Image and video processing

Digital video

Dr. Qianni Zhang (qianni.zhang@qmul.ac.uk)

EBU723U

Agenda

- Digital video
- Video compression
- Video formats and codecs
- MPEG
- Other codecs
- Web video

EBU723U

- 2 -

Digital Video

- Until the arrival of the Pentium processor, in 1993, even the most powerful PCs were limited to capturing images no more than 160 x 120 pixels

Slide 3

- 3 -

Digital Video

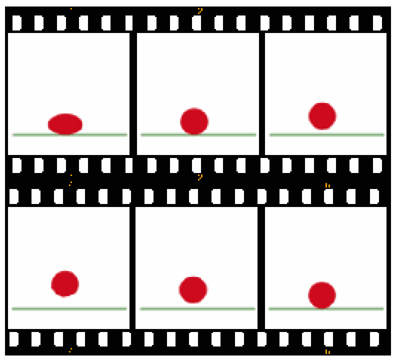
- When processor speeds finally exceeded 200MHz, PCs could handle images up to 320 x 240 without the need for expensive compression hardware

Slide 4

- 4 -

Digital Video Problem

- Digital video is a series of still images



Slide 5

- 5 -

Digital Video Problem

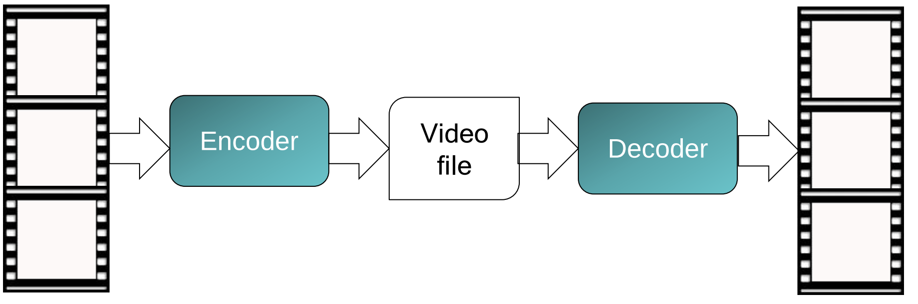
- Images take up storage space
- Lots of images take up lots of storage space
- Raw PAL (Standard Definition) video is 25 images a second at 768 x 576 resolution
- Approx. 33 Mbit/s or 4Mbytes/s

Slide 6

- 6 -

Video Compression

- Video compression relies on the encoding and decoding of the video data
- The algorithms that handle compression and decompression of the video are known as codecs



Slide 7

- 7 -

Redundancy

- Types of redundancy
  - Coding redundancy
    - some grey levels / colours are more common than others
  - Inter-pixel redundancy
    - the same grey level covers large areas
  - Psycho visual redundancy
    - the eye can only resolve 32 grey levels locally

EBU723U

- 8 -

Digital Video

- There are numerous video encoding formats
  - QuickTime
  - WMV
  - M-JPEG, DIVX
  - FLV
  - MPEG
  - AVI
  - H.264
  - DV

Slide 9

- 9 -

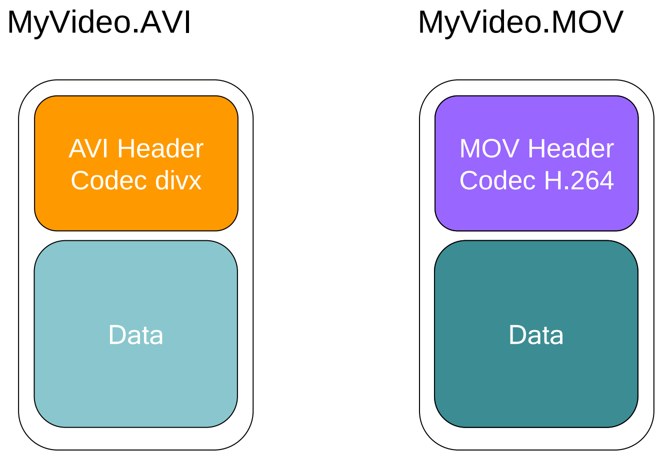
Formats and Codecs

- Formats are required for storage and transmission of video data
- Often a format will specify a single encoding method
- Sometimes the format will specify several different encoding methods
- Every architecture has certain codecs built-in and some codecs are common to many architectures
  - E.g. QuickTime originally used Sorenson video codec, but now supports MPEG-4, Cinepak etc.

