Introduction to Multimedia Computing

Introduction to Adaptive Video Coding
Scalable Video Coding

Topics

- Communicating Video Media
- Need for Adaptive Video Coding
- Transcoding (recap)
- Scalable Video Coding
  - Spatial Scalability
  - Temporal Scalability
  - SNR Scalability
- Multiple Scalability
- Drift Problem

Communicating Multimedia (recap)

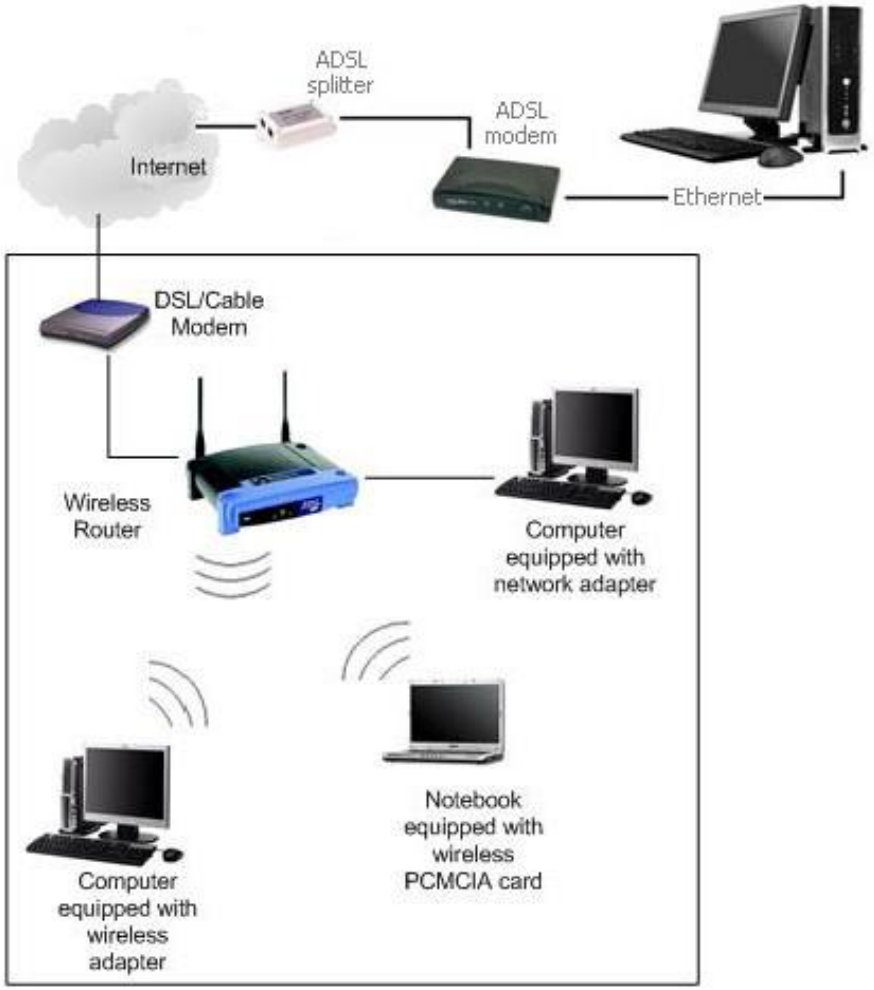
- Multimedia data is used from a distance using computer networks.
- The Internet is a public network that can be used for multimedia transmission.

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.

Scalable Video Coding

- In scalable video coding the video is divided into multiple layers
- First layer is called the **Base Layer**
- Base Layer defines the video in the lowest quality
- Remaining layers add to the quality of the video and are called **Enhancement Layers.**

Types of Video Scalability

- Scalability can be defined in term of:
  - Lower resolution (Spatial Scalability)
  - Lower frame rate (Temporal Scalability)
  - Lower color depth (Signal to Noise (SNR) Scalability)

Lower Resolution Scalability

- The base layer video is a lower resolution video.
- The enhancement layer increases the resolution.
- If the network bandwidth is not sufficient, or the display resolution of the receiver is low, only the base layer is received.

Low Frame Rate Scalability

- The base layer has a lower frame rate (for instance: 15 frames/sec)
- The enhancement layer adds to the frame rate (for instance: increasing the frame rate to 30 frames/sec)
- The frame rate scalability changes the number of frames in unit of time (hence temporal scalability)

Color Depth Scalability

- Each frame pixel is represented by a few bits.
- The typical number of bits/pixel is 24 bits (Red=8 bits, Green=8 bits, Blue=8 bits)
- Lower number of bits means fewer number of colors in the frame.

