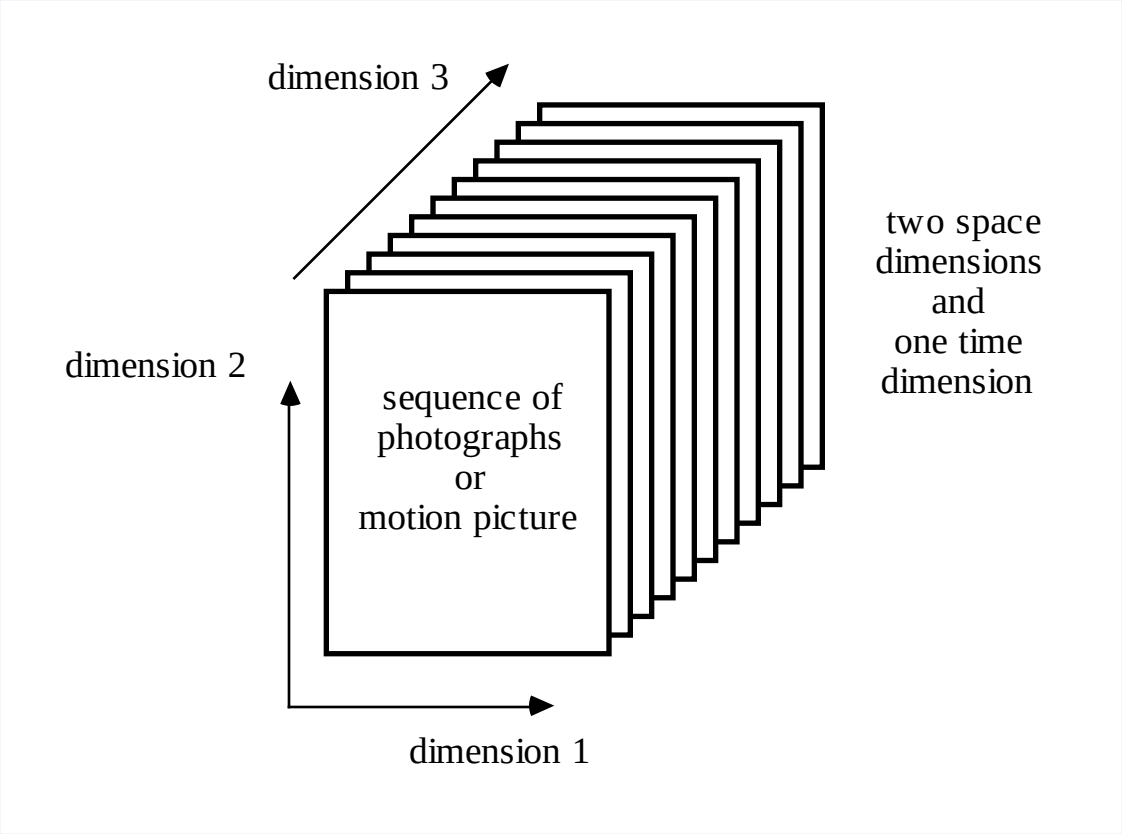
Image Processing and Visual Communications

Motion and Video Processing

Zhou Wang

Dept. of Electrical and Computer Engineering
University of Waterloo

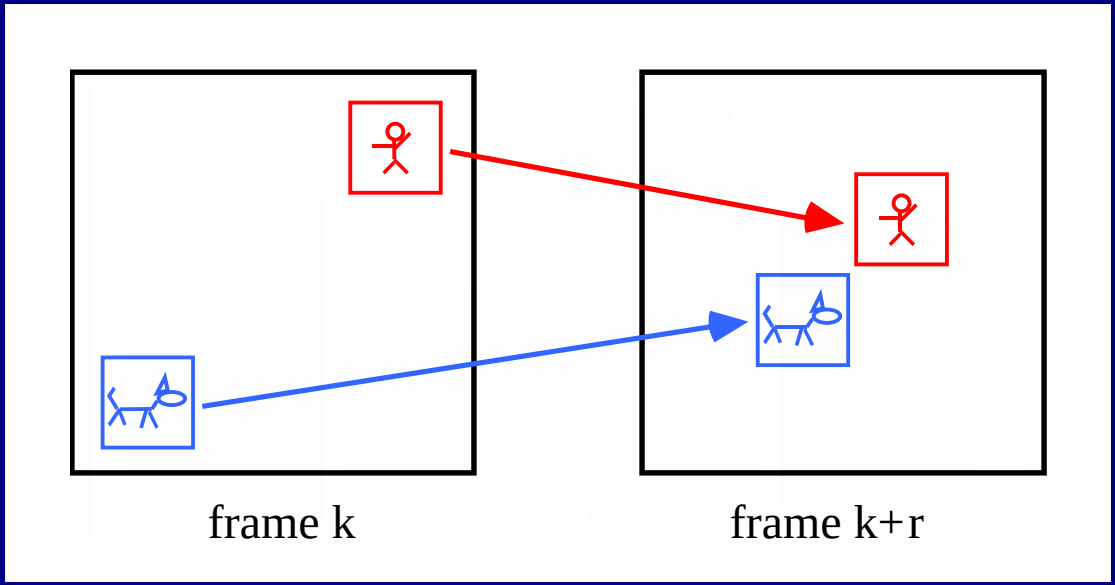
From Image to Video



- **What's new other than just a stack of still images?**
  - Strong **correlation** between frames
  - Representing **motion**

