ECE 634: Digital Video Systems
Predictive coding: 2/23/17

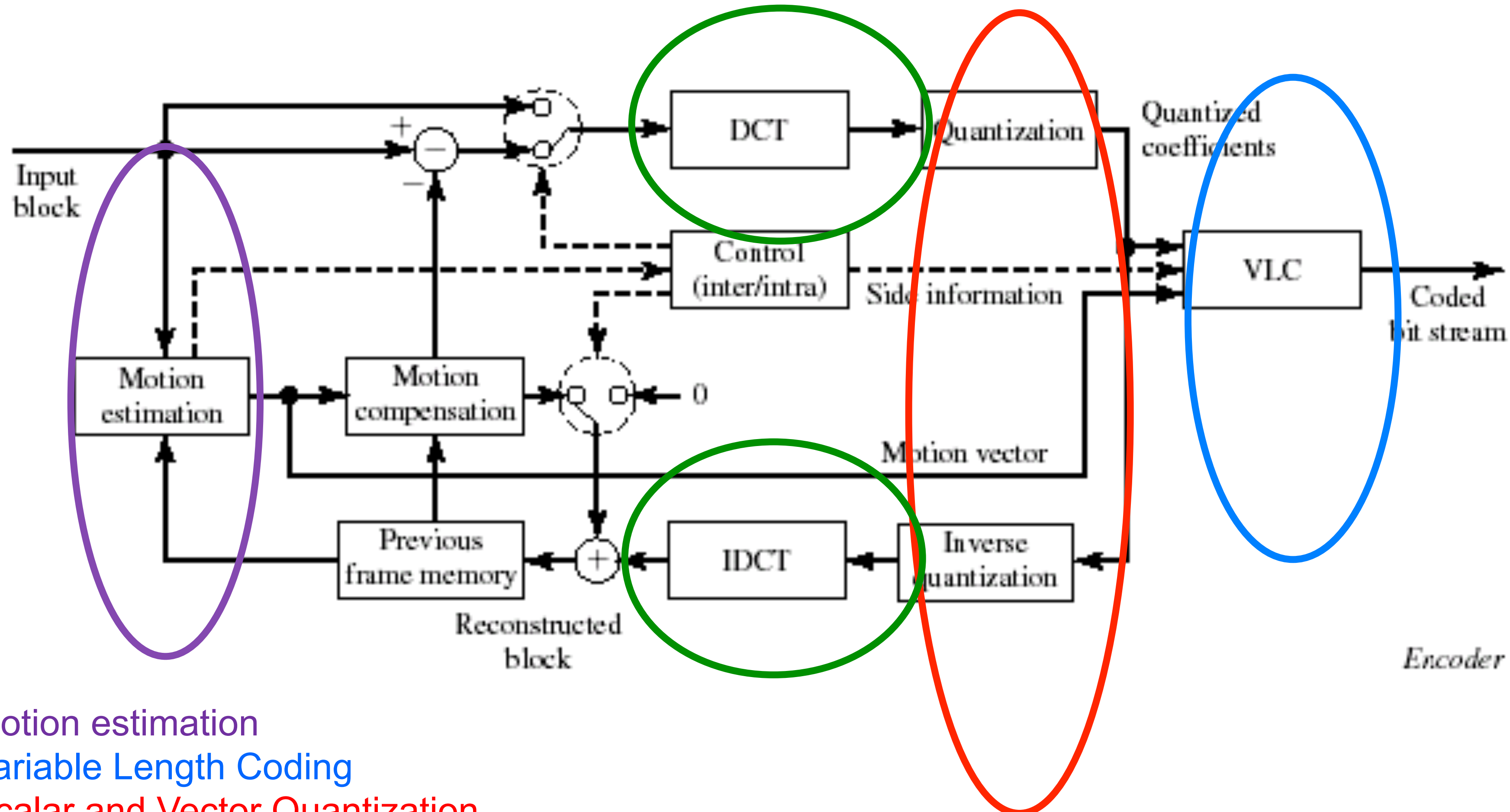
Professor Amy Reibman
MSEE 356
reibman@purdue.edu

http://engineering.purdue.edu/~reibman/ece634/index.html

Outline

- Predictive coding
  - What, how, which, how good?

