Content-Adaptive Improved Error Concealment Methods for H.264/AVC Video Communication

Santosh V. Chapaneri

Signal and Image Lab
Department of Electrical and Computer Engineering
University of Arizona
santoshc@email.arizona.edu

Thesis Advisor: Dr. Jeffrey J. Rodriguez

THE UNIVERSITY OF ARIZONA®

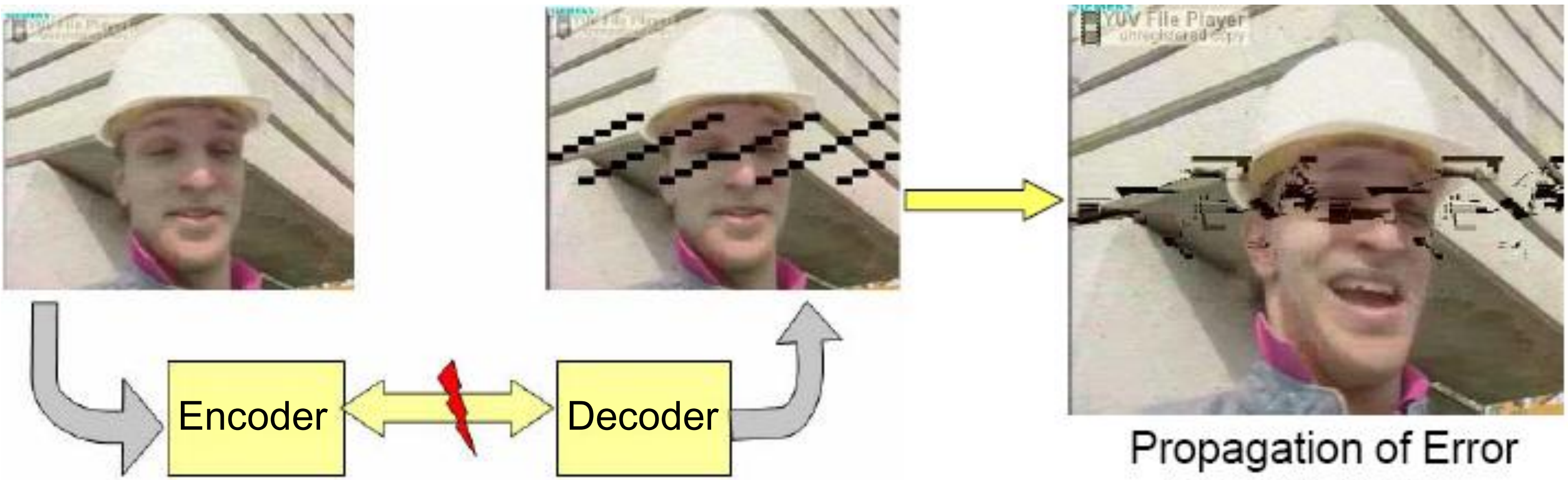
Outline

- Introduction
  - Error concealment (EC)
  - H.264/AVC
  - Video quality metrics
- Non-normative EC
- Previous work
- Directional Spatial EC
  - Suggestions and Results
- Proposed Spatial EC: PMEC
  - Methodology and Results
- Proposed Temporal EC: CAMP
  - Methodology and Results
- Content-Adaptive Refined EC (CAREC)
- Conclusions

Error Concealment (EC)

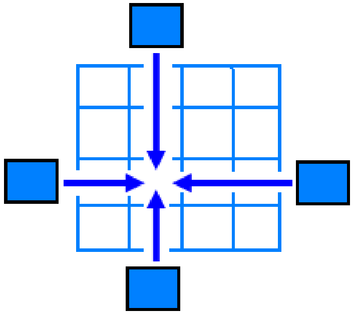
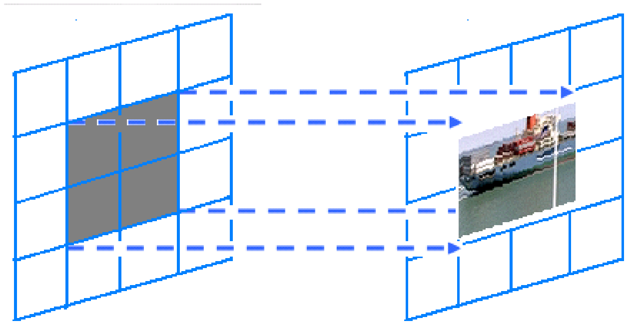
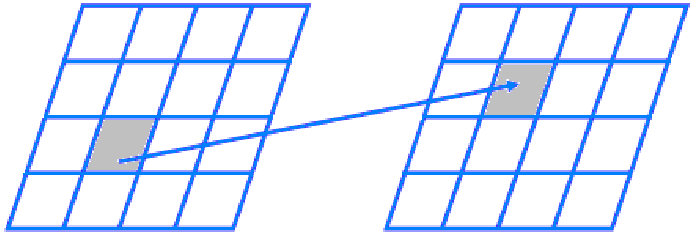
- **Problem**: Transmission errors may result in lost information
- **Goal**: Estimate the lost information in order to conceal the fact that an error has occurred
- **Observation**: Video exhibits a significant amount of correlation along the spatial and temporal dimensions
- **Approach**: Perform some form of spatial/temporal interpolation to estimate the lost information from the correctly received data

Example of error propagation



Propagation of Error
After 14 Frames

Error concealment basic techniques

Spatial interpolation	
Temporal Replacement (TR)	
Motion-compensated temporal interpolation	

What is H.264/AVC?

- Substantial improvement over all previous video coding standards (2x compression, substantial perceptual quality)
- Addresses full range of video applications :

  low bit-rate wireless applications, HD DVD, video streaming over Internet, digital cinema, etc.
- Jointly developed by ITU-T (H.264) and ISO/IEC (MPEG-4); commonly known as H.264/AVC

ITU-T → *H.264*

ISO-IEC → *MPEG-4 AVC*

(Advanced Video Coding)

