Lossy Video Compression

Why Another Video Compression?

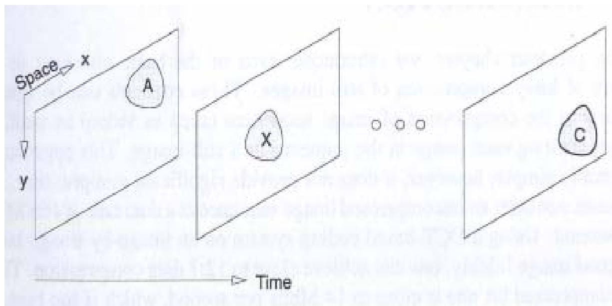
- **Video is a sequence of images.**
  - **Why can't we use lossy image compression schemes?**
- **Lossy image compression is not enough**
  - **Typical output rate of a CD-ROM drive: 1.5 Mbits/second, while**
  - **Full motion video at 30 fps and 720 x 480: 166 Mbits/second**
  - **110:1 compression ratio required**
  - **Hard to achieve 12:1 data compression ratio with lossy image compression alone**

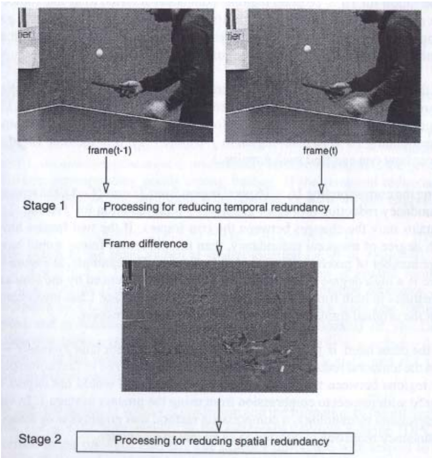
Video Coding Basics

- **Take advantages of temporal redundancy as well as spatial redundancy**
  - **3-D DCT?**
    - » **Computational complexity**
- **Two-stage process**

  Stage 1: Exploit temporal redundancy between frames

  Stage 2: Exploit spatial redundancy within the frame





Stage 1: Reducing Temporal Redundancy

- **Redundancy tracking unit**
  - Pixel vs. Block?
  - Video standards use macroblocks (MBs), 16 x 16 pixel regions
  - Why 16?
    - » Compromise between the temporal redundancy reduction efficiency and the implementation complexity
- **Interframe coding**

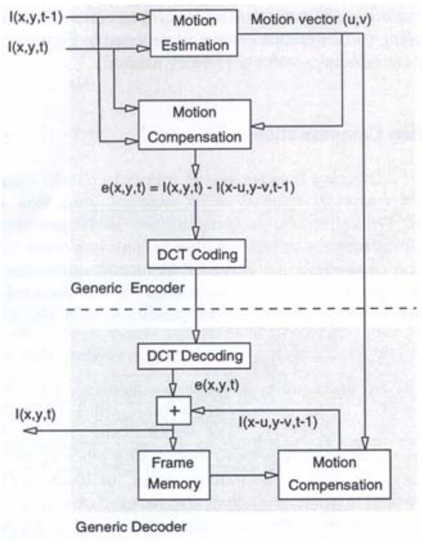
Stage 2: Reducing Spatial Redundancy

- **A DCT coding method is used for reducing spatial redundancy**
- **Intraframe coding**
- **Hybrid coding**

Temporal Prediction with Motion Compensation

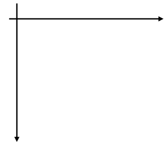
- ***Definition:* the process of computing changes among frames by establishing correspondence between frames**
- ***Motion Compensation*: the process of compensating for the displacement of moving objects from one frame to another**
- **Motion Compensation preceded by *Motion Estimation***
  - The process of finding corresponding pixels among frames
  - Outputs the coordinates (u, v) which defines the relative motion of a block from one frame to another
  - (u, v) called the motion vector for block at (x, y)
  - A hybrid video coding scheme

    = a DPCM method + a DCT coding



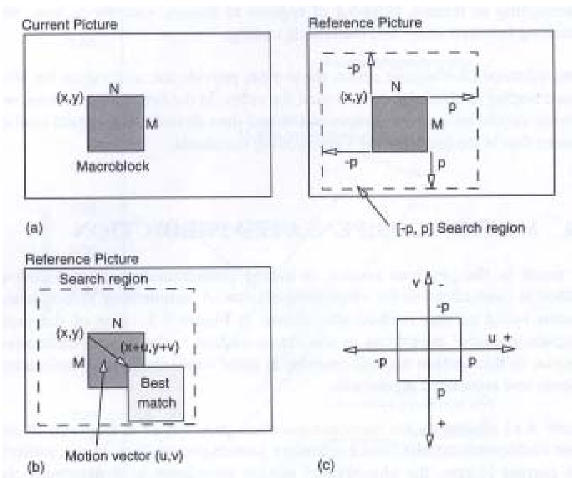
Motion Estimation Problem

- One of the most computing-intensive steps in interframe coding
  - » Search region: [-p, p]
    - p = 15 for head-and-shoulders-type video scenes
    - p = 63 for sporting events
  - » a motion vector is expressed in relative coordinates
    - (x, y) is at location (0, 0) and the motion vector is (u, v)



Target to prediction MB

– In video coding standards, only translatory motion model is assumed.