Slide 10

- 10 -

Formats and Codecs



Slide 11

- 11 -

M-JPEG

- Motion-JPEG not surprisingly based on the JPEG still image format
- Stores every frame as a compressed bitmap image
- Typical compression ratio of between 2:1 and 12:1
- Can be applied in hardware or as a software codec

Slide 12

- 12 -

MPEG

- Moving Picture Expert Group standards for compressing motion video and audio signals using DCT
- The MPEG formats are asymmetrical – it takes longer to compress a frame of video than it does to decompress it

Slide 13

- 13 -

MPEG Compression

- MPEG uses both spatial DCT compression (as JPEG) and temporal compression
  - Temporal compression involves removing data that does not change between consecutive frames
- MPEG video streams consist of a sequence of sets of frames known as a GOP (group of pictures)

Slide 14

- 14 -

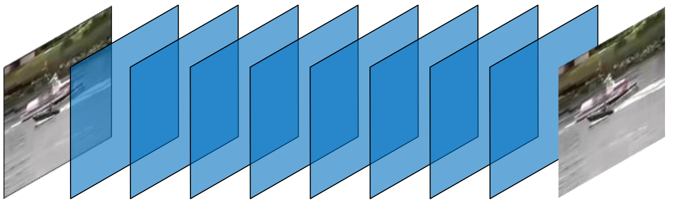
MPEG Compression

- Each GOP (typically 8-24 frames), has only one complete frame represented in full, which is compressed using only spatial compression
  - The spatially compressed frame is just like a JPEG and is known as an I frame (intra coded frame)
  - Around the I frame are temporally-compressed frames called
    - P frames (inter-coded frames) and
    - B frames (bi-directional inter-coded frames)
    - representing only change data

Slide 15

- 15 -

Group of Pictures



Slide 16

- 16 -

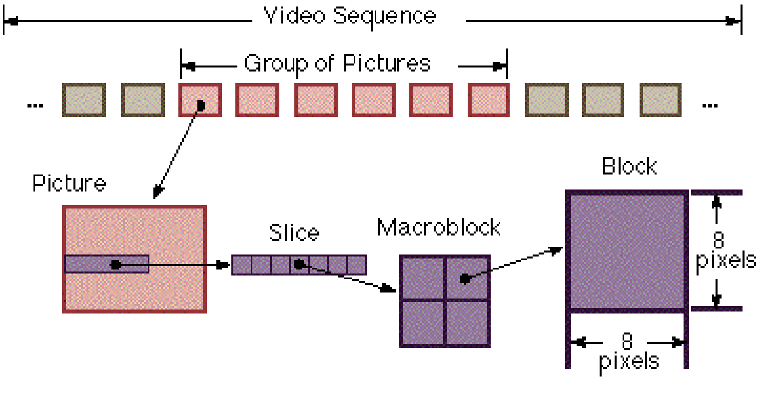
Predictive coding

- Key idea in video compression
  - Predict a new frame from a previous frame and only code the prediction error (using the DCT)
  - Prediction errors have *smaller energy* than the original pixel values and can be coded with fewer bits
  - Regions that cannot be predicted well will be coded directly using DCT
  - Work on each macroblock (MB) independently for reduced complexity
    - Motion compensation done at the MB level (16x16 pixels)
    - DCT coding of error at the block level (8x8 pixels)

EBU723U

- 17 -

Motion compensation



EBU723U

- 18 -

Temporal Compression

- During encoding, prediction techniques compare neighbouring frames and pinpoint areas which move and those that do not change

Slide 19

- 19 -

Temporal Compression

- Frames are split into 8 x 8 macroblocks for comparison

Slide 20

- 20 -

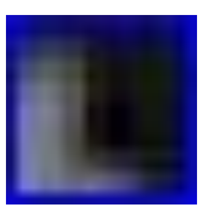
Temporal Compression

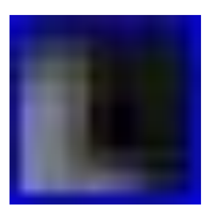
- Frames are split into 8 x 8 macroblocks for comparison

Slide 21

Temporal Compression

- When comparing parts of a frame with a previous I frame there are three things which could happen





1. The part required is identical to the same area in the I-Frame and so no data needs to be stored

Slide 22

Temporal Compression

- When comparing parts of a frame with a previous I frame there are three things which could happen

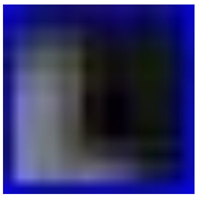
2. The part is different to any part of the I-Frame so the new part is encoded as an I-macroblock

Slide 23

- 23 -

Temporal Compression

- When comparing parts of a frame with a previous I frame there are three things which could happen