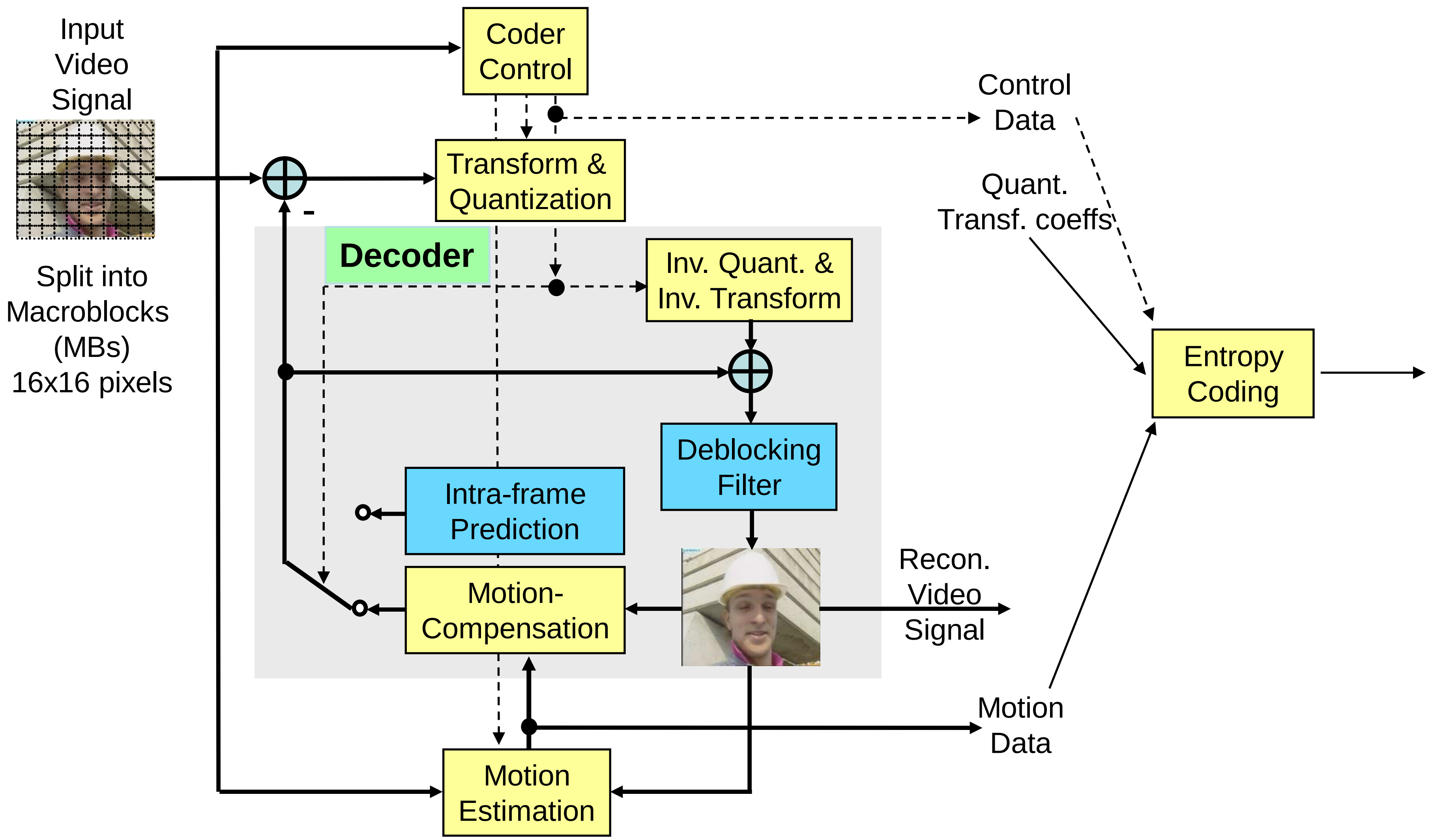
JVT

(Joint Video Team)

-H.264/AVC

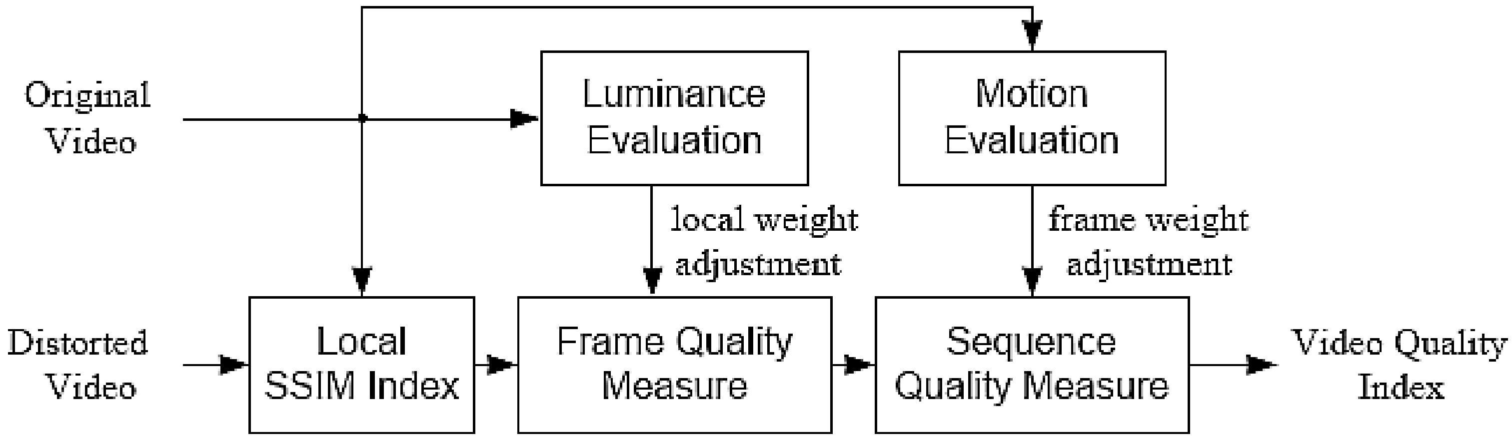
-MPEG 4 Part 10

H.264/AVC Encoder



Video Quality Metrics

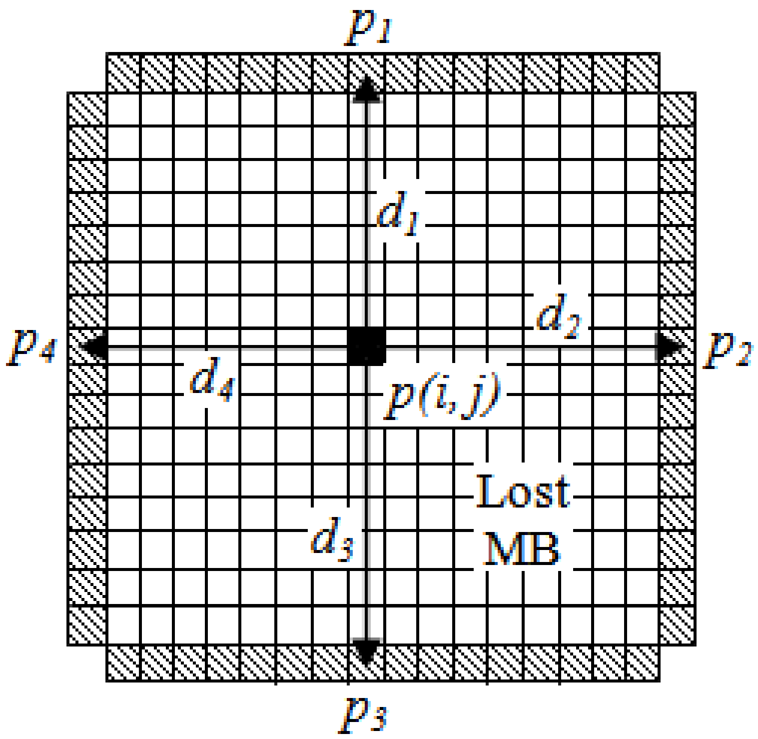
- PSNR
- Structural SIMilarity (SSIM)
  - measures deviations in *luminance*, *contrast*, and *structure* between the reference and concealed frame portions
  - Mean SSIM (MSSIM) indicates quality of overall frame
  - correlates well with the mean opinion score*
- Video quality index (*Q*) based on SSIM



*Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 1, pp. 1-14, Jan. 2004

Non-normative Spatial EC

- Based on weighted-pixel bilinear interpolation

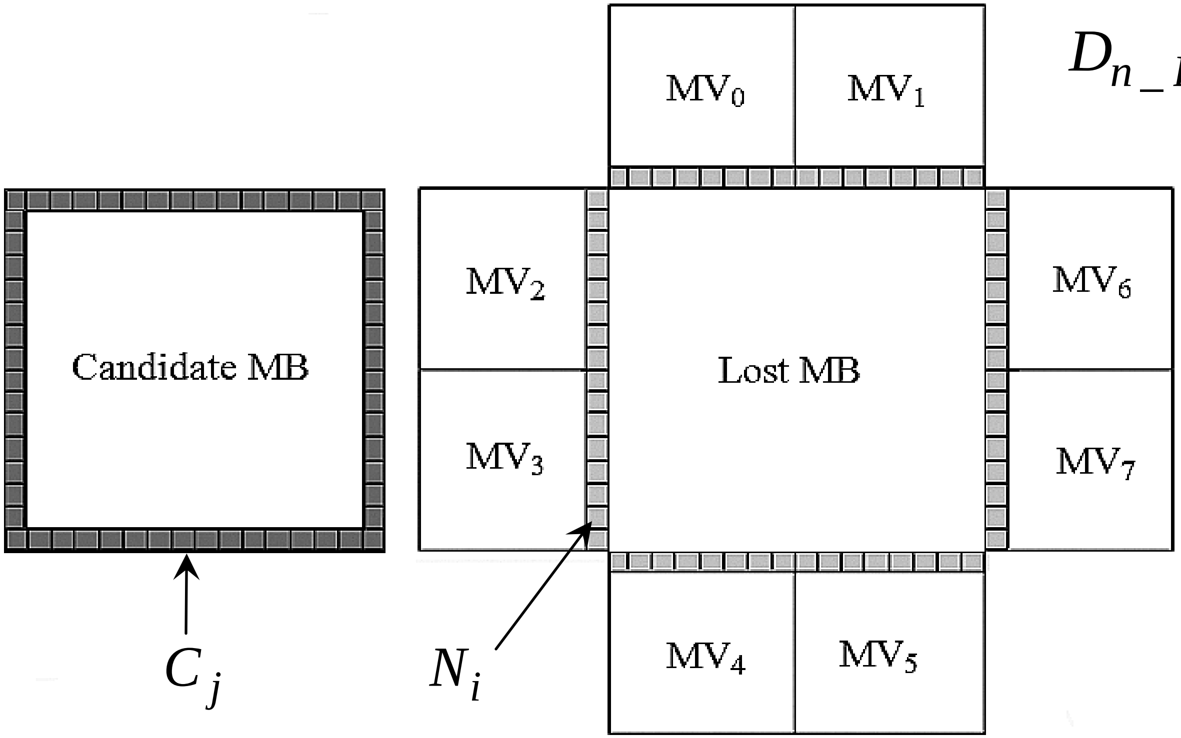


$$p(i,j) = \frac{\sum_{m=1}^{4} p_m (15 - d_m)}{\sum_{m=1}^{4} d_m}$$

- Performs well only when missing MB is in smooth region
- Does not consider edge directions; creates blocking artifacts in reconstructed picture

