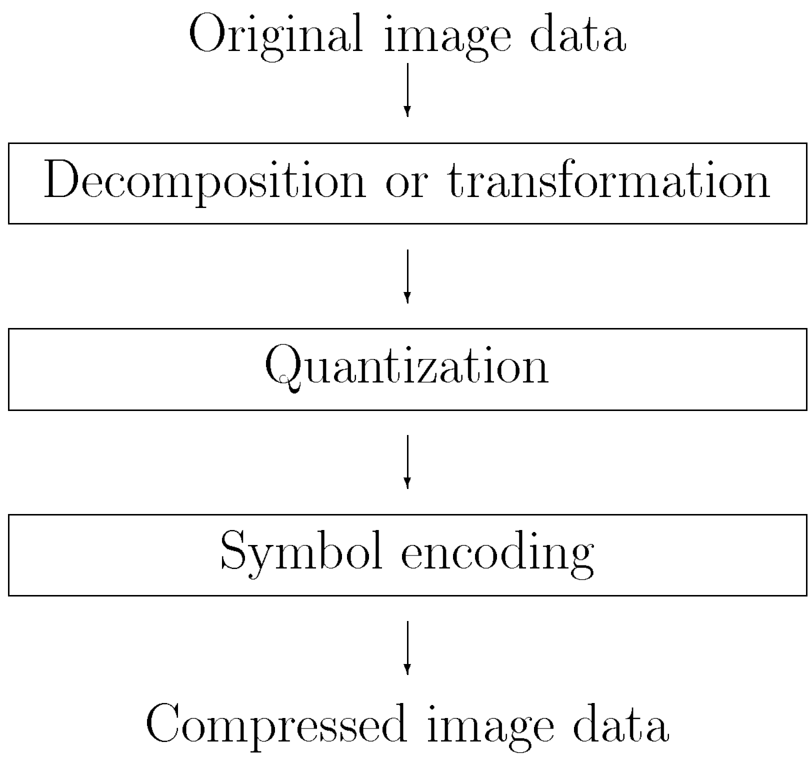
Part IV
Lossy Compression Techniques

- Introduction
  - degradations in reconstructed images in exchange for a reduced bit rate
  - degradations may or may not be visually apparent
  - greater compression achieved by allowing more degradations



- image decomposition or transformation:
  * reducing dynamic range of signal
  * eliminating redundant information
  * in general, providing a representation to be coded more efficiently

    * a reversible operation
    * both representations correspond to the same image
    * but the zeroth-order entropy of the transformed image is significantly lower than that of the original image, more amenable to efficient encoding

- The primary difference between lossy and lossless schemes:
  - inclusion of quantization to reduce the number of possible output symbols

- Any of the components of a lossy scheme may be implemented in an adaptive or non-adaptive mode.
  - adaptive: the structure of a component or its parameters changes locally within an image
  - taking advantage of variations in local statistics
  - causal or non-causal
    - causal: previously reconstructed pixel values
    - non-causal: previous pixel values+future input values

- most popular and widely used techniques
  - predictive coding
  - transform coding
  - block truncation coding (BTC)
  - vector quantization (VQ)
  - subband coding (SBC)
  - hierarchical coding
- Error metrics:

$$\text{RMSE} = \sqrt{\frac{1}{N^2}\sum_{i,j=1}^{N}[f(i,j) - \hat{f}(i,j)]^2}$$

$$\text{PSNR} = 20\log_{10}\left(\frac{255}{RMSE}\right)$$

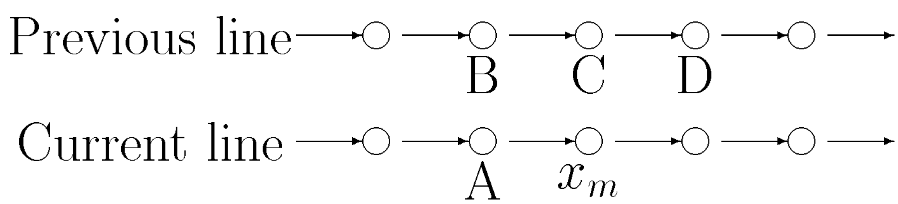
(for an 8-bit image)

