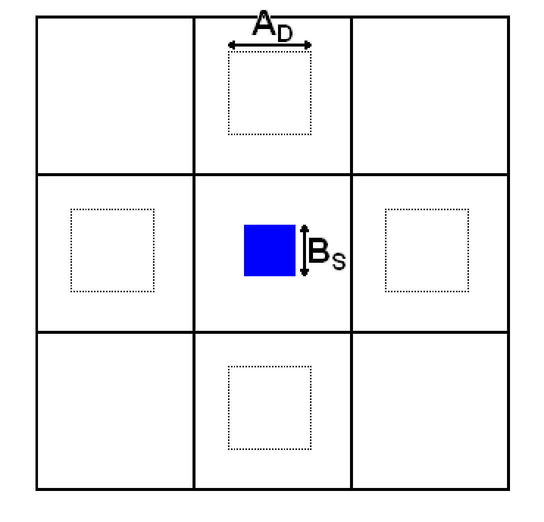
Temporal domain error concealment

Block Matching[16]

- It is a more refined way of using the SAD
- Search for the best match of the missing block (MB_D) among the neighboring top, bottom, left and right blocks

  where $\mathrm{MB}_D : D \in \{T, B, L, R\}$

- The final position of the best match is given by the average over the positions of the best matches



Temporal domain error concealment

- The equation for SAD is as follows:

$$[\hat{x}, \hat{y}]_D = \arg \min_{x,y \in \mathbf{A}_D} \sum_{i,j \in \mathbf{MB}_D} |\mathbf{F}_{n-1}(x+i, y+j) - \mathbf{MB}_D(i,j)|$$

- The equation for final positions is as follows:

$$\hat{x} = \frac{1}{M} \sum_D \hat{x}_D; \quad \hat{y} = \frac{1}{M} \sum_D \hat{y}_D.$$

CIF and QCIF formats

- Common Intermediate Format (CIF) and Quadrature Common Intermediate Format (QCIF) determine the resolution of the frame.

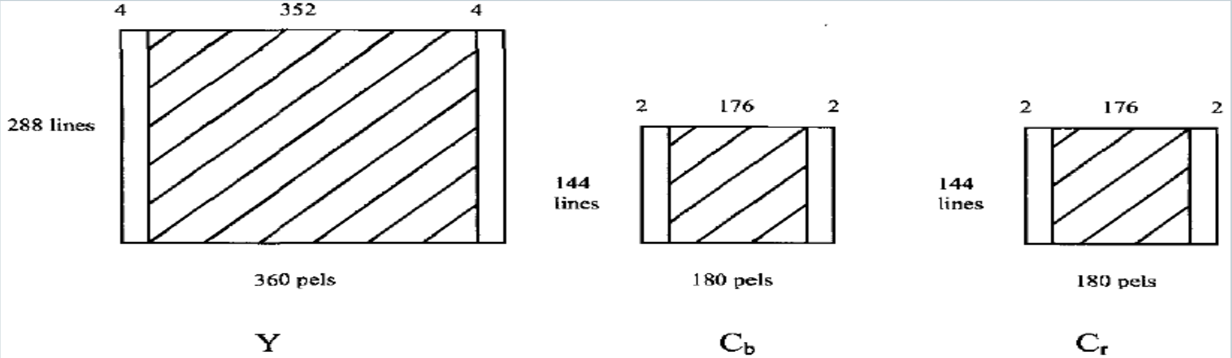
- The resolution of CIF is 352x288 and the resolution of QCIF is 1/4 of CIF, which is 176x144 [26].

Consider the YC_bC_r family of color spaces where Y represents the luminance, C_b represents the blue-difference chroma component and C_r represents the red-difference chroma component [25].

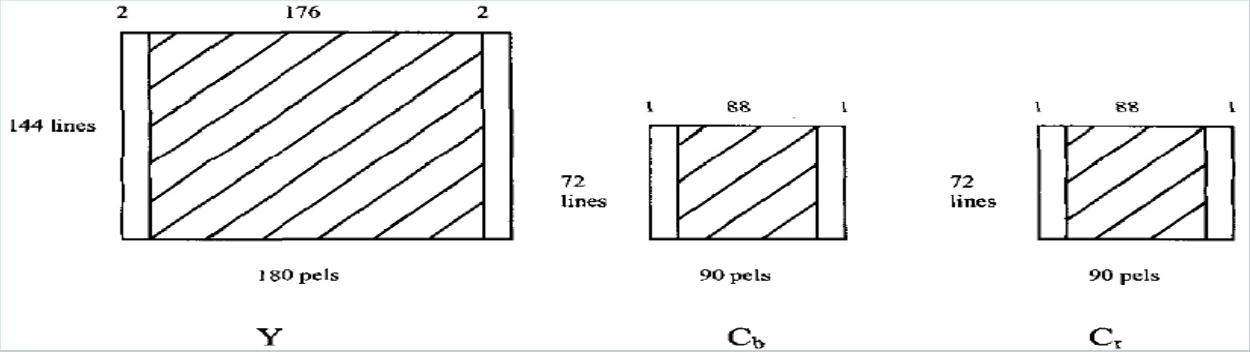
For QCIF and CIF, the luminance Y is equal to the resolution.