Motion Estimation: Block Matching

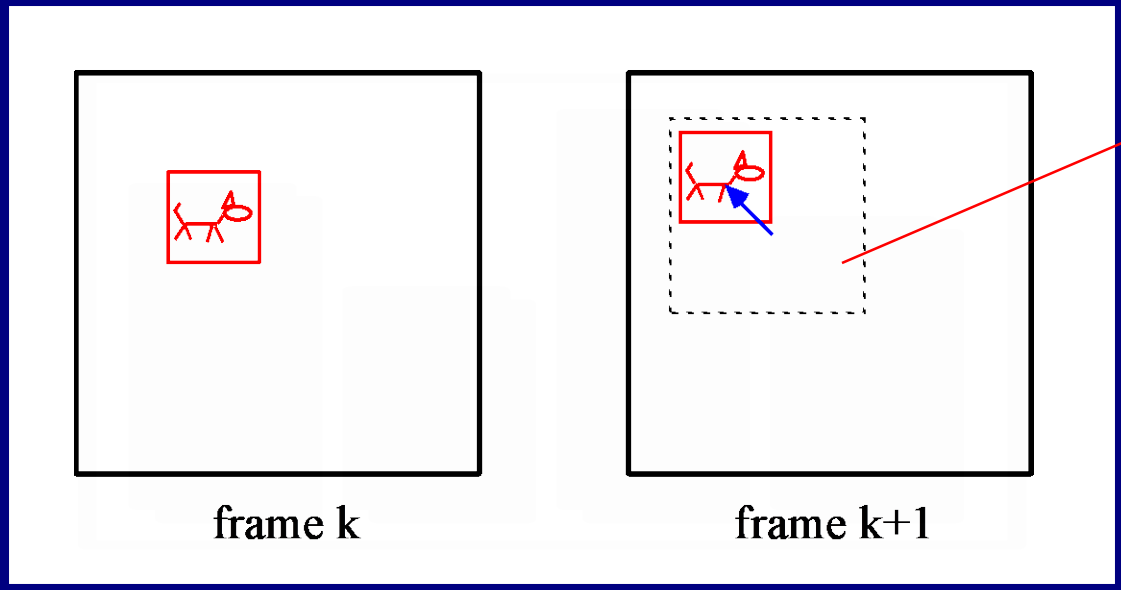
- **Estimate Motion**
  - Equivalently: Find correspondence between frames
- **Block Matching (template matching)**
  - For every block in one frame, find the best match in another
  - Requires search



From Prof. Al Bovik

Motion Estimation: Block Matching

- **Search for Best Matching Block**
  - Search is conducted in a neighborhood regions

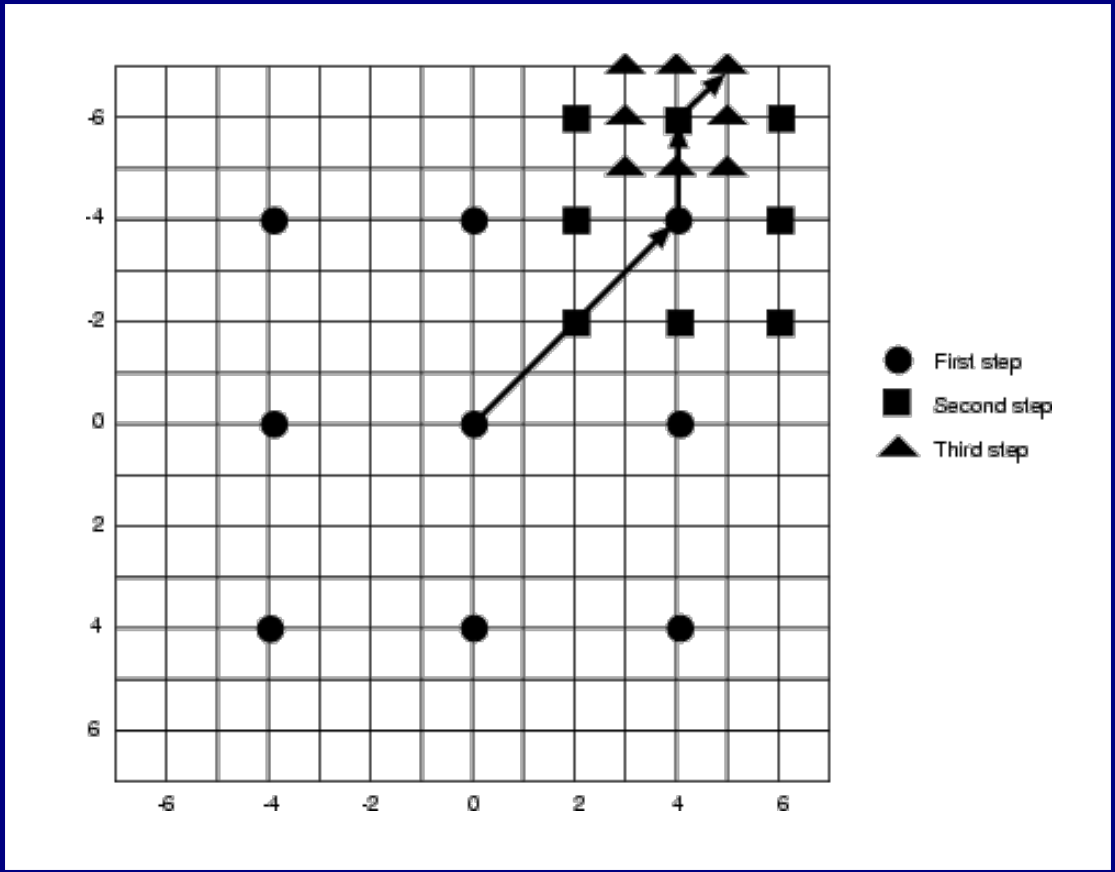


From Prof. Al Bovik

- **Search Algorithm: Exhaustive Search**
  - Try all possible blocks with a sliding window → slow

