Introduction to Multimedia Computing

Overview of Video Coding using an Example
Introduction to Adaptive Video Coding

Topics

- Video Frames
- Color Space Conversion
- Down-sampling
- Intra-frame Coding
  - DCT
  - Quantization
  - Code-words
- Inter-frame Coding
  - Motion Estimation
  - Motion Compensation
- Adaptive Video Coding

Video

- Video is a sequence of still images taken in short time intervals.
- The difference between consecutive video frames is small
- The similarity between the frames is used for better compression

Sample Frames (Frame 1000)

Sample Frames (Frame 1003)

Coding Intra-Frames

- Each frame is divided into macro-blocks of size 16x16
- Each Macro-block contains 16x16 pixels with RGB color components
- The macro-block is converted into YCbCr, down-sampled, DCT transformed, Quantized, zig-zag scanned, and entropy coded.

A Sample Macro-Block (Red)

13	12	13	15	18	17	19	20	21	20	22	21	22	22	23	22
13	14	14	15	18	17	19	19	22	21	22	21	22	22	22	22
13	14	14	15	18	17	19	19	22	21	22	20	21	21	22	22
13	12	14	15	18	17	19	20	20	20	24	21	23	21	22	21
14	13	13	14	16	16	18	19	20	22	24	20	20	20	20	20
14	13	13	14	16	16	18	19	20	22	24	20	20	21	21	20
14	12	12	14	16	16	18	19	20	22	24	20	20	21	21	20
14	12	12	14	16	16	18	19	20	22	24	20	20	20	20	19
14	12	12	14	16	16	18	19	20	22	24	20	20	20	20	19
13	12	13	14	16	16	18	19	20	22	24	20	20	20	20	19
14	13	13	14	16	16	18	19	21	22	24	21	20	20	20	19
14	13	13	14	16	16	18	19	21	22	24	21	20	20	20	19
14	13	13	14	16	16	18	19	21	22	24	20	20	20	20	19
14	14	14	15	18	16	18	19	21	22	23	22	22	20	20	19
14	14	14	15	18	16	18	19	21	22	23	22	22	21	21	20
14	14	14	15	18	16	18	19	21	22	23	22	22	21	21	20

A Sample Macro-Block (Green)

32	31	32	34	37	39	41	43	45	48	50	53	54	54	55	54
32	33	33	34	37	39	41	42	46	49	50	53	54	54	54	54
32	33	33	34	37	39	41	42	46	49	50	52	53	53	54	54
32	31	33	34	37	39	41	43	43	46	50	51	53	53	54	53
33	32	32	33	36	37	40	42	43	47	50	50	50	52	52	52
33	32	32	33	36	37	40	42	43	47	50	50	50	51	51	52
33	31	31	33	36	37	40	42	43	47	50	50	50	51	51	52
33	31	31	33	36	37	40	42	43	47	50	50	50	50	50	51
33	31	31	33	36	37	40	42	43	47	50	50	50	50	50	51
32	31	32	33	36	37	40	42	43	47	50	50	50	50	50	51
33	32	32	33	36	37	40	42	45	47	50	52	50	50	50	51
33	32	32	33	36	37	40	42	45	47	50	52	50	50	50	51
33	32	32	33	36	37	40	42	45	47	50	50	50	50	50	51
33	33	33	34	37	37	40	42	45	47	48	50	50	50	50	51
33	33	33	34	37	37	40	42	45	47	48	50	50	52	52	52
33	33	33	34	37	37	40	42	45	47	48	50	50	52	52	52

A Sample Macro-Block (Blue)

46	45	46	48	51	50	52	59	60	66	69	73	74	76	78	76
46	47	47	48	51	52	54	57	61	67	69	73	74	76	76	76
46	47	47	48	51	52	54	57	61	67	69	72	73	75	76	76
46	45	47	48	51	52	54	59	59	63	67	72	73	73	74	75
47	46	46	47	50	51	53	57	59	64	67	71	71	72	72	74
47	46	46	47	50	51	53	57	59	64	67	69	69	72	72	72
47	45	45	47	50	51	53	57	59	64	67	69	69	72	72	72
47	45	45	47	50	51	53	57	59	64	67	69	69	69	69	69
47	45	45	47	50	51	53	57	59	64	67	69	69	69	69	69
46	45	46	47	50	51	53	57	59	64	67	69	69	69	69	69
47	46	46	47	50	51	53	57	60	64	67	70	69	69	69	69
47	46	46	47	50	51	53	57	60	64	67	70	69	69	69	69
47	46	46	47	50	51	53	57	60	64	67	69	69	69	69	69
47	47	47	48	51	51	53	57	60	64	65	69	69	69	69	69
47	47	47	48	51	51	53	57	60	64	65	69	69	70	70	70
47	47	47	48	51	51	53	57	60	64	65	69	69	70	70	70