- A lower RMSE or a higher PSNR does not necessarily imply a higher subjective reconstructed image quality; these error metrics do not always correlate well with perceived image quality.
  - Actual reconstructed images and error images shown in each implementation
    - error image g:

$$g(i,j) = 2[f(i,j) - \hat{f}(i,j)] + 128$$

Chapter 9
Lossy Predictive Coding

- Differential Pulse Code Modulation(DPCM)



$$\hat{x}_m = \sum_{i=0}^{m-1} \alpha_i x_i$$

– rounded to the nearest integer

– clipped to the range $[0, 2^n - 1]$ for an n-bit image

$$e_m = x_m - \hat{x}_m$$

- the differential image:
  - a greatly reduced variance
  - significantly less correlated
  - a stable histogram well approximated by a Laplacian distribution
- To lower bit rate, the differential image in lossy DPCM quantized before encoding and transmission

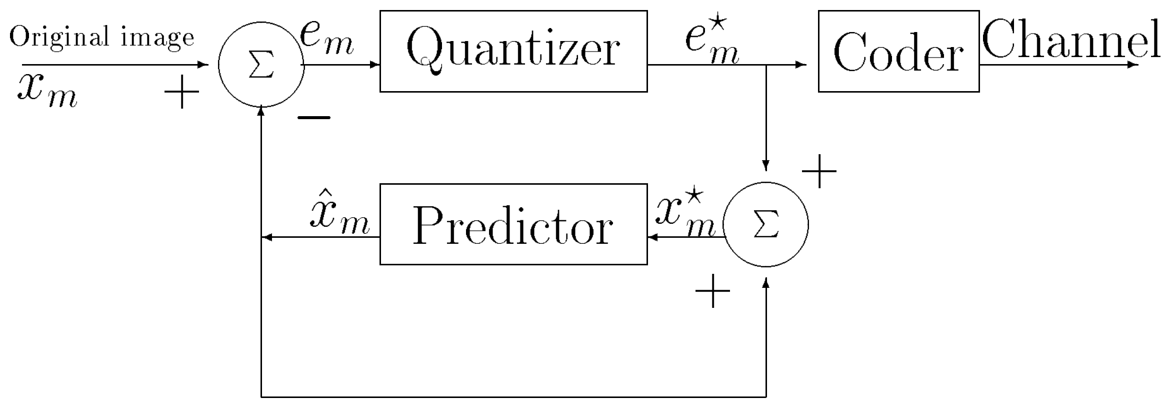
$$e_m \longrightarrow e_m^\star$$

$$e_m^\star + \hat{x}_m \longrightarrow x_m^\star$$

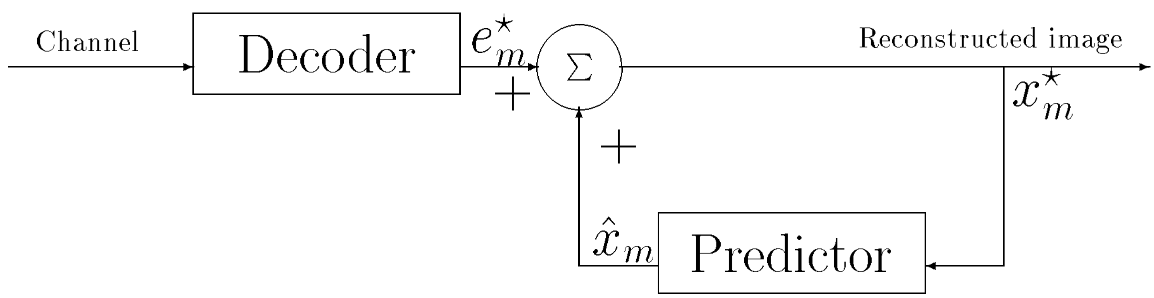
$$x_A^\star, x_B^\star, x_C^\star, x_D^\star \longrightarrow \hat{x}_m$$

$$x_m - \hat{x}_m \longrightarrow e_m$$

Transmitter



Receiver



- Design of DPCM
  - Optimizing the predictor and quantizer

- Predictor optimization

$$\min_{\alpha_0,\ldots,\alpha_{m-1}} \sigma_e^2 = E\{(x_m - \sum_{i=0}^{m-1} \alpha_i x_i)^2\}$$

$$\Rightarrow E\{(x_m - \sum_{i=0}^{m-1} \alpha_i x_i) x_k\} = 0$$

$$k = 0, 1, ..., m-1$$

  – 2-D stationary random field:

$$R_{k,l} = E\{x(i,j)x(i-k, j-l)\}$$

  – Local prediction impractical
  – The performance gain achieved by a local predictor over a global predictor marginal
  – Global prediction more attractive