Non-normative Temporal EC

- Uses Boundary Matching Algorithm (BMA) as the distortion measure
- Motion vector (MV) yielding least distortion is selected for best candidate MB



$$D_{n\_BMA} = \frac{1}{M}\left|\sum_{\forall i \in I} N_i - \sum_{\forall j \in J} C_j\right|$$

M = total number of boundary pixels

Non-normative concealment

Spatial EC –
Blocking artifacts

Temporal EC –
Inaccurate motion vectors

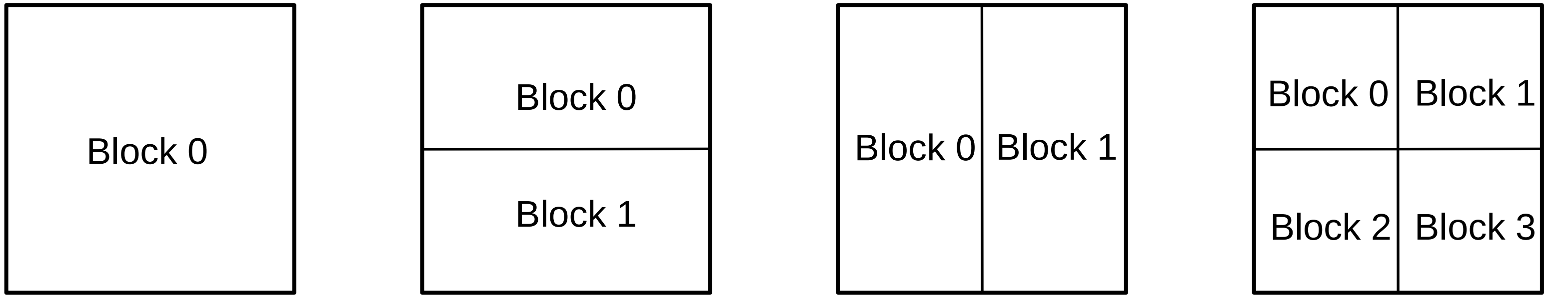
Previous Work: Spatial EC

- Directional interpolation (Xu et al.) [1] – using gradient filters to determine the dominant edge direction
- POCS – iterative procedure for smoothness (Yu et al.) [2]
- DCT domain based concealment (Alkachouh et al.) [3]
- Best neighborhood matching (Wang et al.) [4] – exploiting block-wise similarities in the frame
- Directional entropy of neighboring edges (Agrafiotis et al.) [5] – switch between directional and bilinear interpolation

Previous Work: Temporal EC

- Multi-frame BMA approach (Lee et al.) [6] – distortion measure constrained according to the motion of the succeeding frames
- Refined BMA (RBMA) (Chen et al.) [7] – splits the lost MB into four 8×8 blocks and constrains the BMA procedure
- Adaptive block sizes (ABS) (Kim et al.) [8] – splits the lost MB into four partition types; does not consider the spatial continuity of the concealed MB

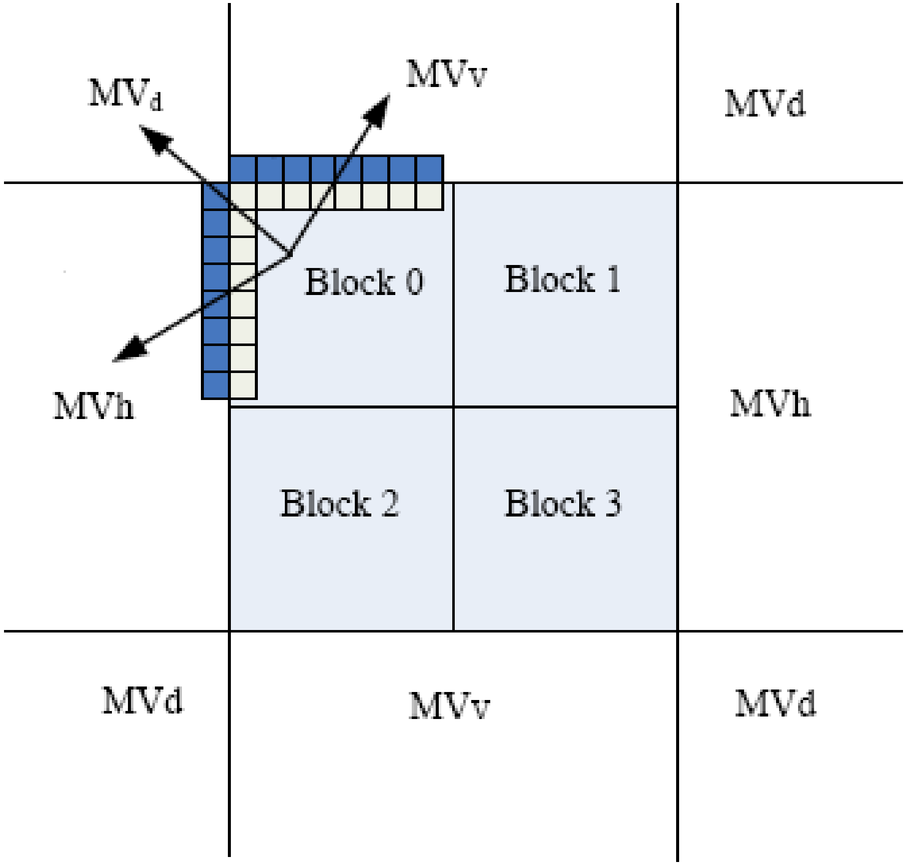
ABS



Previous Work: Temporal EC (contd..)

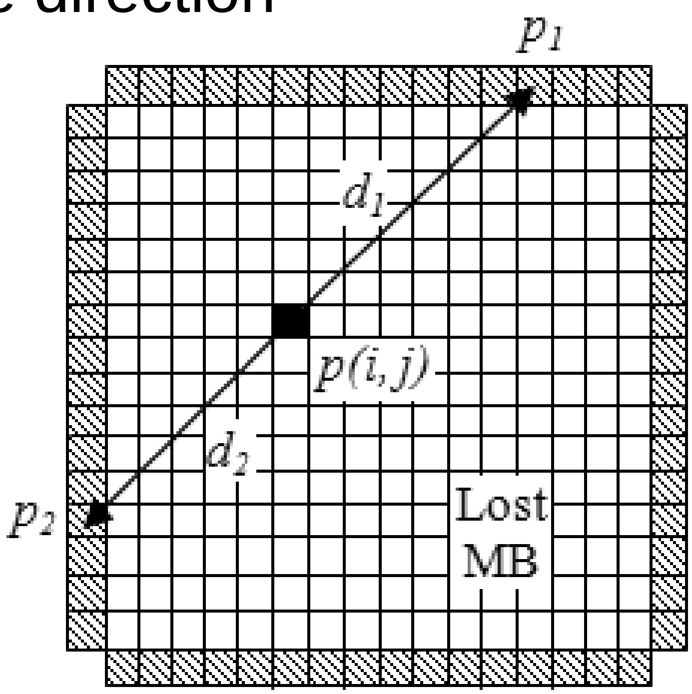
- Refined temporal concealment (RTC) [1] – splits the lost MB into four 8×8 blocks and increases the candidate set of MVs

1) Horizontal MV
2) Vertical MV
3) Diagonal MV
4) Median of all MVs
5) Average of all MVs
6) MV of co-located MB
7) Zero MV



Directional Spatial EC

- To preserve the edge continuity of the missing MB with the neighboring MBs, dominant edge direction is computed using gradient filters
- Pixels of missing MB are interpolated in the dominant edge direction



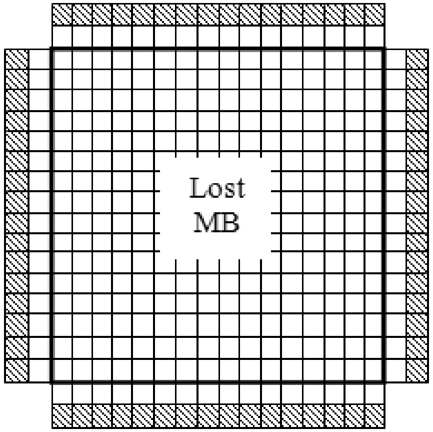
$$p(i, j) = \frac{p_1 d_2 + p_2 d_1}{d_1 + d_2}$$

$$\forall (i, j) \in \text{lost MB}$$

Directional Spatial EC (contd..)