Conversion from RGB to YCbCr

- RGB can be converted to YCbCr linearly by (use rgb2ycbcr in MATLAB)

$$\begin{bmatrix} Y \\ \mathrm{Cb} \\ \mathrm{Cr} \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.14713 & -0.28886 & 0.436 \\ 0.615 & -0.51499 & -0.10001 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

RGB to YCbCr (Y)

40	39	40	42	44	45	47	49	50	52	54	55	56	56	57	56
40	41	41	42	44	45	47	48	51	53	54	55	56	56	56	56
40	41	41	42	44	45	47	48	51	53	54	54	55	55	56	56
40	39	41	42	44	45	47	49	49	50	54	54	56	55	56	55
41	40	40	41	43	44	46	48	49	52	54	53	53	54	54	55
41	40	40	41	43	44	46	48	49	52	54	53	53	54	54	54
41	39	39	41	43	44	46	48	49	52	54	53	53	54	54	54
41	39	39	41	43	44	46	48	49	52	54	53	53	53	53	53
41	39	39	41	43	44	46	48	49	52	54	53	53	53	53	53
40	39	40	41	43	44	46	48	49	52	54	53	53	53	53	53
41	40	40	41	43	44	46	48	50	52	54	54	53	53	53	53
41	40	40	41	43	44	46	48	50	52	54	54	53	53	53	53
41	40	40	41	43	44	46	48	50	52	54	53	53	53	53	53
41	41	41	42	44	44	46	48	50	52	52	54	54	53	53	53
41	41	41	42	44	44	46	48	50	52	52	54	54	54	54	54
41	41	41	42	44	44	46	48	50	52	52	54	54	54	54	54

RGB to YCbCr (Cb)

137	137	137	137	137	136	136	138	138	140	140	142	142	142	143	142
137	137	137	137	137	137	137	138	138	140	140	142	142	142	142	142
137	137	137	137	137	137	137	138	138	140	140	142	142	142	142	142
137	137	137	137	137	137	137	138	138	139	139	142	141	142	142	142
137	137	137	137	137	137	137	138	138	139	139	142	142	142	142	142
137	137	137	137	137	137	137	138	138	139	139	141	141	142	142	142
137	137	137	137	137	137	137	138	138	139	139	141	141	142	142	142
137	137	137	137	137	137	137	138	138	139	139	141	141	141	141	141
137	137	137	137	137	137	137	138	138	139	139	141	141	141	141	141
137	137	137	137	137	137	137	138	138	139	139	141	141	141	141	141
137	137	137	137	137	137	137	138	138	139	139	141	141	141	141	141
137	137	137	137	137	137	137	138	138	139	139	141	141	141	141	141
137	137	137	137	137	137	137	138	138	139	139	141	141	141	141	141
137	137	137	137	137	137	137	138	138	139	139	140	140	141	141	141
137	137	137	137	137	137	137	138	138	139	139	140	140	141	141	141
137	137	137	137	137	137	137	138	138	139	139	140	140	141	141	141

RGB to YCbCr (Cr)

119	119	119	119	119	118	118	117	116	114	114	113	113	112	112	112
119	119	119	119	119	117	117	117	116	114	114	113	113	112	112	112
119	119	119	119	119	117	117	117	116	114	114	113	113	112	112	112
119	119	119	119	119	117	117	117	117	115	115	113	113	113	113	112
119	119	119	119	118	118	117	117	117	116	115	113	113	113	113	112
119	119	119	119	118	118	117	117	117	116	115	113	113	113	113	113
119	119	119	119	118	118	117	117	117	116	115	113	113	113	113	113
119	119	119	119	118	118	117	117	117	116	115	113	113	113	113	113
119	119	119	119	118	118	117	117	117	116	115	113	113	113	113	113
119	119	119	119	118	118	117	117	117	116	115	113	113	113	113	113
119	119	119	119	118	118	117	117	116	116	115	113	113	113	113	113
119	119	119	119	118	118	117	117	116	116	115	113	113	113	113	113
119	119	119	119	118	118	117	117	116	116	115	113	113	113	113	113
119	119	119	119	119	118	117	117	116	116	116	114	114	113	113	113
119	119	119	119	119	118	117	117	116	116	116	114	114	113	113	113
119	119	119	119	119	118	117	117	116	116	116	114	114	113	113	113