- The selection of a robust set of global predictor coefficients for typical images:
  - A Markov model with a separable autocorrelation function:

$$R_{k,l} = \overline{x}^2 + \sigma^2 \rho_v^{|k|} \rho_h^{|l|}$$

    - $\overline{x}, \sigma^2$: mean and variance of the image
    - k, l: vertical and horizontal displacements
    - ρ_v, ρ_h: vertical and horizontal correlation coefficients
    - Mostly, $\rho_v, \rho_h \geq 0.9$

– For example, if $\overline{x} = 0$, then

$$\alpha_A = \rho_h, \alpha_B = -\rho_h\rho_v, \alpha_C = \rho_v, \alpha_D = 0$$

$$\hat{x}_m = \rho_h A - \rho_v\rho_h B + \rho_v C$$

B	C	D
A	x_m	

– The optimal fourth-order predictor has $\alpha_D = 0$ due to

$$R_{k,l} = \sigma^2\rho_v^{|k|}\rho_h^{|l|}$$

- If the image mean $\overline{x} \neq 0$,

$$\begin{aligned} E\{e_m\} &= E\{x_m - \sum_{i=0}^{m-1} \alpha_i x_i\} \\ &= E\{x_m\} - \sum_{i=0}^{m-1} \alpha_i E\{x_i\} \\ &= \overline{x}(1 - \sum_{i=0}^{m-1} \alpha_i) \\ &= 0 \quad \text{as} \quad \sum_{i=0}^{m-1} \alpha_i = 1 \end{aligned}$$

  - $z = 1$ becomes a pole for the linear prediction filter

$$\frac{1}{1 - \Sigma_{i=0}^{m-1} \alpha_i z^{i-m}}$$

  - unstable to channel errors

– some examples of typical predictors:

$$
\begin{aligned}
\hat{x} &= 0.97A, && \text{1st-order,1-D} \\
\hat{x} &= 0.5A + 0.50C, && \text{2nd-order,2-D} \\
\hat{x} &= 0.90A - 0.81B + 0.90C, && \text{3rd-order,2-D} \\
\hat{x} &= 0.75A - 0.50B + 0.75C, && \text{3rd-order,2-D} \\
\hat{x} &= A - B + C, && \text{3rd-order,2-D}
\end{aligned}
$$

- Quantizer optimization
  - Substantial compression by a lossy DPCM scheme is due to quantization of differential image
  - Statistical or visual criteria
  - HVS

- A quantizer: a staircase function:
  - map many input values into a smaller, finite number of output levels
  - e: real, random
  - $P_e(e)$: probability density
  - A quantizer: $e \longrightarrow e^\star \in \{r_i, i = 0, 1, ..., N-1\}$ (reconstruction levels)
  - Decision levels $\{d_i, i = 0, ..., N\}$
  - $e \in (d_i, d_{i+1}] \longrightarrow e^\star = r_i \in (d_i, d_{i+1}]$

- The quantizer design:
  - optimum decision and reconstruction levels for a given $P_e(e)$ and a given optimization criterion
- Depending on whether the quantizer output levels are encoded using variable-length or fixed-length codewords, two different types of quantizers are typically used in a DPCM system.
  - For fixed-length codewords, the DPCM bit rate $\sim \log_2 N$, N : #(quantizer levels)

- For a given N, it is desirable to design a quantizer that minimizes the quantization error.
- If the MSE criterion is used, the approach leads to the Lloyd-Max quantizer.
- Nonuniform decision regions

• For variable-length codewords, the bit rate is lower bounded by the entropy of the quantizer output, leading to the approach:

$$\begin{cases} \text{min quantization error} \\ \text{subject to entropy constraint} \end{cases}$$

- Since the quantizer input distribution is usually highly skewed, the use of variable-length coding seems appropriate.
  - For a Laplacian density and MSE distortion, the optimum quantizer is uniform.
  - It has more levels than a Lloyd-Max quantizer, but it also has a lower output entropy.
  - With a large number of quantizer levels, optimum variable-length coding improves the SNR by 5.6 dB over fixed-length coding at the same bit rate.