CIF and QCIF formats

- If sampling resolution 4:2:0 is used, then for
  CIF, the Cb and Cr are 176 x 144 lines
  QCIF, the Cb and Cr are 88 x 72 lines



Common Intermediate Format (CIF) [14]



Quadrature Common Intermediate Format (QCIF)[14]

YUV File 1: suzie_qcif.yuv

Specifications:

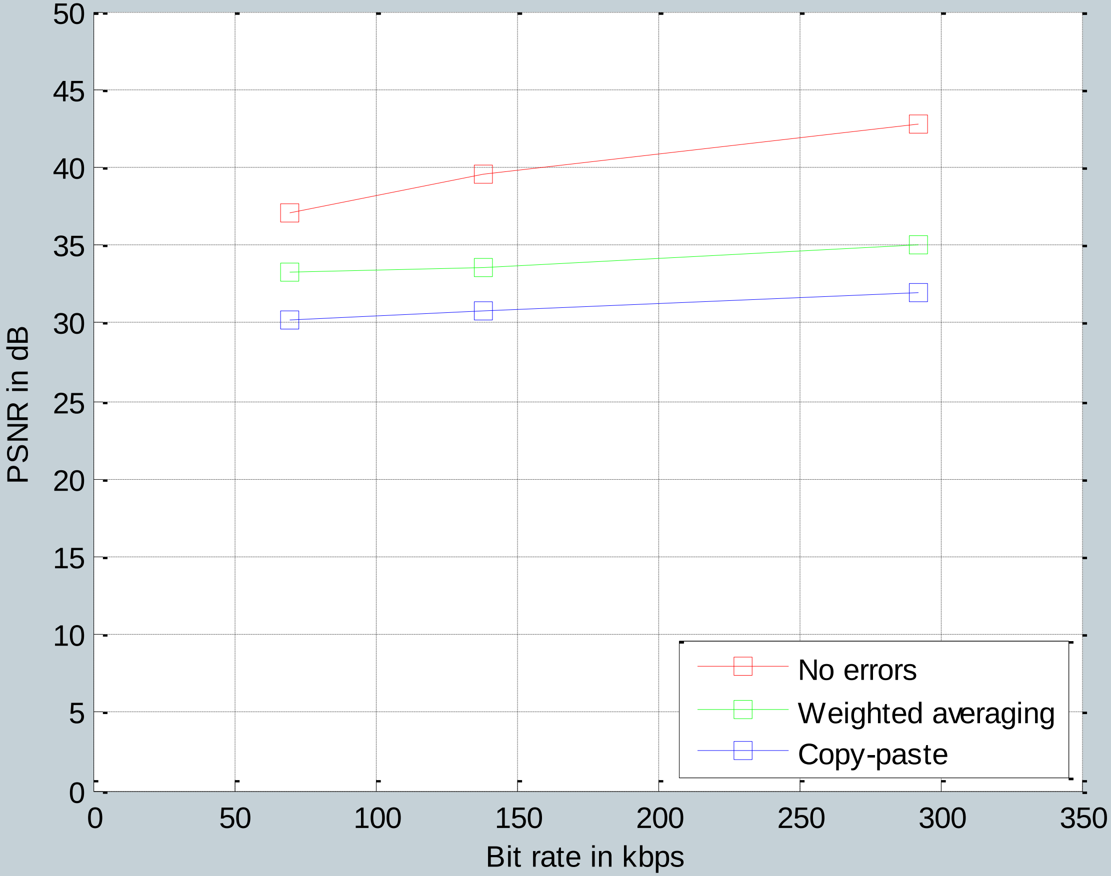
- QCIF sequence: suzie_qcif.yuv
- Total number of frames: 150
- Height: 176; Width: 144
- Total number of frames used: 20
- Frame rate: 30 frames/second

YUV File 1: suzie_qcif.yuv

Quantization Parameter	Bit Rate (Kbps)	Average Original PSNR(dB) (No Errors)	Average PSNR after weighted averaging (dB)	Average PSNR after copy paste (dB)
28	69.92	37.064	33.194	30.23
24	138.05	39.531	33.502	30.783
20	291.98	42.707	34.948	31.984