Motion Estimation: Block Matching

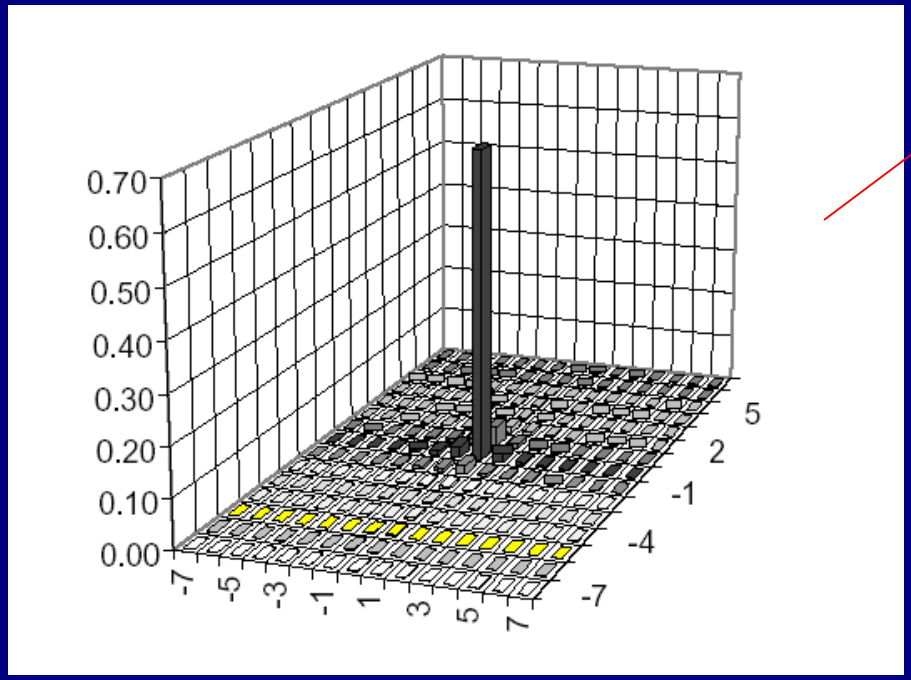
- **Fast Search Algorithm: 3-Step Search**