Encoder Block Diagram of a Typical Block-Based Video Coder (Assuming No Intra Prediction)



Motion estimation
Variable Length Coding
Scalar and Vector Quantization
DCT, (and wavelets)
This lecture: predictive coding

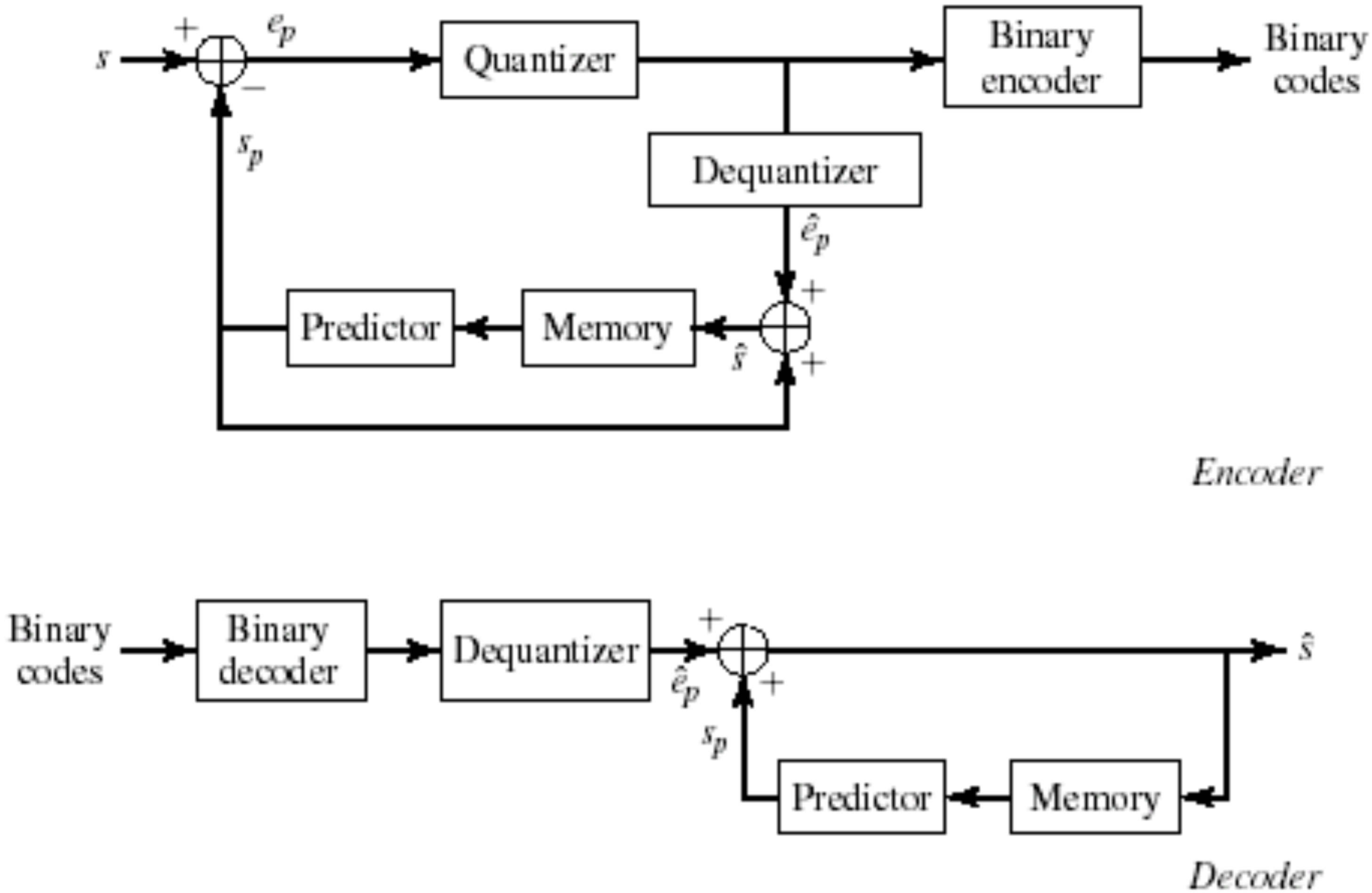
Predictive Coding

- Motivation: Predict a sample from past samples;
  quantize and code the prediction error only
- If the prediction error is typically small, then it can be represented with a lower average bit rate
- Optimal predictor: minimize the prediction error