Down-Sampled Cr

119	119	119	118	116	114	113	112
119	119	119	117	116	114	113	112
119	119	118	117	117	115	113	113
119	119	118	117	117	115	113	113
119	119	118	117	117	115	113	113
119	119	118	117	116	115	113	113
119	119	118	117	116	115	113	113
119	119	119	117	116	116	114	113

Intra-Frame Coding

- Down-sample Cb and Cr
- Apply DCT
- Apply Quantization
- Zig-Zag Scan
- Create Code-Words
- Entropy Encode the Code-Words

Discrete cosine transform

- Cosine transform is applied to each block after down-sampling.

$$G_{u,v} = \alpha(u)\alpha(v) \sum_{x=0}^{7} \sum_{y=0}^{7} g_{x,y} \cos\left[\frac{\pi}{8}\left(x + \frac{1}{2}\right)u\right] \cos\left[\frac{\pi}{8}\left(y + \frac{1}{2}\right)v\right]$$

where

$$\alpha_p(n) = \begin{cases} \sqrt{\frac{1}{8}}, & \text{if } n = 0 \\ \sqrt{\frac{2}{8}}, & \text{otherwise} \end{cases}$$

DCT of the Cr

930.625	18.5341	-2.8953	0.2844	-0.125	-1.3907	0.5228	0.1837
-0.5689	1.1724	-0.6796	-0.6189	0.4714	0.2764	-0.5469	0.6013
0.028	0.6451	-0.0884	-0.7622	-0.9519	1.1911	0.2134	-0.0382
-0.5798	0.594	-0.0329	-0.4721	0.3021	0.3233	-0.2386	-0.2642
0.625	-0.147	-0.3943	0.0345	-0.125	0.1734	0.028	0.0982
-0.0868	-0.0893	-0.022	0.043	0.5025	-0.1618	-0.1595	-0.3631
0.3943	-0.3516	-0.0366	0.1404	-0.0116	0.0125	0.0884	-0.2532
0.1417	-0.179	-0.1352	0.1807	0.3487	-0.2653	-0.1088	-0.0385

Rounded DCT Coefficients

931	19	-3	0	0	-1	1	0
-1	1	-1	-1	0	0	-1	1
0	1	0	-1	-1	1	0	0
-1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Quantization Matrix

$$Q \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Quantized Coefficients

58	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	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

ZigZag Scan and Code-Words

- 58, 1, EOB
- Code-words: (0,58), (0,1)

Inter-Frame Coding

- In inter-frame coding, before DCT transform we should remove temporal redundancy
- To remove temporal redundancy we perform
  - Motion Estimation (ME)
  - Motion Compensation (MC)

Motion Estimation

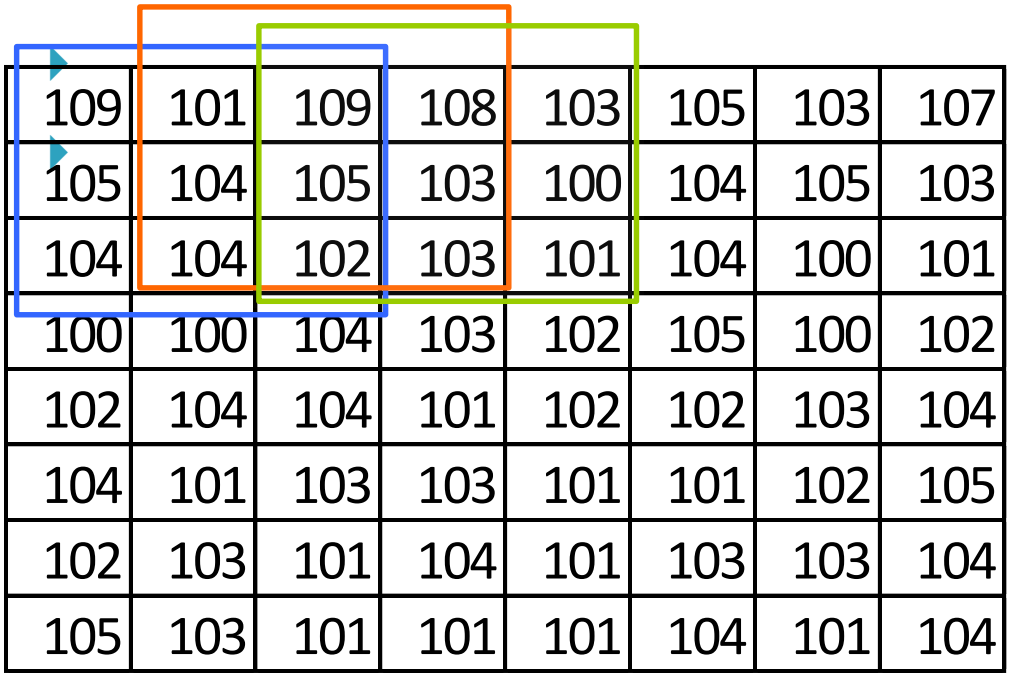
- The goal of motion estimation is finding the most similar area in the reference frame (previous frame) to a block in the current frame.