3. The part is similar to part of the I-Frame so difference values and a motion vector are stored

Slide 24

- 24 -

P Frames

- Each macroblock in a P-frame can be encoded either as an I-macroblock or as a P-macroblock
  - An I-macroblock is encoded in the same way as a macroblock in an I-frame (i.e. similar to JPEG)
  - A P-macroblock is encoded as an area of the past reference frame, plus an error term

Slide 25

- 25 -

P Frames

- To specify where the macroblock came from in the reference frame, a motion vector is used
  - If the motion vector is zero, then the macroblock has not moved between frames
  - If the difference values are also zero, then the area does not need to be coded
- If the macroblock is not exactly the same as the area in the reference frame, then differenece values are also stored

Slide 26

- 26 -

Motion Vector

80	80	60	20
100	100	100	100
100	80	60	100
20	40	80	100

Reference Picture

100	100	80	20
80	60	20	40
100	100	100	100
80	60	100	80

Current P-Picture

Slide 27

- 27 -

Difference Values

100	100	100
100	80	60
20	40	80

100	100	100
80	80	60
20	20	80

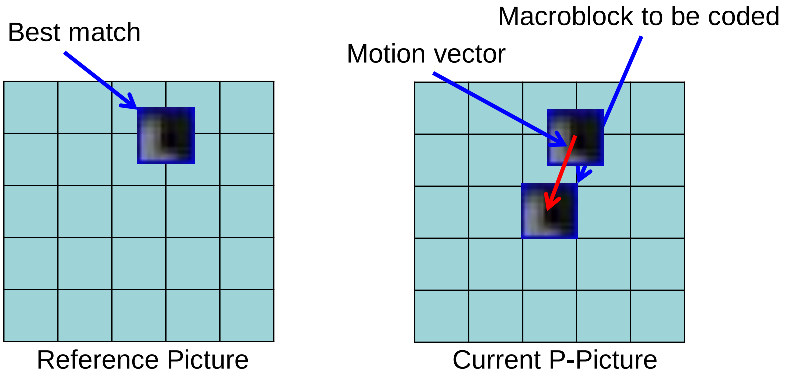
0	0	0
20	0	0
0	20	0

Slide 28

- 28 -

P Frames

- By recording only vectors and change data, the information which needs to be recorded can be substantially reduced



Slide 29

- 29 -

Temporal prediction

- No motion compensation
  - Works well in stationary regions
- Uni-directional motion compensation
  - Does not work well for uncovered regions by object motion
- Bi-directional motion compensation
  - Can handle better covered/uncovered regions

EBU723U

Prediction with motion compensation

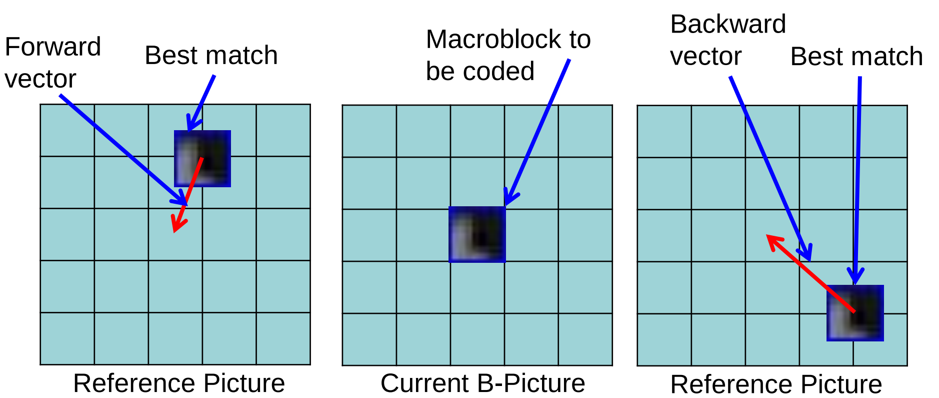
- Backward
  - To predict where the pixels are in a previous image (previous time instant)
- Forward
  - To predict where the pixels will be in a next image (future time instant)

- In MPEG compression, there are two types of vector frames:
  - P frames (predictive), refer only to the previous I or P frame
  - B frames (bi-directional) rely on previous and subsequent I or P frames