Macroblock Dimensions

- Tradeoff among:
  - – Smoothness constraint
    - » Small values for N and M are preferable
  - – Reliability of the motion vector (u, v)
    - » Small values for N and M reduce reliability
  - – Efficient implementation
    - » More efficient for large values of N and M

Block Matching Criterion

- Mean absolute error (MAE) / Mean absolute difference (MAD)

MAE(i, j) =

- Mean squared error or correlation between blocks are also valid choices
  - Expensive to implement
  - MAE does a good job

Algorithms for Motion Estimation

- Assume that MAE is used as a block matching criterion
- Goal:
  - find the motion vector that gives the smallest MAE within the search space
- Alternatives:
  - Full search algorithm
  - Logarithmic search
    - » Three step search
  - Parallel hierarchical one-dimensional search
  - Pixel projections
  - Hierarchical motion estimation

Full Search Algorithm

- Guarantees finding the minimum MAE value
- Computationally expensive
  - (2p + 1)^2 search locations
  - N x M pixels in macroblocks
  - 3 operations per pixel comparison
  - For an I x J resolution picture with a picture rate of F per second,

$$\frac{I\,J\,F}{M\,N}(2p+1)^2 \text{ x } MN \text{ x } 3$$

  - » For N = M = 16, I = 720, J = 480 and F=30,
    - 29.89 GOPS for p=15 and
    - 6.99 GOPS for p=7

How To Reduce The Complexity

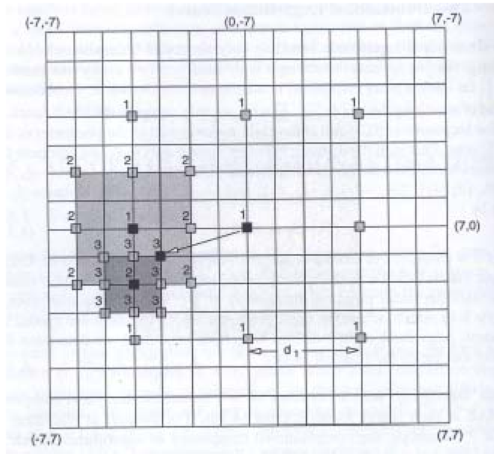
- Decrease the number of search locations
- Compute MAE(i, j) using fewer than N x M pixel differences per search location

2-D Logarithmic Search

- Divide [-p, p] search rectangle into two areas:
  - Inside [-p/2, p/2] rectangle
  - Outside [-p/2, p/2] rectangle
- Compute MAE for nine locations:
  - (0, 0),
  - (0, d1), (0, -d1), (-d1, 0), (d1, 0)
  - (d1, d1), (d1, -d1), (-d1, d1), (-d1, -d1)

d1 = 2^(k-1), and k = ceiling($\log_2 p$)
d1, d2, …, di = 2^(k-1-I), … dk = 2^0 = 1

- Computational cost:
  - 8k + 1 search locations
- Three Step Search where k=3 and p=7

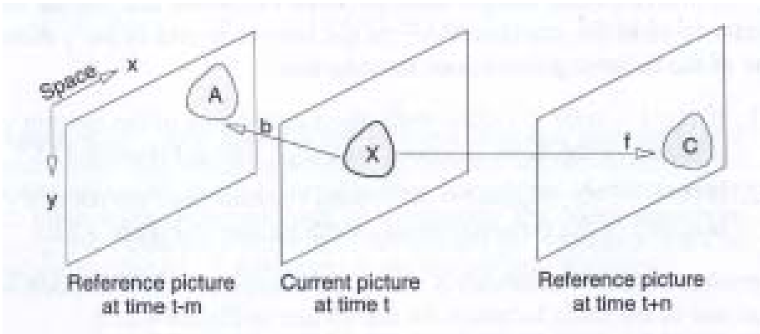


Sub-Pixel Accurate Motion Estimation

- Idea: If displacement estimates were obtained at a finer resolution, prediction accuracy will improve
- Video standards allow motion vectors to be a half-pixel accuracy

Multipicture Motion Estimation

- Bidirectional Prediction
- Prediction Choices:
  - forward direction or
  - backward direction or
  - an average of two directions


Space
x
y
A
b
X
f
C
Reference picture
at time t-m
Current picture
at time t
Reference picture
at time t+n