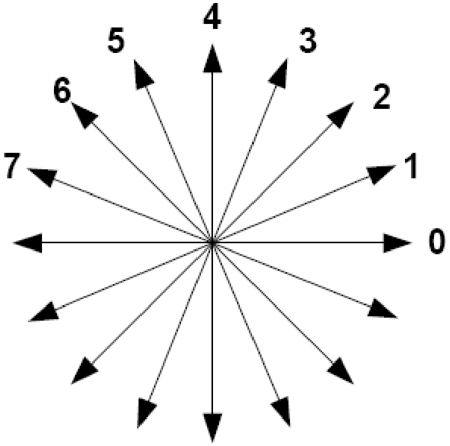
- Two existing techniques, Mean method [9] and Mode method [1], to determine the dominant edge direction

Mean method



$$\theta_d = \frac{\sum \theta(i,j)\left|G(i,j)\right|}{\sum \left|G(i,j)\right|}$$

Mode method



$$C[\theta(i,j)] = 0$$

$$C[\theta(i,j)] += \left|G(i,j)\right|$$

$$\theta_d = \theta(i,j) \text{ with } \max(C[\theta(i,j)])$$

Directional Spatial EC (contd..)

- Mode method should be used to determine the dominant edge direction
- Outperforms the Mean method by enhancing the perceptual quality

Concealed with Mean method

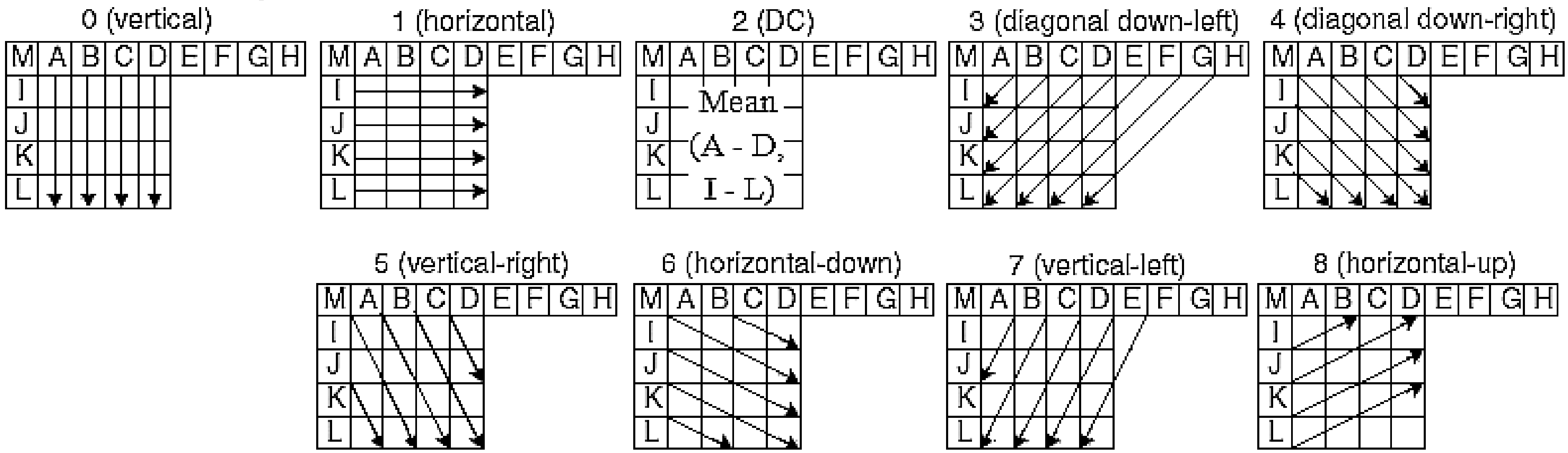
Concealed with Mode method

Proposed Spatial EC: PMEC

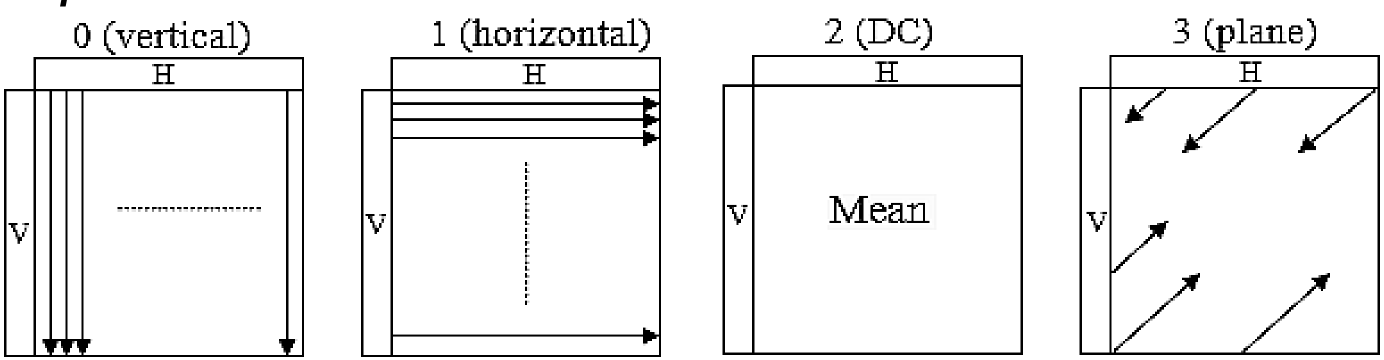
- H.264/AVC uses Intra prediction to reduce the spatial redundancies prior to transformation
- The prediction modes (*pmodes)* implicitly describe the edge orientations
- Use the existing information of *pmodes* of neighboring MBs to determine the dominant edge direction of the missing MB
- The proposed **prediction modes error concealment** (PMEC) algorithm has a reduced computational complexity

Proposed Spatial EC: PMEC (contd..)