EBU723U

- 31 -

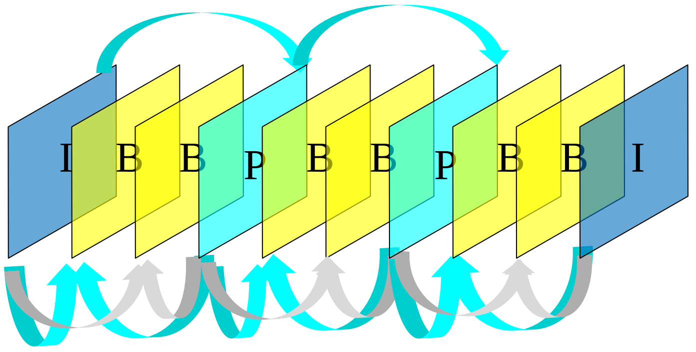
B Frames



Slide 32

- 32 -

Group of Pictures



Slide 33

- 33 -

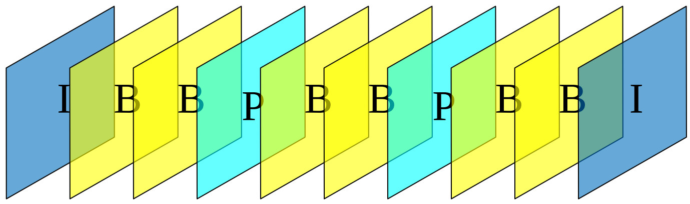
Group of Pictures

- I frames do not rely on other frames
- P frames rely on previous I and P frames
- B frames rely on previous and future I and P frames
- No frames rely on B frames

Slide 34

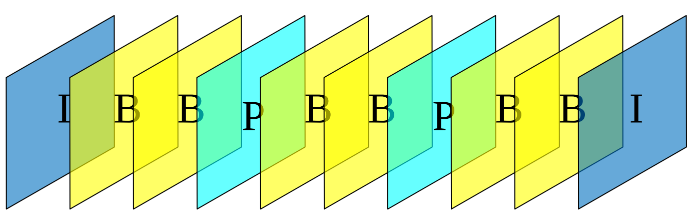
- 34 -

Drop an I Frame



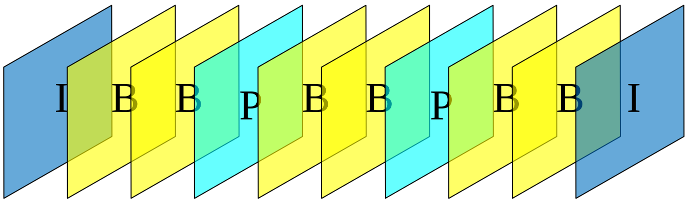
Slide 35

Drop a P Frame



Slide 36

Drop a B Frame



Slide 37

- 37 -

Group of Pictures

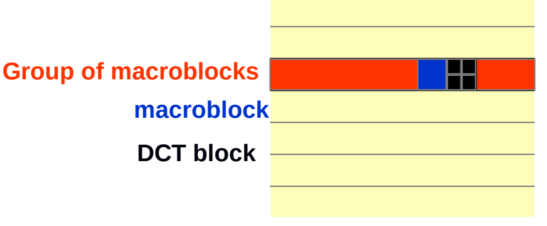
- P frames take significantly less storage space than I-frames (around 20:1)
- B frames are even smaller (around 50:1)
- Using longer GOPs with more B and P frames, reduce data rates

Slide 38

- 38 -

Bitstream syntax

- A video sequence is composed of Images of different types ...IPBBPBBPBBP…I...
  - Each image is composed of:
    - Groups of Macroblocks (synchronisation)
      - Macroblocks (motion+texture)
        - Blocks (DCT)



EBU723U

- 39 -

MPEG-1

- MPEG-1 (White Book Standard) was introduced in 1993
  - Supports video coding at bit-rates up to about 1.5 Mbit/s and virtually transparent stereo audio quality at 192 Kbit/s
  - Provide video resolution of 352 x 288 at 25 fps
- This produces video quality slightly below the quality of conventional VCR videos

EBU723U

- 40 -

MPEG-2

- Resolutions of 720 x 576 and 1280 x 720 at 50 fps, with full CD quality audio
- This is sufficient for all the major TV standards, including PAL, and even HDTV
- MPEG-2 is used by DVD-ROMs, digital satellite and Cable
  - Can compress a 2 hour video into a few gigabytes
  - Decompressing an MPEG-2 data stream requires only modest computing power

EBU723U

- 41 -

MPEG-4

- MPEG-4 is a standardised way to define, encode and playback time based media
- It can be used in numerous applications:
  - Video and streaming video
  - Delivering 2D still images
  - Controlling animated 3D models
  - Handling two-way video conferences
  - Coding of audio-visual objects
  - Objects can be coded and decoded *independently*
  - Object can be flexibly *composed* to create different scenes
  - *Natural and synthetic* object are treated in the same way
  - Excellent error resilience
- Standardised in October 1998

EBU723U

- 42 -

MPEG-4

- Provides standardised ways of representing units of aural, visual or audio-visual content, as discrete “media objects”
- These can be of natural of synthetic origin
- For example, they could be recorded with a camera or microphone, or generated with a computer