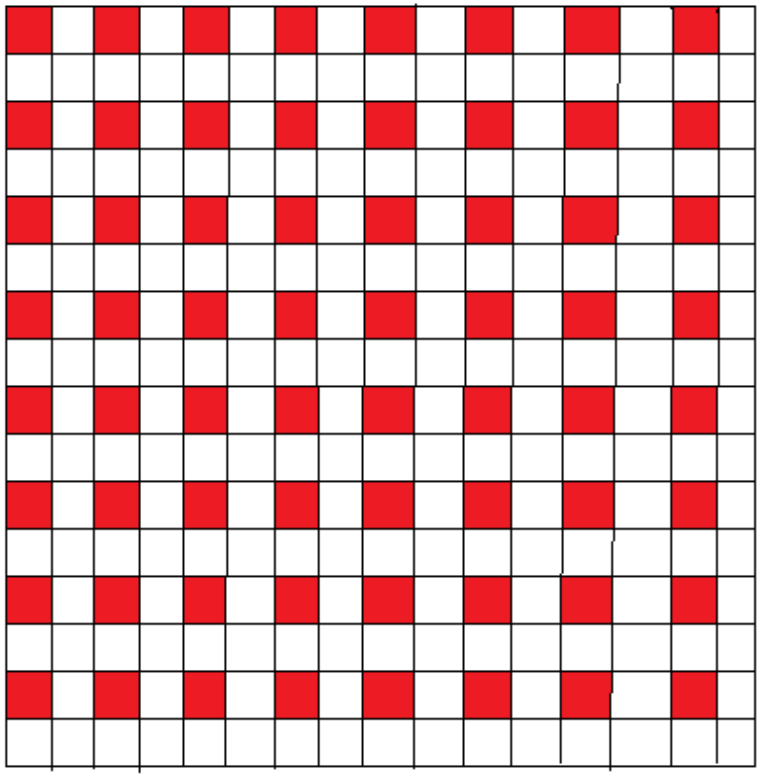
Color Depth Scalability

- The base layer uses fewer number of bits/pixel.
- The enhancement layer adds to the bits/pixel rate, and improves the visual quality of the frames.
- The pixel accuracy in representing colors is referred to as Signal to Noise Ratio.

Spatial Scalability

- Some pixels from each frame are put in the base layer and the remaining in the enhancement layer(s)
- e.g. The low resolution is the base layer, and the high resolution is base +enhancement layer frames