- Lloyd-Max quantizer:

$$\min D = \sum_{i=0}^{N-1} \int_{d_i}^{d_{i+1}} (e - r_i)^2 P_e(e) de$$

w.r.t. $\{d_i\}$ and $\{r_i\}$

$$\Longrightarrow$$

$$d_i = \frac{r_{i-1} + r_i}{2}$$

$$r_i = \frac{\int_{d_i}^{d_{i+1}} e P_e(e) de}{\int_{d_i}^{d_{i+1}} P_e(e) de}$$

- Adaptive DPCM
  - DPCM schemes can be made adaptive in terms of the predictor or quantizer or both.
  - Adaptive prediction reduces the prediction error $\Longrightarrow$ the reduced dynamic range of the quantizer input signals $\Longrightarrow$ less quantization error and better reconstructed image quality for the same bit rate

- Adaptive quantization reduces the quantization error directly by varying the decision and reconstruction levels according to the local image statistics.

- Adaptive prediction
  - Nonadaptive predictors perform poorly at edges
  - Adaptive predictors switch among a set of predictors based upon the most likely edge direction.
  - In particular,
    $\hat{x}_m = A$ or B or C or D

– For a causal decision, the reconstructed values of neighboring pixels are used.

– Improved perceived image quality

– SNR gains $\approx$ 4 dB

– In another causal technique,

$$\hat{x}_m \longrightarrow k\hat{x}_m$$

$$\vdots$$

adaptive coefficient

$*$ k $\sim$ the previous reconstruction level r_i

- Adaptive quantization
  - The quantizer levels for a given pixel are found by scaling the levels for the previous pixel by some factor which depends on the reconstruction level used for the previous pixel.
  - A more sophisticated approach makes use of the visual masking effects in the HVS.
  - It is well known that the luminance sensitivity of the HVS decreases in the picture areas of high-contrast detail.

- In these areas, large quantization errors can be masked by the HVS.
  * Sharma and Netravali (1978):
    quantizers with the minimum number of output levels subject to the constraint that the largest magnitude of the quantization errors from an arbitrary input is less than the visibility threshold.
  * Anastassiou et al. (1986):
    an estimate of the number of bits required to quantize the differential

image is made for each pixel.
The estimate is based on the previously reconstructed differential values and can be tracked by the receiver.

* Ready and Spencer (1975),
  Habibi and Batsen (1978):
  - The distribution of the differential image e_m is generally a function of the neighboring (past and future) pixel values (conditional!)

* Nonadaptive quantizers assume:
  e_m has a Laplacian pdf with a variance equal to the global variance of the differential image.
* For a given set of neighboring pixel values, the actual distribution of e_m may substantially differ from the above assumption:
  - The variance of e_m in flat regions $\ll$ the global variance

- The variance in highly textured areas > the global variance
- Near contours or high-contrast edges, the distribution may not even be symmetric.
- The system switches among a set of quantizers designed to accommodate the varying local statistics.
- In a practical system, the selection of a given quantizer can be made for a block rather than each individual pixel

- Partition each scan line into blocks of k pixels
- Encode the block using each of the m available quantizers
- Measure the distortion resulting from each quantizer
- Select the quantizer with minimum distortion
- Transmit $\log_2 m$ bits of overhead information per k-pixel block to identify the quantizer to the receiver
- Transmit the encoded signal for the block

- The length of the image block
- The number of quantizers
- The structure of each individual quantizer
- The distortion measure
  - $k = 6, 10, 16$?
  - $m = 4, 7$?
  - MSE

- DPCM results
  - Table 9.2:
    nonadaptive and adaptive (switched quantization) DPCM techniques using fixed 1-D and 2-D predictors
    1-D: $\hat{x} = 0.97A$
    2-D: $\hat{x} = 0.75A - 0.50B + 0.75C$
  - For nonadaptive DPCM:
    Lloyd-Max nonuniform quantizers with 2, 4, 8 output levels, fixed-length coding(1.00, 2.00, 3.00 bits/pixel)

– For ADPCM:
$m = 4$ (four quantizers)
$k = 10$ (10 pixel block length)
overhead 2 bits/10 pixels
=0.2 bit/pixel

* Each of the four quantizers was a scaled version of the same Lloyd-Max nonuniform quantizers:
  scale factors:
  0.5, 1.00, 1.75, 2.50
* The resulting bit rates:
  1.20, 2.20, 3.20 bits/pixel

- Implementation Issues and Complexity of ADPCM Algorithm