EBU723U

- 43 -

MPEG-4

- MPEG-4 consists of several standards (parts)
  - Part 2 – Advanced Simple Profile (ASP), used by codecs such as DivX, Xvid and Nero Digital
  - Part 10 – Advanced Video Coding (AVC), used by the x264 codec, Quicktime 7, BBC iPlayer and Blu-ray Disc
  - Part 14 – a multimedia container standard (MP4)

EBU723U

- 44 -

Part 10 and H.264 AVC

- Approx. half the bit rate of MPEG2/MPEG4 Part 2
- Has many new features that allow it to compress video much more effectively, e.g.
  - Up to 16 reference frames can be used, unlike previous limit of 2 for B-frames
  - Variable block-size motion compensation (VBSMC) with block size from 16 x 16 to 4 x 4, enabling precise segmentation of moving regions

H.264 AVC Applications

- The AVC encoder is designed for use in a variety of applications
  - mobile video streaming
  - video conferencing
  - internet video
  - DSL video streaming
  - Digital HD TV
- It is a generic standard
  - supporting a wide range of bit rates
  - and a wide range of resolutions

QuickTime

- A multimedia framework developed by Apple Computer, with 2 main components:
  - The QuickTime framework providing a common set of APIs for encoding and decoding audio and video
  - The QuickTime Moving (.mov) file format, an open source media container

Slide 47

- 47 -

QuickTime

- Its open architecture supports many file formats and codecs, including Motion JPEG, MPEG, H.264 and is extensible to support future codecs
- In February 1998, the ISO standards body gave QuickTime a boost by deciding to use it as the basis for parts of the new MPEG-4 standard

Slide 48

- 48 -

AVI

- Audio Video Interleave
- Video for Windows multimedia container format
- A special case of the RIFF (Resources Interchange File Format), defined by Microsoft
- Video for Windows supports several data compression techniques, including RLE, Indeo, MPEG4 and DivX

Slide 49

- 49 -

WMV

- Windows Media Video
- Compressed video file format incorporating several proprietary Microsoft codecs
- WMV 9 codec is a competitor to H.264, DivX etc.
- Adopted as format for Blu-ray discs

Slide 50

- 50 -

Flash Video

- Container file format used to deliver video over the Internet
  - FLV or F4V
- Can use multiple codecs:
  - Sorenson Spark
  - VP6
  - H.264 video

Slide 51

- 51 -

HTML 5 Video

- A HTML 5 tag introduced for the purpose of playing videos or movies
- Current draft specification does not specify which video formats browsers should support
  - Ogg multimeida container with Theora Video and Vorbis Audio
  - H.264

Slide 52

- 52 -

Agenda

- Digital video
- Video compression
- Video formats and codecs
- MPEG
- Other codecs
- Web video

EBU723U

- 53 -

Video Consumption

- 10 years ago: TV broadcast channels with fixed program
- Nowadays: on-demand program via web

Slide 54

- 54 -

Web Video

- The aim of the web is to provide searchable, immediately available network content
- It is likely that in the next ten years the provision of video content on the web will transform how the entertainment and education industries work

Slide 55

- 55 -

Web Video

- Provision of video over the Internet can be split into 4 main categories:
  - Downloadable file
  - Progressive download
  - Real-time streaming
  - Multicasting

Slide 56

- 56 -

Downloadable File

- Simplest method of providing video on the web
- Uses a file transfer protocol to download the file
  - This cannot provide real-time playback during the download
  - Quality is guaranteed and not dependent on bandwidth
  - The only restriction is how long the user is prepared to wait for the download to finish

Slide 57

- 57 -

Progressive download

- Served from standard Web and FTP servers using lossless transfer protocols
  - This means that all the video is guaranteed to arrive, but we cannot predict how long it will take
- Progressive download systems can start playing the file before it is fully downloaded
- Often called HTTP Streaming or Pseudo Streaming

Slide 58

- 58 -

Progressive using HTTP

- Any HTTP streamed content is simply a file that is streamed and started as soon as possible
- The client media player often displays a progress bar, and at some point (based on your connection speed), the movie begins to play
  - In theory playing won't catch up with the downloading
  - If it does, the movie will stop until more data is received

Slide 59

- 59 -

Progressive Download

- Progressive download gives the user a sample of the content before it is fully downloaded and enables termination of the download
- User often has the ability to navigate and play any of the file that has downloaded
- The user cannot navigate beyond the point that has already downloaded
- Simple to set up and very good when used with a limited number of viewers or a powerful web server

Slide 60

- 60 -

Progressive Download

- Progressive download has no mechanism for varying the download quality based on the bandwidth of the user's connection
  - Thus often have to provide several quality/sizes of downloadable file for the user
- As the user has to download the whole file up to a point to be able to view that point, this can cause additional server load
- YouTube – "Pseudo streaming", allows navigation to parts not yet downloaded

