# Video Phones

Application of Structure from Motion

# Video Compression

Mubarak Shah

## What is Compression?

- Compression is a process of converting data into a form requiring less space to store or less time to transmit, which permits the original data to be reconstructed with acceptable precision at a later time.

## Orange Juice Analogy!

- Freshly squeezed orange juice (uncompressed)
- Remove water (redundancy), convert it to concentrate (encoding)
- Shipped, stored, and sold.
- Add water to concentrate (decoding), tastes like freshly squeezed!!!

## Why is compression necessary?

- Storage space limitations
- Transmission bandwidth limitations.

## Resolution

- QCIF: 180 x 144
- MPEG: 352 x 288
- VGA: 640 x 480
- NTSC 720x486
- Workstation 1280x1024
- HDTV: 1920 x 1080
- 35mm slide: 3072 x 2048

## Floppy Disk

- Floppy disk capacity = 1.44 MB
- A single 1280x1024x24 image= 3.9 MB
- A single 640x480x24=922kB
- Floppy disk holds only one VGA image!

## CD-ROM

- Capacity=600 MB
- A 1280x1024x24 @30 fps=118MB/s
- CD-ROM would hold only about 5 sec of video!
- A 160x120x16 image @30 fps=1.15MB/sec
- CD-ROM now holds 8.7 minutes of video

## DVD-ROM

- Capacity 2.4 GB to 15.9 GB
- Single side/single layer→Double side/dual layers
- 4.4 to 25 times capacity of CD ROM
- 20 sec to 2 minutes of 1280x1024x24 @30 fps
- 3 hours of 160x120x16 image @30 fps

## Bandwidth

- 160X120=1.15 MB/sec
- Quad-speed CD-ROM drive delivers 600 KB/sec (half of the required speed)
- DVD ROM delivers from 4Mbs to 9.8Mbs
- “T1” line delivers 1.54 Mb/sec (192KB/sec)
- Ethernet delivers 10Mb/sec (1.25 MB/sec) (barely fast enough, will use up entire bandwidth, 2-way video not possible)

## Digital TV

- Networks started broadcasting limited DTV programs in Nov 98.
- All commercial stations are supposed to switch to DTV by 2002
- All stations are supposed to switch to DTV by 2003
- Govt wants broadcasters' NTSC channels returned by 2006 for auctioning!

## Digital TV

- CBS carried few NFL games last year
- CBS and ABC plans
  - evening news
  - movies
  - rest of the day up-convert standard TV
- NBC
  - no broadcast yet
  - plans for "Tonight Show" this fall!

## Digital TV

- CBS and NBC use 1080i (1920X1080), which is 995Mb/s at 30 fps
- ABC and Fox use 720p (1280X720), which is 424Mb/s at 30 fps
- 6 MHz channel assigned to each network can carry 19.4Mb/s
- Need 50:1 compression ratio!

## Why is compression acceptable?

- Limitations of visual perception
  - Number of shades (colors, gray levels) we can perceive
  - Reduced sensitivity to noise in high-frequencies (e.g. edges of objects)
  - Reduced sensitivity to noise in brighter areas
- Ability of visual perception
  - Ability of the eye to integrate spatially
  - Ability of the mind to interpolate temporally

## Why is compression acceptable?

- Some type of visual information is less important than others
- Goal is to throw away bits in psycho-visually lossless manner
- We have been conditioned to accept imperfect reproduction
- Limitations of intended output devices

## Why is compression possible?

- Some sample values (gray levels, colors) are more likely to occur at a particular pixel than others.
  - Remove spatial and temporal redundancy that exist in natural video
    - Correlation itself can be removed in a lossless fashion
    - Important to medical applications
    - Only realizes about 2:1 compression

## Why is compression possible?

- No single algorithm can compress all possible data
- Random data cannot be compressed

## Lossless Compression

- Needed when loss is unacceptable or highly undesirable
- Fixed compression ratio is hard to achieve
- Compression/decompression time varies with image

## Lossy Compression

- Used when loss is acceptable or inevitable
- Permits fixed compression ratios
- Better suited for fixed time decompression

## Compression Techniques

- Subsampling
- Quantization
- Delta Coding
- Prediction
- Color space conversion
- Huffman coding
- Run-length encoding
- De-correlation
- Motion Compensation
- Model-based compression

## Other Techniques

- Fractals
- Wavelets
- Vector Quantization
- K-L Transform
- ...