From Prof. Xin Li

Motion Estimation: Block Matching

- **Improved search: use statistics of motion vectors**



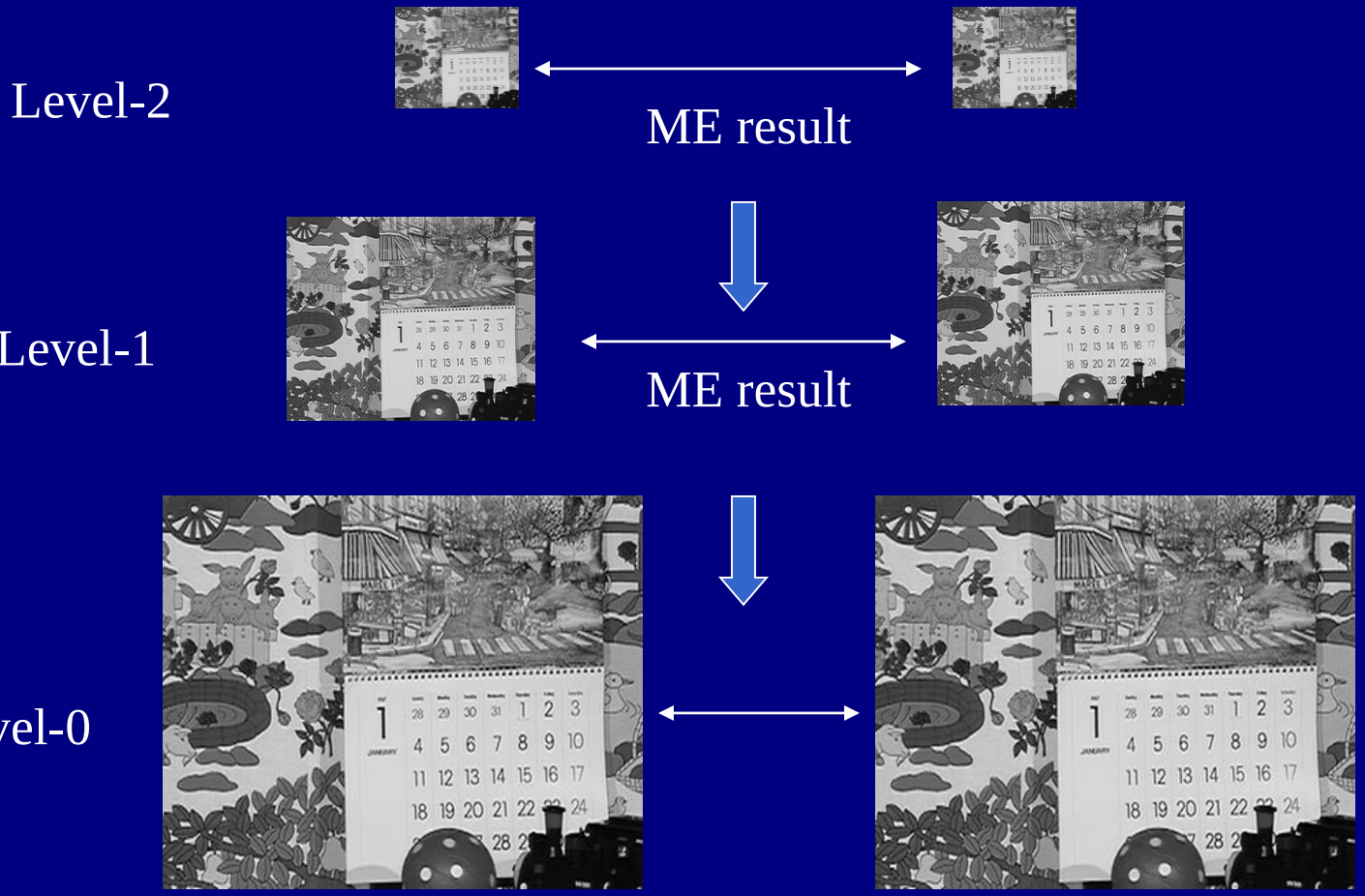
empirical motion vector distribution

Prior knowledge about motion in natural video

Search slow motion earlier

