Presentation Formatting

COS 460/540

1

End to End Data

- Data Formats
- eXtensible Markup Language
- Multimedia Data

2

Data Formats

- Basic data types
- Complex types and data
  - "records", audio, video, ...
- Sequences

3

Complex Types

4

- Compression
  - ...to reduce bandwidth needs
- Error Correction
  - ...to increase reliablilty

Transmitting Data

5

- Encoding
  - ...from model to network
- Decoding
  - ...from network to model

XML

6

eXtensible Markup Language

- Data and Tags/Markup (XML)
- Schema description of documents (XSD)

```
<?xml version="1.0"?>
<catalog>
  <book id="bk101">
    <author>Gambardella, Matthew</author>
    <title>XML Developer's Guide</title>
    <genre>Computer</genre>
    <price>44.95</price>
    <publish_date>2000-10-01</publish_date>
    <description>An in-depth look at applications
    with XML.</description>
  </book>
  <book id="bk102">
    <author>Ralls, Kim</auth
```

7

XML

- Based on Web Technologies
- Data and Markup are TEXT
- XML is a "framework"
- Nested tags/values
- Sequences of tags/values

8

```
<xsd:schema xmlns:xsd="http://www.w3....XMLSchema"
        targetNamespace="urn:books"
        xmlns:bks="urn:books">
 <xsd:element name="books" type="bks:BooksForm"/>
 <xsd:complexType name="BooksForm">
  <xsd:sequence>
   <xsd:element name="book"
            type="bks:BookForm"
            minOccurs="0"
            maxOccurs="unbounded"/>
  </xsd:sequence>
 </xsd:complexType>

 <xsd:complexType name="BookForm">
  <xsd:sequence>
```

9

XSD

10

- Defines valid XML Documents
- Written in XML
- Basic types: integer, string, boolean
- Complex types: nesting, sequences
- Namespaces to avoid name conflicts

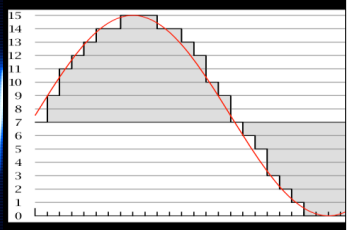
Multimedia Data

11

- The nature of multimedia data
- Compression
  - Lossless (for data)
  - Lossy (for images, video, audio)

Audio

12



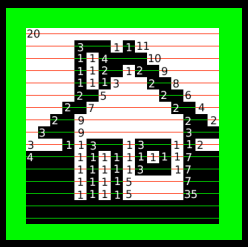
13

Sampling

- Sampling (time)
- Quantization (quantity, e.g. amplitude)

14

Images & Video



video is a sequence of images

15

That's a lot of data!

1080 x 1920 x 24

= 50Mb

24fps = 1.2Gbps

16

Lossless Compression

All the data are important

- Run Length Encoding
- Differential Pulse Code Modulation

17

Run Length Encoding

AAABBCDDDDDDAAAAABCCC 21B

12B

11B

18

DPCM

AAABBCDDDDDDAAAAABCCC 21B

(differential)

12B

(delta) A

11B

(RLE delta)

10B

19

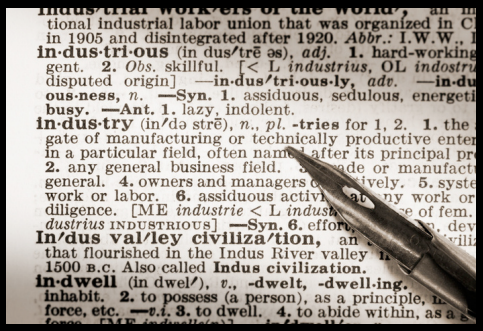
Huffman Code

Morse Code

A ·−	J ·−−−	S ···	1 ·−−−−
B −···	K −·−	T −	2 ··−−−
C −·−·	L ·−··	U ··−	3 ···−−
D −··	M −−	V ···−	4 ····−
E ·	N −·	W ·−−	5 ·····
F ··−·	O −−−	X −··−	6 −····
G −−·	P ·−−·	Y −·−−	7 −−···
H ····	Q −−·−	Z −−··	8 −−−··
I ··	R ·−·		9 −−−−·
			0 −−−−−