## Compression using original source

- For best compression, get the original source material and try to *understand* its properties.
  - Email messages are far smaller than fax, voice mail or video mail.
  - A musical score is far more compact than a digitized recording

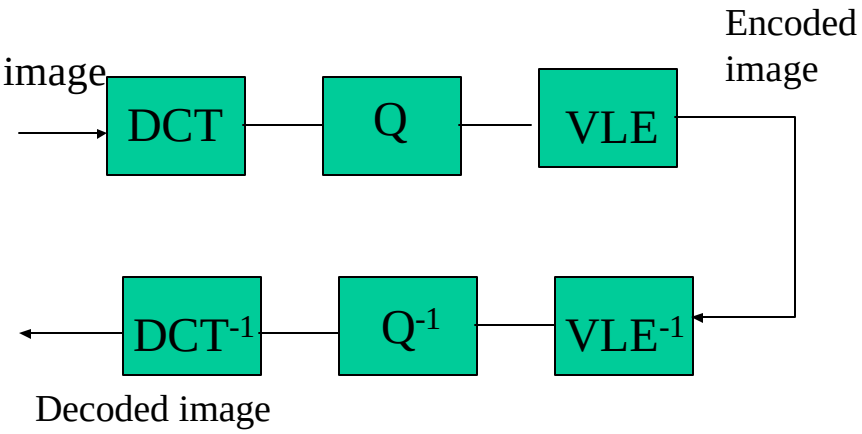
## Compression of Synthesized Image or Video

- For synthesized image or video clip it is far more efficient to transmit original source material and re-synthesized the image or clip at the receiver than to transmit the compressed image or video clip.

## How to Select Compression Scheme?

- High quality reproduction?
- Very high compression ratio?
- Fixed compression ratio?
- Real-time compression?
- Real-time decompression?
- Limited de-compression computer power?

# JPEG BLOCK DIAGRAM




# RLE: Example

8
0, 4, 4
1, 2, 5
1, 5, 2
1, 3, 2, 1, 1
2, 1, 2, 2, 1
0, 4, 1, 1, 2
8

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

## JPEG Baseline Coding

- Divide image into blocks of size 8X8.
- Level shift all 64 pixels values in each block by subtracting $2^{n-1,}$ (where $2^n$ is the maximum number of gray levels).
- Compute 2D DCT of a block.
- Quantize DCT coefficients using quantization table.

## JPEG Baseline Coding

- Zig-zag scan the quantized DCT coefficients to form 1-D sequence.
- Code 1-D sequence (AC and DC) using JPEG Huffman variable length codes.

# Example (Encoding)

$$I = \begin{bmatrix} 52 & 55 & 61 & 66 & 70 & 61 & 64 & 73 \\ 63 & 59 & 66 & 90 & 109 & 85 & 69 & 72 \\ 62 & 59 & 68 & 113 & 144 & 104 & 66 & 73 \\ 63 & 58 & 71 & 122 & 154 & 106 & 70 & 69 \\ 67 & 61 & 68 & 104 & 126 & 88 & 68 & 70 \\ 79 & 65 & 60 & 70 & 77 & 68 & 58 & 75 \\ 85 & 71 & 64 & 59 & 55 & 61 & 65 & 83 \\ 87 & 79 & 69 & 68 & 65 & 76 & 78 & 94 \end{bmatrix} \quad I' = \begin{bmatrix} -76 & -73 & -67 & -62 & -58 & -67 & -64 & -55 \\ -65 & -69 & -62 & -38 & -19 & -43 & -59 & -56 \\ -66 & -69 & -60 & -15 & 16 & -24 & -62 & -55 \\ -65 & -70 & -57 & -6 & 26 & -22 & -58 & -59 \\ -61 & -67 & -60 & -24 & -2 & -40 & -60 & -58 \\ -49 & -63 & -68 & -58 & -51 & -65 & -70 & -53 \\ -43 & -57 & -64 & -69 & -73 & -67 & -63 & -45 \\ -41 & -49 & -59 & -60 & -63 & -52 & -50 & -34 \end{bmatrix}$$

$$DCT = \begin{bmatrix} -415 & -29 & -62 & 25 & 55 & -20 & -1 & 3 \\ 7 & -21 & -62 & 9 & 11 & -7 & -6 & 6 \\ -46 & 8 & 77 & -25 & -30 & 10 & 7 & -5 \\ -50 & 13 & 35 & -15 & -9 & 6 & 0 & 3 \\ -11 & -8 & -13 & -2 & -1 & 1 & -4 & 1 \\ -10 & 1 & 3 & -3 & -1 & 0 & 2 & -1 \\ -4 & -1 & 2 & -1 & 2 & -3 & 1 & -2 \\ -1 & -1 & -1 & -2 & -1 & -1 & 0 & -1 \end{bmatrix} \quad Q' = \begin{bmatrix} -26 & -3 & -6 & 2 & 2 & 0 & 0 & 0 \\ 1 & -2 & -4 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ \mathbf{1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