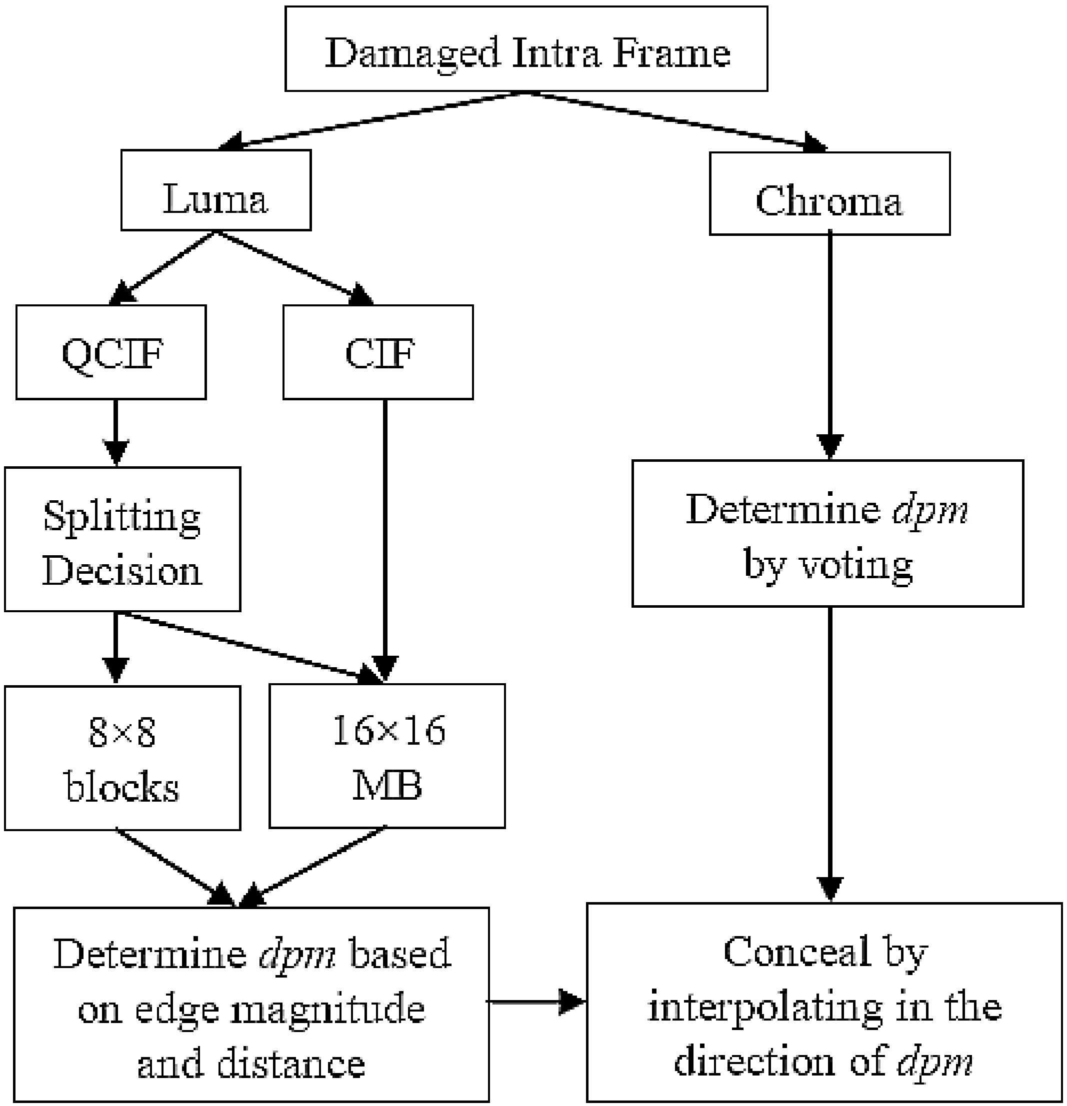
- Nine *pmodes* for 4×4 sub-blocks

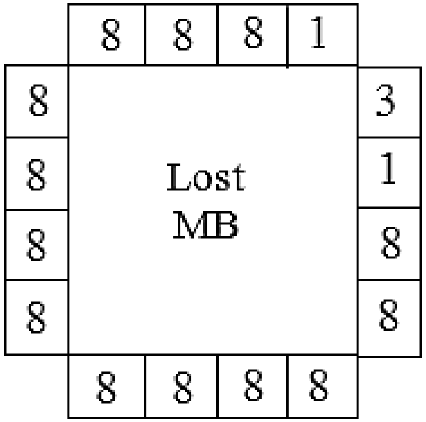


- Four *pmodes* for 16×16 MB

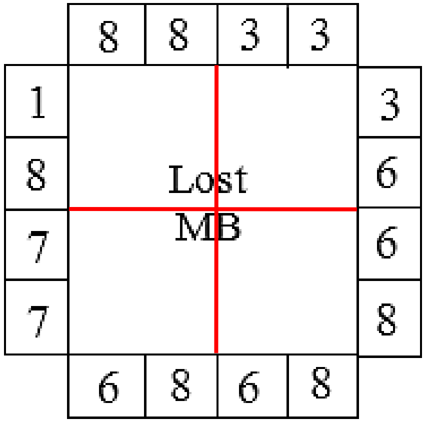


Proposed Spatial EC: PMEC (contd..)





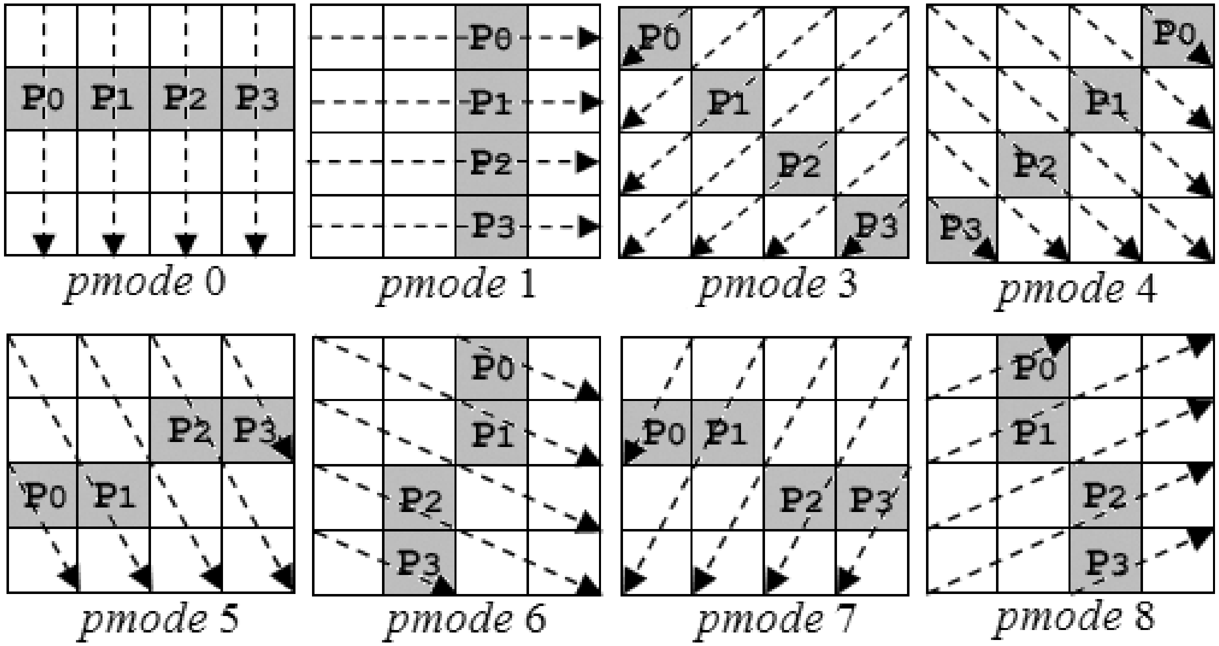
No splitting of lost MB



Split lost MB into 8x8 blocks

dpm: dominant *pmode*

Proposed Spatial EC: PMEC (contd..)



- Edge magnitude is estimated as [max(P) – min(P)]/*d* where P = {P_0, P_1, P_2 ,P_3} and *d* is the distance between max(P) and min(P)
- The *pmode* with the highest edge magnitude is selected as the dominant *pmode* ⟹ dominant edge direction

Computational Complexity of PMEC

Operations	JM	Xu's method	PMEC
Additions	1536	1280+512*	**544**+512*
Multiplications	1024	640	**512**
Divisions	256	320	**272**
Comparisons	0	7	104
Shifts	0	0+512*	0+512*

*Extra computations for half-pixel boundary values

JM = Joint Model H.264/AVC reference software implementation

PMEC Results

- *Foreman* CIF Intra frame @ 10% loss rate

Concealed with Xu's method
PSNR = 33.17 dB, MSSIM = 0.8035

Concealed with PMEC
PSNR = 34.26 dB, MSSIM = 0.9126

PMEC Results (contd..)