Resolution Scalability



- Red Pixels are Base Layer Pixels

Spatial Scalability

99	100	86	82
111	102	70	78
36	45	150	152
23	44	154	160

Original Frame

103	79
37	154

Base Layer

-4	-3	7	3
8	-1	-9	-1
-1	8	-4	-2
-14	7	0	6

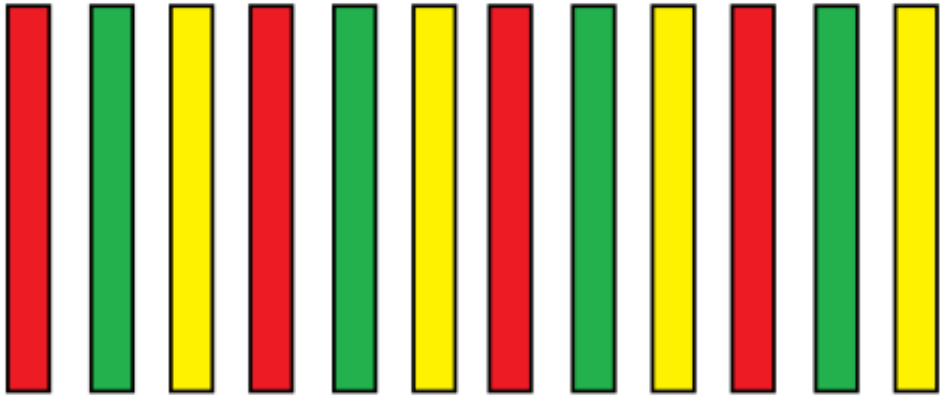
Enhancement Layer

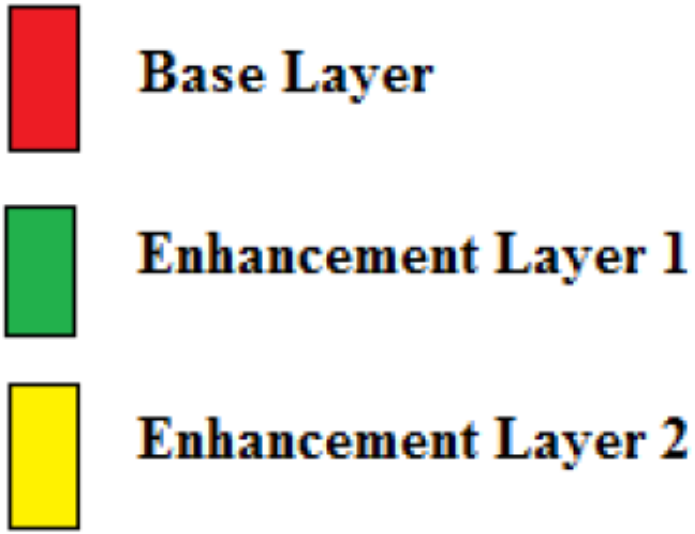
Spatial Scalability at the Receiver

- The receiver follows the steps below to represent and render the video:
  - If only the base layer is used, display it
  - If the base layer and the enhancement layer are used, up-sample the base layer, and add enhancement layer values to get the original frame

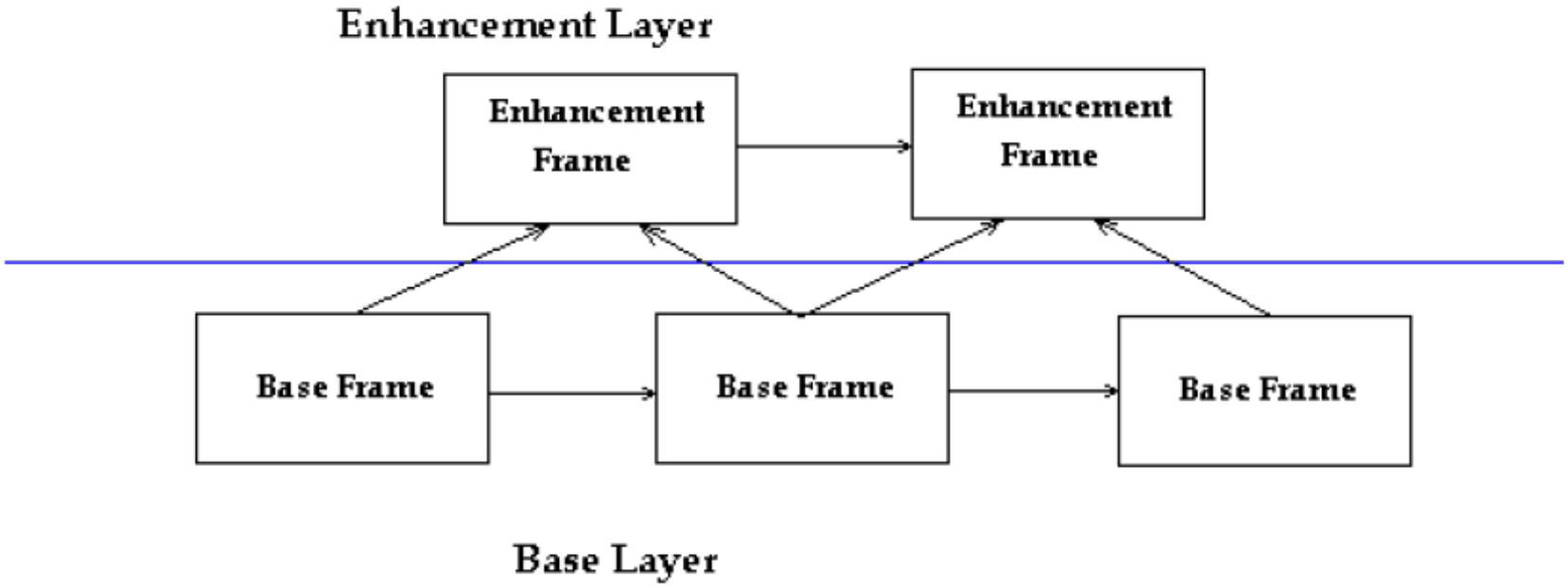
Temporal Scalability

- In temporal scalability, some frames are put in base layer and some in enhancement layers
- The example below is a video with one base layer and two enhancement layers