# JPEG Quantization Table (Luma)

$$Q_{u,v} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 51 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

# Example (Decoding)

$$P = \begin{bmatrix} -26 & -3 & -6 & 2 & 2 & 0 & 0 & 0 \\ 1 & -2 & -4 & 0 & 0 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad P' = \begin{bmatrix} -416 & -33 & -60 & 32 & 48 & 0 & 0 & 0 \\ 12 & -24 & -56 & 0 & 0 & 0 & 0 & 0 \\ -42 & 13 & 80 & -24 & -40 & 0 & 0 & 0 \\ -56 & 17 & 44 & -29 & 0 & 0 & 0 & 0 \\ 18 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$P'' = \begin{bmatrix} -70 & -64 & -61 & -64 & -69 & -66 & -58 & -50 \\ -72 & -73 & -61 & -39 & -30 & -40 & -54 & -59 \\ -68 & -78 & -58 & -9 & 13 & -12 & -48 & -64 \\ -59 & -77 & -57 & 0 & 22 & -13 & -51 & -60 \\ -52 & -71 & -72 & -54 & -54 & -71 & -71 & -54 \\ -42 & -50 & -70 & -68 & -67 & -67 & -61 & -50 \\ -45 & -59 & -70 & -68 & -67 & -67 & -61 & -50 \\ -35 & 47 & -61 & -66 & -60 & -48 & -44 & -44 \end{bmatrix} \quad P'' = \begin{bmatrix} 58 & 64 & 67 & 64 & 59 & 62 & 70 & 78 \\ 56 & 55 & 67 & 89 & 98 & 88 & 74 & 69 \\ 60 & 50 & 70 & 119 & 141 & 116 & 80 & 64 \\ 69 & 51 & 71 & 128 & 149 & 115 & 77 & 68 \\ 74 & 53 & 64 & 105 & 115 & 84 & 65 & 72 \\ 76 & 57 & 56 & 74 & 75 & 57 & 57 & 74 \\ 83 & 69 & 59 & 60 & 61 & 61 & 67 & 78 \\ 93 & 81 & 67 & 62 & 69 & 80 & 84 & 84 \end{bmatrix}$$

# Comparison

$$I = \begin{bmatrix} 52 & 55 & 61 & 66 & 70 & 61 & 64 & 73 \\ 63 & 59 & 66 & 90 & 109 & 85 & 69 & 72 \\ 62 & 59 & 68 & 113 & 144 & 104 & 66 & 73 \\ 63 & 58 & 71 & 122 & 154 & 106 & 70 & 69 \\ 67 & 61 & 68 & 104 & 126 & 88 & 68 & 70 \\ 79 & 65 & 60 & 70 & 77 & 68 & 58 & 75 \\ 85 & 71 & 64 & 59 & 55 & 61 & 65 & 83 \\ 87 & 79 & 69 & 68 & 65 & 76 & 78 & 94 \end{bmatrix} \quad P'' = \begin{bmatrix} 58 & 64 & 67 & 64 & 59 & 62 & 70 & 78 \\ 56 & 55 & 67 & 89 & 98 & 88 & 74 & 69 \\ 60 & 50 & 70 & 119 & 141 & 116 & 80 & 64 \\ 69 & 51 & 71 & 128 & 149 & 115 & 77 & 68 \\ 74 & 53 & 64 & 105 & 115 & 84 & 65 & 72 \\ 76 & 57 & 56 & 74 & 75 & 57 & 57 & 74 \\ 83 & 69 & 59 & 60 & 61 & 61 & 67 & 78 \\ 93 & 81 & 67 & 62 & 69 & 80 & 84 & 84 \end{bmatrix}$$