YUV File 1: suzie_qcif.yuv

Comparison between the Weighted averaging and Copy paste algorithms of Error Concealment



YUV File 2: foreman_qcif.yuv

Specifications:

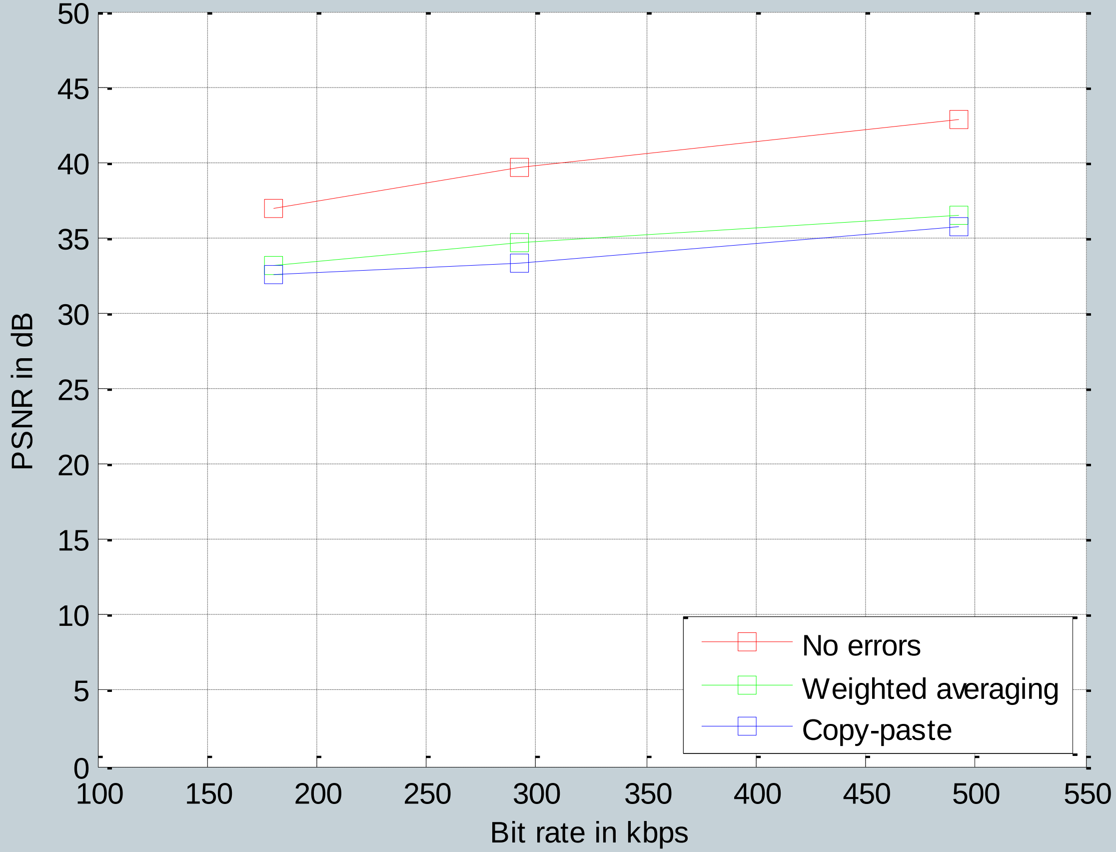
- QCIF sequence: foreman_qcif.yuv
- Total number of frames: 300
- Height: 176; Width: 144
- Total number of frames used: 20
- Frame rate: 30 frames/second

YUV File 2: foreman_qcif.yuv

Quantization Parameter	Bit Rate (Kbps)	Average Original PSNR(dB) (No Errors)	Average PSNR after weighted averaging (dB)	Average PSNR after copy paste (dB)
28	180.65	36.976	33.141	32.54
24	292.03	39.678	34.624	33.233
20	492.14	42.861	36.413	35.784

YUV File 2: foreman_qcif.yuv

Comparison between the Weighted averaging and Copy paste algorithms of Error Concealment



Conclusions

- Spatial error concealment using the weighted averaging algorithm is generally more effective than the temporal error concealment technique using the copy-paste algorithm.
- If the screen cuts are slow and there are no fast movements between the frames then the difference in effectiveness between the two techniques is very small.
- The copy-paste algorithm may be more effective than the weighted averaging algorithm in rare cases such as when the error is contained in the background which does not change over a sequence of frames.
- A Hybrid error concealment technique which adaptively switches between the spatial and temporal error concealment techniques will be most effective at error concealment.