A	B	C	D
E	F	G	H
I	J	K	L

$$\hat{f}_K = a f_F + b f_G + c f_H + d f_J$$

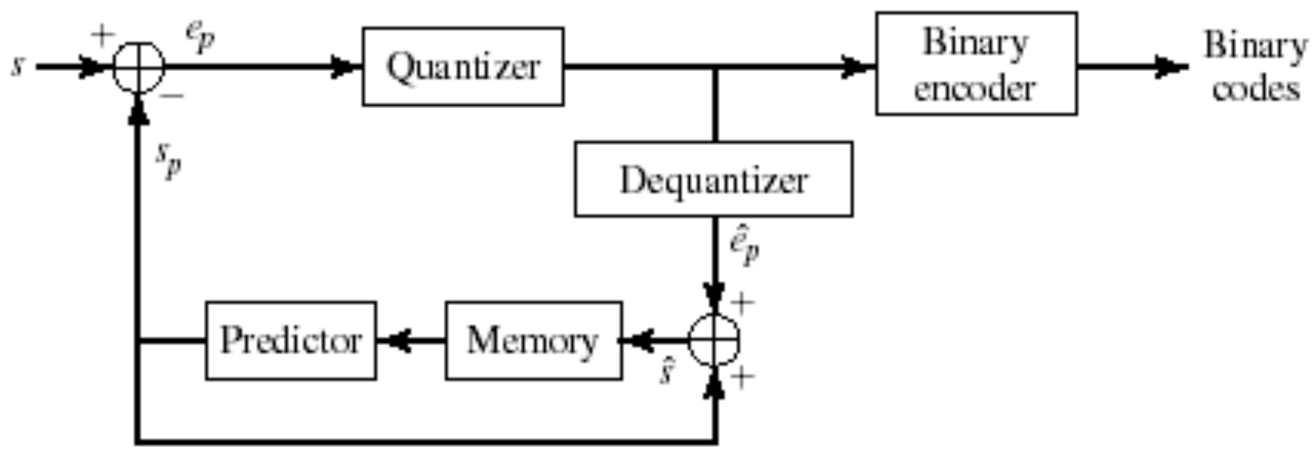
Encoder and Decoder Block Diagram (Closed Loop Prediction)



Distortion in Predictive Coder

- With closed-loop prediction, reconstruction error in a sample is equal to the quantization error for the prediction error

$$\hat{s} = s_p + \hat{e}_p = s_p + e_p - e_q = s - e_q .$$



Efficiency of coder depends on efficiency of predictor

- Question: what predictor should we use?
  - Minimize the bit rate for coding the prediction error
  - Because quantization error with a given bit rate depends on the variance of the signal, minimizing the quantization error equivalent to minimizing the prediction error variance
  - We consider only a linear predictor: $s_p = \sum_{k=1}^{K} a_k s_k$

K: order of the predictor

Linear Minimal MSE Predictor

- Prediction error:

$$\sigma_p^2 = E\{|\mathcal{S}_0 - \mathcal{S}_p|^2\} = E\left\{\left|\mathcal{S}_0 - \sum_{k=1}^{K} a_k \mathcal{S}_k\right|^2\right\}.$$

- Optimal coefficients must satisfy:

$$E\left\{\left(\mathcal{S}_0 - \sum_{k=1}^{K} a_k \mathcal{S}_k\right)\mathcal{S}_l\right\} = 0, \quad l = 1, 2 \ldots, K. \qquad (*)$$

$$\sum_{k=1}^{K} a_k R(k, l) = R(0, l), \quad l = 1, 2 \ldots, K,$$

Note (*) is also known as the orthogonality principle in estimation theory for a linear minimum mean square error estimator

Matrix Form

- The previous equation can be rewritten as:

$$\begin{bmatrix} R(1,1) & R(2,1) & \cdots & R(K,1) \\ R(1,2) & R(2,2) & \cdots & R(K,2) \\ \cdots & \cdots & \cdots & \cdots \\ R(1,K) & R(2,K) & \cdots & R(K,K) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_K \end{bmatrix} = \begin{bmatrix} R(0,1) \\ R(0,2) \\ \cdots \\ R(0,K) \end{bmatrix}$$

$$[\mathbf{R}]\mathbf{a} = \mathbf{r}. \qquad \mathbf{a} = [\mathbf{R}]^{-1}\mathbf{r}.$$

- Optimal solution:

$$\sigma_p^2 = E\{(\mathcal{S}_0 - \mathcal{S}_p)\mathcal{S}_0\} = R(0,0) - \sum_{k=0}^{K} a_k R(k,0)$$

$$= R(0,0) - \mathbf{r}^T\mathbf{a} = R(0,0) - \mathbf{r}^T R^{-1}\mathbf{r}.$$

Predictive Coding Gain

Distortion of optimal scalar quantizer

$$D_{\mathrm{DPCM}} = \epsilon_p^2 \sigma_p^2 2^{-2R} \qquad G_{\mathrm{DPCM}} = \frac{D_{\mathrm{PCM}}}{D_{\mathrm{DPCM}}} = \frac{\epsilon_s^2}{\epsilon_p^2} \frac{\sigma_s^2}{\sigma_p^2}$$

$$\sigma_{p,\min}^2 = \lim_{K \to \infty} \sigma_p^2 = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log_e S(e^{j\omega})\, d\omega \right)$$