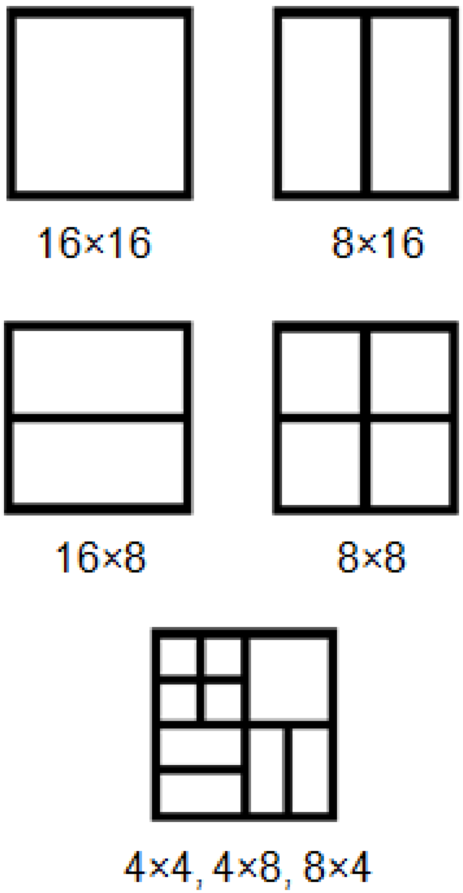
Sequence	Loss rate	JM	Xu's method	PMEC
Carphone	10%	33.6031 dB, 0.5925	33.8173 dB, 0.7201	34.3468 dB, 0.8153
Stefan	10%	31.3147 dB, 0.5387	32.1071 dB, 0.7152	33.4715 dB, 0.7836
Hall	10%	32.9773 dB, 0.6928	33.6285 dB, 0.7818	34.2156 dB, 0.8572

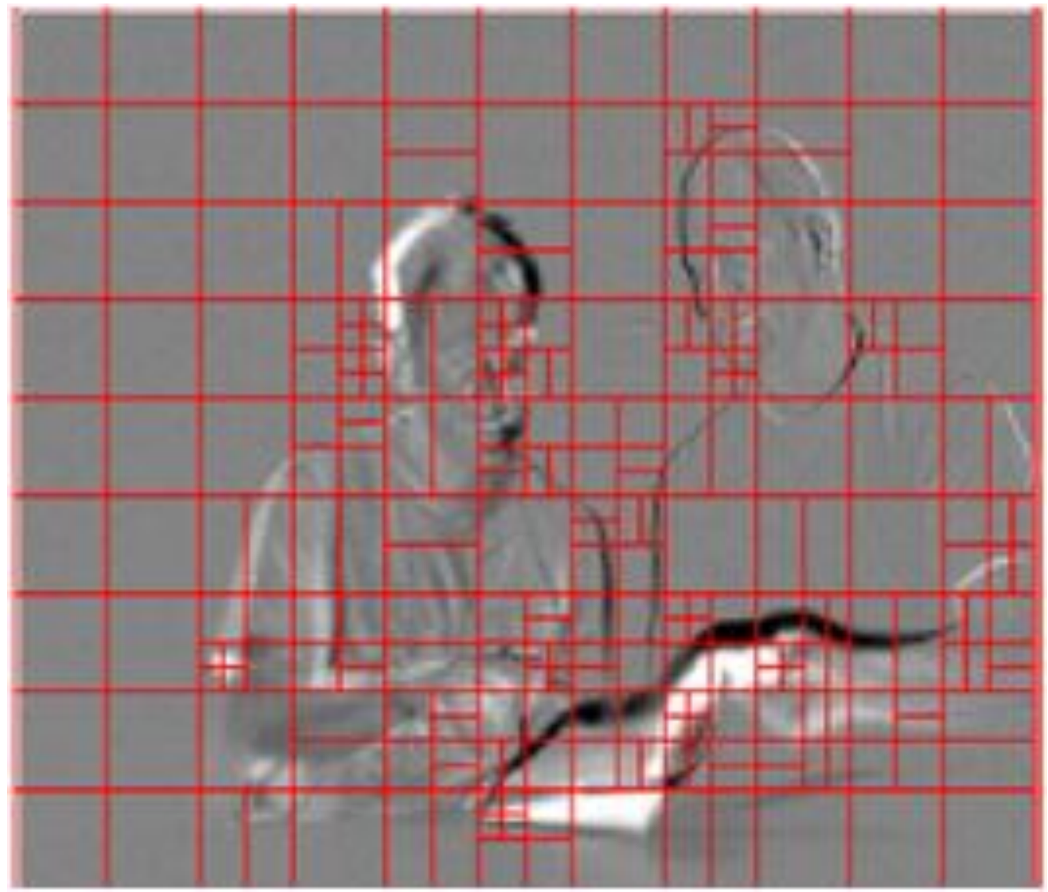
Table shows the values of average PSNR and MSSIM

PSNR improvement: **0.92** dB relative to Xu's method and **2.25** dB relative to JM
MSSIM improvement: **0.1** relative to Xu's method and **0.37** relative to JM

Proposed Temporal EC: CAMP

- H.264/AVC uses tree-structured motion compensation with variable block sizes

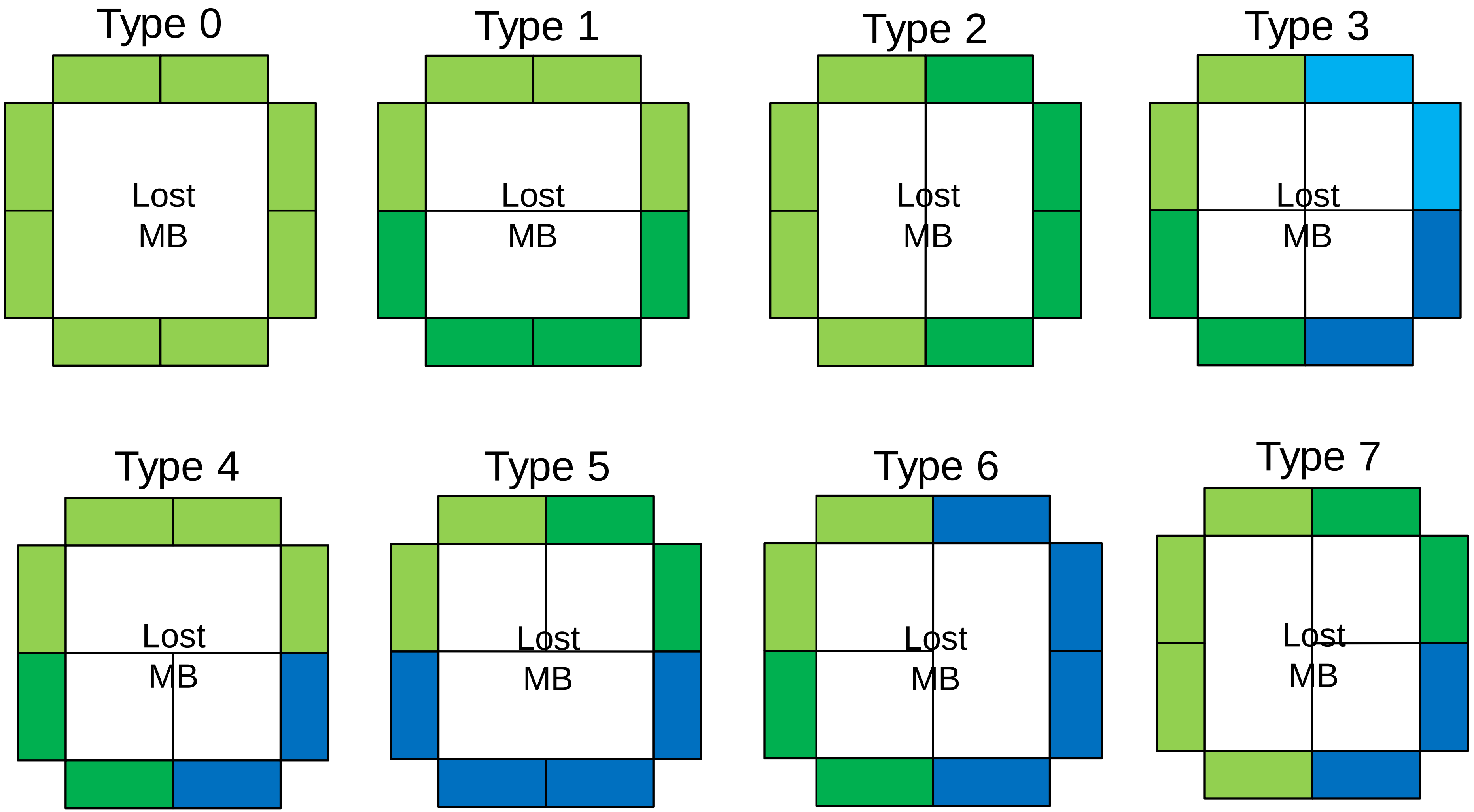




Proposed Temporal EC: CAMP (contd..)