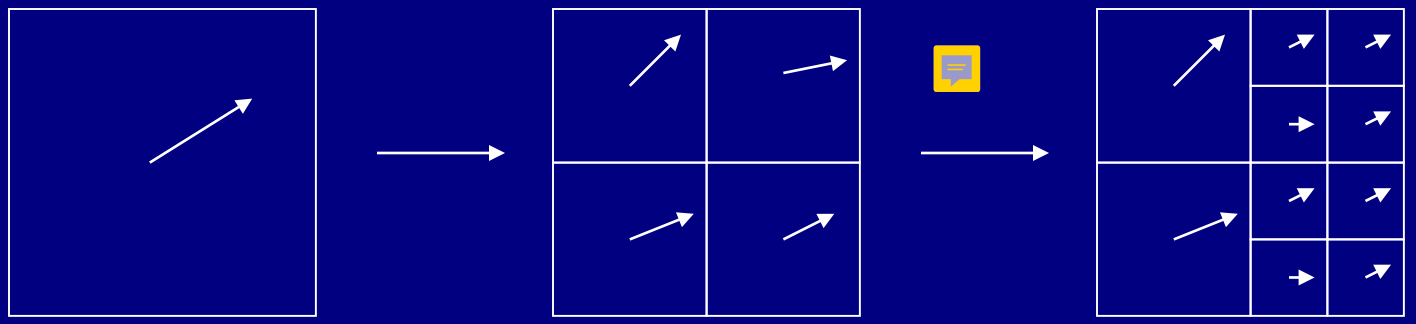
Motion Estimation: Block Matching

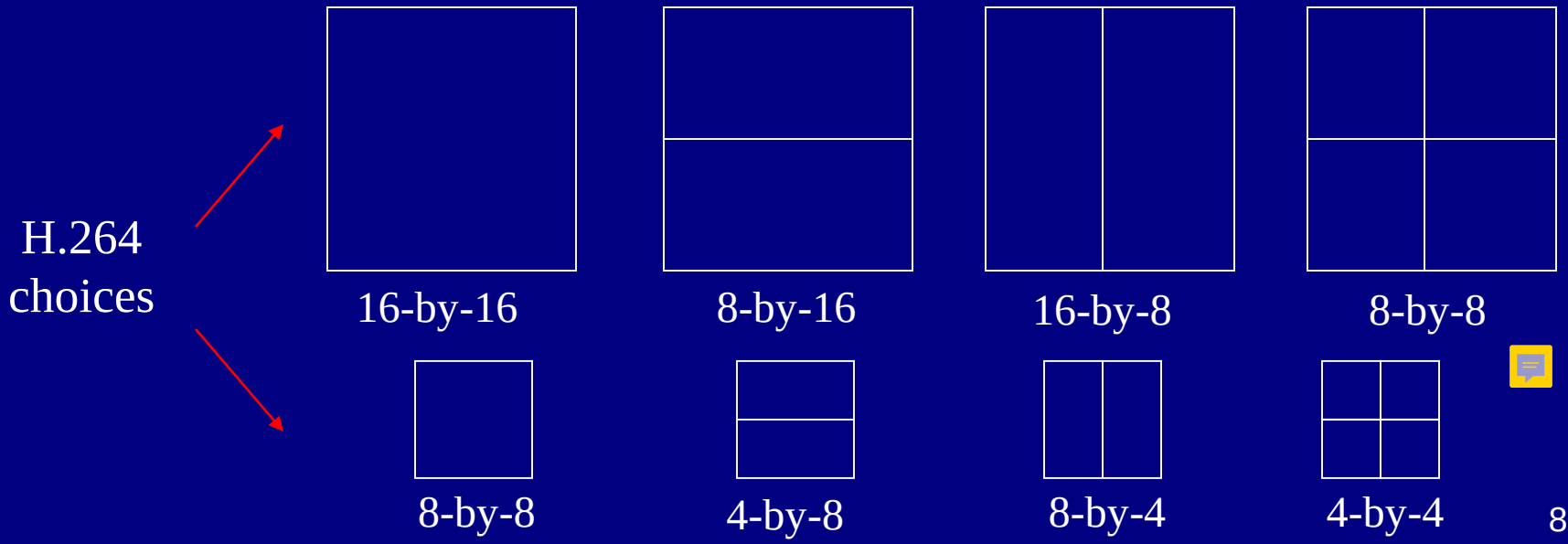
- **Improved search: hierarchical search**



Motion Estimation: Block Matching

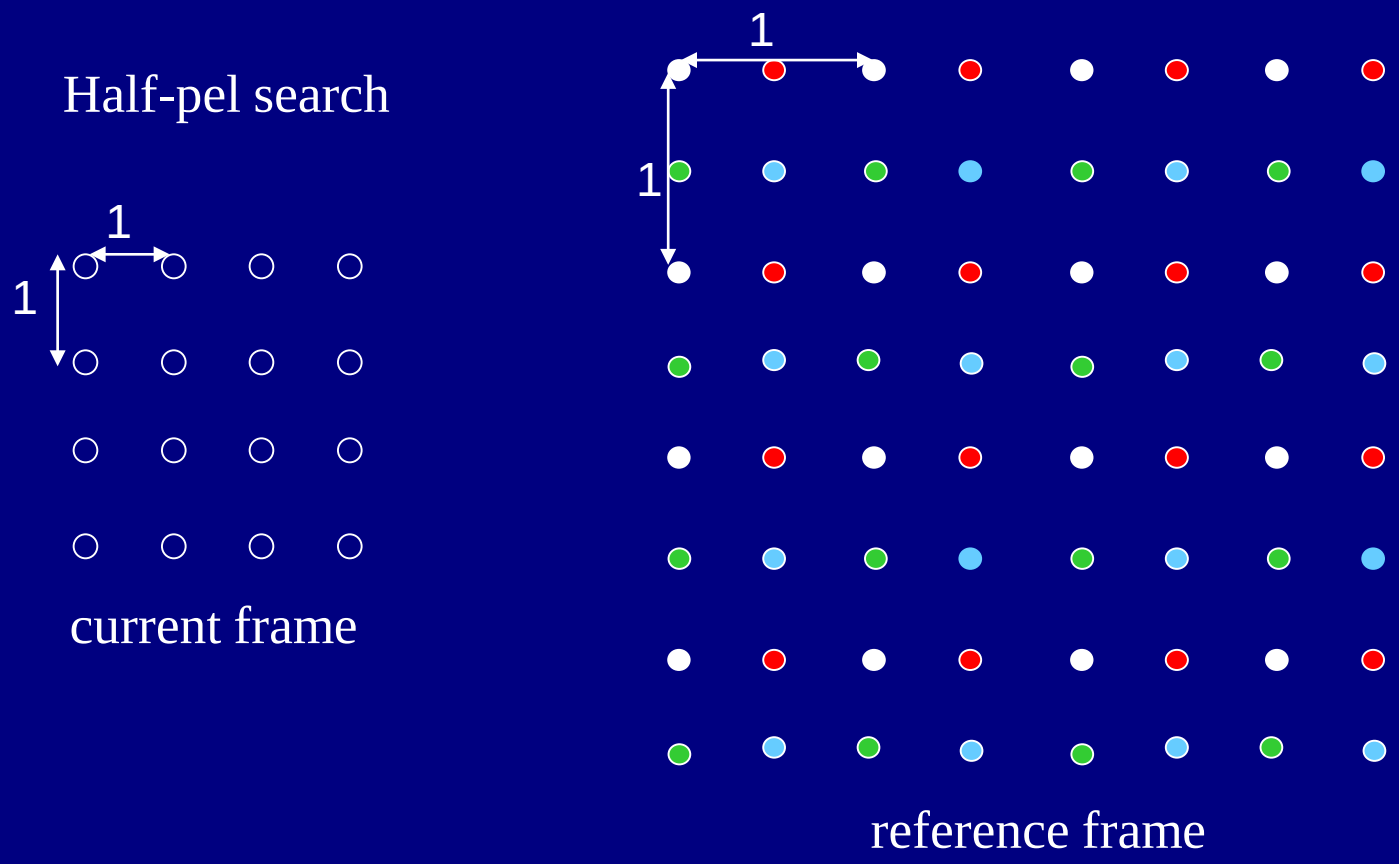
- **Improved search: variable block size search**