MSU Video Quality Measurement Tool [24]

Future Work

- Further enhancing error detection by locating random and burst errors.
- Implementation of a hybrid error concealment technique which combines the advantages of both the spatial and temporal techniques of error concealment.
- Implementation of temporal error concealment using motion vector estimation.
- Error concealment in other codecs like AVS China and Dirac Pro.

References

[1] Y. Xu and Y. Zhou, “H.264 Video Communication Based Refined Error Concealment Schemes”, IEEE Transactions on Consumer Electronics, vol. 50, issue 4, pp. 1135–1141, November 2004.

[2] M. Wada, “Selective Recovery of Video Packet Loss using Error Concealment,” IEEE Journal on Selected Areas in Communication, vol. 7, issue 5, pp. 807-814, June 1989.

[3] Y. Chen, et al, “An Error Concealment Algorithm for Entire Frame Loss in Video Transmission”, Microsoft Research Asia, Picture Coding Symposium, December 2004.

[4] H. Ha, C. Yim and Y. Y. Kim, “Packet Loss Resilience using Unequal Forward Error Correction Assignment for Video Transmission over Communication Networks”, ACM digital library on Computer Communications, vol. 30, pp. 3676-3689, December 2007.

[5] X. Xiu, L. Zhuo and L. Shen, "A hybrid error concealment method based on H.264 standard", 8th International Conference on Signal Processing, vol. 2, April 2006.

[6] G. Sullivan, P. Topiwala and A. Luthra, "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions", SPIE Conference on Applications of Digital Image Processing XXVII, vol. 5, pp. 454-474, November 2004.

[7] R. Schafer, T. Wiegand and H. Schwarz, “The emerging H.264/AVC standard,” EBU Technical Review, Special Issue on Best of 2003, January 2003.

[8] T. Wiegand, et al, “Overview of the H.264/AVC Video Coding Standard” IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, pp. 560-576, June 2003.

[9] S. K. Bandyopadhyay, et al, “An error concealment scheme for entire frame losses for H.264/AVC”, IEEE Sarnoff Symposium, pp. 1-4, March 2006.

[10] Y. Xu and Y. Zhou, "Adaptive Temporal Error Concealment Scheme for H.264/AVC Video Decoder", IEEE Transactions on Consumer Electronics, vol. 54, issue 4, pp. 1846 – 1851, November 2008.

[11] D. Levine, W. Lynch and T. Le-Ngoc, "Observations on Error Detection in H.264", 50th Midwest Symposium on Circuits and Systems, pp. 815-818, August 2007.

References (contd.)

[12] B. Hrušovský, J. Mochná and S. Marchevský, "Temporal-spatial Error Concealment Algorithm for Intra-Frames in H.264/AVC Coded Video", 20th International Conference Radioelektronika, pp. 1-4, April 2010.

[13] W. Kung, C. Kim and C. Kuo "Spatial and Temporal Error Concealment Techniques for Video Transmission Over Noisy Channels", IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, issue 7, pp. 789-803, July 2006.

[14] S. Kwon, A. Tamhankar and K.R. Rao, "Overview of H.264 / MPEG-4 Part 10", J. Visual Communication and Image Representation, vol. 17, pp. 186-216, April 2006.

[15] S. Kumar, et al, "Error Resiliency Schemes in H.264/AVC Standard", IEEE Military Communications Conference, pp. 1-6, October 2006.

[16] Ignacio Cort Todoli, "Performance of Error Concealment Methods for Wireless Video", Diploma Thesis, Vienna University of Technology, 2007.

[17] M.S. Koul, "Error Concealment And Performance Evaluation Of H.264/AVC Video Streams In A Lossy wireless Environment", Department of Electrical Engineering, University of Texas at Arlington, May 2008.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[19] H.264/AVC Reference Software Download: http://iphome.hhi.de/suehring/tml/download/

[20] L. Liu, S. Zhang, X. Ye and Y. Zhang, "Error Resilience Schemes of H.264/AVC for 3G Conversational Video", Proc. IEEE Conf. Computer and Information Technology, pp. 657- 661, Sept. 2005

References (contd.)

[21] M.T. Sun, and A.R. Reibman, Compressed Video over Networks, Marcel Dekker, New York, 2001.

[22] V.S. Kolkeri "Error Concealment Techniques In H.264/AVC ,for Video Transmission Over Wireless Networks", M.S. Thesis, Department of Electrical Engineering, University of Texas at Arlington, December 2009

[23] YUV video sequences; website: http://trace.eas.asu.edu/yuv/

[24] MSU video quality measurement tool : http://compression.ru/video/quality_measure/video_measurement_tool_en.html

[25] JVT"Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T rec. H.264– ISO/IEC 14496-10 AVC)," March 2003, JVT-G050 available on http://ip.hhi.de/imagecom_G1/assets/pdfs/JVT- G050.pdf.