- Similarity is generally defined using Sum of Absolute Difference (SAD)

Motion Estimation Example

- In this example blocks are considered as 3x3 for simplicity.
- The SAD is computed for each position in the reference frame
- The most similar area is shown using a motion vector (MV)

Motion Estimation Example



109	101	109	108	103	105	103	107
105	104	105	103	100	104	105	103
104	104	102	103	101	104	100	101
100	100	104	103	102	105	100	102
102	104	104	101	102	102	103	104
104	101	103	103	101	101	102	105
102	103	101	104	101	103	103	104
105	103	101	101	101	104	101	104

26 28 25

MV = 2,0

104	101	101
101	102	103
101	102	102

- Reference Image

Current Block

Motion Compensation

- Motion Estimation gives us a motion vector
- Using the motion vector the most similar area is found in the reference frame.
- The most similar area is subtracted from the current block
- The difference between them (error) is:
  - DCT transformed
  - Quantized
  - Zig-zag scanned
  - Entropy coded

Adaptive Video Coding

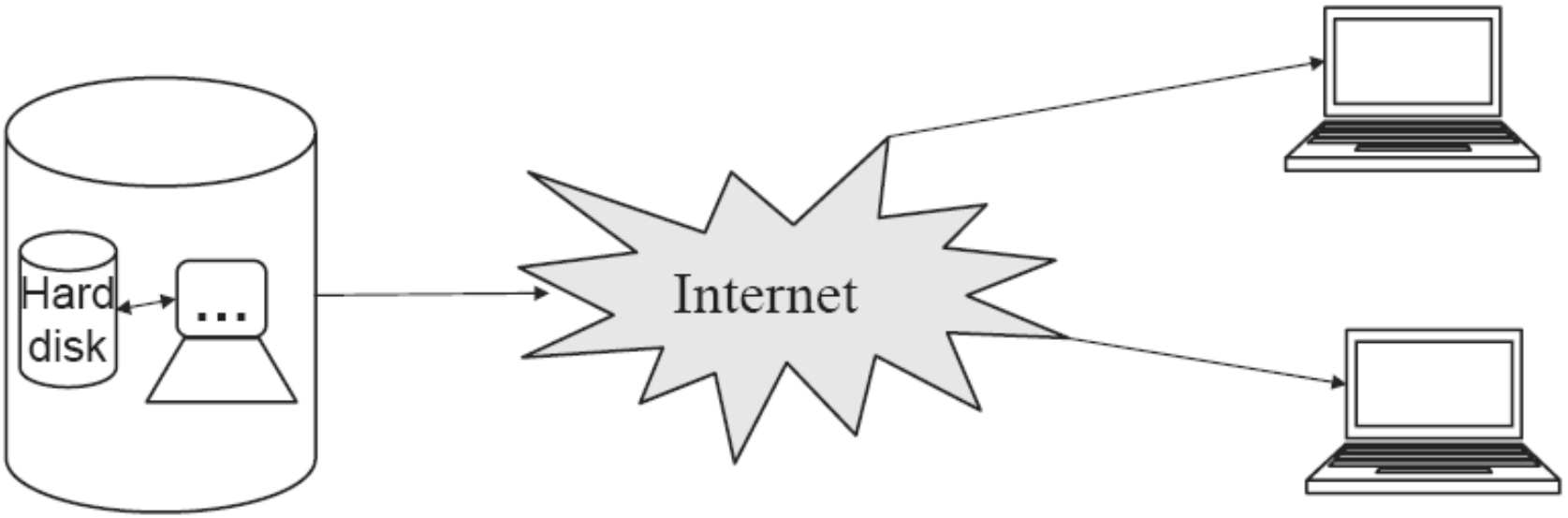
Communicating Multimedia

- Multimedia data is used from a distance using computer networks.
- The Internet is a public network that can be used for multimedia transmission.