Motion Estimation: Block Matching

- **Improved search: fractional pixel search**

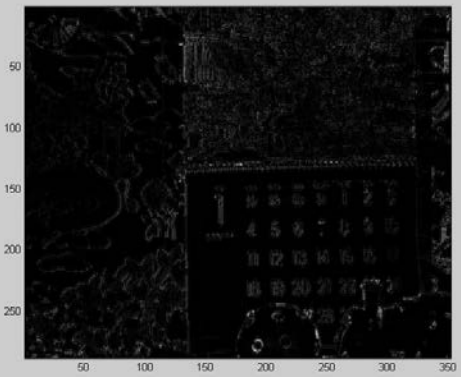


Can go to quarter-pel or even 1/8th-pel

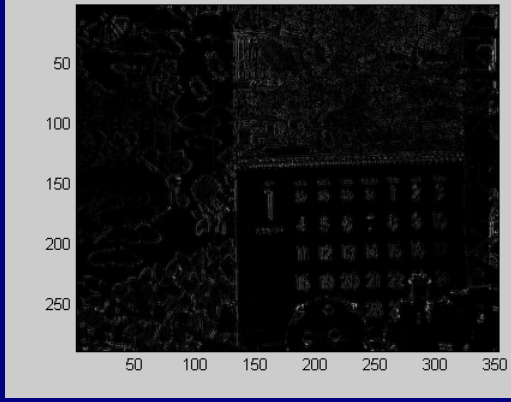
Motion Estimation: Block Matching

- **Improved search: fractional pixel search**

Frame 1

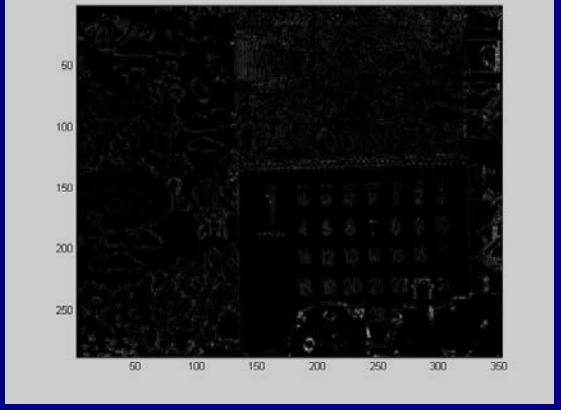


16x16 block, integer-pel, var(e)=272

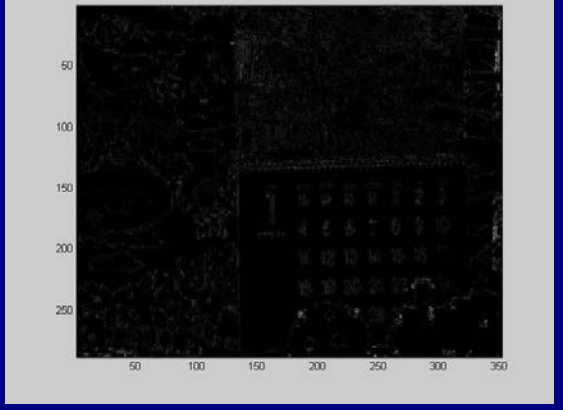


8x8 block, integer-pel, var(e)=221

Frame 2



16x16 block, half-pel, var(e)=164

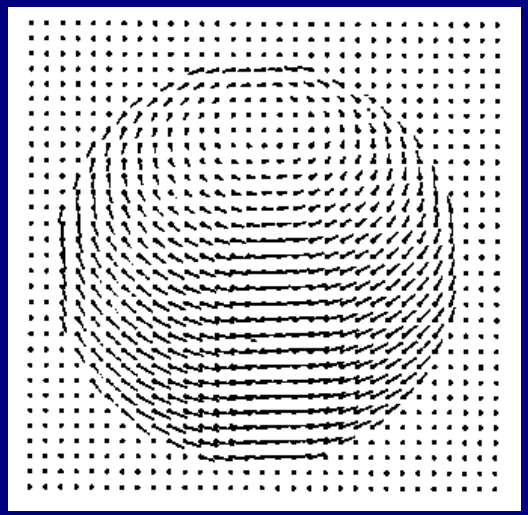


8x8 block, half-pel, var(e)=124

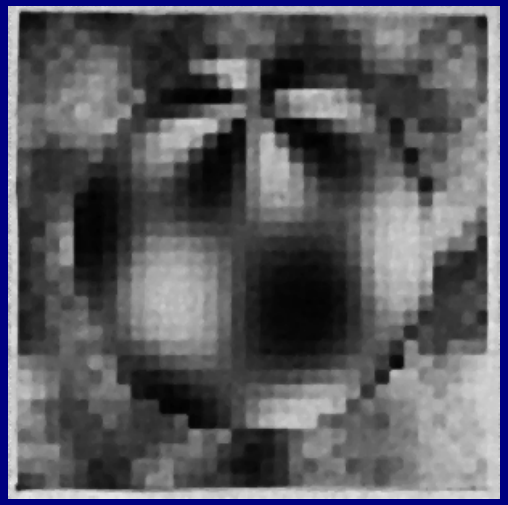
Motion Estimation: Optical Flow

Frame 1

Frame 2



optical flow



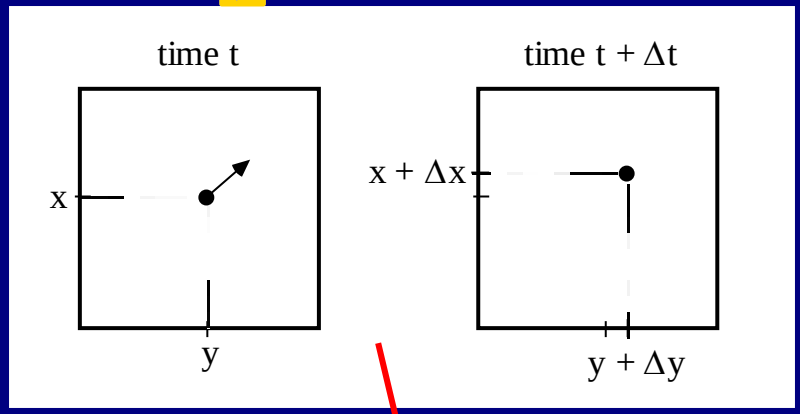
Frame 3

Frame 4

Motion Estimation: Optical Flow

Brightness constant constraint:

$$I(x+\Delta x, y+\Delta y, t+\Delta t) = I(x,y,t)$$



Taylor series expansion
(ignore higher-order terms)

$$I(x,y,t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t \approx I(x,y,t)$$

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y}\frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t} = 0$$

Is this a good assumption?

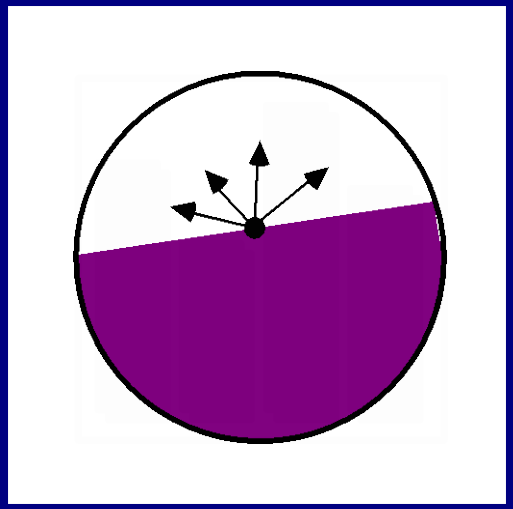
Optical flow equation:

$$I_x v_x + I_y v_y + I_t = 0$$

$$(\nabla I)^T \mathbf{v} + I_t = 0$$

Motion Estimation: Optical Flow

Aperture problem:



The problem of the aperture problem is not the aperture!

$$I_x v_x + I_y v_y + I_t = 0$$

two unknowns, one equation:
need additional constraints

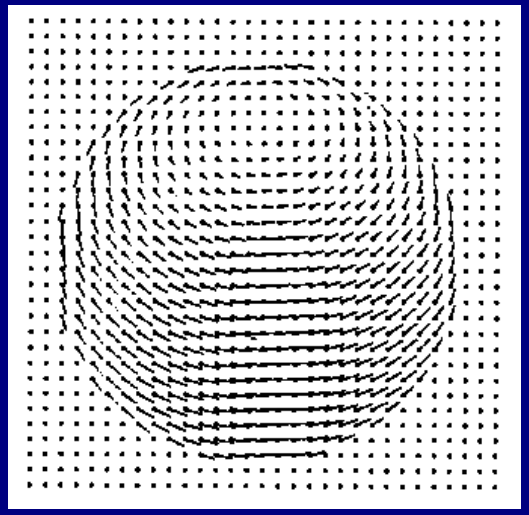
Additional constraints:

- Slow motion (choose the slowest of all solutions)
- Smooth motion (motion field changes smoothly)
- Rigid motion (consistent motion vectors within rigid objects)
- Bayesian motion (prior knowledge about the probability distribution of motion)
- ……