Slide 61

- 61 -

Real-time Streaming

- A technique for transferring data such that it can be processed as a steady and continuous stream
- Real-time streaming should provide random access to the whole streamed file so the user can fast-forward or rewind
- Requires a dedicated streaming server

Slide 62

- 62 -

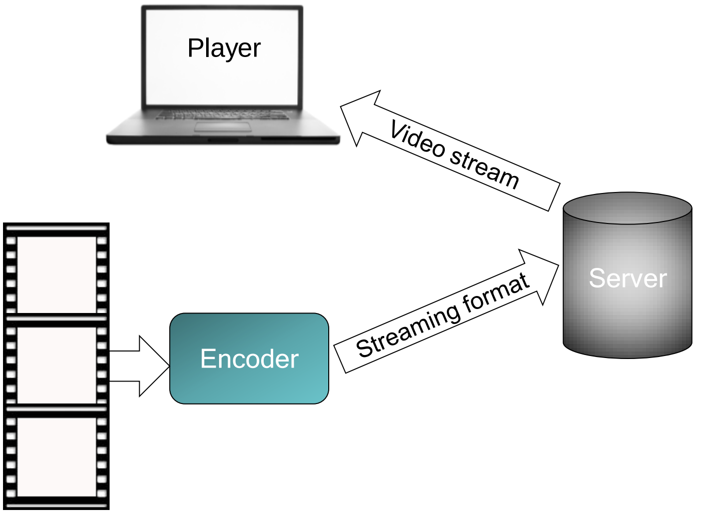
Real-time Streaming

- The multimedia data arrives, is briefly buffered before being played, and is then discarded
- The whole file is never actually stored on the user's computer
- Every time the play head is moved, the video will have re-buffer

Slide 63

- 63 -

Real-time Streaming



Slide 64

- 64 -

Application Protocols

- The most common application level protocol that can be used for real-time streaming of video media is the Real-Time Streaming Protocol (RTSP)
- RTSP is a client-server multimedia presentation control protocol originated by RealNetworks
- The protocol is similar in syntax and operation to HTTP but RTSP adds new functionality

Slide 65

- 65 -

Streaming with RTSP

- The sending of actual streaming data itself is not part of the RTSP protocol, but it acts as a control mechanism to handle client requests such as pause, fast forward, reverse, and absolute positioning
- Most RTSP servers use the Real-time Transport Protocol (RTP) as the transport protocol for the actual audio/video data

Slide 66

- 66 -

Real-Time Streaming

- Breaks data into many packets sized according to the bandwidth available between client and server
- When enough packets have been received by the client, the user's software can be:
  - Playing one packet
  - Decompressing another
  - Downloading the third

Slide 67

- 67 -

Real-Time Streaming

- A streamed clip will be instantly viewable, with few or no interruptions,
- but its video quality may be lower than progressive download
- Clips are streamed a the same time they are being reviewed
- Quality of a streamed clip may vary, but the clip should not stop playing

Slide 68

- 68 -

Real-Time Streaming

- The advantage of streaming are random access and “live” streaming
- Progressive download does not allow you to randomly jump to any position in the movie
- It is also difficult to stream a live performance using progressive download

Slide 69

- 69 -

Multicasting

- Real-time streaming to multiple recipients to all view the same content at the same time

- Unicast: one source to one destination
- Multicast: one source to many destinations

- Two main functions:
  - Efficient data distribution
  - Logical naming of a group

Slide 70

- 70 -

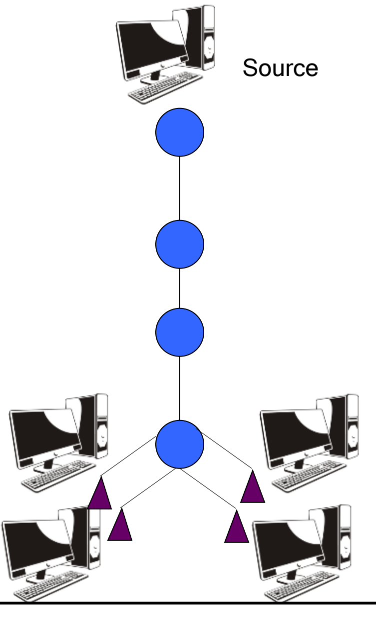
Multicast Routing

Requires users to “tune in” for a particular multicast which is only available to them