Original Image

Decoded Image

# Difference

$$Diff = \begin{bmatrix} -6 & -9 & -6 & 2 & 11 & -1 & -6 & -5 \\ 7 & 4 & -1 & 1 & 11 & -3 & -5 & 3 \\ 2 & 9 & -2 & -6 & -3 & -12 & -14 & 9 \\ -6 & 7 & 0 & -4 & -5 & -9 & -7 & 1 \\ -7 & 8 & 4 & -1 & 11 & 4 & 3 & 2 \\ 3 & 8 & 4 & -4 & 2 & 11 & 1 & 1 \\ 2 & 2 & 5 & -1 & -6 & 0 & -2 & 5 \\ -6 & -2 & 2 & 6 & -4 & -4 & -6 & 10 \end{bmatrix}$$

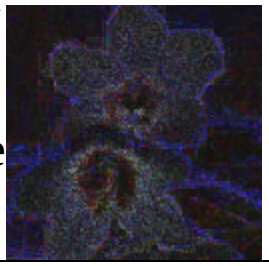
# JPEG

Original 64K

13K

5K

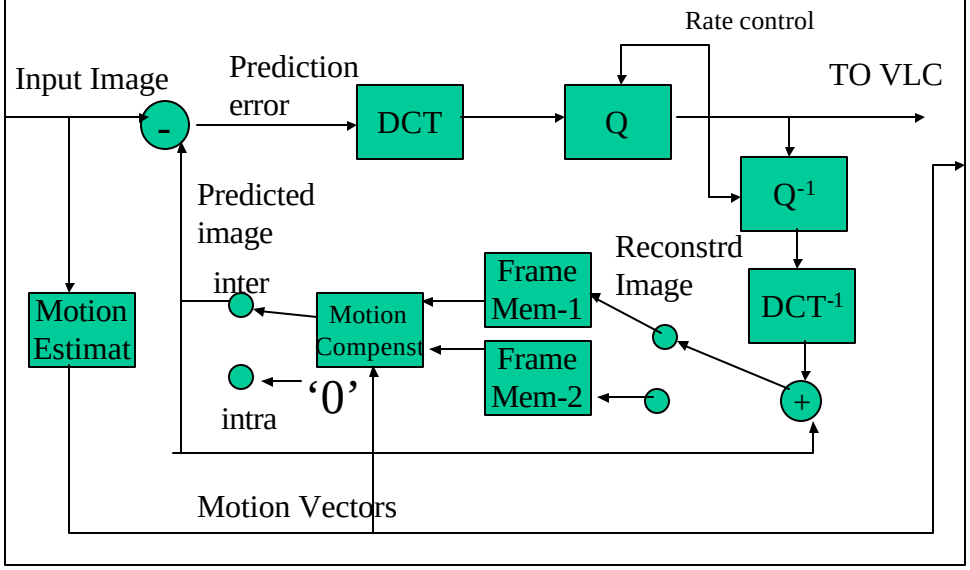
Difference



# Video Compression Standards

- H.261
- H.263
- MPEG-1
- MPEG-2
- MPEG-4
- MPEG-7 (Multimedia Content Description Interface)

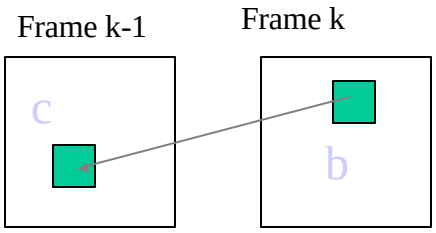
# MPEG-1 Encoder



## Motion Prediction

$b' = c'$

$Error = b - b'$




## Part IV

Video Phones and MPEG-4

## MPEG-1 & MPEG -2 Artifacts

- Blockiness
  - poor motion estimation
  - seen during dissolves and fades
- Mosquito Noises
  - edges of objects (high frequency DCT terms)
- Dirty Window
  - streaks or noise remain stationary while objects move

## MPEG-1 & MPEG -2 Artifacts

- Wavy Noise
  - seen during pans across crowds
  - coarsely quantized high frequency terms cause errors

## Where MPEG-2 will fail?

- Motions which are not translation
  - zooms
  - rotations
  - non-rigid (smoke)
  - dissolves
- Others
  - shadows
  - scene cuts
  - changes in brightness

## Video Compression At Low Bitrate

- The quality of block-based coding video (MPEG-1 & MPEG-2) at low bitrate, e.g., 10 kbps is very poor.
  - Decompressed images suffer from blockiness artifacts
  - Block matching does not account for rotation, scaling and shear