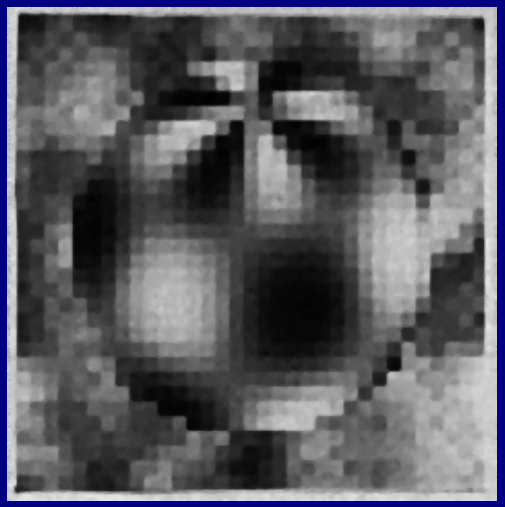
Motion Estimation: Optical Flow

Frame 1

Frame 2

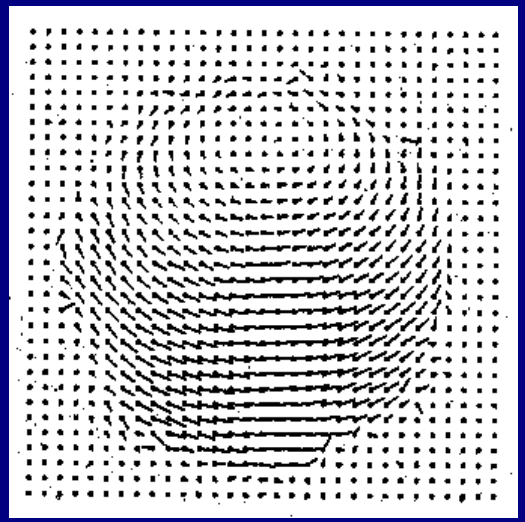


true optical flow



Frame 3

Frame 4



estimated optical flow

Application: Video Compression

- **Frame by Frame Coding VS. Motion Predictive Coding**
  - Frame by frame: code individual frames independently
  - Important observation: high correlation between frames
  - Motion predictive coding: remove **temporal redundancy**
  - Typically higher CR with motion prediction

- **Intraframe and Interframe Coding**
  - Intraframe coding
    Code one video frame independently
  - Interframe coding
    Code a video frame with information about neighboring frames
  - Used jointly in typical video coders

Video Compression: Temporal Predictive Coding

- **Direct Temporal Prediction**

– Intraframe:

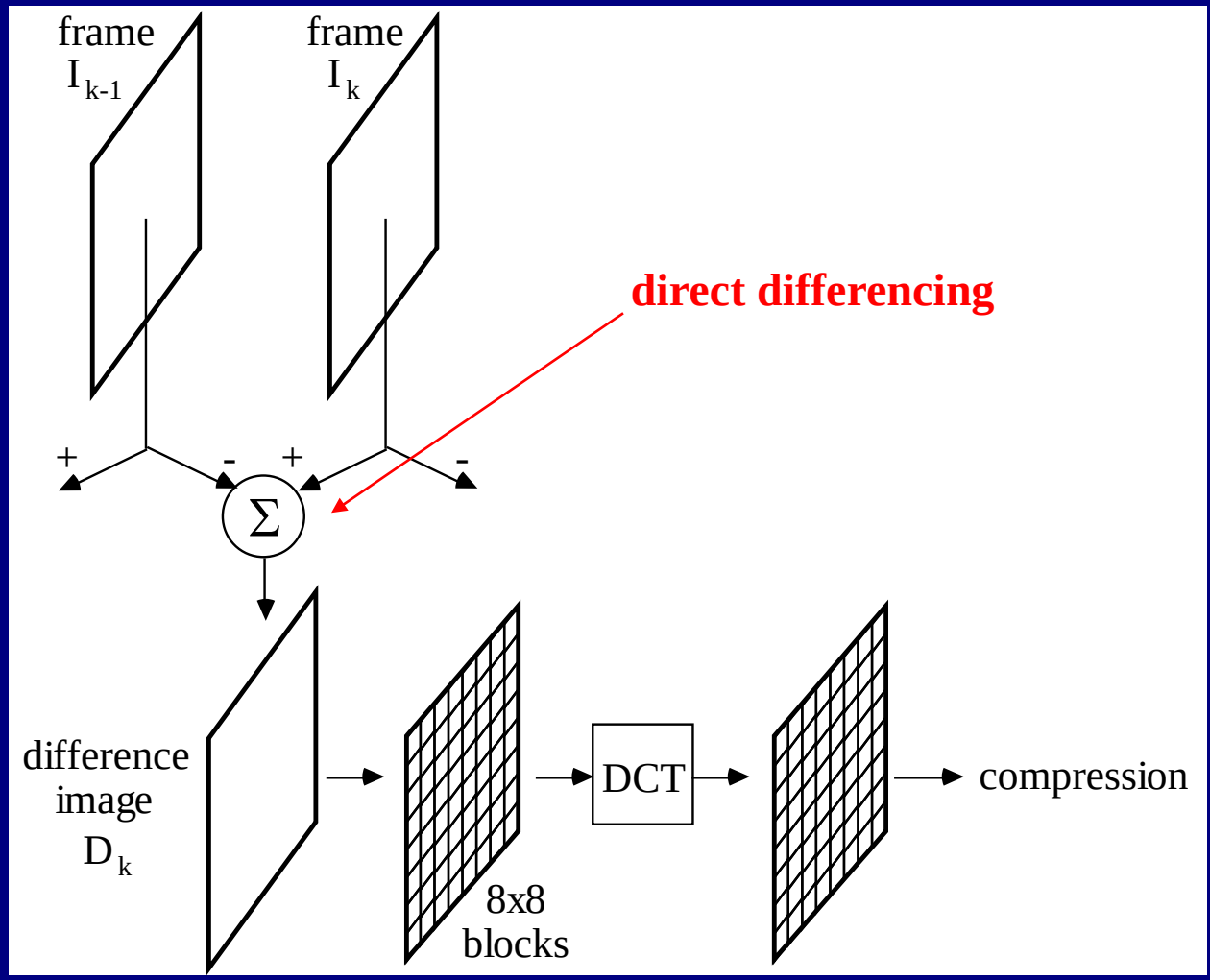
JPEG like coding

– Interframe:

Temporal differencing

+

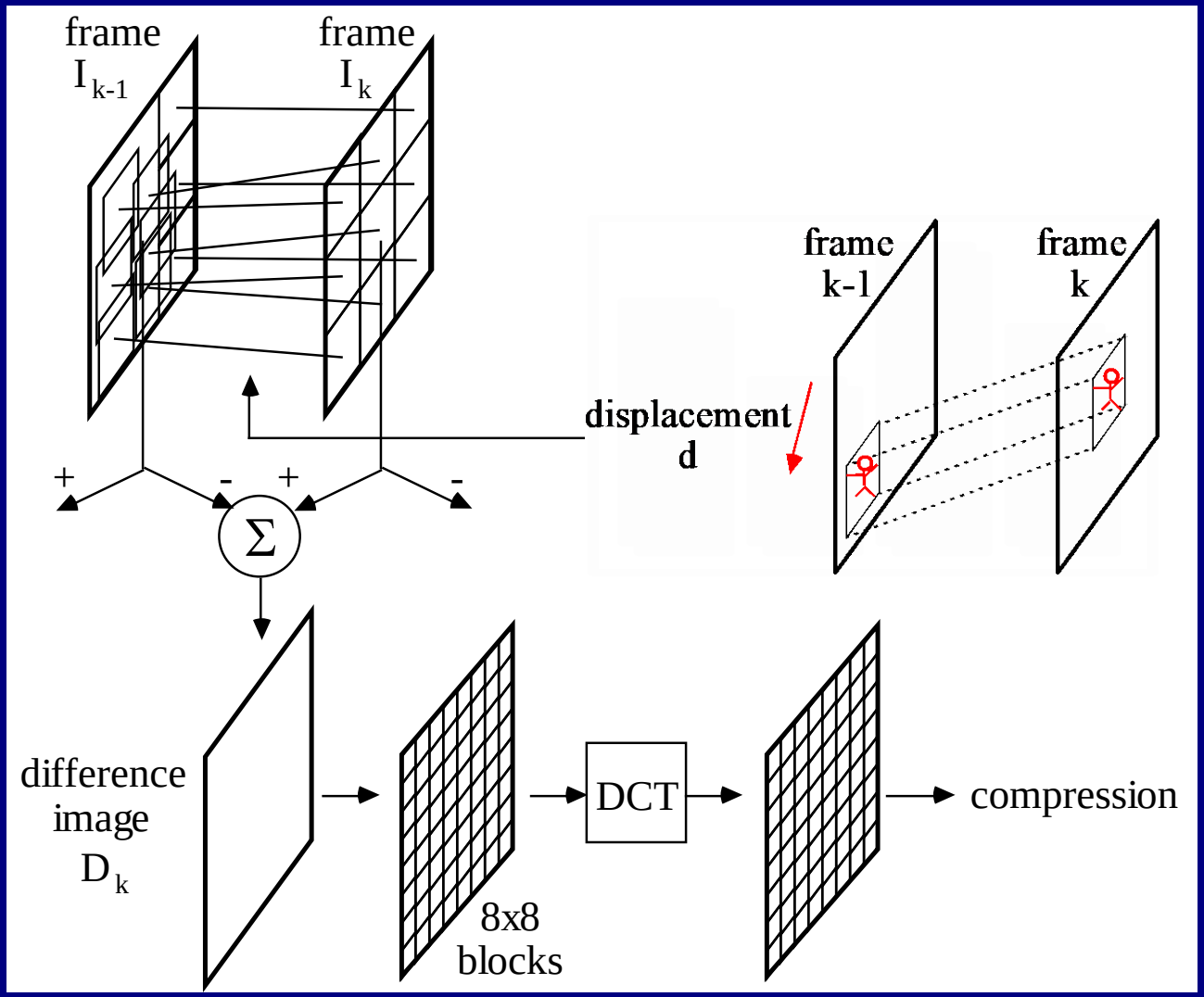
JPEG like coding



From Prof. Al Bovik

Video Compression: Temporal Predictive Coding

- **Motion Estimation/Compensation-Based Coding**



From Prof. Al Bovik

reduced frame prediction error

↓

reduced redundancy

↓

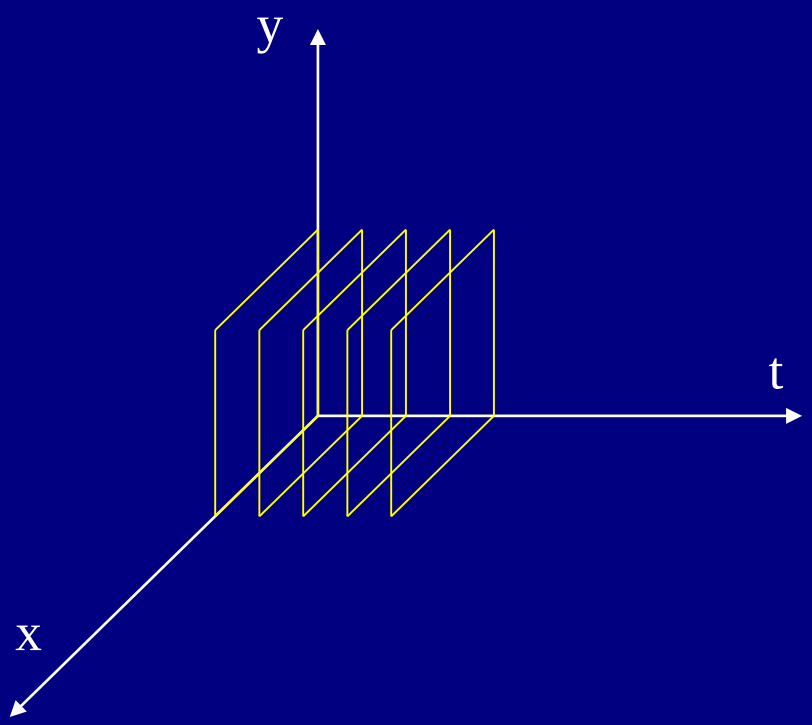
increased CR (and increased computational complexity)

Adopted by most video coding standards:

MPEG-x, H.26x, …

Video Compression: 3-D Transform Coding

- **3-D Transform Coding**



2D spatial +1D temporal transform

– Direct 3-D transform coding, e.g.

8x8x8 cubic DCT/DWT

+

scalar quantization or bitplane coding

– Motion compensated 3-D DWT

reduced frame prediction error
may create holes when filtering
increased computational complexity
increased memory requirement

(on going research …)

More Video Processing

- **Video Filtering/Restoration/Denoising**
  - Intra- and inter-frame (motion compensated filtering)
- **Video Segmentation**
  - Motion provides additional clues for segmentation
  - Segmentation can help refine motion estimation too
- **Target Tracking**
  - Extension of motion estimation (motion + geometric change)
  - Difficulties: joint estimation, occlusion …
- **Video Information Retrieval**