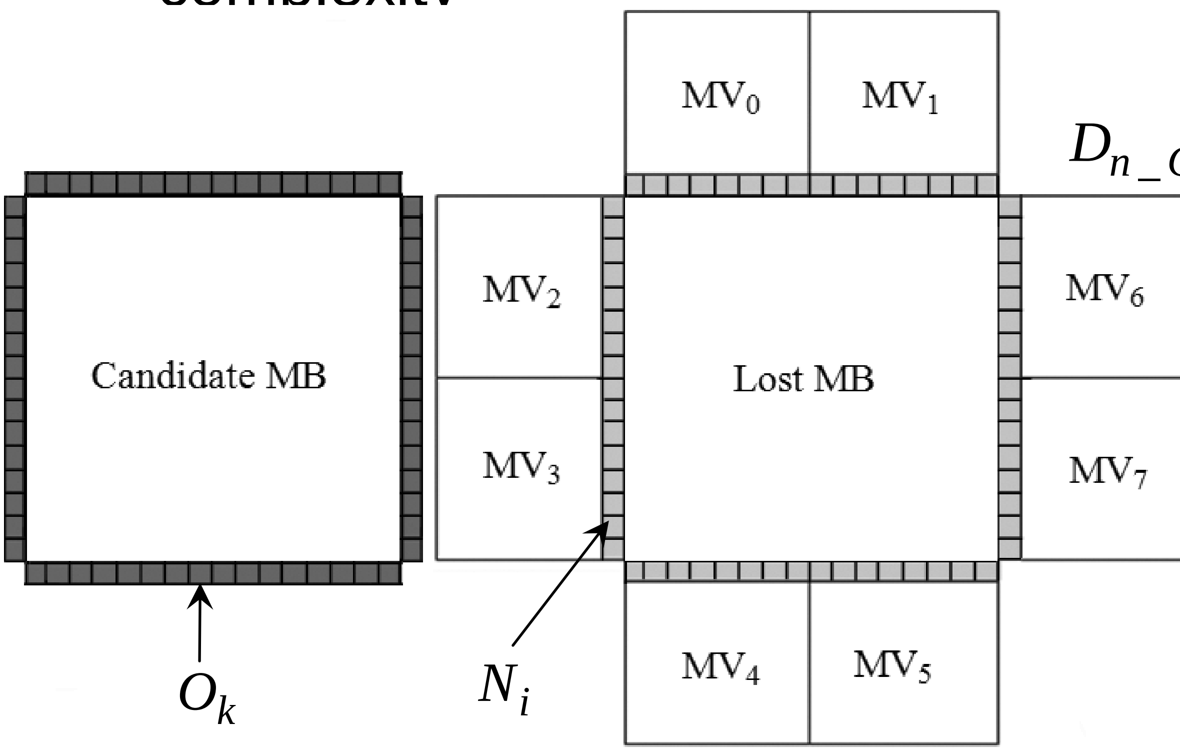
- In the proposed **content-adaptive macroblock partitioning** (CAMP) algorithm, the lost MB is partitioned adaptively into different block sizes
- Using the mode information of the neighboring MBs, the lost MB is suitably partitioned into one out of eight possible types
- Each partition is concealed with different candidate set of motion vectors
- Results in smoother concealment and avoids the blocking artifacts

Proposed Temporal EC: CAMP (contd..)



Proposed Temporal EC: CAMP (contd..)

- Outer Boundary Matching Algorithm is used for distortion computation to perform the initial temporal concealment
- Better performance than BMA at the same level of complexity



$$D_{n\_OBMA} = \frac{1}{M}\left| \sum_{\forall i \in I} N_i - \sum_{\forall k \in K} O_k \right|$$

M = total number of relevant boundary pixels

Proposed Temporal EC: CAMP (contd..)

- Overlapped Block Motion Compensation (OBMC) is used to post-process the initial concealed MB
- Avoids spatial discontinuities between the concealed MB and its neighbors
- Split the initial concealed MB into four 8×8 blocks
- For each 8×8 block, the pixels are modified by a weighted sum of prediction values
- The MVs of the neighboring blocks spatially adjacent to the concerned 8×8 block are used for prediction

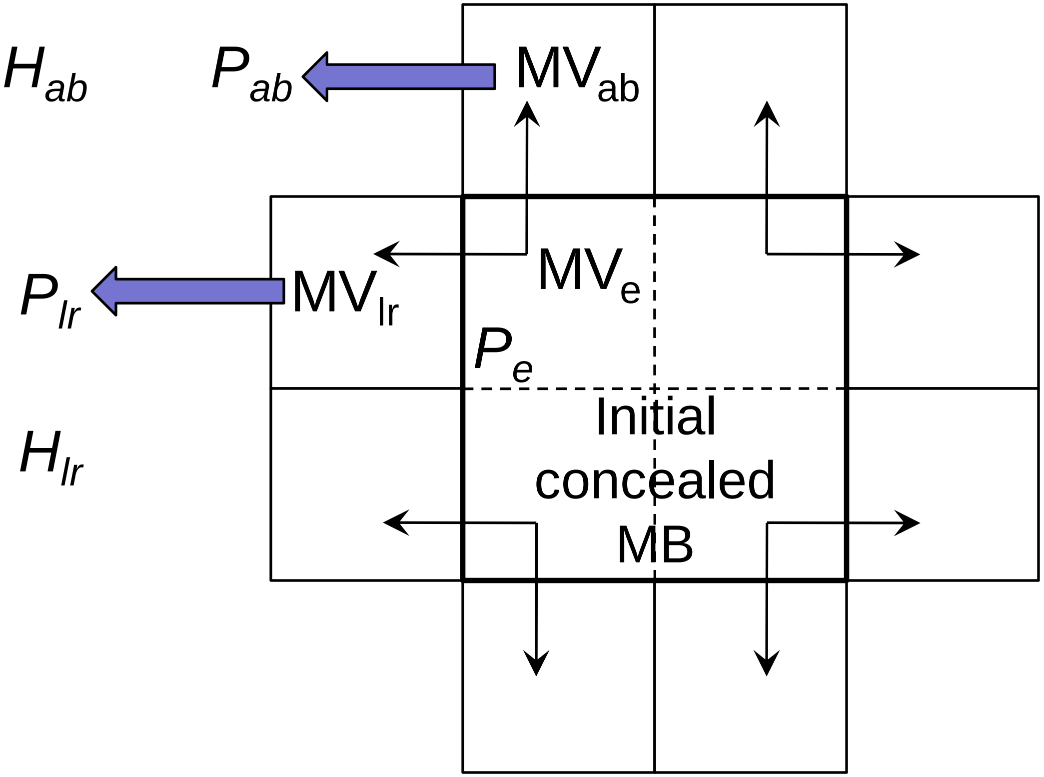
Proposed Temporal EC: CAMP (contd..)

H_{ab}

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

H_{lr}

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



H_e

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

$$P(i, j) = [P_e(i, j)H_e(i, j) + P_{ab}(i, j)H_{ab}(i, j) + P_{lr}(i, j)H_{lr}(i, j) + 4] >> 3$$

$\forall(i, j)$ 8×8 block of initial concealed MB

CAMP Results

- *Stefan* CIF Inter frame @ 10% loss rate

Concealed with ABS
PSNR = 26.35 dB, Q = 0.8726

Concealed with CAMP
PSNR = 27.63 dB, Q = 0.9432

CAMP Results (contd..)

Sequence	Loss rate	JM	ABS	CAMP
Table-tennis	10%	26.2604 dB, 0.7255	27.2978 dB, 0.8189	28.7926 dB, 0.8751
Carphone	10%	31.1764 dB, 0.7677	32.2136 dB, 0.8561	33.8261 dB, 0.9517
Foreman	10%	28.9998 dB, 0.7649	30.3074 dB, 0.8362	31.4613 dB, 0.9258

Table shows the values of average PSNR and video quality index *Q*

PSNR improvement: **1.1** dB relative to ABS and **2.5** dB relative to JM
Q improvement: **0.1** relative to ABS and **0.2** relative to JM

Content-Adaptive Refined EC (CAREC)

- A hybrid spatio-temporal adaptive switching mechanism for H.264/AVC decoder
- Handles errors occurring in both Intra and Inter frames adaptively
- Lost MBs in Intra frame: Use PMEC algorithm for spatial concealment
- Lost MBs in Inter frame: Switch between PMEC & CAMP
  - Determine the neighboring MB encoding modes
  - If number of Inter coded MBs > Intra/Skipped MBs, then use CAMP algorithm for temporal concealment
  - Else use PMEC for spatial concealment, since in this case there is not enough motion information available
  - If any neighboring MB is skipped, replace it with the co-located MB in the reference frame

CAREC Results

- *Carphone* QCIF Inter frame @ 10% loss rate

Concealed with Xu's method (RSTC)
PSNR = 31.82 dB, Q = 0.8674

Concealed with CAREC
PSNR = 33.08 dB, Q = 0.9578

CAREC Results (contd..)