Depends on the source

This is a spectral flatness measure: good prediction with narrow spectrum; poor prediction with nearly flat spectrum

$$\sigma_{p,\min}^2 = \lim_{K \to \infty} \left(\prod_k \lambda_k \right)^{1/N}$$

Typo:1/K

$$\sigma_s^2 = \lim_{K \to \infty} \frac{1}{K} \sum_k \lambda_k$$

$$\lim_{K \to \infty} G_{\mathrm{DPCM}} = \frac{\epsilon_s^2 \lim_{K \to \infty} \frac{1}{K} \sum_k \lambda_k}{\epsilon_p^2 \lim_{K \to \infty} \left(\prod_k \lambda_k \right)^K},$$

Typo: K → 1/K

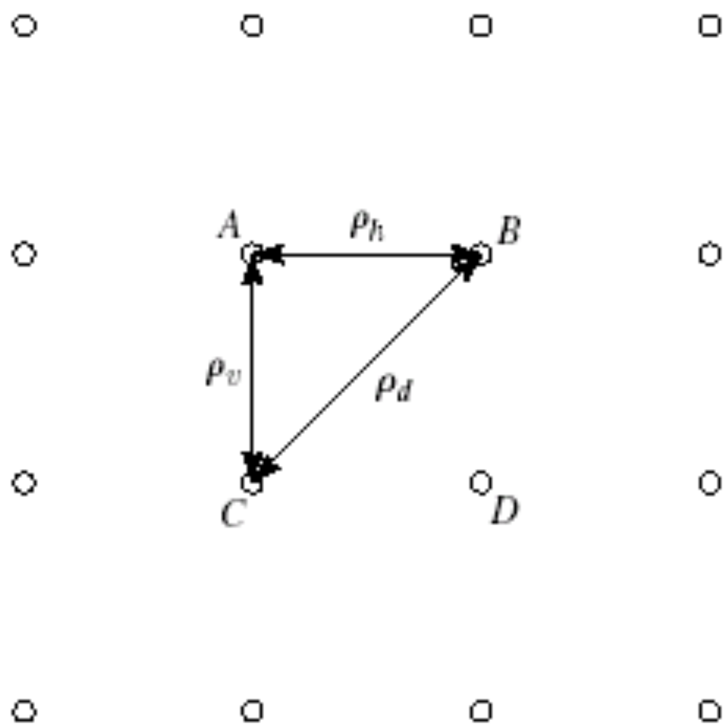
Predictive coding vs. transform coding

$$\lim_{K\to\infty} G_{\mathrm{DPCM}} = \frac{\epsilon_S^2}{\epsilon_P^2} \frac{\lim_{K\to\infty} \frac{1}{K}\sum_k \lambda_k}{\lim_{K\to\infty} \left(\prod_k \lambda_k\right)^K},$$

K → 1/K

- Transform coding and predictive coding have identical coding gains if the block length in TC and the predictive order in PC both go to infinity
- PC is better for any finite length
  - Predictive coding intrinsically captures information through feedback about an arbitrary amount of the past

Example



$$\hat{D} = a_1 C + a_2 B + a_3 A$$

Example Continued

$$\begin{bmatrix} R(C,C) & R(C,B) & R(C,A) \\ R(B,C) & R(B,B) & R(B,A) \\ R(A,C) & R(A,B) & R(A,A) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} R(D,C) \\ R(D,B) \\ R(D,A) \end{bmatrix}$$

$$\begin{bmatrix} 1 & \rho_d & \rho_v \\ \rho_d & 1 & \rho_h \\ \rho_v & \rho_h & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \rho_h \\ \rho_v \\ \rho_d \end{bmatrix}.$$

In the special case of $\rho_h = \rho_v = \rho$, $\rho_d = \rho^2$, the optimal predictor is

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \rho \\ \rho \\ -\rho^2 \end{bmatrix}.$$

The MSE of this predictor, using Equation (9.2.10), is

$$\sigma_p^2 = R(0,0) - [R(0,1) \quad R(0,2) \quad R(0,3)] \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = (1-\rho^2)^2 \sigma_s^2.$$

$$G_{\mathrm{DPCM}} = \frac{\sigma_s^2}{\sigma_p^2} = \frac{1}{(1-\rho^2)^2}$$

(DPCM is better than TC for this case!)
Recall

$$G_{\mathrm{TC}} = \frac{\sigma_s^2}{\sigma_t^2} = \frac{1}{1-\rho^2}$$

Predictive Coding for Video

- If predictive coding is better than transform coding, why don't we use predictive spatial coding?

- For video, we apply prediction both among pixels in the same frame (intra-prediction or spatial prediction), and also among pixels in adjacent frames (inter-prediction or temporal prediction)
- Temporal prediction is done with motion compensation
- More on this subject in the next lecture.