20

LZW - Dictionary



21

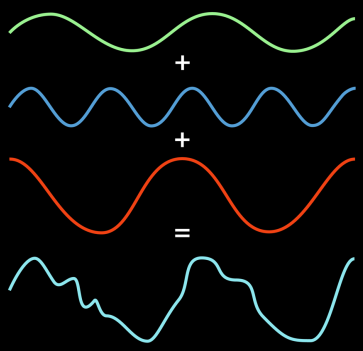
Lossy Compression

All the data are

NOT important

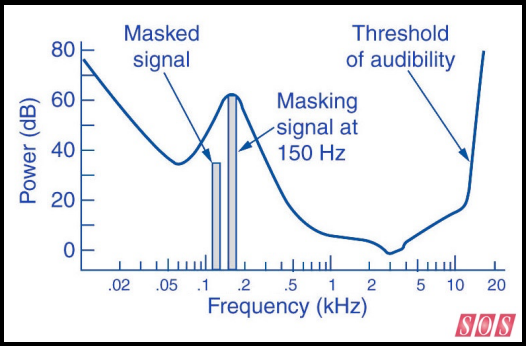
- Single image compression
- Stream compression

MP3 (audio)



22

Signal Reduction (masking)



23

CBR vs VBR (constant vs variable bitrate)

24Kbps = Spoken word (telephone)
48-64Kbps = AM Radio
128Kbps = reasonable for car-radio, falls off over 16KHz (cymbals)
192KBps = 'near CD quality'
>= 256Kbps = identical to original up to about 18KHz

Bitrate?

- 24Kbps = Spoken word (telephone)
- 48-64Kbps = AM Radio
- 128Kbps = reasonable for car-radio, falls off over 16KHz (cymbals)
- 192KBps = 'near CD quality'
- >= 256Kbps = identical to original up to about 18KHz

24

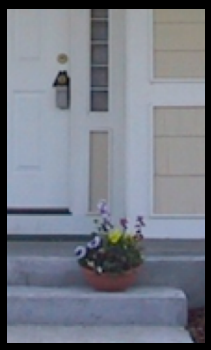
25

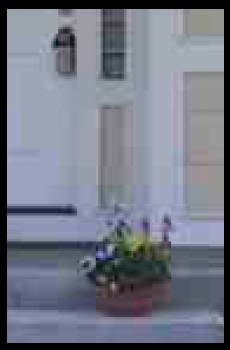
CBR vs VBR

- Constant
  - Same bitrate throughout the stream
- Variable
  - bitrate changes based on content analysis

26

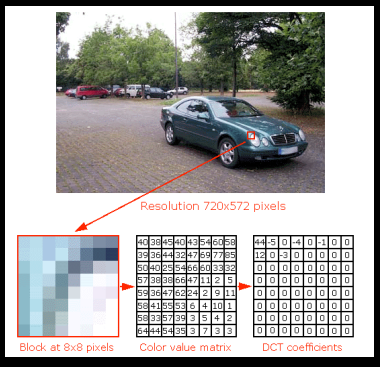
Images & Audio





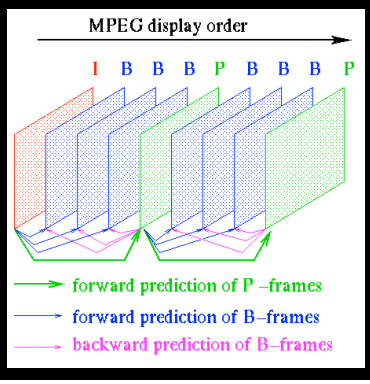
27

JPEG



MPEG

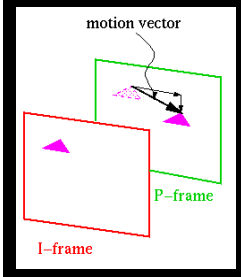
28

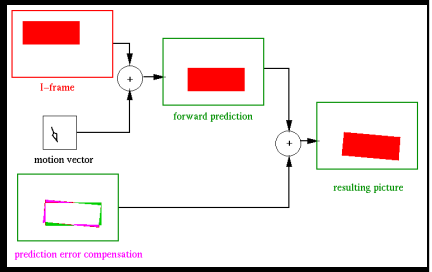


https://vsr.informatik.tu-chemnitz.de/~jan/MPEG/HTML/mpeg_tech.html

Prediction

29





End

30

Presentation Formatting
XML