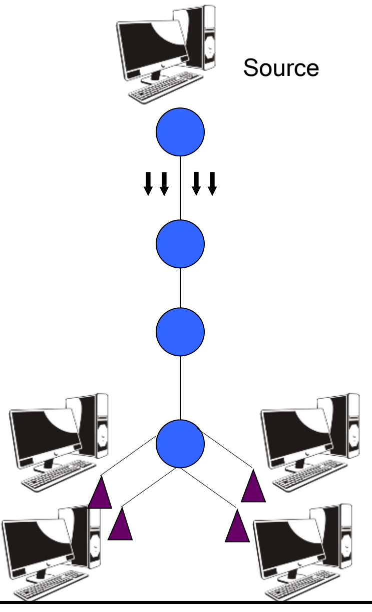
- Normal Internet traffic needs separate connections for each source-distribution pair, IP Multicasting allows many recipients to share the same source
- This means that just one set of packets is transmitted for all the destinations

- 71 -

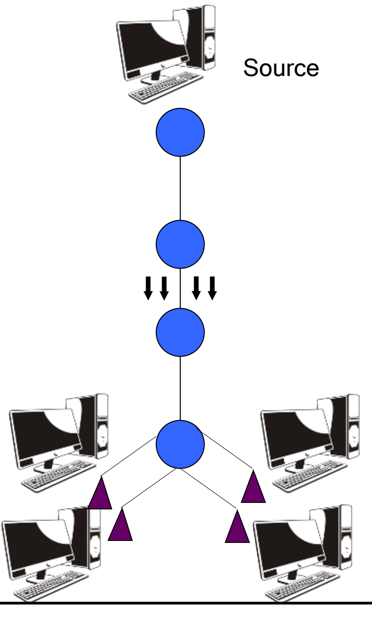
Unicast



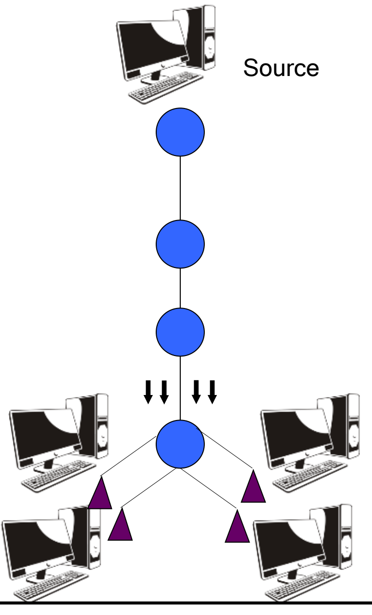
Unicast



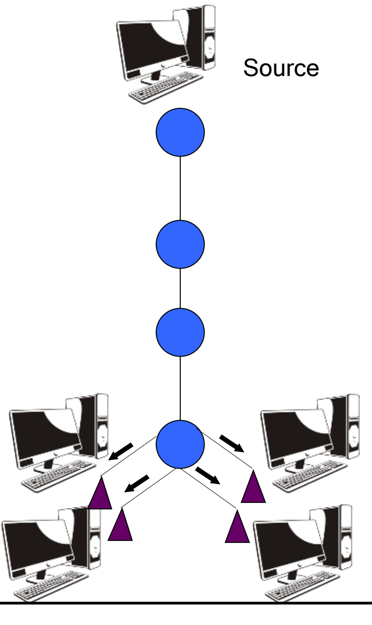
Unicast



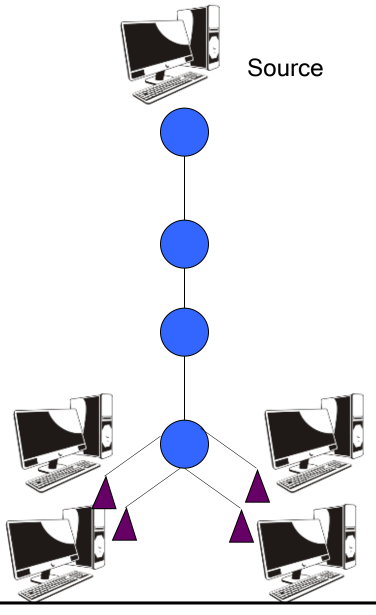
Unicast



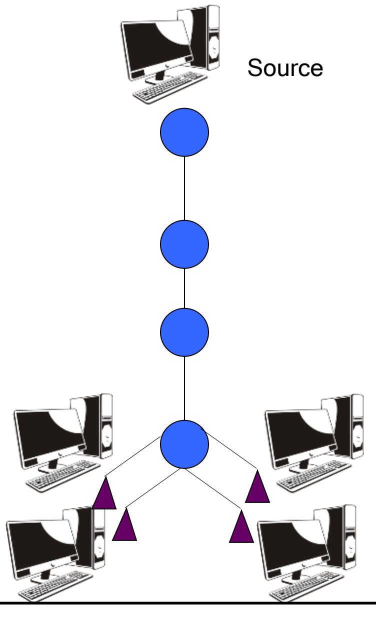
Unicast



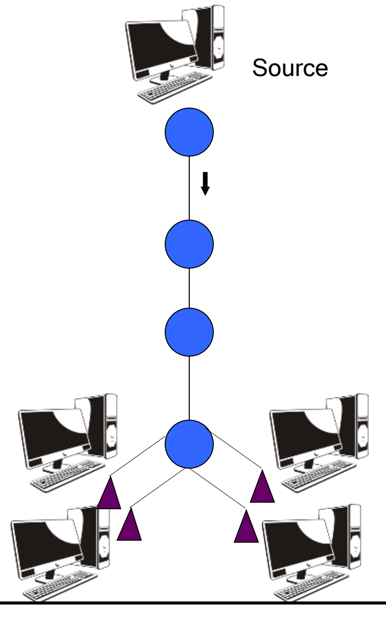
Unicast



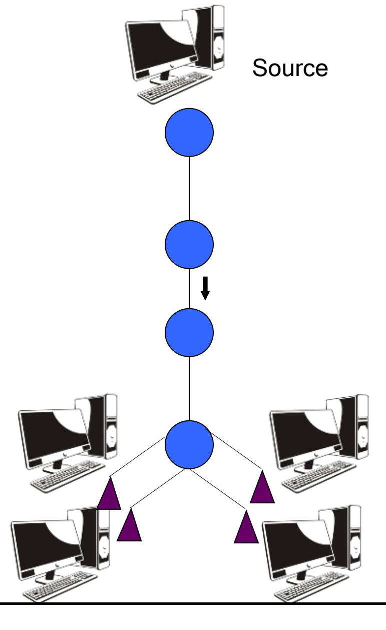