Multimedia over the Internet

- Multimedia data can be used as:
  - Streaming data
  - Data (audio, video, etc.) is presented to the end-user only once (without being saved to a file).
- Downloadable
  - Data is downloaded to a file and presented to the end-user repeatedly.

Video on Demand over the Internet



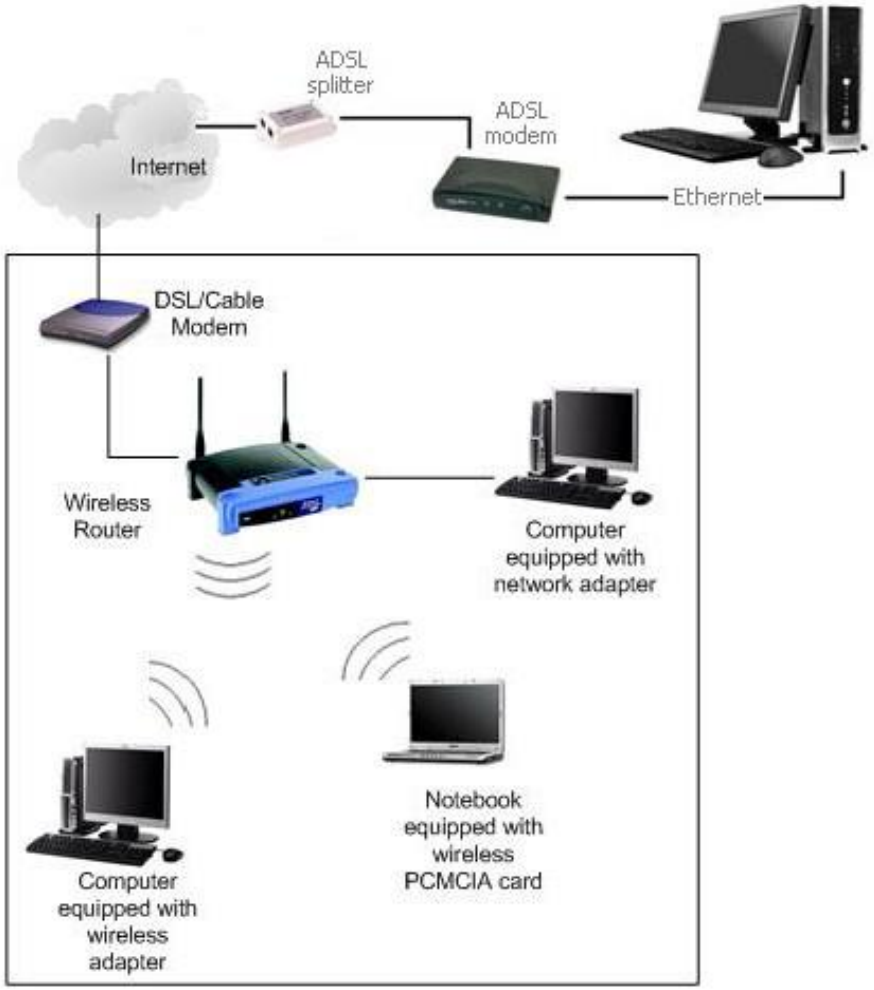
Video On Demand (VOD) streaming server

Computer Networks and the Internet

- The Internet has a heterogeneous structure.
- The Internet is a best-effort network.
- The Internet does not guarantee a fixed data rate over a connection.
- **Multimedia data should adapt itself with network data rate changes.**

Need for Adaptive Videos

- Networks have different bandwidths and data rates



Need for Adaptive Videos

- Display devices have different properties

Adapting Video (1): Transcoding

- Transcoding is defined as changing a video in
  - Resolution (Spatial Transcoding)
  - Frames per second (Temporal Transcoding)
  - Bits per pixel (SNR Transcoding)
  - Inserting additional data into the video (Content Transcoding)
  - Algorithm (Standard Transcoding)

Real-time Transcoding

- The gateways in the network should perform transcoding
  - Transcoding is slow because
  - Video should be decoded (include IDCT)
  - Video should be re-encoded (includes DCT and Motion Estimation)

Adapting Video (2): Scalable Video Coding

- In Scalable Video Coding, the receiver adapts the video to its capabilities.
- Video is coded in a way that the receiver can receive part of it.
- Adapting video should be fast.

Questions?