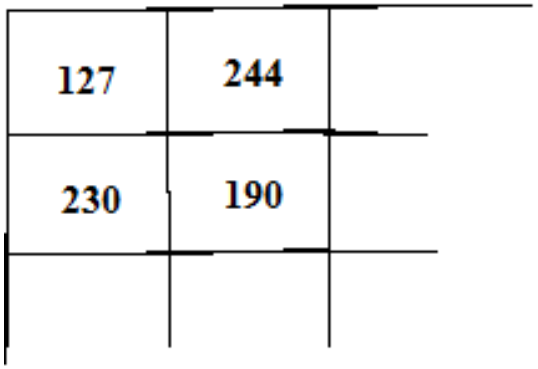

Base Layer
Enhancement Layer 1
Enhancement Layer 2

Temporal Scalability

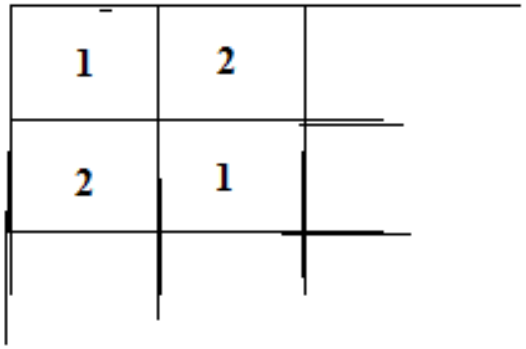


SNR Scalability

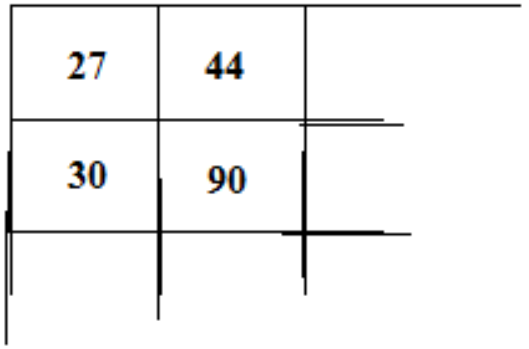
- In Signal To Noise Scalability, significant bits are put in the base layer and the remaining bits in the enhancement layers.
- Receiver concatenates the bits to create high quality video



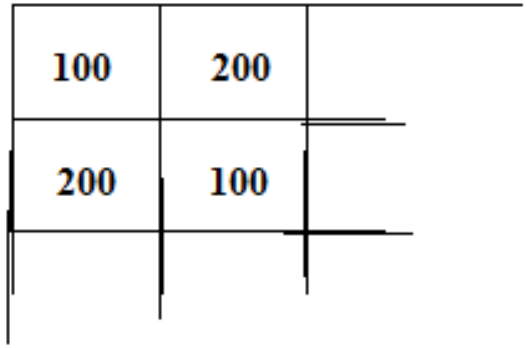
Original Frame



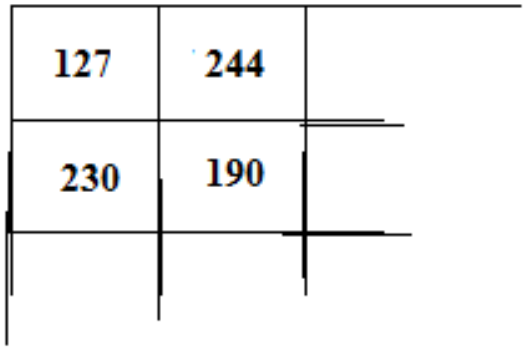
Base Layer



Enhancement Layer



Reconstructed using Base Layer Only



Reconstructed using both Layers

SNR Scalability

99	101	86	82
111	102	70	78
36	45	150	152
23	44	154	160

Original Frame

9	10	8	8
11	10	7	7
3	4	15	15
2	4	15	16

Base Frame

9	1	6	2
1	2	0	8
6	5	0	2
3	4	4	0

Enhancement Layer

Multilayer Scalability

- Videos can be modified to include multiple types of scalability
- For instance, resolution and frame rate scalability can be applied to a video

Multiple Scalability

Drift Problem

- In video coding each frame is obtained from the previous frame.
- Any error in reconstructing a frame causes error in the next frame.
- Accumulated error reduces the quality of the video. (Drift Problem)

Multiple Scalability Problem

- Receiver can receive data partially.
- Partial data is used for reconstructing the next frame.
- The reconstructed frame is slightly different than the original frame
- The reconstructed frame is used as a reference frame for the next frame
- Drift problem happens

Summary

- On Demand Video requires adaptation with network properties
- Transcoding is used for video adaptation but requires a long processing time
- Scalable video coding encodes video in a way that it can adapt without decoding/encoding

Questions?