Multicast



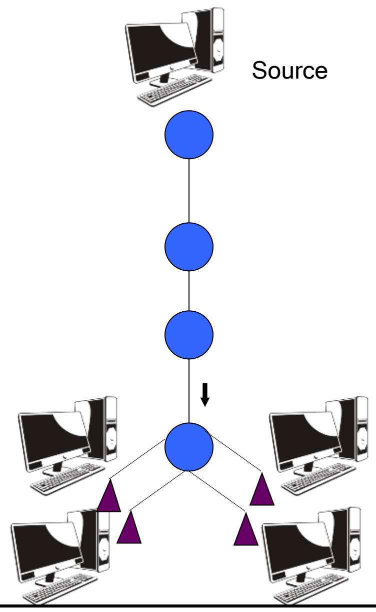
Multicast



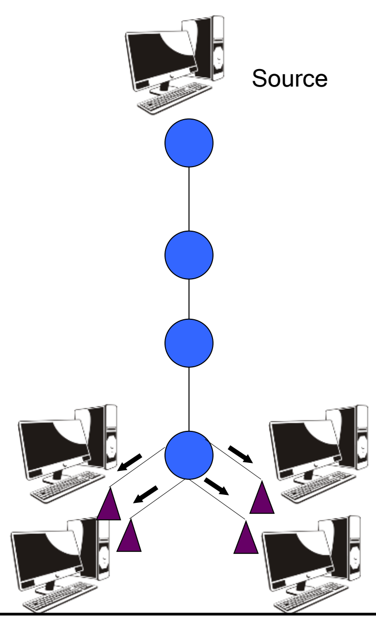
Multicast



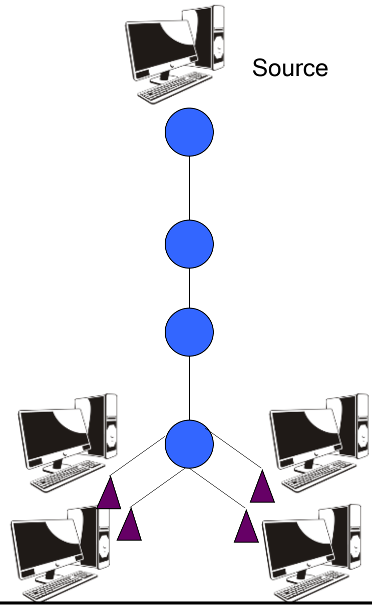
Multicast



Multicast



Multicast



Multicasting

- Provides one to many communication over an IP infrastructure
- Source sends a packet only once, even if it needs to be delivered to a large number of receivers
- The nodes in the network take care of replicating the packet to reach multiple receivers only where necessary

Slide 84

What did we learn?

- Digital video
- Video compression
- Video formats and codecs
- MPEG
- Other codecs