Sequence	Loss rate	JM	RSTC	CAREC
Foreman	10%	28.7335 dB, 0.8766	30.0827 dB, 0.8968	31.4852 dB, 0.9273
Hall	10%	29.8872 dB, 0.8752	31.1054 dB, 0.9023	32.4916 dB, 0.9426
Salesman	10%	29.0715 dB, 0.8315	30.1722 dB, 0.8782	31.8347 dB, 0.9521

Table shows the values of average PSNR and video quality index *Q*

PSNR improvement: **1.4** dB relative to RSTC and **2.9** dB relative to JM
Q improvement: **0.1** relative to RSTC and **0.2** relative to JM

Conclusions

- To determine the dominant edge direction for spatial interpolation, Mode method should be used
- The proposed PMEC algorithm for spatial EC reduces the computational complexity and enhances the concealment performance
- The proposed CAMP algorithm for temporal EC achieves a smoother concealment of the lost MB due to the adaptive partitioning and avoids the spatial discontinuities with its neighbors
- The hybrid CAREC switching mechanism improves the overall concealment performance and reduces the structural degradations by achieving higher video quality values
- *Future Work*: EC methods for Main profile targeting broadcast and HD DVD applications, EC using Scalable video coding

References

1) Y. Xu, and Y. Zhou, “H.264 video communication based refined error concealment schemes,” *IEEE Trans. Consum. Electron.*, vol. 50, no. 4, pp. 1135-1141, Nov. 2004
2) G. S. Yu, M. M. K. Liu, and M. W. Marcellin, “POCS-based error concealment for packet video using multiframe overlap information,” *IEEE Trans. Circuits Syst. Video Tech.*, vol. 8, no. 4, pp. 422-434, Aug. 1998
3) Z. Alkachouh, and M. G. Bellanger, “Fast DCT-based spatial domain interpolation of blocks in images,” *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 729-732, Apr. 2000
4) Z. Wang, Y. Yu, and D. Zhang, “Best neighborhood matching: An information loss restoration technique for block-based image coding systems,” *IEEE Trans. Image Process.*, vol. 7, no. 7, pp. 1056-1061, Jul. 1998
5) D. Agrafiotis, D. R. Bull, and C. N. Canagarajah, “Enhanced error concealment with mode selection,” *IEEE Trans. Circuits Syst. Video Tech.*, vol. 16, no. 8, pp. 960-973, Aug. 2006
6) Y. C. Lee, Y. Altunbasak, and R. Mersereau, “A temporal error concealment method for MPEG coded video using a multi-frame boundary matching algorithm,” *IEEE Intl. Conf. Image Process.*, vol. 1, pp. 990-993, Oct. 2001
7) T. Chen, X. Zhang, and Y. Q. Shi, “Error concealment using refined boundary matching algorithm,” *Proc. IEEE Intl. Conf. ITRE*, pp. 55-59, Aug. 2003
8) D. Kim, S. Yang, and J. Jeong, “A new temporal error concealment method for H.264 using adaptive block sizes,” *IEEE Intl. Conf. Image Process.*, vol. 3, pp. 928-932, Sep. 2005
9) O. Nemethova, A. Al-Moghrabi, and M. Rupp, “Flexible error concealment for H.264 based on directional interpolation,” *IEEE Intl. Conf. Wireless Networks, Communications and Mobile Computing*, vol. 2, pp. 1255-1260, Jun. 2005

Thank You

Back up

FMO

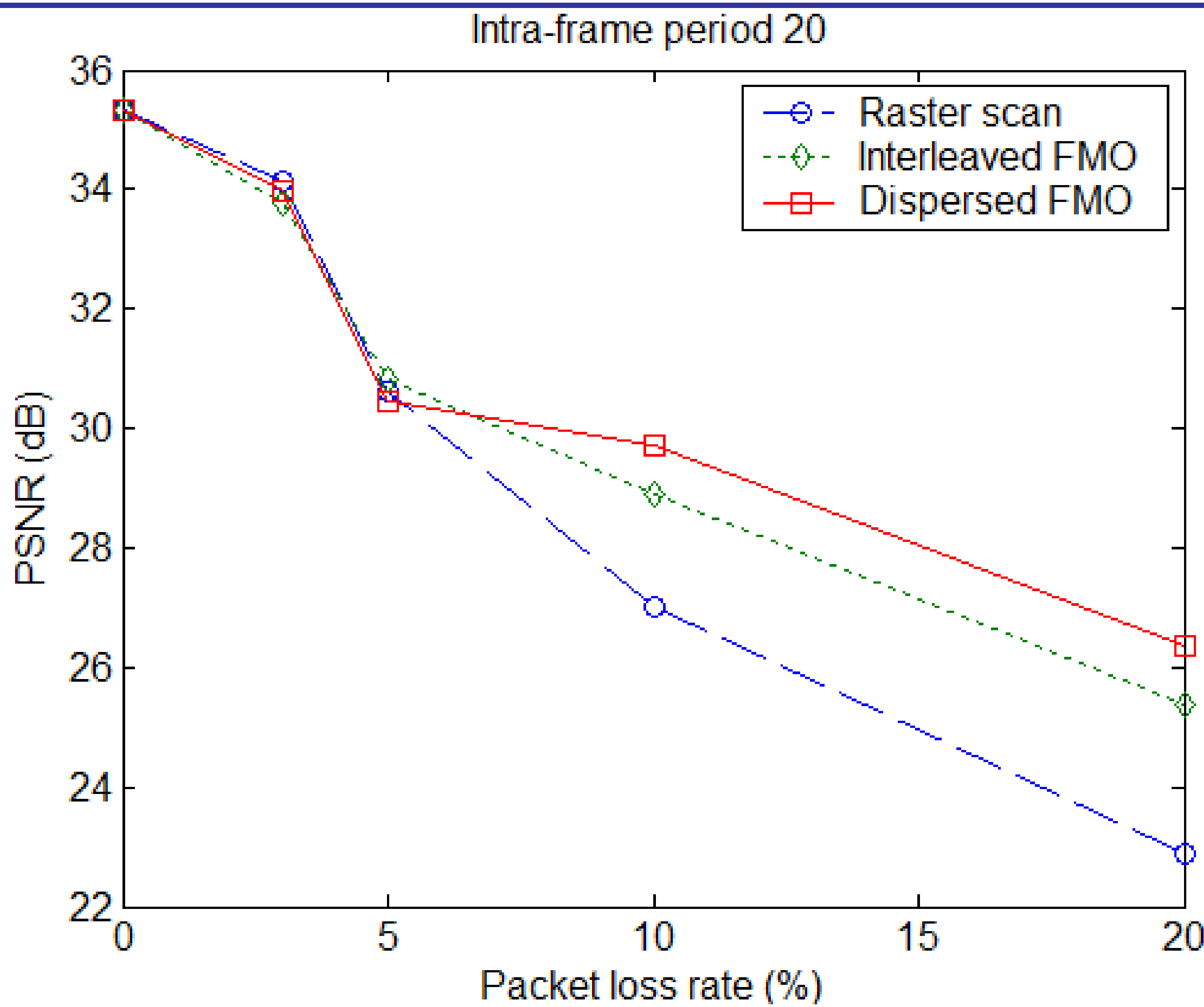
- FMO improves the concealment performance compared to traditional raster scan pattern
- Dispersed FMO type gives better performance than interleaved FMO type – more number of neighboring MBs aid in concealment

Raster scan
@ 10% loss

Interleaved FMO
@ 10% loss

Dispersed FMO
@ 10% loss

FMO



THE UNIVERSITY OF ARIZONA®

Profiles and Tools